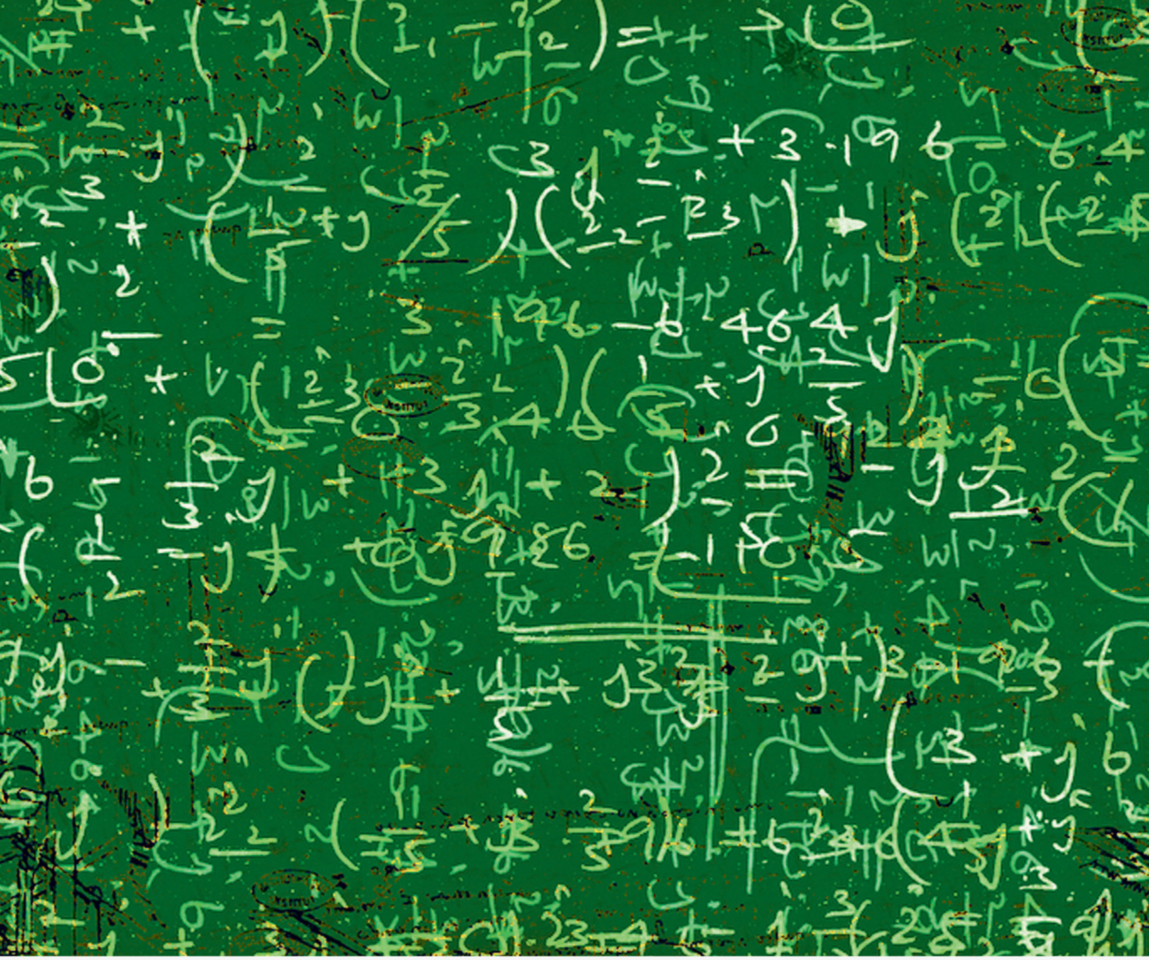


Fundamentals of Advanced Mathematics 3

Henri Bourlès

Differential Calculus, Tensor Calculus, Differential Geometry, Global Analysis

ISTE PRESS

ELSEVIER

Fundamentals of Advanced Mathematics 3

This page intentionally left blank

New Mathematical Methods, Systems and Applications Set

coordinated by
Henri Bourlès

Fundamentals of Advanced Mathematics 3

Differential Calculus, Tensor Calculus, Differential Geometry, Global Analysis

Henri Bourlès

iSTE PRESS

ELSEVIER

First published 2019 in Great Britain and the United States by ISTE Press Ltd and Elsevier Ltd

ISTE Press Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

Elsevier Ltd
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB
UK

www.elsevier.com

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress
ISBN 978-1-78548-250-2

Printed and bound in the UK and US

Contents

This page intentionally left blank

Preface

This third volume of *Fundamentals of Advanced Mathematics* (the first two volumes are referenced by [P1] and [P2] below) is dedicated to differential and integral calculus, examined from both the local and global perspectives. This book is intended for anyone who uses mathematics (mathematicians, but also physicists and engineers, and in particular anyone who needs to understand the control of nonlinear systems). Some local questions of integral calculus were already partially addressed in [P2], Chapter 4, and the natural framework of differential calculus, Banach spaces, is presented in Chapter 1 of this third volume. Nonetheless, we will need to consider a few generalizations to sketch the so-called "convenient" context for more recent developments in global analysis; more on this later. We will also present the "Carathédory conditions", which are finer than the classical Cauchy–Lipschitz existence and uniqueness conditions for solutions of finite-dimensional ordinary differential equations.

Global questions demand another perspective. Seen through our window, the Earth appears entirely flat, yet even the Greeks in the age of Plato knew better, as testified by an excerpt from his *Phaedo* (108, e). Knowledge of the Earth's shape undoubtedly extended back to the Pythagoreans in the 6th Century BC. Thus, Euclid, who was an avid student of Plato's work (consider, for example, his construction of the five Platonic solids in Book XIII of the *Elements*), certainly also knew that the Earth is round. Yet his geometry is quite different from the geometry of a sphere. Euclidean geometry offers a good local approximation of spherical geometry but was of course useless for the lengthy voyages of the Renaissance. Global analysis must therefore be expressed in the framework of *manifolds*, a concept which generalizes curves and surfaces since Riemann.

Ever since the invention of variational calculus in the late 17th Century, it has become common in mathematics to argue about sets of functions. If certain conditions are satisfied, these sets can be endowed with the structure of a manifold, in which case they are called "functional manifolds" and imagined as deformed versions of the usual function spaces (in the same way that the Earth might be imagined as a deformed version of the plane). "Banach" manifolds were the first candidates to be considered as a framework for global analysis in the late 1950s and the 1960s [EEL 66, PAL 68] (these manifolds are deformed versions of a Banach space). However, as became clear in [P2], section 4.3, many of the function spaces encountered in practice are not Banach spaces. For example, the space $\mathcal{E}$ of infinitely differentiable functions on a non-empty open subset of $\mathbb{R}^n$ is a nuclear Fréchet space. Since the 1980s, this has inspired research into manifolds that are deformed versions of spaces of this type; this is the "convenient" context for global analysis mentioned above (which matured around the late 1990s [KRI 97]). Although we will not be able to present it exhaustively in this book, our discussion of manifolds of mappings in section 5.3 will demonstrate the considerable value of this approach.

Chapters 2 to 4 develop the required formalism, with a brief detour in Chapter 3 to introduce a notion that has played a fundamental role ever since É. Cartan, the concept of fiber bundles, and in particular principal bundles. According to general relativity, we live in a pseudo-Riemannian space that is the "base" of a principal bundle; the latter is namely the manifold of orthonormal frame, and its "structural group" (which performs changes of reference frame) is a "Lie group" , namely the Lorentz–Poincaré orthogonal group of matrices leaving invariant the quadratic form $(ds)^2 = c^2 (dt)^2 - (dx)^2 - (dy)^2 - (dz)^2$. Tensor calculus, a staple of physics textbooks since the early 20th Century, is presented in Chapter 4, alongside the theory of differential p-forms.

Our formalism first begins to truly bear fruit in Chapter 5. Distributions, and the generalized notion of currents, may now be defined on manifolds instead of open subsets of $\mathbb{R}^n$. The idea of exterior derivatives of a differential p-form (introduced by É. Cartan) allows us to give highly condensed expressions for the classical formulas of "vector calculus" involving gradients, divergences, Laplacians, etc. The first fundamental result of this chapter is a general formulation of Stokes' theorem encompassing the Ostrogradsky, Gauss, Green–Riemann and Green theorems widely used in physics, as well as the "classical" Stokes' theorem. On a Riemannian manifold, Stokes' theorem enables us to formulate Hodge duality, which simplifies many of our calculations involving vectors. From the perspective of algebraic topology, Stokes' theorem also gives rise to two other types of duality: Poincaré duality for homologies and De Rham duality for cohomologies. On $\mathbb{R}^3$, for instance, we know that the curl of any vector field $\overrightarrow{E}$ that derives from a potential is zero and the divergence of any vector field $\overset{\hookrightarrow}{B}$ that can be expressed as a curl is also zero. Stokes' theorem allows us to prove the converse of each claim. The second fundamental result of Chapter 5 is the Frobenius theorem, which gives necessary and sufficient conditions for the integrability of a "contact distribution". This allows us to establish the concept of foliation. The Frobenius theorem also implies a result by Riemann in Chapter 7 that is essential for general relativity, namely that a Riemannian manifold is flat if and only if its curvature tensor is zero (section 7.4.3, Theorem 7.56).

Lie groups are manifolds but also groups; in Chapter 6, the group structure enables us to perform operations that would not make sense on an ordinary manifold, specifically convolution of functions or distributions. Furthermore, a "taxonomy" of Lie groups can be established from the Lie algebras associated with each group, which are vector spaces and therefore easier to study: as a set, the Lie algebra $\mathfrak{g} = \mathrm{Lie}(\mathbf{G})$ of the Lie group $\mathbf{G}$ is the tangent space $T_e(\mathbf{G})$ of $\mathbf{G}$ at the point e, where e is the neutral element of $\mathbf{G}$. However, the three "fundamental theorems" of S. Lie imply that there exists a "dictionary" that allows us to characterize Lie groups by the properties of their Lie algebras, at least locally (and globally if $\mathbf{G}$ is simply connected). The classification established by Lie is complete in the case of simple or semi-simple Lie groups (or algebras). This is the most important case, since these are the groups frequently encountered in particle physics, where they play an essential role (including the so-called "exceptional" simple Lie algebras). Simple and semi-simple Lie algebras have been studied since Cartan in terms of their "root systems"; the ability to represent these root systems graphically (as proposed by Coxeter and Dynkin, among others) is extremely useful, but cannot be presented in detail in this book[1].

On a reductive Lie group $\mathbf{G}$, we can also fully develop the theory of harmonic analysis (Fourier transforms of functions or tempered distributions). The abelian case will be presented in detail: when $\mathbf{G} = \mathbb{R}^n$, we recover the usual notion of Fourier transform; when $\mathbf{G}$ is the torus $\mathbb{T}^n$, we recover the Fourier series expansion of periodic functions or distributions. The non-abelian case would fill another entire volume and thus will only be briefly mentioned (even though engineering applications have recently been found [CHI 01]). Readers are welcome to refer to the bibliography for the non-abelian case [VAR 77, VAR 89].

Defining a geometry on a manifold is equivalent to equipping this manifold with a connection (Chapter 7). Lie groups are implicitly equipped with a connection. Riemannian or pseudo-Riemannian manifolds are often implicitly equipped with the simplest possible connection: the Levi-Cività connection. This is a special case of a "$\mathbf{G}$-structure" that is frequently used in general relativity. É. Cartan clarified the notion of connection; he studied affine, projective and conformal connections, summarizing his ideas by proposing the concept of "generalized space" [CAR 26];

1 See the Wikipedia article on *Coxeter–Dynkin diagrams*.

these spaces are equipped with connections called *Cartan connections* since Ehresmann (who rephrased these ideas within the context of principal connections). Connections can be equipped with curvature (an idea that should be familiar to relativistic physicists) and in some cases torsion, which attracted considerable interest from Einstein ([EIN 54], Appendix II), who hoped to find a way to unify the theories of gravitation and electromagnetism.

Henri BOURLÈS
June 2019

Errata for Volume 1 and Volume 2

Volume 1 (Cont'd)

1) On p. 12, the fifth line of **(V)** should read $\boldsymbol{R}$ instead of R.

2) On p. 22, on the right-hand side of [1.6], it should read $\varprojlim$ instead of $\varinjlim$.

3) On p. 41, line 10 should read "the cardinal of G/H (equal to the cardinal of $G\backslash H$)" instead of "this cardinal".

4) On p. 45, the first line after [2.12] should read M_3 instead of G.

5) On p. 190, in Definition 3.177, it should read Δ^n instead of Δ_n.

6) On p. 191, in the first line after [3.70], add "also denoted as d_p" after "operator".

7) On p. 193, line 17 should read $\mathfrak{S}_p$ instead of $\mathfrak{S}_n$.

8) On p. 220, line 7 should read "$\pi = \prod_{j \in J} \pi_j$ where the elementary divisors π_j are pairwise non-associated and maximal powers (among all elementary divisors) of irreducible polynomials" instead of "$\pi = \prod_{i=1}^{n} \pi_i$".

Volume 2

1) On p. 12, line 6 should read "$\mathbf{K}' \, (\alpha')$" instead of "$\mathbf{K} \, (\alpha')$".

2) On p. 17, line 19 should read "$\neq$" instead of "$=$"; line 21 should read "does not exist" instead of "exists".

3) On p. 20, line 11 should read "$0 \leq i \leq r$" instead of "$1 \leq i \leq r$".

4) On p. 24, line 28 should read "$x \in \bar{\mathbf{K}}$" instead of "$x \in \mathbf{K}$".

5) On p. 27, line 10 should read "$\bar{\mathbf{K}}$" instead of "$\mathbf{K}$".

6) On p. 32, line 3 should read "a is not of the form $b^n, b \in \mathbf{K}$" instead of "$a^n \notin \mathbf{K}$"; line 5 should read "ζ" instead of "ς".

7) On p. 43, in line 10, after "yields", add "with $j = 0, ..., n-2$"; the last sum of line 11 should read "$u_i^{(j)}$" instead of "$u_i^{(j+1)}$".

8) On p. 51, replace the sentence beginning line 23 by: "The coefficients c and d are free parameters, the former in $\mathbb{C}$, the second in $\mathbb{C}^\times$; indeed, putting $\mathbf{M} = \mathbb{C}\left(t, e^{-t^2/2}\right)$ we have that $\mathrm{Gal}^D(\mathbf{N},\mathbf{M}) = \mathbb{C}$, $\mathrm{Gal}^D(\mathbf{M},\mathbf{K}) = \mathbb{C}^\times$, which yields the short exact sequence of Abelian groups $0 \longrightarrow \mathbb{C} \longrightarrow G \longrightarrow \mathbb{C}^\times \longrightarrow 0$, in other words G is an extension of $\mathbb{C}^\times$ by $\mathbb{C}$ ([P1], section 2.2.2(**II**)).".

9) On p. 57, line 2 should read "nonempty set" instead of "set".

10) On p. 62, lines 27, 29; on p. 63, lines 3, 8; on p. 64, lines 7, 9: it should read "$\left(x'_j\right)_{j\in J}$" instead of "$\left(x'_j\right)_{i\in J}$".

11) On p. 83, line 22 should read "smallest" instead of "largest" and "$=$" instead of "$\neq$".

12) On p. 91, line 28 should read "$j \succeq i_0$" instead of "$j \geq i_0$".

13) On p. 92, in line 19, add after the last sentence: "This extension is then unique.".

14) On p. 95, line 8 should read "$\forall\,(x', x'')$" instead of "$\forall\,(x'; x'')$".

15) On p. 105, line 25 should read "$\forall i \succeq i''_0$" instead of "$\forall i \succeq i_0$".

16) On p. 119, line 17 should read "$\mathbb{K}$" instead of "$\mathbf{K}$".

17) On p. 128, lines 1, 2 should read "(iv)" instead of the second "(iii)" and "(v)" instead of "(iv)".

18) On p. 132, line 20 should read "$\xi x_0 \mapsto \xi$" instead of "$\xi \mapsto \xi x_0$".

19) On p. 133, in line 14, delete "in **Lcsh**".

20) On p. 135, lines 19–21 should read "*reduced* if for all $i \in I$, the projection $\mathrm{pr}_i(E)$, where pr_i is the canonical projection $\prod_i E_i \twoheadrightarrow E_i$, is dense in E_i" instead of "*decreasing* [...] $\mathfrak{T}_j$)"; in lines 27–31, replace Statement 2) of Remark 3.33 by: "Every projective limit can be put into the form of a reduced projective limit: if $F_i = \mathrm{pr}_i(E)$ and $\tilde{\psi}_i^j$ is the restriction of ψ_i^j to $\prod_i F_i$, then E is the subspace $\bigcap_{i\preceq j} \ker\left(\tilde{\psi}_i^j\right)$ of $\prod_i F_i$.".

21) On p. 137, the last line should read "$\sqrt[p]{\sum_{i=1}^n \|x_i\|^p}$" instead of "$\sqrt[p]{\sum_{i=1}^n \|x_i\|^P}$".

22) On p. 141, in lines 24, 26, delete "decreasing".

23) On p. 142, in lines 10, 11, delete "is surjective (*ibid.*) and".

24) On p. 144, line 11 should read "bounded in F" instead of "bounded in M".

25) On p. 168, line 13 should read "$\bar{x} \in A$" instead of "$\bar{x} \in \bar{\mathbb{R}}$".

26) On p. 150, in line 6, delete "decreasing"; line 8 should read "mapping" instead of "surjection" and "$\to$" instead of "$\twoheadrightarrow$".

27) On p. 172, in line 25, it should read "exact" instead of "quasi-exact"; delete "and from $\mathbf{FS}^{\mathrm{op}}$ to **Sil**" and Note 17.

28) On p. 173, in line 2, delete "strict"; line 5 should read "an" instead of "a strict"; in line 7, delete "strict" and read "Silva" instead of "$(\mathcal{FS})$-"; line 8 should read "$\longrightarrow$" instead of "$\hookrightarrow$" and "mapping" instead of "injection"; in line 8, it should read "mapping" instead of "surjection", "$\longrightarrow$" instead of "$\hookrightarrow$" and delete "strict"; in line 10, it should read "$(\mathcal{FS})$-" instead of "Silva"; replace lines 13–27 by: "(1) Since the spaces $(E_i')_b$ are Fréchet, they are bornological, and their inductive limits as well. Thus they are Mackey spaces and the result follows from ([SCF 99], Chapter IV, Section 4.4). (2) The E_i are reflexive $(\mathcal{DF})$ spaces, and the result follows from ([SCF 99], Exercise 24(f), p. 197).".

29) On p. 174, in line 2, delete "increasing"; in line 3, it should read "mapping" instead of "injection" and "$\longrightarrow$" instead of "$\hookrightarrow$".

30) On p. 175, line 14 should read "$\mathcal{L}(R; S)$" instead of "$\mathcal{L}(R, S)$".

31) On p. 177, lines 8, 9 should read "$\bigcup_{u \in H, x_1 \in A_1} u(x_1, A_2)$" instead of "$\bigcup_{u \in H, x_1 \in A_1}$".

32) On p. 185, line 5 should read "such" instead of "any"; lines 15, 23 should read "$\geq$" instead of "$\leq$".

33) On p. 186, line 11 should read "$|\langle e_i|x\rangle|^2$" instead of "$\|\langle e_i|x\rangle\|^2$".

34) On p. 188, in line 27, add "salient" after "closed".

35) On p. 193, in line 12, add "$, v = \sum_{i=1}^{r} x_i' \otimes y_i$" at the end of the line.

36) On p. 200, in line 4, change $\delta_a(a)$ to $\delta_a(A.)$

37) On p. 204, in line 5, suppress the sign $\int$.

38) On p. 204, line 21 should read "$\sqrt[p]{\int |\mathbf{f}|^p .d\mu}$" instead of "$\sqrt[p]{\int |\mathbf{f}| .d\mu}$".

39) On p. 218, the third line of the proof should read "$\left|\frac{1}{h}(f(x+h) - f(x)) - g(c)\right|$" instead of "$\frac{1}{h}|f(x+h) - f(x)|$".

40) On p. 230, in line 16, add "nonzero" after "many".

41) On p. 253, in lines 1, 3, delete "strict"; in line 26, add "denoted by $C_r(0)$" after "plane".

42) On p. 299, line 18 should read “is an **A**-module” instead of “is the **A**-module W”.

43) On p. 204, line 13 should read “equation” instead of “solution”.

44) On p. 247, change lines 1, 2 to “given by the functions with absolute value upper bounded by the absolute value of a polynomial, and their partial derivatives of all orders as well”.

List of Notations

Chapter 1: Differential Calculus

$\infty, \omega, \mathbb{N}_{\mathbb{K}}, \mathbb{N}_{\mathbb{K}}^{\times}$, p. 2

$O(f), o(f)$: Landau notation, p. 2

$\mathbf{E}^{\vee}$: dual of $\mathbf{E}$, p. 2

$|.|_{\gamma}$, $\|.\|_{\gamma}$, p. 2

$\mathcal{L}_{n,s}(E;F)$: space of symmetric elements of $\mathcal{L}_n(\mathbf{E};\mathbf{F})$, p. 2

$\mathbf{u}.\mathbf{h}^n$, p. 2

$D\mathbf{f}(a)$, $d\mathbf{f}(a)$, $\mathbf{f}'(a)$: (Fréchet) differential of $\mathbf{f}$ at the point a, p. 6

$\dot{\mathbf{f}}(a)$: derivative of $\mathbf{f}$ at the point a, p. 6

$\mathcal{D}_a(A;F)$, p. 6

$\mathrm{rk}_a(\mathbf{f})$, p. 6

$\mathrm{grad}_a(f), \nabla_a f$: gradient of f at the point a, p. 6

$D_i\mathbf{f}(a), \partial_i\mathbf{f}(a), \frac{\partial f}{\partial x^i}(a)$: partial differential, partial derivative, p. 8

$\frac{\partial(f^1,\dots,f^n)}{\partial(x^1,\dots,x^n)}(a)$: Jacobian, p. 8

$\mathcal{C}^p(A;F)$, p. 10

$Hf(a)$: Hessian matrix, p. 11

$\mathcal{N}$: Nemitsky operator, p. 11

ev: evaluation operator, p. 11, 186

$\bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$, p. 26

F_B, p. 12

$\int_a^b \mathbf{f}(t)\,dt$, p. 13

D^α: partial differential of order α, p. 16

$\mathcal{S}(\mathbf{E};\mathbf{F})$: space of formal series, p. 16

$\mathbb{S}(\mathbf{E};\mathbf{F})$: space of convergent series, p. 17

$\rho(\mathbf{S})$: radius of convergence, p. 17

$\mathcal{C}^\omega(A;F)$, p. 17

$\mathbf{f}_a$, p. 17

$\delta\mathbf{f}(a), \delta^2\mathbf{f}(a)$: first, second Lagrange variation, p. 28

$D^{\mathcal{G}}\mathbf{f}(a)$: Gateaux differential, p. 28

$\mathfrak{c}^\infty\mathbf{E}$, p. 32

$\mathfrak{c}^\infty$, $\mathfrak{c}^\omega$, p. 33

$\varphi(.;t_0,x_0)$, p. 41

$\Phi(t_2,t_1)$: resolvent, p. 44

$\varphi_\lambda = \varphi(.,t_0(\lambda),x_0(\lambda))$, p. 46

Chapter 2: Differential and Analytic Manifolds

$(U,\varphi,\mathbf{E}), (U,\varphi,m)$: chart, p. 51, 52

$\dim(M), \dim_x(M)$: dimension, p. 52

$T_a(M)$: tangent space, p. 59

$\theta_c : T_a(M) \to \mathbf{E}, \vartheta_c = \theta_c^{-1}$, p. 59

$S_a^r(M)$: space of germs of stationary functions at the point a, p. 61

$\left(\frac{\partial}{\partial \xi^j}\right)_a$, p. 61

$\mathcal{L}_{X_a}$: Lie derivative, p. 62

$\gamma_{i*0}, \partial_i(a), \partial_i|_a$, p. 64

$X_a.f$, p. 64

$T_a(f), f_{*a}, f_*(a)$: tangent linear mapping, p. 65

$\mathrm{rk}_a(f)$: rank of a morphism, p. 66

$d_a\mathbf{f}$: differential, p. 66

$T^i_{(a_1,a_2)}(f)$, p. 75

$M_1 \times_Z M_2$, $f_1 \times_Z f_2$: fiber product, p. 79

${}^t f_a$: cotangent linear mapping, p. 80

$\mathbf{LieGrp}$: category of Lie groups, p. 81

$\mathbf{G}^\circ$: neutral component of $\mathbf{G}$, p. 83

$\mathbb{T}^n$: n-dimensional torus, p. 82

$\mathbf{H} \ltimes \mathbf{K} = \mathbf{K} \rtimes \mathbf{H} = \mathbf{H} \times_\tau \mathbf{K}$: semi-direct product of subgroups ($\mathbf{H}$ normal)

V^∞, p. 84

$\boldsymbol{\lambda}(s), \boldsymbol{\rho}(s)$: left, right translation, p. 88

$\mathrm{GL}(\mathbf{E})$: automorphism group of $\mathbf{E}$, p. 82

$\mathrm{U}_n(\mathbb{C}), \mathrm{O}_n(\mathbb{R}), \mathrm{SL}_n(\mathbb{K}), \mathrm{SO}_n(\mathbb{R}), \mathrm{SU}_n(\mathbb{C})$: unitary, orthogonal, special linear, special orthogonal, special unitary group, p. 86

$\mathfrak{Z}(\mathbf{G})$: center of $\mathbf{G}$, p. 87

$\mathrm{PGL}_n(\mathbb{K}), \mathrm{PSL}_n(\mathbb{R}), \mathrm{PO}_n(\mathbb{K}), \mathrm{PSO}_n(\mathbb{K}), \mathrm{PU}_n(\mathbb{C})$: projective general linear, projective special linear, projective orthogonal, projective special orthogonal, projective unitary group, p. 87

$\mathrm{Sp}_{2n}(\mathbb{K}), \mathrm{USp}_n, \mathrm{A}_n, \mathrm{E}_n, \mathrm{SE}_n$: symplectic, unitary symplectic, general affine, affine orthogonal, special affine orthogonal group, p. 87

$\mathrm{D}_n(\mathbb{K}), \mathrm{T}_n(\mathbb{K}), \mathrm{ST}_n(\mathbb{K}), \mathrm{N}_n(\mathbb{K})$, p. 87

Ad: adjoint mapping, p. 88

$M/\mathbf{G}, \mathbf{G}\backslash M$: manifold of orbits, p. 89

$\mathfrak{g}$: tangent space $T_e(\mathbf{G})$, p. 90

$\mathfrak{Aff}_n(\mathbb{K}), \mathfrak{Euc}_n$: affine, affine Euclidean spaces, p. 91

Chapter 3: Fiber Bundles

$T(B)$: tangent fiber bundle, p. 94

$T^\vee(B)$: cotangent fiber bundle, p. 96

$\lambda = (M, B, \pi)$: fibration, p. 99

$\mathbb{S}^n$: n-dimensional sphere of radius 1, p. 98

$\lambda \times_B \lambda', M \times_B M'$: fiber product of fibrations, p. 101

$\lambda \times \lambda'$: product of fibrations, p. 102

$f^{0*}(\lambda)$: preimage of a fibration, p. 102

$\tilde{\mathbf{G}}$: universal covering of the Lie group $\mathbf{G}$, p. 106

$\mathrm{Spin}_n(\mathbb{K})$: spinor group, p. 106

$\Gamma^{(k)}(U, M), \Gamma(U, M)$: set of sections of class C^k of U in M, of morphisms of class C^k from U into M, p. 108

(M, B, π): vector bundle M with base B and projection π, p. 108

$\mathrm{rk}_b(M)$: rank of the vector bundle M, p. 108

$\mathbf{E}_B$: trivial bundle, p. 109

$M^\vee, M^*$: dual bundle of M, p. 112

$(\mathbf{s}^{*i}) = (d\xi^i)$: dual of the frame $(\mathbf{s}_i) = \left(\frac{\partial}{\partial \xi^i}\right)$, p. 112

M/M': quotient bundle, p. 113

$M' \otimes M''$, $M' \oplus M''$: tensor product, Whitney sum of two fiber bundles, p. 114

$M_{(\mathbb{C})}$: complexification of the real fiber bundle M, p. 115

$\mathbf{VB}$: category of vector bundles, p. 117

$\ker(u), \mathrm{im}(u), \mathrm{coker}(u)$: kernel, image, cokernel of the locally direct morphism u, p. 118

$f^{0*}(M)$: preimage of the vector bundle M under f^0, p. 120

$(P, B, \mathbf{G}, \pi)$: principal bundle P of base B, structural group $\mathbf{G}$ and projection π, p. 121

$(B \times \mathbf{G}, B, pr_1)$: trivial principal bundle, p. 122

$V_q(P)$: space of vertical tangent vectors, p. 123

$P \times^{\mathbf{G}} F, \mathbf{G}\backslash(P \times F), \left(P \times^{\mathbf{G}} F, B, \pi_F\right)$, p. 126

Chapter 4: Tensor Calculus on Manifolds

$\mathbf{T}_q^p(\mathbf{E})$: space of p-times contravariant and q-times covariant tensors (also called tensors of type (p, q)), p. 133

$\mathbf{T}(\mathbf{E})$: tensor algebra of $\mathbf{E}$, p. 133

$c_j^i : \mathbf{T}_q^p(\mathbf{E}) \to \mathbf{T}_{q-1}^{p-1}(\mathbf{E})$: index contraction mapping, p. 134

$\widehat{(.)}$, p. 134

$\sigma.\boldsymbol{t}$: image of the tensor $\boldsymbol{t}$ under the permutation σ, p. 135

$\mathbf{s}.\boldsymbol{t}, \mathbf{a}.\boldsymbol{t}$: symmetrization, antisymmetrization of the tensor t, p. 135

alt, p. 136

$\mathbf{TS}^n(\mathbf{E})$: space of symmetric contravariant tensors of order n, p. 135

$\mathbf{A}^n(\mathbf{E})$: space of antisymmetric contravariant tensors of order n, p. 135

$z_p \wedge z_q$: exterior (or wedge) product of $z_p \in \mathbf{A}^p(\mathbf{E})$ and $z_q \in \mathbf{A}^q(\mathbf{E})$, p. 138

$\bigwedge^{n} \mathbf{E} = \mathbf{A}^{n}(\mathbf{E})$: n-th exterior power of $\mathbf{E}$, p. 138

$\det(\mathbf{E}) = \bigwedge^{m} \mathbf{E}$, p. 138

$\bigwedge \mathbf{E} = \bigoplus_{0 \leq p \leq m} \bigwedge^{p} \mathbf{E}$: exterior algebra of $\mathbf{E}$, p. 139

$\mathcal{A}_n(\mathbf{E}; \mathbf{F})$: vector space of alternating n-linear mappings from E^n into $\mathbf{F}$, p. 140

$\mathrm{Sh}(p, q)$, p.140

$\lrcorner, i_v, i(\mathbf{s})$: interior product, p. 141, 153

$\mathrm{Alt}^{q}(\mathbf{E}'; \mathbf{E})$: space of continuous antisymmetric q-linear mappings from $\mathbf{E}'^{q}$ into $\mathbf{E}$, p. 143, 151

λ: vector functor, p. 144

$\wedge_{\Phi}$, p. 145, 152

$\mathcal{T}_0^1(U)$: space of vector fields of class C^r, p. 145

$\mathcal{T}_1^0(U), \Omega^1(U)$: space of covector fields of class C^r (Pfaff forms), p. 146

$\mathcal{T}_q^p(U)$: tensor field of type (p, q), p. 147

$\Omega^p(U; N), \Omega^p(U; \mathbf{F})$: space of p-forms, p. 148, 151

$f^*(\omega)$: preimage, p. 149, 154, 164

$\Omega(U; A)$: de Rham algebra, p. 149, 153

$\mathrm{Or}_{T(B)} = \tilde{B}$: orientation covering, p. 155

$\overset{\frown}{B}$: oriented manifold, p. 157

$\int_{\overset{\frown}{B}} \omega$, p. 160

$[\omega]$: volume form, p. 161

$\tilde{f}$: orientation of the morphism f, p. 162

$\tilde{O}$: canonical orientation of $\tilde{B}$, p. 163

Chapter 5: Differential and Integral Calculus on Manifolds

Der $(\mathbf{A})$: algebra of derivations of the algebra $\mathbf{A}$, p. 189

ad: adjoint linear representation, p. 190

$\mathfrak{z}(\mathfrak{g})$: center of $\mathfrak{g}$, p. 190

$\mathcal{L}_X$: Lie derivative, p. 191, 193, 194

d: exterior differential, p. 195

grad, $\overrightarrow{\nabla}$: gradient, p. 198, 209

div: divergence, p. 199, 211

∇^2, Δ: Beltrami Laplacian, de Rham Laplacian, p. 199, 212

curl, p. 199

$\partial_{\mathbf{n}}\varphi$: normal derivative of φ, p. 204

$*$: Hodge operator, p. 208

$^\flat, ^\sharp$: musical operators, p. 208

$\langle .|. \rangle$: bilinear symmetric form induced by the metric, p. 208

$\boldsymbol{\delta}$: codifferential, p. 210

$\langle .|. \rangle_2$: bilinear symmetric form, p. 210

$L^2\left(B; \mathrm{Alt}^k(B)\right)$: Hilbert space of k-forms, p. 210

$\mathbf{Z}_c^p(B;\mathbb{R}), \mathbf{B}_c^p(B;\mathbb{R})$: space of p-cocycles, of compactly supported p-coboundaries, p. 215

$\mathbf{H}_c^p(B;\mathbb{R})$: p-th compactly supported de Rham cohomology space, p. 215

$\mathbf{Z}'_p(B), \mathbf{B}'_p(B)$: space of p-cycles, of p-boundaries (for currents), p. 216

$\mathbf{H}'_p(B)$: p-th homology space (for currents), p. 216

$\mathbf{H}_p(B;\mathbb{R}) = \mathbf{H}_p^\infty(B;\mathbb{R})$: p-th homology space (for chains of smooth simplexes), p. 216

$b_p(B)$: p-th Betti number, p. 221

Chapter 6: Analysis on Lie Groups

$Z_X : q \mapsto q.X$: field of fundamental (or Killing) vectors associated with $X \in \mathrm{Lie}(\mathbf{G})$, p. 275

$\exp_{\mathbf{G}}$: exponential mapping, p. 280

Lie: functor $\mathbf{LieGrp} \to \mathbf{LieAl}$, p. 281

$\mathcal{F}, \overline{\mathcal{F}}$: Fourier transform, cotransform, p. 287, 300, 302

$\mathcal{C}_b^k(\mathbb{R}^n), \mathcal{C}_0^k(\mathbb{R}^n)$, p. 289

$\chi_a : t \mapsto \exp(2\pi i a.t)$, p. 287

$W^{k,p}(\mathbb{R}^n), W_0^{k,p}(\mathbb{R}^n)$: Sobolev spaces, p. 289

$\hat{q}_{s,m}$, p. 290

M^α, p. 289

$\mathcal{F}_{\mathbf{n}}$: normalized Fourier transform, p. 292

$\mathcal{O}_M(\mathbb{R}^n), \mathcal{O}'_c(\mathbb{R}^n)$: space of multipliers, of convolutors, p. 293

$v.p.$: Cauchy principal value, p. 295

$\mathbb{I} = [-1/2, 1/2]^n$: unit hypercube, p. 296

$\mathcal{E}_{\mathbb{I}}(\mathbb{R}^n), \mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$: space of C^∞ functions, of periodic distributions, p. 298

$\dot{U}$ (respectively $\dot{f}$): distribution (respectively function) on the torus associated with the periodic distribution (respectively function) U, p. 299

$\mathcal{S}(\mathbb{Z}^n)$ (respectively $\mathcal{S}'(\mathbb{Z}^n)$): space of rapidly decreasing (respectively slowly increasing) sequences, p. 299

q_m, p. 299

$\int_Z U_z.d\pi(z)$: continuous sum of unitary representations, p. 303

χ: unitary character, p. 304

$\mathbf{U}$: group of complex numbers with unit modulus, p. 304

$\widehat{\mathbf{G}}$: dual group or space, p. 304, 313

$\langle \widehat{x}, x \rangle$: duality bracket, p. 305

$x \mapsto x^*, f \mapsto \tilde{f}$: involution, p. 306

$A(\mathbf{G})$, p. 306

$\mathcal{S}'(\mathbf{G})$: space of tempered distributions on the group $\mathbf{G}$, p. 309

$\mathbf{H}^{\perp}$: orthogonal complement of the subgroup $\mathbf{H}$ of $\mathbf{G}$, p. 310

π: Plancherel measure, p. 313

$\mathfrak{F}\left(\widehat{\mathbf{G}}\right)$, p. 313

$\Theta_{\widehat{x}}$: character, p. 314

Chapter 7: Connections

$\Gamma_{ij}^{k} = \left\{ \begin{array}{c} k \\ ij \end{array} \right\}, \gamma_{ij}^{k}$: coefficients of the connection, Christoffel symbol of the second kind, p. 319, 324

$\omega_j^k, \underline{\omega}$, gauge potential, p. 319, 336

∇_j, ∇_Z: covariant derivative, p. 320, 324

g_{ij}, p. 320

$\Gamma_{pq,j} = [pq, j]$: Christoffel symbol of the first kind, p. 321

g^{ij}, p. 321

$\nabla, \mathbf{d}$: covariant differential of the linear connection $\mathbf{C}$, p. 323, 326, 330

∇_0: covariant differential of the trivial connection, p. 324

$a_{,j}^k, a_{:j}^k, a_{;j}^k$, p. 322

$\mathbf{r}$: curvature, p. 331

$R_{i,jk}^l$: Riemann–Christoffel tensor, p. 332

$\mathbf{t}, T_{ij}^k$: torsion, p. 332

$\underline{\sigma}$, p. 335

$\underline{\mathbf{\Omega}}_U, \Omega_i^l$: curvature 2-form, p. 339

$\underline{\mathbf{\Theta}}^U, \Theta^k$: torsion 2-form, p. 339

$R(B), R_{\mathbf{G}}(B)$: bundle of frames, of $\mathbf{G}$-frames, p. 334, 345, 351

$\rho_g: q \longmapsto q.g$: p. 346

H_q: space of horizontal vectors, p. 347

ω: connection 1-form of a principal connection $\mathbf{P}$, p. 347

$\widehat{c_q}$: horizontal lifting of the curve c, p. 349

$[\omega, \omega]$, p. 350

$\mathbf{D}$: covariant differential of a principal connection $\mathbf{P}$, p. 351

$\mathbf{\Omega}$: curvature 2-form, p. 352

σ: soldering 1-form, p. 353

$\mathbf{\Theta}$: torsion 2-form, p. 353

$\mathfrak{b}(\xi)$: basic vector field corresponding to ξ, p. 354

R_{ik}: components of the Ricci tensor $\mathbf{Ric}$, p. 363

R: Ricci curvature, p. 363

K: Gaussian curvature, p. 363

S_{ik}: components of the Einstein tensor $\mathbf{S}$, p. 363

1

Differential Calculus

1.1. Introduction

We could cite various forerunners of differential calculus, including Descartes, Fermat and Cavalieri, but Newton and Leibniz should be remembered as the true pioneers of the field. This dual paternity created terrible priority disputes where the only certainty is the complexity of the controversy [HAL 80]. Newton's "fluxions" and Leibniz's "vanishing quantities" are analogous to our modern concepts of derivative and differential, respectively. The extension of these ideas to functions of several variables was due to Euler (who introduced partial derivatives) and Clairaut. After contributions by many other mathematicians, including Volterra and Hadamard, the concept of a derivative of arbitrary order (Lemma-Definition 1.4) was ultimately introduced by Fréchet between 1909 and 1925. The inverse mapping theorem was proved by Lagrange in 1770, and the simplest case of the implicit function theorem was proved by Cauchy around 1833, followed by the case of vector-valued functions of several variables by Dini in 1877. Gateaux then extended Fréchet's earlier ideas to develop his concept of differential, which was presented in a posthumous publication in 1919. The "convenient" form of differentials, one of their most recent incarnations, was introduced by Frölicher, Kriegl and Michor in the early 1980s ([KRI 97], p. 73). These differentials were intended for mappings taking values in locally convex non-normable spaces, in particular nuclear Fréchet spaces (see, in particular, section 5.3.2 on manifolds of mappings). Wherever possible, this chapter therefore chooses to present differential calculus for mappings which take values in locally convex spaces rather than the normed vector spaces considered by the standard approach (nothing essential changes). Nonetheless, for the inverse mapping theorem and the implicit function theorem (Theorems 1.29 and 1.30), we will restrict attention to the Banach case (for a more general context, see [GLÖ 06], which uses yet another concept of differentiability distinct from any of those mentioned above). Both theorems are proved in full detail, including the Banach analytic case, by explicitly filling in the hints from [WHI 65]. Together with the Carathéodory theorems mentioned below, these two theorems are the most profound results of this chapter.

The classical Cauchy–Lipschitz conditions for the existence and uniqueness of solutions of ordinary differential equations were considerably weakened by Carathéodory ([CAR 27], section 576 and following) using measure and integration theory ([P2], section 4.1), which was an extremely recent development at the time. It was judged useful to give a full proof of Carathéodory's theorems in section 1.5.1, since the available literature seems to expect the reader to reassemble this proof from a scattered collection of isolated results and special cases, at the expense of clarity. The parameter dependence of solutions is studied in section 1.5.3. Again, proofs are given in full, with the exception of one result: the differentiability of solutions with respect to the initial condition when the classical hypothesis (Corollary 1.82) is replaced by Lusin's condition. This generalization is important (Remark 1.84) and the proof is not uninteresting ([ALE 87], Chapter 2, section 2.5.6); it is omitted from these pages not because it is too difficult but because it is too long.

1.2. Fréchet differential calculus

1.2.1. *General conventions*

The conventions presented below apply throughout the entire volume. The conventions listed under **(II)** are motivated by tensor calculus (Chapter 4).

(I) $\mathbb{K}$ denotes the field of real or complex numbers. Two elements ∞ and ω are adjoined to the set of integers $\mathbb{Z}$. The usual order relation on $\mathbb{Z}$ is extended by the convention $n < \infty < \omega$ for every integer $n \in \mathbb{Z}$, with $\infty + n = n + \infty = \infty$, $\omega + n = n + \omega = \omega$. We define $\mathbb{N}_{\mathbb{K}} = \mathbb{N} \cup \{\infty, \omega\}$ if $\mathbb{K} = \mathbb{R}$ and $\mathbb{N}_{\mathbb{K}} = \{0, \omega\}$ if $\mathbb{K} = \mathbb{C}$, and $\mathbb{N}_{\mathbb{K}}^{\times} = \mathbb{N}_{\mathbb{K}} - \{0\}$. See also the convention **(C)** in section 2.2.1**(II)**.

For $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$, recall that $\alpha! = \alpha_1!...\alpha_n!$. For $\alpha \in \mathbb{Z}^n$, $|\alpha| = \alpha_1 + ... + \alpha_n$.

The locally convex vector spaces considered below are defined over $\mathbb{K}$ and are always *Hausdorff*, with the exception of semi-normed spaces. We will use the Landau notation: let X be a topological space, a some point in X, $\mathbf{F}$ a locally convex space, and $f : X \to \mathbb{R}_+$, $\mathbf{g} : X \to \mathbf{F}$ two mappings; suppose that $f(x) > 0$ in some neighborhood of a. We write $\mathbf{g} = O(f)\ (x \to a)$ (respectively $\mathbf{g} = o(f)\ (x \to a)$) and say that $\mathbf{g}$ is dominated by f (respectively is negligible with respect to f) in the neighborhood of a if the function $\mathbf{g}/f$ is bounded in some neighborhood of a (respectively tends to 0 as the variable x tends to a). If there are several such mappings, we can write $O_1(f)$, $O_2(f)$ (respectively $o_1(f)$, $o_2(f)$), etc.

(II) Unless otherwise stated, the dual of a locally convex space $\mathbf{E}$ is denoted by $\mathbf{E}^{\vee}$[1]. Let (e_i) be a basis of the finite-dimensional vector space $\mathbf{E}$, $\left(e^{\vee i}\right)$ the dual basis, (e'_i) some other basis of $\mathbf{E}$ and $\left(e'^{\vee i}\right)$ its dual basis ([P1], section 3.1.3(**VI**)). In practice, we can unambiguously write $e_{i'}$ for $e'_{i'}$, $e^{\vee i'}$ for $e'^{\vee i'}$. Similarly, given a change-of-basis matrix $A = \left(A_i^{i'}\right)$ (where i ranges over the rows and i' ranges over the columns), we can write $A^{-1} = \left(A_{i'}^{i}\right) = \frac{1}{\det(A)}\alpha^T$, where α is the matrix of cofactors of A ([P1], section 2.3.11(**V**)); thus, $\sum_{i'} A_i^{i'} A_{i'}^{j} = \delta_i^j$[2]. Let $x = \sum_i x^i e_i \in \mathbf{E}$; $\mathbf{E}$ is typically a left $\mathbb{K}$-vector space and so the vector x can be represented by the *row* $\left(x^i\right)$ with respect to the basis (e_i) ([P1], section 3.1.3(**II**)). By contrast, $\mathbf{E}^{\vee}$ is typically a right $\mathbb{K}$-vector space; thus, if $x^{\vee} = \sum_i x_i^{\vee} e^{\vee i} \in \mathbf{E}^{\vee}$, the covector $x^{\vee}$ can be represented by the *column* $(x_i^{\vee})$ with respect to the basis $\left(e^{\vee i}\right)$ ([P1], section 3.1.3(**IV**)).

MEMORY AID:– Indices such as i, j, k refer to the old basis; primed indices such as i', j', k' refer to the new basis and are written as superscripts for the change-of-basis matrix A or as subscripts for its inverse A^{-1}. The indices of the components of a vector are usually written as superscripts; the indices of the components of a covector are usually written as subscripts. The indices of a sequence of vectors are always subscripts; the indices of a sequence of covectors are always superscripts. This is motivated by "Einstein's summation convention" (see Remark 4.2 in section 4.2.1).

If $x \in \mathbf{E}$, $x^{\vee} \in \mathbf{E}^{\vee}$, the duality bracket of these two vectors is written as $\langle x, x^{\vee}\rangle$. A change of basis in $\mathbf{E}$ and $\mathbf{E}^{\vee}$ may therefore be written as:

$$e_i = \sum_{j'} A_i^{j'} e_{j'}, \quad e_{j'} = \sum_j A_{j'}^{i} e_i,$$
$$e^{\vee i} = \sum_{j'} e^{\vee j'} A_{j'}^{i}, \quad e^{\vee j'} = \sum_i e^{\vee i} A_i^{j'}. \qquad [1.1]$$

1 We cannot write $e'_{i'}$ for the dual of $\mathbf{F}$ in this volume as we did in [P2] (and as is widely done in practice) without creating ambiguity. Instead, we will use the notation $\mathbf{F}^{\vee}$, which is borrowed from [LAN 99b].

2 This abuse of notation (strictly speaking, we should, for example, write $\tilde{A}_{j'}^{j}$ instead of $A_{j'}^{j}$) is perfectly standard in mathematical physics (e.g. [LIC 55]) and is not ambiguous if we remember that $A_{i'}^{i}$ is the matrix that transforms from a coordinate system indexed by letters like i to the primed coordinate system indexed by symbols like i'. This relaxed notation makes calculations much more straightforward.

Let $x = \sum_i x^i e_i \in \mathbf{E}$ and $x^\vee = \sum_i e^{\vee i} x_i^\vee \in \mathbf{E}'$. Then, $x = \sum_{j'} x^{j'} e_{j'}$ and $x^\vee = \sum_{j'} e^{\vee j'} x_{j'}^\vee$, where:

$$x_i^\vee = \sum_{j'} A_i^{j'} x_{j'}^\vee, \quad x_{j'}^\vee = \sum_i A_{j'}^i x_i^\vee,$$

$$x^i = \sum_{j'} x^{j'} A_{j'}^i, \quad x^{j'} = \sum_i x^i A_i^{j'}. \qquad [1.2]$$

Since $\mathbb{K}$ is commutative, we do not need to distinguish between left $\mathbb{K}$-vector spaces and right $\mathbb{K}$-vector spaces. Therefore, the duality bracket of $x \in \mathbf{E}$, $x^\vee \in \mathbf{E}^\vee$ can also be written as $\langle x^\vee, x \rangle$. For example, if $f : \Omega \to \mathbb{K}$ is a differentiable function in some non-empty open subset Ω of $\mathbf{E}$, then, at each point a of Ω, we have $Df(a) \in \mathbf{E}^\vee$ (see Lemma-Definition 1.4). If $\mathbf{h}_a \in \mathbf{E}$, it might seem more convenient to write $\langle Df(a), \mathbf{h}_a \rangle$ instead of $\langle \mathbf{h}_a, Df(a) \rangle$ for the quantity $Df(a).\mathbf{h}_a$, but the reverse notation is also perfectly justifiable (section 2.2.4(**IV**)).

(III) Recall the following fact ([P2], section 3.9.3(**II**)): let $\mathbf{E}_1, ..., \mathbf{E}_n$ be normed vector spaces, each equipped with a norm $|.|$, and suppose that $\mathbf{F}$ is a semi-normed vector space equipped with a semi-norm $|.|_\gamma$. The space $\mathcal{L}(\mathbf{E}_1, ..., \mathbf{E}_n; \mathbf{F})$ of continuous n-linear mappings from $\mathbf{E}_1 \times ... \times \mathbf{E}_n$ into $\mathbf{F}$ is a semi-normed vector space when equipped with the semi-norm $\|.\|_\gamma$ defined for any mapping $\mathbf{u} \in \mathcal{L}(\mathbf{E}_1, ..., \mathbf{E}_n; \mathbf{F})$ by:

$$\boxed{\|\mathbf{u}\|_\gamma = \sup_{|x_1|,...,|x_n| \leq 1} |\mathbf{u}(x_1, ..., x_n)|_\gamma.} \qquad [1.3]$$

If $\mathbf{F}$ is a *Hausdorff* locally convex space whose topology is induced by the family of semi-norms $\left(|.|_\gamma\right)_{\gamma \in \Gamma}$ ([P2], section 3.3.3), then $\mathcal{L}(\mathbf{E}_1, ..., \mathbf{E}_n; \mathbf{F})$ is a *Hausdorff* locally convex space when equipped with the family of semi-norms $\left(\|.\|_\gamma\right)_{\gamma \in \Gamma}$. This space is quasi-complete whenever $\mathbf{F}$ is quasi-complete ([P2], section 3.4.2(**II**)).

REMARK 1.1.– *If $\mathbf{F}$ is a normed vector space with norm $|.|$, the index γ and the set Γ have only to be omitted. This simplification only causes problems in sections 1.3.1 and 1.3.3. In sections 1.2.4 and 1.2.5, read "Banach space" instead of "quasi-complete locally convex space" if the simplification has been made.*

Write $\mathcal{L}_n(\mathbf{E}; \mathbf{F})$ for the space $\mathcal{L}(\mathbf{E}_1, ..., \mathbf{E}_n; \mathbf{F})$ whenever $\mathbf{E}_i = \mathbf{E}$ for all $i \in \{1, ..., n\}$. An element $\mathbf{u}$ of $\mathcal{L}_n(\mathbf{E}; \mathbf{F})$ is said to be *symmetric* if, for all $\left(h^1, ..., h^n\right) \in \mathbf{E}^n$ and every permutation $\sigma \in \mathfrak{S}_n$, we have $\mathbf{u}(h^1, ..., h^n) = \mathbf{u}(h^{\sigma(1)}, ..., h^{\sigma(n)})$. The set of symmetric elements of $\mathcal{L}_n(\mathbf{E}; \mathbf{F})$ is written as $\mathcal{L}_{n,s}(\mathbf{E};\mathbf{F})$ and is a vector subspace of $\mathcal{L}_n(\mathbf{E};\mathbf{F})$; this subspace is equipped with the family of semi-norms $\left(\|.\|_\gamma\right)_{\gamma\in\Gamma}$. If $\mathbf{h}\in\mathbf{E}$ and $\mathbf{u}\in\mathcal{L}_{n,s}(\mathbf{E};\mathbf{F})$, set:

$$\mathbf{u}.\mathbf{h}^n = \mathbf{u}.\left(\underbrace{\mathbf{h},...,\mathbf{h}}_{n\text{ terms}}\right).$$

If $\mathbf{E}$ is finite-dimensional, then we have the following canonical isomorphism:

$$\mathcal{L}_{n,s}(\mathbf{E};\mathbf{F}) \cong \operatorname{Hom}_{\mathbb{K}}\left(\underbrace{\mathbf{E}\otimes...\otimes\mathbf{E}}_{n\text{ terms}};\mathbf{F}\right). \qquad [1.4]$$

(IV) See also the conventions **(C1)** (section 2.2.1**(II)**, p. 55), **(C2)** (section 2.2.2**(III)**, p. 58), (**C3**) (section 6.2.1(**II**), p. 253).

1.2.2. *Fréchet differential*

(I) Let $\mathbf{E},\mathbf{F}$ be two locally convex spaces and ϕ a mapping from U into $\mathbf{F}$, where U is some neighborhood of 0 in $\mathbf{E}$.

DEFINITION 1.2.– *We say that ϕ is tangent to 0 if, for every neighborhood W of 0 in $\mathbf{F}$, there exists a balanced neighborhood $V\subset U$ of 0 in $\mathbf{E}$ such that, for every $t\in\mathbb{K}$ with sufficiently small $|t|$, we have $\phi(tV)\subset o(t)\,W$ (where $o:\mathbb{K}\to\mathbb{K}$).*

LEMMA 1.3.– *i) If ϕ is linear and tangent to 0, then $\phi=0$*

ii) If $\mathbf{E}$ and $\mathbf{F}$ are normed vector spaces with norm $|.|$, then ϕ is tangent to 0 if and only if $|\phi(x)| = o(|x|)$.

PROOF.– (i) With the notation of Definition 1.2, we have $\phi(V)\subset\frac{o(t)}{t}W$, so $\phi(V)\subset\{0\}$ and $\phi=0$. (ii): **exercise**. ■

From this lemma, we immediately deduce the following claim (ii):

LEMMA-DEFINITION 1.4.– *Let a be some point of $\mathbf{E}$, U some neighborhood of 0 in $\mathbf{E}$, and $f\colon U+a\to\mathbf{F}$ some mapping where $a+U:=\{a+x : x\in U\}$.*

i) We say that $\mathbf{f}$ is (Fréchet) differentiable *at a if there exists a continuous linear mapping $\mathbf{L}\in\mathcal{L}(\mathbf{E};\mathbf{F})$ such that the mapping $\mathbf{h}\mapsto\mathbf{f}(a+\mathbf{h})-\mathbf{f}(a)-\mathbf{L}.\mathbf{h}$ (defined in U) is tangent to 0.*

ii) This mapping $\mathbf{L}$ is uniquely determined.

iii) This mapping is written as $\mathbf{L} = D\mathbf{f}(a)$ *(or* $\mathbf{L} = d_a\mathbf{f}$ *or* $\mathbf{L} = \mathbf{f}'(a)$*) and is called the (Fréchet)* differential *of* $\mathbf{f}$ *at the point* a*.*

The notion of the Fréchet differential is especially fruitful when $\mathbf{E}$ is a normed vector space. In this case, if A is an open subset of $\mathbf{E}$, $\mathbf{F}$ is a locally convex space, a is a point of A, and $\mathbf{f} : A \to \mathbf{F}$ is a mapping, then $\mathbf{f}$ is differentiable at a with differential $D\mathbf{f}(a)$ at this point if and only if

$$\lim_{|\mathbf{h}|\to 0, |\mathbf{h}|\neq 0} \frac{\mathbf{f}(a+\mathbf{h}) - \mathbf{f}(a) - D\mathbf{f}(a).\mathbf{h}}{|\mathbf{h}|} = 0.$$

The canonical isomorphism $\mathcal{L}(\mathbb{K};\mathbf{F}) \cong \mathbf{F} : \mathbf{u} \mapsto \mathbf{u}.1$ ([P2], section 3.3.8) reintroduces the usual notion of derivative, including the notion of a "complex derivative" when $\mathbb{K} = \mathbb{C}$ ([P2], section 4.2.1, Theorem-Definition 4.45). Indeed, keeping the same notation as before, we obtain:

COROLLARY-DEFINITION 1.5.– *If* $\mathbf{E} = \mathbb{K}$*, the function* $\mathbf{f} : U + a \to \mathbf{F}$ *is differentiable at the point* $a \in \mathbb{K}$ *(in the above sense) if and only if it is differentiable at* a *(in the usual sense of differentiability of functions of one real or complex variable), in which case the element* $\dot{\mathbf{f}}(a) := D\mathbf{f}(a).1$ *of* $\mathbf{F}$ *is the* derivative *of* $\mathbf{f}$ *at* a*.*

THEOREM 1.6.– (Rolle) *Let* $[a,b]$ *be a compact interval of* $\mathbb{R}$ *with non-empty interior and suppose that* $\varphi : [a,b] \to \mathbb{R}$ *is a continuous function in* $[a,b]$ *that is differentiable in* $]a,b[$ *and which satisfies* $\varphi(a) = \varphi(b)$*. Then, there exists* $c \in]a,b[$ *such that* $\dot{\varphi}(c) = 0$*.*

PROOF.– If φ is constant in $[a,b]$, then it is differentiable and has zero derivative on $]a,b[$. Otherwise, it attains its lower and upper bounds ([P2], section 2.3.7, Theorem 2.42). One of these values is attained at some point $c \in]a,b[$, which must therefore satisfy $\dot{\varphi}(c) = 0$ (**exercise**). ■

If $\mathbf{f}$ is differentiable at the point a, then it is also continuous at this point (**exercise**). Let A be a non-empty open subset of $\mathbf{E}$ and write $\mathcal{D}_a(A;\mathbf{F})$ for the set of mappings from A into $\mathbf{F}$ that are differentiable at the point $a \in A$. The mapping $\mathcal{D}_a(A;\mathbf{F}) \to \mathcal{L}(\mathbf{E};\mathbf{F}) : \mathbf{f} \mapsto D\mathbf{f}(a)$ is known as the *differentiation operator* at the point a and is $\mathbb{K}$-linear.

DEFINITION 1.7.– *Let* $\mathbf{f} \in \mathcal{D}_a(A;\mathbf{F})$*. The rank* $rk(D\mathbf{f}(a))$ *([P1], section 3.1.10, Definition 3.38(ii)) is said to be the* rank *of* $\mathbf{f}$ *at the point* a *and is written as* $rk_a(\mathbf{f})$*.*

(II) Let $\mathbf{E}$ be a *real* Hilbert space, A some non-empty open subset of $\mathbf{E}$, and $f : A \to \mathbb{R}$ a function that is differentiable at the point $a \in A$. Then, $Df(a) \in \mathbf{E}^{\vee}$. By Riesz's theorem ([P2], section 3.10.2**(IV)**, Theorem 3.15(1)), there exists some uniquely determined element $x^* \in \mathbf{E}$ such that $Df(a)$ coincides with the linear form $h \mapsto \langle x^*|h\rangle$.

DEFINITION 1.8.– *The element* x^* *specified above is written as* $\operatorname{grad}_a(f)$ *or* $\nabla_a f$ *and is called the* gradient *of* f *at the point* a. *It satisfies* $\langle \nabla_a f | h \rangle := Df(a).h$ *for all* $h \in \mathbf{E}$.

(III) Some of the classical results of differential calculus ([DIE 93], Volume 1, Chapter 8) are reproduced below. A few are stated in a slightly more general form than the cited reference; however, the reasoning is identical and entirely straightforward in each case.

THEOREM 1.9.– (chain rule) *Let* $\mathbf{E}, \mathbf{F}$ *be normed vector spaces,* $\mathbf{G}$ *a locally convex space,* A *some non-empty open subset of* $\mathbf{E}$, a *some point of* A, B *an open subset of* $\mathbf{F}$ *containing* $\mathbf{f}(a)$, $\mathbf{f} \in \mathcal{D}_a(A;\mathbf{F})$, *and* $\mathbf{g} \in \mathcal{D}_{\mathbf{f}(a)}(B;\mathbf{G})$. *Then,* $\mathbf{g} \circ \mathbf{f} \in \mathcal{D}_a(A;\mathbf{G})$ *and*

$$D(\mathbf{g} \circ \mathbf{f})(a) = D\mathbf{g}(\mathbf{f}(a)) \circ D\mathbf{f}(a). \qquad [1.5]$$

EXAMPLE 1.10.– *Let* $\mathbf{E}$ *be a Banach space. The set* $\mathfrak{H} \subset \mathcal{L}(\mathbf{E})$ *of continuous linear bijections is open in* $\mathcal{L}(\mathbf{E})$ *([P2], section 3.4.1(**II**), Lemma 3.50). The mapping* $\mathbf{u} \mapsto \mathbf{u}^{-1}$ *from* $\mathfrak{H}$ *onto* $\mathfrak{H}$ *is differentiable, and its differential at the point* $\mathbf{u}_0$ *is the continuous linear mapping* $\mathbf{s} \mapsto -\mathbf{u}_0^{-1} \circ \mathbf{s} \circ \mathbf{u}_0$ *(from* $\mathcal{L}(\mathbf{E})$ *into* $\mathcal{L}(\mathbf{E})$*). The reader may wish to prove this result as an* ***exercise*** *or refer to Lemma 1.24 of section 1.2.5(**II**).*

(IV) Let $\mathbf{E}_1, ..., \mathbf{E}_n$ be normed vector spaces. Then, $\mathbf{E} = \mathbf{E}_1 \times ... \times \mathbf{E}_n$ can be canonically equipped with the structure of a normed vector space ([P2], section 3.4.1(**I**)). Let A be a non-empty open subset of $\mathbf{E}$ and suppose that $a = \left(a^1, ..., a^n\right) \in A$. Suppose further that $\mathbf{F}$ is a locally convex space. If $\mathbf{f} : A \to \mathbf{F} : \left(x^1, ..., x^n\right) \mapsto \mathbf{f}\left(x^1, ..., x^n\right)$ is differentiable at the point a, then the mapping

$$\mathbf{h}_i \mapsto \mathbf{f}\left(a^1, ..., a^{i-1}, a^i + \mathbf{h}^i, a^{i+1}, ..., a^n\right)$$

is defined in some open neighborhood of 0 in $\mathbf{E}_i$ and differentiable at 0 for all $i \in \{1, ...n\}$ (**exercise**); its differential at this point is written as $D_i\mathbf{f}(a)$ (or $d_i\mathbf{f}(a)$ or $\frac{\partial \mathbf{f}}{\partial x^i}(a)$).

DEFINITION 1.11.– *The mapping* $D_i\mathbf{f}(a)$ *is called the* i*-th* partial differential *of* $\mathbf{f}$ *at the point* a.

Given the conditions stated above, the differential $D\mathbf{f}(a)$ can be expressed in terms of the partial differentials $D_i\mathbf{f}(a)$ as follows:

$$D\mathbf{f}(a).\mathbf{h} = \sum\nolimits_{i=1}^{n} D_i\mathbf{f}(a).\mathbf{h}^i, \qquad [1.6]$$

where $\mathbf{h} = \left(\mathbf{h}^1, ..., \mathbf{h}^n\right)$. If $\mathbf{E}_i = \mathbb{K}$, the element $\partial_i \mathbf{f}(a) = D_i \mathbf{f}(a).1 \in \mathbf{F}$ is called the i-th *partial derivative* of $\mathbf{f}$ at the point a; writing $(\mathbf{e}_i)_{1 \leq i \leq n}$ for the canonical basis of $\mathbb{K}^n$, we have

$$\partial_i \mathbf{f}(a) = D_i \mathbf{f}(a).1 \in \mathbf{F}. \qquad [1.7]$$

If $\mathbb{K} = \mathbb{R}$, the existence of the partial differentials $D_i \mathbf{f}(a)$ $(i = 1, ..., n)$ does not imply the existence of the differential $D\mathbf{f}(a)$ (however, see Theorem 1.16(ii)).

If $\mathbf{E}_j = \mathbb{K}$ for all $j \in \{1, ..., n\}$ and $\mathbf{F} = \mathbb{K}^m$, then the linear mapping $D\mathbf{f}(a)$ can be represented with respect to the canonical bases by the *Jacobian matrix* $\left[\partial_i f^j(a)\right]_{1 \leq i \leq m, 1 \leq j \leq n}$, where $\mathbf{f} = \left(f^1, ..., f^n\right)$. If $n = m$, the determinant of this matrix is called the *Jacobian* of $\mathbf{f}$ at the point a and is written as $\frac{\partial\left(f^1, ..., f^n\right)}{\partial(x^1, ..., x^n)}(a)$.

(V) In the following, $\mathbb{K} = \mathbb{R}$. Let $\mathbf{F}$ be a locally convex space and $|.|_\gamma$ a continuous semi-norm on $\mathbf{F}$.

THEOREM 1.12.– (mean value theorem, 1st version) *Let $I = [\alpha, \beta]$ be a compact interval of $\mathbb{R}$, $\mathbf{f}$ a continuous mapping from I into $\mathbf{F}$ and φ_γ a continuous mapping from I into $\mathbb{R}$. Suppose that there exists a countable subset $D \subset \mathring{I}$ such that $\mathbf{f}$ and φ_γ both have a derivative at every point $\xi \in D$ and $\left|\dot{\mathbf{f}}(\xi)\right|_\gamma \leq \dot{\varphi}_\gamma(\xi)$. Then, $|\mathbf{f}(\beta) - \mathbf{f}(\alpha)|_\gamma \leq \varphi_\gamma(\beta) - \varphi_\gamma(\alpha)$.*

PROOF.– This is a classical result[3] when $\mathbf{F}$ is a normed vector space with norm $|.|$; we simply need to replace $|.|$ par $|.|_\gamma$ and φ by φ_γ. ■

Let $\mathbf{E}$ be a normed vector space with norm $|.|$. With the notation of [1.3] (with $n = 1$), we have the following result:

THEOREM 1.13.– (mean value theorem, 2nd version) *Let $\mathbf{f}$ be a mapping taking values in $\mathbf{F}$ that is defined and continuous in some neighborhood of the closed segment $S = [a, a + \mathbf{h}]$ joining the points a, $a + \mathbf{h}$ of $\mathbf{E}$ ([P2], section 3.3.1). Suppose that $\mathbf{f}$ is differentiable at every point of S.*

i) Given any subset Θ of $[0, 1]$ with countable complement in $[0, 1]$,

$$|\mathbf{f}(a + \mathbf{h}) - \mathbf{f}(a)|_\gamma \leq |\mathbf{h}| . \sup_{t \in \Theta} \|D\mathbf{f}(a + t.\mathbf{h})\|_\gamma .$$

ii) Let $\mathbf{L} \in \mathcal{L}(\mathbf{E}; \mathbf{F})$. Given any subset Ξ of $[a, a + \mathbf{h}]$ with countable complement in $[a, a + \mathbf{h}]$, we have $|\mathbf{f}(a + \mathbf{h}) - \mathbf{f}(a) - \mathbf{L}.\mathbf{h}|_\gamma \leq M_\gamma . |\mathbf{h}|$ with $M_\gamma = \sup_{x \in \Xi} \|D\mathbf{f}(x) - \mathbf{L}\|_\gamma$. In particular, if $M_\gamma = \sup_{x \in \Xi} \|D\mathbf{f}(x) - D\mathbf{f}(a)\|_\gamma$, then

$$|\mathbf{f}(a + h) - \mathbf{f}(a) - D\mathbf{f}(a).\mathbf{h}|_\gamma \leq M_\gamma . |\mathbf{h}| . \qquad [1.8]$$

3 See [DIE 93], Volume 1, (8.5.1) or the Wikipedia article on *Mean value theorem*.

iii) Let $\mathbf{F} = \mathbb{K} = \mathbb{R}$*. There exists* $\theta \in \left]0,1\right[$ *such that*

$$\mathbf{f}(a+\mathbf{h}) - \mathbf{f}(a) = D\mathbf{f}(a+\theta.\mathbf{h}).\mathbf{h}.$$

PROOF.– i) Let $\mathbf{g} : [0,1] \to \mathbf{F} : t \mapsto \mathbf{f}(a+t.\mathbf{h})$. By Theorem 1.9, $\dot{\mathbf{g}}(t) = D\mathbf{f}(a+t.\mathbf{h}).\mathbf{h}$, so $|\dot{\mathbf{g}}(t)|_\gamma \leq N_\gamma.|\mathbf{h}|$, where $N_\gamma = \sup_{t\in\Theta} \|D\mathbf{f}(a+t.\mathbf{h})\|_\gamma$. The upper bound of (i) is therefore obtained by applying Theorem 1.12 with $\mathbf{f}$ replaced by $\mathbf{g}$ and $\varphi_\gamma(t) = N_\gamma.t.|\mathbf{h}|$.

ii) Can be deduced by applying (i) to the function $\xi \mapsto \mathbf{f}(\xi) - \mathbf{L}.(\xi - x)$.

iii) Let $g : [0,1] \to \mathbb{R} : t \mapsto \mathbf{f}(a+t.\mathbf{h}) - \mathbf{f}(a) - (f(a+\mathbf{h} - f(a)).t$. Since $g(0) = g(1) = 0$, Rolle's theorem (Theorem 1.6) implies the existence of a number $\theta \in \left]0,1\right[$ such that $\dot{g}(\theta) = 0$. ■

Theorem 1.13 has the following two corollaries:

COROLLARY 1.14.– *Let* A *be a* connected *open subset of* $\mathbf{E}$ *and suppose that* $\mathbf{f} : A \to \mathbf{F}$ *is differentiable in* A*. If the differential* $D\mathbf{f} : x \mapsto D\mathbf{f}(x)$ *is zero in the complement of a countable subset of* A*, then* $\mathbf{f}$ *must be constant in* A*.*

COROLLARY 1.15.– *Let* $\left(|.|_\gamma\right)_{\gamma\in\Gamma}$ *be a family of semi-norms that induce the topology of* $\mathbf{F}$, A *some* convex *open subset of* $\mathbf{E}$ *and* $\mathbf{f} : A \to \mathbf{F}$ *a differentiable mapping in* A*. If, for all* $\gamma \in \Gamma$*, there exists a real number* $k_\gamma > 0$ *such that* $\sup_{x\in A} \|D\mathbf{f}(x)\|_\gamma \leq k_\gamma$*, then* $\mathbf{f}$ *is Lipschitz ([P2], section 2.4.3(**II**)) and hence uniformly continuous.*

PROOF.– Let $x', x'' \in A$. Since A is convex, the segment $[x', x'']$ is contained in A, so by Theorem 1.13(i) $|\mathbf{f}(x') - \mathbf{f}(x'')|_\gamma \leq k_\gamma |x' - x''|$ for all $\gamma \in \Gamma$, which means that $\mathbf{f}$ is Lipschitz. ■

1.2.3. *Mappings of class* C^p

In this section, $\mathbb{K} = \mathbb{R}$.

(I) Let $\mathbf{E}$ be a normed vector space and suppose that $\mathbf{F}$ is a locally convex space. Write $\mathcal{L}_n(\mathbf{E};\mathbf{F})$ for the space of continuous n-linear mappings from $\mathbf{E}^n$ into $\mathbf{F}$ equipped with the family of semi-norms [1.3].

Let A be a non-empty open subset of $\mathbf{E}$ and suppose that $\mathbf{f} : A \to \mathbf{F}$ is a mapping. We say that $\mathbf{f}$ is of class C^0 if it is continuous in A. We say that it is of class C^1 if it is differentiable in A and its differential $D\mathbf{f} : A \to \mathcal{L}(\mathbf{E};\mathbf{F})$ is continuous. We say that $\mathbf{f}$ is twice differentiable (respectively is of class C^2) in A if its differential

$D\mathbf{f}$ is differentiable (respectively is of class C^1) in A. At any given point $a \in A$, the second differential $D(D\mathbf{f})(a)$, written as $D^2\mathbf{f}(a)$, is an element of $\mathcal{L}(\mathbf{E};\mathcal{L}(\mathbf{E};\mathbf{F}))$; this space can be identified with $\mathcal{L}_2(\mathbf{E};\mathbf{F})$ by ([P2], section 3.9.3**(II)**). Similarly, we may recursively define a p-times differentiable mapping (respectively a mapping of class C^p) in A for any integer $p \geq 1$, as well as the p-th order differential $D^p\mathbf{f}(a) \in \mathcal{L}_p(\mathbf{E};\mathbf{F})$ for any point $a \in A$. We say that $\mathbf{f}$ is of class C^∞ if $\mathbf{f}$ is of class C^p for every $p \geq 1$. We write $\mathcal{C}^p(A;\mathbf{F})$ for the $\mathbb{K}$-vector space of mappings of class C^p from A into $\mathbf{F}$ $(0 \leq p \leq \infty)$.

From ([DIE 93], Volume 1, (8.12.1), (8.16.6)), *mutatis mutandis*, we have the following result:

THEOREM 1.16.– *Let* $\mathbf{F}$ *be a locally convex space.*

i) (Schwarz's theorem) *Let* $\mathbf{E}$ *be a normed vector space,* A *some non-empty open subset of* $\mathbf{E}$ *and* $\mathbf{f}: A \to \mathbf{F}$ *a* p*-times differentiable mapping* $(p \geq 2)$ *at some point* $a \in A$*. Then, the differential* $D^p\mathbf{f}(a)$ *is symmetric and therefore belongs to* $\mathcal{L}_{p,s}(\mathbf{E};\mathbf{F})$.

ii) Let $\mathbf{E}_1, ..., \mathbf{E}_n$ *be normed vector spaces* $(n \geq 2)$ *and suppose that* A *is a non-empty subset of* $\mathbf{E}_1 \times ... \times \mathbf{E}_n$*. For any* p *such that* $1 \leq p \leq \infty$*,* $\mathbf{f}$ *is of class* C^p *in* A *if and only if the* n^p *partial derivatives of order* p *of* $\mathbf{f}$ *exist and are continuous in* A *(where* $n^\infty := \infty$*).*

By Theorem 1.16(ii), a mapping f from a non-empty open subset Ω of $\mathbb{R}^n$ into $\mathbb{C}$ is of class C^p if and only if its n^p partial derivatives exist and are continuous, i.e. $f \in \mathcal{E}^{(p)}(\Omega)$ ([P2], section 4.3.1**(I)**).

(II) In the context of Theorem 1.16(i), when $\mathbf{E} = \mathbb{K}^n$, introducing "symbolic powers" gives us an easy way to express $D^p\mathbf{f}(a).\mathbf{h}^p$ (with the notation of section 1.2.1**(III)**) in terms of the partial derivatives of order p:

$$\partial_{i_1}...\partial_{i_p}\mathbf{f}(a) := D_{i_1}...D_{i_p}\mathbf{f}(a).\left(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_p}\right),$$

where D_i denotes the partial differential in the i-th variable. Writing $\mathbf{h} = \left(h^1, ..., h^n\right)$, we have $D^2\mathbf{f}(a).\mathbf{h}^2 = \sum_{1\leq i,j\leq n} \partial_i\partial_j\mathbf{f}(a).h^i h^j$[4]. This can be expressed as $\left(\sum_{1\leq i\leq n} \partial_i\mathbf{f}(a).h^i\right)^{[2]}$ by developing the latter like any other squared

4 Here, i, j are indices and $\partial_{i_1}...\partial_{i_p}\mathbf{f}(a)$ denotes a partial derivative. However, in some contexts, this symbol is used by abuse of notation to denote a partial *differential* (see section 1.2.2**(IV)**).

parentheses and replacing $\left(\partial_i \mathbf{f}(a).h^i\right)\left(\partial_j \mathbf{f}(a).h^j\right)$ by $\partial_i\partial_j \mathbf{f}(a).h^i h^j$. With the same conventions, we can continue inductively to obtain the following result:

$$D^p \mathbf{f}(a).\mathbf{h}^p = \left(\sum_{1\leq i\leq n} \partial_i \mathbf{f}(a).h^i\right)^{[p]}.$$

If $\mathbf{E}_i = \mathbb{K}$ $(i = 1, ..., n)$, $\mathbf{F} = \mathbb{K}$, and $f : A \to \mathbb{K}$ is twice differentiable at some point a of A, then the matrix $Hf(a) = (\partial_i\partial_j f(a))_{1\leq i,j\leq n}$ is called the *Hessian matrix*[5] of f at the point a. Identifying the vectors of $\mathbb{K}^n$ with the columns of this matrix gives $D^2 f(a).\mathbf{h}^2 = \mathbf{h}^T.Hf(a).\boldsymbol{h}$.

(III) Suppose that $1 \leq p \leq \infty$. The following results can be shown by induction (**exercise**). The composition of two mappings of class C^p is also of class C^p. If $\mathbf{E}$ is a normed vector space, A some non-empty open subset of $\mathbf{E}$, $\mathbf{F}$ and $\mathbf{G}$ locally convex spaces, $\mathbf{u}$ a continuous linear mapping from $\mathbf{F}$ into $\mathbf{G}$ and $\mathbf{f} : A \to \mathbf{F}$ a mapping of class C^p in A, then $\mathbf{u} \circ \mathbf{f}$ is of class C^p in A and $D^p(\mathbf{u} \circ \mathbf{f}) = \mathbf{u} \circ D^p\mathbf{f}$. With the same hypotheses on $\mathbf{E}$ and A, if $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{G}$ are locally convex spaces, $[.,.]$ is a continuous bilinear mapping from $\mathbf{F}_1 \times \mathbf{F}_2$ into $\mathbf{G}$, and $\mathbf{f}_i$ is a mapping of class C^p from A into $\mathbf{F}_i$ for $i = 1, 2$, then $[\mathbf{f}_1, \mathbf{f}_2]$ is of class C^p from A into $\mathbf{G}$ and the so-called Leibniz rule holds:

$$D[\mathbf{f}_1, \mathbf{f}_2] = [D\mathbf{f}_1, \mathbf{f}_2] + [\mathbf{f}_1, D\mathbf{f}_2]. \qquad [1.9]$$

(IV) Nemytskii operator Let $\mathbf{F}$ be a Banach space, J a compact interval of $\mathbb{R}$ with non-empty interior, and $\mathcal{C}^p(J; \mathbf{F})$ the space of mappings of class C^p $(p \geq 1)$ from J into $\mathbf{F}$. Then, $\mathcal{C}^p(J; \mathbf{F})$ is a Banach space when equipped with the norm

$$\|\varphi\|_p = \sup_{0\leq k\leq p, t\in J} \left|\varphi^{(k)}(t)\right|$$

(**exercise**). Let U be an open neighborhood of 0 in $\mathbf{F}$, $\mathbf{f} : U \times J \to \mathbf{F}$ a mapping of class C^p and $\mathcal{N} : \mathcal{U} \times \mathring{J} \to \mathbf{F}$ the so-called *Nemytskii* operator, defined by $\mathcal{N}(\varphi, t) = \mathbf{f}(\varphi(t), t)$, where $\mathcal{U} := \left\{\varphi \in \mathcal{C}^p(J; \mathbf{F}) : \varphi(t) \in U, \forall t \in \mathring{J}\right\}$.

THEOREM 1.17.– *The Nemytskii operator $\mathcal{N}$ is of class C^p and*

$$D\mathcal{N}(\varphi, t).(h, \tau) = D_1\mathbf{f}(\varphi(t), t).(h(t) + \dot{\varphi}(t).\tau) + D_2\mathbf{f}(\varphi(t), t).\tau. \qquad [1.10]$$

---

5 In this Hessian matrix, i ranges over the rows and j ranges over the columns or vice versa by Theorem 1.16(i).

PROOF.– The operator $\mathcal{N}$ is the composition

$$\mathcal{U} \times \overset{\circ}{J} \overset{\mathrm{ev}\times 1_{\overset{\circ}{J}}}{\longrightarrow} \mathbf{F} \times \overset{\circ}{J} \overset{\mathbf{f}}{\longrightarrow} \mathbf{F},$$

where ev : $C^p(J;\mathbf{F}) \times \overset{\circ}{J} \to \mathbf{F}$ is the *evaluation operator* defined by $\mathrm{ev}(\varphi, t) = \varphi(t)$. But ev is of class C^p and

$$D\mathrm{ev}(\varphi, t) \cdot (h, \tau) = h(t) + \dot{\varphi}(t) \cdot \tau \qquad [1.11]$$

(**exercise***: see [ABR 83], Proposition 2.4.17). Since **f** is of class C^p, so is $\mathcal{N}$ by **(III)**, and [1.10] follows from [1.5]. ■

REMARK 1.18.– *The operator $\mathcal{N}$ defined above is stated slightly more generally than the classical Nemytskii operator, which does not allow t to vary (to recover the classical case, set the increase τ to 0 in [1.10]).*

1.2.4. *Taylor's formulas*

Throughout this section, $\mathbb{K} = \mathbb{R}$, and $\mathbf{F}$ denotes a quasi-complete locally convex space. Readers who are only interested in mappings taking values in a Banach space (see Remark 1.1) may skip **(I)** and **(II)** below.

(I) MACKEY CONVERGENCE Let B be a balanced convex subset (also known as an "absolutely convex" subset ([KÖT 79], Volume 1, section 16.1(2))) that is closed and bounded in $\mathbf{F}$, and write $\mathbf{F}_1 = \bigcup_{n=1}^{\infty} nB$. The gauge of p_B of B in $\mathbf{F}_1$ ([P2], section 3.3.2(**I**)) is a norm on $\mathbf{F}_1$ (**exercise**) and $\mathbf{F}_1$ is a Banach space, written as $\mathbf{F}_B$, when equipped with this norm.

DEFINITION 1.19.– *We say that a sequence (x_n) of elements of $\mathbf{F}$ is Mackey convergent to 0 (or locally convergent to 0 ([KÖT 79], Volume 1, section 28.3)), if there exists a balanced, closed and bounded subset B containing 0 and all the x_n such that (x_n) converges to 0 in $\mathbf{F}_B$.*

Mackey convergence of a sequence of elements of $\mathbf{F}$ implies the usual notion of convergence and is equivalent to it whenever $\mathbf{F}$ is a Fréchet space or the strong dual of an infrabarreled Schwartz space ([HOG 71], p. 8, Example 3), such as the distribution spaces studied in ([P2], section 4.4.1).

(II) GENERALIZED RIEMANN INTEGRAL To state Taylor's formulas in a general framework, it will be useful to introduce the generalized Riemann integral of a function of a real variable taking values in a quasi-complete locally convex space[6].

6 We will use the terminology adopted in [KRI 97]. It would be more true to the history of mathematics to call it the *generalized Cauchy integral*.

LEMMA-DEFINITION 1.20.– *Let* $\mathbf{F}$ *be a quasi-complete locally convex space,* I *a non-empty open interval of* $\mathbb{R}$*,* c *some point of* I *and* $\mathbf{f} : I \to \mathbf{F}$ *a continuous mapping.*

i) There exists a unique differentiable mapping $\int \mathbf{f} : I \to \mathbf{F}$ *such that* $\left(\int \mathbf{f}\right)(c) = 0$ *and* $\left(\int \mathbf{f}\right)' = \mathbf{f}$.

ii) Let $a, b \in I$*. Write* $\int_a^b \mathbf{f}(t)\,dt = \left(\int \mathbf{f}\right)(b) - \left(\int \mathbf{f}\right)(a)$*; this quantity is called the* (generalized) Riemann integral *of* $\mathbf{f}$ *from* a *to* b*.*

iii) If $\mathbf{f} : [a, b] \to \mathbf{F}$ *is Lipschitz, consider the points* $a = t_0 < t_1 < ... < t_n = b$ *of the interval* $[a, b]$*, and the Riemann sum* $S_n = \sum_{i=0}^{n-1} (t_{i+1} - t_i)\,\mathbf{f}(t_i)$*. If* $n \to +\infty$ *and* $|t_{i+1} - t_i| \to 0$ *for all* $i \in \{1, ..., n-1\}$*, then* $S_n \to \int_a^b \mathbf{f}(t)\,dt$ *in the sense of Mackey convergence.*

PROOF.– See [KRI 97], Chapter 1, Lemma 2.5 and Proposition 2.7. ■

If $\mathbf{F}$ is a Banach space, this recovers the classical notion of the integral of a continuous function. The integration operator $\int_a^b$ defined above satisfies the usual properties ([KRI 97], Chapter 1, Corollary 2.6). In particular, we have the following result:

THEOREM 1.21.– (integral mean value theorem) *Let* $I = [\alpha, \beta]$ *be a compact interval of* $\mathbb{R}$ *and let* $\mathbf{f} : I \to \mathbf{F}$ *be a continuous mapping. For every continuous semi-norm* $|.|_\gamma$ *of* $\mathbf{F}$*,*

$$\left|\int_\alpha^\beta \mathbf{f}(t)\,.dt\right|_\gamma \leq \int_\alpha^\beta |\mathbf{f}(t)\,.dt|_\gamma \leq (\beta - \alpha)\,.\sup_{t\in I} |\mathbf{f}(t)|_\gamma\,.$$

PROOF.– Let $\mathbf{g}(t) = \int_\alpha^t \mathbf{f}(t)\,.dt$. By the definition of the integral, $\mathbf{g}$ is differentiable in I and $|\dot{\mathbf{g}}(t)|_\gamma = \dot{\varphi}(t)$, where $\dot{\varphi}(t) = |\mathbf{f}(t)|_\gamma$. Applying the mean value theorem (Theorem 1.12) to $\mathbf{g}$ therefore gives the desired inequalities. ■

(III) TAYLOR'S FORMULAS Let $\mathbf{E}$ be a normed vector space with norm $|.|$, A some non-empty open subset of $\mathbf{E}$ and $\mathbf{F}$ a quasi-complete locally convex space whose topology is defined by a family of semi-norms $\left(|.|_\gamma\right)_{\gamma\in\Gamma}$. With the conventions of section 1.2.1 and $p \geq 1$, we have the following result:

THEOREM 1.22.– *i) Let* $\mathbf{f} : A \to \mathbf{F}$ *be a mapping of class* C^p *in* A*. If the segment* $[a, a + \mathbf{h}]$ *is contained in* A*, then* $\mathbf{f}$ *satisfies Taylor's formula*

$$\mathbf{f}(a + \mathbf{h}) = \sum_{k=0}^{p-1} \frac{1}{k!} D^k \mathbf{f}(a)\,.\mathbf{h}^k + r_p(\mathbf{h})\,, \qquad [1.12]$$

where the residual $r_p(\mathbf{h})$ *is the* (Laplace residual) *integral*

$$r_p(\mathbf{h}) = \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p \mathbf{f}(a + t.\mathbf{h}) .\mathbf{h}^p .dt.$$

ii) As $|\mathbf{h}| \to 0$*, we can alternatively express the residual in the form of* Young's residual*:*

$$r_p(\mathbf{h}) = \frac{1}{p!} D^p \mathbf{f}(a) .\mathbf{h}^p + o(|\mathbf{h}|^p).$$

iii) Let $\gamma \in \Gamma$ *and suppose that there exists some real number* $M > 0$ *such that* $\sup_{x \in]a, a+\mathbf{h}[} \|D^p \mathbf{f}(x)\|_\gamma \leq M$*. Then,* $|r_p(\mathbf{h})|_\gamma$ *is upper bounded by the* Lagrange residual*:*

$$|r_p(\mathbf{h})|_\gamma \leq \frac{M}{p!} |\mathbf{h}|^p.$$

iv) If $\mathbb{K} = \mathbf{F} = \mathbb{R}$*, suppose that* $\mathbf{f}$ *is of class* C^{p-1} *in* A *and admits a differential of order* p*,* $D^p \mathbf{f}(x)$*, at every point of the open segment* $]a, a+\mathbf{h}[$ *. The* Lagrange residual *can be expressed as follows: there exists* $\theta \in]0,1[$ *such that*

$$r_p(\mathbf{h}) = \frac{D^p \mathbf{f}(a + \theta.\mathbf{h})}{p!} |\mathbf{h}|^p.$$

PROOF.– i) For $p = 1$,

$$\begin{aligned} \mathbf{f}(a + 1.\mathbf{h}) - \mathbf{f}(a) &= \int_0^1 \frac{d}{dt}(\mathbf{f}(a + t.\mathbf{h}))\, dt \\ &= \int_0^1 D\mathbf{f}(a + t.\mathbf{h}).\mathbf{h}.dt. \end{aligned}$$

For $p = 2$, we can simply evaluate the above integral by parts. This integral is of the form $\int_0^1 \mathbf{u}.dv$ with $\mathbf{u} = D\mathbf{f}(a + t\mathbf{h}).\mathbf{h}$ and $v = 1 - t$; hence, $[\mathbf{u}.v]_0^1 - \int_0^1 v.d\mathbf{u} = D\mathbf{f}(a) + \int_0^1 (1-t) D^2 \mathbf{f}(a + t\mathbf{h}) .\mathbf{h}^2.dt$. Continuing inductively gives (i).

ii) Since $D^p\mathbf{f}$ is continuous, for all $\varepsilon > 0$, there exists a real number $r_\gamma > 0$ such that $\|D^p\mathbf{f}(a+t.\mathbf{h}) - D^p\mathbf{f}(a)\|_\gamma \leq p!\varepsilon$ whenever $|\mathbf{h}| \leq r$ and $0 \leq t \leq 1$. The first inequality of the integral mean value theorem (Theorem 1.21) then implies:

$$\begin{aligned}
&\left|r_p(\mathbf{h}) - \frac{1}{p!}D^p\mathbf{f}(a).\mathbf{h}^p\right|_\gamma \\
&= \left|\int_0^1 \left(\frac{(1-t)^{p-1}}{(p-1)!}\left(D^p\mathbf{f}(a+t.\mathbf{h})\,dt - \frac{1}{p!}D^p\mathbf{f}(a)\right)\right).\mathbf{h}^p.dt\right|_\gamma \\
&\leq \int_0^1 \left\|\frac{(1-t)^{p-1}}{(p-1)!}\left((D^p\mathbf{f}(a+t.\mathbf{h})\,dt - D^p\mathbf{f}(a))\right)\right\|_\gamma .dt.\,|\mathbf{h}|^p \leq \varepsilon.\,|\mathbf{h}|^p\,.
\end{aligned}$$

iii) This again follows from (i) and the integral mean value theorem, since

$$|r_p(\mathbf{h})|_\gamma \leq \int_0^1 \left|\frac{(1-t)^{p-1}}{(p-1)!}D^p\mathbf{f}(a+t.\mathbf{h}).\mathbf{h}^p\right|_\gamma dt \leq \frac{M_\gamma.\,|\mathbf{h}|^p}{(p-1)!}\int_0^1 (1-t)^{p-1}\,dt.$$

iv) First, consider the case of a function g that is defined and of class C^{p-1} in some open neighborhood of a compact interval $[a,b]$ of $\mathbb{R}$ and assumed to have a p-th derivative $g^{(p)}(t)$ in $]a,b[$. Set

$$\varphi_{p-1}(t) = g(b) - \sum_{k=1}^{p-1}\frac{(b-t)^k}{k!}g^{(k)}(t)$$

and $\psi_p(t) = \varphi_{p-1}(t) - \lambda.\frac{(b-t)^p}{p!}$, where λ is chosen so that $\psi_p(a) = 0$. Since $\psi_p(b) = 0$, Rolle's theorem (Theorem 1.6) implies that there exists a point $c \in\,]a,b[$ such that $\psi_p^{(p)}(c) = 0$, and hence

$$\varphi_{p-1}(a) - \frac{(b-a)^p}{p!}g^{(p)}(c) = 0.$$

We now simply need to apply this result to the mapping $g(t) = \mathbf{f}(a+t.\mathbf{h})$, $t \in [0,1]$. ■

COROLLARY 1.23.– *Let* $\mathbf{E}$ *be a real normed vector space,* A *some non-empty subset of* $\mathbf{E}$*,* $f : A \to \mathbb{R}$ *a mapping of class* C^p $(2 \leq p \leq \infty)$ *and a some point of* A*.*

i) For f *to have a* relative minimum *(or* local minimum*) at the point* a*, the smallest integer* $n \leq p$ *such that* $D^n f(a)$ *is non-zero (if any such integer* n *exists) must necessarily be even; furthermore,* $D^n f(a).\mathbf{h} \geq 0$ *for every non-zero vector* $\mathbf{h} \in \mathbf{E}$*.*

ii) Conversely, if A is convex, $n \leq p$ is even, $D^i f(a) = 0$ for all $i \in \{1, ..., n-1\}$, and $D^n f(x).\mathbf{h}^n > 0$ for all $x \in A$ and every $\mathbf{h} \neq 0$, then f has a strict relative minimum at the point a.

iii) If $D^i f(a) = 0$ for all $i \in \{1, ..., n-1\}$, $n \leq p$ is even, and there exists $\varepsilon > 0$ such that $D^n f(x).\mathbf{h}^n > \varepsilon |\mathbf{h}|^n$ for every $x \in A$ and every $\mathbf{h} \neq 0$, then f has a strict relative minimum at the point a (if $\mathbf{E}$ is finite-dimensional, this sufficient condition is still valid with $\varepsilon = 0$).

PROOF.– i) follows from Taylor's formula with Young's residual; ii) and iii) follow from Taylor's formula with the Lagrange residual. If $\mathbf{E}$ is finite-dimensional, the unit sphere $\mathbb{S}^1 := \{\mathbf{h} \in E : |\mathbf{h}| = 1\}$ is compact by Riesz's theorem ([P2], section 3.2.3, Theorem 3.11); thus, $\mathbb{S}^n$ is compact by Tychonov's theorem ([P2], section 2.3.7, Theorem 2.43) and $D^n f(x).\mathbf{h}^n$ has a minimum $\varepsilon > 0$ in $\mathbb{S}^n$ (*ibid.*, Theorem 2.39). ■

(IV) If $\mathbf{E} = \mathbf{E}_1 \times ... \times \mathbf{E}_n$, where each $\mathbf{E}_i$ is a normed vector space, let α be the multi-index $(\alpha_1, ..., \alpha_n)$, $\alpha_i \in \mathbb{N}$ $(i \in \{1, ..., n\})$. Define

$$D^\alpha := D_1^{\alpha_1} ... D_n^{\alpha_n}, \quad \mathbf{h}^\alpha := (\mathbf{h}_1^{\alpha_1}, ..., \mathbf{h}_n^{\alpha_n}) \quad (\mathbf{h} = (\mathbf{h}_1, ..., \mathbf{h}_n));$$

the operator D^α is called the *partial differential of order* α. If each $\mathbf{E}_i$ is equal to $\mathbb{R}$, write ∂^α for D^α ([P2], section 4.3.1**(I)**). Let U be an open neighborhood of a in $\mathbf{E}$ and suppose that the function $\mathbf{f} : U \to \mathbf{F}$ is of class C^p in U. Then, Taylor's formula can be rewritten as follows:

$$\mathbf{f}(a + \mathbf{h}) = \sum_{|\alpha| \leq p-1} \frac{1}{\alpha!} D^\alpha \mathbf{f}(a).\mathbf{h}^\alpha + r_p(\mathbf{h}). \quad [1.13]$$

1.2.5. *Analytic functions*

(I) POWER SERIES Let $\mathbf{E}$ be a normed vector space with norm $|.|$ and suppose that $\mathbf{F}$ is a Hausdorff quasi-complete locally convex space (see Remark 1.1 for the case where $\mathbf{F}$ is normable). With the notation of section 1.2.1, let $\mathcal{S}(\mathbf{E}; \mathbf{F})$ be the $\mathbb{K}$-vector space of formal power series $\mathbf{S} = \sum_p \mathbf{S}_p$, where $\mathbf{S}_p = \mathbf{c}_p.X^p$ and $\mathbf{c}_p \in \mathcal{L}_{p,s}(\mathbf{E}; \mathbf{F})$. Let $\left(|.|_\gamma\right)_{\gamma \in \Gamma}$ be a family of semi-norms which induces the topology of $\mathbf{F}$ and let $r > 0$; we write that $\|\mathbf{S}\|_{\gamma,r} = \sum_p r^p \|\mathbf{c}_p\|_\gamma$ and

$$\mathbb{S}_r(\mathbf{E}; \mathbf{F}) = \left\{\mathbf{S} \in \mathcal{S}(\mathbf{E}; \mathbf{F}) : \|\mathbf{S}\|_{\gamma,r} < \infty, \forall \gamma \in \Gamma\right\};$$

$$\mathbb{S}(\mathbf{E}; \mathbf{F}) = \bigcup_{r>0} \mathbb{S}_r(\mathbf{E}; \mathbf{F}).$$

The set $\mathbb{S}(\mathbf{E};\mathbf{F})$ is a $\mathbb{K}$-vector space called the space of *convergent power series*. If $\mathbf{S}\in\mathbb{S}(\mathbf{E};\mathbf{F})$, we say that $\rho(\mathbf{S}):=\inf\{r>0:\mathbf{S}\in\mathbb{S}_r(\mathbf{E};\mathbf{F})\}$ is the *radius of convergence* of $\mathbf{S}$. Suppose that $\rho(\mathbf{S})>0$; if we replace the indeterminate X with an element $\mathbf{h}\in\mathbf{E}$ such that $|\mathbf{h}|<\rho(\mathbf{S})$, then the family $\mathbf{c}_p.\mathbf{h}^p$ is summable ([P2], section 3.2.1**(III)**), as can be seen by adapting the proof of ([P2], section 3.4.1**(I)**, Theorem 3.41), and the mapping $\mathbf{S}:\mathbf{S}\mapsto\mathbf{S}(\mathbf{h})$ is continuous in the open set $|\mathbf{h}|<\rho(\mathbf{S})$. If $\mathbf{F}$ is a Banach space and $\rho(\mathbf{S})>0$, then the power series $\mathbf{S}$ is absolutely convergent in $|\mathbf{h}|<\rho(\mathbf{S})$ and normally convergent in $|\mathbf{h}|\leq r'$ for every r' such that $0<r'<\rho(\mathbf{S})$ ([P2], section 4.3.2**(I)**).

(II) ANALYTIC FUNCTIONS Let A be a non-empty open subset of $\mathbf{E}$. We say that a function $\mathbf{f}$ from A into $\mathbf{F}$ is analytic (or is a mapping of class C^ω) if, for each point $a\in A$, there exists a convergent series $\mathbf{S}\in\mathbb{S}(\mathbf{E};\mathbf{F})$, denoted by $\mathbf{f}_a$, such that $\mathbf{f}(a+\mathbf{h})=\mathbf{f}_a(\mathbf{h})$ for every $\mathbf{h}\in\mathbf{E}$ with sufficiently small norm. This definition generalizes ([P2], section 4.3.2**(I)**, Definition 4.74). Write $\mathcal{C}^\omega(A;\mathbf{F})$ for the $\mathbb{K}$-vector space of analytic functions from A into $\mathbf{F}$. If $\mathbb{K}=\mathbb{R}$ and $\mathbf{f}\in\mathcal{C}^\omega(A;\mathbf{F})$, then $\mathbf{f}$ is of class C^∞ in A, and so is each of its differentials $D^p\mathbf{f}$ $(p\geq 1)$. Every mapping $\mathbf{f}\in\mathcal{C}^\omega(A;\mathbf{F})$ admits the following Taylor series expansion at the point a, which converges in $|\mathbf{h}|<\rho(\mathbf{f}_a)$:

$$\mathbf{f}(a+\mathbf{h})=\sum_{p=0}^{\infty}\frac{1}{p!}D^p\mathbf{f}(a).\mathbf{h}^p. \qquad [1.14]$$

If $|a|<\rho(\mathbf{f}_a)$, then the radius of convergence of the Taylor expansion of $\mathbf{f}$ at the point a is greater than or equal to $\rho(\mathbf{f}_a)-|a|$. If $A=\mathbf{E}$ and $\rho(\mathbf{f}_a)=+\infty$, then the function $\mathbf{f}$ is said to be *entire*.

Let $\mathbf{E},\mathbf{F}$ be Banach spaces, $\mathbf{G}$ a quasi-complete locally convex space, A an open subset of $\mathbf{E}$, $\mathbf{f}:A\to\mathbf{F}$ an analytic function, B an open subset of $\mathbf{F}$ containing $\mathbf{f}(A)$ and $\mathbf{g}:B\to\mathbf{G}$ an analytic function. Then, $\mathbf{g}\circ\mathbf{f}$ is analytic (**exercise**); see ([BOU 82a], 3.2.7), ([WHI 65], p. 1079).

The principle of analytic continuation ([P2], section 4.3.2, Theorem 4.76) can be generalized as follows (**exercise**: see [WHI 65], p. 1080): let $\mathbf{E}$ and $\mathbf{F}$ be two Banach spaces, Ω a *connected* open subset of $\mathbf{E}$ and $\mathbf{f},\mathbf{g}$ two analytic functions from Ω into $\mathbf{F}$. If $\mathbf{f}$ and $\mathbf{g}$ coincide in any non-empty open subset of Ω, then they must be equal.

LEMMA 1.24.– *Let* $\mathbf{E}$ *be a Banach space and write* $\mathfrak{H}$ *for the subset of invertible operators in* $\mathcal{L}(\mathbf{E})$. *Let* $\mathcal{I}:\mathfrak{H}\to\mathfrak{H}:\mathbf{u}\mapsto\mathbf{u}^{-1}$. *The mapping* $\mathcal{I}$ *is analytic and satisfies* $D\mathcal{I}(\mathbf{u}_0).\mathbf{h}=-\mathbf{u}_0^{-1}.\mathbf{h}.\mathbf{u}_0^{-1}$ *for every* $\mathbf{u}_0\in\mathfrak{H}$.

PROOF.– We know that $\mathfrak{H}$ is open in $\mathcal{L}(\mathbf{E})$ ([P2], section 3.4.1**(II)**, Corollary 3.49). Let $\mathbf{u}_0\in\mathfrak{H}$ and $\mathbf{s}\in\mathcal{L}(\mathbf{E})$. We have $\mathbf{u}_0+\mathbf{s}=\mathbf{u}_0(1_E-\mathbf{v})$, where $\mathbf{v}=-\mathbf{u}_0^{-1}\mathbf{s}$. If $\|\mathbf{v}\|<1$, then $1_E-\mathbf{v}$ is invertible with inverse $\sum_{n\geq 0}\mathbf{v}^n$ (*ibid.*). Hence, if $\|\mathbf{s}\|<$

$\frac{1}{\|\mathbf{u}_0\|}$, $\mathbf{u}_0 + \mathbf{s}$ has inverse $\sum_{n\geq 0}\left(-\mathbf{u}_0^{-1}.\mathbf{s}\right)^n \mathbf{u}_0^{-1}$, which shows that $\mathcal{I}$ is analytic. Furthermore, $\sum_{n\geq 0}\left(-\mathbf{u}_0^{-1}.\mathbf{s}\right)^n \mathbf{u}_0^{-1} = \mathbf{u}_0^{-1} - \mathbf{u}_0^{-1}.\mathbf{s}.\mathbf{u}_0^{-1} + o\left(\|\mathbf{s}\|\right).$ ■

(III) HOLOMORPHIC FUNCTIONS Let $\mathbf{E}$ be a normed complex vector space with norm $|.|$, A some non-empty open subset of $\mathbf{E}$, and $\mathbf{F}$ a complex quasi-complete locally convex space. *Goursat's theorem* ([P2], section 4.2.4, Proposition 4.56) can be generalized as follows ([BOU 82a], 3.1.1): the function $\mathbf{f} : A \to \mathbf{F}$ is analytic if and only if it is *holomorphic* (i.e. complex-differentiable). If this condition is satisfied for $\mathbf{E} = \mathbf{E}_1 \times ... \times \mathbf{E}_n$, let $a = (a_1, ..., a_n) \in A$, $\mathbf{r} = (r_1, ..., r_n)$, where $r_i > 0$, and $\mathbf{c}_\alpha = \frac{1}{\alpha!} D^\alpha \mathbf{f}(a)$, where α is the multi-index $(\alpha_1, ..., \alpha_n)$. The Cauchy inequalities ([P2], section 4.3.2(**II**), Lemma-Definition 4.78(2)) can be generalized as follows (**exercise**): for $\mathbf{r}^\alpha := r_1^{\alpha_1}...r_n^{\alpha_n}$,

$$\|\mathbf{c}_\alpha\|_\gamma \leq \frac{M_\gamma}{\mathbf{r}^\alpha} \text{ if } M_\gamma := \sup_{|\xi_i - a_i| \leq r_i : i=1,...,n} |\mathbf{f}(\xi)|_\gamma < \infty.$$

Hence ([P2], section 4.3.2(**II**), Theorem-Definition 4.81(3)), if $\mathbf{f}$ is entire in $\mathbf{E}$ and bounded in $\mathbf{F}$, then it must be constant (*Liouville's theorem*). The statement of Hartogs' theorem ([P2], section 4.3.2(**II**), Corollary 4.80) also holds, *mutatis mutandis*, for a function $\mathbf{f} : A \to \mathbf{F}$, where A is an open subset of $\mathbf{E}_1 \times ... \times \mathbf{E}_n$, and each $\mathbf{E}_i$ is a complex normed vector space: any such function is analytic if and only if it is analytic in each of its variables when the others are held fixed.

THEOREM 1.25.– (maximum modulus) *Let* $\mathbf{E}$ *(respectively* $\mathbf{F}$*) be a* complex *Banach space (respectively quasi-complete Hausdorff locally convex space),* A *some connected non-empty open subset of* $\mathbf{E}$ *and* $\mathbf{f} : \mathbf{E} \to \mathbf{F}$ *a holomorphic function. Let* $|.|_\gamma$ *be a continuous semi-norm on* $\mathbf{F}$*. If the function* $|\mathbf{f}|_\gamma : x \mapsto |\mathbf{f}(x)|_\gamma$ *is not constant, then it does not have a maximum in* A*.*

PROOF.– 1) Let us begin by showing the result by contradiction when $\mathbf{E} = \mathbf{F} = \mathbb{C}$. Suppose that $\mathbf{f}$ has a maximum in A. By translation, we may assume that $0 \in A$ and that this maximum is attained at 0. Let $c_0 = \mathbf{f}(0)$. If $\mathbf{f}$ is not constant, then there exists $b_m \neq 0$ such that $\mathbf{f}(z) = c_0(1 + b_m z^m + z^m.h(z))$, where h is holomorphic in A and satisfies $h(0) = 0$. Choose $r > 0$ such that $|z| \leq r$ implies $z \in A$ and $h(z) \leq \frac{1}{2} b_m$. Let $t \in \mathbb{R}$ be such that $e^{mit} = \frac{|b_m|}{b_m}$. For $z = re^{it}$, we have

$$|1 + b_m z^m + z^m h(z)| \geq 1 + \frac{1}{2}|b_m| z^m,$$

which is a contradiction.

2) In the case where $\mathbf{E} = \mathbb{C}$, we can similarly argue by contradiction by assuming that there exist $z_0, z_1 \in A$ such that $|\mathbf{f}(z)|_\gamma \leq |\mathbf{f}(z_0)|_\gamma$ for all $z \in A$ and $|\mathbf{f}(z_1)|_\gamma < |\mathbf{f}(z_0)|_\gamma$. Let $V = \{\lambda.\mathbf{f}(z_0) : \lambda \in \mathbb{C}\}$ and $\eta : V \to \mathbb{C} : \lambda.\mathbf{f}(z_0) \longmapsto \lambda.|\mathbf{f}(z_0)|_\gamma$.

Then, $|\eta|_\gamma = 1$, where $|\eta|_\gamma := \sup_{\mathbf{y}\in\mathbf{F},|\mathbf{y}|_\gamma\leq 1} |\langle \eta, \mathbf{y}\rangle|$. By the Hahn–Banach theorem ([P2], section 3.3.4(II), Theorem 3.25), there exists a continuous linear form $\xi \in \mathbf{F}^\vee$ extending η such that $|\xi|_\gamma = 1$. Therefore, for all $x \in A$, $|\xi \circ \mathbf{f}(z)| \leq |\mathbf{f}(z)|_\gamma \leq |\mathbf{f}(z_0)|_\gamma = |\xi \circ \mathbf{f}(z_0)|$, so $\xi \circ \mathbf{f}$ is constant by (1). Hence, $|\xi \circ \mathbf{f}(z_1)| = |\xi \circ \mathbf{f}(z_0)| = |\mathbf{f}(z_0)|_\gamma$ and $|\xi \circ \mathbf{f}(z_1)| \leq |\mathbf{f}(z_1)|_\gamma < |\mathbf{f}(z_0)|_\gamma$, contradiction.

3) In the general case, let $\mathbf{g}(\xi) = \mathbf{f}(z_0 + \xi(z - z_0))$ and suppose that $|\mathbf{f}(z)|_\gamma \leq |\mathbf{f}(z_0)|_\gamma$ for all $z \in A$. Then, $\mathbf{g}$ is holomorphic in $\Omega = \{\xi \in \mathbb{C} : |\xi| < 1 + r\}$ for sufficiently small $r > 0$ and z sufficiently close to z_0. Therefore, $|\mathbf{g}(\xi)|_\gamma \leq |\mathbf{f}(z_0)|_\gamma = |\mathbf{g}(0)|_\gamma$, and $\mathbf{g}$ is constant in Ω by (2). Thus, $\mathbf{g}(0) = \mathbf{g}(1)$, so $\mathbf{f}(z) = \mathbf{f}(z_0)$. The set of $z \in A$ satisfying this condition is non-empty, open and closed in A, and so must be equal to A ([P2], section 2.3.8). ■

1.2.6. *The implicit function theorem and its consequences*

(I) Banach–Caccioppoli fixed point theorem

DEFINITION 1.26.– *Let (X, d) be a metric space and $f : X \to X$ a mapping.*

i) We say that a point $\xi \in X$ is a fixed point of f if $f(\xi) = \xi$.

ii) We say that f is a contraction if there exists some constant k, $0 \leq k < 1$, such that, for all $x, x' \in X$, $d(f(x), f(x')) \leq k.d(x, x')$.

THEOREM 1.27.– (Banach–Caccioppoli fixed point theorem) *Every contraction in a* complete *metric space has a unique fixed point.*

PROOF.– a) *Uniqueness*: If $f(\xi) = \xi$ and $f(\xi') = \xi'$, then $d(f(\xi), f(\xi')) = d(\xi, \xi')$ and $d(f(\xi), f(\xi')) \leq k.d(\xi, \xi')$, so $d(\xi, \xi') \leq k.d(\xi, \xi')$, and thus $(1 - k).d(\xi, \xi') \leq 0$. Since $1 - k > 0$, this implies that $d(\xi, \xi') = 0$ and $\xi = \xi'$.

b) *Existence*: We will use the method of successive approximation: let (x_n) be the sequence of elements of X defined from some arbitrary starting point $x_0 \in X$ by the recurrence relation $x_{n+1} = f(x_n)$. For all $n \geq 0$,

$$d(x_{n+1}, x_n) \leq k.d(x_n, x_{n-1}) \leq ... \leq k^n.d(x_1, x_0),$$

and so for all $p \geq 1$, by the triangle inequality,

$$d(x_{n+p}, x_n) \leq \sum_{i=0}^{p-1} d(x_{n+p-i}, x_{n+p-i-1}) \leq \underbrace{\left(\sum_{i=0}^{p-1} k^{p-i-1}\right)}_{1/(1-k)} k^n.d(x_1, x_0).$$

Hence, (x_n) is a Cauchy sequence. Since X is complete, (x_n) must have some limit ξ in X; but f is continuous, so $\xi = f(\xi)$. ■

(II) INVERSE MAPPING AND IMPLICIT FUNCTION THEOREMS Below, we assume that $0 < p \leq \omega$.

DEFINITION 1.28.– *A* diffeomorphism *of class C^p (or a C^p-diffeomorphism) is a bijection of class C^p whose inverse bijection is also of class C^p.*

This definition can be localized in the obvious way as we did for homeomorphisms ([P2], section 2.3.4(**III**)). Every diffeomorphism (respectively local diffeomorphism) is clearly a homeomorphism (respectively local homeomorphism). A local diffeomorphism of class C^p is also known as an *étale mapping* of class C^p (see [P2], section 5.2.3(**II**)).

THEOREM 1.29.– (inverse mapping theorem) *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces and suppose that* $\mathbf{f}$ *is a mapping of class C^p taking values in* $\mathbf{F}$ *and defined in a neighborhood of some point $a \in$* $\mathbf{E}$*. Let $b = \mathbf{f}(a)$ and suppose that $D\mathbf{f}(a)$ is bijective. Then,* $\mathbf{f}$ *is a* local diffeomorphism *of class C^p from some neighborhood U of a into some neighborhood V of b; the inverse diffeomorphism* $\mathbf{g} : V \to U$ *(of class C^p) satisfies $D\mathbf{g}(b) = D\mathbf{f}(a)^{-1}$.*

PROOF.– 1) *Preliminary*: Since $D\mathbf{f}(a) \in \mathcal{L}(\mathbf{E};\mathbf{F})$ is bijective, it is a linear homeomorphism by the Banach inverse operator theorem ([P2], section 3.2.3, Theorem 3.12(2)(i)). Hence, $\mathbf{E}$ and $\mathbf{F}$ are isomorphic and can be identified. We can also assume that $D\mathbf{f}(a) = 1_{\mathbf{E}}$ (left-multiplying $\mathbf{f}$ by $D\mathbf{f}(a)^{-1}$ if necessary) and reduce to the case where $a = 0$ by translation. Let $\phi(x) = x - \mathbf{f}(x)$. We have $D\phi(0) = 0$, and, since $D\phi$ is continuous, there exists $r > 0$ such that $|x| \leq r \Rightarrow \|D\phi(x)\| \leq \frac{1}{2}$. The mean value theorem (Theorem 1.13) then implies that $|\phi(x)| \leq \frac{1}{2}|x|$ whenever $|x| \leq r$, i.e. $\phi(B_r^c(0)) \subset B_{r/2}^c(0)$, where $B_r^c(0) := \{x \in \mathbf{E} : |x| \leq r\}$.

2) *Existence of an inverse mapping* $\mathbf{g} : B_{r/2}^c(0) \to B_r^c(0)$:

Let $y \in B_{r/2}^c(0)$. We will show that there exists a unique element $x \in B_r^c(0)$ such that $\mathbf{f}(x) = y$. Let $\psi_y(x) = y + x - \mathbf{f}(x)$. If $|y| \leq \frac{r}{2}$ and $|x| \leq r$, then $|\psi_y(x)| \leq r$, so ψ_y is a mapping from $B_r^c(0)$ into $B_r^c(0)$. For all $x_1, x_2 \in B_r^c(0)$,

$$|\psi_y(x_1) - \psi_y(x_2)| = |\phi(x_1) - \phi(x_2)| \leq \frac{1}{2}|x_1 - x_2|.$$

Since $B_r^c(0)$ is a complete metric space ([P2], section 2.4.4(**II),** Lemma 2.77), Theorem 1.27 implies that ψ_y has a unique fixed point in this set. This fixed point x satisfies $y + x - \mathbf{f}(x) = x$, so $\mathbf{f}(x) = y$, and $x = \mathbf{g}(y)$.

3) *Continuity of* $\mathbf{g}$: For $x_1, x_2 \in B_r^c(0)$, we have:

$$x_1 - x_2 = \underbrace{x_1 - \mathbf{f}(x_1)}_{\mathbf{g}(x_1)} + \mathbf{f}(x_1) - \mathbf{f}(x_2) - \left(\underbrace{x_2 - \mathbf{f}(x_2)}_{\mathbf{g}(x_2)}\right);$$

$$|x_1 - x_2| \leq |\mathbf{f}(x_1) - \mathbf{f}(x_2)| + |\mathbf{g}(x_1) - \mathbf{g}(x_2)| \leq |\mathbf{f}(x_1) - \mathbf{f}(x_2)| + \frac{1}{2}|x_1 - x_2|.$$

Therefore, $|x_1 - x_2| \leq 2|\mathbf{f}(x_1) - \mathbf{f}(x_2)|$.

4) *Differentiability of* $\mathbf{g}$: Let $y_i = \mathbf{f}(x_i)$, $y_i \in B_{r/2}^c(0)$, $x_i \in B_r^c(0)$ $(i = 1, 2)$. Then:

$$\begin{aligned}&\left|\mathbf{g}(y_1) - \mathbf{g}(y_2) - D\mathbf{f}(x_2)^{-1}.(y_1 - y_2)\right| \\ &\quad = \left|x_1 - x_2 - D\mathbf{f}(x_2)^{-1}.(\mathbf{f}(x_1) - \mathbf{f}(x_2))\right|.\end{aligned}$$

By taking sufficiently small $r > 0$, we can guarantee that $\left\|D\mathbf{f}(x_2)^{-1}\right\| \leq 1$. Therefore,

$$\left|\mathbf{g}(y_1) - \mathbf{g}(y_2) - D\mathbf{f}(x_2)^{-1}.(y_1 - y_2)\right| = o_1(|x_1 - x_2|) = o_2(|y_1 - y_2|),$$

which shows that $\mathbf{g}$ is differentiable and $D\mathbf{g}(y) = D\mathbf{f}(x)^{-1}$ in $B_{r/2}(0)$.

5) *Class of* $\mathbf{g}$: If $\mathbb{K} = \mathbb{C}$, the generalized Goursat theorem (section 1.2.5(**III**)) implies that $p = \omega$. Consider the case where $\mathbb{K} = \mathbb{R}$. Since $D\mathbf{f}$ and $\mathbf{g}$ are continuous and $\mathcal{I}$ is analytic (Lemma 1.24), $D\mathbf{g} = \mathcal{I} \circ D\mathbf{f} \circ \mathbf{g}$ is continuous, and $\mathbf{g}$ is therefore of class C^1. By induction, it follows that if $\mathbf{f}$ is of class C^p $(1 \leq p \leq \infty)$, then $\mathbf{g}$ is of class C^p.

Suppose that $\mathbf{f}$ is analytic and therefore can be expressed as an absolutely convergent series $\mathbf{f}(x) = \sum_p \mathbf{c}_p.x^p$ $(|x| < \rho)$. By (1), $\mathbf{c}_0 = 0$, $\mathbf{c}_1 = 1_{\mathbf{E}}$, so $y = x + \mathbf{c}_2.x^2 + \mathbf{c}_3.x^3 + ...$ According to a classical procedure dating back to Newton, x can be expressed in terms of y as a *formal* series $x = y + \sum_{i \geq 2} \mathbf{d}_i.y^i$. The terms $\mathbf{d}_i$ $(i \geq 2)$ are determined recursively: $y^2 = x^2 + 2\mathbf{c}_2.x^3 + ...$, $y^3 = x^3 + ...$, which gives $y = x + \mathbf{c}_2.\left(y^2 - 2\mathbf{c}_2.x^3\right) + ...$, and $x = y - \mathbf{c}_2.y^2 + (2\mathbf{c}_2^2 - \mathbf{c}_3).y^3 + ...$, where $\mathbf{c}_2^2.y^3 := \mathbf{c}_2(y, \mathbf{c}_2(y, y))$. We now need to study the convergence of this series. There exists $K > 0$ such that $\|\mathbf{c}_i\| \leq \gamma_i$, where $\gamma_i = \frac{K}{\rho^i}$ $(i \geq 2)$. Thus, the majorant series

$$\eta = \xi - \sum_{i \geq 2} \gamma^i.\xi^i = \xi - \frac{K}{\rho^2} \sum_{i \geq 0} \frac{\xi^i}{\rho^i} \qquad [1.15]$$

converges for $|\xi| < \rho$, with sum $\varphi(\xi) = \xi - \frac{K}{\rho}\frac{\xi^2}{\rho-\xi}$. As before, we can construct an inverse formal series $\xi = \eta + \sum_{i\geq 2}\delta^i.\eta$. Cauchy showed that this series has a non-zero radius of convergence as follows. By elementary arithmetic, the relation $\eta = \varphi(\xi)$ is invertible and we may write $\xi = \psi(\eta)$, for $\xi \in]-\infty, \xi_1[$, where $\xi_1 = \rho\left(1 - \sqrt{\frac{K}{K+\rho}}\right) > 0$, and $\eta \in]-\infty, \eta_1[$, where $\eta_1 = 2K + \rho - 2\sqrt{K(K+\rho)} > 0$. It is easy to check that ψ can be expanded into an entire series in a neighborhood of 0 ([KNO 51], section 107). By section 1.2.5(**I**), the formal series $y + \sum_{i\geq 2}\mathbf{d}_i.y^i$ therefore has a non-zero radius of convergence. ■

THEOREM 1.30.– (implicit function theorem) *Let* $\mathbf{E}$*,* $\mathbf{F}$ *and* $\mathbf{G}$ *be Banach spaces,* A *some non-zero open subset of* $\mathbf{E}\times\mathbf{F}$ *and* $\mathbf{f} : A \to \mathbf{G}$ *a mapping of class* C^p *in* A*. Let* $(a,b) \in A$ *be such that* $\mathbf{f}(a,b) = 0$ *and suppose that* $D_2\mathbf{f}(a,b) \in \mathcal{L}(\mathbf{F};\mathbf{G})$ *is bijective. There exists some neighborhood* U_0 *of* a *in* $\mathbf{E}$ *such that, for every* connected *open set* $U \subset U_0$ *containing* a*, there is a unique continuous mapping* u *from* U *into* $\mathbf{F}$ *satisfying* $u(a) = b$*,* $(x, \mathbf{u}(x)) \in A$*, and* $\mathbf{f}(x, \mathbf{u}(x)) = 0$ *for all* $x \in U$*. Furthermore,* u *is of class* C^p *in* U *and, for all* $x \in U$*,*

$$D\mathbf{u}(x) = -\left(D_2\mathbf{f}(x, \mathbf{u}(x))\right)^{-1} \circ \left(D_1\mathbf{f}(x, \mathbf{u}(x))\right). \qquad [1.16]$$

PROOF.– 1) *Existence of an implicit function:* Since $D_2\mathbf{f}(a,b)$ is bijective, it is an isomorphism from $\mathbf{F}$ onto $\mathbf{G}$, and we may consider that $\mathbf{F} = \mathbf{G}$; furthermore, replacing $\mathbf{f}$ by $D_2\mathbf{f}(a,b)^{-1}.\mathbf{f}$ if necessary, we may assume that $D_2\mathbf{f}(a,b) = 1_{\mathbf{F}}$. Let $\varphi : A \to \mathbf{E}\times\mathbf{F} : (x,y) \mapsto (x, \mathbf{f}(x,y))$. We have

$$D\varphi(a,b) = \begin{bmatrix} 1_{\mathbf{E}} & 0 \\ D_1\mathbf{f}(a,b) & 1_{\mathbf{F}} \end{bmatrix},$$

so $D\varphi(a,b)$ is invertible in $\mathcal{L}(\mathbf{E}\times\mathbf{F})$. By Theorem 1.29, φ is a local diffeomorphism of class C^p and admits an inverse local diffeomorphism ψ of class C^p.

Write $\psi(x,z) = (x, \mathbf{h}(x,z))$, where $\mathbf{h}$ is defined and of class C^p in some neighborhood of (a,b) and takes values in $\mathbf{F}$. Finally, set $\mathbf{u}(x) = \mathbf{h}(x,0)$. Then, $\mathbf{u}$ is of class C^p in some neighborhood of a and takes values in $\mathbf{F}$. Thus, there exists some neighborhood U_0 of a in $\mathbf{E}$ such that, for all $x \in U_0$,

$$(x, \mathbf{f}(x, \mathbf{u}(x))) = \varphi(x, \mathbf{u}(x)) = \varphi(x, \mathbf{h}(x,0)) = \varphi(\psi(x,0)) = (x,0)$$

and $\mathbf{u}(a) = \mathbf{h}(a,0)$, so $(a, \mathbf{u}(a)) = \psi(a,0) = \varphi^{-1}((a,0)) = (a,b)$, and therefore $\mathbf{u}(a) = b$. Hence, $\mathbf{u}$ is an "implicit function" of class C^p.

2) *Uniqueness of the implicit function:* Since φ is a local homeomorphism, there exist a neighborhood U_0' of a and a neighborhood V_0 of b such that there is a unique (x,y) in $U_0' \times V_0$ satisfying $\varphi(x,y) = (x,0)$ (see Figure 1.1, where Γ is the graph

of $\mathbf{f}$). We may assume that U'_0 is the same U_0 as above (replacing U_0 by $U_0 \cap U'_0$ if necessary).

If a *continuous* mapping $\mathbf{v} : U_0 \to \mathbf{F}$ satisfies $\mathbf{v}(a) = b$ and $\mathbf{f}(x, \mathbf{v}(x)) = 0$ for all $x \in U_0$, then we may assume that $\mathbf{v}(x) \in V_0$ for all $x \in U_0$, further reducing the neighborhood U_0 of a if necessary. We may similarly assume that $\mathbf{u}$, like $\mathbf{v}$, is defined in U_0. Let $U \subset U_0$ be a connected neighborhood of a and suppose that $M = \{x \in U : \mathbf{u}(x) = \mathbf{v}(x)\}$. Then, $a \in M$ and M is closed in U ([P2], section 2.3.3**(II)**, Lemma 2.30). We will show that M is also open in U. By the hypotheses, the mapping $x \mapsto D_2\mathbf{f}(x, \mathbf{u}(x))$ is continuous and $D_2\mathbf{f}(a, b) = 1_{\mathbf{F}}$, so (again reducing the neighborhood U_0 if necessary) we may assume that $D_2\mathbf{f}(x, \mathbf{u}(x))$ is invertible for all $x \in U_0$. Let $a' \in M$. There exist a neighborhood $U_{a'} \subset U$ of a' and a neighborhood $V_{a'} \subset V_0$ of $b' = \mathbf{u}(a')$ such that, for all $x \in U_{a'}$, $\mathbf{u}(x)$ is the only solution y of $\mathbf{f}(x, y) = 0$ satisfying $y \in V_{a'}$. Given that $\mathbf{v}$ is continuous at a' and $\mathbf{v}(a') = \mathbf{u}(a')$, there exists a neighborhood $W \subset U_{a'}$ of a' such that $\mathbf{v}(x) \in V_{a'}$ whenever $x \in W$. Therefore, $\mathbf{v}(x) = \mathbf{u}(x)$ for all $x \in W$, which proves that M is open. The set M is non-empty, open, and closed in the connected space U, which implies that $M = U$ ([P2], section 2.3.8).

3) *Calculation of* $D\mathbf{u}(x)$: Since $\mathbf{f}(x, \mathbf{u}(x)) = 0$ in U, the chain rule (Theorem 1.9) implies that

$$D_1\mathbf{f}(x, \mathbf{u}(x)) + D_2\mathbf{f}(x, \mathbf{u}(x)) \circ D\mathbf{u}(x) = 0, \qquad [1.17]$$

which gives us [1.16]. ■

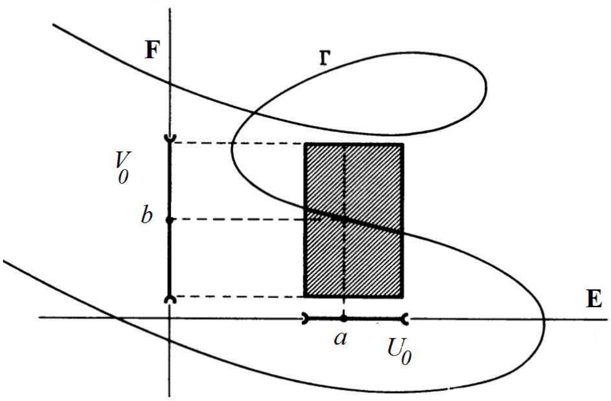


Figure 1.1. *Implicit function theorem*

REMARK 1.31.– *i) The implicit function theorem answers the following question: what condition makes it possible to express* $y = b + \Delta y$ *uniquely as a function of* $x = a + \Delta x$ *(so that* $y = \mathbf{u}(x)$*) whenever* $\mathbf{f}(x, y) = 0$ *in an open neighborhood* A *of* (a, b)*? First of all, with suitable continuity hypotheses, neglecting second-order terms, and assuming that* Δx*,* Δy *are sufficiently small, we have* $0 = \mathbf{f}(x, y) \simeq \frac{\partial \mathbf{f}}{\partial x}(x, y).\Delta x + \frac{\partial \mathbf{f}}{\partial y}(x, y).\Delta y$*. Hence:*

$$\Delta y \simeq \underbrace{D_2\mathbf{f}(x, y)^{-1} \circ D_1\mathbf{f}(x, y)}_{D\mathbf{u}(x)}.\Delta x$$

if $D_2\mathbf{f}(x, y)$ *is invertible, or equivalently if* $D_2\mathbf{f}(a, b)$ *is invertible by continuity of* $D_2\mathbf{f}$*. Figure 1.1 shows that the functional relation* $y = \mathbf{u}(x)$ *might only be valid in a sufficiently small neighborhood of* (a, b) *. If* x *belongs to a* connected *neighborhood* U_0 *of* a*, the variable* y *such that* $\mathbf{f}(x, y) = 0$ *can be made arbitrarily close to* b*, provided that* U_0 *is chosen small enough.*

ii) The reader may wish to find an expression for [1.16] in the finite-dimensional case using Jacobian matrices ([DIE 93], Volume 1, (10.2.2)); the composition of two linear mappings translates to the product of their matrices.

iii) The statement of Theorem 1.30 no longer holds when $\mathbf{F}$ *and* $\mathbf{G}$ *are arbitrary Fréchet spaces [SER 72]; however, it remains valid when* $\mathbf{E}$ *is a non-complete normed vector space (see [SCH 93], Volume 2, Theorem 3.8.5).*

(III) IMMERSIONS, SUBMERSIONS, SUBIMMERSIONS, THE RANK THEOREM In the following, we assume that $0 < p \leq \omega$.

COROLLARY-DEFINITION 1.32.– *1) Let* $\mathbf{E}, \mathbf{F}$ *be Banach spaces,* A *an open subset of* $\mathbf{E}$*,* $a \in A$*, and* $\mathbf{i} : A \to \mathbf{F}$ *a mapping of class* C^p *such that* $\mathbf{i}(a) = 0$*,* $D\mathbf{i}(a)$ *is injective and its image* $\mathbf{F}_1 = im(D\mathbf{i}(a))$ splits *in* $\mathbf{F}$ *([P2], section 3.2.2(**IV**)), i.e. admits a topological complement* $\mathbf{F}_2$ *(*ibid.*). Then, there exist a local homeomorphism* $\mathbf{r} : \mathbf{F} \to \mathbf{F}_1 \times \mathbf{F}_2$ *in some neighborhood of* 0 *and an open neighborhood* $U \subset A$ *of* a *in* $\mathbf{E}$ *such that* $\mathbf{r} \circ \mathbf{i}$ *induces a diffeomorphism of class* C^p *from* U *onto an open subset of* $\mathbf{F}_1$*. The local homeomorphism* $\mathbf{r}$ *is a local diffeomorphism of class* C^p*.*

2) The mapping $\mathbf{i}$ *defined above is called an* immersion *of class* C^p*.*

PROOF.– The Banach space $\mathbf{F}$ can be identified with $\mathbf{F}_1 \times \mathbf{F}_2$, and since $D\mathbf{i}(a)$ is an isomorphism from $\mathbf{E}$ onto $\mathbf{F}_1$, $\mathbf{E}$ can be identified with $\mathbf{F}_1$. By translation, we may assume that $a = 0$. Let

$$\varphi : U \times \mathbf{F}_2 \to \mathbf{F}_1 \times \mathbf{F}_2 : (x, y_2) \mapsto \mathbf{i}(x) + (0, y_2).$$

We have $\varphi(x,0) = \mathbf{i}(x)$ and $D\varphi(0,0) = D\mathbf{i}(0) + (0, 1_{\mathbf{F}_2})$. Since $D\mathbf{i}(0)$ is an isomorphism from $\mathbf{F}_1$ onto $\mathbf{F}_1 \times 0$, $D\varphi(0,0)$ is an isomorphism from $\mathbf{F}_1 \times \mathbf{F}_2$ onto $\mathbf{F}_1 \times \mathbf{F}_2$. By Theorem 1.29, φ is a local diffeomorphism of class C^p that admits an inverse diffeomorphism $\mathbf{r}$ of class C^p. Hence, there exists an open neighborhood U of a in $\mathbf{F}_1$ such that, for all $x \in U$, $\mathbf{r}(\mathbf{i}(x)) = x$. ■

Recall that every *closed* subspace of a Hilbert space $\mathbf{E}$ splits in $\mathbf{E}$ ([P2], section 3.10.2**(II),** Theorem 3.147(2)). Corollary 1.33 shows that any immersion of class C^p admits a local retraction $\mathbf{r}$ of class C^p. Immersions are therefore local sections (see [P1], section 1.1.1**(III)**).

COROLLARY-DEFINITION 1.33.– *1) Let* $\mathbf{E}, \mathbf{F}$ *be Banach spaces,* A *a non-empty open subset of* $\mathbf{E}$*,* $a \in A$ *and* $\mathbf{s} : A \to \mathbf{F}$ *a mapping of class* C^p *such that* $D\mathbf{s}(a)$ *is surjective with a kernel* $\mathbf{E}_2$ *that* splits *in* $\mathbf{E}$ *(i.e.* $\mathbf{E} \cong \mathbf{E}_1 \times \mathbf{E}_2$*). Then, there exist an open neighborhood* $U \subset A$ *of* a *and a diffeomorphism* ψ *of class* C^p

$$\psi : V_1 \times V_2 \to U,$$

where $V_1 \times V_2 \subset A$ $(V_i \subset \mathbf{E}_i)$ *is an open neighborhood of* a *in* $\mathbf{E}$*, such that* $\mathbf{s} \circ \psi$ *is the first projection* $V_1 \times V_2 \twoheadrightarrow V_1$.

2) The mapping $\mathbf{s}$ *defined above is called a* submersion *of class* C^p.

PROOF.– By translation, we may assume that $a = (a_1, a_2) = (0,0)$. Thus, $D_1\mathbf{s}(0,0)$ is an isomorphism from $\mathbf{E}_1$ onto $\mathbf{F}$ that allows these two Banach spaces to be identified. Let

$$\varphi : A \to \mathbf{E}_1 \times \mathbf{E}_2 : (x_1, x_2) \mapsto (\mathbf{s}(x_1, x_2), x_2).$$

Then, $D\varphi(0)$ is represented by the matrix

$$\begin{pmatrix} D_1\mathbf{s}(0,0) & D_2\mathbf{s}(0,0) \\ 0 & 1_{\mathbf{E}_2} \end{pmatrix},$$

which is an automorphism of $\mathbf{E}_1 \times \mathbf{E}_2$. By Theorem 1.29, there exist a neighborhood $U \subset A$ of 0 in $\mathbf{E}_1 \times \mathbf{E}_2$ and a diffeomorphism $\varphi : U \to V_1 \times V_2$ of class C^p with inverse diffeomorphism $\psi : V_1 \times V_2 \to U$ of class C^p. For all $y_1 \in V_1$, we have $\mathbf{f}(\psi(y_1, x_2)) = y_1$. ■

([P2], section 3.2.2**(IV)**, Theorem 3.5(3)) implies the following result:

COROLLARY-DEFINITION 1.34.– *(1) Let* $\mathbf{E}, \mathbf{F}$ *be Banach spaces,* A *a non-empty open subset of* $\mathbf{E}$*,* $a \in A$*, and* $\mathbf{f} : A \to \mathbf{F}$ *a mapping of class* C^p*. The following conditions are equivalent:*

i) There exist an open neighborhood U *of* a*,* $U \subset A$*, a Banach space* $\mathbf{G}$*, a submersion* $\mathbf{s}$ *of* U *into* $\mathbf{G}$ *of class* C^p *and an immersion* $\mathbf{i}$ *of* $\mathbf{G}$ *into* $\mathbf{F}$ *of class* C^p *such that* $\mathbf{f}|_U = \mathbf{i} \circ \mathbf{s}$.

ii) There exist an open neighborhood $U \subset A$ *of* a*, a diffeomorphism* $\xi : U \tilde{\rightarrow} U'$ *of class* C^p*, where* U' *is an open subset of* $\mathbf{E}$*, a mapping* $\mathbf{\Phi} \in \mathcal{L}(\mathbf{E}; \mathbf{F})$ *and a diffeomorphism* $\psi : V \tilde{\rightarrow} V'$ *of class* C^p*, where* V *is an open neighborhood of* $\mathbf{f}(a)$ *in* $\mathbf{F}$ *and* $V' = \mathbf{\Phi}(U')$ *is an open subset of* $\mathbf{F}$*, such that*

$$\mathbf{f}(U) \subset V, \quad \mathbf{\Phi}(\xi(U)) \subset \psi(V), \quad \mathbf{f}|_U = \psi^{-1} \circ \mathbf{\Phi} \circ \xi, \qquad [1.18]$$

and the kernel and image of $\mathbf{\Phi}$ *split.*

2) The mapping $\mathbf{f}$ *defined above is called a* subimmersion *of class* C^p*.*

The composition of two immersions is an immersion, the composition of two submersions is a submersion and, given a subimmersion $\mathbf{f}$, an immersion $\mathbf{i}$, and a submersion $\mathbf{s}$, the mapping $\mathbf{i} \circ \mathbf{f} \circ \mathbf{s}$ is a subimmersion (**exercise**). However, the composition of two subimmersions is not always a subimmersion ([DIE 93], Volume 3, section 16.8, Problem 1(b)).

THEOREM 1.35.– (rank) *Let* $\mathbf{E}, \mathbf{F}$ *be Banach spaces,* A *some non-empty open subset of* $\mathbf{E}$ *and* $\mathbf{f} : A \to \mathbf{F}$ *a mapping of class* C^1*.*

i) If $\mathbf{f}$ *is a subimmersion, there exists an open neighborhood* $U \subset A$ *of* a *in* $\mathbf{E}$ *such that* $rk_x(\mathbf{f}) = rk_a(\mathbf{f})$ *for all* $x \in U$*.*

ii) Conversely, if there exists an open neighborhood $U \subset A$ *of* a *in* $\mathbf{E}$ *such that* $rk_x(\mathbf{f}) = rk_a(\mathbf{f})$ *for all* $x \in U$ *and if the spaces* $\mathbf{E}, \mathbf{F}$ *are finite-dimensional, then* $\mathbf{f}$ *is a subimmersion at the point* a*.*

iii) Write $rk_x(\mathbf{f}) = +\infty$ *if* $rk_x(\mathbf{f})$ *is not finite. The mapping* $x \mapsto rk_x(\mathbf{f})$ *from* A *into the discrete subspace* $\bar{\mathbb{N}} := \mathbb{N} \cup \{+\infty\}$ *of the extended real line* $\bar{\mathbb{R}}$ *is lower semi-continuous ([P2], section 2.3.3(**III**)) in* A*.*

PROOF.– (i): By [1.18], $\mathrm{rk}_x(\mathbf{f}) = \mathrm{rk}(\mathbf{\Phi})$ for all $x \in U$. (ii): See [DIE 93], Volume 1, (10.3). (iii): If $r = \mathrm{rk}_a(\mathbf{f}) < +\infty$, we may assume that $\mathbf{E} = \mathbb{K}^r$ and extract a square submatrix M_x that has rank r at $x = a$ from the Jacobian matrix of $\mathbf{f}$ at the point x. The determinant Δ_x of this submatrix is therefore non-zero for $x = a$. But the mapping $x \mapsto \Delta_x$ is continuous, so there exists a neighborhood U of a in which $\Delta_x \neq 0$, and so $\mathrm{rk}_x(\mathbf{f}) \geq r$. If $\mathrm{rk}_a(\mathbf{f}) = +\infty$, then, for all $r \in \mathbb{N}$, there exist vectors $\mathbf{e}_1, ..., \mathbf{e}_r \in \mathbf{E}$ such that the vectors $D\mathbf{f}(a).\mathbf{e}_1, ..., D\mathbf{f}(a).\mathbf{e}_r$ generate a subspace of $\mathbf{F}$ of dimension r. Arguing by contradiction, we deduce that there exists a neighborhood U of a in which $\mathrm{rk}_x(\mathbf{f}) = +\infty$. ■

1.3. Other approaches to differential calculus

1.3.1. *Lagrange variations and Gateaux differentials*

In this section, $\mathbb{K} = \mathbb{R}$.

(I) AFFINE SPACES Given a vector space $\mathbf{E}$, an *affine space* $\mathfrak{E}$ attached to the space $\mathbf{E}$ is a homogeneous space of the additive group $\mathbf{E}$ ([P1], section 2.2.8**(II)**) such that the (transitive) action of $\mathbf{E}$ on $\mathfrak{E}$ is *faithful*, i.e. such that the neutral element 0 is the only element of $\mathbf{E}$ that fixes every element of $\mathfrak{E}$. The action of $x \in \mathbf{E}$ on $P \in \mathfrak{E}$ is written as $P + x$. We say that $\mathbf{E}$ is the space of translations of $\mathfrak{E}$, its elements are the translations of $\mathfrak{E}$ and, if $\dim(\mathbf{E}) < \infty$, this quantity is called the dimension of $\mathfrak{E}$. Given some origin O chosen from $\mathfrak{E}$, the elements of $\mathfrak{E}$ are all of the form $O + x$ $(x \in \mathbf{E})$.

REMARK 1.36.– *The mapping $O + x \mapsto x$ is a bijection from $\mathfrak{E}$ onto $\mathbf{E}$ that allows these two sets to be identified.*

If $Q = P + x$, we write $x = \overrightarrow{PQ}$ (the bipoint of origin P and endpoint Q)[7]. If $\mathbf{E}$ is a locally convex space, the sets $O + U = \{O + x : x \in U\}$, where the U are the open sets of $\mathbf{E}$, define a topology on $\mathfrak{E}$. When equipped with this topology, $\mathfrak{E}$ is called a *locally convex affine space*. Every point of such a space admits a fundamental system of convex neighborhoods. We can similarly define the concepts of affine topological space, affine normed space, affine pre-Hilbert space, etc.

(II) LAGRANGE VARIATIONS Let $\mathfrak{E} = O + \mathbf{E}$ be a locally convex affine space, $\mathbf{F}$ a locally convex space, A some non-empty subset of $\mathfrak{E}$, a some point of A and $\mathbf{f} : A \to \mathbf{F}$ a mapping.

We say that $\mathbf{f}$ admits a *Lagrange first variation* $\delta\mathbf{f}(a) : \mathbf{E} \to \mathbf{F} : \mathbf{h} \mapsto \delta\mathbf{f}(a)[\mathbf{h}]$ at the point a if, for all $\mathbf{h} \in \mathbf{E}$,

$$\lim_{t \to 0, t \neq 0} \frac{\mathbf{f}(a + t.\mathbf{h}) - \mathbf{f}(a) - t.\delta\mathbf{f}(a)[\mathbf{h}]}{t} = 0.$$

If this condition is satisfied, we say that $\mathbf{f}$ admits a Lagrange *second variation* $\delta^2\mathbf{f}(a) : \mathbf{E} \to \mathbf{F}$ if

$$\lim_{t \to 0, t \neq 0} \frac{\mathbf{f}(a + t.\mathbf{h}) - \mathbf{f}(a) - t.\delta\mathbf{f}(a)[\mathbf{h}] - t^2.\delta^2\mathbf{f}(a)[\mathbf{h}]}{t^2} = 0.$$

7 As we learn in secondary school, any vector can more precisely be viewed as an equivalence class of bipoints under the equipollence relation.

The Lagrange variation of order n, $\delta^n \mathbf{f}(a) : \mathbf{E}^n \to \mathbf{F} : \mathbf{h} \mapsto \delta^n \mathbf{f}(a)[\mathbf{h}]$, is defined inductively in the same way.

(III) GATEAUX DIFFERENTIABILITY It is easy to show using the generalized Goursat theorem (section 1.2.5**(III)**) that, if $\mathbf{E}$ and $\mathbf{F}$ are *complex* locally convex spaces and $\mathbf{f}$ admits a Lagrange first variation at the point a, then $\delta\mathbf{f}(a)$ is linear ([HIL 57], Theorem 26.3.2). In general, we have the following result:

LEMMA 1.37.– *If* $\mathbf{f}$ *admits a Lagrange first variation* $\delta\mathbf{f}(a)$*, then the mapping* $\delta\mathbf{f}(a) : \mathbf{E} \to \mathbf{F}$ *is homogeneous, i.e.* $\delta\mathbf{f}(a)[\lambda.\mathbf{h}] = \lambda.\delta\mathbf{f}(a)[\mathbf{h}]$ *for all* $\lambda \in \mathbb{R}$ *and all* $\mathbf{h} \in \mathbf{E}$ ***(exercise****).*

DEFINITION 1.38.– *We say that* $\mathbf{f}$ *is* Gateaux differentiable (or G-differentiable) *at the point* a *if it admits a Lagrange first variation* $\delta\mathbf{f}(a)$ *at* a *and* $\delta\mathbf{f}(a)$ *is a* bounded linear *mapping from* $\mathbf{E}$ *into* $\mathbf{F}$ *([P2], section 3.4.4(****I****), Definition 3.63). If so,* $\delta\mathbf{f}(a)$ *is written as* $D^{\mathbf{G}}\mathbf{f}(a)$ *and is called the* Gateaux differential *of* $\mathbf{f}$ *at the point* a*.*

The set $\mathcal{D}_a^{\mathbf{G}}(A;\mathbf{F})$ of G-differentiable mappings at the point a is an affine space. A G-differentiable mapping is not necessarily continuous. If $\mathbf{E}$ is a normed vector space and $\mathbf{f}$ is differentiable at the point a, then it is also G-differentiable at this point and $D\mathbf{f}(a) = D^{\mathbf{G}}\mathbf{f}(a)$ (**exercise**). On product spaces, we may define the Gateaux partial differential $D_1^{\mathbf{G}}\mathbf{f}(a_1, a_2)$ in the first variable and, similarly, in the second variable, etc. In the conditions of Theorem 1.9, where $\mathbf{f} \in \mathcal{D}_a^{\mathbf{G}}(A;\mathbf{F})$ and $\mathbf{g} \in \mathcal{D}_{\mathbf{f}(a)}(B;\mathbf{G})$ (and B denotes an open subset of $\mathbf{F}$ containing $\mathbf{f}(A)$), we have $\mathbf{g} \circ \mathbf{f} \in \mathcal{D}_a^{\mathbf{G}}(A;\mathbf{G})$ and (instead of [1.5])

$$D^{\mathbf{G}}(\mathbf{g} \circ \mathbf{f}) = D\mathbf{g}(\mathbf{f}(a)) \circ D^{\mathbf{G}}\mathbf{f}(a) \qquad [1.19]$$

(**exercise***: see [ALE 87] section 2.2.2).

Let $\mathfrak{E}$ be a normed affine space, $\mathbf{F}$ a locally convex space, A a non-empty open subset of $\mathfrak{E}$, $[a, a+\mathbf{h}]$ a segment contained in A and $\mathbf{f} : A \to \mathbf{F}$ a G-differentiable mapping. Then, the mean value theorem (Theorem 1.13(i)) remains valid (**exercise***: see [ALE 87], section 2.2.3) in the form

$$|\mathbf{f}(a+\mathbf{h}) - \mathbf{f}(a)|_\gamma \leq |\mathbf{h}| . \sup_{t\in\Theta} \left\| D^{\mathbf{G}}\mathbf{f}(a+t\boldsymbol{h}) \right\|_\gamma , \qquad [1.20]$$

where $|.|_\gamma$ is a continuous semi-norm on $\mathbf{F}$. The claims (ii) and (iii) of this theorem also hold after making analogous adjustments. From this, we deduce the following result:

COROLLARY 1.39.– *Let* $\mathbf{E}$ *be a normed vector space,* A *some non-empty open subset of* $\mathbf{E}$*,* $\mathbf{F}$ *a locally convex space and* $\mathbf{f} : \mathbf{E} \to \mathbf{F}$ *a mapping that is G-differentiable in* A*. If* $D^{\mathbf{G}}\mathbf{f} : A \to \mathcal{L}(\mathbf{E};\mathbf{F})$ *is continuous, then* $\mathbf{f}$ *is differentiable in* A*.*

PROOF.– Let $\varepsilon > 0$ and $\gamma \in \Gamma$, where $(|.|)_{\gamma\in\Gamma}$ is the family of semi-norms which defines the topology of $\mathbf{F}$. There exists $\delta = \delta(\varepsilon, \gamma) > 0$ such that, $\forall x \in A$,

$$|x - a| < \delta \Longrightarrow \left\| D^{\mathbf{G}}\mathbf{f}(x) - D^{\mathbf{G}}\mathbf{f}(a) \right\|_{\gamma} < \varepsilon.$$

Let $\mathbf{h} \in \mathbf{E}$ be such that $|\mathbf{h}| < \delta$. The relation [1.8] with the modifications stated above implies that $\left|\mathbf{f}(a + \mathbf{h}) - \mathbf{f}(a) - D^{\mathbf{G}}\mathbf{f}(a).\mathbf{h}\right|_{\gamma} \leq \varepsilon |\mathbf{h}|$. ■

Thus, we do not need to distinguish between Fréchet and Gateaux differentiation when talking about mappings of class C^p $(p > 0)$.

1.3.2. *Calculus of variations: elementary concepts*

In this section, $\mathbb{K} = \mathbb{R}$.

(I) EULER CONDITION Let $\mathfrak{E}$ be a locally convex affine space, A some non-empty open subset of $\mathfrak{E}$, a some point of A and $J : A \to \mathbb{R}$ a mapping that admits a Lagrange first variation $\delta J(a)$ at the point a.

THEOREM 1.40.– (Euler) *For the mapping J to have a* relative extremum *(or* local extremum*) at the point a, the* Euler stationarity condition $\delta J(a) = 0$ *must necessarily be satisfied.*

PROOF.– By translating the origin, we may assume that $a = 0$. Similarly, we can assume that J has a minimum at 0 for clarity. We will argue by contradiction. Assume that $\delta J(0) \neq 0$. There exists $\mathbf{h} \in A$ such that $\beta := \delta J(0)[\mathbf{h}] \neq 0$. By Lemma 1.37, we may assume without loss of generality that $\beta < 0$ (replacing $\mathbf{h}$ by $-\mathbf{h}$ if necessary). There exists a function φ defined in some open neighborhood of 0 in $\mathbb{R}$ satisfying $\lim_{t\to 0}\varphi(t) = 0$ and $J(t.\mathbf{h}) = J(0) + t.\beta + t.\varphi(t)$. Pick $t > 0$ sufficiently small that $\beta + \varphi(t) < 0$. Then, $J(t.\mathbf{h}) < J(0)$, contradiction. ■

DEFINITION 1.41.– *We say that a is an* extreme point *(or an* extremal point*) of J if $\delta J(a) = 0$.*

THEOREM 1.42.– *Suppose that the Euler stationarity condition is satisfied. A necessary condition for J to admit a relative minimum at the point a is given by $\delta^2 J(a)[\mathbf{h}] \geq 0$ for all $\mathbf{h} \neq 0$.*

PROOF.– We have $f(a + t.\mathbf{h}) - f(a) = t^2.\delta^2 J(a)[\mathbf{h}] + o(t^2)$. ■

(II) EULER–LAGRANGE EQUATION The Euler–Lagrange equation is essentially the Euler condition applied to calculus of variations. A full treatment would require another volume[8]; we will content ourselves with briefly mentioning the simplest part,

8 A very comprehensive presentation of classical variational calculus is given in the Wikipedia article on *Calculus of variations*.

which is sufficient for our purposes. Let $\mathbf{E}$ be a Banach space and suppose that $t_1, t_2 \in \mathbb{R}$, $t_1 < t_2$.

LEMMA 1.43.– (fundamental lemma of the calculus of variations) *Let* $\mathbf{f} : [t_1, t_2] \to \mathbf{E}^\vee$ *be a continuous mapping. If every mapping* $\mathbf{h} : [t_1, t_2] \to \mathbf{E}$ *of class* C^1 *such that* $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$ *satisfies* $\int_{t_1}^{t_2} \langle \mathbf{f}(t), \mathbf{h}(t) \rangle \, dt = 0$, *then* $\mathbf{f} = 0$.

PROOF.– We will argue by contradiction from the assumption that $\mathbf{f}(t_0) \neq 0$. There must therefore exist $\mathbf{z} \in \mathbf{E}$ such that $\langle \mathbf{f}(t_0), \mathbf{z} \rangle > 0$. Since $\mathbf{f}$ is continuous, there exists an interval $[\alpha, \beta] \subset [t_1, t_2]$ containing t_0 such that $\alpha < \beta$ and $\langle \mathbf{f}(t), \mathbf{z} \rangle > 0$ for all $t \in [\alpha, \beta]$. There exists a function $\varphi : [t_1, t_2] \to \mathbb{R}$ of class C^1 that is zero in $\complement_{[t_1,t_2]}]\alpha, \beta[$ and which satisfies $\varphi(t) > 0$ for $t \in]\alpha, \beta[$; for example, $\varphi(t) = (t-\alpha)^2 (\beta - t)^2$ if $t \in]\alpha, \beta[$ and $\varphi(t) = 0$ if $t \in \complement_{[t_1,t_2]}]\alpha, \beta[$. Therefore, setting $\mathbf{h}(t) = \varphi(t).\mathbf{z}$, we have $\int_{t_1}^{t_2} \langle \mathbf{f}(t), \mathbf{h}(t) \rangle \, dt > 0$, contradiction. ■

Let Ω_1, Ω_2 be non-empty open subsets of $\mathbf{E}$ and let

$$\mathcal{L} : [t_1, t_2] \times \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R} : (t, x, u) \mapsto \mathcal{L}(t, x, u)$$

be a mapping (known as the *Lagrangian* in mechanics) of class C^1 with partial differential $D_3 \mathcal{L} = \frac{\partial \mathcal{L}}{\partial u}$. Let $x_1, x_2 \in \Omega_1$ and write $\mathcal{X}$ for the set of mappings $x : [t_1, t_2] \to \Omega_1$ of class C^1 satisfying $x(t_1) = x_1$, $x(t_2) = x_2$ whose derivatives $\dot{x}$ take values in Ω_2. The set $\mathcal{X}$ is an open subset of the normed affine space $O + \mathbf{X}$, where $\mathbf{X}$ is equipped with the norm $\|\mathbf{h}\|_1 := \sup_{t \in [t_1,t_2]} \left(|\mathbf{h}(t)| + \left| \dot{\mathbf{h}}(t) \right| \right)$; $\mathbf{X}$ is a Banach space whose elements satisfy the condition $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$ (**exercise**). Let

$$J : \mathcal{X} \to \mathbb{R} : x \mapsto \int_{t_1}^{t_2} \mathcal{L}(t, x(t), \dot{x}(t)) \, .dt.$$

THEOREM 1.44.– *Let* $x^* \in \mathcal{X}$ *be a mapping of class* C^2[9]. *The relation* $\delta J(x^*) = 0$ *holds, i.e.* x^* *is an extreme point of* J, *if and only if* x^* *is a solution of the* Euler–Lagrange equation*:*

$$\boxed{\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x} = 0.} \qquad [1.21]$$

PROOF.– Let $\mathbf{h} \in \mathbf{X}$. If $\varepsilon > 0$ is sufficiently small, then $x^* + \varepsilon.\mathbf{h} \in \mathcal{X}$. Thus,

$$\frac{1}{\varepsilon}\left(J(x^* + \varepsilon.\mathbf{h}) - J(x^*)\right) = \frac{1}{\varepsilon} \int_{t_1}^{t_2} \left(\mathcal{L}\left(t, x^* + \varepsilon.\mathbf{h}, \dot{x}^* + \varepsilon.\dot{\mathbf{h}}\right) - \mathcal{L}(t, x^*, \dot{x}^*) \right) .dt.$$

---

9 Using a slightly different proof based on the du Bois–Reymond lemma instead of the fundamental lemma of the calculus of variations, it can be shown that the statement of this theorem remains true if we only assume that x^* is of class C^1 (see the Wikipedia article on the *Fundamental lemma of calculus of variations*).

Let $f(\varepsilon,t) = \mathcal{L}\left(t, x^*(t) + \varepsilon.\mathbf{h}(t), \dot{x}^*(t) + \varepsilon.\dot{\mathbf{h}}(t)\right)$. The mapping $(\varepsilon,t) \mapsto f(\varepsilon,t)$ has a partial derivative $\frac{\partial f}{\partial \varepsilon}$ whose absolute value is upper bounded by a function $g : t \mapsto g(t)$ that is integrable in $[t_1, t_2]$. Therefore, as $\varepsilon \to 0$, $\frac{1}{\varepsilon}(J(x^* + \varepsilon.\mathbf{h}) - J(x^*))$ converges ([P2], section 4.1.2(**II**), Theorem 4.11) to

$$\begin{aligned}
\delta J(x^*)[\mathbf{h}] &= \int_{t_1}^{t_2} \frac{\partial}{\partial \varepsilon}\left(\mathcal{L}\left(t, x^*(t) + \varepsilon.\mathbf{h}(t), \dot{x}^*(t) + \varepsilon.\dot{\mathbf{h}}(t)\right)\right)\Big|_{\varepsilon=0}.dt \\
&= \int_{t_1}^{t_2}\left(\frac{\partial \mathcal{L}}{\partial x}\Big|_{x=x^*}.\mathbf{h}(t) + \frac{\partial \mathcal{L}}{\partial \dot{x}}\Big|_{x=x^*}.\dot{\mathbf{h}}(t)\right).dt \\
&= \int_{t_1}^{t_2}\left(\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right)\right)\Big|_{x=x^*}.\mathbf{h}(t).dt,
\end{aligned}$$

which gives the desired result by Lemma 1.43. ■

The Lagrangian $\mathcal{L}$ is said to be *regular* at $x = x^*$ if $\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2}(t, x^*(t), \dot{x}^*(t))$ is invertible in $\mathcal{L}_{2s}(\mathbf{E};\mathbb{R})$ for every $t \in [t_1, t_2]$.

COROLLARY 1.45.– (Hilbert) *If $x^* \in \mathcal{X}$ is a solution of the Euler–Lagrange equation [1.21] and the Lagrangian is regular at $x = x^*$, then x^* is of class C^2.*

PROOF.– The Euler–Lagrange equation can be written as $\varphi(t,x,u,q) = 0$, $\varphi(t,x,u,q) = \frac{\partial \mathcal{L}}{\partial u} - q$, $q = \int \dfrac{\partial \mathcal{L}}{\partial x}.dt + C^{te}$; φ is of class C^1 and such that $\frac{\partial \varphi}{\partial u} = \frac{\partial^2 \mathcal{L}}{\partial u^2}$. If the Lagrangian is regular at the point x^*, then, for any $t \in [t_0, t_1]$, the implicit function theorem (Theorem 1.30) implies that there exists an open neighborhood of $(t, x^*(t), \dot{x}^*(t))$ in $[t_1, t_2] \times \Omega_1 \times \Omega_2$, in which the relation $\varphi(t,x,u,q) = 0$ is equivalent to $u = \psi(t,x,q)$, where ψ is of class C^1; thus, $\dot{x}^*(t) = \psi(t, x^*(t), \dot{x}^*(t))$ is of class C^1, and x^* is of class C^2. ■

(III) LEGENDRE CONDITION Suppose that $\mathcal{L}$ is of class C^2 and x^* satisfies the Euler–Lagrange equation. Expanding to second order, we have

$$\begin{aligned}
\delta^2 J(x^*)[\mathbf{h}] &= \int_{t_1}^{t_2}\left(\left(\frac{\partial^2 \mathcal{L}}{\partial x^2}\right).(\mathbf{h},\mathbf{h}) + 2\frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}}\left(\mathbf{h},\dot{\mathbf{h}}\right) + \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2}\right).\left(\dot{\mathbf{h}},\dot{\mathbf{h}}\right)\right).dt \\
&= \int_{t_1}^{t_2}\left(\left(\underbrace{\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2}}_{P}\right).\left(\dot{\mathbf{h}},\dot{\mathbf{h}}\right) + \left(\underbrace{\left(\frac{\partial^2 \mathcal{L}}{\partial x^2}\right) - \frac{d}{dt}\left(\frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}}\right)}_{Q}\right).(\mathbf{h},\mathbf{h})\right).dt,
\end{aligned}$$

where P and Q are evaluated at the point $(t, x^*(t))$ and $\mathbf{h}$ is evaluated at the point t.

THEOREM 1.46.– (Legendre) *A* necessary condition *for x^* to minimize J over $\mathcal{X}$ is given by the* weak Legendre condition*:*

$$\boxed{P(t, x^*) \geq 0, \quad t_1 \leq t \leq t_2.}$$

PROOF.– Suppose that there exists $t_0 \in \left]t_1, t_2\right[$ such that $P(t_0, x^*(t_0)) \ngeq 0$. This means that there is $\mathbf{v} \in \mathbf{E}^\times$ such that $P(t_0, x^*(t_0)).(\mathbf{v}, \mathbf{v}) = -2\beta$ with $\beta > 0$. Since $t \mapsto P(t, x^*(t)).(\mathbf{v}, \mathbf{v})$ is continuous, there exists a real number $\alpha > 0$ such that $t_1 \leq t_0 - \alpha < t_0 + \alpha \leq t_2$ and $P(t, x^*(t)).(\mathbf{v}, \mathbf{v}) \leq -\beta$ for all $t \in [t_0 - \alpha, t_0 + \alpha]$. We will construct a function $\rho : [t_1, t_2] \to \mathbb{R}$ for which the above integral is < 0. Let $\mathbf{h}(t) = \lambda(t).\mathbf{v}$ with $\lambda(t) = \sin^2\left(\frac{\pi(t-t_0)}{\alpha}\right)$ for $t \in [t_0 - \alpha, t_0 + \alpha]$ and $\lambda(t) = 0$ for $t \in [t_1, t_2] - [t_0 - \alpha, t_0 + \alpha]$. It is easy to see that λ is of class C^1 in $[t_1, t_2]$ and **(exercise)**

$$\int_{t_1}^{t_2} \left(P.(\mathbf{v}, \mathbf{v}).\dot{\lambda}^2 + Q.(\mathbf{v}, \mathbf{v}).\lambda^2\right) dt < -\frac{\beta.\pi^2}{\alpha} + 2M.\alpha,$$

where $M = \sup_{t \in [t_1, t_2]} Q(t, x^*(t)).(\mathbf{v}, \mathbf{v})$. For α sufficiently small, the right-hand side is < 0. After replacing $\mathbf{h}$ by $\varepsilon.\mathbf{h}$ if necessary, where $\varepsilon > 0$ is sufficiently small, $x^*(t) + \varepsilon.\mathbf{h}(t) \in \Omega_1$ and $\dot{x}^*(t) + \varepsilon.\dot{\mathbf{h}}(t) \in \Omega_2$ for all $t \in [t_1, t_2]$. The necessity of the weak Legendre condition then follows from Theorem 1.42. ■

REMARK 1.47.– *If* $\mathbf{E}$ *is finite-dimensional, the* strong Legendre condition *can be stated as* $P(t, x^*(t)) > 0$ $(t_1 \leq t \leq t_2)$*, i.e.* $P(t, x^*(t)).(\mathbf{v}, \mathbf{v}) > 0$ *for all* $\mathbf{v} \neq 0$. *If the strong Jacobi condition*[10] *is also assumed, we obtain the sufficient condition for a "weak" local minimum proved by Weierstrass in 1879.*

1.3.3. *"Convenient" differentials*

(I) $\mathfrak{c}^\infty\mathbf{E}$ TOPOLOGY Let $\mathbf{E}$ be a *real* locally convex space and suppose that A is a non-empty subset of $\mathbf{E}$. A *smooth curve* in A is defined as a mapping $c : I \to A$ of class C^∞, where I is a non-empty open interval of $\mathbb{R}$, for example $\left]-1, 1\right[$. Any such curve is said to be *analytic* if it is of class C^ω.

If $\mathbf{E}$ is a *complex* locally convex space, suppose again that A is a non-empty subset of $\mathbf{E}$. Then, A can be viewed as a subset A_0 of the real locally convex space $\mathbf{E}_0$ obtained from $\mathbf{E}$ by restriction of the field of scalars ([P2], section 3.2.2**(II)**), and any smooth curve in A_0 is said to be a smooth curve in A.

DEFINITION 1.48.– *[KRI 97] Let* $\mathbf{E}$ *be a locally convex space. The topology* $\mathfrak{c}^\infty$ *of* $\mathbf{E}$ *is the finest topology that makes all smooth curves in* $\mathbf{E}$ *(section 1.3.3) continuous. When equipped with this topology, the space* $\mathbf{E}$ *is denoted by* $\mathfrak{c}^\infty\mathbf{E}$.

10 See the Wikipedia article quoted above.

The topology of $\mathfrak{c}^\infty\mathbf{E}$ is finer than the topology of $\mathbf{E}$, so every open subset of $\mathbf{E}$ is open in $\mathfrak{c}^\infty\mathbf{E}$. If the space $\mathbf{E}$ is bornological ([P2], section 3.4.4**(I)**, Definition 3.61), it has the finest locally convex topology of all locally convex topologies coarser than the topology of $\mathfrak{c}^\infty\mathbf{E}$ ([KRI 97], Corollary 4.6). "Convenient" differential calculus is performed with mappings defined in open sets of $\mathfrak{c}^\infty\mathbf{E}$.

Recall that every Fréchet space and every Silva space is bornological and complete ([P2], sections 3.4.4**(I)** and 3.8.2**(II)**). We have the following result ([KRI 97], Theorem 4.11):

THEOREM 1.49.– *Let* $\mathbf{E}$ *be a metrizable locally convex space or a Silva space. Then, the topology of* $\mathbf{E}$ *coincides with the topology of* $\mathfrak{c}^\infty\mathbf{E}$*.*

REMARK 1.50.– *If a locally convex space* $\mathbf{E}$ *is bornological but* non-normable*, then* $\mathfrak{c}^\infty\left(\mathbf{E}\times\mathbf{E}^\vee\right)$ is not *a topological vector space ([KRI 97], Corollary 4.21).*

(II) $(\mathcal{KM})$ **SPACES**

DEFINITION 1.51.– *[KRI 97] A locally convex space* $\mathbf{F}$ *is said to be* convenient *if it is Mackey-complete, i.e. if the normed vector space* $\mathbf{F}_B$ *(Definition 1.19) is a Banach space for every bounded, balanced and convex set* $B\subset\mathbf{F}$*.*

The following definition will be useful:

DEFINITION 1.52.– *A locally convex space* $\mathbf{E}$ *is called a* $(\mathcal{KM})$ *space if it is quasi-complete and bornological, and the topologies of* $\mathbf{E}$ *and* $\mathfrak{c}^\infty\mathbf{E}$ *coincide.*

Theorem 1.49 shows that Fréchet (and in particular Banach spaces) and Silva spaces are $(\mathcal{KM})$ spaces. Any quasi-complete locally convex space, and hence any $(\mathcal{KM})$ space, is convenient (**exercise**). Nonetheless, $(\mathcal{KM})$ spaces are sufficiently general for our purposes, so we will restrict attention to them to simplify the statements of results.

(II) MAPPINGS OF CLASS $\mathfrak{c}^r$ $(r\in\{\infty,\omega\})$

DEFINITION 1.53.– *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be real locally convex* $(\mathcal{KM})$ *spaces and let* A *be an open subset of* $\mathbf{E}$*. A mapping* $\mathbf{f}$ *from* A *into* $\mathbf{F}$ *is said to be of class* $\mathfrak{c}^\infty$ *if* $\mathbf{f}\circ c$ *is a smooth curve in* $\mathbf{F}$ *for any smooth curve* c *in* A*. This mapping* $\mathbf{f}$ *is said to be of class* $\mathfrak{c}^\omega$ *if it is of class* $\mathfrak{c}^\infty$ *and* $\mathbf{f}\circ c$ *is an analytic curve in* $\mathbf{F}$ *for every analytic curve* c *in* A*.*[11]

If $\mathbf{E}$ is a Banach space and $\mathbf{f}$ is of class C^r $(r\in\{\infty,\omega\})$, then $\mathbf{f}$ is of class $\mathfrak{c}^r$. J. Boman showed the following result in 1967 ([KRI 97], Corollary 3.14):

11 In [KRI 97], the classes $\mathfrak{c}^\infty$ and $\mathfrak{c}^\omega$ are written as C^∞ and C^ω, respectively. In this book, we need to choose different notation to avoid ambiguity.

THEOREM 1.54.– (Boman) *Let A be a non-empty open subset of $\mathbb{R}^n$. The mapping $\mathbf{f} : A \to \mathbb{R}^m$ is of class C^∞ if and only if it is of class $\mathfrak{c}^\infty$.*

THEOREM 1.55.– *Let $\mathbf{E}$ and $\mathbf{F}$ be $(\mathcal{KM})$ spaces and suppose that A is an open subset of $\mathbf{E}$.*

1) Suppose that $\mathbf{f} : A \to \mathbf{F}$ is of class $\mathfrak{c}^\infty$. Then, $\mathbf{f}$ has a Gateaux differential $D^{\mathbf{G}}\mathbf{f}(a) \in \mathcal{L}(\mathbf{E};\mathbf{F})$ at every point of A and the mapping $D^{\mathbf{G}}\mathbf{f} : A \to \mathcal{L}(\mathbf{E};\mathbf{F})$ is of class $\mathfrak{c}^\infty$.

2) Let $\mathbf{G}$ be a locally convex space and suppose that $\mathbf{g} : B \to \mathbf{G}$ is a mapping of class $\mathfrak{c}^\infty$, where B is an open subset of $\mathbf{F}$ containing $\mathbf{f}(A)$. Then, the following chain differentiation rule holds (compare with [1.5] and [1.19]):

$$D^{\mathbf{G}}(\mathbf{g} \circ \mathbf{f})(a) = D^{\mathbf{G}}\mathbf{g}(\mathbf{f}(a)) \circ D^{\mathbf{G}}\mathbf{f}(a).$$

PROOF.– (1) Let $\mathbf{h} \in \mathbf{E}$; there exists $\varepsilon > 0$ such that $a + t\mathbf{h} \in A$ for all $t \in]-\varepsilon, +\varepsilon[$ and the curve $t \mapsto a + t\boldsymbol{h}$ is smooth. Therefore, $\mathbf{f}$ has a Lagrange first variation $\delta\mathbf{f}(a)$. It can be shown that $\delta\mathbf{f}(a)$ is bounded linear ([KRI 97], Chapter I, 3.18), so $\delta\mathbf{f}(a) = D^{\mathbf{G}}\mathbf{f}(a)$. Furthermore, $D^{\mathbf{G}}\mathbf{f}(a)$ is continuous linear, since $\mathbf{E}$ and $\mathbf{F}$ are bornological ([P2], section 3.4.4**(I)**, Theorem 3.62). If c is a smooth curve in A, then $\delta\mathbf{f}(a) \circ c$ is a smooth curve in $\mathcal{L}(\mathbf{E};\mathbf{F})$, which gives the stated result by induction (*ibid.*). (2): *ibid.* ■

In the real analytic case, the following result ([KRI 97], Chapter II, section 10.4) generalizes Boman's theorem:

THEOREM 1.56.– *Let $\mathbf{E}$ and $\mathbf{F}$ be real $(\mathcal{KM})$ spaces, U a non-empty open subset of $\mathbf{E}$ and $\mathbf{f} : U \to \mathbf{F}$ a mapping. The following conditions are equivalent: (i) $\mathbf{f}$ is of class $\mathfrak{c}^\omega$; (ii) $\mathbf{f}$ is of class $\mathfrak{c}^\infty$ and $\lambda \circ \mathbf{f} \circ \mu$ is analytic from Ω into $\mathbb{R}$ for any linear form $\lambda \in \mathbf{F}^\vee$ and any affine mapping $\mu : \Omega \to U$ ($\Omega = \mu^{-1}(U)$).*

(III) HOLOMORPHIC MAPPINGS Assume that $\mathbb{K} = \mathbb{C}$.

A *holomorphic curve* in a non-empty open subset A of a *complex* $(\mathcal{KM})$ space $\mathbf{E}$ is a holomorphic mapping from a disk in the complex plane, for example the open disk $\mathbb{D}$ of center 0 and radius 1, into A. A $\mathfrak{c}$-*holomorphic* mapping (or a mapping of class $\mathfrak{c}^\omega$) from A into $\mathbf{F}$, where $\mathbf{F}$ is a complex $(\mathcal{KM})$ space, is a mapping that transforms the holomorphic curves in A into holomorphic curves in $\mathbf{F}$. With this notation, the mapping $\mathbf{f} : A \to \mathbf{F}$ is holomorphic if and only if $\lambda \circ \mathbf{f} \circ c$ is a holomorphic function for every continuous linear form $\lambda \in \mathbf{F}^\vee$ and every holomorphic curve $c : \mathbb{D} \to A$. The following result ([KRI 97], Chapter II, section 7.9) generalizes the classical Hartogs theorem ([P2], section 4.3.2**(II)**, Corollary 4.80):

THEOREM 1.57.– (generalized Hartogs theorem) *Let $\mathbf{E}_1, \mathbf{E}_2, \mathbf{F}$ be complex $(\mathcal{KM})$ spaces and suppose that $A = A_1 \times A_2$ is a non-empty open subset of $\mathbf{E}_1 \times \mathbf{E}_2$. The* *mapping* $\mathbf{f} : A \to \mathbf{F}$ *is* $\mathfrak{c}$*-holomorphic if and only if it is* $\mathfrak{c}$*-holomorphic separately in each variable, i.e.* $\mathbf{f}(., z_2)$ *and* $\mathbf{f}(z_1, .)$ *are* $\mathfrak{c}$*-holomorphic for every* $z_i \in A_i$ $(i = 1, 2)$.

1.4. Smooth partitions of unity

In this section, $\mathbb{K} = \mathbb{R}$.

1.4.1. C^∞-paracompactness of Banach spaces

Let $\mathbf{E}$ be a normed vector space. The topological notions of normal space and paracompact space ([P2], sections 2.3.10 and 2.3.11) inspire the following definitions:

DEFINITION 1.58.– *i)* $\mathbf{E}$ *is said to be* C^r*-normal* $(0 \le r \le \infty)$ *if, given any two disjoint closed subspaces* A *and* B *of* $\mathbf{E}$*, there exists a mapping* $f : X \to [0,1]$ *of class* C^r *taking the values* 1 *in* A *and* 0 *in* B*.*

ii) $\mathbf{E}$ *is said to be* C^r*-paracompact* $(0 \le r \le \infty)$ *if it admits partitions of unity of class* C^r*, i.e. given any open covering* $(U_i)_{i \in I}$ *of* X*, there exists a subordinate partition of unity* $(\psi_i)_{i \in I}$ *such that each* ψ_i *is of class* C^r*.*

Recall that $(\psi_i)_{i \in I}$ is a subordinate partition of unity of $(U_i)_{i \in I}$ if and only if $\operatorname{supp}(\psi_i) \subset U_i$, the family $(\operatorname{supp}(\psi_i))_{i \in I}$ is locally finite and $\sum_{i \in I} \psi_i = 1$ ([P2], section 2.3.12).

THEOREM 1.59.– *The normed vector space* $\mathbf{E}$ *is* C^r*-paracompact if and only if it is* C^r*-normal.*

PROOF.– The necessary condition is clear. We will show the sufficient condition: let $(U_i)_{i \in I}$ be an open covering of $\mathbf{E}$. If $\mathbf{E}$ is paracompact, there exists a locally finite open covering $(V_j)_{i \in J}$ finer than $(U_i)_{i \in I}$ such that each V_j is contained in some $U_{i(j)}$. There also exist other open refinements $(W_j)_{j \in J}$ and $(Z_j)_{j \in J}$ of $(V_j)_{i \in J}$ such that $\overline{Z_j} \subset W_j \subset \overline{W_j} \subset V_j$. If $\mathbf{E}$ is C^r-normal, then, for all j, there exists a function $\psi_j : X \to [0,1]$ of class C^r that is equal to 1 in $\overline{Z_j}$ and 0 in $\complement_{\mathbf{E}} V_j$. Let $\psi = \sum_j \psi_j$ and $\theta_j = \psi_j / \psi$. Thus, $(\theta_j)_{j \in J}$ is a subordinate partition of unity of class C^r of $(V_j)_{i \in J}$ and hence of $\left(U_{i(j)}\right)_{j \in J}$, which shows that $\mathbf{E}$ is C^r-paracompact. ■

REMARK 1.60.– *If a Banach space* $\mathbf{E}$ *is* C^r*-paracompact, then every vector subspace of* $\mathbf{E}$ *and every open set of* $\mathbf{E}$ *is* C^r*-paracompact ([ABR 83], Proposition 5.5.20; [TOR 73], Corollary 1 of Theorem 1).*

As a metrizable space, every Banach space is paracompact by Stone's theorem ([P2], section 2.3.10, Theorem 2.57). The next result ([ABR 83], Propositions 5.5.18

and 5.5.19), whose proof was established by Bonic and Frampton in 1966, follows from the fact that any separable Banach space is a Lindelöf space ([P2], section 2.6.3) (**exercise**).

THEOREM 1.61.– *For a* separable *Banach space to be C^r-paracompact* $(1 \leq r \leq \infty)$, *it is sufficient for its norm* $|.| : x \mapsto |x|$ *to be of class* C^r *in* $\mathbf{E} - \{0\}$.

COROLLARY 1.62.– *Every separable Hilbert space* $\mathbf{E}$ *is* C^∞*-paracompact.*

PROOF.– Write N for the norm of $\mathbf{E}$. We have $N(x)^2 = \langle x|x\rangle$, so $2N(x)\,DN(x).\mathbf{h} = 2\langle x, |\mathbf{h}\rangle$, and if $x \neq 0$, $DN(x).\mathbf{h} = \langle x, |\mathbf{h}\rangle / N(x)$. Hence, N is differentiable in $\mathbf{E} - \{0\}$ and $DN(x) = \langle x|.\rangle / N(x)$ $(x \neq 0)$. This mapping DN is continuous from $\mathbf{E} - \{0\}$ into $\mathbf{E}^\vee \cong \mathbf{E}$. It is easy to see that it is differentiable from $\mathbf{E} - \{0\}$ into $\mathbf{E}^\vee$, and so we can deduce by induction (**exercise**) that N is of class C^∞. ■

It was shown in [TOR 73] that every reflexive (not necessarily separable) Banach space is C^1-paracompact and that the separability condition is not required in Corollary 1.62:

THEOREM 1.63.– *Every Hilbert space* $\mathbf{E}$ *is* C^∞*-paracompact.*

This implies the result stated in ([P2], section 4.4.1, Theorem 4.88):

COROLLARY 1.64.– (Whitney's theorem) *The space* $\mathbb{R}^n$ *is* C^∞*-paracompact.*

1.4.2. $\mathfrak{c}^\infty$*-paracompactness*

We can define $\mathfrak{c}^\infty$-regularity, $\mathfrak{c}^\infty$-normality and $\mathfrak{c}^\infty$-paracompactness of a locally convex space in the obvious ways; the statement of Theorem 1.59 still holds if C^∞ is replaced by $\mathfrak{c}^\infty$; moreover, if $\mathfrak{c}^\infty\mathbf{E}$ is a $\mathfrak{c}^\infty$-regular Lindelöf space, then it is $\mathfrak{c}^\infty$-paracompact ([KRI 97], Proposition 16.2). We already know that any nuclear space E has a topology defined by a family of pre-Hilbert norms ([P2], section 3.11.3(**III**)). Each of these semi-norms is of class $\mathfrak{c}^\infty$ in $E - \{0\}$. Nuclear Fréchet spaces (also known as $(\mathcal{FN})$ spaces), nuclear Silva spaces (also known as $(\mathcal{SN})$ spaces), and countable products of $(\mathcal{FN})$ spaces and $(\mathcal{SN})$ spaces are paracompact Lindelöf spaces (*ibid.*). The following result is analogous to Corollary 1.62 ([KRI 97], Theorem 16.10):

THEOREM 1.65.– *Every* $(\mathcal{FN})$ *space, every strict inductive limit of* $(\mathcal{FN})$ *spaces and every* $(\mathcal{SN})$ *space is* $\mathfrak{c}^\infty$*-paracompact.*

1.5. Ordinary differential equations

1.5.1. *Existence and uniqueness theorems*

(I) Notion of the solution of a differential equation Let I be an interval of $\mathbb{R}$ with non-empty interior $\mathring{I}$, Ω a non-empty open subset of $\mathbf{E} = \mathbb{R}^n$ and $\mathbf{f}$ a mapping from $I \times \Omega$ into $\mathbf{E}$. We say that a mapping $\varphi : I \to \mathbf{E}$ is a solution (or integral) of the differential equation

$$\boxed{\dot{x} = \mathbf{f}(t, x)} \qquad [1.22]$$

if the conditions $(\mathbf{ODE}_{1,2,3})$ below are satisfied:

$(\mathbf{ODE}_1)$ $\varphi(t) \in \Omega$ for all $t \in I$;

$(\mathbf{ODE}_2)$ φ is locally absolutely continuous in I (i.e. each of its components with respect to the canonical basis of $\mathbb{R}^n$ is locally absolutely continuous: see [P2], section 4.1.7**(III)**);

$(\mathbf{ODE}_3)$ $\dot{\varphi}(t) = \mathbf{f}(t, \varphi(t))$ λ-almost everywhere in I, where λ is the Lebesgue measure on $\mathbb{R}$ ([P2], section 4.1.1**(II)**).

A *Cauchy problem* involves determining a function φ that is a solution of [1.22] and that satisfies the *Cauchy condition*:

$$\varphi(t_0) = x_0 \quad (t_0 \in \mathring{I}, x_0 \in \Omega). \qquad [1.23]$$

If φ satisfies the conditions $(\mathbf{ODE}_{1,2,3})$ and [1.23], then, for all $t \in I$,

$$\varphi(t) = x_0 + \int_{t_0}^{t} \mathbf{f}(\tau, \varphi(\tau))\, d\tau. \qquad [1.24]$$

Conversely, suppose that [1.24] holds, $(\mathbf{ODE}_1)$ is satisfied and $t \mapsto \mathbf{f}(t, \varphi(t))$ is locally λ-integrable. Then, [1.23] holds; furthermore, Lusin's measurability criterion ([P2], section 4.1.6**(II)**) implies that the function $t \mapsto \mathbf{f}(t, \varphi(t))$ is λ-measurable for every locally absolutely continuous function $\varphi : I \to \mathbf{E}$ if the conditions $(\mathbf{Cat}_{1,2})$ below are satisfied:

$(\mathbf{Cat}_1)$ the function $x \mapsto \mathbf{f}(t, x)$ from Ω into $\mathbf{E}$ is continuous for every $t \in I$;

$(\mathbf{Cat}_2)$ the function $t \mapsto \mathbf{f}(t, x)$ from I into $\mathbf{E}$ is λ-measurable for every $x \in \Omega$.

Suppose further that:

(Cat$_3$) For all $x_0 \in \Omega$ and every $r > 0$ such that $B_r(x_0) \subset \Omega$, where $B_r(x_0)$ is the open ball of center x_0 and radius r in **E**, there exists a locally λ-integrable function m from I into $\mathbb{R}_+$ such that $|\mathbf{f}(t,x)| \leq m(t)$ for all $(t,x) \in I \times B_r(x_0)$.

Then, for every locally absolutely continuous function $\varphi : I \to \mathbf{E}$ satisfying $\varphi(I) \subset B_r(x_0)$, the mapping $t \longmapsto |\mathbf{f}(t,\varphi(t))|$ is locally λ-integrable in I, so $t \longmapsto \mathbf{f}(t,\varphi(t))$ is locally λ-integrable in I ([P2], section 4.1.2(**I**)).

DEFINITION 1.66.– *The conditions (**Cat**$_{1,2,3}$) are known as the Carathéodory conditions.*

(II) EXISTENCE THEOREM

THEOREM 1.67.– (Carathéodory) *Suppose that the Carathéodory conditions (**Cat**$_{1,2,3}$) are satisfied. Then, for all $(t_0,x_0) \in I \times \Omega$, there exists an interval $J \subset I$ with some point t_0 in its interior and a mapping $\varphi : J \to \mathbf{E}$ such that $\varphi(J) \subset B_r(x_0)$, φ is a solution of [1.22] and this solution satisfies the Cauchy condition [1.23].*

PROOF.– We will show that there exist an interval $J_\beta = [t_0, t_0+\beta] \subset I$ $(\beta > 0)$ and an absolutely continuous mapping $\varphi : J_\beta \to \mathbf{E}$ such that $\varphi(J_\beta) \subset B_r(x_0)$ and φ satisfies [1.24]. The same argument works on an interval $J'_\alpha = [t_0-\alpha, t_0]$ $(\alpha > 0)$. Let $M : [t_0, +\infty[\cap I \to \mathbb{R}_+$ be the mapping:

$$M(t) = 0 \quad (t < t_0), \quad M(t) = \int_{t_0}^{t} m(\tau)\, d\tau \quad (t \in]t_0, +\infty[\cap I).$$

Since M is continuous and non-decreasing, there exists an interval J_β as specified above satisfying the property that, for all $t \in J_\beta$,

$$0 \leq M(t) < r. \qquad [1.25]$$

We can now inductively define a sequence of absolutely continuous mappings $\varphi_i : J_\beta \to \mathbf{E}$ using the conditions:

$$\varphi_i(t) = x_0 \quad (t_0 \leq t \leq \beta/i),$$
$$\varphi_i(t) = x_0 + \int_{t_0}^{t-\beta/i} \mathbf{f}(\tau, \varphi_i(\tau))\, d\tau \quad (t_0 + \beta/i < t \leq t_0 + \beta).$$

By (**Cat**$_3$), the second equation implies that, for all $t \in J_\beta$,

$$|\varphi_i(t) - x_0| \leq \int_{t_0}^{t-\beta/i} |\mathbf{f}(\tau, \varphi_i(\tau))|\, d\tau \leq \int_{t_0}^{t_0+\beta} m(\tau)\, d\tau < r,$$

so $\varphi_i(t) \in B_r(x_0)$ for all $t \in J_\beta$. If $t_1, t_2 \in J_\beta$, then

$$|\varphi_i(t_2) - \varphi_i(t_1)| \leq |M(t_2 - \beta/i) - M(t_1 - \beta/i)|,$$

and M is uniformly continuous in the compact set J_β by Heine's theorem ([P2], section 2.4.5, Theorem 2.86), so $|M(t_2 - \beta/i) - M(t_1 - \beta/i)| \to 0$ uniformly in i if $t_2 - t_1 \to 0$, and the set $H := \{\varphi_i : i \in \mathbb{N}^\times\}$ is equicontinuous ([P2], section 2.7.3). Since $H(t) := \{\varphi_i(t) : i \in \mathbb{N}^\times\}$ is contained in $B_r(x_0)$ for all $t \in J_\beta$, the third Ascoli–Arzelà theorem (*ibid.*) implies that H is relatively compact in $\mathcal{C}(J_\beta; \mathbf{E})$ equipped with the uniform structure of uniform convergence. Hence, there exists a subsequence (φ_{i_k}) that converges uniformly to some mapping $\varphi \in \mathcal{C}(J_\beta; \mathbf{E})$. Moreover, $|\mathbf{f}(t, \varphi_{i_k}(t))| \leq m(t)$ $(t_0 \leq t \leq t_0 + \beta)$ and $\mathbf{f}(t, \varphi_{i_k}(t)) \to \mathbf{f}(t, \varphi(t))$ for $i_k \to \infty$ by $(\mathbf{Cat}_1)$; furthermore, as we saw earlier, $(\mathbf{Cat}_2)$ implies that $t \mapsto \mathbf{f}(t, \varphi(t))$ is measurable. The Lebesgue dominated convergence theorem ([P2], section 4.1.2**(II)**, Theorem 4.9) therefore implies that, for all $t \in J_\beta$,

$$\int_{t_0}^{t} \mathbf{f}(\tau, \varphi_{i_k}(\tau))\, d\tau \to \int_{t_0}^{t} \mathbf{f}(\tau, \varphi(\tau))\, d\tau \quad (i_k \to \infty).$$

But, for all $t \in J_\beta$,

$$\varphi_{i_k}(t) = x_0 + \int_{t_0}^{t} \mathbf{f}(\tau, \varphi_{i_k}(\tau))\, d\tau - \int_{t-\beta/i}^{t} \mathbf{f}(\tau, \varphi_{i_k}(\tau))\, d\tau,$$

and the second integral tends to 0 as $i_k \to \infty$, so the equality [1.24] is satisfied for all $t \in J_\beta$. Finally, $(\varphi_{i_k}) \to \varphi$ in the Banach space $AC(J_\beta; \mathbf{E})$ of absolutely continuous mappings from J_β into $\mathbf{E}$ ([P2], section 4.1.7**(III)**), so φ is absolutely continuous. ■

COROLLARY 1.68.– (Peano's theorem) *Suppose that* $\mathbf{f}$ *is continuous in* $I \times \Omega$*. Let* J *be a compact interval that is a neighborhood of* t_0 *in* I *and let* $m = \sup_{t \in J, x \in B_r(x_0)} |\mathbf{f}(t, x)|$*. For every compact interval* $[t_0, t_0 + \beta]$ *contained in* J *satisfying* $\beta < r/m$*, there exists a solution of [1.24] that takes values in* $B_r(x_0)$*.*

PROOF.– We have $M = m\beta$, so the inequality [1.25] is satisfied if and only if $\beta < r/m$. ■

REMARK 1.69.– *Corollary 1.68 (and hence Theorem 1.67) fails if* $\mathbf{E}$ *is replaced by an arbitrary Banach space ([BOU 76], Chapter 4, section 1, Exercise 18). We can define an absolutely continuous mapping* $\varphi : J_\beta \to \mathbf{E}$ *as we did in ([P2], section 4.1.7**(I)**), but it is not true in general that* $\varphi(t) - \varphi(t_0) = \int_{t_0}^{t} \dot{\varphi}(\tau)\, d\tau$ *for all* $t \in J_\beta$ *(however, this property does hold if* $\mathbf{E}$ *is assumed to be reflexive). Furthermore, since the ball* $B_r(x_0)$ *is not relatively compact when* $\mathbf{E}$ *is infinite-dimensional, the proof of Theorem 1.67 no longer works.*

(III) UNIQUENESS THEOREM The fourth Carathéodory condition can be stated as follows (reusing some of the notation of $\mathbf{Cat}_3$):

($\mathbf{Cat}_4$) For all $x_0 \in \mathbf{E}$ and every real number $r > 0$ such that $B_r(x_0) \subset \Omega$, there exists a locally λ-integrable function k from I into $\mathbb{R}_+$ such that

$$|\mathbf{f}(t, x') - \mathbf{f}(t, x'')| \leq k(t) \,.\, |x' - x''|\,, \quad \forall (t, x'), (t, x'') \in I \times B_r(x_0)\,.$$

THEOREM 1.70.– (Carathéodory) *Suppose that the Carathéodory conditions* $\boldsymbol{Cat}_{1,2,3,4}$ *are satisfied. Then, for all* $(t_0, x_0) \in \mathring{I} \times \Omega$, *there exists an interval* $J \subset I$ *with interior point* t_0 *and a* unique *mapping* $\varphi : J \to \Omega$ *that is a solution of [1.22] and which satisfies the Cauchy condition [1.23].*

PROOF.– Let

$$K(t) = \int_{t_0}^{t_0+t} k(\tau)\, d\tau \quad (0 \leq t \leq b)$$

and choose the real number $\beta > 0$ in the proof of Theorem 1.67 to satisfy the additional condition $K(\beta) < 1$. Furthermore, let

$$\mathcal{X} = \{\psi \in \mathcal{C}([0, \beta], \mathbf{E}) : \psi(0) = 0 \,\&\, \psi(t) \in B_r^c(0), \forall t \in [0, \beta]\}$$

and, for every function $\psi \in \mathcal{X}$, set

$$(T.\psi)(t) := \int_{t_0}^{t_0+t} \mathbf{f}(\tau, \psi(\tau - t_0) + x_0)\, d\tau \quad (0 \leq t \leq \beta)\,.$$

Clearly, $T(\mathcal{X}) \subset \mathcal{X}$. The set $\mathcal{X}$ is a closed subset of the space $\mathcal{C}([0, \beta], \mathbf{E})$ equipped with the uniform structure of uniform convergence; the norm of $\mathcal{C}([0, \beta], \mathbf{E})$ is $\|\psi\|_\infty := \sup_{t \in [0, \beta]} |\psi(t)|$. The space $\mathcal{C}([0, \beta], \mathbf{E})$ is complete ([P2], section 2.7.2, Corollary 2.115), so $\mathcal{X}$ is complete ([P2], section 2.4.4(**II**), Lemma 2.77). Furthermore, $\psi \in \mathcal{C}([0, \beta], \mathbf{E})$ is a fixed point of T if and only if the function $\varphi \in \mathcal{C}([t_0, t_0 + \beta], \mathbf{E})$ defined by

$$\varphi(t) := \psi(t - t_0) + x_0 \quad (0 \leq t \leq \beta)$$

satisfies [1.24]; finally, $\psi(t - t_0) \in B_r(0)$ if and only if $\varphi(t) \in B_r(x_0)$. If $\psi_1, \psi_2 \in \mathcal{X}$ for all $t \in [0, \beta]$, then

$$\begin{aligned}|(T.\psi_1)(t) - (T.\psi_2)(t)| &\leq \int_{t_0}^{t_0+t} k(\tau) \,.\, |\psi_1(\tau - t_0) - \psi_2(\tau - t_0)|\, d\tau \\ &\leq K(b)\, \|\psi_1 - \psi_2\|_\infty\end{aligned}$$

and hence $\|T.\psi_1 - T.\psi_2\|_\infty \leq K(b)\, \|\psi_1 - \psi_2\|_\infty$. Therefore, T has a unique fixed point by Theorem 1.27. ■

DEFINITION 1.71.– *1) We say that* $\mathbf{f} : I \times \Omega \to \mathbf{E}$ *is* locally Lipschitz *in the second variable in* $I \times \Omega$ *if, for every* $(t, x) \in I \times \Omega$*, there exists a neighborhood* V *of* t*, a neighborhood* S *of* x *and a* constant $k_{V,S} > 0$ *such that* $\mathbf{f}(., x)$ *is* regulated *on* V *([DIE 93], Volume 1, section 7.6) and* $\mathbf{f}$ *satisfies the Lipschitz condition:*

$$|\mathbf{f}(t, x') - \mathbf{f}(t, x'')| \leq k_{V,S}. |x' - x''|, \quad \forall t \in V, \forall x', x'' \in S.$$

2) We say that $\mathbf{f}$ *is* Lipschitz *with constant* $k > 0$ *in the second variable if the above statement is satisfied when* $V = I$*,* $S = \Omega$ *and* $k_{V,S} = k$*.*

Any locally Lipschitz function in the second variable clearly satisfies the conditions $\mathbf{Cat}_{1,2,3,4}$, and we may therefore apply Theorem 1.70 to this function. We also have the following result:

COROLLARY 1.72.– (Cauchy–Lipschitz theorem) *If* $\mathbf{f}$ *is Lipschitz with constant* k *in the second variable in* $I \times \Omega$*, let* J *be a compact interval contained in* I *with non-empty interior,* t_0 *a point of* $\mathring{I}$*,* x_0 *a point of* $\mathbf{E}$*,* $r > 0$ *a real number such that* $B_r(x_0) \subset \Omega$*,* $m = \sup_{t \in J, x \in B_r(x_0)} |\mathbf{f}(t, x)|$ *and* $\rho = \min\{r/m, 1/k\}$*. For every compact interval* K *contained in* $J \cap]t_0 - \rho, t_0 + \rho[$*, there exists a unique mapping* φ *that is a solution of [1.22] and which satisfies the Cauchy condition [1.23].*

REMARK 1.73.– *i) Theorem 1.70 fails if* $\mathbf{E}$ *is replaced by an arbitrary Banach space. However, Corollary 1.72 remains valid, with an identical proof. Furthermore,* $\rho = r/m$ *([BOU 76], Chapter 4, section 1.5, Theorem 1)*[12]*. Interested readers can find additional existence and uniqueness results for the solutions of [1.22] in infinite dimensions in [DEI 77], section 8, and [SCH 89]. Note, however, that "infinite-dimensional systems"* are not *governed by a functional differential equation of the form [1.22], where* $\mathbf{E}$ *is a Banach space: see [HAL 77].*

ii) By the mean value theorem (Theorem 1.13), in order for f *to be locally Lipschitz, it is sufficient for it to be of class* C^1*.*

THEOREM 1.74.– *Let* $\mathbf{E}$ *be the space* $\mathbb{R}^n$ *(respectively any Banach space),* $I \subset \mathbb{R}$ *an interval with some interior point* t_0*,* Ω *a non-empty open subset of* $\mathbf{E}$ *and* $\mathbf{f} : I \times \Omega \to \mathbf{E}$ *a mapping satisfying the conditions* $\mathbf{Cat}_{1,2,3,4}$ *(respectively a locally Lipschitz function in the second variable). For all* $x_0 \in \Omega$*, there exists a maximal interval* $J(t_0, x_0) \subset I$ *with interior point* t_0 *in which [1.22] has a solution* φ *satisfying the Cauchy condition [1.23] and such that* $\varphi(J(t_0, x_0)) \subset \Omega$*. This solution* $\varphi(.; t_0, x_0)$ *is unique.*

PROOF.– Let $\mathfrak{M}$ be the set of intervals $L \subset I$ with non-empty interior containing t_0 as an interior point and such that there exists a solution φ of [1.22] in L satisfying the

12 The value $\rho = r/m$ is obtained from the method of comparison of approximate solutions using Gronwall's lemma (*loc. cit.*).

Cauchy condition [1.23] and $\varphi(L) \subset \Omega$. The set $\mathfrak{M}$ is not empty by Theorem 1.70 (respectively Corollary 1.72 and Remark 1.73). Let $L, L' \in \mathfrak{M}$ with $L \subset L'$. If φ, φ' are solutions of [1.22] in L, L', respectively, satisfying the Cauchy condition [1.23], then it follows from Theorem 1.70 (respectively Corollary 1.72 and Remark 1.73) (arguing by contradiction: **exercise**) that φ' is a continuation of φ. Let $J(t_0, x_0) = \bigcup_{L \in \mathfrak{M}} L$; there must therefore exist a unique solution φ of [1.22] in $J(t_0, x_0)$ satisfying the Cauchy condition [1.23] and $\varphi(J(t_0, x_0)) \subset \Omega$. ■

DEFINITION 1.75.– *The solution $\varphi(.; t_0, x_0)$ defined in J is called the* maximal integral *of [1.22] satisfying [1.23].*

REMARK 1.76.– *Suppose that $\mathbf{f}$ is locally Lipschitz in the second variable.*

1) Let $t_f := \sup(J(t_0, x_0)) \leq +\infty$ (the argument below also works when $t_i := \inf(J(t_0, x_0)) \geq -\infty$, mutatis mutandis*). If $\mathbf{f}(t, \varphi(.; t_0, x_0))$ is bounded in $J(t_0, x_0)$, then $\varphi(t; t_0, x_0)$ admits a limit $\mathbf{c} := \varphi(t_f - 0; t_0, x_0)$; furthermore, $\mathbf{c}$ is a frontier point of Ω if $J(t_0, x_0) \cap [t_0, +\infty[\neq I \cap [t_0, +\infty[$. This inequality, together with the condition $\Omega = \mathbf{E}$, implies that $\lim_{t \to t_f} |\varphi(t; t_0, x_0)| = +\infty$; by contrast, if $t_f \in I$ and $\Omega = \mathbf{E}$, then $J(t_0, x_0) \cap [t_0, +\infty[= I \cap [t_0, +\infty[$ ([BOU 76], Chapter 4, section 1.5, Theorems 2 and Corollaries 1 and 2). In addition to this remark, see Theorem 5.67 in section 5.7.1.*

2) Let (τ, ξ) be an arbitrary point of $I \times \Omega$. There exists an interval $K \subset I$, a neighborhood of τ in I and a neighborhood S of ξ in Ω, such that, for every point $(t_0, x_0) \in K \times S$, there is a unique solution $\varphi(., t_0, x_0)$ of [1.22] satisfying [1.23], defined in K (i.e. $J(t_0, x_0) \supset K$). The mapping $(t, t_0, x_0) \mapsto \varphi(t; t_0, x_0)$ from $K \times K \times S$ into Ω is uniformly continuous ([BOU 76], Chapter 4, section 1.7, Theorem 4).

(IV) DIFFERENTIAL EQUATIONS IN IMPLICIT FORM Let I be an open interval of $\mathbb{R}$, t_0 a point of I, $\mathbf{F}$ a Banach space and Ω a non-empty open subset of $\mathbf{F}^{n+1}$, where n is a natural integer. Let $\mathbf{g} : I \times \Omega \to \mathbf{F}$ be a mapping of class C^1 and consider the differential equation:

$$\mathbf{g}\left(t, y, \dot{y}, ..., y^{(n-1)}, y^{(n)}\right) = 0. \qquad [1.26]$$

Let $\eta_0 = \left(\eta_0^0, \eta_0^1, ..., \eta_0^{n-1}\right) \in \mathbf{F}^n$ and $\xi_0 \in \mathbf{F}$ such that $(\eta_0, \xi_0) \in \Omega$, $\mathbf{g}(t_0, \eta_0, \xi_0) = 0$, and suppose that the following condition **(Inv)** is satisfied:

(Inv) $\frac{\partial \mathbf{g}}{\partial \xi}(t_0, \eta_0, \xi_0)$ is invertible in $\mathcal{L}(\mathbf{F})$.

By the implicit function theorem (Theorem 1.30), there exists an open neighborhood $J \times U$ of (t_0, η_0) in $I \times \mathbf{F}^n$, an open neighborhood V of ξ_0 in $\mathbf{F}$, where $U \times V \subset \Omega$, and a mapping $\mathbf{h} : J \times U \to V$ of class C^1, such that, for all $(t, \eta, \xi) \in J \times U \times V$,

$$\mathbf{g}(t, \eta, \xi) = 0 \Longleftrightarrow \xi = \mathbf{h}(t, \eta).$$

Therefore, any mapping $\psi : J \to \mathbf{F}$ such that $\left(\psi(t), \dot{\psi}(t), ..., \psi^{(n-1)}(t)\right) \in U$ and $\psi^{(n)}(t) \in V$ for all $t \in J$ is a solution of [1.26] in J if and only if ψ is a solution of the following differential equation on J, said to be in *explicit form*:

$$y^{(n)} = \mathbf{h}\left(t, y, \dot{y}, ..., y^{(n-1)}\right). \quad [1.27]$$

Let $x_i = y^{(i-1)}$ $(i = 1, ..., n)$ and $x = (x_1, ..., x_n)$. For $(t, x) \in J \times U \times V$, the differential equation [1.27] is equivalent to [1.22], where $\mathbf{f} = (\mathbf{f}_1, ..., \mathbf{f}_n)$ and

$$\mathbf{f}_1(t, x) = x_2, ..., \mathbf{f}_{n-1}(t, x) = x_n, \mathbf{f}_n(t, x) = \mathbf{h}(t, x).$$

The mapping $\mathbf{f} : J \times U \to \mathbf{F}^n$ is of class C^1 and hence locally Lipschitz in x by the mean value theorem (Theorem 1.13(i)). We can therefore apply Theorem 1.72 according to Remark 1.73. Let $x_0 = \left(\eta_0^1, ..., \eta_0^{n-1}\right)$; any solution $\varphi = (\varphi_1, ..., \varphi_n)$ of [1.22] of class C^1 in J satisfies the Cauchy condition [1.23] if and only if $\psi = \varphi_1$ is a solution of [1.26], of class C^n in J and $\psi(t_0) = \eta_0^1, ..., \psi^{(n-1)}(t_0) = \eta_0^{n-1}$.

REMARK 1.77.– *If the condition* ***(Inv)*** *is not satisfied, the differential equation [1.26] is singular. We already encountered this situation in the linear case ([P2], section 5.4.6), where it was necessary to introduce solutions in the form of hyperfunctions. The nonlinear case does not have an equivalent general theory.*

1.5.2. *Linear differential equations*

(I) Let I be an interval of $\mathbb{R}$ with non-empty interior $\mathring{I}$ and let $\mathbf{E} = \mathbb{R}^n$. Consider the linear differential equation

$$\dot{x} = A(t).x + \mathbf{b}(t), \quad [1.28]$$

where $A : I \to \mathcal{L}(\mathbf{E})$ and $\mathbf{b} : I \to \mathbf{E}$ are locally λ-integrable. The Carathéodory conditions $\mathbf{Cat}_{1,2,3,4}$ are all satisfied. (In particular, setting $\mathbf{f}(t, x) = A(t).x + b(t)$, we have

$$|\mathbf{f}(t, x') - \mathbf{f}(t, x'')| \leq \|A(t)\|.|x' - x''|,$$

which shows that $\mathbf{Cat}_4$ is satisfied.) Hence, for all $t_0 \in \mathring{I}$ and every $x_0 \in \mathbf{E}$, there exists a unique solution $\varphi(.; t_0, x_0)$ of [1.28] in I satisfying [1.23].

The linear equation [1.28] is said to be *homogeneous* if $\mathbf{b} = 0$, in which case

$$\dot{x} = A(t).x. \tag{1.29}$$

(II) The set of solutions of [1.29] (where A is locally λ-integrable) is an $\mathbb{R}$-vector space. Let $\varphi_0(.;t_0,x_0)$ be a solution of [1.29] in I satisfying [1.23]. The mapping $x_0 \mapsto \varphi_0(t;t_0,x_0)$ is a bijective linear mapping $\Phi(t,t_0)$ from $\mathbf{E}$ onto $\mathbf{E}$, and $\Phi(.,t_0)$ is identical to the solution of the differential equation

$$\frac{dU}{dt} = A(t)U \tag{1.30}$$

for $U(t_0) = 1_{\mathbf{E}}$. For all $t_1, t_2, t_3 \in I$, we have $\Phi(t_3,t_1) = \Phi(t_3,t_2) \circ \Phi(t_2,t_1)$ and $\Phi(t_1,t_2) = \Phi(t_2,t_1)^{-1}$.

DEFINITION 1.78.– *The mapping Φ is called the* resolvent *of the equation [1.29]. The matrix representing this resolvent with respect to the canonical basis of $\mathbf{E}$ is called the* transition matrix.

THEOREM 1.79.– *We have*

$$\boxed{\det(\Phi(t,t_0)) = \exp\left(\int_{t_0}^{t} Tr(A(\tau)).d\tau\right).}$$

PROOF.– Set $\Phi(t,t_0) = U(t), \Delta(t) = \det(U(t))$ and write $U(t+h) = U(t) + h.V + o(h)$, where $V = \dot{U}(t)$. Then, $\Delta(t+h) = \Delta(t).\det\left(I + h.V.U^{-1}\right.$ $\left.(t) + o(h)\right)$. But

$$\det\left(I + h.V.U^{-1}(t) + o(h)\right) = h^n \det\left(h^{-1}.I + V.U^{-1}(t) + o(1)\right).$$

Furthermore, by [P1], section 2.3.11**(VII)**,

$$\det\left(h^{-1}.I + V.U^{-1}(t) + o(1)\right) = h^{-n}I + h^{-n+1}\mathrm{Tr}\left(V.U^{-1}(t)\right) + o\left(h^{-n+1}\right),$$

so $\det\left(I + h.V.U^{-1}(t) + o(h)\right) = 1 + h.\mathrm{Tr}\left(V.U^{-1}(t)\right) + o(h)$. Hence, $\Delta(t+h) = \Delta(t).\left(1 + h.\mathrm{Tr}\left(V.U^{-1}(t_0)\right) + o(h)\right)$, so

$$\dot{\Delta}(t) = \mathrm{Tr}\left(\dot{U}(t).U^{-1}(t)\right).\Delta(t).$$

We now simply apply the relation $\dot{U}(t)U^{-1}(t) = A(t)$, which follows from [1.30]. ■

(III) The above shows that the general solution of [1.29] is of the form $t \mapsto \Phi(t, t_0)\,\xi$. The "variation of constants" method (**exercise**) allows us to obtain the following solution (defined in I) of [1.28] and the Cauchy condition [1.23]:

$$\boxed{\varphi(t; t_0, x_0) = \Phi(t, t_0).x_0 + \int_{t_0}^{t} \Phi(\tau, t_0).\mathbf{b}(\tau).d\tau.}$$

(IV) The integration of linear differential equations with constant coefficients is a classical problem and is performed using the Jordan normal form ([P1], section 3.4.3(**IV**)): see, for example, [BOU 10], section 12.5.2.

1.5.3. *Parameter dependence of solutions*

Let I be an interval of $\mathbb{R}$ with non-empty interior, $\mathbf{E}$ a Banach space (see Remark 1.73), Ω a non-empty open subset of $\mathbf{E}$, Λ a topological space and $\mathbf{f}$ a mapping from $I \times \Omega \times \Lambda$ into $\mathbf{E}$. Write $\mathbf{f}_\lambda(t, x)$ for the value of $\mathbf{f}$ at the point $(t, x, \lambda) \in I \times \Omega \times \Lambda$. It is possible to show the following result ([BOU 76], Chapter 4, section 1.6, Theorem 3):

THEOREM 1.80.– (parameter dependence of solutions) *Suppose that, for all $\lambda \in \Lambda$, $(t, x) \mapsto \mathbf{f}_\lambda(t, x)$ is Lipschitz in the second variable x in $I \times \Omega$ and that $\mathbf{f}_\lambda(t, x) \to \mathbf{f}_{\lambda_0}(t, x)$ uniformly in $I \times \Omega$ as $\lambda \to \lambda_0$. Let φ_{λ_0} be a solution of $\dot{x} = \mathbf{f}_{\lambda_0}(t, x)$ satisfying the Cauchy condition $\varphi_{\lambda_0}(t_0) = x_0$ $\left(t_0 \in \mathring{I}, x_0 \in \Omega\right)$, defined on an interval $J = [t_0, t_0 + \beta[\; \subset I$ and taking values in Ω. For every compact interval $[t_0, t_1] \subset J$, there exists a neighborhood V of λ_0 in Λ such that, for all $\lambda \in V$, the differential equation*

$$\dot{x} = \mathbf{f}_\lambda(t, x) \qquad [1.31]$$

admits a unique solution φ_λ defined in $[t_0, t_1]$ satisfying the Cauchy condition $\varphi_\lambda(t_0) = x_0$ and taking values in Ω; furthermore, as $\lambda \to \lambda_0$, $\varphi_\lambda \to \varphi_{\lambda_0}$ uniformly in $[t_0, t_1]$.

Suppose now that Λ is an open subset of a normed vector space $\mathbf{F}$ and that the mapping $(t, x, \lambda) \mapsto \mathbf{f}_\lambda(t, x)$ is continuous with continuous partial differentials[13] $(t, x, \lambda) \mapsto \frac{\partial \mathbf{f}_\lambda}{\partial x}(t, x)$ and $(t, x, \lambda) \mapsto \frac{\partial \mathbf{f}_\lambda}{\partial \lambda}(t, x)$. Then, $\mathbf{f}_\lambda$ is locally Lipschitz in the second variable x (see section 1.5.1(**V**)). Suppose further that the mappings $x_0 : \Lambda \to \Omega : \lambda \mapsto x_0(\lambda)$ and $t_0 : \Lambda \to I : \lambda \mapsto t_0(\lambda)$ are of class C^1 in Λ. For all $\lambda_0 \in \Lambda$, Theorem 1.80 implies that there exist an open neighborhood V of λ_0 in Λ and an open interval $J \subset I$ such that $t_0(\lambda_0) \in J$, where the sets V and J

13 See footnote 4, p. 10.

satisfy the following condition: for all $\lambda \in V$, $t_0(\lambda) \in J$ and there exists a solution $\varphi_\lambda = \varphi(., t_0(\lambda), x_0(\lambda))$ of [1.31] in J taking the value $x_0(\lambda) \in \Omega$ at the point $t_0(\lambda)$. Then, we have the following result:

THEOREM 1.81.– *For every $t \in J$, the mapping $\lambda \mapsto \varphi_\lambda(t)$ from V into Ω has a differential $\boldsymbol{h}(t) \in \mathcal{L}(\mathbf{F}; \mathbf{E})$ at the point λ_0 of V; $\boldsymbol{h}$ is of class C^1 in J and is the unique solution of the linear equation:*

$$\dot{\boldsymbol{y}} = A(t) \circ \boldsymbol{y} + \boldsymbol{b}(t) \qquad [1.32]$$

where

$$A(t) := \left.\frac{\partial \mathbf{f}_\lambda}{\partial x}(t, \varphi_\lambda(t))\right|_{\lambda=\lambda_0}, \quad \boldsymbol{b}(t) := \left.\frac{\partial \mathbf{f}_\lambda}{\partial \lambda}(t, \varphi_\lambda(t))\right|_{\lambda=\lambda_0}.$$

This differential $\boldsymbol{h}$ satisfies the Cauchy condition:

$$\boldsymbol{h}(t_0(\lambda_0)) = Dx_0(\lambda_0) - \mathbf{f}_\lambda(t_0(\lambda), x_0(\lambda))|_{\lambda=\lambda_0} \circ Dt(\lambda_0).$$

PROOF.– For all $t \in J$,

$$\varphi_\lambda(t) = x_0(\lambda) + \int_{t_0(\lambda)}^{t} \mathbf{f}_\lambda(\tau, \varphi_\lambda(\tau))\, d\tau.$$

Hence, by [DIE 93], Volume 1, section 8.11, (8.11.2) and Problem 1,

$$\begin{aligned}\boldsymbol{h}(t) = &\int_{t_0(\lambda_0)}^{t} \left[\frac{\partial \mathbf{f}_\lambda}{\partial x}(\tau, \varphi_\lambda(\tau))\, \boldsymbol{h}(\tau) + \frac{\partial \mathbf{f}_\lambda}{\partial \lambda}(\tau, \varphi_\lambda(\tau))\right]_{\lambda=\lambda_0} d\tau \\ &+ Dx_0(\lambda_0) - \mathbf{f}_\lambda(\tau, \varphi_\lambda(\tau))|_{\lambda=\lambda_0} \circ Dt(\lambda_0).\end{aligned}$$

Therefore, $\boldsymbol{h}$ is of class C^1 in J and

$$\begin{aligned}\dot{\boldsymbol{h}}(t) &= \left.\frac{\partial \mathbf{f}_\lambda}{\partial x}(t, \varphi_\lambda(t))\right|_{\lambda=\lambda_0} \boldsymbol{h}(t) + \left.\frac{\partial \mathbf{f}_\lambda}{\partial \lambda}(\tau, \varphi_\lambda(\tau))\right|_{\lambda=\lambda_0}, \\ \boldsymbol{h}(t_0(\lambda_0)) &= Dx_0(\lambda_0) - \mathbf{f}_\lambda(\tau, \varphi_\lambda(\tau))|_{\lambda=\lambda_0} \circ Dt(\lambda_0).\end{aligned}$$

■

Note that, in the linear equation [1.32], $A(t) \in \mathcal{L}(\mathbf{E})$; moreover, $\boldsymbol{y}$, $A(t) \circ \boldsymbol{y}$, and $\boldsymbol{b}(t)$ belong to $\mathcal{L}(\mathbf{F}; \mathbf{E})$. In the Cauchy condition, $\mathbf{f}_\lambda(t_0(\lambda), x_0(\lambda))|_{\lambda=\lambda_0} \in \mathbf{E} \cong \mathcal{L}(\mathbb{R}; \mathbf{E})$, $Dx_0(\lambda_0) \in \mathcal{L}(\mathbf{F}; \mathbf{E})$, $Dt(\lambda_0) \in \mathcal{L}(\mathbf{F}, \mathbb{R}) = \mathbf{F}^\vee$, and $\mathbf{f}_\lambda(t_0(\lambda), x_0(\lambda))|_{\lambda=\lambda_0} \circ Dt(\lambda_0) \in \mathcal{L}(\mathbf{F}; \mathbf{E})$.

COROLLARY 1.82.– (differentiability of the solution with respect to x_0) *Suppose that* $\mathbf{F} = \mathbf{E}$*,* $\Lambda = \Omega$*, and* $\lambda = x_0$*, and adopt the same hypotheses as Theorem 1.81 (*mutatis mutandis*). Write* $\mathbf{f}$ *and* φ *for* $\mathbf{f}_\lambda$ *and* φ_λ *respectively. Then:*

$$\boxed{\frac{\partial \varphi}{\partial x_0}(t; t_0, x_0) = \Phi(t, t_0)}$$

where Φ *is the resolvent (Definition 1.78) of the linear differential equation*

$$\dot{\boldsymbol{y}} = A(t) \circ \boldsymbol{y}, \quad A(t) := \frac{\partial \mathbf{f}}{\partial x}(t, \varphi(t; t_0, x_0)). \tag{1.33}$$

PROOF.– By Theorem 1.81, $\dot{\boldsymbol{h}}(t) = A(t)\,\boldsymbol{h}(t)$ and $\boldsymbol{h}(t_0) = 1_{\mathbf{E}}$. ■

COROLLARY 1.83.– (differentiability with respect to the initial time) *Suppose that* $\mathbf{F} = \mathbb{R}$*,* $\Lambda = I$*, and* $\lambda = t_0$*. With the notation of Corollary 1.82 and the same hypotheses (**exercise**):*

$$\boxed{\frac{\partial \varphi}{\partial t_0}(t; t_0, x_0) = -\Phi(t, t_0)\,\mathbf{f}(t_0, x_0).}$$

REMARK 1.84.– *Let* $\mathbf{E}$ *be finite-dimensional. The statement of Corollary 1.82 still holds under the weaker condition ([ALE 87], Chapter 2, section 2.5.6) that* $\mathbf{f}$ *satisfies the Carathéodory conditions (**Cat**$_{1,2,3}$) and* Lusin's condition *(**L**) holds:*

*(**L**): For all* $t \in I$*, the mapping* $x \mapsto \mathbf{f}(t, x)$ *is continuously differentiable on* Ω *and, for every compact set* $K \subset \Omega$*, there exists a locally* λ*-integrable function* $k : I \to \mathbb{R}_+$ *such that*

$$\left\| \frac{\partial \mathbf{f}}{\partial x}(t, x) \right\| \leq k(t), \quad \forall (t, x) \in I \times K.$$

*It is clear that condition (**L**) implies (**Cat**$_4$) by the mean value theorem. This condition does not require* $\mathbf{f}$ *to be continuous in* t*, which is very important in optimal control theory (Pontryagin maximum principle): see* op. cit. *and [PON 62]. We can further weaken (**L**) by considering generalized gradients and differential inclusions ([CLA 90], Theorem 7.4.1), but this exceeds the scope of the present book.*

This page intentionally left blank

2

Differential and Analytic Manifolds

2.1. Introduction

The notion of a differential manifold proposed by Riemann in his inaugural lecture in 1854 ([SPI 99], Volume 2, Chapter 4, section A) is a generalization of curves in the plane, or curves or surfaces in everyday space; these curves and surfaces (extensively studied by Gauss, who supervised Riemann's thesis), and more generally manifolds, are assumed to be regular. For a curve, being regular means having a tangent at every point; for a surface, it means having a tangent plane at every point. Locally, each curve "looks like" its tangent and each surface "looks like" its tangent plane. Manifolds can be studied locally using charts in the same way that we study regions of our planet geographically. It was Gauss who suggested charts to study curves and surfaces locally. We cannot represent the whole Earth with a single chart unless we are willing to accept significant deformations of some of its regions, for example the poles in the classical Mercator projection[1]. Hence, we require more than one chart; these charts are assembled into an atlas.

In general, manifolds do not have any algebraic properties. Lie groups (section 2.4.1) are an exception: they have a group structure that is compatible with their manifold structure. A Lie group $\mathbf{G}$ can act on a manifold, determining orbits in the latter, and has a canonical action on its homogeneous spaces $\mathbf{G}/\mathbf{H}$ (where $\mathbf{H}$ is a subgroup of $\mathbf{G}$) ([P1], section 2.2.8(**II**)). Lie groups originate from the prodigious but obscure works of S. Lie, which began around 1873 after fruitful discussions with F. Klein, the author who proposed the renowned Erlangen program[2] in 1872. After 1888, S. Lie clarified some of his work with help from F. Engel. Over time, an entire cohort of mathematicians led by W. Killing, É. Cartan and H. Weyl expanded this work into one of the most imposing monuments of mathematics. Readers who are

1 See the Wikipedia article on the *Mercator projection*.

2 See the Wikipedia article on the *Erlangen program*.

interested in the historical details can refer to the excellent historical notes in [BOU 82b] (Chapters 3–4 and Chapters 5–6) and [HAW 00]. Lie groups are omnipresent in contemporary theoretical physics [GEO 82], mechanics, and of course differential geometry, as we will see in Chapter 7.

Manifolds were originally studied in finite dimensions before mathematicians, including J. Eells [EEL 58], showed that the Banach framework (systematically presented in [LAN 62, BOU 82a]) was appropriate for studying global analysis, at the turn of the 1960s. This perspective was later widely embraced [EEL 66, PAL 68, MAR 74]. Infinite-dimensional Hilbert manifolds have better properties than general Banach manifolds: whenever they are paracompact, they admit C^∞ partitions of unity (Corollary 2.15); subbundles are straightforward to define whenever the base is a Hilbert manifold (Lemma-Definition 3.40(1)), since every closed subspace of a Hilbert space splits ([P2], section 3.10.2(**II**)), and thus any such subbundle always splits (Theorem 3.47). Infinite-dimensional Hilbert manifolds play an essential role in global analysis, especially for applications of Morse theory in global calculus of variations [PAL 63, PAL 64]; nonetheless, whenever these manifolds are separable, their structure is relatively poor, since they are always parallelizable (Theorem 3.29). Moreover, if $\mathbf{E}$ and $\mathbf{F}$ are two Hilbert spaces, U is an open subset of $\mathbf{E}$, and $\mathbf{f}$ is a mapping of class C^r $(r \geq 1)$ from U into $\mathbf{F}$, then the differential $D\mathbf{f}$ is a mapping of class C^{r-1} from U into $\mathcal{L}(\mathbf{E};\mathbf{F})$, but the latter space is not a Hilbert space in general. The Hilbert framework is therefore unsuitable for differential calculus of order ≥ 2. Banach Lie groups were introduced by Birkhoff in 1936 [BIR 38] and are systematically presented in [BOU 82b] (Chapter 3).

The manifolds modeled by Fréchet nuclear spaces provide a highly relevant alternative approach to problems of global analysis [KRI 97]. We will only briefly mention this approach and its key results in order to avoid excessively cluttering the text, but interested readers are welcome to explore it further in the cited reference. Diffieties are a particular type of Fréchet manifold that are still the object of extensive theoretical research by Vinogradov and his collaborators; they have recently been applied to the theory of dynamical systems [FLI 97], but are beyond the scope of this book.

2.2. Manifolds: tangent space of a manifold at a point

2.2.1. *Notion of a manifold*

(I) Topological manifolds Let M be a set and $\mathfrak{S}$ a set of topological vector spaces over $\mathbb{K}$. An *atlas* $\mathcal{A}$ of type $\mathfrak{S}$ and class C^0 on M is a family of pairs $(U_i, \varphi_i)_{i\in I}$ satisfying the following conditions (see Figure 2.1):

(AT1) Each U_i is a subset of M and $(U_i)_{i\in I}$ is a covering of M.

(AT2) Each φ_i is a bijection from U_i onto an open set $\Omega_i = \varphi_i(U_i)$ of a topological vector space $E_i \in \mathfrak{S}$ and, for all $i, j \in I$, the set $\varphi_i(U_i \cap U_j)$ is open in E_i.

(AT3) The mapping $\varphi_i^j = \varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is a homeomorphism, called a *transition homeomorphism*.

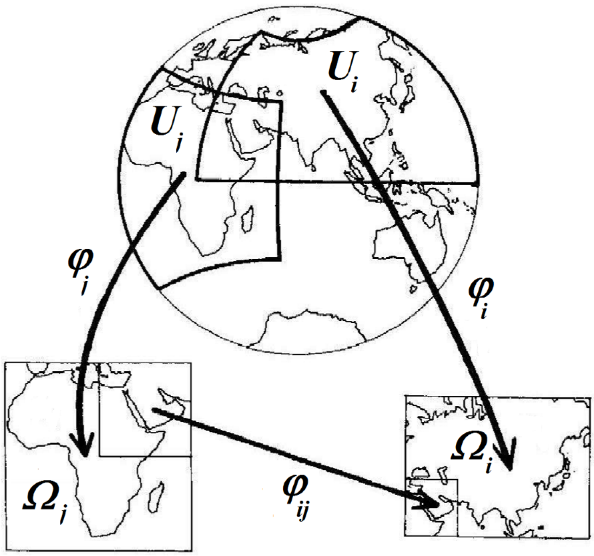


Figure 2.1. *Charts and transition morphisms*

Each triple $(U_i, \varphi_i, \mathbf{E}_i)$ is called a *chart* of $\mathcal{A}$. This chart is said to be centered on an element x of M if $\varphi_i(x) = 0$. The subsets of M that can be expressed as a union of the U_i are the open sets of a topology on M. Each φ_i is then a homeomorphism from U_i onto the open set $\Omega_i = \varphi_i(U_i)$ of $\mathbf{E}_i$.

Two atlases $\mathcal{A}$ and $\mathcal{B}$ are said to be C^0-equivalent if $\mathcal{A} \cup \mathcal{B}$ is also an atlas. C^0-equivalence is an equivalence relation on atlases of class C^0, and the topology on M defined above only depends on the C^0-equivalence class of the atlas $\mathcal{A}$.

DEFINITION 2.1.– *The set M, equipped with the topology defined above, is called a* topological $\mathbb{K}$-manifold *(or simply a* topological manifold *if $\mathbb{K}$ is clear from context) and $\mathcal{A}$ is said to be an atlas of the manifold M. The U_i are called the* affine open sets *of M. We say that the manifold M is* modeled *on the topological vector spaces in $\mathfrak{S}$, and that these spaces* model *M.*

LEMMA 2.2.– *Let M be a manifold with an atlas of type $\mathfrak{S}$, where $\mathfrak{S}$ is a non-empty set of locally convex spaces. The manifold M is locally connected. If the topological vector spaces in $\mathfrak{S}$ are* Hausdorff, *then M is connected if and only if it is arc-connected.*

PROOF.– The manifold M is a locally connected topological space, since each $\mathbf{E}_i$ is a connected space. Every arc-connected topological space is connected ([P2], section 2.3.8, Corollary 2.47). Conversely, suppose that M is connected and the $\mathbf{E}_i$ are Hausdorff. Let $x \in M$ and write A_x for the set of points of M that can be reached by finitely many arcs $B_i = [x_i, x_{i+1}]$, each contained in some affine open set U_i. Since B_i and B_{i+1} intersect at x_{i+1}, A_x must be arc-connected. Furthermore, A_x must be closed, since, if $y \in \bar{A}_x$ and V is an affine open neighborhood of y, then V intersects non-trivially with A_x, so there exists $z \in A_x \cap V$ such that z can be reached from y by an arc, and hence $y \in A_x$. Similarly, A_x is open, since, if $y \notin A_x$, there exists an affine open set V containing y such that $A_x \cap V = \varnothing$. Thus, $A_x = M$. ■

A manifold M is a uniformizable space ([P2], section 2.4.1(**II**)); thus, it is metrizable if and only if every point of M admits a countable fundamental system of neighborhoods.

Let $x \in M$ and suppose that $(U_i, \varphi_i, \mathbf{E}_i)$, $(U_j, \varphi_j, \mathbf{E}_j)$ are two charts of M such that $x \in U_i \cap U_j$. We know from **(AT3)** that $\varphi_i^j := \varphi_j \circ \varphi_i^{-1}$ is a homeomorphism from $\mathbf{E}_i$ onto $\mathbf{E}_j$. If one of these spaces has a finite dimension n, the Brouwer invariance of dimension theorem[3] implies that the other must also have dimension n. The dimension of M at the point x, written as $\dim_x(M)$, is defined as the dimension of $\mathbf{E}_i$ whenever this is finite. If not, $\dim_x(M) = +\infty$. If $x, y \in U_i$, then $\dim_x(M) = \dim_y(M)$, so the mapping $x \mapsto \dim_x(M)$ from M into the discrete space $\bar{\mathbb{N}}$ (see Theorem 1.35(iii)) is continuous. This mapping is therefore constant in each connected component M^0 of M. The *dimension* of M is defined as the quantity $\dim(M) := \sup_{x \in M} \dim_x(M)$. A manifold satisfying $\dim_x(M) < +\infty$ at every point $x \in M$ is said to be *locally finite-dimensional*. If $\mathbf{E} = \mathbb{K}^m$, the chart $(U, \varphi, \mathbf{E})$ is written as (U, φ, m).

DEFINITION 2.3.– *A manifold M is said to be* pure *if it has an atlas of type $\mathfrak{S}$ where each element of $\mathfrak{S}$ is isomorphic to the same topological vector space $\mathbf{E}$ (any such atlas is also said to be an $\mathbf{E}$-atlas). A manifold M is said to be* locally pure *if every connected component of M is pure.*

EXAMPLE 2.4.– *1) Let $\mathbf{E}$ be a topological vector space. Then, $(\mathbf{E}, 1_{\mathbf{E}}, \mathbf{E})$ is a chart of $\mathbf{E}$ which has the canonical structure of a pure manifold of type $\mathbf{E}$.*

*2) Let M be a set. There exists precisely one manifold structure on M for which the underlying topological space is discrete ([P2], section 2.3.1(**III**)); this is namely the structure of a pure manifold of dimension* 0.

3 See the Wikipedia article on *Invariance of domain*.

If $\mathfrak{S}$ consists of Hilbert spaces (respectively Banach spaces, respectively Fréchet spaces), then M is said to be a Hilbert manifold (respectively Banach manifold, respectively Fréchet manifold). Similarly, if $\mathfrak{S}$ consists of $(\mathcal{KM})$ spaces (respectively $(\mathcal{FN})$ spaces, respectively $(\mathcal{SN})$ spaces, etc.), then M is said to be a $(\mathcal{KM})$ manifold (respectively $(\mathcal{FN})$ manifold, respectively $(\mathcal{SN})$ manifold, etc.).

Any Fréchet manifold M is a locally metrizable Baire space ([P2], section 2.6.1); Stone's theorem and Smirnov's theorem ([P2], section 2.3.10) imply the following result:

COROLLARY 2.5.– *A Fréchet manifold M is metrizable if and only if it is paracompact.*

The following result was shown by R. Palais in 1966 ([PAL 66], Corollary of Theorem 3):

THEOREM 2.6.– (R. Palais) *Every paracompact Banach manifold is homeomorphic to a complete metrizable space.*

LEMMA 2.7.– *A* Hausdorff *topological manifold M is locally finite-dimensional if and only if it is locally compact. If this condition is satisfied and M is also paracompact and connected, then M is furthermore countable at infinity.*

PROOF.– The Hausdorff manifold M is locally finite-dimensional if and only if it is locally compact by Riesz's theorem ([P2], section 3.2.3, Theorem 3.11). Let $(U_x)_{x\in M}$ be a family of relatively compact open sets in M such that $x \in M$. By paracompactness, there exists a locally finite open covering $(V_i)_{i\in I}$ that is finer than $(U_x)_{x\in M}$. Let V_1 be one of these open sets and consider $K_1 = \overline{V_1}$. Since K_1 is closed and contained in one of the compact sets $\overline{U_x}$, it must also be compact ([P2], section 2.3.7). For every integer $p \geq 2$, let $I_p \subset I$ be a set of indices i such that $V_i \cap K_{p-1} = \varnothing$, and write $K_p = \bigcup_{i\in I_p} \overline{V_i}$. Then, K_p is compact, as a finite union of compact sets. Write $M_1 = \bigcup_{p\geq 1} K_p$. We have $K_{p-1} \subset \bigcup_{i\in I_p} V_i \subset K_p$, so $M_1 = \bigcup_{p\geq 1} V_i$ is open in M. Furthermore, M_1 is closed in M : if $x \in \overline{M_1}$, then x belongs to one of the V_i, say V_n, and therefore $x \in \bigcup_{p\geq 1} V_i = M_1$. Hence, since M is connected, $M = M_1$, which is a countable union of compact sets. ■

More precisely, it can be shown ([BOU 74], Chapter I, section 9.10, Proposition 18) that a locally compact topological space (e.g. a locally finite-dimensional Hausdorff manifold) is paracompact if and only if each of its connected components is countable at infinity.

COROLLARY 2.8.– *If a Hausdorff locally finite-dimensional manifold M is separable, then it is paracompact, its connected components are at most countable, and each connected component is locally compact and countable at infinity.*

PROOF.– Since the manifold M is Hausdorff and locally finite-dimensional, it is locally compact (Lemma 2.7). If it is separable, it is paracompact ([P2], section 2.3.10). Since M is locally connected, its connected components form a partition of open sets (**exercise***: see [BOU 74], Chapter I, section 11.6, Proposition 11); each of these open sets contain at least one open set of a base of the topology of M. This base is countable because M is separable ([P2], section 2.3.1(**IV**)), so the connected components of M are at most countable. Furthermore, each of these connected components is locally compact and countable at infinity by Lemma 2.7. ■

THEOREM 2.9.– *A manifold M is paracompact if and only if each of its connected components is paracompact.*

PROOF.– If the manifold M is paracompact, then it is metrizable (Corollary 2.5), so each of its connected components is metrizable and hence paracompact. Conversely, we saw in the proof of Corollary 2.8 that the connected components M_i $(i \in I)$ of M are open in M. Let $(U_\lambda)_{\lambda\in\Lambda}$ be an open covering of M. The covering formed by the open sets $M_i \cap U_\lambda$ is finer than $(U_\lambda)_{\lambda\in\Lambda}$. If M_i is paracompact for each $i \in I$, there exists a locally finite open covering $(V_{i,j})_{j\in J_i}$ that is finer than $(M_i \cap U_\lambda)_{\lambda\in\Lambda}$, and the open covering of M formed by the $V_{i,j}$ $(i \in I, j \in J_i)$ is locally finite and finer than $(U_\lambda)_{\lambda\in\Lambda}$. Hence, M is paracompact. ■

(II) DIFFERENTIAL OR ANALYTIC MANIFOLDS Topological manifolds are very "soft", like the watches painted by Salvador Dalí. They are even more elastic than rubber: both the hexagonal border of France and the boot-shaped border of Italy are homeomorphic to a circle. The manifolds of class C^r or $\mathfrak{c}^r$ $(r > 0)$ which are studied in the next section are barely any more rigid (they will only rigidify once equipped with a connection to determine their geometry: see Chapter 7); but they are no longer elastic and are smoother (a hedgehog's back is not diffeomorphic to a disk of the plane as a differential manifold but is homeomorphic to this disk as a topological manifold), which allows them to serve as a basis for differential calculus, unlike topological manifolds.

Let $r \in \mathbb{N}_{\mathbb{K}}^{\times}$ (section 1.2.1(**I**)). We can define an atlas $\mathcal{A}$ of type $\mathfrak{S}$ and class C^r on a set M when $\mathfrak{S}$ is a set of Banach spaces: we simply assume that the transition homeomorphisms φ_j^i are *diffeomorphisms* of class C^r, i.e. both these homeomorphisms and their inverse homeomorphisms are of class C^r. By adapting **(I)** accordingly, we can obviously deduce the notions of C^r-equivalence of atlases and (Banach) manifolds of class C^r.

DEFINITION 2.10.– *i) Let M be a* real *Banach manifold of class C^r $\left(r \in \mathbb{N}_{\mathbb{R}}^{\times}\right)$. If $r \leq \infty$, we say that M is a* differential manifold *of class C^r. If $r = \omega$, we say that M is an* analytic manifold.

ii) A complex *Banach manifold of class C^ω is said to be* holomorphic.

If $\mathfrak{S}$ is a set of $(\mathcal{KM})$ spaces (Definition 1.52), then we can obviously define atlases and real manifolds of class $\mathfrak{c}^r$ for $r \in \{\infty, \omega\}$, as well as complex manifolds of class $\mathfrak{c}^\omega$.

In the following, to simplify the presentation, differential manifolds are always assumed to be of class C^∞ unless otherwise stated. For each result, the reader may wish to determine the minimum value of $r \in \mathbb{N}_{\mathbb{R}}^{\times}$ for which the calculations are valid as an exercise. In other words, we will adopt the following convention:

(C1) Unless otherwise stated, the letter r denotes some element of $\mathbb{N}_{\mathbb{K}}^{\times} \cap \{\infty, \omega\}$. All diffeomorphisms are of class C^r.

LEMMA 2.11.– *Every differential, analytic or holomorphic manifold M is locally pure (Definition 2.3).*

PROOF.– For clarity, suppose that M is Banach. Let $(U_i, \varphi_i, \mathbf{E}_i)$ and $(U_j, \varphi_j, \mathbf{E}_j)$ be two charts of M such that $U_i \cap U_j \neq \varnothing$. Let x be a point of $U_i \cap U_j$. Then, $D\varphi_i^j(x) \in \mathcal{L}(\mathbf{E}_i; \mathbf{E}_j)$ is an isomorphism (in the category $\mathbf{Tvs}$ of topological vector spaces) that identifies $\mathbf{E}_i$ and $\mathbf{E}_j$ with the same Banach space $\mathbf{E}$. By a reasoning similar to the proof of Lemma 2.2, the set of points $x \in M$ for which there exists a chart $(U_i, \varphi_i, \mathbf{E}_i)$ centered on M and such that the Banach space $\mathbf{E}_i$ is isomorphic to $\mathbf{E}$ is open and closed in M. ■

EXAMPLE 2.12.– *The Earth is flat to the first order of approximation (as was believed to be the case in ancient times). To the second order of approximation, it is a sphere of radius $R \cong 6{,}500$ km (according to the calculation performed by Eratosthenes around 200 BC). Any point P on the Earth's surface can be referenced by its longitude λ and its latitude φ (see Figure 2.2). If we choose the usual origin O of space to be the center of the Earth, then the Cartesian coordinates of the point P (in an orthonormal frame of reference) are given by:*

$$x = -R.\cos\varphi.\cos\lambda, \quad y = -R.\cos\varphi.\sin\lambda, \quad z = R.\sin\varphi. \qquad [2.1]$$

A geographic map (i.e. a chart *in the mathematical terminology introduced above) of the region U of the Earth located between the longitudes λ_0 and λ_1 and the latitudes φ_0 and φ_1, where $-180^\circ < \lambda_0 < \lambda_1 < 180^\circ$ and $-90^\circ < \varphi_0 < \varphi_1 < 90^\circ$, is therefore a bijection $\Phi : U \to]\lambda_0, \lambda_1[\times]\varphi_0, \varphi_1[$ that determines a chart $(U, \Phi, 2)$. If the Earth were perfectly spherical, any two such charts could be chosen to be compatible of class C^ω, in which case the Earth would be a real analytic manifold. This no longer holds if we take its surface irregularities into account.*

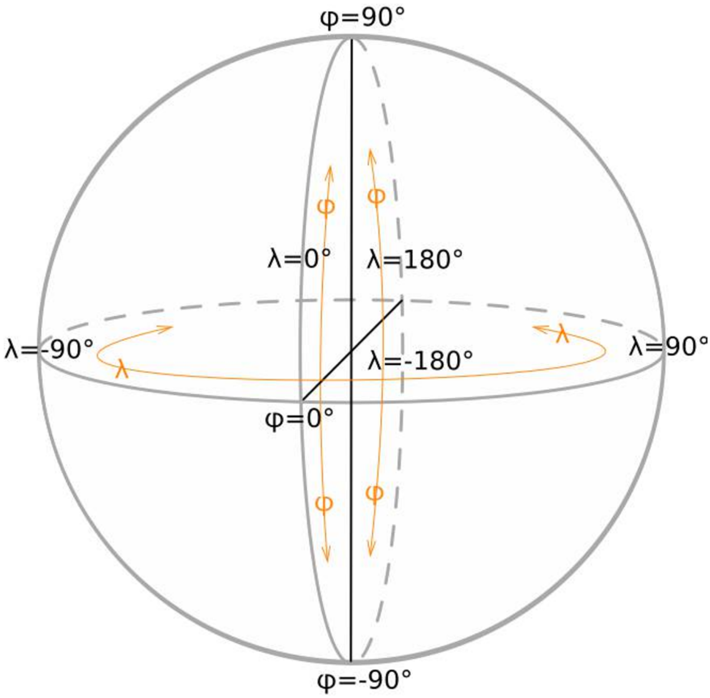


Figure 2.2. *Longitude and latitude. For a color version of this figure, see www.iste.co.uk/bourles/fundamentals3.zip*

2.2.2. *Morphisms of manifolds*

(I) Let M and N be Banach manifolds (differential, analytic, or holomorphic) and suppose that $f : M \to N$ is a mapping. Consider two charts $(U, \xi, \mathbf{E})$ and $(V, \eta, \mathbf{F})$, of M and N respectively, satisfying $f(U) \subset V$. The mapping

$$\Phi = \eta \circ (f|_U) \circ \xi^{-1} : \mathbf{E} \supset \xi(U) \to \eta(V) \subset \mathbf{F}$$

is called the *expression of* f *in these charts*. We say that f is of class C^p in U for $p \in \mathbb{N}_{\mathbb{K}}^{\times}, p \leq r$, if Φ is of class C^p in $\xi(U)$. If $p = r = \omega$, we say that f is *analytic* (or *holomorphic* if $\mathbb{K} = \mathbb{C}$). The composition of two mappings of class C^r is itself of class C^r. Hence, the $\mathbb{K}$-manifolds of class C^r form a *category* whose *morphisms* are the mappings of class C^r ([P1], section 1.1.1(**I**)). An *isomorphism* of manifolds of class C^r is simply a diffeomorphism of class C^r ([P1], section 1.1.1(**III**)). The Banach manifolds of class C^r form a concrete category with base $\mathbf{Set}$. This observation gives meaning to the idea of the *structure* of a Banach manifold of class C^r ([P1], section 1.3.1).

(II) If M and N are $(\mathcal{KM})$ manifolds, we can similarly define mappings of class $\mathfrak{c}^r$, morphisms of manifolds of class $\mathfrak{c}^r$ and the notion of a structure of manifolds of class $\mathfrak{c}^r$.

The notion of C^∞-paracompactness (respectively $\mathfrak{c}^\infty$-paracompactness) (section 1.4) can obviously be extended to Banach manifolds (respectively $(\mathcal{KM})$ manifolds).

(III) The same reasoning as the proof of Theorem 1.59 gives the following result ([ABR 83], Proposition 5.5.16):

THEOREM 2.13.– (Palais) *A paracompact manifold of class C^∞ modeled on a Banach space $\mathbf{E}$ is C^∞-paracompact if and only if $\mathbf{E}$ is C^∞-paracompact. If a manifold M is C^∞-paracompact, every submanifold of M is C^∞-paracompact.*

The sufficient condition of Theorem 2.13 can be improved as indicated in (1) below:

THEOREM 2.14.– *1) Any paracompact manifold of class C^∞ that has an atlas of type $\mathfrak{S}$ consisting of C^∞-paracompact Banach spaces is itself C^∞-paracompact.*

2) Every metrizable manifold of class $\mathfrak{c}^\infty$ and type $(\mathcal{FN})$ or $(\mathcal{SN})$ is $\mathfrak{c}^\infty$-paracompact.

PROOF.– 1) By the same reasoning as the proof of Theorem 2.9, a manifold is C^∞-paracompact if and only if each of its connected components is C^∞-paracompact. But each of these connected components is pure (Lemma 2.11); for each component to be C^∞-paracompact, it is sufficient for each of the Banach spaces $\mathbf{E} \in \mathfrak{S}$ to be C^∞-paracompact.

2) See [KRI 97], section 27.4, Corollary. ■

COROLLARY 2.15.– *Every paracompact Hilbert manifold of class C^∞ is C^∞-paracompact.*

COROLLARY 2.16.– *Every separable locally compact manifold of class C^∞ is C^∞-paracompact.*

PROOF.– Any finite-dimensional vector space is C^∞-paracompact by Corollary 1.62 (or Whitney's theorem: see Corollary 1.64), so the result follows from Corollary 2.8.■

The notion of a C^∞-normal manifold can be defined in the same way as the notion of a C^∞-normal vector space, *mutatis mutandis* (Definition 1.58). The same reasoning as Theorem 1.59 gives the following result:

COROLLARY 2.17.– *A Banach manifold is C^∞-paracompact if and only if it is C^∞-normal.*

Up to and including section 5.5, we will adopt the following convention:

(C2) Unless otherwise stated, the word *manifold* denotes a *differential or analytic Banach manifold* (rather than a topological manifold), i.e. a Banach manifold of class C^r in accordance with the convention **(C1)** of section 2.2.1**(III)**. The word *morphism*, unless otherwise stated, denotes a *mapping of class* C^r. In particular, all diffeomorphisms are of class C^r.

2.2.3. *Tangent mappings*

Let M, N be two manifolds, $a \in M$, and $f_{1,}f_2 : M \to N$ two morphisms. These morphisms are said to be *tangent* at the point a if:

i) $f_1(a) = f_2(a) = b$;

ii) there exist a chart $c = (U, \xi, \mathbf{E})$ of M centered on a and a chart $c' = (V, \eta, \mathbf{F})$ of N centered on b such that $f_1(U) \cap f_2(U) \subset V$ and the mappings

$$\Phi_i = \eta \circ f_i \circ \xi^{-1} : \mathbf{E} \supseteq \xi(U) \to \eta(V) \subset \mathbf{F} \quad (i = 1, 2)$$

have the same differential at the point a. This is an equivalence relation.

REMARK 2.18.– *(i) Unless the manifolds M and N are pure,* $\mathbf{E}$ *and* $\mathbf{F}$ *depend on a and b, respectively. (ii) Naturally, $D\Phi_i(a) \in \mathcal{L}(\mathbf{E}; \mathbf{F})$, the space of continuous linear mappings from* $\mathbf{E}$ *into* $\mathbf{F}$.

DEFINITION 2.19.– *The mapping $\Phi_i : \xi(U) \to \eta(V)$ is called the expression of f_i in the charts c, c'.*

Adapting the above to the case where M, N are $(\mathcal{KM})$ manifolds is left to the reader. In this case, the Fréchet differential operator D is replaced by the Gateaux differential operator $D^{\mathcal{G}}$.

2.2.4. *Tangent vectors*

We will give four equivalent definitions of a tangent vector. Each of them is useful.

(I) FIRST DEFINITION

DEFINITION 2.20.– *Let M be a manifold of class C^r (respectively $\mathfrak{c}^r$). A* curve *of class C^p (respectively $\mathfrak{c}^p$) ($p \in \mathbb{N}_{\mathbb{K}}, p \leq r$) in M is a mapping of class C^p (respectively $\mathfrak{c}^p$) from a non-empty connected open subset of $\mathbb{K}$ into M.*

In the following, every curve is of class C^r or $\mathfrak{c}^r$, depending on the context.

Let M be a manifold and γ_1, γ_2 two curves in M passing through $a = \gamma_1(0) = \gamma_2(0)$. These curves are tangent at 0 if and only if there exists a chart $(U, \xi, \mathbf{E})$ of M centered on a such that

$$D(\xi \circ \gamma_1)(0) = D(\xi \circ \gamma_2)(0).$$

This defines an equivalence relation $\mathcal{R}_1$ of curves in M passing through a.

DEFINITION 2.21.– *The equivalence classes of the relation* $\mathcal{R}_1$ *are called the* tangent vectors *at the point* a.

Write $T_a(M)$ for the set of all tangent vectors at the point $a \in M$.

Let γ be a curve in M passing through a and write $\mathring{\gamma}$ for the equivalence class of this curve. Clearly, $D(\eta \circ \gamma)(0)$ only depends on $\mathring{\gamma}$, so there exists a well-defined mapping

$$\theta_c : T_a(M) \to \mathbf{E} : \mathring{\gamma} \mapsto D(\xi \circ \gamma)(0).$$

LEMMA-DEFINITION 2.22.– *The mapping* θ_c *is linear and bijective. Its inverse bijection is written as* ϑ_c.

PROOF.– It is clear that θ_c is linear and injective. We need to show that it is also surjective. Let $\vartheta_c : \mathbf{E} \to T_a(M)$ be the mapping that sends each $\mathbf{h} \in \mathbf{E}$ to the equivalence class (mod.$\mathcal{R}_1$) of the curve $\gamma : t \mapsto \xi^{-1}(t\mathbf{h})$ (where t belongs to a sufficiently small neighborhood of 0 in $\mathbb{K}$). Then, $\mathring{\gamma} \in T_a(M)$ and $\mathbf{h} = D(\xi \circ \gamma)(0) = \theta_c(\mathring{\gamma})$, so θ_c is bijective. ■

Hence, the vector spaces $T_a(M)$ and $\mathbf{E}$ are isomorphic, and we can transport the topology of $\mathbf{E}$ onto $T_a(M)$ using ϑ_c, which makes $T_a(M)$ a Banach space isomorphic to $\mathbf{E}$.

DEFINITION 2.23.– *The Banach space* $T_a(M)$ *is called the* tangent space *of the manifold* M *at the point* a *(see Figure 2.3).*

A similar construction can be performed in the case where M is a $(\mathcal{KM})$ manifold by replacing the Fréchet differential D by the Gateaux differential $D^{\mathcal{G}}$.

(II) SECOND DEFINITION Let M be a manifold and $a \in M$. Consider the pairs $(c, \mathbf{h})$, where $c = (U, \xi, \mathbf{E})$ is a chart of M centered on a and $\mathbf{h} \in \mathbf{E}$. Any two such pairs $(c, \mathbf{h}), (c', \mathbf{h}')$ are said to be equivalent (mod.$\mathcal{R}$) if

$$D\left(\xi' \circ \xi^{-1}\right)(\xi(a)).\mathbf{h} = \mathbf{h}'.$$

It should immediately be clear that $\mathcal{R}$ is an equivalence relation.

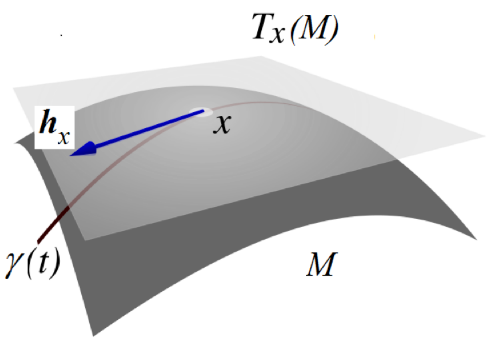


Figure 2.3. *Tangent space and tangent vector*

DEFINITION 2.24.– *The equivalence class (mod.$\mathcal{R}$) of a pair $(c, \mathbf{h})$ is called a tangent vector of M at the point a.*

As before, write $T_a(M)$ for the set of tangent vectors of M at a. Write also ϑ_c for the mapping from $\mathbf{E}$ into $T_a(M)$ that sends each vector $\mathbf{h}$ to the tangent vector represented by the pair $(c, \mathbf{h})$. The set $T_a(M)$ can be equipped with a Banach space structure isomorphic to $\mathbf{E}$ by transporting its structure.

REMARK 2.25.– *After fixing the chart c centered on a, picking a vector $\mathbf{h} \in \mathbf{E}$ is equivalent to picking a curve $\gamma : t \mapsto \xi^{-1}(t\mathbf{h})$ defined in a neighborhood of 0 in $\mathbb{K}$. Thus, Definitions 2.21 and 2.24 of a tangent vector are equivalent.*

Adapting the above to the case where M is a $(\mathcal{KM})$ manifold is left to the reader.

(III) THIRD DEFINITION Let M be a manifold and $a \in M$. Let $C_a^r(M)$ be the set of germs of functions taking values in $\mathbb{K}$ that are of class C^r in a neighborhood of a ([P2], section 2.3.13). In the following, we will reference a germ by one of its representatives (Ω, f), where Ω is an open neighborhood of a and $f : \Omega \to \mathbb{K}$ is of class C^r.

Let $(U, \xi, \mathbf{E})$ be a chart of M centered on a. If (Ω_1, f_1) and (Ω_2, f_2) are two representatives of the same germ, then we clearly have

$$D\left(f_1 \circ \xi^{-1}\right)(0) = D\left(f_2 \circ \xi^{-1}\right)(0) \in \mathbf{E}^{\vee},$$

a quantity that is therefore "intrinsic" in the sense that it only depends on the choice of germ and not on any particular representative of this germ.

LEMMA-DEFINITION 2.26.– *i) Let* (Ω, f) *be a germ of a function of class* C^r *at* $a \in M$*. This germ is said to be* stationary *at* a *if there exists a chart* $(U, \xi, \mathbf{E})$ *of* M *centered on* a *such that*

$$D\left(f \circ \xi^{-1}\right)(0) = 0.$$

ii) If $(U_1, \xi_1, \mathbf{E}_1)$ *is another chart of* M *centered on* a *(which implies that* $\mathbf{E} \cong \mathbf{E}_1$*), then we also have* $D\left(f \circ \xi_1^{-1}\right)(0) = 0$.

iii) The set of germs of functions of class C^r *that are stationary at* a *is written as* $S_a^r(M)$.

iv) $C_a^r(M)$ *and* $S_a^r(M) \subset C_a^r(M)$ *are* $\mathbb{K}$*-vector spaces.*

Claims (ii) and (iv) are immediate.

DEFINITION 2.27.– *Suppose that* M *is* locally finite-dimensional*; let* $c = (U, \xi, m)$ *be a chart of* M *centered on a,* $p^j : \mathbb{K}^m \to \mathbb{K}$ *the* j*-th projection, and* $\xi^j = p^j \circ \xi : U \to \mathbb{K}$*. We say that* $\left(\xi^j\right)_{1 \leq j \leq m}$ *is the system of* local coordinates *associated with the chart c.*

Let $f : U \to \mathbb{K}$ be a function of class C^r and $F = f \circ \xi^{-1} : \mathbb{K}^m \supset \xi(U) \to \mathbb{R}$. For every $x \in U$, we have $f(x) = F(\xi(x))$, where $\xi(x) \in \xi(U)$. Let $(\mathbf{e}_j)_{1 \leq j \leq m}$ be the canonical basis of $\mathbb{K}^m$; then $p^j(\mathbf{e}_i) = \delta_i^j$ and $\xi(x) = \sum_{1 \leq j \leq m} \xi^j(x)\, \mathbf{e}_j$. The ξ^j can either be viewed as variables or as functions of x, so that $f(x) = F(\xi)$, and hence $\frac{\partial F}{\partial \xi^j}(0) = DF(0).\mathbf{e}_j$. In the following, set

$$\boxed{\left(\frac{\partial}{\partial \xi^j}\right)_a f := \frac{\partial F}{\partial \xi^j}(0).}$$

This quantity only depends on the *germ* of the function f at a.

LEMMA 2.28.– *i) The mappings* $f \mapsto \left(\frac{\partial}{\partial \xi^j}\right)_a f$ *from* $C_a^r(M)$ *into* $\mathbb{K}$ *are linear forms on* $C_a^r(M)$ *that vanish on* $S_a^r(M)$.

ii) Any linear form $L \in C_a^r(M)^{\vee}$[4] *that vanishes on* $S_a^r(M)$ *can be uniquely written in the form*

$$L(f) = \sum_{1 \leq j \leq m} \alpha^j \left(\frac{\partial}{\partial \xi^j}\right)_a f \quad \left(\alpha^j \in \mathbb{K}\right). \qquad [2.2]$$

4 $C_a^r(X)$ is finite-dimensional, so its algebraic dual $C_a^r(X)^*$ coincides with its (topological) dual $C_a^r(X)^{\vee}$.

In other words, the $\left(\frac{\partial}{\partial \xi^j}\right)_a$ $(1 \leq j \leq m)$ *form a basis of the* $\mathbb{K}$*-vector space of linear forms on* $C_a^r(M)$ *that vanish on* $S_a^r(M)$*.*

PROOF.– (i) is clear. (ii): (a) We will show that any linear form $L \in C_a^r(M)^\vee$ that vanishes on $S_a^r(M)$ can be written in this form. To do this, consider the function g defined in the neighborhood of a by

$$g(x) = f(x) - f(a) - \sum\nolimits_{1 \leq j \leq m} \left(p^j \circ \xi\right)(x) \left(\frac{\partial}{\partial \xi^j}\right)_a f.$$

With $z = \xi(x)$, it follows that

$$\begin{aligned} \left(g \circ \xi^{-1}\right)(z) &= \left(f \circ \xi^{-1}\right)(z) - f(a) - \sum\nolimits_{1 \leq j \leq m} z^j \left(\frac{\partial}{\partial z^j}\right)_{z=0} f \\ &\Longrightarrow \frac{\partial\left(g \circ \xi^{-1}\right)}{\partial z^j}(0) = \frac{\partial\left(f \circ \xi^{-1}\right)}{\partial z^j}(0) - \left(\frac{\partial}{\partial z^j}\right)_{z=0} f = 0. \end{aligned}$$

Hence, g is stationary at a (and so is the germ of g). The same is true for the function $x \mapsto f(a)$ defined in an open neighborhood of a and the germ of this function. Therefore, L vanishes at g and at $x \mapsto f(a)$, which implies [2.2] with $\alpha^j = L\left(\xi^j\right)$.

(b) It remains to be shown that the linear forms $\left(\frac{\partial}{\partial \xi^j}\right)_a \in C_a^r(M)^\vee$ are linearly independent. Suppose that there exist elements $\mu_j \in \mathbb{K}$ such that $\sum_{1 \leq j \leq m} \mu^j \left(\frac{\partial}{\partial \xi^j}\right)_a = 0$. By applying this relation to the local coordinate functions ξ^j, it follows that $\mu^j = 0$ $(1 \leq j \leq m)$. ■

LEMMA-DEFINITION 2.29.– *Let* M *be a manifold,* $a \in M$ *and* $X_a = \mathring{\gamma}$ *be a tangent vector of* M *at* a *(in the sense of Definition 2.21).*

i) The Lie derivative *along the tangent vector* X_a *(at the point* a*) is defined by*

$$\mathcal{L}_{X_a} : f \mapsto \frac{d}{dt}(f \circ \gamma)(0)$$

for any function $f : M \to \mathbb{K}$ *of class* C^r *in the neighborhood of* a *and is a continuous form on* $C_a^r(M)$ *that vanishes on* $S_a^r(M)$*. The mapping* $\mathcal{L} : X_a \mapsto \mathcal{L}_{X_a}$ *is a linear injection from* $T_a(M)$ *into the space of continuous linear forms on* $C_a^r(M)$ *that vanish on* $S_a^r(M)$*.*

ii) Suppose that M *is finite-dimensional in the neighborhood of* a*. Then, the mapping* $\mathcal{L} : X_a \mapsto \mathcal{L}_{X_a}$ *defined above is bijective.*

iii) With the same hypothesis as (ii), the $\mathbb{K}$-vector space of linear forms on $C_a^r(M)$ that vanish on $S_a^r(M)$ is isomorphic to $(C_a^r(M)/S_a^r(M))^\vee$.

PROOF.– i) The definition of $\mathcal{L}_{X_a}$ only uses the values taken by f in the neighborhood of a and so depends solely on the germ of this function at a. Let $(U, \xi, \mathbf{E})$ be a chart of M centered on a and set $\psi(t) = \xi(\gamma(t))$. Then, ψ is a curve in $\mathbf{E}$ that passes through 0. If the germ of f at a is stationary,

$$\mathcal{L}_{X_a}(f) = \frac{d}{dt}\left\{\left(f \circ \xi^{-1}\right)(\xi(\gamma(t)))\right\}_{t=0} = \left(f \circ \xi^{-1}\right)'(0)\,\psi'(0) = 0. \quad [2.3]$$

We have $\mathcal{L}_{X_a}(f) \in \mathbb{K}$ and the first two equalities above show that $\mathcal{L}$ is linear and injective. Indeed, if $\left(f \circ \xi^{-1}\right)'(0)\,\psi'(0) = 0$ for every germ $f \in C_a^r(M)$, then $\psi'(0) = 0$ by the Hahn–Banach theorem ([P2], section 3.3.4(**II**), Corollary 3.26), and hence $X_a = 0$.

ii) Let L be a linear form on $C_a^r(M)$ that vanishes on $S_a^r(M)$ and let (U, ξ, m) be a chart of M centered on a. Then, by Lemma 2.28, L can be uniquely expressed in terms of the $\left(\frac{\partial}{\partial \xi^j}\right)_a$ in the form $L(f) = \sum_{1 \le j \le m} \alpha^j \left(\frac{\partial}{\partial \xi^j}\right)_a f$. Consider the curve γ in M and passing through a defined by $\gamma(t) = \xi^{-1}(t(\alpha_1, ..., \alpha_m))$. Then, $f(\gamma(t)) = \left(f \circ \xi^{-1}\right)(t(\alpha_1, ..., \alpha_m))$, so

$$\mathcal{L}_{X_a}(f) = \frac{d}{dt}(f \circ \gamma)_{t=0} = \left(f \circ \xi^{-1}\right)'(0)(\alpha_1, ..., \alpha_m) = L(f),$$

which shows that $\mathcal{L}$ is surjective.

iii) is clear. ■

This lemma leads to the following result:

COROLLARY-DEFINITION 2.30.– *Suppose that M is* locally finite-dimensional.

1) The tangent vector space of M at a is

$$T_a(M) = (C_a^r(M)/S_a^r(M))^\vee.$$

2) A basis of this tangent space is given by the $\left(\partial/\partial \xi^j\right)_a$, with

$$\left(\partial/\partial \xi^j\right)_a : C_a^r(M) \ni f \mapsto \left(\frac{\partial}{\partial \xi^j}\right)_a f \in \mathbb{K}.$$

3) Consider the expression [2.2], where $L : f \mapsto L(f)$ is an arbitrary tangent vector $\in T_a(M)$. Choosing $f = \xi^i$ gives $L\left(\xi^i\right) = \alpha^j$. The α^j are said to be the components of the tangent vector *$L = \sum_{1 \le j \le m} \alpha^j \partial_j(a)$ relative to the local coordinates ξ^j.*

REMARK 2.31.– *1) Let* (U, ξ, m) *be a chart of* M *centered on* a, $\mathbf{h} \in \mathbb{K}^m$, *and consider in* M *the curve* $\gamma : V \ni t \to \xi^{-1}(t\mathbf{h})$ *of class* C^r *(where* V *is an open neighborhood of* 0 *in* $\mathbb{R}$ *such that* $\{t.\mathbf{h} : t \in V\}$ *is contained in* $\xi(U)$*). Let* $f \in C_a^r(M)$ *and define the mapping*

$$\gamma_{*0} : f \mapsto \frac{d}{dt}(f \circ \gamma)(0).$$

Setting $\varepsilon_{\mathbf{h}} : V \ni t \mapsto t.\mathbf{h} \in \mathbb{K}^m$ *gives* $f \circ \gamma = f \circ \xi^{-1} \circ \varepsilon_{\mathbf{h}}$.

Hence,

$$(\gamma_{*0})(f) = D\left(f \circ \xi^{-1}\right)(0).\frac{d\varepsilon_{\mathbf{h}}}{dt}(0) = D\left(f \circ \xi^{-1}\right)(0).\mathbf{h}.$$

In particular, $(\gamma_{*0})(f) = 0$ *if* $f \in S_a^r(M)$. *By taking quotients,* γ_{*0} *induces an element of* $\left(C_a^r(M)/S_a^r(M)\right)^\vee = T_a(M)$, *which shows the equivalence between Definitions 2.30(1) and 2.23 in the local finite-dimensional case.*

2) Let $(\mathbf{e}_i)_{1 \le i \le m}$ *be the canonical basis of* $\mathbb{K}^m$ *and choose* $\mathbf{h} = \mathbf{e}_i$ *in the above. Let* $\gamma_i : t \to \xi^{-1}(t\mathbf{e}_i)$. *Then,* $(\gamma_{i*0})(f) = D\left(f \circ \xi^{-1}\right)(0).\mathbf{e}_i = \partial_i\left(f \circ \xi^{-1}\right)(0)$. *Hence,* γ_{i*0} *is the tangent vector* L *defined by*

$$f \mapsto \sum\nolimits_{1 \le j \le m} \alpha^j \left(\frac{\partial}{\partial \xi^j}\right)_a f = \sum\nolimits_{1 \le j \le m} \alpha^j \left(\frac{\partial}{\partial \xi^j}\right)_a f,$$

where $\alpha^j = \delta_i^j$. *Finally, with the notation of Corollary-Definition 2.30(2),*

$$\boxed{\gamma_{i*0} = \left(\frac{\partial}{\partial \xi^i}\right)_a.}$$

(IV) OPERATIONAL TANGENT VECTORS Let M be a manifold, a some point of M, and consider the Lie derivative $\mathcal{L}_{X_a}$ (Lemma-Definition 2.29(ii)). The mapping $X_a \mapsto \mathcal{L}_{X_a}$ is injective, so we can identify the tangent vector X_a with the linear form $\mathcal{L}_{X_a}$. The latter satisfies the Leibniz rule $\mathcal{L}_{X_a}(f.g) = \mathcal{L}_{X_a}(f).g + f.\mathcal{L}_{X_a}(g)$ (where f, g are mappings of class C^r from M into $\mathbb{K}$) and is therefore a derivation ([P1], section 2.3.12). With this identification, the Leibniz rule can also be written as $X_a(f.g) = (X_a f)\,g + f.X_a g$. The derivation $X_a : f \mapsto X_a.f$ is sometimes called an "operational tangent vector"[5].

5 As opposed to the tangent vectors introduced by Definitions 2.21 and 2.24, which are called "kinematic tangent vectors".

EXAMPLE 2.32.– *Let* M *be the helix in* $\mathbb{R}^3$ *defined by* $x = \cos\theta, y = \sin\theta, z = \theta$. *Any curve of class* C^r *on this helix is of the form* $\gamma(t) = \cos(\theta(t))\,\mathbf{e}_1 + \sin\theta(t)\,\mathbf{e}_2 + \theta(t)\,\mathbf{e}_3$*, where* $\theta : t \mapsto \theta(t)$ *is a mapping of class* C^r*. Setting* $\theta_0 = \theta(0)$ *and* $\dot{\theta}(0) = 1$ *gives* $\dot{\gamma}(0) = -\sin(\theta_0)\,\mathbf{e}_1 + \cos(\theta_0)\,\mathbf{e}_2 + \mathbf{e}_3$*. Let* $f : M \to \mathbb{R}$ *be a mapping of class* C^r *and let* P *be the point* $(x_0, y_0, z_0) = (\cos(\theta_0), \sin(\theta_0), \theta_0)$ *of* M*. Then, by [1.5], [1.7]:*

$$\frac{d}{dt}(f \circ \gamma)(0) = \left(-\sin(\theta_0)\frac{\partial}{\partial x}|_P + \cos(\theta_0)\frac{\partial}{\partial y}|_P + \frac{\partial}{\partial z}|_P\right) f,$$

so the tangent vectors of M *at the point* P *are collinear with* $-\sin(\theta_0)\frac{\partial}{\partial x}|_P + \cos(\theta_0)\frac{\partial}{\partial y}|_P + \frac{\partial}{\partial z}|_P = -y_0\frac{\partial}{\partial x}|_P + x_0\frac{\partial}{\partial y}|_P + \frac{\partial}{\partial z}|_P$.

2.3. Tangent linear mappings; submanifolds

2.3.1. *Tangent linear mapping; rank*

Let M, N be two manifolds, a some point of M, and $f : M \to N$ a morphism (section 2.2.2(**III**)). Let $c = (U, \xi, \mathbf{E})$ be a chart of M centered on a, $c' = (V, \eta, \mathbf{F})$ a chart of N centered on $f(a)$ and $\Phi : \xi(U) \to \eta(V)$ the expression of f in the charts c, c' (Definition 2.19). Let $\Phi'(0) = D\Phi(0) \in \mathcal{L}(\mathbf{E}; \mathbf{F})$ be the differential of Φ at 0 and, with the notation of Lemma-Definition 2.22,

$$\boxed{T_a(f) = \theta_{c'}^{-1} \circ \Phi'(0) \circ \theta_c.}$$

It is easy to check that $T_a(f)$ does not depend on the choice of charts. Furthermore, it follows from the definitions of $T_a(M)$ and $T_{f(a)}(N)$ (Definition 2.23) that $T_a(f) \in \mathcal{L}\left(T_a(M); T_{f(a)}(N)\right)$.

DEFINITION 2.33.– $T_a(f)$ *(also written as* f_{*a} *or* $f_*(a)$*) is the* tangent linear mapping *of* f *at the point* a*.*

LEMMA 2.34.– *i) Let* $X_a \in T_a(M)$*. With the notation of Remark 2.31, there exists a curve* $\gamma : \mathbb{K} \supset V \to M$ *of class* C^r *such that* $\gamma(0) = a$ *and* $X_a = \gamma_{*0}$*. Then,* $f \circ \gamma$ *is a curve of class* C^r *in* N *and* $(f \circ \gamma)(0) = f(a)$*. Let* $f_{*a}.X_a := (f \circ \gamma)_*(0) \in T_{f(a)}(N)$*. Then, the mapping* $f_{*a} : T_a(M) \to T_{f(a)}(N) : X_a \mapsto f_{*a}.X_a$ *is identical to* $T_a(f)$*.*

ii) For all $X_a = \gamma_{*0} \in T_a(M)$*,* $f_{*a}.X_a \in T_{f(a)}(N)$ *acts on* $C^r_{f(a)}(N)$ *as follows: for every function* $\psi \in C^r_{f(a)}(N)$*,*

$$(f_{*a}.X_a)\,\psi = (\gamma_{*0})(\psi \circ f). \qquad [2.4]$$

PROOF.– The mapping $f \circ \gamma$ is a curve λ of class C^r on N such that $\lambda(0) = f(a)$ and $(\lambda_{*0})\psi = \frac{d}{dt}(\psi \circ \lambda)(0) = \frac{d}{dt}(\psi \circ f \circ \gamma)(0)$. We have $\lambda_{*0} = (f \circ \gamma)_{*0} = f_{*a} X_a$. Furthermore, by definition (Remark 2.31),

$$(f \circ \gamma)_{*0} : C^r_{f(a)}(M) \ni \psi \mapsto \frac{d}{dt}(\psi \circ f \circ \gamma)(0),$$

which gives the desired result. ■

DEFINITION 2.35.– *The rank of f at the point a (written as rk$_a(f)$) is defined as the quantity rk$(T_a(f))$ (see Definition 1.7).*

Charts and Theorem 1.35(iii) give us the following result:

THEOREM 2.36.– *The mapping $x \mapsto$ rk$_x(f)$ from M into the discrete set $\bar{\mathbb{N}}$ is lower semi-continuous in M.*

Furthermore, Theorem 1.9 implies:

THEOREM 2.37.– *Let M, N, Y be three manifolds, a some point of M and $f : M \to N$, $g : N \to Y$ two mappings of class C^r. Then, $g \circ f$ is of class C^r and*

$$\boxed{T_a(g \circ f) = T_{f(a)}(g) \circ T_a(f).} \qquad [2.5]$$

Adapting Definition 2.33 and Lemma 2.34 to the case of $(\mathcal{KM})$ manifolds is left to the reader. When C^r is replaced by $\mathfrak{c}^r$, Theorem 2.36 is inapplicable, but the statement of Theorem 2.37 remains valid.

2.3.2. *Differential*

Let $\mathbf{F}$ be a Hausdorff locally convex space, M a manifold and $\mathbf{f} : M \to \mathbf{F}$ a mapping. We say that $\mathbf{f}$ is of class C^p $(p \in \mathbb{N}_{\mathbb{K}},\ p \leq r)$ if, for every chart $c = (U, \xi, \mathbf{E})$ of M, $\mathbf{f}|_U \circ \xi^{-1} : \xi(U) \to \mathbf{F}$ is of class C^p. If $a \in U$, write $d_a\mathbf{f}$ or $d\mathbf{f}_a$ for the continuous linear mapping $D\left(\mathbf{f}|_U \circ \xi^{-1}\right)(a) \in \mathcal{L}(T_a(M); \mathbf{F})$.

DEFINITION 2.38.– *$d_a\mathbf{f}$ is the* differential *of $\mathbf{f}$ at a.*

EXAMPLE 2.39.– *Let M be a manifold and $c = (U, \xi, \mathbf{E})$ a chart of M centered on a. Writing $\iota : U \hookrightarrow M$ for the canonical injection, the expression of ξ in the charts c, c', where $c' = (U, \iota, \mathbf{E})$, is given by $\xi \circ \xi^{-1} = 1_U$. Thus, $d_a\xi = 1_U \circ \theta_c = \theta_c$. The relation*

$$\boxed{\theta_c = d_a\xi} \qquad [2.6]$$

expresses the bijection $\theta_c : T_a(M) \to \mathbf{E}$ explicitly in terms of ξ.

Suppose that $\mathbf{E} = \mathbb{K}^m$ *and let* $(\mathbf{e}_j)_{1\leq i\leq n}$ *be its canonical basis. The* $\theta_c^{-1}(\mathbf{e}_j) = (d_a\xi)^{-1}.\mathbf{e}_j$ $(1 \leq j \leq m)$ *form a basis of* $T_a(M)$. *Then (section 2.2.4(**III**)),* $d_a\xi.\left(\frac{\partial}{\partial\xi^j}\right)_a = \theta_c\left(\frac{\partial}{\partial\xi^j}\right)_a = \mathbf{e}_j$. *Hence,*

$$\boxed{\left(\frac{\partial}{\partial\xi^j}\right)_a = (d_a\xi)^{-1}.\mathbf{e}_j.}$$

REMARK 2.40.– *1) The differential can therefore be viewed as a special case of a tangent linear mapping when* $\mathbf{F}$ *is a Banach space by identifying the tangent space* $T_{f(a)}(\mathbf{F})$ *with* $\mathbf{F}$. *With this convention, let* M, N *be two manifolds,* $a \in M$, $f: M \to N$ *a morphism, and* $c = (U, \xi, \mathbf{E})$ *and* $c' = (V, \eta, \mathbf{F})$ *charts of* M *and* N *centered on* a *and* $b = f(a)$, *respectively. Furthermore, let* $\Phi : \xi(U) \to \eta(V)$ *be the expression of* f *in these charts, i.e.* $\Phi = \eta \circ (f|_U) \circ \xi^{-1}$. *Then,* $f|_U = \eta^{-1} \circ \Phi \circ \xi$. *The equality* $T_a(f) = \theta_{c'}^{-1} \circ \Phi'(0) \circ \theta_c$ *can be rewritten as* $T_a(f) = (d_b\eta)^{-1} \circ \Phi'(0) \circ d_a\xi$. *Let* $\psi = \eta^{-1}$, *which gives* $f|_U = \psi \circ \Phi \circ \xi$. *Writing* $(d_b\eta)^{-1} = (\eta_{*b})^{-1} = (\eta^{-1})_{*\eta(b)}$ *gives* $f_{*a} = \psi_{*\eta(b)} \circ \Phi_{*0} \circ \xi_{*a}$, *resulting in the following expression, similar to [2.5]:*

$$f_{*a} = (\psi \circ \Phi \circ \xi)_{*a} = \psi_{*\psi^{-1}(b)} \circ \Phi_{*0} \circ \xi_{*a}.$$

2) The set $\mathcal{C}^p(M; \mathbf{F})$ *of mappings of class* C^p $(p \in \mathbb{N}_{\mathbb{K}}^{\times}, p \leq r)$ *from* M *into* $\mathbf{F}$ *is a* $\mathbb{K}$*-vector space and the operator* $d_a : \mathcal{C}^p(M; \mathbf{F}) \to \mathcal{L}(T_a(M); \mathbf{F}) : \mathbf{f} \mapsto d_a\mathbf{f}$ *is* $\mathbb{K}$*-linear.*

3) Let $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces,* U *an open subset of* $\mathbf{E}$, a *some point of* U *and* $\mathbf{f} : U \to \mathbf{F}$ *a mapping of class* C^p $(p \in \mathbb{N}_{\mathbb{K}}^{\times}, p \leq r)$. *After identifying* $T_a(U)$ *with* $\mathbf{E}$ *and* $T_{f(a)}(\mathbf{F})$ *with* $\mathbf{F}$, *the* tangent linear mapping $T_a(\mathbf{f})$ *and the* differential $d_a(\mathbf{f})$ *are identical.*

4) To simplify the terminology, the tangent linear mapping $T_a(f)$ *of a mapping* $f : M \to N$ *can simply be called its* differential *at the point* a, *written as* d_af *or* $df(a)$.

5) If M *is a* discrete *manifold, then every mapping from* M *into* $\mathbf{F}$ *is of class* C^ω *and* $d_a\mathbf{f} = 0$ *for all* $a \in M$. *Indeed,* $T_a(M) = \{0\}$ *and* $\mathcal{L}(\{0\}; \mathbf{F}) = \{0\}$.

Adapting the above to the case of a $(\mathcal{KM})$ manifold is entirely straightforward.

2.3.3. *Submanifolds*

Let M be a manifold and N a non-empty subset of M. Suppose that, at every point $y \in N$, there exists a chart $(V, \psi, \mathbf{E}_1 \times \mathbf{E}_2)$ of M centered on y such that $\psi(V) = W_1 \times W_2$, where W_i is an open set of the Banach space $\mathbf{E}_i$ $(i = 1, 2)$ and

$$\psi(N \cap V) = W_1 \times \{0\}.$$

Then:

– N is *locally closed* in M, i.e. every $y \in N$ has an open neighborhood V in M such that $V \cap N$ is closed in V. Furthermore, ψ induces a continuous bijection $\psi_1 : N \cap V \to W_1$.

– The thus obtained collection of triples $(N \cap V, \psi_1, \mathbf{E}_1)$ forms an atlas of N of class C^r.

DEFINITION 2.41.– *When equipped with this atlas, the set N is called a* submanifold *of M.*

In particular, every non-empty open subset of M is a submanifold of M. Any curve in M is a one-dimensional submanifold of M.

COROLLARY-DEFINITION 2.42.– *Let M be a manifold, N a submanifold of M and $\iota : N \to M$ the canonical injection. With the above notation, for all $y \in N$, we have:*

$$T_y(M) = \mathbf{E}_1 \times \mathbf{E}_2, \quad T_y(N) = \mathbf{E}_1,$$

so the mapping $T_y(\iota) : T_y(N) \to T_y(M)$ is injective and $T_y(M) / T_y(N) \cong \mathbf{E}_2$ is a Banach space, said to be transversal *to N (in M) at y. The dimension n_2 of this space (finite or infinite) is said to be the* codimension *of N in M at the point y.*

Corollary-Definition 2.42 further implies that:

$$0 \longrightarrow T_y(N) \xrightarrow{T_y(\iota)} T_y(M) \xrightarrow{\varphi} T_y(M) / T_y(N) \longrightarrow 0$$

is a *short exact sequence* that *splits* in **Vec** ([P1], section 3.1.4**(I)**, Lemma-Definition 3.15), where φ denotes the canonical surjection.

Consider the special case where M and N are finite-dimensional. Each point y of N has a neighborhood V in M with local coordinates $\left(\xi^1, ..., \xi^m\right)$ such that the $m - n$ last coordinates of the points of N in V are zero. These local coordinates are therefore of the form:

$$\left(\xi^1, ..., \xi^n, 0, ..., 0\right)$$

and the codimension of N in M is $m - n = \dim_y(M) - \dim_y(N)$.

2.3.4. *Immersions and embeddings*

LEMMA-DEFINITION 2.43.– *Let M, N be two manifolds, $f : M \to N$ a morphism, $a \in M$ and $b = f(a)$.*

1) The following properties are equivalent:

i) There exists an open neighborhood U of a in M such that $f|_U$ induces an isomorphism (of manifolds) from U onto a submanifold of N.

ii) The linear mapping $T_a(f) : T_a(M) \to T_b(N)$ is injective and its image splits *(see section 1.2.6).*

2) The morphism f is called an immersion *at a if it satisfies any of the equivalent conditions stated above. The set of points of M at which f is an immersion is an open subset of M; if this open set is equal to the whole of M, we say that f is an* immersion, *and $f(M)$ is called an* immersed submanifold *of N.*

PROOF.– This is a consequence (and a reformulation) of Corollary-Definition 1.32. ■

REMARK 2.44.– *Let M be a topological space, N a manifold and f a mapping from M into N. It is possible to show the following result: for a manifold structure that makes f an immersion to exist on M (section 2.2.2(**I**)), it is necessary and sufficient for the following condition **(Inv)** to be satisfied:*

***(Inv)** For all $a \in M$, there exist an open neighborhood U of a and a chart $(V, \varphi, \mathbf{E})$ of N at $f(a)$ such that $f(U) \subset V$ and $\varphi \circ f$ induces a homeomorphism from U onto $\varphi(V) \cap \mathbf{F}$, where $\mathbf{F}$ is a subspace of $\mathbf{E}$ that splits ([P2], section 3.2.2(**IV**)).*

If it exists, this manifold structure of M is unique and is called the preimage *of the manifold structure of N under f.*

COROLLARY 2.45.– *Suppose that M and N are finite-dimensional. Then, a morphism $f : M \to N$ is an immersion if and only if one of the following equivalent properties is satisfied:*

iii) There exist an open neighborhood U of a in M, a coordinate system $(\eta^1, ..., \eta^n)$ of N in an open neighborhood V of b containing $f(U)$, and an integer $r \leq n$ such that $\eta^j \circ f = 0$ $(r < j \leq n)$ and the functions $\eta^1 \circ f, ..., \eta^r \circ f$ form a coordinate system of M in U.

iv) $rk_a(f) = \dim_a(M)$.

PROOF.– (iii) is a special case of (i) from Lemma-Definition 2.43. The equivalence of (ii) and (iv) is clear. ■

DEFINITION 2.46.– *An isomorphism $f : M \to N$ from a manifold M onto a submanifold N of a manifold Z is said to be an* embedding *of M in Z.*

REMARK 2.47.– *1) An immersion $f : M \to N$ is an embedding if and only if $f(M)$, equipped with the topology induced by the topology of N, is homeomorphic to M; if so, $f(M)$ is a submanifold of M, and f induces a diffeomorphism from M onto $f(M)$.*

2) An immersion is a "local embedding"; however, from a global *perspective, there are injective immersions that are not isomorphisms onto a submanifold. An injective immersion $f : M \to N$ is an embedding if it is proper (i.e. the preimage of any compact set of N under f is compact in M) (**exercise***: see [LEB 82], Chapter I, section 4, Proposition 2). The manifold shown in Figure 2.4 is not embedded in $\mathbb{R}^2$. However, the Earth and the Klein bottle*[6] *can be embedded in $\mathbb{R}^3$. The* Whitney embedding theorem *states that any real paracompact p-dimensional manifold can be embedded in $\mathbb{R}^{2p}$ [WHI 44].*

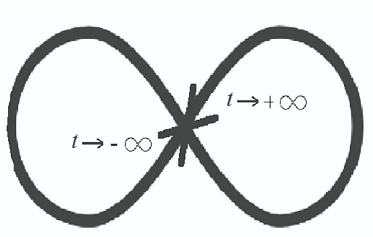


Figure 2.4. *An immersion that is not an embedding*

THEOREM 2.48.– *Suppose that N is locally compact and let $f : M \to N$ be a morphism. The graph Gr (f) of f ([P2], section 2.3.3(**II**)) is a closed submanifold of $M \times N$, the mapping $g : x \mapsto (x, f(x))$ from M into $M \times N$ is an embedding, and $T_{(x,f(x))}(Gr(f))$ is the Banach subspace of $T_{(x,f(x))}(M \times N)$ given by the graph of $T_x(f)$.*

PROOF.– The mapping g is a continuous bijection from M onto Gr (f) whose inverse bijection is the restriction of $pr_1 : M \times N \twoheadrightarrow N$ to Gr (f). Hence, g is a homeomorphism. Since N is Hausdorff, Gr (f) is closed in $M \times N$ ([P2], section 2.3.3(**II**)). Thus, it simply remains to be shown that g is an immersion (Remark 2.47(1)). But $T_x(g) = (T_x(1_M), T_x(f))$, so $T_x(g)$ is injective. Since im $(T_x(g))$ has finite codimension in $T_{(x,f(x))}(M \times N)$, im $(T_x(g))$ splits in this space ([P2], section 3.2.2(**IV**)). ■

It is possible to show the following result ([EEL 70], Theorem 1):

THEOREM 2.49.– *Let M be a real Hausdorff paracompact manifold of class C^∞ modeled on a separable Hilbert space* **E**. *There exists an embedding of class C^∞ from M onto an open set of* **E**.

6 See the Wikipedia article on the *Klein bottle*.

2.3.5. *Submersions, subimmersions and* étale *mappings*

LEMMA-DEFINITION 2.50.– *Let M, N be two manifolds, $f : M \to N$ a morphism, $a \in M$ and $b = f(a)$.*

1) The following properties are equivalent:

i) There exist an open neighborhood U of a, an open neighborhood V of b containing $f(U)$ and a morphism g from U into a manifold Z such that the mapping $(f, g) : U \to V \times Z$ is an isomorphism.

ii) The linear mapping $T_a(f) : T_a(M) \to T_b(N)$ is surjective and its kernel splits.

2) The morphism f is called a submersion at *a if it satisfies one of the equivalent conditions stated above. The set of points of M where f is a submersion is an open subset of M; when this open set is equal to the whole of M, we say that f is a* submersion.

PROOF.– This is a consequence (and a reformulation) of Corollary-Definition 1.33. ■

COROLLARY 2.51.– *Suppose that M and N are locally finite-dimensional. A morphism $f : M \to N$ is a submersion if and only if:*

iii) There exist an open neighborhood U of a, an open neighborhood V of b containing $f(U)$ and a coordinate system $(\eta^1, ..., \eta^n)$ of N in V, such that the functions $\eta^i \circ f$ form a coordinate system of M in U.

PROOF.– The equivalence of conditions (ii) and (iii) is clear. ■

Note that submersions are open mappings.

DEFINITION 2.52.– *Let N be a locally finite-dimensional manifold and A some subset of N. We say that A is locally negligible in N if, for any chart (U, φ, n) of N, $\varphi(A \cap U)$ has zero Lebesgue measure in $\mathbb{R}^n$ ([P2], sections 4.1.1**(III)**, 4.1.3**(II)**).*

We will assume (2) below without proof. It was shown by A. Sard in 1943 (see [LAN 99b], Chapter 14, Theorem 1.4, and [BOU 82a], 10.1.3(d)).

THEOREM-DEFINITION 2.53.– *1) Let M and N be two locally finite-dimensional differential manifolds, $f : M \to N$ a morphism and $a \in M$. We say that a is a* regular point *of f if f is a submersion, i.e. if $T_a(f) : T_a(M) \to T_{f(x)}(N)$ is surjective. Otherwise, we say that a is a* singular *or* critical point *of f.*

2) (Sard's theorem) *Let Z be the set of singular points of f. If M is Hausdorff (and hence locally compact) and countable at infinity, then $f(Z)$ is locally negligible in N.*

LEMMA-DEFINITION 2.54.– *Let M, N be two manifolds, a some point of M, $f : M \to N$ a morphism and $b = f(a)$.*

1) The following conditions are equivalent:

i) There exist an open neighborhood U of a, a manifold Z, a submersion S from U onto Z and an immersion i from Z into N such that $f|_U = i \circ s$:

$$M \supset U \xrightarrow{s} Z \xrightarrow{i} N.$$

ii) There exist a chart $(U, \xi, \mathbf{E})$ of M centered on a, a chart $(V, \eta, \mathbf{F})$ of N centered on $f(a)$ and a mapping $\mathbf{g} \in \mathcal{L}(\mathbf{E}; \mathbf{F})$ such that

$$f(U) \subset V, \quad \mathbf{g}(\xi(U)) \subset \eta(V), \quad f|_U = \eta^{-1} \circ \mathbf{g} \circ \xi$$

and the kernel and image of $\mathbf{g}$ *split.*

2) We say that f is a subimmersion *at a if one of the equivalent conditions stated above is satisfied. The set of points of M where f is a subimmersion is an open subset Ω of M (and Ω is dense in M if M is finite-dimensional); when this set is equal to the whole of M, we say that f is a* subimmersion.

3) If $f : M \to N$ is a subimmersion, then, for every $b \in N$, $f^{-1}(\{b\})$ is a closed submanifold Z of M, and $T_a(Z) = \ker(T_a(f))$. Writing $\iota : T_a(Z) \hookrightarrow T_a(M)$ for the canonical injection, the following exact sequence splits in **Vec***:*

$$0 \longrightarrow T_a(Z) \xrightarrow{\iota} T_a(M) \xrightarrow{T_a(f)} T_b(N).$$

4) If $f : M \to N$ is a submersion and Y is a submanifold of N, then $f^{-1}(Y)$ is a submanifold of M and the mapping $f^{-1}(Y) \to Y$ induced by f is a submersion.

PROOF.– (1) is a reformulation of Corollary-Definition 1.34. (2): Using the charts $(U, \xi, \mathbf{E}), (V, \eta, \mathbf{F})$, we can reduce to the case where $f = \mathbf{g}$. (3): **exercise**. ■

The *rank theorem* (Theorem 1.35) can be expressed in its most natural form in the context of manifolds (the full proof of (3) below is an **exercise***: see, for example, [SCH 93], Volume 2, Theorem 3.8.22):

THEOREM 2.55.– *1) Let M, N be two manifolds, a some point of M and $f : M \to N$ a morphism.*

1) If f is a subimmersion at a, then $rk_a(f)$ is constant in a neighborhood of a.

2) Conversely, if rk$_a$ $(f) < +\infty$ and rk$_x$ (f) is constant in a neighborhood of a, then f is a subimmersion at a.

3) Suppose that M and N are locally finite-dimensional. Then, $f : M \to N$ is a subimmersion at a if and only if there exist a coordinate system $\left(\xi^1, ..., \xi^m\right)$ of M at a, a coordinate system $\left(\eta^1, ..., \eta^n\right)$ of N at $f(a)$ and an integer $s \leq \min(m, n)$ such that

$$\eta^j \circ f = \xi^j, \quad 1 \leq j \leq s$$
$$\eta^j \circ f = 0, \quad s + 1 < j \leq m,$$

i.e. $\Phi = \eta \circ (f|_U) \circ \xi^{-1}$ (the local expression of f in the relevant charts) is the restriction to $\xi(U)$ of the mapping

$$\left(\xi^1, ..., \xi^m\right) \mapsto \left(\xi^1, ..., \xi^s, 0, ..., 0\right).$$

Furthermore, f is an immersion (respectively a submersion) if and only if $s = m$ (respectively $s = n$).

REMARK 2.56.– *The composition of two immersions is an immersion, the composition of two submersions is a submersion. However, the composition of two subimmersions is not a subimmersion in general (**exercise**).*

Theorem 1.29 leads to the following result:

LEMMA-DEFINITION 2.57.– *Let M, N be manifolds, $f : M \to N$ a morphism and $a \in M$.*

1) The following conditions are equivalent:

i) $T_a(f)$ is bijecitve.

ii) There exist an open neighborhood U of a and an open neighborhood V of $f(a)$ such that f induces an isomorphism from U onto V.

iii) f is both an immersion and a submersion at a.

2) Whenever these conditions are satisfied, we say that f is étale *or is a* local diffeomorphism *at a (section 1.2.6(**II**)). If these conditions hold for every $a \in M$, we say that f is* étale *(or an* étale *mapping or a local diffeomorphism) and that the manifold M is* étale over *N (relative to f).*

2.3.6. *Submanifolds of* $\mathbb{K}^n$

(I) Parametric representation of a manifold Let $\mathbf{E}$ be an n-dimensional vector space over $\mathbb{K}$. The following lemma is a reformulation of Corollary-Definition 2.42:

LEMMA 2.58.– *A set $M \subset \mathbf{E}$ is a p-dimensional manifold if and only if, for every point $a \in M$, there exist a neighborhood V of a in $\mathbf{E}$ and a mapping $\mathbf{z} : U \to \mathbf{E}$, where U is an open subset of $\mathbb{K}^p$ $(p \leq n)$, and $\mathbf{z}$ is a homeomorphism from U onto $M \cap V$ such that $D\mathbf{z}(u)$ has rank p for every $u \in U$. If so:*

$$M \cap V = \{\mathbf{z}(u) : u \in U\}.$$

Given u_0 such that $a = \mathbf{z}(u_0)$, the affine *tangent space $T_a(M)$ of M at the point a is:*

$$\{x \in \mathbf{E} : \exists u \in \mathbb{K}^p, x - a = D\mathbf{z}(u_0).(u - u_0)\}.$$

For example, let M be the circle of center 0 and radius $R > 0$. Let $\mathbf{z} : \mathbb{R} \to \mathbb{R}^2 : u \mapsto (x_1, x_2)$ with $x_1 = R\cos(u)$, $x_2 = R\sin(u)$, and $a_1 = R\cos(u_0)$, $a_2 = R\sin(u_0)$. The tangent to M at the point $a = (a_1, a_2)$ is the set of points $x = (x_1, x_2)$ such that $x_1 - a_1 = -R\sin(u_0)(u - u_0)$ and $x_2 - a_2 = R\cos(u_0)(u - u_0), u \in \mathbb{R}$.

(II) Implicit representation of a manifold

THEOREM 2.59.– *A set $M \subset \mathbf{E}$ $(\mathbf{E} \cong \mathbb{K}^n)$ is a p-dimensional manifold $(p \leq n)$ if and only if, for each point $a \in M$, there exist an open neighborhood Ω of a in $\mathbf{E}$ and a submersion $f : \Omega \to \mathbb{K}^{n-p}$ such that*

$$M \cap \Omega = \{X \in \Omega : f(X) = 0\}.$$

The affine tangent space $T_a(M)$ is then given by

$$\{X \in \mathbf{E} : Df(a).(X - a) = 0\}. \qquad [2.7]$$

PROOF.– This follows from Lemma 2.58 and the implicit function theorem (Theorem 1.30). ■

EXAMPLE 2.60.– *(1) In three-dimensional Euclidean space, consider a surface, i.e. a real submanifold of dimension 2. This surface can be described in the neighborhood of the point $a = (x_0, y_0, z_0)$ by some equation $f(X) = 0$, $X = (x, y, z)$, where the function f is real, is of class C^r and has non-zero gradient (Definition 1.8) $\mathrm{grad}_a(f)$ (i.e. a is a* regular point *of f); $\mathrm{grad}_a(f)$ is the vector with coordinates $\frac{\partial f}{\partial x}(a)$, $\frac{\partial f}{\partial y}(a)$,*

$\frac{\partial f}{\partial z}(a)$. *The tangent plane to the surface at the point* a *is orthogonal to grad*$_a(f)$ *and is therefore generated by the vectors* $\mathbf{h}$ *such that* $\langle grad_a(f)|\mathbf{h}\rangle = 0$, *namely* $Df(a).\mathbf{h} = 0$. *Thus, the* affine *tangent plane has the equation* $Df(a).(X-a) = 0$.

For example, the sphere of center 0 *and radius* $R > 0$ *has the equation* $f(x,y,z) = 0$, *where* $f(x,y,z) = x^2 + y^2 + z^2 - R^2$ *and* $grad_a(f) = 2(x_0, y_0, z_0) = 2a$. *The tangent plane at the point* a *therefore has the equation*

$$x_0(x-x_0) + y_0(y-y_0) + z_0(z-z_0) = 0.$$

2) Under the same conditions, a curve in three-dimensional Euclidean space is the intersection of two surfaces. It is locally described by the equation $f(X) = 0$, $X = (x,y,z)$, *where* $f = (f_1, f_2)$ *is of* C^r, *and the tangent to this curve at the point* a *is given by the equation [2.7].*

2.3.7. *Products of manifolds*

(I) Let M_1, M_2 be two manifolds and let $c_i = (U_i, \varphi_i, \mathbf{E}_i)$ be a chart of M_i $(i = 1, 2)$. The triple $c_1 \times c_2 := (U_1 \times U_2, \varphi_1 \times \varphi_2, \mathbf{E}_1 \times \mathbf{E}_2)$ is a chart on the set $M_1 \times M_2$. On $M_1 \times M_2$, there exists a unique manifold structure such that, for any charts c_1, c_2 of M_1, M_2 respectively, $c_1 \times c_2$ is a chart of $M_1 \times M_2$. The manifold $M_1 \times M_2$ is the *product* of M_1, M_2 in the category of manifolds ([P1], Section 1.2.6(**I**)).

The canonical projection $pr_i : M_1 \times M_2 \twoheadrightarrow M_i$ is a morphism $(i = 1, 2)$ and, if π_i is its tangent linear mapping at the point (a_1, a_2), the mapping

$$(\pi_1, \pi_2) : T_{(a_1,a_2)}(M_1 \times M_2) \to T_{a_1}(M_1) \times T_{a_2}(M_2)$$

is an isomorphism that identifies these two spaces. The injection $T_{a_1}(M_1) \hookrightarrow T_{(a_1,a_2)}(M_1 \times M_2)$ resulting from this identification is the tangent linear mapping of the morphism $x_1 \to (x_1, a_2)$ at a_1. An analogous result holds for $T_{a_2}(M_2) \hookrightarrow T_{(a_1,a_2)}(M_1 \times M_2)$.

(II) Let M_1, M_2, N be three manifolds, $f : M_1 \times M_2 \to N$ a morphism and $(a_1, a_2) \in M_1 \times M_2$. The tangent linear mapping of the partial mapping $f(., a_2)$ (respectively $f(a_1, .)$) is denoted by $T^1_{(a_1,a_2)}(f)$ (respectively $T^2_{(a_1,a_2)}(f)$). For all $\mathbf{u}_i \in T_{a_i}(M_i)$ $(i = 1, 2)$, we have (see [1.6])

$$\boxed{T_{(a_1,a_2)}(f).(\mathbf{u}_1, \mathbf{u}_2) = T^1_{(a_1,a_2)}(f).\mathbf{u}_1 + T^2_{(a_1,a_2)}(f).\mathbf{u}_2.}$$

(III) IMPLICIT FUNCTION THEOREM The following result can be deduced immediately from Theorem 1.30:

THEOREM 2.61.– (Implicit function) *(1) Suppose that $T^2_{(a_1,a_2)}(f)$ is bijective. There exists an open neighborhood U_i of a_i in M_i $(i = 1, 2)$ satisfying the following property: for all $x_1 \in U_1$, there exists a unique point $g(x_1)$ of U_2 such that $f(x_1, g(x_1)) = f(a_1, a_2)$ and g is a morphism from U_1 into U_2. For all $x_1 \in U_1$,*

$$\boxed{T_{x_1}(g) = -\left(T^2_{(x_1,g(x_1))}(f)\right)^{-1} \circ T^1_{(x_1,g(x_1))}(f).}$$

2) In particular, let M, N be manifolds, $f : M \to N$ a morphism and x some point of M. The following conditions are equivalent:

i) The tangent linear mapping $T_x(f)$ is an isomorphism.

ii) f is a local diffeomorphism (Lemma-Definition 2.57(2)).

If M and N are finite-dimensional, these conditions are also equivalent to:

iii) rk$_x(f) = \dim_x(M) = \dim_{f(x)}(N)$.

2.3.8. *Transversal morphisms and manifolds*

The following generalizes the notion of transversal space at a point to manifolds (Corollary-Definition 2.42).

LEMMA-DEFINITION 2.62.– *Let M, N be manifolds, $f : M \to N$ a morphism and N_1 a submanifold of N.*

1) Let $x \in M$ such that $f(x) \in N_1$. The following conditions are equivalent:

i) Let $(V, \eta, \mathbf{F})$ be a chart of N centered on $f(x)$ such that $\eta : V \to W_1 \times W_2$ is an isomorphism onto a product, with

$$\eta(N_1 \cap V) = V_1 \times \{0\}.$$

Then, there exists an open neighborhood U of x in M such that the composition

$$U \xrightarrow{f|_U} V \xrightarrow{\eta} W_1 \times W_2 \xrightarrow{pr_2} W_2$$

is a submersion.

ii) The composition

$$T_x(M) \xrightarrow{T_x(f)} T_{f(x)}(N) \longrightarrow T_{f(x)}(N) / T_{f(x)}(N_1)$$

is surjective and its kernel splits.

2) If the equivalent conditions in (1) are satisfied, then the morphism f is said to be transversal over N_1 at x. *If they are satisfied for every $x \in M$, f is said to be* transversal over N_1.

3) Suppose that N is a submanifold of M and let $y \in N$. Then, the canonical injection $\iota : N \to M$ is transversal over $\{y\}$; in other words, the Banach space $T_y(M)/T_y(N)$ is transversal to N at y (Corollary-Definition 2.42).

4) If f is transversal over N_1, then $f^{-1}(N_1)$ is a submanifold of M. (In (3), this is trivial, since $\iota^{-1}(\{y\}) = \{y\}$.)

5) Let N_1 and N_2 be two submanifolds of N and let $y \in N_1 \cap N_2$. The following conditions are equivalent:

i') The canonical injection $\iota_1 : N_1 \hookrightarrow N$ is transversal over N_2 at y.

ii') $T_y(N) = T_y(N_1) + T_y(N_2)$ and $T_y(N_1) \cap T_y(N_2)$ splits in $T_y(N)$ (the latter part is automatically satisfied when N is a Hilbert manifold).

6) If the conditions of (5) are satisfied, we say that N_1 and N_2 are transversal *at the point $y \in N_1 \cap N_2$. This concept is* symmetric *in N_1, N_2. If N_1 and N_2 are* transversal *at every point of $N_1 \cap N_2$, we say that N_1 and N_2 are* transversal.

PROOF.– (1): This is a consequence of Corollary 1.33. (4): we have $(pr_2 \circ \eta \circ f|_U)^{-1}(\{0\}) = \eta^{-1}(N_1 \cap V)$. Since $pr_2 \circ \eta \circ f|_U$ is a submersion and hence a subimmersion, $\eta^{-1}(N_1 \cap V)$ is a submanifold of M by Lemma-Definition 2.54(3), and the same is true for $\eta^{-1}(N_1)$.

5) The surjectivity of the composition defined in (1)(ii) is equivalent to having

$$T_{f(x)}(N) = T_x(f)(T_x(M)) + T_{f(x)}(N_1). \qquad [2.8]$$

This condition can be taken as the definition of the transversality of f over N_1 at x when the manifolds M, N and N_1 are locally finite-dimensional ([DIE 93], Volume 3, section 16.8, Problem 9). In the situation considered in (5)(i'), consider the composition

$$T_y(N_1) \xrightarrow{T_y(\iota_1)} T_y(N) \xrightarrow{\varphi_2} T_y(N)/T_y(N_2),$$

where φ_2 is the canonical epimorphism. Then, [2.8] becomes $T_y(N) = T_y(N_1) + T_y(N_2)$, and moreover

$$\ker(\varphi_2 \circ T_y(\iota_1)) = \{\mathbf{h}_y \in T_y(N_1) : \varphi_2(\mathbf{h}_y) = 0\} = T_y(N_1) \cap T_y(N_2).$$

This completes the proof of (5). ■

EXAMPLE 2.63.– *In the plane, two curves* N_1, N_2 *are transversal at the point* $y \in N_1 \cap N_2$ *if and only if their tangents at* y *are not collinear. In everyday space, a sphere* N_1 *and a plane* N_2 *such that* $N_1 \cap N_2 \neq \varnothing$ *are transversal if and only if* N_2 *is not tangent to* N_1.

It is possible to show the following result (**exercise***; see [BOU 82a], 5.11.8):

THEOREM 2.64.– *Let* N_1, N_2 *be transversal submanifolds of* N. *Then,* $N_1 \cap N_2$ *is a submanifold of* N *and* $T_x(N_1 \cap N_2) = T_x(N_1) \cap T_x(N_2)$.

LEMMA 2.65.– *Let* $\mathbf{E}$ *be a Banach space. The diagonal* $\Delta_{\mathbf{E}}$ *in* $\mathbf{E} \times \mathbf{E}$ *(i.e. the set of* (x, x), $x \in \mathbf{E}$*) is a subspace that splits and hence a closed submanifold of* $\mathbf{E} \times \mathbf{E}$. *Similarly, if* M *is a manifold, the diagonal* Δ_M *in* $M \times M$ *is a submanifold of* $M \times M$.

PROOF.– We have $\Delta_{\mathbf{E}} = f^{-1}(\{0\})$, where $f : \mathbf{E} \times \mathbf{E} \to \mathbf{E}$ is defined by $f(x, y) = x - y$. Since f is continuous linear, Δ is a closed subspace of $\mathbf{E} \times \mathbf{E}$. The factor $\mathbf{E} \times \{0\}$ of $\mathbf{E} \times \mathbf{E}$ is a complementary subspace Δ, since

$$\{(x, y) : (x, y) \in \mathbf{E} \times \mathbf{E}\} = \{(x, 0) : x \in \mathbf{E}\} \oplus \{(y, y) : y \in \mathbf{E}\}.$$

However, $\mathbf{E} \times \{0\}$ splits in $\mathbf{E} \times \mathbf{E}$. It follows that $\Delta_{\mathbf{E}}$ splits in $\mathbf{E} \times \mathbf{E}$ and hence is a closed submanifold of $\mathbf{E} \times \mathbf{E}$; similarly, Δ_M is a submanifold of $M \times M$ by Definition 2.41. ■

DEFINITION 2.66.– *Let* M, N, Z *be manifolds and let* $f : M \to Z$, $g : N \to Z$ *be morphisms. We say that* f *and* g *are* transversal *if the morphism*

$$f \times g : M \times N \to Z \times Z$$

is transversal over the diagonal.

This implies that, for any pair $(x, y) \in M \times N$ such that $z =: f(x) = g(y)$,

$$T_z(Z) = T_x(f)(T_x(M)) + T_y(g)(T_y(N)).$$

This condition is also sufficient to guarantee that f, g are transversal whenever M, N, and Z are finite-dimensional.

REMARK 2.67.– *If* f *or* g *is a submersion, then the morphisms* f *and* g *are transversal.*

2.3.9. *Fiber product of manifolds*

(I) Consider manifolds M_1, M_2, Z and *transversal* morphisms $f_1 : M_1 \to Z$, $f_2 : M_2 \to Z$ (Definition 2.66). Then, $\{(x_1, x_2) \in M_1 \times M_2 : f_1(x_1) = f_2(x_2)\} =$

$(f_1 \times f_2)^{-1}(\Delta_Z)$. This set P is a submanifold of $M \times N$ by Lemma 2.65 and Lemma-Definition 2.62(4). Equipped with the natural morphisms into M_1 and M_2 (obtained from the projections), P is a fiber product ([P1], section 1.2.7) $M_1 \times_Z M_2$ of f_1, f_2 over Z in the category of manifolds. Write $f_1 \times_Z f_2$ for the morphism $h : P \to Z$ such that $f_1(x_1) = f_2(x_2)$ for all $(x_1, x_2) \in P$.

(II) GLUING OF LOCAL FIBER PRODUCTS Fiber products can be obtained by gluing together locally constructed fiber products: simply consider an open covering $(W_i)_{i\in I}$ of Z. If the fiber product $P_i = U_i^1 \times_{W_i} U_i^2$ exists for every $i \in I$, with $U_i^j = f_j^{-1}(W_i)$ $(i = 1, 2)$, then the union P of the P_i is the fiber product $M_1 \times_Z M_2$ and P_i is open in P ([LAN 99b], Chapter 2, Proposition 2.6).

2.3.10. *Covectors and cotangent spaces*

Let M be a manifold and a some point of M.

DEFINITION 2.68.– *i) A* cotangent vector *(or* covector*) at a is an element of the dual $T_a^\vee(M)$ of the tangent space $T_a(M)$.*

ii) The space $T_a^\vee(M)$ is called the cotangent space *of M at a.*

LEMMA 2.69.– *Let $f : M \to \mathbb{K}$ be a morphism. Then, the differential $d_a f$ is a cotangent vector at a.*

REMARK 2.70.– *If M is locally finite-dimensional, we know that $T_a(M) = (C_a^{\boldsymbol{r}}(M)/S_a^{\boldsymbol{r}}(M))^\vee$ (Corollary-Definition 2.30), so (with the canonical identifications)*

$$T_a^\vee(M) = C_a^{\boldsymbol{r}}(M)/S_a^{\boldsymbol{r}}(M).$$

THEOREM 2.71.– *Let M be a locally finite-dimensional manifold, $a \in M$, $c = (U, \xi, m)$ a chart of M centered on a, and $(\xi^j)_{1\leq j\leq m}$ the system of local coordinates associated with c. Then, a basis of the cotangent space $T_a^\vee(M)$ is given by $(d_a\xi^j)_{1\leq j\leq m}$, the dual basis of $\left(\left(\frac{\partial}{\partial\xi^j}\right)_a\right)_{1\leq 1\leq m}$, where the duality bracket is defined by the expression [2.10] below.*

PROOF.– We know that $\left(\left(\frac{\partial}{\partial\xi^j}\right)_a\right)_{1\leq 1\leq m}$ is a basis of the tangent space $T_a(M)$ (Lemma 2.28). Every element of $T_a(M)$ is a linear form $f \mapsto L\left(\mathring{f}\right)$, where $L\left(\mathring{f}\right)$ is of the form [2.2], and $\mathring{f} \in C_a^{\boldsymbol{r}}(M)/S_a^{\boldsymbol{r}}(M)$ denotes the canonical image of a germ $(\Omega, f) \in C_a^{\boldsymbol{r}}(M)$ at the point $a \in U$ in the quotient space $C_a^{\boldsymbol{r}}(M)/S_a^{\boldsymbol{r}}(M)$.

This canonical image $\mathring{f}$ can be identified with the differential $d_a f$. In particular, for every $j \in \{1, ..., m\}$, $d_a \xi^j$ belongs to $T_a^\vee(M)$. We have:

$$\boxed{d_a f = \sum_{1 \le j \le m} \left(\left(\frac{\partial}{\partial \xi^j} \right)_a f \right) .d_a \xi^j,} \qquad [2.9]$$

so $\left(d_a \xi^j\right)_{1 \le j \le m}$ is a basis of $T_a^\vee(M)$. Finally, in the dual space $T_a^\vee(M)$, every element $v^\vee$ can be written as $v^\vee = \sum_{1 \le j \le m} v_j \, \mathbf{h}_{aj}^\vee$, where $v_j = \langle v^\vee, \mathbf{h}_{aj} \rangle$, $(\mathbf{h}_{aj})_{1 \le j \le m}$ is a basis of $T_a(M)$, and $\left(\mathbf{h}_{aj}^\vee\right)_{1 \le j \le m}$ is the dual basis. Let $\langle -, - \rangle : T_a^\vee(M) \times T_a(M) \to \mathbb{K}$ be the duality bracket. The expression [2.9] leads us to write

$$\left\langle d_a f, \left(\tfrac{\partial}{\partial \xi^j} \right)_a \right\rangle = \left(\tfrac{\partial}{\partial \xi^j} \right)_a f.$$

More generally,

$$X_a = \sum_{1 \le j \le n} \alpha^j \left(\tfrac{\partial}{\partial \xi^j} \right)_a \in T_a(M) \Rightarrow \langle d_a f, X_a \rangle = \sum_{1 \le j \le n} \alpha^j \left(\tfrac{\partial}{\partial \xi^j} \right)_a f. \quad [2.10]$$

With $X_a = \left(\frac{\partial}{\partial \xi^j} \right)_a$ and $f = \xi^i$, it is straightforward to check that

$$\langle d_a f, X_a \rangle = \left\langle d_a \xi^i, \left(\tfrac{\partial}{\partial \xi^j} \right)_a \right\rangle = \left(\tfrac{\partial}{\partial \xi^j} \right)_a \xi^i = \delta^i_j,$$

or in other words,

$$\boxed{\left\langle d_a \xi^i, \left(\frac{\partial}{\partial \xi^j} \right)_a \right\rangle = \delta^i_j,}$$

which means that $\left(d_a \xi^j\right)_{1 \le j \le m}$ is the dual basis of $\left(\left(\frac{\partial}{\partial \xi^j} \right)_a \right)_{1 \le 1 \le m}$. ■

2.3.11. *Cotangent linear mapping*

Let M, N be two manifolds, $f : M \to N$ a morphism, and $a \in M$. We have $f_{*a} : T_a(M) \to T_b(N)$ with $b = f(a)$. The transpose ([P2], section 3.5.3) ${}^t f_a : T_b(N)^\vee \to T_a(M)^\vee$ of this continuous linear mapping is called the *cotangent linear mapping* at the point b.

2.4. Lie groups and their actions

2.4.1. *Lie groups*

(I) NOTION OF A LIE GROUP A topological group $\mathbf{G}$ is a group equipped with a topology that makes the mapping $\mathbf{G} \times \mathbf{G} \ni (x, y) \to x.y^{-1} \in \mathbf{G}$ continuous ([P2], section 2.8.1). If this group $\mathbf{G}$ is a finite-dimensional $\mathbb{K}$-manifold (respectively Banach manifold, respectively Fréchet manifold, respectively $(\mathcal{KM})$ manifold, etc.) and the mapping $(x, y) \to xy^{-1}$ is a morphism of manifolds of the same type, $\mathbf{G}$ is said to be a finite-dimensional *Lie* $\mathbb{K}$*-group* (or simply a *Lie group* if $\mathbb{K}$ is clear from context) of class C^r or $\mathfrak{c}^r$, depending on the case. We will see in **(V)** below that $\mathbf{G}$ is a pure manifold; hence, it is locally finite-dimensional if and only if it is finite-dimensional.

If $\mathbb{K}$ is the field of real numbers, the above still holds if we simply assume that $r \geq 1$; however, it can be shown that a Banach Lie group is analytic if and only if it is of class C^r with $r \geq 3$ [MAI 62]. Throughout the rest of this chapter, every (real or complex) Lie group is of class C^ω. The morphisms of the category $\mathbf{LieGrp}$ of Lie groups are therefore the morphisms of groups of class C^ω.

THEOREM 2.72.– *Any Banach Lie group* $\mathbf{G}$ *is a metrizable, complete and locally connected topological group.*

PROOF.– Since the neutral element e of $\mathbf{G}$ has an open neighborhood that is homeomorphic to an open ball of a Banach space, $\{e\}$ is closed in $\mathbf{G}$, so $\mathbf{G}$ is Hausdorff; furthermore, e has a countable fundamental system of neighborhoods, which means that $\mathbf{G}$ is metrizable ([P2], section 2.8.1**(II)**). It is complete by Theorem 2.6. Finally, like any other manifold (Lemma 2.2), $\mathbf{G}$ is locally connected. ■

It can be shown that, if $\mathbf{G}, \mathbf{G}'$ are two real Lie groups and $\varphi : \mathbf{G} \to \mathbf{G}'$ is a morphism of topological groups, then φ is analytic, and hence a morphism of Lie groups ([BOU 82b], Chapter 3, section 8, Theorem 1). This result fails for complex Lie groups (*ibid.*, exercise 1).

COROLLARY 2.73.– *Every finite-dimensional Lie group* $\mathbf{G}$ *is locally compact; if it is connected, its topology has a countable base.*

PROOF.– By Theorem 2.72, $\mathbf{G}$ is Hausdorff, and hence locally compact by Lemma 2.7. If $\mathbf{G}$ is connected, it is countable at infinity by Corollary 2.8, so its topology has a countable base by ([P2], section 2.3.9**(II)**, Theorem 2.53). ■

A Lie group of dimension n is also known as an *n-parameter Lie group*. A complex Lie group can be viewed as a real Lie group by restricting attention to its real analytic manifold structure, which is described as the *underlying* real Lie group.

EXAMPLE 2.74.– *i) Any group equipped with discrete topology is a discrete topological group (e.g.* $\mathbb{Z}$ *with the discrete topology). Any such group also has a discrete manifold structure (Example 2.4(2)) and is therefore a Lie group (of dimension* 0*).*

ii) Any Banach space $\mathbf{E}$ *is a Lie group (with respect to its abelian group structure).*

iii) The n*-dimensional torus* $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$ *is an abelian Lie group that coincides with the circle* $\mathbb{S}^1$ *in the case* $n = 1$.

It is possible to show the following result ([BOU 82b], Chapter 3, section 6.3, Proposition 6):

LEMMA 2.75.– *Every connected, complex, compact Lie group is commutative.*

(II) LIE SUBGROUPS AND LIE QUOTIENT GROUPS A *Lie subgroup* $\mathbf{H}$ of a Banach Lie group $\mathbf{G}$ is a subgroup of $\mathbf{G}$ that is also a submanifold of $\mathbf{G}$ (section 2.3.3). This implies that $\mathbf{H}$ is closed in $\mathbf{G}$ ([BOU 82b], Chapter 3, section 1.3, Proposition 5). The quotient $\mathbf{G}/\mathbf{H}$ of a Lie group $\mathbf{G}$ by a *normal* Lie subgroup $\mathbf{H}$ is a Lie group $\mathbf{G}/\mathbf{H}$ and the sequence

$$\{1\} \longrightarrow \mathbf{H} \longrightarrow \mathbf{G} \longrightarrow \mathbf{G}/\mathbf{H} \longrightarrow \{1\}$$

is said to be exact in **LieGrp**. If $\mathbf{H}$ and $\mathbf{L}$ are normal Lie subgroups of $\mathbf{G}$ with $\mathbf{L} \subset \mathbf{H}$, then $\mathbf{H}/\mathbf{L}$ is a normal Lie subgroup of $\mathbf{G}/\mathbf{L}$ and the group isomorphism $\mathbf{G}/\mathbf{H} \cong (\mathbf{G}/\mathbf{L}) / (\mathbf{H}/\mathbf{L})$ (Noether's third isomorphism: see ([P1], section 2.2.3(**II**))) is an isomorphism of Lie groups (**exercise***: see [BOU 82b], Chapter 3, section 1.6, Corollary).

In finite dimensions, we have the following theorem. The first part (the Cartan–von Neumann theorem) is classical; the second is less widely known and was established by Kuranishi and Yamabé in 1949:

THEOREM 2.76.– *1)* Cartan–von Neumann theorem*: Any closed subgroup* $\mathbf{H}$ *of a* real *Lie group* $\mathbf{G}$ *of finite dimension*[7] *is a Lie group (when* $\mathbf{G}$ *is a complex Lie group, any closed subgroup* $\mathbf{H}$ *of* $\mathbf{G}$ *is a* real *Lie subgroup of* $\mathbf{G}$ *but not necessarily a complex Lie subgroup).*

2) Kuranishi-Yamabé theorem*: Any* arc-connected *subgroup ([P1], section 3.3.8(**VII**)) of a real finite-dimensional Lie group is a Lie subgroup.*

7 This result was given by von Neumann in 1927 in the case where $\mathbf{G} = \mathrm{GL}_n(\mathbb{R})$ and extended by Cartan [CAR 52] to the general case considered here in 1930. It fails in the infinite-dimensional Banach case ([BOU 82b], Chapter 3, section 8, Exercise 2).

PROOF.– (1): See ([BOU 82b], Chapter 3, section 8.2, Theorem 2). (2): See [YAM 50]. ■

Let $\mathbf{G}$ be a topological group and e its neutral element. The connected component of e ([P2], section 3.3.8) is said to be the *neutral component* of $\mathbf{G}$ and is written as $\overset{\circ}{\mathbf{G}}$. It is a closed and normal topological subgroup of $\mathbf{G}$ (**exercise***: see [BOU 74], Chapter 3, section 2.2, Proposition 8). The quotient group $\mathbf{G}/\overset{\circ}{\mathbf{G}}$ is totally discontinuous ([P2], section 1.1.12**(III)**) and, if $\mathbf{G}$ is locally connected, $\overset{\circ}{\mathbf{G}}$ is open in $\mathbf{G}$ and $\mathbf{G}/\overset{\circ}{\mathbf{G}}$ is discrete.

COROLLARY 2.77.– *The neutral component* $\overset{\circ}{\mathbf{G}}$ *of a real finite-dimensional Lie group is a normal Lie subgroup such that* $\mathbf{G}/\overset{\circ}{\mathbf{G}}$ *is a discrete Lie group.*

(III) PRODUCTS AND SEMI-DIRECT PRODUCTS OF LIE GROUPS If $\mathbf{H}$ and $\mathbf{H}'$ are two Banach Lie groups, the product $\mathbf{H} \times \mathbf{H}'$ of these two groups ([P1], section 2.2.1**(III)**) is a Lie group by section 2.3.7, called the *direct product* of $\mathbf{H}$ and $\mathbf{H}'$. In this case, any two elements (h, e') and (e, h') of $\mathbf{H}$ and $\mathbf{H}'$ (viewed as subgroups of $\mathbf{G} = \mathbf{H} \times \mathbf{H}'$) commute. A more general situation is when the Lie group $\mathbf{G}$ is of the form $\mathbf{H}.\mathbf{K}$ where $\mathbf{H} \cap \mathbf{K} = \{e\}$ and $\mathbf{H}$ is a normal subgroup of $\mathbf{G}$.

LEMMA 2.78.– *In this situation, every element* g *of* $\mathbf{G}$ *can be uniquely written in the form* $h.k$ $(h \in \mathbf{H}, k \in \mathbf{K})$*, and* $\mathbf{G}$ *is an extension of* $\mathbf{K}$ *by* $\mathbf{H}$ *([P1], section 2.2.2**(II)**), i.e. the short sequence in* $\mathbf{Grp}$ *given by*

$$\{1\} \longrightarrow \mathbf{H} \longrightarrow \mathbf{G} \longrightarrow \mathbf{K} \longrightarrow \{1\}$$

is exact.

PROOF.– By Noether's second isomorphism theorem ([P1], section 2.2.3**(II)**), $\mathbf{G}/\mathbf{H} \cong \mathbf{K}/(\mathbf{K} \cap \mathbf{H}) = \mathbf{K}$. Hence, writing $\varphi : \mathbf{G} \twoheadrightarrow \mathbf{G}/\mathbf{H}$ for the canonical epimorphism, we have $\varphi(\mathbf{G}) \cong \mathbf{K}$. Suppose that $h.k = h'.k'$; then $\varphi(h.k) = \varphi(h'.k')$, so $k = k'$, and therefore $h = h'$. The rest is clear. ■

DEFINITION 2.79.– *In the above conditions, the group* $\mathbf{G}$ *is called the* semi-direct product *of* $\mathbf{K}$ *by* $\mathbf{H}$ *and is written in the form* $\mathbf{H} \ltimes \mathbf{K}$ *or* $\mathbf{K} \rtimes \mathbf{H}$.

If $\mathbf{H}$ and $\mathbf{K}$ are Lie groups, their semi-direct product is a Lie group (**exercise**). If $k \in \mathbf{K}$, $\tau_k : h \mapsto h.k^{-1}.h$ is an automorphism of $\mathbf{H}$ and $k \mapsto \tau_k$ is a homomorphism from $\mathbf{K}$ into $\mathrm{Aut}(\mathbf{H})$ that determines the composition law of $\mathbf{G}$ from the laws of $\mathbf{H}$ and $\mathbf{K}$, since $(h.k).(h'.k') = h.\left(k.h'.k'^{-1}\right).(k.k') = h.\tau_k(h').(k.k')$. We sometimes write that $\mathbf{H} \times_\tau \mathbf{K}$ instead of $\mathbf{H} \ltimes \mathbf{K}$.

LEMMA 2.80.– *Let* $\mathbf{G}$ *be a* connected *topological group,* e *its neutral element, and* V *a neighborhood of* e.

i) The topological group $\mathbf{G}$ *is generated by* V*, i.e.* $\mathbf{G} = V^{\infty}$*, where* V^{∞} *is the set of* $\prod_{i=1}^{n} x_i \ (x_i \in V, n \in \mathbb{N})$.

ii) Every Hausdorff topological group with the property (i) is locally compact and countable at infinity.

PROOF.– (i): By replacing V with the interior of $V \cap V^{-1}$ if necessary, we may assume that V is open and symmetric. Let $x_0 = \prod_{i=1}^{n} x_i$; then $x_0.V$ is open ([P2], section 2.8.1(**I**)) and contained in V^{∞}; V^{∞} is therefore open, since it is a neighborhood of each of its points. Let $x \in \overline{V^{\infty}}$; the set $x.V$ is open and $x.V \cap V^{\infty} \neq \varnothing$. Let $y_0 \in x.V \cap V^{\infty}$; then y_0 is of the form $x. \prod_{i=1}^{n} x_i$, $x_i \in V$. Therefore, $x = y_0. \left(\prod_{i=1}^{n} x_i\right)^{-1} \in V^{\infty}$, and V^{∞} is closed, since it is equal to its closure. Since $\mathbf{G}$ is connected, $V^{\infty} = \mathbf{G}$ ([P2], section 2.3.8). (ii): **exercise**. ■

(IV) LOCAL LIE GROUPS

DEFINITION 2.81.– *An open neighborhood of the neutral element of a Lie group is called a* local Lie group[8]. *A morphism of manifolds of class* C^{ω}*,* $\varphi : U \to U'$*, where* U *and* U' *are local Lie groups, is said to be a* local morphism (of Lie groups) *if* $\varphi(x.y) = \varphi(x).\varphi(y)$ *for all* $x, y \in U$.

The local Lie groups and local morphisms form a category. We can similarly define local isomorphisms, monomorphisms and epimorphisms. If U, U' are open neighborhoods of the neutral elements of the Lie groups $\mathbf{G}, \mathbf{G}'$, respectively, and φ is a local morphism (respectively epimorphism) from U into U', then we say that φ is a local morphism (respectively epimorphism) from $\mathbf{G}$ to $\mathbf{G}'$.

DEFINITION 2.82.– *Two Lie groups* $\mathbf{G}, \mathbf{G}'$ *are said to be* locally isomorphic *if there exists a local isomorphism from* $\mathbf{G}$ *to* $\mathbf{G}'$.

By section 2.3.3, we have the following result:

LEMMA 2.83.– *If two Lie groups* $\mathbf{G}, \mathbf{G}'$ *are locally isomorphic, then* $T_e(\mathbf{G}) \cong T_{e'}(\mathbf{G}')$.

COROLLARY 2.84.– *Let* $\mathbf{G}$ *be a* simply connected *Lie group. If* f *is a local morphism from* $\mathbf{G}$ *to* $\mathbf{G}'$*, then there exists a unique morphism of Lie groups* $\hat{f} : \mathbf{G} \to \mathbf{G}'$ *that coincides with* f *in a neighborhood of* e.

8 This notion coincides with that of a Lie group germ ([BOU 82b], Chapter 3, section 1.10) in finite dimensions. In the Banach case, however, there exist Lie group germs that cannot be embedded in a Lie group ([EEL 66], section 3, p. 762), thus it is necessary to take a more complicated definition of a Lie group germ ([BOU 82b], Chapter 3, section 1.10, Definition 5).

PROOF.– The uniqueness follows from Lemma 2.80, since $\hat{f}$ is determined by a set of generators of $\mathbf{G}$. For a proof of existence, see ([HAU 68], Part I, section 2, Theorem 7). ■

THEOREM 2.85.– *If two simply connected Lie groups* $\mathbf{G}, \mathbf{G}'$ *are locally isomorphic, then they are isomorphic.*

PROOF.– Let f be a local isomorphism from $\mathbf{G}$ to $\mathbf{G}'$, f_1 its local inverse, and $\hat{f} : \mathbf{G} \tilde{\rightarrow} \mathbf{G}'$ and $\hat{f}_1 : \mathbf{G}' \tilde{\rightarrow} \mathbf{G}$ the isomorphisms deduced from it by Corollary 2.84. It is immediate that $\hat{f}_1$ is the inverse of $\hat{f}$, and hence the latter is an isomorphism. ■

(V) TRANSLATIONS The left and right *translations*

$$\boxed{\boldsymbol{\lambda}(s) : x \mapsto s.x, \quad \boldsymbol{\rho}(s) : x \mapsto x.s^{-1}} \qquad [2.11]$$

of a topological group $\mathbf{G}$ are analytic diffeomorphisms when $\mathbf{G}$ is a Lie group, and hence every Lie group is a pure manifold. These translations are also automorphisms of the group $\mathbf{G}$, since[9]

$$\boldsymbol{\lambda}(s.t) = \boldsymbol{\lambda}(s) \circ \boldsymbol{\lambda}(t), \quad \boldsymbol{\rho}(s.t) = \boldsymbol{\rho}(s) \circ \boldsymbol{\rho}(t).$$

LEMMA 2.86.– *The operators* $\boldsymbol{\lambda}(s), \boldsymbol{\rho}(t)$ *commute. The following conditions are equivalent: (a)* $\mathbf{G}$ *is commutative; (b)* $\boldsymbol{\lambda}(s), \boldsymbol{\lambda}(t)$ *commute for all* $s, t \in \mathbf{G}$; *(c)* $\boldsymbol{\rho}(s), \boldsymbol{\rho}(t)$ *commute for all* $s, t \in \mathbf{G}$.

PROOF.– For every $x \in \mathbf{G}$, $(\boldsymbol{\lambda}(s) \circ \boldsymbol{\rho}(t))(x) = s.x.t^{-1} = (\boldsymbol{\rho}(t) \circ \boldsymbol{\lambda}(s))(x)$. Furthermore, $(\boldsymbol{\lambda}(s) \circ \boldsymbol{\lambda}(t))(x) = s.t.x$ and $(\boldsymbol{\lambda}(t) \circ \boldsymbol{\lambda}(s))(x) = t.s.x$, so $\boldsymbol{\lambda}(s) \circ \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}(t) \circ \boldsymbol{\lambda}(s)$ if and only if $s.t = t.s$. ■

(VI) CLASSICAL LIE GROUPS Let $\mathbf{K}$ be a field; ever since H. Weyl [WEY 53], the term “classical groups” has been used to describe the group $\mathrm{GL}_n(\mathbf{K})$ and its subgroups that fix quadratic forms (orthogonal group), Hermitian forms (unitary group), or alternating forms (symplectic group). We will also include the algebraic linear groups ([P2], section 1.3.3) and their quotients by their respective centers under the same terminology. For a more detailed presentation of these groups when $\mathbf{K}$ is an arbitrary (possibly non-commutative) division ring, see [DIE 63, DIE 73] (see also [PER 96] for the case where $\mathbf{K}$ is commutative). The next section lists the classical groups and their key properties when $\mathbf{K} = \mathbb{K}$; in this case, they are all Lie groups.

Let $\mathbf{E}$ be a Banach space over $\mathbb{K}$; then the group $\mathrm{GL}(\mathbf{E}) \subset \mathcal{L}(\mathbf{E})$ of automorphisms of $\mathbf{E}$ is a Banach Lie group. Let $\mathbf{E}$ be a complex (respectively real) Hilbert space; then the unitary group (respectively orthogonal group) $\mathrm{U}(\mathbf{E})$

---

9 This explains why the expressions [2.11] seem non-symmetric at first glance. Greek-to-English memory aid: $\boldsymbol{\lambda}$ *left*, $\boldsymbol{\rho}$ *right*.

(respectively $\mathrm{O}(\mathbf{E})$) ([P2], section 3.10.3**(I)**) is a Lie subgroup of $\mathrm{GL}(\mathbf{E})$. If $\mathbf{E}$ has dimension n, $\mathrm{GL}(\mathbf{E})$ is the general linear group $\mathrm{GL}_n(\mathbb{K})$ ([P1], section 2.3.11**(I)**). This group is not compact, since the continuous mapping $\det : \mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}$ is surjective and $\mathbb{K}$ is not compact; $\mathrm{U}(\mathbf{E})$ (respectively $\mathrm{O}(\mathbf{E})$) is written as $\mathrm{U}_n(\mathbb{C})$ (respectively $\mathrm{O}_n(\mathbb{R})$). All of these groups can be identified with groups of $n \times n$ matrices with respect to an orthonormal basis. The subgroup of $\mathrm{GL}_n(\mathbb{K})$ formed by the matrices of determinant 1 is the special linear group $\mathrm{SL}_n(\mathbb{K})$; this is the derived group of $\mathrm{GL}_n(\mathbb{K})$ ([P1], section 2.3.11**(III)**). The groups $\mathrm{GL}_n(\mathbb{C})$ and $\mathrm{U}_n(\mathbb{C})$ are connected ([GOD 17], section 3.11), and $\mathrm{SL}_n(\mathbb{C})$ is simple and simply connected ([DIE 93], Volume 5, (21.18.11)). It can be shown that $\mathrm{GL}_n(\mathbb{K})/\mathrm{SL}_n(\mathbb{K}) \cong \mathbb{K}^\times$ (**exercise**).

Every algebraic subgroup of $\mathrm{GL}_n(\mathbb{R})$ ([P2], section 1.3.3) is a Lie subgroup by the Cartan–von Neumann theorem (Theorem 2.76(1)). The Lie group $\mathrm{GL}_n(\mathbb{R})$ has two connected components: the component written as $\mathrm{GL}_n^+(\mathbb{R})$ of matrices with determinant > 0 (this is the neutral component $\mathrm{GL}_n(\mathbb{R})^\circ$) and the component of matrices with determinant < 0. The *general orthogonal group* $\mathrm{O}_n(\mathbb{K})$ of orthogonal $n \times n$ matrices (i.e. the set of matrices $A \in \mathfrak{M}_n(\mathbb{K})$ satisfying $A.A^T = I_n$) is also a Lie group. If $A \in \mathrm{O}_n(\mathbb{K})$ and $d = \det(A)$, then $d = \pm 1$, so $\mathrm{O}_n(\mathbb{K})$ has two connected components, namely $\mathrm{SO}_n(\mathbb{K}) = \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K})$ (the *special orthogonal group* of rotation matrices, also written as $\mathrm{O}_n^+(\mathbb{K})$) and $\mathrm{O}_n^-(\mathbb{K})$; $\mathrm{SO}_n(\mathbb{K})$ is a normal subgroup of $\mathrm{O}_n(\mathbb{K})$ but not of $\mathrm{SL}_n(\mathbb{K})$ (**exercise**). The groups $\mathrm{O}_n(\mathbb{K})$ and $\mathrm{SO}_n(\mathbb{K})$ are compact if and only if $\mathbb{K} = \mathbb{R}$ (**exercise**); furthermore, $\mathrm{SO}_n(\mathbb{R})$ is connected (but not simply connected), abelian, and isomorphic to the torus $\mathbb{T}$ (Example 2.74(iii)) for $n = 2$, as well as simple for $n = 3$ and $n \geq 5$. The group $\mathrm{SL}_n(\mathbb{K})$ is not compact for $n \geq 2$; for example, for all $z \in \mathbb{K}$, the matrix $A(z) = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}$ belongs to $\mathrm{SL}_2(\mathbb{K})$ but $\lim_{|z| \to +\infty} \|A(z)\| = +\infty$.

The *special unitary group* $\mathrm{SU}_n(\mathbb{C})$ of matrices $A \in \mathrm{U}_n(\mathbb{C})$ satisfying $\det(A) = 1$ is compact, as is $\mathrm{U}_n(\mathbb{C})$ (**exercise**). It is furthermore a simple Lie group and is simply connected for every $n \geq 2$ ([DIE 93], Volume 3, (16.30.6)).

The general and special linear groups act canonically on the affine space $\mathfrak{Aff}(\mathbb{K})$, and the general and special orthogonal groups act canonically on the Euclidean space $\mathfrak{Euc}_n$. The projective groups considered below act canonically on the projective space $\mathbf{P}_n(\mathbb{K})$ (see section 3.5.7). The *general projective group* $\mathrm{PGL}_n(\mathbb{K}) = \mathrm{GL}_n(\mathbb{K}) / (\mathfrak{Z}(\mathrm{GL}_n(\mathbb{K})))$, where $\mathfrak{Z}(.)$ denotes the center of the group in parentheses, is also a Lie group by **(II)**; furthermore, $\mathfrak{Z}(\mathrm{GL}_n(\mathbb{K})) = \mathbb{K}^\times . I_n$. The *special projective group* $\mathrm{PSL}_n(\mathbb{K})$ is the canonical image of $\mathrm{SL}_n(\mathbb{K})$ in $\mathrm{PGL}_n(\mathbb{K})$, i.e. $\mathrm{SL}_n(\mathbb{K}) / (\mathfrak{Z}(\mathrm{SL}_n(\mathbb{K})))$; we have $\mathfrak{Z}(\mathrm{SL}_n(\mathbb{K})) \cong \boldsymbol{\mu}_n$, the group of n-th roots of unity in $\mathbb{K}$ ([P2], section 1.1.5**(III)**); $\mathrm{PSL}_n(\mathbb{K})$ is therefore a Lie group and $\mathrm{PSL}_n(\mathbb{K}) = \mathrm{PGL}_n(\mathbb{K})^\circ$. But $\mathrm{PGL}_n(\mathbb{C})$ is connected, so $\mathrm{PSL}_n(\mathbb{C}) = \mathrm{PGL}_n(\mathbb{C})$, whereas $\mathrm{PGL}_n(\mathbb{R})$ is connected if n is odd and has two connected components if n is even ([SHA 97], Chapter 8, Lemma 1.2). The group $\mathrm{PSL}_n(\mathbb{K})$ is simple if $n \geq 2$ ([DIE 63], Chapter 2, section 2; [PER 96], Theorem 4.1). The Lie groups $\mathrm{PGL}_n(\mathbb{K})$ and $\mathrm{PSL}_n(\mathbb{K})$ are locally isomorphic to $\mathrm{SL}_n(\mathbb{K})$ (**exercise**). The projective orthogonal group $\mathrm{PO}_n(\mathbb{K}) = \mathrm{O}_n(\mathbb{K}) / \{\pm I_n\}$ is also a Lie group by **(II)**, as is the *projective special orthogonal group* $\mathrm{PSO}_n(\mathbb{K}) = \mathrm{SO}_n(\mathbb{K}) / (\mathfrak{Z}(\mathrm{SO}_n(\mathbb{K})))$, where $\mathfrak{Z}(\mathrm{SO}_n(\mathbb{K})) = \{I_n\}$ if n is odd and $\{\pm I_n\}$ if n is even; $\mathrm{PO}_n(\mathbb{K})$ and $\mathrm{PSO}_n(\mathbb{K})$ are locally isomorphic to $\mathrm{O}_n(\mathbb{K})$ (**exercise**). Note that the group $\mathrm{PSO}_n(\mathbb{R})$ is simple if $n = 3$ or $n \geq 5$ and that $\mathrm{PSO}_4(\mathbb{R}) \cong \mathrm{SO}_3(\mathbb{R}) \times \mathrm{SO}_3(\mathbb{R})$ ([PER 96], Chapter 6, Theorem 7.1 and Chapter 7, Theorem 3.2). Finally, the *projective unitary group* is $\mathrm{PU}_n(\mathbb{C}) = \mathrm{SU}_n(\mathbb{C}) / (\mathfrak{Z}(\mathrm{SU}_n(\mathbb{C})))$, where $\mathfrak{Z}(\mathrm{SU}_n(\mathbb{C})) = (\mathbb{Z}/n\mathbb{Z}) .I_n$; this group is simple ([DIE 63], Chapter 2, section 5) and locally isomorphic to $\mathrm{SU}_n(\mathbb{C})$.

The symplectic group $\mathrm{Sp}_{2n}(\mathbb{K})$ is the (simple and non-compact) subgroup of $\mathrm{GL}_{2n}(\mathbb{K})$ consisting of the matrices A that satisfy $A^T.J.A = J$, where $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. Note that $J^T = J^{-1} = -J$ and $\mathrm{Sp}_{2n}(\mathbb{K}) \subset \mathrm{SL}_{2n}(\mathbb{K})$ with equality for $n = 1$; $\mathrm{Sp}_{2n}(\mathbb{C})$ is simply connected. The *unitary symplectic group* $\mathrm{USp}_{2n} = \mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{U}_{2n}(\mathbb{C})$ is simply connected for $n \geq 3$ ([WEY 53], Theorem 8.12B). The projective symplectic group $\mathrm{PSp}_{2n}(\mathbb{K}) = \mathrm{Sp}_{2n}(\mathbb{K}) / \{\pm I\}$ ($\{\pm I\} = \mathfrak{Z}(\mathrm{Sp}_{2n}(\mathbb{K}))$) is simple ([DIE 63], Chapter 2, section 5) and locally isomorphic to $\mathrm{Sp}_{2n}(\mathbb{K})$.

The *general affine group* A_n is the subgroup of $\mathrm{GL}_{n+1}(\mathbb{R})$ consisting of matrices of the form $\begin{bmatrix} \Omega & \nu \\ 0 & 1 \end{bmatrix}$, where $\Omega \in \mathrm{GL}_n(\mathbb{R})$. Identifying $\mathbb{R}^n$ with the subgroup of A_n of matrices of the form $\begin{bmatrix} I_n & \nu \\ 0 & 1 \end{bmatrix}$ gives $\mathrm{A}_n = \mathrm{GL}_n(\mathbb{R}) \rtimes \mathbb{R}^n$. We can similarly define the *affine orthogonal group* $\mathrm{E}_n = \mathrm{O}_n(\mathbb{R}) \rtimes \mathbb{R}^n$ and the *affine special orthogonal group* $\mathrm{SE}_n = \mathrm{SO}_n(\mathbb{R}) \rtimes \mathbb{R}^n$.

Write $\mathrm{D}_n(\mathbb{K})$ (respectively $\mathrm{T}_n(\mathbb{K})$, respectively $\mathrm{ST}_n(\mathbb{K})$, respectively $\mathrm{N}_n(\mathbb{K})$) for the subgroup of $\mathrm{GL}_n(\mathbb{K})$ of diagonal matrices (respectively of upper triangular matrices, respectively of upper triangular matrices with determinant 1, respectively of upper triangular unipotent matrices – i.e. whose diagonal elements are equal to 1). These are Lie subgroups of $\mathrm{GL}_n(\mathbb{K})$.

(VII) REPRESENTATION OF A LIE GROUP Representations of a group in a finite-dimensional space were defined in ([P2], section 1.3.2). This concept can be generalized as follows:

DEFINITION 2.87.– *Let* $\mathbf{G}$ *be a Lie group and* $\mathbf{F}$ *a Banach space. A linear representation of* $\mathbf{G}$ *in* $\mathbf{F}$ *is a morphism from* $\mathbf{G}$ *into* $GL(\mathbf{F})$*, or in other words an analytic mapping* $\rho : \mathbf{G} \to GL(\mathbf{F})$ *such that* $\rho(g.g') = \rho(g).\rho(g')$ *for all* $g, g' \in \mathbf{G}$.

The mapping $x \mapsto g^{-1}.x.g$ is called the interior automorphism $\text{Int}(g)$ of $\mathbf{G}$. The tangent linear mapping $T_e(\text{Int}(g))$ (from $T_e(\mathbf{G})$ into $T_e(\mathbf{G})$) is written as $\text{Ad}(g)$. We therefore have $\text{Ad}(g) : \mathbf{h}_e \mapsto g^{-1}.\mathbf{h}_e.g$ and $\text{Ad} : g \mapsto \text{Ad}(g)$ is a homomorphism from $\mathbf{G}$ into $\text{GL}(T_e(\mathbf{G}))$.

DEFINITION 2.88.– *The homomorphism Ad* : $\mathbf{G} \to GL(T_e(\mathbf{G}))$ *is called the* adjoint representation *of* $\mathbf{G}$.

2.4.2. *Manifolds of orbits and homogeneous manifolds*

(I) MANIFOLDS OF ORBITS Let $\mathbf{G}$ be a $\mathbb{K}$-Lie group and M a $\mathbb{K}$-manifold of class C^r $(r \in \{\infty, \omega\} \cap \mathbb{N}_{\mathbb{K}})$. We say that $\mathbf{G}$ acts *left differentiably* (respectively *left analytically*) on M (or that its left action on M is *infinitely differentiable*, respectively *analytic*) if $\mathbb{K} = \mathbb{R}$ and $r = \infty$ (respectively if $r = \omega$) if there is left action of $\mathbf{G}$ on M ([P1], section 2.2.8(**I**))

$$\mathbf{G} \times M \ni (s, x) \to s.x = \boldsymbol{m}(s, x) \in M$$

that is of class C^r as a mapping from $\mathbf{G} \times M$ into M. It is always assumed that $e.x = x$ for every $x \in M$, where e is the neutral element of $\mathbf{G}$. We can similarly define *right* differentiable or analytic actions of $\mathbf{G}$ on M. The adverbs *differentiably* and *analytically* can be omitted whenever they are implicitly clear from context. Left and right continuous actions of a topological group on a topological space can be defined similarly. Like in [2.11], we set $\boldsymbol{\lambda}(s)\, x = s.x$, $\boldsymbol{\rho}(s)\, x = x.s^{-1}$ and:

$$T_x(\boldsymbol{\lambda}(s)).\mathbf{h}_x = s.\mathbf{h}_x, \quad T_x(\boldsymbol{\rho}(s)).\mathbf{h}_x = \mathbf{h}_x.s^{-1}. \qquad [2.12]$$

LEMMA 2.89.– *The mappings* $\mathbf{h}_x \mapsto s.\mathbf{h}_x$ *and* $\mathbf{h}_x \mapsto \mathbf{h}_x.s^{-1}$ *are isomorphisms of* $\mathbf{Tvs}$ *from* $T_x(M)$ *onto* $T_{s.x}(M)$ *and from* $T_x(M)$ *onto* $T_{x.s^{-1}}(M)$ *respectively.*

PROOF.– We have $\boldsymbol{\lambda}(s^{-1}).\boldsymbol{\lambda}(s).x = x$, so $\boldsymbol{\lambda}(s)$ is an automorphism of M; by Theorem 2.37, $T_x(\boldsymbol{\lambda}(s))$ is therefore an isomorphism of $\mathbf{Tvs}$ from $T_x(M)$ onto $T_{s.x}(M)$. Similarly, $T_x(\boldsymbol{\lambda}(s))$ is an isomorphism of $\mathbf{Tvs}$ from $T_x(M)$ onto $T_{x.s^{-1}}(M)$. ■

Suppose that the Lie group $\mathbf{G}$ has a left action on the manifold M. Write

$$\mathbf{h}_x \mapsto s.\mathbf{h}_x, \quad \mathbf{g}_s \mapsto \mathbf{g}_s.x \qquad [2.13]$$

for the tangent linear mappings $T_x(\boldsymbol{m}(s, \bullet)) : T_x(M) \to T_{s.x}(M)$ and $T_x(\boldsymbol{m}(\bullet, x)) : T_s(\mathbf{G}) \to T_{s.x}(M)$. By the Leibniz rule [1.9], the tangent linear mapping of $\boldsymbol{m}$ at the point (s, x) is then given by:

$$T_{(s,x)}(\boldsymbol{m})(\mathbf{g}_s, \mathbf{h}_x) = s.\mathbf{h}_x + \mathbf{g}_s.x. \qquad [2.14]$$

LEMMA 2.90.– *If M is locally finite-dimensional, $\boldsymbol{m}$ is a submersion from $\mathbf{G} \times M$ onto M.*

PROOF.– By [2.14], $T_{(s,x)}(\boldsymbol{m})$ is surjective from $T_s(\mathbf{G}) \times T_x(M)$ onto $T_{s.x}(M)$. If $T_{s.x}(M)$ is finite-dimensional, then $\ker\left(T_{(s,x)}(\boldsymbol{m})\right)$ has finite codimension and hence splits in $T_s(\mathbf{G}) \times T_x(M)$ ([P2], section 3.2.2(**IV**)). ■

For every $x, x' \in M$, write $x \sim x'$ if x' belongs to the orbit $\mathbf{G}.x$ of x, i.e. if there exists $s \in \mathbf{G}$ such that $x' = s.x$ ([P1], section 2.2.8(**I**)). Then, $\sim$ is an equivalence relation and the set of equivalence classes, written as $M/\mathbf{G}$, is called the *set of orbits* (if $\mathbf{G}$ instead has a *right* action on M, the set of orbits is written as $\mathbf{G}\backslash M$). Let $\pi : M \to M/\mathbf{G}$ be the canonical surjection and let $\mathfrak{O}$ be the set of subsets U of $M/\mathbf{G}$ such that $\pi^{-1}(U)$ is open in M. Then, $\mathfrak{O}$ is the finest topology on $M/\mathbf{G}$ that makes π continuous. Equipped with this topology, $M/\mathbf{G}$ is called the *space of orbits*.

LEMMA 2.91.– *Suppose that M is locally finite-dimensional.*

1) For every $x \in M$, the mapping $\boldsymbol{m}_x : s \mapsto s.x$ from $\mathbf{G}$ into M is a subimmersion of constant rank and the stabilizer $\mathbf{S}_x$ *of x (i.e. $\mathbf{S}_x = \{s \in \mathbf{G} : s.x = x\}$) is a Lie subgroup of* $\mathbf{G}$.

*2) The stabilizers of any two points belonging to the same orbit are conjugate ([P1], section 2.2.2(**III**)).*

3) A necessary and sufficient condition for the space of orbits $M/\mathbf{G}$ to have a Hausdorff differential (respectively analytic) manifold structure that makes π a submersion is for the graph Gr$(\sim)$ of $\sim$ to be a closed submanifold of $M \times M$. This differential (respectively analytic) manifold structure of $M/\mathbf{G}$ is then unique; the manifold $M/\mathbf{G}$ is called the manifold of orbits *of the action of* $\mathbf{G}$ *on M.*

*4) Suppose that the manifold of orbits $M/\mathbf{G}$ exists. Then, for all $x \in M$, the orbit $\mathbf{G}.x$ is a manifold and the canonical isomorphism $\varphi_x : \mathbf{G}/\mathbf{S}_x \to \mathbf{G}.x$ ([P1], section 2.2.8(**I**)) is a diffeomorphism.*

PROOF.– (1) The mappings $u : \mathbf{g}_s \mapsto \left(ts^{-1}\right).\mathbf{g}_s$ and $v : \mathbf{h}_{s.x} \mapsto \left(ts^{-1}\right).\mathbf{h}_{s.x}$ are isomorphisms from $T_s(\mathbf{G})$ onto $T_t(\mathbf{G})$ and from $T_{s.x}(M)$ onto $T_{t.x}(M)$ respectively. Let $f_s = T_s(\boldsymbol{m}_x)$ and $f_t = T_t(\boldsymbol{m}_x)$; then the following diagram commutes:

$$\begin{array}{ccc} T_s(\mathbf{G}) & \xrightarrow{u} & T_t(\mathbf{G}) \\ f_s \downarrow & & \downarrow f_t \\ T_{s.x}(M) & \xrightarrow{v} & T_{t.x}(M) \end{array}$$

so $\boldsymbol{m}_x$ has finite and constant rank and hence is a subimmersion by Theorem 2.55(2). $\mathbf{S}_x$ is therefore a Lie subgroup of $\mathbf{G}$ by Lemma-Definition 2.54(3). (2): Let $g \in \mathbf{G}$ be such that $y = g.x$. Since $\mathbf{S}_x.x = x$, we have $\mathbf{S}_x.g^{-1}.y = g^{-1}.y$, so $g.\mathbf{S}_x.g^{-1}.y = y$. Hence, $\mathbf{S}_y = g.\mathbf{S}_x.g^{-1}$, which means that $\mathbf{S}_x$ and $\mathbf{S}_y$ are conjugate. (3) If π is a submersion, the same is true for

$$\pi \times \pi : M \times M \to (M/\mathbf{G}) \times (M/\mathbf{G})$$

and we have $\mathrm{Gr}(\sim) = (\pi \times \pi)^{-1}(\Delta)$, where Δ is the diagonal of $(M/\mathbf{G}) \times (M/\mathbf{G})$. But Δ is a submanifold of $(M/\mathbf{G}) \times (M/\mathbf{G})$ (Lemma 2.65), so $\mathrm{Gr}(\sim)$ is a submanifold of $M \times M$ (Lemma-Definition 2.54(4)). The equivalence relation $\sim$ is open (by the definition of the topology of $M/\mathbf{G}$), so $M/\mathbf{G}$ is Hausdorff if and only if $\mathrm{Gr}(\sim)$ is closed in $M \times M$ ([P2], section 2.3.4**(IV)**). For the converse and (4), see ([DIE 93], Volume 3, (16.10.3)), and ([BOU 82b], Chapter 3, section 1.5, Proposition 10), noting that M is locally finite-dimensional. ■

(II) ACTION OF A LIE GROUP ON A LIE GROUP Any Lie group $\mathbf{G}$ acts on itself by left and right translation. By [2.14], $T_{(x,y)}(\boldsymbol{m})(\mathbf{h}_x, \mathbf{h}_y) = x\,\mathbf{h}_y + \mathbf{h}_x\,y$.

COROLLARY 2.92.– *The tangent linear mappings $T_s\left(\gamma\left(s^{-1}\right)\right)$ and $T_s(\boldsymbol{\rho}(s))$ are isomorphisms from the Banach space $T_s(\mathbf{G})$ onto $T_e(\mathbf{G})$.*

In the following, $T_e(\mathbf{G})$ is written as $\mathfrak{g}$. Let $u : \mathbf{G} \to \mathbf{G}'$ be a morphism of real Lie groups, and suppose that $\mathbf{G}'$ is finite-dimensional. Then, $\mathbf{G}$ acts left differentiably on $\mathbf{G}'$ by $s.x' = u(s)\,x'$, and $\mathbf{S}_{e'} := \{s \in \mathbf{G} : s.e' = e'\} = \{s \in \mathbf{G} : u(s) = e'\} = \ker u$. Lemma-Definition 2.91 therefore allows us to deduce the following result:

COROLLARY 2.93.– *The morphism $u : \mathbf{G} \to \mathbf{G}'$ is a subimmersion of constant rank and* $\ker u$ *is a normal Lie subgroup of* $\mathbf{G}$.

(III) HOMOGENEOUS MANIFOLDS Consider a topological group with a left continuous action on a topological space M. This action is said to be *effective* if the only element $s \in \mathbf{G}$ such that $s.x = x, \forall x \in M$, is the neutral element e. By quotienting $\mathbf{G}$ by the closed and normal subgroup of elements that fix the points of M if necessary, we may always assume that the action of $\mathbf{G}$ on M is effective in practice. If the action of $\mathbf{G}$ on M is *transitive*, then M is a homogeneous space and there exists a bijection $\varphi_x : \mathbf{G}/\mathbf{S}_x \tilde{\to} M$ ([P1], section 2.2.8**(I)**); conversely, if $\mathbf{H}$ is a subgroup of $\mathbf{G}$, then the set of left cosets $\mathbf{G}/\mathbf{H}$ is homogeneous ([P1], section 2.2.8**(II)**). The group $\mathbf{G}$ acts *freely* on M if and only if $\mathbf{S}_x = \{e\}$ for all $x \in \mathbf{G}$, or equivalently if the mapping $s \mapsto x.s$ from $\mathbf{G}$ into $x.\mathbf{G}$ is bijective. As a result of Lemma 2.91, we have the following result:

COROLLARY 2.94.– *Let* $\mathbf{G}$ *be a Lie group.*

i) If $\mathbf{G}$ *acts differentiably* and transitively *on a locally finite-dimensional manifold M, then, for all $x \in M$, the canonical mapping $\varphi_x : \mathbf{G}/\mathbf{S}_x \to M$ is a diffeomorphism (between homogeneous spaces).*

ii) If $u : \mathbf{G} \to \mathbf{G}'$ is a finite-dimensional epimorphism of Lie groups and $\mathbf{H} = \ker u$, then the canonical mapping $\mathbf{G}/\mathbf{H} \to \mathbf{G}'$ is an isomorphism of Lie groups.

It can be shown that, given a Lie group $\mathbf{G}$ and a Lie subgroup $\mathbf{H}$ of $\mathbf{G}$, there exists a unique manifold structure on $\mathbf{G}/\mathbf{H}$ that makes the canonical surjection $\mathbf{G} \to \mathbf{G}/\mathbf{H}$ a submersion ([BOU 82b], Chapter 3, section 1.6, Proposition 11). The result stated below, established by H. Omori in 1978 [OMO 78], shows that there are only a few cases where an infinite-dimensional Lie group can act on a finite-dimensional manifold:

THEOREM 2.95.– (Omori) *Let $\mathbf{G}$ be a real connected Lie group acting differentiably, transitively, and effectively on a finite-dimensional manifold M. If M is compact, then $\mathbf{G}$ is finite-dimensional. If M is not compact and $\mathbf{G}$ is infinite-dimensional, then $\mathbf{G}$ is almost solvable*[10].

(IV) EXAMPLES (i) The Lie group $\mathrm{A}_n(\mathbb{K})$ acts transitively on the n-dimensional affine space $\mathfrak{Aff}_n(\mathbb{K})$ over $\mathbb{K}$ (section 1.3.1**(I)**). We write that $\begin{bmatrix} \Omega & \nu \\ 0 & 1 \end{bmatrix} .(P + x) = P + \nu + \Omega.x$ (with the identification from Remark 1.36, this equality can be rewritten as $\begin{bmatrix} \Omega & \nu \\ 0 & 1 \end{bmatrix} .x = \nu + \Omega.x$, taking $P = O$).

From $\mathfrak{Aff}_n(\mathbb{R})$, we can construct the Euclidean space $\mathfrak{Euc}_n$ by equipping the translation space $\mathbb{R}^n$ with the usual scalar product $\langle x|y\rangle = \sum_{1\leq i\leq n} x_i.y_i$. The Lie groups $\mathrm{A}_n(\mathbb{R})$ and $\mathrm{ASO}_n(\mathbb{R})$ act transitively on $\mathfrak{Euc}_n$, and the group $\mathrm{ASO}_n(\mathbb{R})$ conserves orientations (we will return to discuss orientations in more depth in section 4.4.4). This can be interpreted geometrically as follows: any displacement in $\mathfrak{Euc}_n$ can be decomposed into an orientation-preserving rotation $(x \mapsto \Omega.x)$ and a translation $(P \mapsto P + \nu)$.

(ii) The *Poincaré half-plane* is the region $\mathbf{P}_+ = \{z \in \mathbb{C} : \Im(z) > 0\}$ of the complex plane. The groups $\mathrm{SL}_2(\mathbb{R})$ and $\mathrm{PSL}_2(\mathbb{R})$ act transitively (and for the second group effectively) on $\mathbf{P}_+$ by $\begin{bmatrix} a & b \\ c & d \end{bmatrix} .z = \frac{a.z+b}{c.z+d}$ and $\mathbf{P}_+ \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cong \mathrm{PSL}_n(\mathbb{R})/\mathrm{PSO}_2(\mathbb{R})$ (**exercise***); $\mathbf{P}_+$ can be equipped with a hyperbolic Riemannian manifold structure that makes this action analytic[11].

10 A group $\mathbf{G}$ is said to be *almost solvable* if there exists a tower of subgroups $\mathbf{G} = \mathbf{G}_0 \supset \ldots \supset \mathbf{G}_m = \{e\}$ such that, for every $i \in \{1, \ldots, m\}$, $(\mathbf{G}_{i-1} : \mathbf{G}_i) < \infty$ or $\mathbf{G}_{i-1}/\mathbf{G}_i$ is abelian.

11 For help solving this exercise and some details regarding the manifold structure of $\mathbf{P}_+$, readers are welcome to visit the Wikipedia article on the *Poincaré half-plane model*.

This page intentionally left blank

3

Fiber Bundles

3.1. Introduction

(I) In the previous chapter, we assigned a tangent space $T_b(B)$ to each point b of a manifold B. This idea is effective for first order differential calculus, allowing us to define the tangent linear mapping $T_b(f) \in \mathcal{L}\left(T_b(B); T_{f(b)}(B')\right)$ of a morphism $f : B \to B'$ at every point $b \in B$. Our next task is to define the mapping $T(f) : b \mapsto T_b(f)$. To do this, we need to assemble the various tangent spaces $T_b(B)$ of B $(b \in B)$ into a single object, written as $T(B)$. If B is a surface in everyday space, the set of its tangent planes still belongs to this space. However, it is useful to reference points in the tangent plane $T_b(B)$ by four coordinates: two for the point $b = (b_1, b_2)$ on the surface (e.g. longitude and latitude for the Earth) and two determining the tangent vector $\mathbf{h}_b$ in $T_b(B)$ (its components with respect to a basis of $T_b(B)$). The set $T(B)$ of planes tangent to a surface B is therefore the union of $\{b\} \times T_b(B)$, i.e. a disjoint union of all $T_b(B)$. After establishing these definitions, in order to perform more advanced differential calculus, we still need to equip $T(B)$ with a manifold structure. We can then define the tangent linear mapping $T_x^2(f) := T_x(T(f))$ of $T(f) : T(B) \to T(B')$ at a point $x \in T(B)$[1]. The set $T(B)$ has a fiber bundle structure ([P2], section 5.2.3(**I**)) with base B and projection $\pi : T(B) \to B : (b, \mathbf{h}_b) \mapsto b$ (here, the canonical surjection). In $T_b(B)$, each tangent vector $\mathbf{h}_b$ is referenced by its coordinates $(\mathbf{t}_1, \mathbf{t}_2) = \mathbf{t}$, and any point $x \in T(B)$ is a pair $(b, \mathbf{t})$, where $b = \pi(x)$. The fiber bundle M is characterized by its projection $\pi : M \twoheadrightarrow B$. Once the manifold $T(B)$ has been constructed, we can iterate the process and define the fiber bundle $T(T(B)) = T^2(B)$, etc., *ad lib.*, which allows us to perform differential calculus of arbitrary order within the framework of manifolds. However, we need to expand this framework slightly further still by making the fibers themselves manifolds rather than vector spaces, leading to

1 $T_x^2(f)$ should not be confused with the partial tangent linear mapping $T^2_{(a_1,a_2)}(f)$ from section 2.3.7(**II**).

the notion of *fibration*. The details of this construction are somewhat tedious but are given in sections 3.2, 3.3 and 3.4. Most of the proofs are omitted to lighten the presentation. The majority of these proofs are entirely trivial; they are listed in full detail in [LAN 99b], Chapter 3, and [DIE 93], Volume 3, Chapter 16, for the finite-dimensional case.

(II) The notion of a Lie group (section 2.4.1**(I)**) enables us to introduce the fundamental idea of a principal bundle (which is a fibration instead of a vector bundle). Returning to the previous example, we are naturally led to choose a frame r_b of origin b in the tangent plane $T_b(B)$, where b is a point on the surface B. We can then perform a change of reference in this plane $T_b(B)$, provided that certain conditions are satisfied. For example, if B is a two-dimensional Riemannian manifold (section 4.5), every frame r_b with origin b must be orthonormal, so the change-of-frame matrices must be orthogonal matrices $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, i.e. each matrix belongs to the Lie group $\mathrm{O}_2(\mathbb{R})$ (section 2.4.1**(II)**). The set P_b of every "admissible" frame in $T_b(B)$ is therefore isomorphic to the group $\mathbf{G} = \mathrm{O}_2(\mathbb{R})$. The disjoint union of all (b, P_b) is a *fiber bundle* P acted upon by $\mathbf{G}$; we say that P is a *principal bundle*, $\mathbf{G}$ is its *structural group* and P_b is the *fiber* over b. Any point $q \in P$ is a *frame* r_b and there exists a surjection $\pi : P \twoheadrightarrow B : r_b \mapsto b$, where b is the origin of the frame and π is the *projection*. The fiber bundle P has a manifold structure; thus, if $q \in P$, we can consider the tangent vector $\mathbf{h}_q \in T_q(P)$. This vector can be interpreted as an "infinitesimal displacement" of the frame q. Among these displacements, we can distinguish the *vertical tangent vectors*, which leave the origin $\pi(q)$ of the moving frame unchanged and only rotate the frame by some infinitesimal rotation, as well as the *horizontal tangent vectors*, which conversely move the origin without rotating the frame, lying in some sense "parallel to the base". The vertical tangent vectors are defined in section 3.5 of this chapter. For a definition of the horizontal tangent vectors, we will need to wait until Chapter 7, since specifying these vectors is equivalent to specifying a *connection*. We will see that the frames r_b "move" with $b = \pi(r_b)$ not only by experiencing a change in origin but also by rotating; the fiber bundle P over these frames is therefore called the *bundle of frames*.

3.2. Tangent bundle and cotangent bundle

3.2.1. *Tangent bundle*

Let B be a manifold,

$$T(B) := \dot{\bigcup_{b\in B}} T_b(B) \qquad [3.1]$$

the disjoint union of all $T_b(B)$, i.e. the union of all $\{b\} \times T_b(B)$ ([P1], section 1.2.6**(II)**), and set $M = T(B)$. Let $c = (U, \xi, \mathbf{E})$ be a chart of B such that $b \in U$.

A tangent vector of B at the point b (section 2.2.4) is written as $\mathbf{h}_b = \vartheta_c(\mathbf{h})$. The natural projection $\pi : M \to B : (b, \mathbf{h}_b) \mapsto b$ is clearly surjective. For every $b \in U$, $d_b\xi : T_b(B) \to \mathbf{E}$ is an isomorphism. Let

$$\psi_c : U \times \mathbf{E} \to \pi^{-1}(U) : (b, \mathbf{h}) \mapsto \left(b, (d_b\xi)^{-1}.\mathbf{h}\right) = (b, \mathbf{h}_b). \quad [3.2]$$

LEMMA-DEFINITION 3.1.– *1) The $\pi^{-1}(U)$ constructed above has the following properties:*

i) Let $\tau_U : \pi^{-1}(U) \to U \times \mathbf{E} : (b, \mathbf{h}_b) \mapsto (b, d_b\xi.\mathbf{h}_b)$; then τ_U is a homeomorphism that commutes with the projection onto U; in other words, the following diagram commutes:

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\tau_U} & U \times \mathbf{E} \\ & \searrow \quad \swarrow & \\ & U & \end{array}$$

ii) Let $(U_i, \xi_i, \mathbf{E}_i)$ $(i = 1, 2)$ be two charts. If $U_i \cap U_j \neq \varnothing$, then $\mathbf{E} := \mathbf{E}_i \cong \mathbf{E}_j$; thus, let $\xi_i^j = \xi_j \circ \xi_i^{-1} : \xi_i(U_i \cap U_j) \to \xi_j(U_i \cap U_j)$ and

$$\tau_i^j : \xi_i(U_i \cap U_j) \times \mathbf{E} \to \xi_j(U_i \cap U_j) \times \mathbf{E} : (b, \mathbf{h}) \mapsto \left(\xi_i^j(b), d_b\xi_i^j.\mathbf{h}\right).$$

The differential $D\xi_i^j : U_i \cap U_j \to \mathcal{L}(\mathbf{E}) : b \mapsto D\xi_i^j(b)$ is of class C^r and $\tau_i^j = \tau_j \circ \tau_i^{-1}$, where $\tau_i = \tau_{U_i}$.

*2) The triple $\lambda = (T(B), B, \pi)$ is a fiber bundle with base B and projection π ([P2], section 5.2.3**(I)**), said to be the* tangent bundle. *The fiber of M at $b \in B$ is $\pi^{-1}(\{b\}) = \{b\} \times T_b(B)$, also written as $T(B)_b$. We say that M is the* space *of λ.*

3) Let $c = (U, \xi, \mathbf{E})$ be a chart of B. Then:

$$t_c =: \left(\pi^{-1}(U), \zeta_c, \mathbf{E} \times \mathbf{E}\right), \quad \zeta_c =: (\xi \times 1_{\mathbf{E}}) \circ \psi_c^{-1}$$

is called the vector chart *(or* fibered chart*) of M associated with c. The charts t_c endow M with a manifold structure (see (6) below). Since the fibers $T(B)_b$ have a vector space structure, the fiber bundle λ (or, erroneously, M itself) is called a* vector bundle.

4) If $c_2 = (U, \xi_2, \mathbf{E})$ is another chart of B defined in the same domain and $\mathbf{u} = \xi \circ \xi_2^{-1} : \xi_2(U) \to \xi(U)$ is the transition diffeomorphism*, the* transition diffeomorphism for the fibered charts of $T(B)$ associated with *c, c_2 is*

$$(\xi, \mathbf{h}) \mapsto (\mathbf{u}(\xi), D\mathbf{u}(\xi).\mathbf{h}). \quad [3.3]$$

5) Since $\pi|_{\pi^{-1}(U)} = pr_1 \circ \psi_c^{-1}$, π *is a submersion.*

6) (a) If B *is a Banach manifold (respectively a Hilbert manifold), then so is* $T(B)$. *(b) If* B *is locally compact, metrizable and separable, then so is* $T(B)$.

PROOF.– Claims (1)–(5) are tautological. Claim (6)(a) follows from (3) and [P2], sections 3.4.1**(I)**, 3.11.3**(III)**. For claim (6)(b), see [DIE 93], Volume 3, section 16.13.3. ■

Note that, for all $(b, \mathbf{h}_b) \in \pi^{-1}(U)$, $\psi_c^{-1}(b, \mathbf{h}_b) = (b, \mathbf{h})$ with $\mathbf{h} = d_b\xi.\mathbf{h}_b \in \mathbf{E}$ and

$$\left((\xi \times 1_{\mathbf{E}}) \circ \psi_c^{-1}\right)(b, \mathbf{h}_b) = (\xi(b), \mathbf{h}).$$

REMARK 3.2.– *We can also define the tangent bundle* $T(B)$ *of a manifold of type* $(\mathcal{FN})$ *or* $(\mathcal{SN})$ *([KRI 97], Chapter 6, section 28): it is the space of "kinematic" tangent vectors (there is also an "operational tangent bundle"* $D(B)$*: see footnote 5, section 2.2.4). See Remark 3.35.*

The tangent bundle of a circle is shown in Figure 3.1: the tangent spaces (top) do not intersect, as shown by the bottom of the figure.

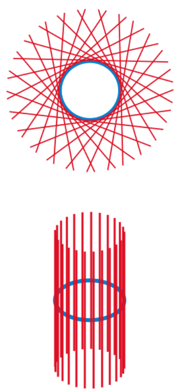


Figure 3.1. *Tangent bundle of a circle. For a color version of this figure, see www.iste.co.uk/bourles/fundamentals3.zip*

3.2.2. *Cotangent bundle*

Let B be a manifold. This time, write

$$M = T^{\vee}(B) = \dot{\bigcup_{b \in B}} T_b^{\vee}(B).$$

Let $\pi : M \to B : (b, \mathbf{h}_b^\vee) \mapsto b$ be the natural projection and $(U, \xi, \mathbf{E})$ a chart of B. For all $b \in U$, ${}^t d_b\xi : \mathbf{E}^\vee \to T_b^\vee (B)$ is an isomorphism. Let

$$\psi_c^\vee : U \times \mathbf{E}^\vee \to \pi^{-1}(U) : (b, \mathbf{h}^\vee) \mapsto \left(b,\ {}^t d_b\xi.\mathbf{h}^\vee\right). \qquad [3.4]$$

Analogously to Lemma-Definition 3.1, we have the following result:

LEMMA-DEFINITION 3.3.– *1) The $\pi^{-1}(U)$ constructed above have the following properties:*

i) Let $\sigma_U : \pi^{-1}(U) \to U \times \mathbf{E}^\vee : (b, \mathbf{h}_b^\vee) \mapsto \left(b,\ {}^t(d_b\xi)^{-1}.\mathbf{h}_b^\vee\right)$; then σ_U is a bijection that commutes with the projection onto U, i.e. the following diagram commutes:

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\sigma_U} & U \times \mathbf{E}^\vee \\ & \searrow \quad \swarrow & \\ & U & \end{array}$$

ii) Let $(U_i, \xi_i, \mathbf{E}_i)$ $(i = 1, 2)$ be two charts of B. If $U_i \cap U_j \neq \varnothing$, then $\mathbf{E} := \mathbf{E}_1 \cong \mathbf{E}_2$; thus, let $\xi_j^i = \xi_j \circ \xi_i^{-1} : \xi_i(U_i \cap U_j) \to \xi_j(U_i \cap U_j)$ and

$$\sigma_i^j : \xi_i(U_i \cap U_j) \times \mathbf{E}^\vee \to \xi_j(U_i \cap U_j) \times \mathbf{E}^\vee : (b, \mathbf{h}^\vee) \mapsto \left(\xi_j^i(y),\ {}^t\left(d_b\xi_j^i\right)^{-1}.\mathbf{h}^\vee\right).$$

Then, $\sigma_i^j = \sigma_j \circ \sigma_i^{-1}$, where $\sigma_i = \sigma_{U_i}$.

2) The triple $\lambda = (T^\vee(B), B, \pi)$ is a fiber bundle with base B and projection π, known as the cotangent bundle. *The fiber of M at $b \in B$ is $\pi^{-1}(\{b\}) = \{b\} \times T_b^\vee(B)$, also written as $T^\vee(B)_b$.*

3) Let $c = (U, \xi, \mathbf{E})$ be a chart of B. Then:

$$t_c^\vee := \left(\pi^{-1}(U), \eta_c, \mathbf{E} \times \mathbf{E}^\vee\right), \quad \eta_c := (\xi \times 1_{\mathbf{E}^\vee}) \circ \psi_c^{-1}$$

is a vector chart *(or* fibered chart*) of M associated with c. These charts endow M with a manifold structure. The fibers $T^\vee(B)_b$ have a vector space structure, so the fiber bundle λ (or erroneously M) is a* vector bundle.

4) If $c_2 = (U, \xi_2, \mathbf{E})$ is another chart of B defined in the same domain and $\mathbf{u} = \xi \circ \xi_2^{-1} : \xi_2(U) \to \xi(U)$ is the transition diffeomorphism, *then the* transition diffeomorphism for the fibered charts of $T^\vee(B)$ associated with *c, c_2 is*

$$(\xi, \mathbf{h}^\vee) \mapsto \left(\mathbf{u}(\xi),\ {}^t(D\mathbf{u}(\xi))^{-1}.\mathbf{h}^\vee\right). \qquad [3.5]$$

5) Since $\pi|_{\pi^{-1}(U)} = pr_1 \circ \psi_c^{-1}$, π *is a submersion.*

6) (a) If B *is a Banach manifold (respectively a Hilbert manifold), then so is* $T^{\vee}(B)$. *(b) If* B *is locally compact, metrizable and separable, then so is* $T^{\vee}(B)$.

3.2.3. *Tangent bundle and cotangent bundle functors*

The *tangent bundle functor* T from the category of manifolds in itself is defined as follows: let B, B' be two manifolds, and let $f : B \to B'$ be a morphism. Then:

– $T(B)$ is the tangent bundle defined above;

– $Tf : T(B) \to T(B')$ is the mapping whose restriction to the fiber $T(B)_b = \{b\} \times T_b(B)$ is $\mathbf{h}_b \mapsto (f(b), T_b(f).\mathbf{h}_b)$ (this restriction can be identified with the tangent linear mapping $T_b(f)$). This is a covariant functor ([P1], section 1.2.1(**I**)).

The *cotangent bundle functor* $T^{\vee}$ from the category of manifolds in itself is defined as follows: let B, B' be two manifolds and let $f : B \to B'$ be a morphism. Then:

– $T^{\vee}(B)$ is the cotangent bundle defined above;

– $T^{\vee}f : T^{\vee}(B') \to T^{\vee}(B)$ is the mapping whose restriction to the fiber $\{b\} \times T^{\vee}_{f(b)}(B')$ is $\mathbf{h}^{\vee}_{b'} \mapsto (b, T^{\vee}_b(f).\mathbf{h}^{\vee}_{b'})$; where $b' = f(b)$ and $T^{\vee}_b(f) = {}^t f_b$ is the cotangent linear mapping at b (section 2.3.11). This is a contravariant functor ([P1], section 1.2.1(**II**)).

3.3. Fibrations

Fibrations are fiber bundles whose projections are surjective and which satisfy a "local trivialization condition" **(T)** (see Lemma-Definition 3.4(1)). They provide a general framework for working with coverings (section 3.3.3), vector bundles (section 3.4) and principal bundles (section 3.5).

If B and F are two manifolds, the product space $B \times F$ adjoined by the first projection $\pi : pr_1 : B \times F \twoheadrightarrow B$ gives a typical example of a fibration with base B whose fibers are diffeomorphic to F. This fibration written as $(B \times F, B, pr_1)$ is said to be *trivial*.

The cylinder $\mathbb{S}^1 \times [0,1]$ with the unit circle $B = \mathbb{S}^1$ as is base is a trivial bundle whose fibers are diffeomorphic to $F = [0,1]$[2]. The Möbius strip[3] in Figure 3.2 is a

2 $\mathbb{S}^n$ is the n-dimensional sphere of center 0 and radius 1.

3 See the Wikipedia article on the *Möbius strip*.

fiber bundle with base $\mathbb{S}^1$ and fibers diffeomorphic to $[0,1]$. However, it is not trivial, since it cannot be deformed into the cylinder considered above without tearing (even if made from a pliable material).

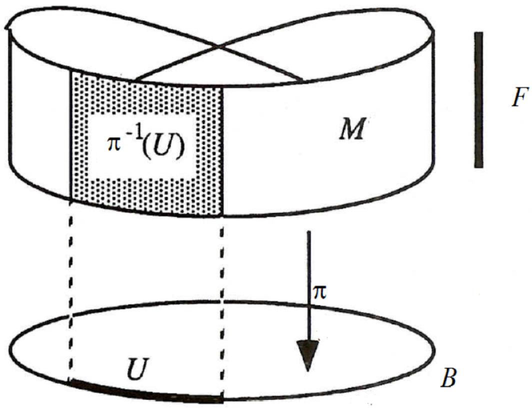


Figure 3.2. *Fibration*

The structure of trivial bundles is far too specific to be fruitful. The general notion of a fibration can be deduced from the notion of a trivial bundle by *localizing* the triviality condition, as discussed in the next section.

3.3.1. *Notion of a fibration*

LEMMA-DEFINITION 3.4.– *1) A* fibration *of class C^r is a* fiber bundle $\lambda = (M, B, \pi)$ *([P2], section 5.2.3(**I**) and section 3.2.1) such that M and B are manifolds (namely, Banach manifolds of class C^r, in accordance with the conventions **(C1)** and **(C2)**), respectively called the* space *and the* base *of the fibration; $\pi : M \to B$ is a* surjective morphism *and the following* local trivialization condition ***(T)** is satisfied:*

***(T)** For every $b \in B$, there exists an open neighborhood U of b, a manifold F and a so-called* trivializing *diffeomorphism τ defined by*

$$\tau^{-1} = \psi : U \times F \tilde{\to} \pi^{-1}(U)$$

such that $\pi(\psi(y,t)) = y$ for every $y \in U$ and every $t \in F$ (see Figure 3.2).

2) The projection π is a submersion. For every $b \in B$, the fiber $M_b = \pi^{-1}(\{b\})$ over $\{b\}$ is a closed submanifold of M.

3) Let $x \in M$ and $b = \pi(x)$. Since $M_b \subset M$, $T_x(M_b) \subset T_x(M)$ (by the canonical identification). More precisely, $T_x(M_b) = \ker T_x(\pi)$. The elements of $T_x(M_b)$ are called the vertical tangent vectors at the point x.

4) Let $\lambda = (M, B, \pi)$ be a fibration and B' a submanifold (e.g. an open set) of B. Then, $\left(\pi^{-1}(B'), B', \pi\left|_{\pi^{-1}(B')}\right.\right)$ is another fibration, said to be induced *by the fibration λ on B' and denoted by $\lambda|_{B'}$.*

PROOF.– (2) The restriction $\pi^{-1}(U) \to U$ of π is $pr_1 \circ \tau$, so π is a submersion and M_b is a submanifold of M by Lemma-Definition 2.54(3). (4) We can reduce to the case, where M and M_b are Banach spaces, $x = 0, b = 0$ and $\pi : M \to B$ is linear. Then, $M_0 = \pi^{-1}(\{0\}) = \ker(\pi) = \ker T_0(\pi)$. ■

EXAMPLE 3.5.– *a) Let B be a manifold, $T(B)$ its tangent bundle and π its projection (section 3.2.1). Then, $\lambda = (T(B), B, \pi)$ is a fibration such that $F = \mathbf{E} \cong T_b(B)$ in **(T)** by canonically identifying the Banach space $\mathbf{E}$ and the corresponding manifold F defined by the chart $(F, 1_F, \mathbf{E})$. The fiber $T(B)_b = \{b\} \times T_b(B)$ is canonically diffeomorphic to the tangent space $T_b(B)$.*

*b) Let B be a manifold, $T^\vee(B)$ its cotangent bundle and π' its projection (section 3.2.2). Then, $\lambda' = (T^\vee(B), B, \pi')$ is a fibration such that $F = \mathbf{E}^\vee$, $\mathbf{E} \cong T_b(B)$ in **(T)**. The fiber $T^\vee(B)_b = \{b\} \times T_b^\vee(B)$ is canonically diffeomorphic to the cotangent space $T_b^\vee(B)$.*

As a result of the canonical diffeomorphisms $T(B)_b \cong T_b(B)$ and $T^\vee(B)_b \cong T_b^\vee(B)$, the tangent space (respectively cotangent space) $T_b(B)$ (respectively $T_b^\vee(B)$) of B at the point b is called the fiber *of $T(B)$ (respectively $T^\vee(B)$) at the point $b \in B$.*

DEFINITION 3.6.– *Let A be a set and $f^0 : A \to B$. We say that $f : A \to M$ is a* lifting *of f^0 into M if $\pi \circ f = f^0$, i.e. the following diagram commutes:*

$$\begin{array}{ccc} A & \xrightarrow{f^0} & B \\ & f \searrow & \pi \uparrow \\ & & M \end{array}$$

LEMMA-DEFINITION 3.7.– *Let $\lambda = (M, B, \pi)$ and $\mu = (M', B', \pi')$ be two fibrations.*

i) A morphism *from λ into λ' is defined as a pair (f^0, f), where $f^0 : B \to B'$ and $f : M \to M'$ are morphisms of manifolds and $\pi' \circ f = f^0 \circ \pi$. The following diagram commutes:*

$$\begin{array}{ccc} M & \xrightarrow{\pi} & B \\ f \downarrow & & \downarrow f^0 \\ M' & \xrightarrow{\pi'} & B' \end{array}$$

*ii) The composition of two morphisms of fibrations is a morphism of fibrations; thus, the fibrations form a category ([P1], section 1.1.1(**I**)). This category is concrete with base* $\mathbf{Set}$*, which gives us the notion of* fibration structure *([P1], section 1.3.1).*

iii) Isomorphisms *of fibrations are defined obviously.*

iv) When $B = B'$ *and* $(1_B, f)$ *is a morphism (respectively an isomorphism), we say that* f *is a* B-morphism *from* λ *into* λ'*, or from* M *into* M' *(respectively a* B*-isomorphism from* λ *onto* λ'*, or from* M *onto* M'*). This means that* $f : M \to M'$ *is a morphism (respectively an isomorphism) of manifolds and* $\pi' \circ f = \pi$*. The following diagram then commutes:*

$$\begin{array}{lll} M & \xrightarrow{\pi} & B \\ f \downarrow & \nearrow \pi' & \\ M' & & \end{array} \qquad [3.6]$$

LEMMA 3.8.– *For every* B*-morphism* $f : M \to M'$ *and every* $b \in B$*, there exists a morphism of manifolds* $f_b : M_b \to M'_b$ *such that* $f_b(x) = f(x)$ *for all* $x \in M_b$*.*

PROOF.– Let $\tilde{f}_b : M_b \to M' : x \mapsto f(x)$. Then, $\tilde{f}_b$ is a morphism of manifolds by Lemma-Definition 3.4(2), since M_b is a submanifold of M. Furthermore, $\pi'\left(\tilde{f}_b(x)\right) = \pi(x) = b$, since $x \in M_b = \pi^{-1}(\{b\})$, so $\tilde{f}_b(x) \in \pi'^{-1}(\{b\}) = M'_b$. Hence, $\tilde{f}_b$ induces a mapping $f_b : M_b \to M'_b$; this is a morphism of manifolds because M'_b is a submanifold of M'. ■

A B-morphism $f : M \to M'$ is an isomorphism if and only if $f_b : M_b \to M'_b$ is an isomorphism for every $b \in B$ (**exercise**).

3.3.2. *Fiber product and preimage of fibrations*

LEMMA-DEFINITION 3.9.– *Let* $\lambda = (M, B, \pi)$ *and* $\lambda' = (M', B, \pi')$ *be two fibrations with the same base* B*. Then, the projections* $\pi : M \twoheadrightarrow B$ *and* $\pi' : M' \twoheadrightarrow B$ *are submersions and hence are transversal (Remark 2.67), and the fiber product* $M \times_B M'$ *over* B *exists (section 2.3.9). Therefore,*

$$\lambda \times_B \lambda' := (M \times_B M', B, \pi \times_B \pi')$$

is a fibration, said to be the fiber product of λ and λ' *over* B *([P1], section 1.2.7).*

REMARK 3.10.– *We can also define the (non-fiber)* product *of two fibrations* $\lambda = (M, B, \pi)$ *and* $\lambda' = (M', B', \pi')$ *with different bases as follows:*

$$\lambda \times \lambda' := (M \times M', B \times B', \pi \times \pi') .$$

LEMMA-DEFINITION 3.11.– *i) Let* $\lambda = (M, B, \pi)$ *be a fibration,* B' *a manifold and* $f^0 : B' \to B$ *a morphism. Let* $\pi' : B' \times_B M \to B'$ *be the canonical morphism, namely*

$$B' \times_B M = \left\{ (b', x) \in B' \times M : f^0 (b') = \pi (x) \right\} ,$$
$$\pi' \ : \ B' \times_B M \to B' : (b', x) \mapsto b'.$$

Then, the triple

$$f^{0*} (\lambda) := (B' \times_B M, B', \pi')$$

is a fibration called the preimage of λ under f^0. *We write that* $M' = B' \times_B M$.

ii) The projection π' *is the restriction to* M' *of* $pr_1 : B' \times M \twoheadrightarrow B'$; *let* $\lambda' := f^{0*} (\lambda)$. *Then,* λ' *is a fibration such that the fiber* $M'_{b'}$ *is canonically diffeomorphic to* $M_{f^0(b')}$ *for every* $b' \in B'$.

iii) Let $f' : M' \to M$ *be the restriction to* $M' := B' \times_B M$ *of* $pr_2 : B' \times M \to M$, *i.e. the canonical mapping defined for all* $(b', x) \in M'$ *by*

$$f' ((b', x)) = x. \qquad [3.7]$$

The pair $\left(f^0, f'\right)$ *is a morphism from* $\lambda' = f^{0*} (\lambda)$ *into* λ, *said to be the* canonical morphism *because it satisfies the following universal property:*

If $\left(f^0, g'\right)$ is a morphism from a fibration $\nu = (M'', B', \sigma)$ with base B' into the fibration λ, then there exists a unique B'-morphism $u : M'' \to B' \times_B M$ such that $g' = f' \circ u$.

We say that f' *is an* f^0-morphism.

iv) The following diagram commutes:

$$\begin{array}{ccc} B' & \xrightarrow{f^0} & B \\ \uparrow \pi' & & \uparrow \pi \\ M' := B' \times_B M & \xrightarrow{f'} & M \end{array}$$

PROOF.– (ii) Let $y \in B$. Then:

$$\begin{aligned}(B' \times_B M)_{b'} &= \pi'^{-1}(\{b'\}) = \left\{(b', x) : f^0(b') = \pi(x)\right\} \\ &= \{b'\} \times \pi^{-1}\left(f^0(b')\right) = \{b'\} \times M_{f^0(b')}.\end{aligned}$$

The proof of the other claims is left to the reader as an **exercise**. ■

EXAMPLE 3.12.– *Let B, B' be two manifolds and $f^0 : B' \to B$ a morphism. Consider the fibrations $\lambda = (T^\vee(B), B, \pi)$ and $f^{0*}(\lambda) = (B' \times_B T^\vee(B), B', \pi')$. The fiber of $f^{0*}(\lambda)$ at the point $b' \in B'$ is canonically diffeomorphic to $\{b'\} \times T^\vee_{f^0(b')}(B)$. This is the fiber considered in section 3.2.2.*

3.3.3. *Coverings*

(I) COVERING OF A MANIFOLD The proof of (1) is obvious.

LEMMA-DEFINITION 3.13.– *Let $\lambda = (M, B, \pi)$ be a fibration.*

1) The following conditions are equivalent:

i) The fibers of λ are discrete.

ii) For every $b \in B$, there exists an open neighborhood U of b in B such that $\pi^{-1}(U)$ is the union of a (finite or infinite) sequence (V_n) of pairwise disjoint open subsets of M and each restriction $\pi_n : V_n \to U$ of π is a diffeomorphism from V_n onto U.

2) If the conditions of (1) are satisfied, λ is said to be a covering *of B.*

(II) RIEMANN SURFACES Let $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ be a holomorphic submersion. Let

$$\begin{aligned}&Y = \left\{(z_1, z_2) \in \mathbb{C}^2 : f(z_1, z_2) = 0\right\}, \\ &Y_0 = \{(z_1, z_2) \in Y : D_2 f(z_1, z_2) \neq 0\}.\end{aligned}$$

By Lemma-Definition 2.54(3), Y is a submanifold of $\mathbb{C}^2$. Furthermore, Y_0 is an open subset of Y, since $D_2 f$ is continuous. Hence, Y_0 is a submanifold of $\mathbb{C}^2$.

DEFINITION 3.14.– *We say that Y_0 is the* Riemann surface *(relative to the second coordinate) defined by the function f.*

EXAMPLE 3.15.– *Previously, the complex logarithm was only determined up to mod.*$2\pi i$ *([P2], section 4.4.2(**VI**)). The notion of a Riemann surface eliminates this difficulty by considering each determination on its own leaf isomorphic to an open set* $U \subset \mathbb{C}^{\times}$ *as in Lemma-Definition 3.13(1)(ii). The leaves are sliced and then glued together one above the other in a vertical spiral to form a covering (Figure 3.3). More precisely, let* $f(z_1, z_2) = z_1 - e^{z_2}$*. With the above notation,* $Y = Y_0$ *and* Y *is an analytic subgroup of the complex analytic group* $\mathbb{C}^{\times} \times \mathbb{C}$ *whose composition law is*

$$(x, y)(x', y') = (xx', y + y').$$

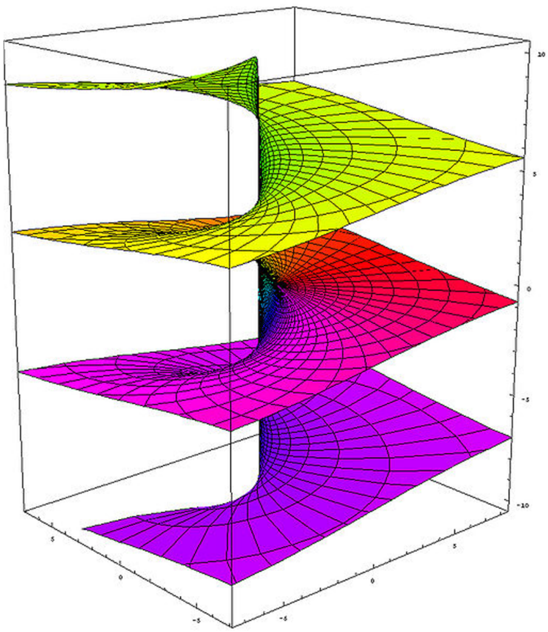


Figure 3.3. *Riemann surface of the logarithm. For a color version of this figure, see www.iste.co.uk/bourles/fundamentals3.zip*

We say that Y *is the* Riemann surface of the logarithm. *For every* $t \in Y$*, we set* $\ln(t) = pr_2(t)$*, so that*

$$\ln(tt') = \ln(t) + \ln(t')$$

and $t \mapsto \ln(t)$ *is a holomorphic mapping from* Y *into* $\mathbb{C}$.

The Riemann surface Y of the logarithm is a covering of $\mathbb{C}^\times$ of fiber type $\mathbb{Z}$ *(i.e. whose fibers are isomorphic to* $\mathbb{Z}$*) for the restriction π of pr_1 to Y (this projection π should of course not be confused with the number $\pi = 3.1415...$). Indeed, every point $z_0 = r_0 e^{i\theta_0} \in \mathbb{C}^\times$ (with $r_0 > 0$, $\theta_0 \in \mathbb{R}$) has some neighborhood U in $\mathbb{C}^\times$ that is the image under the bijection $(r, \theta) \mapsto re^{i\theta}$ of the open set $V \subset \mathbb{R}^2$ defined by $r > 0$, $\theta_0 - \pi < \theta < \theta_0 + \pi$. It immediately follows that*

$$\left(re^{i\theta}, k\right) \mapsto \ln(r) + i\theta + 2k\pi i$$

is a diffeomorphism from $U \times \mathbb{Z}$ onto $\pi^{-1}(U)$ and that $\lambda = (Y, \mathbb{C}^\times, \pi)$ is a fibration and hence a covering *of* $\mathbb{C}^\times$.

See also Example 3.52.

(III) UNIVERSAL COVERING The *universal covering* of a finite-dimensional connected manifold B is a covering (M, B, π) that makes M simply connected ([P1], section 3.3.8(**VII**)). It can be shown that this universal covering exists and is unique up to B-isomorphism ([DIE 93], Volume 3, (16.29.1)).

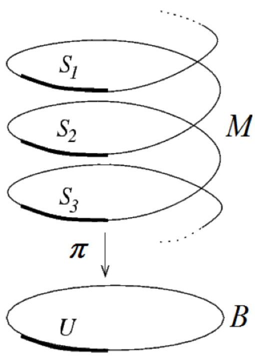


Figure 3.4. *Universal covering of the circle $B = \mathbb{S}^1$*

EXAMPLE 3.16.– *The universal covering of the circle $B = \mathbb{S}^1$ is a helix isomorphic to $\mathbb{R}$ (Figure 3.4). The Riemann surface of the logarithm (Example 3.15) is a universal covering of $\mathbb{C}^\times$, and the Poincaré group $\pi_1(\mathbb{C}^\times)$ ([P1], section 3.3.8(**VII**)) of $\mathbb{C}^\times$ is isomorphic to $\mathbb{Z}$*[4].

(IV) UNIVERSAL COVERING OF A LIE GROUP If $\mathbf{G}$ is a connected Lie group, its universal covering $\rho : \tilde{\mathbf{G}} \twoheadrightarrow \mathbf{G}$ can be constructed as follows: let V be a simply connected symmetric open neighborhood of the neutral element e and φ a bijection

4 Roughly speaking, the Riemann surface of the logarithm "looks like" $\mathbb{C}^\times \times \mathbb{Z}$.

from V onto a set $\tilde{V}$. The bijection φ transports to $\tilde{V}$ the analytic manifold structure of V. For every $x \in V$, define $\tilde{x} = \varphi(x)$. The group $\tilde{\mathbf{G}}$ whose generators are the elements of $\tilde{V}$ and whose relations are given by $\tilde{x}.\tilde{y} = \tilde{z}$ for all $x, y, z \in V$ such that $x.y = z$ is a Lie group. Let $\rho : \tilde{\mathbf{G}} \to \mathbf{G}$ be the unique morphism of groups that extends φ^{-1}. Then, $\rho|_{\tilde{V}} : \tilde{V} \to V$ is a local isomorphism from $\tilde{\mathbf{G}}$ to $\mathbf{G}$ (Definition 2.82) and $\tilde{\mathbf{G}}$ is simply connected ([COH 65], Theorem 7.4.5). Hence, $\tilde{\mathbf{G}}$ is the universal covering of $\mathbf{G}$ (Corollary 2.84).

It can be shown ([BOU 16], Chapter 4, section 3.5, Proposition 8) that the kernel $\ker(\rho)$ is discrete and isomorphic to the Poincaré group $\pi_1(\mathbf{G})$. However, every discrete normal subgroup of a topological group $\mathbf{H}$ is contained in the center $\mathfrak{Z}(\mathbf{H})$ of $\mathbf{H}$ (**exercise**), so $\pi_1(\mathbf{G}) \cong \ker(\rho) \subset \mathfrak{Z}\left(\tilde{\mathbf{G}}\right)$; in particular, $\ker(\rho)$ is abelian and $\mathbf{G} \cong \tilde{\mathbf{G}} / \ker(\rho)$.

If $\mathbf{G}$, $\mathbf{G}'$ are two locally isomorphic Lie groups, the above construction shows that their universal coverings are isomorphic and thus can be identified. Conversely, if $\mathbf{G}$, $\mathbf{G}'$ are two Lie groups with the same universal covering $\tilde{\mathbf{G}}$, then there exist local isomorphisms from $\tilde{\mathbf{G}}$ to $\mathbf{G}$ and from $\tilde{\mathbf{G}}$ to $\mathbf{G}'$, so $\mathbf{G}$ and $\mathbf{G}'$ are locally isomorphic. In summary:

THEOREM 3.17.– *Two connected Lie groups are locally isomorphic if and only if they have the same universal covering.*

It is also possible to show the following result (**exercise***: see [DIE 93], Volume 3, (16.30.4)):

LEMMA 3.18.– *If $\mathbf{G}$ is a connected Lie group, then any given covering $(\mathbf{G}_1, \rho_1)$ of $\mathbf{G}$ is isomorphic to some quotient group $\tilde{\mathbf{G}}/\mathbf{D}$, where $\mathbf{D}$ is a subgroup of* $\ker(\rho)$*, and* $\ker(\rho_1) \cong \ker(\rho)/\mathbf{D}$ *is isomorphic to the quotient group $\pi_1(\mathbf{G})/\psi(\mathbf{D})$ of $\pi_1(\mathbf{G})$, where ψ :* $\ker(\rho) \tilde{\to} \pi_1(\mathbf{G})$ *is the isomorphism specified above.*

As we saw earlier, the universal covering of the circle $\mathbb{S}^1$ is (isomorphic to) $\mathbb{R}$, and $\pi_1\left(\mathbb{S}^1\right) \cong \mathbb{Z}$. It can be shown that the universal covering of $\mathrm{SO}_3(\mathbb{R})$ is $\mathrm{SU}_2(\mathbb{C})$. For all $n \geq 3$, the universal covering of $\mathrm{SO}_n(\mathbb{K})$ is the "spinor group", written as $\mathrm{Spin}_n(\mathbb{K})$ ([DIE 93], Volume 5, (21.16.10)), and $\pi_1(\mathrm{SO}_n(\mathbb{K})) = \mathbb{Z}/2\mathbb{Z}$ ([DIE 93], Volume 3, section 16.30, Problem 10)[5]. We therefore have $\mathrm{Spin}_3(\mathbb{R}) \cong \mathrm{SU}_2(\mathbb{C})$, and $\mathrm{Spin}_4(\mathbb{R}) \cong \mathrm{SU}_2(\mathbb{C}) \times \mathrm{SU}_2(\mathbb{C})$. The special linear group $\mathrm{SL}_2(\mathbb{R})$ is not simply connected, and $\pi_1(\mathrm{SL}_2(\mathbb{R})) = \mathbb{Z}$; for a construction of the universal covering of $\mathrm{SL}_2(\mathbb{R})$ (which cannot be expressed as a group of matrices), see, for example, [GOD 17], section 2.7. The universal covering of the group $\mathrm{GL}_n(\mathbb{C})$ is isomorphic to $\mathrm{SL}_n(\mathbb{C}) \times \mathbb{R}^2$, and $\pi_1(\mathrm{GL}_n(\mathbb{C})) \cong \mathbb{Z}$.

5 The spinor groups were introduced by É. Cartan in 1913. After 1925, they took on a fundamental role in quantum mechanics.

3.3.4. *Sections*

THEOREM 3.19.– *1) A* section *of a fibration* $\lambda = (M, B, \pi)$ *(or* section of the fiber bundle M*) is defined as a section of* π *([P1], section 1.1.1(**III**)), i.e. a mapping* $\mathrm{s} : B \to M$ *such that* $\pi \circ \mathrm{s} = 1_B$*. A section is therefore a lifting of* 1_B *(Definition 3.6). The sections of class* C^r *(or "morphic sections") of a fibration are defined clearly.*

2) A morphic section s *of a fibration* $\lambda = (M, B, \pi)$ *is an embedding from* B *into* M *whose image is closed in* M*. If* λ *is a covering, then* s *is a diffeomorphism onto a submanifold that is both open and closed in* M*.*

3) If f *is a* B*-morphism from* (M, B, π) *into* (M', B, π') *(Lemma-Definition 3.7(iv)), then the mapping* $f \circ \mathrm{s} : B \to M'$ *is a morphism of manifolds for any morphic section* s*.*

4) Let B' *be a submanifold of* B *and let* $\lambda\,|_{B'} = \left(\pi^{-1}(B'), B', \pi'\right)$ *be the fibration induced by* λ *on* B' *(Lemma-Definition 3.4(4)). The sections of* $\lambda\,|_{B'}$ *are the restrictions* $\mathrm{s}\,|_{B'}$ *of the sections* s *of* λ*, viewing* $\mathrm{s}\,|_{B'}$ *as mappings* $B' \to \pi^{-1}(B')$*.*

5) More generally, let $\lambda = (M, B, \pi)$ *be a fibration,* B' *a manifold and* $f^0 : B' \to B$ *a morphism. Consider the preimage fibration* $f^{0*}(\lambda) = (M', B', \pi')$ *(Lemma-Definition 3.11) and the canonical morphism* $\left(f^0, f\right) : f^{0*}(\lambda) \to \lambda$*. For every section,* $\mathrm{s} : M \to B$ *of* λ*,*

$$\mathrm{s}' : B' \ni b' \mapsto \left(b', \mathrm{s}\left(f^0(b')\right)\right)$$

is a section of $f^{0*}(\lambda)$*, called the* preimage *of* s *under* f^0 *and written as* $f^{0*}(\mathrm{s})$*. If* s *is morphic, so is* $f^{0*}(\mathrm{s})$*. The sections of* $f^{0*}(\lambda)$ *are the mappings* $\mathrm{s}' : B' \to M'$ $(M' = B' \times_B M')$ *such that* $\pi' \circ \mathrm{s}' = 1_{B'}$ *and the following diagram commutes:*

$$\begin{array}{ccccccc} B' & = & B' & \xrightarrow{f^0} & B & = & B \\ & \searrow \mathrm{s}' & \uparrow \pi' & & \uparrow \pi & \mathrm{s} \swarrow & \\ & & M' & \xrightarrow{f} & M & & \end{array} \qquad [3.8]$$

where $f^0(b') = \pi(x)$ *for all* $(b', x) \in B' \times M$*.*

PROOF.– (2): Since this is a local question, we can reduce to the case where λ is a trivial fibration. The remainder of the argument is an **exercise*** (see [DIE 93], Volume 3, (16.12.7)). (3)–(5): Immediate. ■

REMARK 3.20.– *1) If* $M \ni x = \mathrm{s}(b)$*, then* $\pi(x) = b$*, i.e.* $x \in \pi^{-1}(\{b\}) = M_b$*. Thus,* $\mathrm{s}(b) \in M_b$*, and every mapping* $\mathrm{s} : B \to M$ *satisfying this condition is a section of* (M, B, π)*.*

*2) The condition **(T)** from Lemma-Definition 3.4(1) implies that, for every* $b \in B$*, there exists an open subset* U *of* b *such that* $\lambda\,|_U$ *is* trivializable*, i.e. diffeomorphic to a trivial fibration.*

Consider the situation of Lemma-Definition 3.19. Any open subset U of B is a submanifold of B.

COROLLARY-DEFINITION 3.21.– *The mapping* $U \mapsto \pi^{-1}(U)$ *is a sheaf on* B *taking values in the category of manifolds ([P2], section 5.2.2(**III**)), namely the sheaf of sections of the fiber bundle* (M, B, π). *Write* $\Gamma^{(k)}(U, M)$ *for the set of sections of class* C^k $(k \leq r)$ *and* $\Gamma(U, M)$ *for the set of morphic sections over an open subset* U *of* B.

3.4. Vector bundles

3.4.1. *Vector bundles*

DEFINITION 3.22.– *i) Let* $\lambda = (M, B, \pi)$ *be a fibration with the property that the fiber* $M_b = \pi^{-1}(\{b\})$ *has a Banach space structure for every* $b \in B$ *(see Example 3.5(a)). We say that* M*, equipped with the structure defined by the fibration* λ *and the Banach spaces* M_b*, is a* vector bundle *if it satisfies the condition* ***(V)*** *stated below:*

(V) *For every* $b_0 \in B$*, there exist an open neighborhood* U *of* b_0 *in* B*, a Banach space* $\mathbf{E}$ *(which may depend on* b_0*) and a* trivializing isomorphism τ_U

$$\tau_U^{-1} = \psi : U \times \mathbf{E} \tilde{\rightarrow} \pi^{-1}(U)$$

such that $\pi(\psi(b, \mathbf{t})) = b$ *for every* $(b, \mathbf{t}) \in U \times \mathbf{E}$*,* and such that, *for every* $b \in U$*, the partial mapping* $t_b := \psi(b, .)$ *is an isomorphism of Banach spaces from* M_b *onto* $\mathbf{E}$*. We say that* B *is the* base *of the vector bundle* M*.*

ii) We say that $(U, \tau_U, \mathbf{E})$ *is a* vector chart *of* M *at* b_0*. The restriction* τ_{Ub_0} *of* τ_U *to* $\pi^{-1}(\{b_0\})$ *is an isomorphism* $\pi^{-1}(\{b_0\}) \tilde{\rightarrow} \{b_0\} \times \mathbf{E}$ *onto the fiber* $\pi^{-1}(\{b_0\})$.

iii) The rank *of* M b_0*, written as* $rk_{b_0}(M)$*, is defined as the (finite or infinite) dimension of the Banach space* M_{b_0}.

The mapping $b \mapsto \mathrm{rk}_b(M)$ is constant on each connected component of B (**exercise**). If this mapping is constant, $\mathrm{rk}_b(M)$ is said to be the *rank* of M. We say that M is *locally of finite rank* if $\mathrm{rk}_b(M) < \infty$ for every $b \in M$.

LEMMA-DEFINITION 3.23.– *Suppose that* M *is* locally of finite rank. *The mapping* $\psi = \tau_U^{-1}$ *defined above is said to be a* frame mapping *over* U*, and the condition* ***(V)*** *is equivalent to the following condition* ***(V')****:*

(V') *For every* $b \in B$*, there exist an open neighborhood* U *of* b *in* B*, an integer* m *(which may depend on* b*), and* m *morphic sections* $\mathbf{s}_i : U \rightarrow M$ *such that* $\pi \circ \mathbf{s}_i = 1_U$ *for all* $i \in \{1, ..., m\}$ *and the mapping*

$$\tau_U^{-1} : \left(b, a^1, ..., a^m\right) \longmapsto a^1 \mathbf{s}_1(b) + ... + a^m \mathbf{s}_m(b)$$

is a diffeomorphism from $U \times \mathbb{K}^m$ *onto* $\pi^{-1}(U)$.

We say that the morphic sections $\mathbf{s}_i$ $(1 \leq i \leq m)$ *form a* frame *of the vector bundle* λ *over* U *(we sometimes also say that a base of* M_b *over* $\mathbb{K}$ *is a* frame *of* λ *at the point* b *of* B*). The* m *morphic sections* $\mathbf{s}_i$ $(1 \leq i \leq m)$ *of* λ *on* U *form a frame if and only if the* $\mathbf{s}_i(b)$ *are linearly independent over* $\mathbb{K}$ *for every* $b \in U$.

REMARK 3.24.– *Suppose that* $M = T(B)$*, where* B *is locally finite-dimensional. Let* $c = (U, \xi, m)$ *be a chart of* B*, where* $\xi = (\xi^1, ..., \xi^m)$ *is the system of local coordinates associated with* c*. Let* $\boldsymbol{o}_M$ *be the projection mapping from* $T(B)$ *onto* B *given by* $\mathbf{h}_b \mapsto b$*. Then, the fiber bundle* $\boldsymbol{o}_M : T(B) \twoheadrightarrow B$ *is the tangent bundle and, for every chart* c *specified above, the mapping*

$$\psi_c : U \times \mathbb{R}^n \ni (b, \mathbf{h}) \mapsto (d_b\xi)^{-1} .\mathbf{h} \in \boldsymbol{o}_M^{-1}(U)$$

is a diffeomorphism. This also holds when

$$\mathbf{s}_i = X_i, \quad X_i(b) := (d_b\xi)^{-1} .\mathbf{e}_i, \quad \pi = \boldsymbol{o}_M.$$

A section X *of the fiber bundle* $\boldsymbol{o}_M$ *over a subset* A *of* B *is said to be a* vector field *over* A.

LEMMA-DEFINITION 3.25.– *1) Suppose that* U *and* $\tau = \tau_U$ *satisfy Definition 3.22 and let* $(V, \xi, \mathbf{F})$ *be a chart of* B *at the point* b_0 *such that* $V \subset U$*. Then:*

$$\left(\pi^{-1}(V), (\xi \times 1_{\mathbf{E}}) \circ \tau, \mathbf{F} \times \mathbf{E}\right)$$

is a vector chart of M *at every point of* $\pi^{-1}(\{b_0\})$*, called a* fibered chart.

2) There exists a unique manifold structure on M *for which these fibered charts are also charts of the manifold* M.

EXAMPLE 3.26.– *a) If* B *is a manifold and* $\mathbf{E}$ *is a Banach space, then* $pr_1 : B \times \mathbf{E} \twoheadrightarrow B$ *is a vector bundle* $\mathbf{E}_B$ *whose fibers are canonically isomorphic to* $\mathbf{E}$; *we say that this fiber bundle is* trivial. *More generally, a vector bundle* $\pi : M \twoheadrightarrow B$ *is said to be* trivializable *if there exists a* B*-isomorphism (Lemma-Definition 3.7(iv)) from* M *onto a trivial bundle.*

b) Let B *be a pure* n*-dimensional manifold,* $T(B)$ *its tangent bundle, and* π *its projection (section 3.2.1). Then,* $T(B)$*, equipped with the fibration* $\lambda = (T(B), B, \pi)$*, is a vector bundle such that* $T(B)_b = \{b\} \times \mathbb{K}^n \cong T_b(B)$ *for every* $b \in B$.

c) Let B *be a pure* n*-dimensional manifold,* $T^{\vee}(B)$ *its cotangent bundle, and* $\pi^{\vee}$ *its projection (section 3.2.2). Then,* $T^{\vee}(B)$*, equipped with the fibration* $\lambda^{\vee} =$

$(T^\vee(B), B, \pi^\vee)$*, is a vector bundle such that* $T^\vee(B)_b = \{b\} \times (\mathbb{K}^n)^\vee \cong T_b^\vee(B)$ *for every* $b \in B$*.*

d) Let $\lambda = (M, B, \pi)$ *be a vector bundle,* B' *a submanifold (e.g. an open set) of* B*, and* $\lambda|_{B'}$ *the fibration induced by* λ *on* B' *(Lemma-Definition 3.4(5)). Then,* $\lambda|_{B'}$ *is a vector bundle, said to be* induced *by* λ *on* B'*. The sections of this vector bundle are described in Lemma-Definition 3.19(4).*

e) If B *is modeled on Hilbert spaces (respectively separable Hilbert spaces) and the fibers* M_b *are Hilbert spaces (respectively separable Hilbert spaces), then the manifold* M *is modeled on Hilbert spaces (respectively separable Hilbert spaces).*

REMARK 3.27.– *The sections (respectively the morphic sections) of the trivial bundle* $\mathbf{E}_B$ *are the mappings* $b \mapsto (b, f^0(b))$*, where* f^0 *is a mapping (respectively a mapping of class* C^r*) from* B *into* $\mathbf{E}$ *(see Remark 3.20(1)).*

DEFINITION 3.28.– *A manifold* B *is said to be* parallelizable *if the tangent bundle* $\boldsymbol{o}_M : T(B) \twoheadrightarrow B$ *is trivial.*

Kuiper showed the following result in 1965 [KUI 65] (this result can be deduced from Theorem 2.49 but historically preceded it):

THEOREM 3.29.– (Kuiper) *Every vector bundle* $\pi : M \twoheadrightarrow B$ *such that* B *is a separable paracompact manifold and that the* M_b *fibers are infinite-dimensional separable Hilbert spaces is trivial. Hence, every locally infinite-dimensional, paracompact and separable Hilbert manifold is parallelizable.*

DEFINITION 3.30.– *Let* $\lambda = (M, B, \pi)$ *and* $\lambda' = (M', B', \pi')$ *be two vector bundles.*

i) A morphism of vector bundles *is a morphism* (f^0, f) *from* λ *into* λ' *such that, for all* $b \in B$*, the restriction* f_b *of* f *to* X_b *belongs to* $\mathcal{L}\left(M_b; M'_{f^0(b)}\right)$*. The composition of two morphisms of vector bundles is a morphism of vector bundles, so the vector bundles form a concrete* category *with base* $\mathbf{Set}$*, which gives us the notion of vector bundle* structure*. When* $B = B'$ *and* $f^0 = 1_B$*, we also say that* f *is a* B*-morphism of vector bundles, writing Hom* (M, M') *for the set of* B*-morphisms from* M *into* M'*.*

ii) The isomorphisms *and* B-isomorphisms *of vector bundles are defined similarly.*

LEMMA 3.31.– *The set Hom* (M, M') *can be canonically equipped with a vector bundle structure with base* B *associated with a fibration* $\nu = (Hom(M, M'), B, \sigma)$ *such that* $\sigma^{-1}(\{b\}) = \mathcal{L}\left(M_b; M'_{f^0(b)}\right)$ *for every* $b \in B$*.*

PROOF.– See Lemma 3.8. ■

COROLLARY 3.32.– *i) Let* (M, B, π) *and* (M', B, π') *be two vector bundles with the same base* B *and* $u', u'' \in Hom(M, M')$ *(Definition 3.30(i)). For every* $b \in B$*, the restriction* u'_b *of* u' *to the fiber* M_b *belongs to* $\mathcal{L}(M_b; M'_b)$*. Let* $u' + u''$ *be the mapping from* M *into* M' *defined by* $(u' + u'')_b = u'_b + u''_b$*. Then,* $u' + u'' \in Hom(M, M')$*, and this set may therefore be equipped with an abelian group structure.*

ii) Let $\mathcal{C}^r(B)$ *be the ring of functions of class* C^r *from* B *into* $\mathbb{K}$*. For every function* $f \in \mathcal{C}^r(B)$ *and every* $u \in Hom(M, M')$*, define the mapping* $f.u$ *from* M *into* M' *by the relation* $(f.u)_b = f(b)\, u_b$*. Then,* $f.u \in Hom(M, M')$ *and the set* $Hom(M, M')$ *may therefore be equipped with the structure of a* $\mathcal{C}^r(B)$*-module.*

Recall that a B-morphism $u : M \to M'$ is a morphism of manifolds for which the diagram [3.6] commutes (section 3.3.1).

COROLLARY 3.33.– *i) Let* $\mathcal{S}_M(B) = \Gamma(B, M)$ *be the set of morphic sections of* M*, i.e. the set of mappings* $\mathbf{s} : B \to M$ *such that*

a) $\mathbf{s}$ *is of class* C^r*,*

b) for every $b \in B$*,* $\mathbf{s}(b) \in M_b$*.*

The set $\mathcal{S}_M(B)$ *can be viewed as the set of* B*-morphisms from the trivial bundle* $(B \times \{0\}, B, pr_1)$ *into* M*. Therefore, by Corollary 3.32,* $\mathcal{S}_M(B)$ *is a* $\mathcal{C}^r(B)$*-module.*

Furthermore, the three following conditions are equivalent*:*

a') M is locally of finite rank and trivializable (Example 3.26(a));

b') there exists a frame $(\mathbf{s}_i)_{1 \leq i \leq m}$ *of* M *over* B *(Lemma-Definition 3.23);*

c) $(\mathbf{s}_i)_{1 \leq i \leq m}$ *is a basis of the* $\mathcal{C}^r(B)$*-module* $\mathcal{S}_M(B)$ *(which is therefore free and finitely generated).*

ii) In particular, let B' *be a submanifold (e.g. an open subset) of* B *and* $\left(\pi^{-1}(B'), B', \pi|_{B'}\right)$ *the vector bundle induced by* (M, B, π) *on* B'*. Then,* $\mathcal{S}_M(B')$*, i.e. the set of morphic sections of* M *over* B'*, has the structure of a* $\mathcal{C}^r(B')$*-module.*

REMARK 3.34.– *It follows from the above that the mapping* $U \mapsto \mathcal{C}^r(U)$ *(where each* U *is an open subset of* B*) is a sheaf of rings, written as* $\mathcal{C}^r$*, and that the mapping* $\mathcal{S}_M : U \mapsto \mathcal{S}_M(U)$ *is a sheaf of* $\mathcal{C}^r$*-Modules ([P2], section 5.3.1). If* $V \subset U$*, the restriction* $\rho_V^U : \mathcal{C}^r(U) \to \mathcal{C}^r(V)$ *(respectively* $r_V^U : \mathcal{S}_M(U) \to \mathcal{S}_M(V)$*) is surjective, so the sheaves* $\mathcal{C}^r$ *and* $\mathcal{S}_M$ *are flabby ([P2], section 5.4.1).*

REMARK 3.35.– *Modifying Definition 3.22 in an obvious way gives the notion of a vector bundle $\pi : M \twoheadrightarrow B$ of class $\mathfrak{c}^{\infty}$. If B is a metrizable manifold modeled on a space $\mathbf{E}$ of type $(\mathcal{FN})$ (respectively $(\mathcal{SN})$) and $\mathbf{F}$ is a space of type $(\mathcal{FN})$ (respectively $(\mathcal{SN})$), then the trivial bundle $B \times \mathbf{F} \twoheadrightarrow B$ is a $\mathfrak{c}^{\infty}$-paracompact manifold modeled on the space $\mathbf{E} \times \mathbf{F}$, and both this product space and this manifold are of type $(\mathcal{FN})$ (respectively $(\mathcal{SN})$) ([P2], sections 3.4.6(**III**), 3.8.2(**II**), and 3.11.3(**III**); Theorem 2.14(2) and section 1.4.2). In particular, whenever B is a manifold satisfying these conditions, the tangent bundle $T(B)$ is a $\mathfrak{c}^{\infty}$-paracompact manifold modeled on the space $\mathbf{E} \times \mathbf{E}$, and both this product space and this manifold are of type $(\mathcal{FN})$ (respectively $(\mathcal{SN})$).*

3.4.2. *Dual of a vector bundle*

DEFINITION 3.36.– *Let M be a vector bundle with base B and $\mathbb{K}_B$ the trivial bundle associated with the fibration $(B \times \mathbb{K}, B, pr_1)$ (Example 3.26(a)). Then, Hom $(M, \mathbb{K}_B)$ is said to be the dual of M and is written as $M^{\vee}$.*

The fibers of $M^{\vee}$ are therefore the $M_b^{\vee} := \mathcal{L}(M_b; \mathbb{K})$, $b \in B$. If $\mathbf{s}$ (respectively $\mathbf{t}$) is a section of M (respectively $M^{\vee}$) over an open subset U of B, then the mapping

$$b \mapsto (b, \langle \mathbf{t}(b), \mathbf{s}(b) \rangle)$$

is a section, written as $\langle \mathbf{t}, \mathbf{s} \rangle$, of the trivial bundle $\mathbb{K}_U$.

For a morphism of fiber bundles $u : E \to F$, we can define the transpose ${}^t u : F^{\vee} \to E^{\vee}$ by

$$\left\langle {}^t u_U \circ \mathbf{s}^{\vee}, \mathbf{r} \right\rangle = \left\langle \mathbf{s}^{\vee}, u_U \circ \mathbf{r} \right\rangle$$

for any two sections $\mathbf{r}$ of E and $\mathbf{s}^{\vee}$ of $E^{\vee}$ over the open set $U \subset B$.

EXAMPLE 3.37.– *Consider a manifold B, its tangent bundle $T(B)$ and its cotangent bundle $T^{\vee}(B)$. We know that $T(B)$ is the vector bundle associated with the fibration $\lambda = (T(B), B, \pi)$ whose fibers are the $T(B)_b \cong T_b(B)$, $b \in B$; for every $b \in B$, there exists a Banach space $\mathbf{F}$ such that $T(B)_b = \{b\} \times \mathbf{F}$. The cotangent bundle $T^{\vee}(B)$ is the vector bundle associated with the fibration $\lambda' = (T^{\vee}(B), B, \pi')$ whose fibers are the $T^{\vee}(B)_b \cong T_b^{\vee}(B)$, $b \in B$; for every $b \in B$, $T(B)_b = \{b\} \times \mathbf{F}^{\vee}$. Hence, $T^{\vee}(B)$* is the dual of the tangent bundle $T(B)$.

LEMMA-DEFINITION 3.38.– *Suppose that M is locally of finite rank.*

i) Let $(\mathbf{s}_i)_{1 \leq i \leq m}$ be a frame of the fiber bundle M over an open subset U of B (Lemma-Definition 3.22(2)); there exists a unique frame $\left(\mathbf{s}^{\vee i}\right)_{1 \leq i \leq m}$ of $M^{\vee}$ over U such that $\left\langle \mathbf{s}^{\vee i}, \mathbf{s}_j \right\rangle = \delta_j^i$, and $\left(\mathbf{s}^{\vee i}\right)$ is said to be the dual coframe *of the frame $(\mathbf{s}_i)$.*

ii) In particular, let $M = T(B)$, $c = (U, \xi, m)$ be a chart of B, and consider the system of local coordinates ξ^i $(1 \leq i \leq m)$ associated with c (Remark 3.24). Then, the $\mathbf{s}_i = \frac{\partial}{\partial \xi^i}$ $(1 \leq i \leq m)$ form the natural frame *over U (relative to c) and the $\mathbf{s}^{\vee i} = d\xi^i$ $(1 \leq i \leq m)$ form the* natural coframe *(Theorem 2.71).*

REMARK 3.39.– *The cotangent bundle of an $(\mathcal{FN})$ or $(\mathcal{SN})$ manifold can be defined as the dual of its tangent bundle (Lemma-Definition 3.1). There are two possible cotangent bundles: the kinematic cotangent bundle $T^{\vee}(M)$ and the operational cotangent bundle $D^{\vee}(M)$ (Remark 3.2). Since $\mathfrak{c}^{\infty}(\mathbf{E} \times \mathbf{E}^{\vee})$ is not a topological vector space (Remark 1.50), none of these cotangent bundles have a notion of fibered chart (see Lemma-Definition 3.3(3)) and hence they do not have a manifold structure. This is one reason why the "convenient" framework complements but does not replace the Banach framework according to the perspective adopted here.*

3.4.3. *Subbundles and quotient bundles*

LEMMA-DEFINITION 3.40.– *Let $\pi : M \twoheadrightarrow B$ be a vector bundle.*

1) We say that $M' \subset M$ is a vector subbundle *of M if, for every $b \in B$, there exist a vector chart $(U, \varphi, \mathbf{E})$ of M at b (Definition 3.22) and a subspace $\mathbf{F}$ that splits in $\mathbf{E}$ ([P2], section 3.2.2(**IV**)) such that*

$$\varphi^{-1}\left(\pi^{-1}(U) \cap M'\right) = U \times \mathbf{F}.$$

2) If so, there exists a unique vector bundle structure on M' that makes the canonical injection $M' \hookrightarrow M$ a morphism. In this case, for each $b \in B$, $M'_b = M' \cap M_b$, M'_b is a Banach subspace of M_b and M' is a closed submanifold of M.

3) If M is locally of finite rank, then, for every $b \in B$, there exist an open neighborhood U of b in B, a frame $(\mathbf{s}_1, ..., \mathbf{s}_n)$ of M over U, and an integer $m \leq n$ such that, for every $y \in U$, the vectors $\mathbf{s}_1(y), ..., \mathbf{s}_m(y)$ form a basis of M'_y.

LEMMA-DEFINITION 3.41.– *Let M' be a vector subbundle of M and, for every $b \in B$, set $M''_b = M_b / M'_b$ (this is a quotient of Banach spaces and hence a Banach space). Let*

$$M'' = \dot{\bigcup_{b \in B}} M''_b.$$

Let $\pi'' : M'' \twoheadrightarrow B : M''_b \mapsto b$, and let $p : M \to M''$ be the mapping such that, for every $b \in B$, $p|_{M_b}$ is the canonical surjection $M_b \twoheadrightarrow M''_b$. Then, there exists a *unique vector bundle structure on M'' with base B and projection π'' that makes p a morphism of vector bundles. The fiber bundle thus defined is said to be the* quotient *of M by M', written as M/M'. The canonical morphism $p : M \to M/M'$ is a submersion.*

3.4.4. *Whitney sum and tensor product*

Let M', M'' be vector bundles with base B and projections π', π'', respectively. *If M' and M'' are locally of finite rank*, the *tensor product $M' \otimes M''$* is defined by

$$M' \otimes M'' = \dot{\bigcup_{b \in B}} M'_b \otimes M''_b,$$

where $M'_b \otimes M''_b$ is the tensor product of M'_b and M''_b ([P1], section 3.1.5(**I**)). Let U be a non-empty open subset of B and let s', s'' be sections of M', M'' over U. Write $\mathrm{s}' \otimes \mathrm{s}''$ for the mapping $b \mapsto \mathrm{s}'(b) \otimes \mathrm{s}''(b)$ from U into $M' \otimes M''$.

We can proceed in the same way for direct sums instead of tensor products (without the hypothesis of finite rank on M' or M''). With these conditions:

LEMMA-DEFINITION 3.42.– *1) There exists a unique vector bundle structure on $M' \otimes M''$ (respectively $M' \oplus M''$) with base B and projection μ (respectively σ) satisfying the following condition:*

For every open subset U of B and every pair of morphic sections s', s'' of M', M'' over U, $\mathrm{s}' \otimes \mathrm{s}''$ (respectively $\mathrm{s}' \oplus \mathrm{s}''$) is a morphic section of $M' \otimes M''$ (respectively $M' \oplus M''$) over U.

We say that $M' \otimes M''$ (respectively $M' \oplus M''$) is the tensor product *(respectively the* Whitney sum*) of the vector bundles M', M''.*

2) The fibration $(M' \oplus M'', B, \sigma)$ is isomorphic to the fiber product $(M' \times_B M'', B, \pi' \times_B \pi'')$ (Lemma-Definition 3.9).

Let $u' : E' \to F'$ and $u'' : E'' \to F''$ be two morphisms. We define $u' \otimes u'' : E' \otimes E'' \to F' \otimes F''$ as follows:

$$(u' \otimes u'')|_U (\mathrm{s}' \otimes \mathrm{s}'') = (u'|_U \circ \mathrm{s}') \otimes (u''|_U \circ \mathrm{s}'')$$

for all sections s' of E' and s'' of E'' over some arbitrary open subset U of B.

REMARK 3.43.– *The sheaf of sections of M/M' (respectively $M' \oplus M''$, respectively $M' \otimes M''$) is the quotient (respectively direct sum, respectively tensor product) of the sheaves of sections of M' and M'' ([P2], section 5.3).*

If B is a real manifold and M is a real vector bundle of finite rank and base B, the *complexification* of M is defined as the tensor product $M_{(\mathbb{C})} := M \otimes \mathbb{C}_B$, where $\mathbb{C}_B = B \times \mathbb{C}$ is the trivial bundle of base B whose fibers are equal to $\mathbb{C}$ as an $\mathbb{R}$-vector space.

3.4.5. *The category of vector bundles*

(I) THE ADDITIVE CATEGORY OF VECTOR BUNDLES The tangent bundle functor and the cotangent bundle functor were introduced earlier (section 3.2.3). These functors act on the category of manifolds, but we still need to specify their "codomains" ([P1], section 1.2.1), which leads us to define the category of vector bundles. The objects of this category are of course the vector bundles and the morphisms are the morphisms of fiber bundles (Definition 3.30). Let us restate this idea below:

DEFINITION 3.44.– *(A) Let $\lambda = (M, B, \pi)$ and $\lambda' = (M', B', \pi')$ be two fibrations such that M and M' are vector bundles. A morphism of vector bundles $\lambda \to \lambda'$ is a pair of morphisms of manifolds*

$$f^0 : B \to B', \quad f : M \to M'$$

satisfying the following conditions:

1) The diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & M' \\ \pi \downarrow & & \downarrow \pi' \\ B & \xrightarrow{f^0} & B' \end{array}$$

commutes, and the induced mapping

$$f_b : M_b \to M'_{f^0(b)}$$

belongs to $\mathcal{L}\left(M_b; M'_{f^0(b)}\right)$ for each $b \in B$.

2) For every $b_0 \in B$, there exist trivializing mappings (Definition 3.22)

$$\tau : \pi^{-1}(U) \to U \times \mathbf{E}$$

$$\varsigma : \pi'^{-1}(V) \to V \times \mathbf{F}$$

at b_0 and $f^0(b_0)$, respectively, such that $f(U) \subset V$ and the mapping $\mathcal{L}(\mathbf{E}; \mathbf{F})_V$ from U into the trivial bundle defined by

$$b \mapsto \varsigma_{f^0(b)} \circ f_b \circ \tau_b^{-1}$$

is of class C^r.

(B) Suppose that the bases B and B' are equal. Then (Lemma-Definition 3.7(iv)), $f : M \to M'$ is a B-morphism if:

1') The diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & M' \\ \pi\downarrow & & \downarrow\pi' \\ B & \xrightarrow{1_B} & B \end{array}$$

commutes, and the induced mapping

$$f_b : M_b \to M'_b$$

belongs to $\mathcal{L}(M_b; M'_b)$ for each $b \in B$.

2') For every $b_0 \in B$, there exist trivializing mappings

$$\tau : \pi^{-1}(U) \to U \times \mathbf{E}$$
$$\varsigma : \pi'^{-1}(U) \to U \times \mathbf{F}$$

at b_0 such that the mapping $\mathcal{L}(\mathbf{E}; \mathbf{F})_U$ from U into the trivial bundle defined by

$$b \mapsto \varsigma_b \circ f_b \circ \tau_b^{-1}$$

is of class C^r.

The mapping defined in (2) can be decomposed as follows:

$$\{b\} \times \mathbf{E} \xrightarrow{\tau_b^{-1}} \pi^{-1}(\{b\}) \subset M_b \xrightarrow{f_b} M'_{f^0(b)} \xrightarrow{\varsigma_{f^0(b)}} \{f^0(b)\} \times \mathbf{F}.$$

By abuse of notation, a morphism of vector bundles such as the one stated above is written as $f : \pi \to \pi'$. We have the following result:

PROPOSITION 3.45.– *Let π, π' be two vector bundles on the manifolds B, B' respectively. Let $f^0 : B \to B'$ be a morphism of manifolds and suppose that, for every $b \in B$, there exists a continuous linear mapping*

$$f_b : \pi_b \to \pi'_{f^0(b)}$$

such that condition (2) in Definition 3.44 is satisfied for every b_0. Then, the mapping $f : \pi \to \pi'$ defined by f_b on each fiber is a canonical morphism of vector bundles.

PROOF.– We return to the case where π and π' are the trivial bundles $U \times \mathbf{E}$ and $V \times \mathbf{F}$, respectively, so that the trivializing mappings are equal to the identity function. Then, f is given by

$$(b, v) \mapsto \left(f^0(b), f_b(v)\right). \qquad \blacksquare$$

Let $\mathbf{VB}$ be the category of vector bundles with base B.

1) Each set $\mathrm{Mor}_{\mathbf{VB}}(E, F)$ is an abelian group (Corollary 3.32). The composition of morphisms of $\mathbf{VB}$ (i.e. B-morphisms) is clearly bilinear, so the category $\mathbf{VB}$ is *preadditive* ([P1], section 3.3.7(**I**)).

2) The trivial bundle $B \times \{0\}$ is a 0 element of the category $\mathbf{VB}$. Furthermore, we have already defined the biproduct $E \oplus F$ of two fibers E and F : their Whitney sum (Lemma-Definition 3.42). Hence, the category $\mathbf{VB}$ is *additive* ([P1], section 3.3.7(**II**)).

3) Let $u \in \mathrm{Mor}_{\mathbf{VB}}(E, F)$. For every $b \in B$, write $u_b : E_b \to F_b$ for the restriction of u to E_b (so $u_b \in \mathcal{L}(E_b; F_b)$). Let $N_b = \ker u_b$, $I_b = \mathrm{im} u_b$ and

$$N = \dot{\bigcup_{b \in B}} N_b, \quad I = \dot{\bigcup_{b \in B}} I_b.$$

LEMMA-DEFINITION 3.46.– *1) The following conditions are equivalent ([BOU 82a], (7.5.5)):*

i) u is a subimmersion;

ii) N is a vector subbundle of E and I is a vector subbundle of F.

2) The morphism u is said to be locally direct *if one of the equivalent conditions stated above is satisfied. If so, N and I are called the* kernel *and the* image *of the morphism u, respectively, written as* $\ker(u)$ *and im* (u).

3) If u is locally direct, then taking quotients determines an isomorphism $\nu : E/N \tilde{\to} I$.

4) Suppose that E is of finite rank. *Then, the mapping $b \mapsto rku_b$ is lower semi-continuous and the equivalent conditions in (1) are also equivalent to the following condition:*

iii) The function $b \mapsto rku_b$ is locally constant (i.e. u is a subimmersion, by Theorem 2.55).

We say that a sequence

$$M \xrightarrow{f} M' \xrightarrow{g} M''$$

is *exact locally direct* ([BOU 82a], (7.5.6)) if it is exact and the morphisms of vector bundles f and g are locally direct.

In Lemma-Definition 3.46, we defined the image and the kernel of a locally direct morphism $u : M \to M'$; any such morphism also has a coimage $\operatorname{coim}(u) = M/\ker(u)$ and a cokernel $\operatorname{coker}(u) = M'/\operatorname{im}(u)$ ([P1], section 3.3.7**(III)**); Lemma-Definition 3.46(3) shows that $\operatorname{coim}(u)$ and $\operatorname{im}(u)$ are isomorphic and hence that locally direct morphisms are analogous to *strict morphisms* in a preabelian category (*ibid.*)[6]. Note that the composition of two locally direct morphisms is not always locally direct (Remark 2.56); see Corollary 3.49. A sequence of vector bundles

$$0 \to M' \overset{\iota}{\to} M \overset{\varphi}{\to} M'' \to 0$$

is exact locally direct if and only if ι is an isomorphism from M' onto the vector subbundle $\iota(M')$ of M (which enables us to identify M' with $\iota(M') \subset M$) and the diagram

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M'' \\ \eta \downarrow & \nearrow \sigma & \\ \operatorname{coker}(\iota) = M/\ker(\varphi) & & \end{array}$$

commutes, where $\eta : M \twoheadrightarrow M/\ker(\varphi)$ is the canonical surjection and σ is an isomorphism of vector bundles.

(II) SPLIT VECTOR SUBBUNDLES Let $\pi : B \to M$ be a vector bundle and M' a vector subbundle of M. There exists an exact locally direct sequence

$$0 \to M' \overset{\iota}{\to} M,$$

where ι is the inclusion mapping. By definition, this means that, for every $b \in B$, there exist an open neighborhood U_b of b and a morphism of vector bundles φ_b such that the following sequence splits:

$$0 \to M'|U_b \overset{\iota_b}{\to} M|U_b \overset{\varphi_b}{\to} M''|U_b \to 0, \qquad [3.9]$$

where $M'|U_b$ (respectively $M|U_b$) is the vector fiber induced by M' (respectively M) on U_b, in the sense that there exists an isomorphism of vector bundles

$$\iota_b(M'|U_b) \oplus M''|U_b = M|U_b\ . \qquad [3.10]$$

6 Nevertheless, the kernel and cokernel are not defined for *arbitrary* morphisms, so $\mathbf{VB}$ is *not* a preabelian category, much less an abelian category.

By ([P1], section 3.1.4), this is equivalent to saying that ι_b is a monomorphism that splits, i.e. a monomorphism that admits a retraction $r_b : M\,|U_b \twoheadrightarrow M'\,|U_b$. This means that $r_b \circ \iota_b = 1_{M'|U_b}$, and thus $M''\,|U_b := \ker(r_b)$ satisfies [3.10].

THEOREM 3.47.– *Suppose that the manifold B is differential and paracompact (or equivalently metrizable, by Corollary 2.5). If the Banach spaces modeling B (Definition 2.1) are all C^∞-paracompact, for example Hilbert spaces (Theorem 1.63), then there exists a fiber subbundle M'' of M such that $M = M' \oplus M''$; in other words, M'* splits *in M.*

PROOF.– By Theorem 2.14(1), there exists a locally finite open covering $\left(U_{b_j}\right)_{j\in J}$ of B and a C^∞ partition of unity $(\psi_j)_{j\in J}$ subordinate to $\left(U_{b_j}\right)_{j\in J}$ for which the split exact sequence [3.9] holds with b replaced by b_j. Let r_j be a retraction associated with ι_{b_j} $(j \in J)$ and define $r := \sum_{j\in J} \psi_j.r_j$. Then, $r : M \to M'$ is a morphism of vector bundles and

$$r \circ \iota = \sum_{j\in J} \psi_j.r_j \circ \iota_{b_j} = 1_{M'}.$$

Furthermore, if we write $M'' := \ker(r)$, then $M''\left|U_{b_j}\right.$ satisfies [3.10] with b replaced by b_j; thus, M'' is a vector subbundle of M such that $M = M' \oplus M''$. ■

(III) TANGENT BUNDLE AND COTANGENT BUNDLE FUNCTORS Let us return to the tangent bundle and cotangent bundle functors (section 3.2.3). The tangent bundle functor T is defined as follows:

$$\begin{aligned} B &\xrightarrow{T} T(B) \\ \left(f^0 : B \longrightarrow B'\right) &\overset{T}{\mapsto} \left(Tf^0 : T(B) \longrightarrow T(B')\right) \\ Tf^0 : (b, \mathbf{h}_b) &\mapsto \left(f^0(b), T_b\left(f^0\right).\mathbf{h}_b\right) \end{aligned}$$

Thus, T sends any given morphism of manifolds $f^0 : B \to B'$ to a morphism of fiber bundles $f = Tf^0 : T(B) \to T(B')$.

Similarly, the cotangent bundle functor $T^\vee$ is defined as follows:

$$\begin{aligned} B &\xrightarrow{T^\vee} T^\vee(B) \\ \left(f^0 : B \longrightarrow B'\right) &\overset{T^\vee}{\mapsto} \left(Tf^0 : T^\vee(B') \longrightarrow T^\vee(B)\right) \\ T^\vee f^0 \;:\; T^\vee(B')_{f^0(b)} &\ni \left(y = f^0(b), \mathbf{h}_y^\vee \in T_b^\vee(B')\right) \\ &\mapsto \left(b, T_b^\vee\left(f^0\right).\mathbf{h}_y^\vee\right) \in T^\vee(B)_b \end{aligned}$$

Again, $T^{\vee}$ sends any given morphism of manifolds $f^0 : B \to B'$ to a morphism of fiber bundles $f' = T^{\vee} f^0 : T^{\vee}(B') \to T^{\vee}(B)$, which clearly demonstrates the contravariant nature of this functor.

3.4.6. *Preimage of a fiber bundle*

Let M be a vector bundle with base B and $\lambda = (M, B, \pi)$ the associated fibration. Let B' be a manifold and $f^0 : B' \to B$ a morphism of manifolds. The preimage fibration was defined earlier by $f^{0*}(\lambda) = (M', B', \pi')$, where $M' := B' \times_B M$ and the canonical morphism is $\left(f^0, f'\right) : f^{0*}(\lambda) \to \lambda$ (Lemma-Definition 3.11).

LEMMA-DEFINITION 3.48.– *1) Suppose that the restriction $f'_{b'} : M'_{b'} \to M_{f^0(b')}$ (which is a diffeomorphism) is linear for every $b' \in B'$. Then, $\left(f^0, f'\right)$ is a* morphism of vector bundles *(Definition 3.44(A)) and $f'^{-1}_{b'}$ transports the Banach space structure of $M_{f^0(b')}$ to $M'_{b'}$; thus, M' can be equipped with a vector bundle structure with base B', called the* preimage *of M under f^0 and written as $f^{0*}(M)$. The projection π' is $M_{b'} \mapsto b'$. The associated fibration is $f^{0*}(\lambda) = (M', B', \pi')$ (the vector bundle $f^{0*}(M)$ is also called the* vector bundle induced *by f^0).*

We say that rk$_{b'}(f'_b) \leq \infty$ is the vector rank *of $f'_{b'} : f^{0*}(M)_{b'} \tilde{\to} M_{f^0(b')}$.*

2) Let N' be a vector bundle with base B' and suppose that $g' : N' \to M$ is an f^0-morphism (Lemma-Definition 3.11(iii)). There exists a uniquely determined f^0-morphism u from N' into $f^{0*}(M)$ such that $g' = f' \circ u$ (see Lemma-Definition 3.11(iii)).

Consider the situation of Lemma-Definition 3.48(2) with $M = T(B)$, $N' = T(B') = f^{0*}(T(B))$. Then, $T\left(f^0\right) : N' \to M$ (this is the mapping denoted g' earlier), so there exists a unique f^0-morphism $\delta f^0 : T(B') \to T(B)$ (denoted u earlier) such that $T\left(f^0\right) = f' \circ \delta f^0$. With this notation, we have the following result:

COROLLARY 3.49.– *i) f^0 is an immersion (respectively a submersion) if and only if the sequence*

$$0 \to T(B') \overset{\delta f^0}{\to} T(B) \to 0$$

is exact locally direct at $T(B')$ (respectively $T(B)$). If so, the cokernel vector bundle coker $\left(\delta f^0\right)$ is said to be normal *(or* transversal*) of f^0 (respectively the kernel vector bundle* ker $\left(f^0\right)$ *is said to be* tangent to the fibers *of f^0 and is written as $T(B'/B)$).*

ii) If f^0 is a submersion, then the fiber $T(B'/B)_{b'}$ is the tangent space of the submanifold $\left(f^0\right)^{-1}\left(f^0(\{b'\})\right)$ at $b' \in B'$ and the sequence

$$0 \to T(B'/B)_{b'} \overset{\iota}{\to} T_{b'}(B') \overset{T_{b'}\left(f^0\right)}{\to} T_{f^0(b')}(B) \to 0$$

(where ι is the canonical injection) is exact.

iii) f^0 is étale (respectively a subimmersion) if and only if δf^0 is an isomorphism (respectively is locally direct).

If M is locally of finite rank and the sections $\mathrm{s}_1, ..., \mathrm{s}_n$ form a frame of M over $U \subset B$, then the preimage sections (Lemma-Definition 3.19(5)) $\mathrm{s}'_1, ..., \mathrm{s}'_n$ form a frame of M' over $\left(f^0\right)^{-1}(U)$ and the diagram [3.8] (section 3.3.4) commutes.

LEMMA 3.50.– *Let $f^0 : B' \to B$ be a morphism of manifolds and $T^\vee(B)$ the cotangent bundle of B. Consider its preimage $f^{0*}(T^\vee(B))$, with base B'. There exists a unique B'-morphism*

$$w : f^{0*}(T^\vee(B)) \to T^\vee(B')$$

such that, for $b' \in B'$, $w_{b'}$ is the continuous linear mapping formed by the composition

$$f^{0*}(T^\vee(B))_{b'} \xrightarrow{f_{b'}} T^\vee_{f^0(b')}(B) \xrightarrow{{}^tT_{b'}(f^0)} T^\vee_{b'}(B')$$

where $f_{b'}$ is the canonical morphism (Proposition 3.45). If $\Phi : \xi(U) \to \xi'(U')$ is the local expression of f^0, then the local expression of w is

$$(b', \mathbf{h}^\vee) \mapsto \left(b', {}^t D\Phi(b') . \mathbf{h}^\vee\right).$$

3.5. Principal bundles

3.5.1. *Notion of a principal bundle*

(I) GENERAL CASE

LEMMA-DEFINITION 3.51.– *Let P be a differential (respectively analytic) manifold and $\mathbf{G}$ a Lie group with a differentiable (respectively analytic) and free* right *action on P (section 2.4.2). Suppose that the manifold of orbits $\mathbf{G}\backslash P$ exists (see Lemma 2.91 for the finite-dimensional case)[7] and write $\pi : P \to \mathbf{G}\backslash P$ for the canonical surjection. Then, $(P, \mathbf{G}\backslash P, \pi)$ is a fibration (called a* principal fibration*) and P is said to be a* principal bundle *with* structural group $\mathbf{G}$. *In accordance with the general definitions (Lemma-Definition 3.4), P is the space of this principal bundle, $B = \mathbf{G}\backslash P$ is its base, and π is its projection. The principal bundle itself is written as $(P, B, \mathbf{G}, \pi)$.*

Every element $b \in \mathbf{G}\backslash P$ is an orbit $q.\mathbf{G}$ $(q \in P)$, and the fibers are the orbits; each fiber is a manifold diffeomorphic to $\mathbf{G}$ under the diffeomorphism $q.\mathbf{G} \tilde{\to} \mathbf{G} : q.g \mapsto g$ (Figure 3.5). Notationally, we express this by writing $\mathbf{G}_b$ instead of b for any such fiber. Thus, the action of $\mathbf{G}$ sends each fiber to itself, and its restriction to each fiber is simply transitive.

7 The set of orbits is denoted $P/\mathbf{G}$ for a left action of $\mathbf{G}$ on P, and $P\backslash\mathbf{G}$ for a right action.

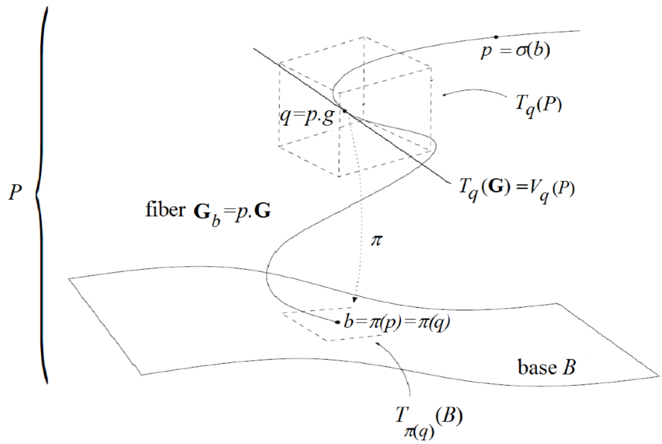


Figure 3.5. *Principal bundle*

EXAMPLE 3.52.– *The Riemann surface of the logarithm is a principal bundle with base $\mathbb{C}^\times$ and structural group $\mathbb{Z}$. Indeed, write Y for this surface and $(Y, \mathbb{C}^\times, \pi)$ for the associated fibration, which is a covering of $\mathbb{C}^\times$ (Example 3.15). Let $y = (z_1, z_2) \in \mathbb{C}^\times \times \mathbb{C}$; then $y \in Y$ if and only if $z_1 - e^{z_2} = 0$, and the fiber Y_{z_1} $(z_1 \in \mathbb{C}^\times)$ can be identified with $\{z_2 \in \mathbb{C} : e^{z_2} = z_1\}$. Consider the action of $\mathbb{Z}$ on Y defined by $(z_1, z_2).k = (z_1, z_2 + 2k\pi i)$. The right action of the group $\mathbb{Z}$ is free on Y, since the mapping $k \mapsto (z_1, z_2 + 2k\pi i)$ is injective. Let $\boldsymbol{R}$ be the relation defined by: $(z_1, z_2)\, \boldsymbol{R}\, (z_1', z_2')$ if there exists $k \in \mathbb{Z}$ such that $(z_1', z_2') = (z_1, z_2 + 2k\pi i)$. Let $z_1 = re^{i\theta}$, $r > 0$, $-\pi < \theta \leq \pi$. Then, $z_1 - e^{z_2} = 0$ if and only if there exists $k \in \mathbb{Z}$ such that $e^{\ln r + i(\theta + 2k\pi) - z_2} = 1$, i.e. $z_2 = \ln r + i(\theta + 2k\pi)$. Hence, the equivalence class of $y = (z_1, z_2)$ (mod.$\boldsymbol{R}$), i.e. its orbit, is fully determined by z_1, and the set of orbits $\mathbb{Z}\backslash Y$ can be identified with the manifold $\mathbb{C}^\times = pr_1(Y)$. The orbit of y is isomorphic to $\mathbb{Z}$. Thus, $(Y, \mathbb{Z}, \mathbb{Z}\backslash Y, \pi')$ is a principal bundle, where π' is the canonical surjection $Y \twoheadrightarrow \mathbb{Z}$.*

(II) TRIVIAL PRINCIPAL BUNDLE

LEMMA-DEFINITION 3.53.– *Let B be a differential (respectively analytic) manifold, $\mathbf{G}$ a Lie group, and consider the right action of $\mathbf{G}$ on $B \times \mathbf{G}$ defined by $(b, t).s = (b, ts)$ $(t, s \in \mathbf{G})$. The orbits are the fibers $pr_1^{-1}(b)$ $(b \in B)$, pr_1 is a submersion, the manifold of orbits $\mathbf{G}\backslash(B \times \mathbf{G})$ exists and can be identified with B, and $\mathbf{G}$ acts freely on $B \times \mathbf{G}$.*

Thus, $(B \times \mathbf{G}, B, pr_1)$ *is a principal bundle, said to be* trivial. *A principal bundle is said to be* trivializable *if it is isomorphic to a trivial principal bundle.*

(III) PRINCIPAL BUNDLE OF A QUOTIENT LIE GROUP Let $\mathbf{K}, \mathbf{H}, \mathbf{G}$ be Lie groups such that $\mathbf{K} \subset \mathbf{H} \subset \mathbf{G}$ and $\mathbf{K} \lhd \mathbf{H}$ (i.e. $\mathbf{K}$ is normal in $\mathbf{H}$: see [P1], section 2.2.2(**I**)). Consider the canonical surjection $\pi : \mathbf{G}/\mathbf{K} \twoheadrightarrow \mathbf{G}/\mathbf{H}$ that sends each coset mod.$\mathbf{K}$ to the coset mod.$\mathbf{H}$ containing it. By Noether's third isomorphism theorem ([P1], section 2.2.3(**II**)), $\mathbf{G}/\mathbf{H} \cong (\mathbf{G}/\mathbf{K}) / (\mathbf{H}/\mathbf{K})$ and $(\mathbf{G}/\mathbf{K}, \mathbf{G}/\mathbf{H}, \pi)$ is a principal bundle with base $\mathbf{G}/\mathbf{H}$ and structural group $\mathbf{H}/\mathbf{K}$ whose fibers are diffeomorphic to the quotient Lie group $\mathbf{H}/\mathbf{K}$.

3.5.2. *Vertical tangent vectors*

In a principal bundle $(P, B, \mathbf{G}, \pi)$, each fiber $\mathbf{G}_b$ over $\{b\}$ is a copy of $\mathbf{G}$ *as a differentiable manifold but not as a group*. Nevertheless, choosing a local morphic section $\sigma \in \Gamma(U, P)$

$$B \supset U \ni b \mapsto \sigma(b) \in P \qquad [3.11]$$

$\left(\sigma \circ \pi = 1_{\pi^{-1}(U)}\right)$ specifies an origin $p = \sigma(b)$ on each fiber $\mathbf{G}_b$ and therefore identifies $\mathbf{G}_b$ with $\mathbf{G}$ *as a Lie group* under the isomorphism (of Lie groups) $\mathbf{G} \tilde{\rightarrow} \mathbf{G}_b : g \mapsto \sigma(b).g$, which depends on σ. The neutral element of the *Lie group* $\mathbf{G}_b$ is $e(b, \sigma) = \sigma(b).e = p$. Recall that $\pi(\mathbf{G}_b) = \{b\}$.

We already defined the vertical tangent spaces of a fibration (Lemma-Definition 3.4(3)). We will now restate this idea in the context of principal bundles:

DEFINITION 3.54.– *Let $q \in P$. The space $V_q(P) = \ker(T_q(\pi)) = T_q\left(\mathbf{G}_{\pi(q)}\right) \subset T_q(P)$, where $\mathbf{G}_{\pi(q)}$ is the fiber over $\pi(q)$, is called the* space of vertical tangent vectors *at the point q, and the subbundle $V(P) \subset P$ is called the* bundle of vertical fields. *A* vertical tangent vector $\mathbf{v} = \mathbf{v}_q$ *is an element of $V_q(P)$ and therefore a* tangent vector of the fiber $\mathbf{G}_{\pi(q)}$ *(Figures 3.5 and 3.6).*

Let $q \in P$ and $b = \pi(q)$. Since both q and $p = \sigma(b)$ (where σ is the local section [3.11]) belong to $\mathbf{G}_b$ and $\mathbf{G}$ acts transitively on the fibers, there exists a unique element $g_\sigma \in \mathbf{G}$ such that $q = \sigma(b).g_\sigma$. Hence:

COROLLARY 3.55.– *The point q is fully determined by the two "components" $b = \pi(q) \in B$ and $g_\sigma \in \mathbf{G}$.*

REMARK 3.56.– *Consider again the example from section 3.1(**II**), where B is a pure n-dimensional Riemannian manifold (see section 4.5). We saw that each element $q \in P$ is an orthonormal frame (written as r_b) at the point $b = \pi(q)$. Specifying a local section σ as above amounts to choosing an "orthonormal frame of reference" $p = \sigma(b)$; a fiber $\mathbf{G}_b$ over $b \in B$ is the set of all orthonormal frames at the point b*

and is diffeomorphic to the structural group $\mathbf{G}$*; the latter is the group of orthonormal changes of reference, i.e. the group of orthogonal matrices* $O_n(\mathbb{R})$ *(section 2.4.1(**VI**)). The other orthonormal frames are the* $p.g$ $(g \in \mathbf{G})$*, which form the orbit* $\mathbf{G}_b = p.\mathbf{G}$*. The mapping* $\mathbf{G} \to p.\mathbf{G} : g \mapsto p.g = \sigma(b).g$ *is a diffeomorphism (that depends on* σ*).*

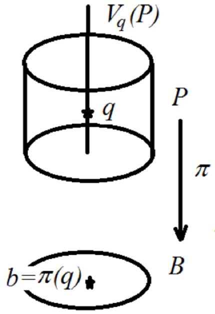


Figure 3.6. *Space of vertical tangent vectors*

3.5.3. *Morphisms of principal bundles*

Let $\lambda = (P, \mathbf{G}/P, \mathbf{G}, \pi)$, $\lambda' = (P', \mathbf{G}'/P', \mathbf{G}', \pi')$ be two principal bundles. A morphism from λ into λ' is a triple $\Phi := (f, \varphi, f^0)$, where $f : P \to P'$, $f^0 : B \to B'$ are morphisms of manifolds, $\varphi : \mathbf{G} \to \mathbf{G}'$ is a morphism of Lie groups (section 2.4.1(**I**)), and $\pi' \circ f = f^0 \circ \pi$, $f(p.g) = f(p).\varphi(g)$ for every $p \in P$ and every $g \in \mathbf{G}$. The principal bundles form a concrete category with base **Set** ([P1], section 1.3.1), which gives us the notion of the *structure* of a principal bundle. The morphism f^0 is determined by f, since $f^0(\pi(p)) = \pi'(f(p))$, so f^0 can be left implicit and Φ is simply written as (f, φ). When $B = B'$ and $f^0 = 1_B$ (respectively when $\mathbf{G} = \mathbf{G}'$ and $\varphi = 1_{\mathbf{G}}$), we say that Φ is a B*-morphism* (respectively a $\mathbf{G}$*-morphism*). A B-morphism that is a $\mathbf{G}$-morphism (when $B = B'$ and $\mathbf{G} = \mathbf{G}'$) is said to be a $\mathbf{G}$-*B*-*morphism*. A $\mathbf{G}$-*B*-*isomorphism* (or *an isomorphism of principal bundles* if this is not ambiguous) is a $\mathbf{G}$-B-morphism $(f, 1_{\mathbf{G}})$ such that f is an isomorphism of manifolds.

3.5.4. *Principal bundles defined by cocycles*

(I) Let $\mathcal{U} = (U_i)_{i \in I}$ be an open covering of B and, for every $i \in I$, let $\Gamma(U_i, P) \ni \sigma_i : U_i \to P$ be a section of P over U_i. If $U_i \cap U_j \neq \varnothing$, let $b \in U_i \cap U_j$ and $p_i = \sigma_i(b)$; then $p_i, p_j \in \mathbf{G}_b$, so there exists a unique element $g_{ij}(b) \in \mathbf{G}$ such that $p_j = p_i.g_{ij}(b)$. The mapping $g_{ij} : B \to \mathbf{G}$ is uniquely determined by the condition

$$\sigma_j(b) = \sigma_i(b).g_{ij}(b), \quad \forall b \in U_i \cap U_j \qquad [3.12]$$

and is a morphism of manifolds called the *transition morphism*. It immediately follows that

$$g_{ik}(b) = g_{ij}(b)\, g_{jk}(b), \quad \forall b \in U_i \cap U_j \cap U_j, \qquad [3.13]$$

i.e. the diagram

$$\begin{array}{ccc} i & \xrightarrow{g_{ij}} & j \\ & g_{ik} \searrow & \downarrow g_{jk} \\ & & k \end{array}$$

commutes. A family of morphisms $(g_{ij})_{i,j \in I \times I}$ satisfying this condition is called a *cocycle* on B taking values in $\mathbf{G}$ subordinate to the covering $\mathcal{U}$.

The cocycle condition [3.13] is analogous to a well-known property of bipoints in the plane, namely that $\overrightarrow{ij} + \overrightarrow{jk} = \overrightarrow{ik}$. The *cycle* of the three letters i, j, k is $i \mapsto j \mapsto k \mapsto i$; (g_{ik}) is a *cocycle* because the last arrow is reversed.

In the case where P is a frame (section 3.1**(II)**), a transition morphism g_{ij} is a change of frame.

(II) Conversely, let B be a manifold, $\mathbf{G}$ a Lie group, and $(g_{ij})_{i,j \in I \times I}$ a cocycle on B taking values in $\mathbf{G}$. It can be shown that there exist a principal bundle $\lambda = (P, \mathbf{G} \backslash P, \mathbf{G}, \pi)$ and a family of sections $(\sigma_i)_{i \in I}$ of λ on the U_i such that [3.12] holds. This principal bundle is *unique up to isomorphism* in the sense that, if $\lambda' = (P', \mathbf{G} \backslash P', \pi')$ is another principal bundle and $(\sigma'_i)_{i \in I}$ is another family of sections of λ' satisfying the same conditions, then there exists a $\mathbf{G}$-B-isomorphism f such that

$$\sigma'_i = f \circ s_i \quad (i \in I). \qquad [3.14]$$

The principal bundle λ is said to be *defined by the cocycle* $(g_{ij})_{i,j \in I \times I}$.

(III) We say that two cocycles $(g_{ij})_{i,j \in I \times I}$, $\left(g'_{ij}\right)_{i,j \in I \times I}$ subordinate to the covering $\mathcal{U}$ of B are *cohomologous* if there exists a family $(h_i)_{i \in I}$ of morphisms from U_i into $\mathbf{G}$ such that $g'_{ij}(b) = h_i(b)^{-1} \circ g_{ij}(b) \circ h_j(b)$ for all $b \in B$. We have the following result ([BOU 16], Chapter 1, section 5.7):

THEOREM 3.57.– *Two principal bundles are* $\mathbf{G}$-B-*isomorphic if and only if the cocyles that determine them are cohomologous.*

3.5.5. *Fiber bundle associated with a principal bundle*

Let $\lambda = (P, B, \mathbf{G}, \pi)$ be a principal bundle with base $B = \mathbf{G} \backslash P$. Furthermore, let F be a manifold and suppose that the structural group $\mathbf{G}$ has a differentiable (if $r = \infty$), respectively analytic (if $r = \omega$) *left* action on F. Then, $\mathbf{G}$ has a differentiable (respectively analytic) and free *right* action on $P \times F$ defined by

$$(p, y).g = \left(p.g, g^{-1}.y\right), \quad (p, y) \in P \times F, g \in \mathbf{G}.$$

We have the following result ([DIE 93], Volume 3, (16.14.7); [BOU 82a], (6.5.1)):

LEMMA-DEFINITION 3.58.– *i) The manifold of orbits* $\mathbf{G}\backslash(P \times F)$ *exists and is written as* $P \times^{\mathbf{G}} F$; *write* $(p, y) \mapsto p.y$ *for the projection from* $P \times F$ *onto* $P \times^{\mathbf{G}} F$.

ii) For every orbit $z \in \mathbf{G}\backslash(P \times F)$, *let* $\pi_F(z)$ *be the element of* P *equal to* $\pi(x)$ *for all* $(x, y) \in z$. *Then,* $\left(P \times^{\mathbf{G}} F, B, \pi_F\right)$ *is a fibration whose fibers are diffeomorphic to (and said to be of type)* F. *More precisely, if* U *is an open subset of* B *such that* $\pi^{-1}(U)$ *is trivializable and* $\sigma : U \to \pi^{-1}(U)$ *is a morphic section of* $\pi^{-1}(U)$, *then* $(b, y) \mapsto \sigma(b).y$ *is a* U*-isomorphism from* $U \times F$ *onto* $\pi_F^{-1}(U)$.

We say that $P \times^{\mathbf{G}} F$ *is the* fiber bundle associated with λ with fibers of type F.

PROOF.– (ii): We return to the case where $P = B \times \mathbf{G}$ is trivial (Lemma-Definition 3.53) and $\sigma(b) = (b, e)$ (where e denotes the neutral element of $\mathbf{G}$). Then, $P \times^{\mathbf{G}} F$ can be identified with $B \times F$, π_F can be identified with pr_1, and $\sigma(b).y$ can be identified with (b, y). ∎

COROLLARY 3.59.– *Let* $(P, B, \mathbf{G}, \pi)$ *be a principal bundle with structural group* $\mathbf{G}$ *that has a differential (respectively analytic) right action on* P, $\mathbf{F}$ *a Banach space and* $\rho : \mathbf{G} \to GL(\mathbf{F})$ *a linear representation of* $\mathbf{G}$ *in* $\mathbf{F}$ *(Definition 2.87). Let* $\pi_{\mathbf{F}} : P \times^{\mathbf{G}} \mathbf{F} \twoheadrightarrow B$ *be the projection defined as in Lemma-Definition 3.58,* mutatis mutandis. *Then,* $\left(P \times^{\mathbf{G}} \mathbf{F}, B, \pi_{\mathbf{F}}\right)$ *has a vector bundle structure whose fibers are Banach spaces isomorphic to* $\mathbf{F}$.

PROOF.– The group $\mathbf{G}$ has an analytic left action on $\mathbf{F}$ defined by $(g, y) \mapsto \rho(g)\, y = g.y$. Let U be a neighborhood of b in B over which P is trivializable (Lemma-Definition 3.53) and let $\sigma : U \to \pi_{\mathbf{F}}^{-1}(U)$ be a morphic section of P on U. The Banach space structure of $\{b\} \times \mathbf{F}$ is transported to $\pi_{\mathbf{F}}^{-1}(b)$ by $(b, y) \mapsto \sigma(b)\, y$. The Banach space structure thus obtained on $\pi_{\mathbf{F}}^{-1}(b)$ does not depend on the choice of morphic section σ (**exercise**). It is clear that $\left(P \times^{\mathbf{G}} \mathbf{F}, B, \pi_{\mathbf{F}}\right)$ is a vector bundle that satisfies the stated properties (Lemma-Definition 3.58(ii)). ∎

3.5.6. *Extension, restriction, quotientization of the structural group*

These operations are analogous to the extension and restriction of the ring of scalars of a module ([P1], section 3.1.5(**VI**)).

(I) Let $\lambda = (P, B, \mathbf{G}, \pi)$ be a principal bundle with base $B = P/\mathbf{G}$ (where $\mathbf{G}$ acts on P *on the left*[8]), $\varphi : \mathbf{G} \to \mathbf{H}$ a morphism of Lie groups, $g \,.\, h = \varphi(g)\,.h$ a left action of $\mathbf{G}$ on $\mathbf{H}$, and $P \times^{\mathbf{G}} \mathbf{H}$ the fiber bundle associated with λ with fibers of type $\mathbf{H}$ (Lemma-Definition 3.58). The group $\mathbf{H}$ has a right action on $E := P \times^{\mathbf{G}} \mathbf{H}$ defined by $(p.g)\,.h = p.(g\,.\,h)$. Write $\varphi(\lambda)$ for the principal bundle $(E, \mathbf{H}\backslash E, \mathbf{H}, \pi_{\mathbf{H}})$. We have $\mathbf{H}\backslash E = \mathbf{H}\backslash \left(P \times^{\mathbf{G}} \mathbf{H}\right) = \mathbf{H}\backslash (P \times \mathbf{H})/\mathbf{G} \cong P/\mathbf{G}$; thus, the base of $\varphi(\lambda)$ can be identified with the base B of λ. We say that the principal bundle $\varphi(\lambda) = (E, B, \mathbf{H}, \pi_{\mathbf{H}})$ is deduced from λ by the morphism φ.

If $\mathbf{G} \subseteq \mathbf{H}$, we say that $\varphi(\lambda)$ is deduced from λ by *extension of the structural group to* $\mathbf{H}$.

If λ is defined using an open covering $\mathcal{U} = (U_i)_{i \in I}$ of $B = P/\mathbf{G}$ and a cocycle $(g_{ij})_{(i,j) \in I \times I}$ subordinate to $\mathcal{U}$ (section 3.5.4), then $\varphi(\lambda)$ can be defined using $\mathcal{U}$ and the cocycle (h_{ij}) subordinate to $\mathcal{U}$ for which $h_{ij} = \varphi \circ g_{ij}$.

EXAMPLE 3.60.– *Given a real pure n-dimensional manifold, this situation occurs when $P = T(B)$, $\mathbf{G} = O_n(\mathbb{R})$, $\mathbf{H} = GL_n(\mathbb{R})$ (extending from orthogonal frames to arbitrary linear frames).*

(II) Assuming again that $\mathbf{G} \subseteq \mathbf{H}$, the principal bundle λ is deduced from $\varphi(\lambda)$ by restriction of the structural group to $\mathbf{G}$. We have seen that λ and $\varphi(\lambda)$ have the same base B.

Let a principal bundle $\lambda' = (P', B', \mathbf{H}, \pi')$ be given. If there exists a principal bundle $\lambda = (P, B, \mathbf{G}, \pi)$ with the same base $B = B'$ as λ' and such that $\lambda' = \varphi(\lambda)$, the bundle λ is not unique in general: see Theorem 7.43 in section 7.3.8.

Nevertheless, $\mathbf{G}$ has a right action on P' defined by $p' \mapsto p'.g$ $(g \in \mathbf{G})$, so we obtain a principal bundle $(P', \mathbf{G}\backslash P', \mathbf{G}, \pi)$, $\pi : P' \twoheadrightarrow \mathbf{G}\backslash P'$ that is deduced from λ' by restriction of the structural group to $\mathbf{G}$ but whose base does not coincide with B'.

(III) Let $\lambda = (P, B, \mathbf{H}, \pi)$ be a principal bundle with base $B = \mathbf{H}\backslash P$ (where $\mathbf{H}$ acts on P on the right) and $\mathbf{G} \subseteq \mathbf{H}$; then $\mathbf{H}$ has a left differentiable action on the homogeneous space $\mathbf{H}/\mathbf{G}$, namely the canonical surjection. The fiber bundle λ' associated with the principal bundle λ with fibers of type $\mathbf{H}/\mathbf{G}$ (which is a Lie group whenever $\mathbf{G}$ is a normal subgroup of $\mathbf{H}$) is $E = P \times^{\mathbf{H}} \mathbf{H}/\mathbf{G}$, and we have $E = \mathbf{H}\backslash (P \times \mathbf{H}/\mathbf{G})$ (Lemma-Definition 3.58). In addition, $\mathbf{H}\backslash (P \times \mathbf{H}/\mathbf{G}) \cong \mathbf{G}\backslash P$ (**exercise**), thus $\lambda' = (\mathbf{G}\backslash P, \mathbf{H}\backslash P, \pi')$ where $\pi' : \mathbf{G}\backslash P \twoheadrightarrow \mathbf{H}\backslash P$ sends every orbit of the action on $\mathbf{G}$ to the unique orbit of the action on $\mathbf{H}$ that contains it.

8 See footnote 7, section 3.5.1.

3.5.7. *Examples of trivial principal bundles*

Parts **(I)** and **(II)** summarize some of what was said in section 3.1 and Remark 3.56. Part **(III)** prepares for the notion of a projective space.

(I) AFFINE SPACE Consider the affine space $B = \mathfrak{Aff}_n(\mathbb{R})$ (section 2.4.2**(IV)**), a frame q with origin $b = \pi(q)$, and the bundle of frames P (section 3.1**(II)**). Write q in the form (b, h), where h is a basis $\left[e^1, ..., e^n\right]$ (written as a row of n column vectors) of $T_b(B) \cong \mathbb{R}^n$, so $h \in \mathrm{GL}_n(\mathbb{R})$. The group $\mathbf{G} = \mathrm{GL}_n(\mathbb{R})$ is the structural group of P and acts freely and analytically on P by $(b, h).g = (b, h.g)$ for all $g \in \mathbf{G}$.

(II) EUCLIDEAN SPACE If B is the Euclidean space $\mathfrak{Euc}_n$, the above remarks remain valid when the structural group $\mathrm{GL}_n(\mathbb{R})$ is replaced by $\mathrm{O}_n(\mathbb{R})$ or $\mathrm{SO}_n(\mathbb{R})$, and $h \in \mathrm{O}_n(\mathbb{R})$. If $\mathbf{G} = \mathrm{SO}_n(\mathbb{R})$, the action of this group leaves the orientation of $T_b(B)$ unchanged.

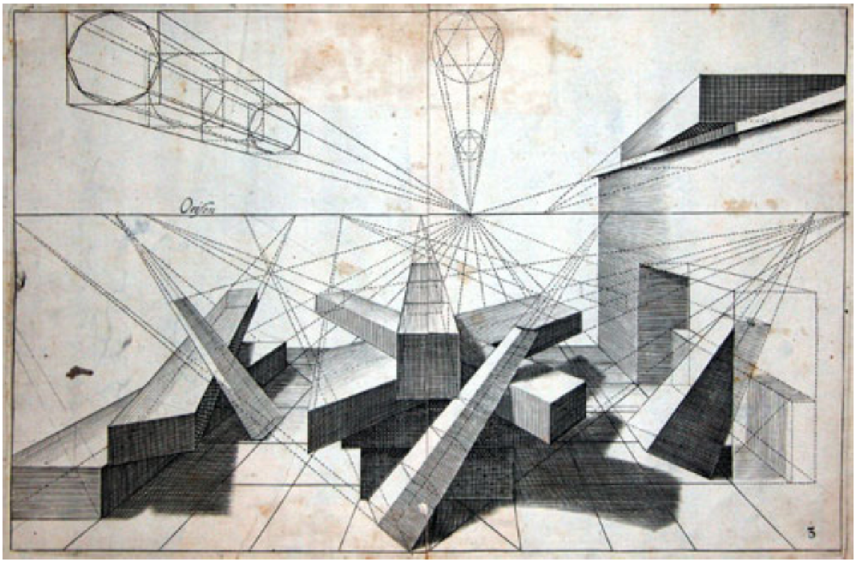
(III) PROJECTIVE SPACE When solving geometric problems, the *projective space* is sometimes more suitable than the affine space [PER 95]. Let $\mathbf{E} = \mathbb{K}^{n+1}$ and $\sim$ the equivalence relation defined as follows on $\mathbf{E}^\times = \mathbf{E} - \{0\} : x \sim y$ whenever there exists $\lambda \neq 0$ such that $y = \lambda.x$. The n-dimensional projective space $\mathbf{P}_n(\mathbb{K})$ is defined as the quotient $\mathbf{E}^\times / \sim$. This set, equipped with the quotient structure, is a connected compact analytic manifold. A projective line (respectively a projective plane) is defined as an element of $\mathbf{P}_1(\mathbb{K})$ (respectively $\mathbf{P}_2(\mathbb{K})$).

Let $x = (x_0, x_1, ..., x_n)$ be an arbitrary point of $\mathbb{K}^{n+1}$ and write $\bar{x}$ for its canonical image in $\mathbf{P}_n(\mathbb{K})$. We say that $\lambda.x_0, \lambda.x_1, ..., \lambda.x_n$ are the *homogeneous coordinates* of $\bar{x}$ for every $\lambda \neq 0$. Let H be the hyperplane with equation $x_0 = 0$, $\bar{H}$ the corresponding projective hyperplane (i.e. its canonical image), and $U = \mathbf{P}_n(\mathbb{K}) - \bar{H}$. The mapping $\varphi : U \to \mathbb{K}^n : x \mapsto (x_1/x_0, ..., x_n/x_0)$ is bijective. By choosing an origin O in the affine space $\mathfrak{Aff}_n(\mathbb{K})$, we can identify the latter with $\varphi^{-1}(\mathbb{K}^n) = U$, which amounts to embedding $\mathfrak{Aff}_n(\mathbb{K})$ in the disjoint union $\mathbf{P}_n(\mathbb{K}) = U \biguplus \bar{H}$ ([P1], section 1.2.6**(II)**). The points of U are said to be "at finite distance", whereas the points of $\bar{H}$ are said to be "at infinity".

For $n = 1$, $\bar{H} = \{\infty\}$. Topologically, $\mathbf{P}_1(\mathbb{K})$ is the Alexandroff complexification $\mathbb{S}^1$ of $\mathbb{R}$ when $\mathbb{K} = \mathbb{R}$ and the Riemann sphere when $\mathbb{K} = \mathbb{C}$ ([P2], section 2.3.9).

For $n = 2$, $\bar{H}$ is the "line at infinity" D_∞ with homogeneous coordinates $(0, x_1, x_2)$, whereas the affine space can be identified with the points with homogenous coordinates $(1, x_1, x_2)$. A line $\bar{D}$ in $\mathbf{P}_2(\mathbb{K})$ is the canonical image of a line D passing through the origin in $\mathbb{K}^3$ with equation $\sum_{0 \leq i \leq 2} \alpha^i.x_i = 0$ (where the α^i are not all zero). If $\alpha^1 = \alpha^2 = 0$, then $D = D_\infty$. Otherwise, $\bar{D} \cap \mathfrak{Aff}_2(\mathbb{K})$ is the set of points whose homogeneous coordinates $(1, x_1, x_2)$ satisfy the relation $\alpha^0 + \alpha^1.x_1 + \alpha^2.x_2 = 0$, and this intersection is therefore an *affine line*.

Moreover, $\bar{D} \cap D_\infty$ is the set of points whose homogeneous coordinates $(0, x_1, x_2)$ satisfy the relation $\alpha^1.x_1 + \alpha^2.x_2 = 0$. There exists a unique point satisfying these conditions; it has homogeneous coordinates $\left(0, \alpha^2, -\alpha^1\right)$. This point at infinity determines the direction of D (see Figure 3.7).



Jan Vredeman de Vries, *in* **Opera Mathematica** by Samuel Marolois (circa 1620)

Figure 3.7. *Projective plane*

A frame q of the projective space $B = \mathbf{P}_n(\mathbb{K})$ can be chosen to be of the form $\left(b, \bar{h}\right)$, $b = \pi(q)$, where $\bar{h}$ is a line $[\bar{e}_0, \bar{e}_1, ..., \bar{e}_n] \in \mathrm{PSL}_{n+1}(\mathbb{K})$, $h = [e_0, e_1, ..., e_n]$ is a basis of $\mathbb{K}^{n+1}$, and $\bar{e}_i$ is the canonical image of e_i $(i = 0, 1, ..., n)$. The projective group $\mathrm{PSL}_{n+1}(\mathbb{K})$ has a simple transitive and analytic action on $\mathbf{P}_n(\mathbb{K})$ defined by $\left(b, \bar{h}\right).\bar{g} = \left(b, \bar{h}.\bar{g}\right)$ $(\bar{g} \in \mathrm{PSL}_{n+1}(\mathbb{K}))$.

This page intentionally left blank

4

Tensor Calculus on Manifolds

4.1. Introduction

Tensor calculus was the culmination of pioneering work by B. Christoffel and G. Ricci, in 1869 and 1887–1896, respectively. It reached maturity in a joint publication by Ricci-Curbastro and Levi-Cività in 1900 [RIC 00]. It is difficult to overstate its importance as a field – general relativity could not exist without it. Tensor calculus plays an essential role in every area of physics; it is also crucial for continuum mechanics and many other engineering sciences, and it, of course, lies at the heart of differential geometry. Section 4.2 of this chapter is purely algebraic (with some algebraic topology in section 4.2.6). The mathematical objects that physicists call *tensors* are more precisely *tensor fields*; they are only introduced in section 4.3. Historically, "tensors" were presented to students as exotic mathematical millipedes of the form $t^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}$ (or more briefly) that behave in a certain way under change of coordinates (see [4.3, 4.4]). This stemmed from the decision to only define the *components* of the tensor (see [4.2]). As a result, students could complete an entire course of tensor calculus without finding a satisfactory answer to the question: "But what actually is a tensor?" We will, of course, choose a different approach by *defining* tensor fields before attempting to study them.

The fields of contravariant tensors of order 1 are just the vector fields. The fields of *antisymmetric* contravariant tensors of order p are called p-vector fields. The fields of covariant tensors of order 1 are the covector fields, or fields of differential forms of degree 1; the fields of *antisymmetric* covariant tensors of order p are the fields of differential forms of degree p, also called p-forms or p-covectors ([P1], section 3.3.8(**VII**)). The integral of an m-form over an m-dimensional manifold (line integral if $m = 1$, surface integral if $m = 2$, volume integral if $m = 3$) is a key idea that has various applications in physics (see, for example, section 5.6.4). These forms also allow us to define the *orientation* of a manifold (section 4.4.4). The differential forms of degree 1 go back to Euler [SAM 01]; higher-order differential forms were introduced by É. Cartan [CAR 99], who made abundant use of them [CAR 22a], whereas the differential forms "of odd type" were developed by G. de Rham [DER 84]. Differential forms are easy to define on $(\mathcal{FN})$ or $(\mathcal{SN})$ manifolds (Remark 4.33), and their exterior and interior products have properties analogous to those stated in section 4.4.3 ([KRI 97], Chapter VII, section 33).

The fields of non-degenerate *symmetric* covariant tensors of order 2, on the other hand, allow us to define pseudo-Riemannian manifolds, which lead to the Riemannian and Lorentz manifolds of general relativity as special cases. The notion of a Riemann space (or Riemannian manifold) is, of course, due to Riemann; they were introduced in his inaugural lecture, mentioned in section 2.1. Pseudo-Riemannian manifolds (and in particular Lorentz manifolds) are the result of a work by Minkowski, Einstein, and Grossmann ([PAU 58], Chapter II, section 7 and Chapter IV, section 50).

4.2. Tensor calculus

As a reminder, we are using the conventions outlined in section 1.2.1, especially those which relate to the representations of vectors and covectors.

4.2.1. *Tensors*

(I) NOTION OF A TENSOR Let $\mathbf{K}$ be a field. Recall that, if $\mathbf{E}$ and $\mathbf{F}$ are $\mathbf{K}$-vector spaces, the tensor product $\mathbf{E} \otimes \mathbf{F}$ is defined as the $\mathbf{K}$-vector space generated by the products $x_i \otimes y_j$ $(x_i \in \mathbf{E},\ y_j \in \mathbf{F})$, where $\otimes : \mathbf{E} \times \mathbf{F} \to \mathbf{E} \otimes \mathbf{F}$ is a canonical bilinear mapping that is the solution of a universal problem ([P1], section 3.1.5). In particular, if $(e_i)_{i \in I}$ and $(f_j)_{j \in J}$ are bases of $\mathbf{E}$ and $\mathbf{F}$, respectively, then $\mathbf{E} \otimes \mathbf{F}$ consists of the elements of the form $\sum_{\substack{i,j \\ \text{finite}}} a^{ij} e_i \otimes f_j$ $\left(a^{ij} \in \mathbf{K}\right)$. If $\mathbf{G}$ is a third $\mathbf{K}$-vector space, then $(\mathbf{E} \otimes \mathbf{F}) \otimes \mathbf{G} = \mathbf{E} \otimes (\mathbf{F} \otimes \mathbf{G})$; thus, this space can be written as $\mathbf{E} \otimes \mathbf{F} \otimes \mathbf{G}$.

Let $\mathbf{E}$ be a finite-dimensional vector space over $\mathbf{K}$, and let $\mathbf{E}^{\vee}$ be its dual.[1] Write $\mathbf{T}_0^n(\mathbf{E}) = \mathbf{E}^{\otimes n}$ for the n-th tensor product of $\mathbf{E}$, namely $\bigotimes_{i=1}^n \mathbf{E}_i$, where $\mathbf{E}_i = \mathbf{E}$ $(i = 1, ..., n)$, with the convention that $\mathbf{E}^{\otimes 0} = \mathbf{K}$; furthermore, write $\mathbf{T}_m^0(\mathbf{E}) = (\mathbf{E}^{\vee})^{\otimes m}$ and $\mathbf{T}_q^p(\mathbf{E}) = \mathbf{T}_q^0(\mathbf{E}) \otimes \mathbf{T}_0^p(\mathbf{E})$. The elements of the vector space $\mathbf{T}_q^p(\mathbf{E})$ are called p-times contravariant and q-times covariant *tensors* (or tensors of type $\begin{pmatrix} p \\ q \end{pmatrix}$

1 (a) Since $\mathbf{E}$ is finite-dimensional, its dual ([P2], section 3.6.1) coincides with its algebraic dual ([P1], section 3.1.2). (b) Up to and including section 4.2.5, we can replace $\mathbf{K}$ by a commutative ring $\mathbf{A}$ and the finite-dimensional $\mathbf{K}$-vector spaces by finitely generated $\mathbf{A}$-modules ([P1], section 2.3.1(**II**)).

or $(p,q))^2$; the *order* of any such tensor is the integer $p+q$. Thus, a *vector* of $\mathbf{E}$ is an element of $\mathbf{T}_0^1(\mathbf{E})$ and hence a tensor of type $(1,0)$, i.e. a once contravariant tensor. A linear form on $\mathbf{E}$ (also called a *covector*) is an element of $\mathbf{T}_1^0(\mathbf{E})$ and hence a tensor of type $(0,1)$, i.e. a once covariant tensor. The vector space $\mathbf{T}_0^n(\mathbf{E})$ is also written as $\mathbf{T}^n(\mathbf{E})$.

If $\mathbf{E}$ and $\mathbf{F}$ are finite-dimensional $\mathbf{K}$-vector spaces, we know that the space $\mathbf{E}^\vee \otimes \mathbf{F}$ can be identified with $\mathcal{L}(\mathbf{E};\mathbf{F})$ by identifying $x^\vee \otimes y$ with the $\mathbf{K}$-linear mapping $f : x \mapsto \langle x, x^\vee\rangle y$ ([P1], section 3.1.5(**I**)). If $(e_i^\vee)_{1\leq i\leq n}$ is a basis of $\mathbf{E}^\vee$ and $(f_i)_{1\leq i\leq m}$ is a basis of $\mathbf{F}$, the elements of $\mathbf{E}^\vee$ can be represented by columns of n elements, the elements of $\mathbf{F}$ can be represented by rows of m elements, and so the elements of $\mathbf{E}^\vee\otimes\mathbf{F}$ can be represented by $n\times m$ matrices, like the elements of $\mathcal{L}(\mathbf{E};\mathbf{F})$. An element of $\mathbf{E}^\vee\otimes\mathbf{F}$ of the form $x^\vee\otimes y$ can be identified with an $n\times m$ matrix of rank 1. Recall that any arbitrary element of $\mathbf{E}^\vee\otimes\mathbf{F}$ is a *finite sum* of terms of this form.

LEMMA-DEFINITION 4.1.– *i) The* $\mathbf{K}$*-vector space* $\mathbf{T}(\mathbf{E}) = \bigoplus_{n=0}^{\infty}\mathbf{T}^n(\mathbf{E})$ *is a graduated* $\mathbf{K}$*-algebra ([P1], section 2.3.12) called the* tensor algebra *of* $\mathbf{E}$.

ii) Let $\mathbf{E}$ *and* $\mathbf{F}$ *be two finite-dimensional* $\mathbf{K}$*-vector spaces,* $f:\mathbf{E}\to\mathbf{F}$ *a linear mapping, and* $\mathbf{T}(f):\mathbf{T}(\mathbf{E})\to\mathbf{T}(\mathbf{F})$ *the mapping uniquely determined by the condition*

$$\mathbf{T}(f)(x_1\otimes\ldots\otimes x_n) = f(x_1)\otimes\ldots\otimes f(x_n)$$

for every integer $n\geq 0$ *and all* $x_1,\ldots,x_n\in\mathbf{E}$*. Then,* $\mathbf{T}(f)$ *is a morphism of algebras and* $\mathbf{T}$ *is a covariant functor ([P1], section 1.2.1) from the category of* $\mathbf{K}$*-vector spaces into the category of graduated* $\mathbf{K}$*-algebras (**exercise**).*

(II) DUALITY AND INDEX CONTRACTIONS The space $\mathbf{T}_q^p(\mathbf{E})^\vee$ can be identified with $\mathbf{T}_p^q(\mathbf{E})$ using the relation

$$\Bigg\langle \underbrace{\bigotimes_{1\leq i\leq p} x_i \otimes \bigotimes_{1\leq j\leq q} x^{\vee j}}_{\in\mathbf{T}_q^p(\mathbf{E})}, \underbrace{\bigotimes_{1\leq j\leq q} y_j \otimes \bigotimes_{1\leq i\leq p} y^{\vee i}}_{\in\mathbf{T}_p^q(\mathbf{E})} \Bigg\rangle = \prod_{\substack{1\leq i\leq p\\ 1\leq j\leq q}} \left\langle x_i, y^{\vee i}\right\rangle \left\langle y_j, x^{\vee j}\right\rangle. \quad [4.1]$$

2 This convention has been adopted by various authors, including [DIE 93, BOU 12, BIS 68], but is not universal. For example, the opposite convention is found in ([SPI 99], Vol. 1, p. 123), where p-times contravariant and q-times covariant tensors are said to be of type $\begin{pmatrix} q \\ p \end{pmatrix}$.

An arbitrary element $\boldsymbol{t}$ of $\mathbf{T}_q^p(\mathbf{E})$ can be expressed in the following form with respect to the basis $(e_i)_{1\leq i\leq m}$ of $\mathbf{E}$ and its dual basis $\left(e^{\vee i}\right)_{1\leq i\leq m}$:

$$\boldsymbol{t} = \sum_{i_1,\dots,i_p,j_1,\dots,j_q} t_{j_1\dots j_q}^{i_1\dots i_p}.e_{i_1}\otimes\dots e_{i_p}\otimes e^{\vee j_1}\otimes\dots e^{\vee j_q}, \quad t_{j_1\dots j_q}^{i_1\dots i_p}\in\mathbf{K}. \qquad [4.2]$$

The $t_{j_1\dots j_q}^{i_1\dots i_p}$ are the components of the tensor $\boldsymbol{t}$.

REMARK 4.2.– Einstein's summation convention, *which is ubiquitous in physics, omits the symbol* $\sum$. *We will avoid this convention below for clarity; nevertheless, whenever bounds are not stated for the indices, it should be implicitly understood that these indices range from 1 to the dimension of the vector space or manifold being considered.*

Write c_j^i for the $\mathbf{K}$-linear *index contraction* mapping $\mathbf{T}_q^p(\mathbf{E})\to\mathbf{T}_{q-1}^{p-1}(\mathbf{E})$, $i\in\{i_1,\dots,i_q\}$ and $j\in\{j_1,\dots,j_p\}$, defined by[3]

$$c_j^i\left(x_1\otimes\dots\otimes x_p\otimes x^{\vee 1}\otimes\dots\otimes x^{\vee q}\right)$$
$$=\left\langle x_i,x^{\vee j}\right\rangle.x_1\otimes\dots\otimes\widehat{x_i}\otimes\dots\otimes x_p\otimes x^{\vee 1}\otimes\dots\otimes\widehat{x^{\vee j}}\otimes\dots\otimes x^{\vee q}.$$

In particular, if $x=\sum_i x^i e_i$ and $x^\vee=\sum_i x_i^\vee e^{i\vee}$, then

$$c_1^1\left(x\otimes x^\vee\right)=\sum_i x^i.x_i^\vee=\left\langle x,x^\vee\right\rangle\in\mathbf{K}.$$

(III) CHANGE OF BASIS Let $A=\left(A_i^{j'}\right)$ be a change-of-basis matrix in $\mathbf{E}$. Consider the tensor [4.2]. By [1.1] and [1.2], with the conventions of section 1.2.1**(II)** and noting that $\mathbf{K}$ is commutative,

$$\boldsymbol{t} = \sum_{i'_1,\dots,i'_p,j'_1,\dots,j'_q} t_{j'_1\dots j'_q}^{i'_1\dots i'_p}.e_{i'_1}\otimes\dots e_{i'_p}\otimes e^{\vee j'_1}\otimes\dots e^{\vee j'_q}$$

with

$$t_{j'_1\dots j'_q}^{i'_1\dots i'_p} = \sum_{i_1,\dots,i_p,j_1,\dots,j_q} A_{i_1}^{i'_1}\dots A_{i_p}^{i'_p}A_{j'_1}^{j_1}\dots A_{j'_q}^{j_q}.t_{j_1\dots j_q}^{i_1\dots i_p}, \qquad [4.3]$$

$$t_{j_1\dots j_q}^{i_1\dots i_p} = \sum_{i'_1,\dots,i'_p,j'_1,\dots,j'_q} A_{i'_1}^{i_1}\dots A_{i'_p}^{i_p}A_{j_1}^{j'_1}\dots A_{j_q}^{j'_q}.t_{j'_1\dots j'_q}^{i'_1\dots i'_p}. \qquad [4.4]$$

3 The notation $\widehat{(.)}$ indicates that the symbol $(.)$ should be omitted.

REMARK 4.3.– *There is an obvious analogy between [1.1] and [1.2]. Ever since Ricci-Curbastro and Levi-Cività, we say that a* vector *is a* contravariant tensor *and a* covector *is a* covariant tensor. *This is motivated by the observation that the* components *of a vector in a given basis transform like covectors under change of basis, whereas the components of a covector in the dual basis transform like vectors. For Ricci-Curbastro and Levi-Cività, the* components *were more important than the* vectors *and* covectors *themselves. Tensor calculus has been around for too long to change this terminology now!*

4.2.2. *Symmetric tensors and antisymmetric tensors*

(I) CASE OF CONTRAVARIANT TENSORS Let $\mathbf{E}$ be an m-dimensional $\mathbf{K}$-vector space. Consider the action of the permutation group $\mathfrak{S}_n$ on space of contravariant tensors $\mathbf{T}_0^n(\mathbf{E})$ defined by

$$\sigma.(x_1 \otimes ... \otimes x_n) = x_{\sigma^{-1}(1)} \otimes ... \otimes x_{\sigma^{-1}(n)} \tag{4.5}$$

for $\sigma \in \mathfrak{S}_n$ $(1 \leq n \leq m)$.

Given a basis $(e_i)_{1 \leq i \leq m}$ of $\mathbf{E}$, the $e_{i_1} \otimes ... \otimes e_{i_n}$ form a basis of the vector space $\mathbf{T}_0^n(\mathbf{E})$. Thus, let:

$$\boldsymbol{t} = \sum_{i_1,\ldots,i_n} t^{i_1,\ldots,i_n}.e_{i_1} \otimes ... \otimes e_{i_n} \in \mathbf{T}_0^n(\mathbf{E}).$$

Now, define:

$$\begin{aligned}\sigma.\boldsymbol{t} &= \sum_{i_1,\ldots,i_n} t^{i_1,\ldots,i_n}.\sigma(e_{i_1} \otimes ... \otimes e_{i_n}) \\ &= \sum_{i_1,\ldots,i_n} t^{i_1,\ldots,i_n}.e_{\sigma^{-1}(i_1)} \otimes ... e_{\sigma^{-1}(i_n)} = \sum_{i_1,\ldots,i_n} t^{\sigma(i_1),\ldots,\sigma(i_n)}.e_{i_1} \otimes ... \otimes e_{i_n}.\end{aligned}$$

This tensor $\boldsymbol{t}$ is said to be *symmetric*, respectively *antisymmetric* if $\sigma(\boldsymbol{t}) = \boldsymbol{t}$, respectively $\sigma.\boldsymbol{t} = \varepsilon_\sigma \boldsymbol{t}$ for every $\sigma \in \mathfrak{S}_n$, where ε_σ is the signature of the permutation σ.

This is equivalent to saying that its coordinates satisfy:

$$\boxed{t^{\sigma(i_1),\ldots,\sigma(i_n)} = t^{i_1,\ldots,i_n}, \quad \text{resp. } t^{\sigma(i_1),\ldots,\sigma(i_n)} = \varepsilon_\sigma t^{i_1,\ldots,i_n}.}$$

Let $\boldsymbol{t} \in \mathbf{T}_0^n(\mathbf{E})$. The *symmetrization*, respectively *antisymmetrization* ([P1], section 3.3.8(**VII**)), of this tensor is defined by

$$\mathrm{s}.\boldsymbol{t} = \sum_{\sigma \in \mathfrak{S}_n} \sigma.\boldsymbol{t}, \quad \text{resp. } \mathbf{a}.\boldsymbol{t} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma (\sigma.\boldsymbol{t}). \qquad [4.6]$$

We sometimes write ([SPI 99], Volume 1, Chapter 7):

$$\mathrm{alt} = \frac{1}{n!}\mathbf{a}.$$

If $\boldsymbol{t}$ is already symmetric (respectively antisymmetric), then $\mathrm{s}.\boldsymbol{t} = n!\boldsymbol{t}$ (respectively $\mathbf{a}.\boldsymbol{t} = n!\boldsymbol{t}$, so $\mathrm{alt}\boldsymbol{t} = \boldsymbol{t}$). Write $\mathbf{TS}^n(\mathbf{E}) \subset \mathbf{T}_0^n(\mathbf{E})$ (respectively $\mathbf{A}^n(\mathbf{E}) \subset \mathbf{T}_0^n(\mathbf{E})$) for the space of symmetric (respectively antisymmetric) contravariant tensors of order n on $\mathbf{E}$.

Consider the sets $H = \{i_1, ..., i_n\} \subset \{1, ..., m\}$, $i_1 < i_2 < ... < i_n$. Then, the elements $e_H = \mathbf{a}.(e_{i_1} \otimes ... \otimes e_{i_n})$ form a basis of $\mathbf{A}^n(\mathbf{E})$, and this basis has $\binom{m}{n}$ elements.

The space $\mathbf{TS}^n(\mathbf{E})$ can also be defined as a quotient of $\mathbf{T}_0^n(\mathbf{E})$: consider the vector subspace $\mathfrak{b}_n$ of $\mathbf{T}_0^n(\mathbf{E})$ generated by all the elements of the form

$$x_1 \otimes ... \otimes x_n - \sigma(x_1 \otimes ... \otimes x_n)$$

for every $x_i \in \mathbf{E}$ and $\sigma \in \mathfrak{S}_n$. Then, $\mathbf{TS}^n(\mathbf{E})$ can be identified with $\mathbf{T}_0^n(\mathbf{E}) / \mathfrak{b}_n$.

LEMMA-DEFINITION 4.4.– *1) The direct sum*

$$\mathbf{TS}(\mathbf{E}) = \bigoplus_{n=0}^{\infty} \mathbf{TS}^n(\mathbf{E})$$

is called the symmetric algebra *of* $\mathbf{E}$.

2) The direct sum

$$\mathfrak{b} = \bigoplus_{n=0}^{\infty} \mathfrak{b}_n$$

is a graduated ideal ([P1], section 2.3.12) of the graduated algebra $\mathbf{T}(\mathbf{E})$*, so* $\mathbf{TS}(\mathbf{E}) \cong \mathbf{T}(\mathbf{E}) / \mathfrak{b}$ *is a graduated* $\mathbf{K}$*-algebra, called the* symmetric algebra *of* $\mathbf{E}$.

3) Proceeding in the same way as in Lemma-Definition 4.1, we can define a covariant functor $\mathbf{TS}$ *from the category of* $\mathbf{K}$*-vector spaces into the category of* *commutative graduated* $\mathbf{K}$*-algebras. The canonical image of* $(x_1, ..., x_n) \in \mathbf{E}^n$ *in* $\mathbf{T}^n(\mathbf{E})$ *is written as* $x_1...x_n$.

4) $\mathbf{TS}(\mathbf{E})$ *is isomorphic to the algebra* $\mathbf{K}[X_1, ..., X_m]$ *of polynomials in* m *variables over* $\mathbf{K}$ *(where* $m := \dim_{\mathbf{K}}(\mathbf{E})$*).*

DEFINITION 4.5.– *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be finite-dimensional* $\mathbf{K}$*-vector spaces. A mapping* $f : \mathbf{E} \to \mathbf{F}$ *is said to be a* homogeneous polynomial *of degree* n *if there exists an* n*-linear mapping* $u : \mathbf{E}^n \to \mathbf{F}$ *such that* $f(x) = u(x, ..., x)$.

LEMMA 4.6.– *i) If* $f : \mathbf{E} \to \mathbf{F}$ *is a homogeneous polynomial, then the* n*-linear mapping* $u : \mathbf{E}^n \to \mathbf{F}$ *defined by* $f(x) = u(x, ..., x)$ *can be chosen to be symmetric, and this choice is unique.*

ii) Let $f : \mathbf{E} \to \mathbf{F}$ *be a homogeneous polynomial mapping of degree* n *and consider* $\gamma_n : \mathbf{E}^n \to \mathbf{T}^n(\mathbf{E}) : x \mapsto x \otimes ... \otimes x$*. There exists a unique* $\mathbf{K}$*-linear mapping* $\tilde{f} : \mathbf{T}^n(\mathbf{E}) \to \mathbf{F}$ *such that* $f = \tilde{f} \circ \gamma_n$.

PROOF.– (i) It can be checked that

$$u(x_1, ..., x_n) = \frac{(-1)^n}{n!} \sum_{H \subset \{1,2,...,n\}} (-1)^{\text{card}(H)} f\left(\sum_{i \in H} x_i\right).$$

ii) By the universality of the tensor product ([P1], section 3.1.5**(I)**), there exists a unique $\mathbf{K}$-linear mapping f' from $\mathbf{TS}^n(\mathbf{E})$ into $\mathbf{F}$ such that $u(x_1, ..., x_n) = f'(x_1 \otimes ... \otimes x_n)$ for all $x_i \in \mathbf{E}$. Thus:

$$f(x) = u(x, ..., x) = f'(x \otimes ... \otimes x) = f'(\gamma_n(x)),$$

which gives that $\tilde{f} = f' \left|_{\mathbf{TS}^n(\mathbf{E})}\right.$. ■

(II) CASE OF COVARIANT TENSORS We saw above that $\mathbf{T}_n^0(\mathbf{E})$ can be identified with the dual of $\mathbf{T}_0^n(\mathbf{E})$ (section 4.2.1**(II)**). By [4.1], it follows that:

$$\langle \sigma.\boldsymbol{t}, \boldsymbol{t}^\vee \rangle = \left\langle \boldsymbol{t}, \sigma^{-1}.\boldsymbol{t}^\vee \right\rangle.$$

Furthermore, $\mathbf{T}_n^0(\mathbf{E})$ can be identified with the space of n-multilinear forms on $\mathbf{E}$; thus, the symmetric (respectively antisymmetric) covariant tensors can be identified with the symmetric (respectively alternating) multilinear forms on $\mathbf{E}$.

4.2.3. *Exterior algebra*

(I) EXTERIOR n-TH POWER Let $\mathbf{E}$ be an m-dimensional $\mathbf{K}$-vector space and let $z_p \in \mathbf{A}^p(\mathbf{E})$, $z_q \in \mathbf{A}^q(\mathbf{E})$ (section 4.2.2(**I**)). The *exterior* or *wedge product* ([P1], section 3.3.8(**VII**)) $z_p \wedge z_q \in \mathbf{A}^{p+q}(\mathbf{E})$ is defined by:

$$\boxed{z_p \wedge z_q = \tfrac{1}{p!q!}\mathbf{a}.\left(z_p \otimes z_q\right) = \tfrac{(p+q)!}{p!q!}\mathrm{alt}\left(z_p \otimes z_q\right).} \qquad [4.7]$$

This product is associative and anticommutative, i.e.

$$z_p \wedge z_q = \left(-1^{pq}\right) z_q \wedge z_p.$$

In particular, if $x, y \in \mathbf{E}$, then $x \wedge y = \mathbf{a}.\left(x \otimes y\right) = x \otimes y - y \otimes x$. If $\mathbf{T}$ is a twice covariant antisymmetric tensor, then

$$\mathbf{T}\left(x \wedge y\right) = \mathbf{T}\left(x \otimes y\right) - \mathbf{T}\left(y \otimes x\right) = 2\mathbf{T}\left(x \otimes y\right). \qquad [4.8]$$

More generally, if $z_{p_i} \in \mathbf{A}^{p_i}(\mathbf{E})$ $(i = 1, ..., k)$, then

$$z_{p_1} \wedge ... \wedge z_{p_k} = \frac{1}{p_1!...p_k!}\mathbf{a}.\left(z_{p_1} \otimes z_{p_2} \otimes ... \otimes z_{p_k}\right).$$

Applying the above to *vectors* $z_{p_i} = x_i \in \mathbf{E}$ gives

$$x_{\sigma(1)} \wedge x_{\sigma(2)} \wedge ... \wedge x_{\sigma(n)} = \varepsilon_\sigma . x_1 \wedge x_2 \wedge ... \wedge x_n. \qquad [4.9]$$

As a result of [4.9], the vector space $\mathbf{A}^n(\mathbf{E})$ is also called the exterior n-th power of $\mathbf{E}$ and can be written as $\bigwedge^n \mathbf{E}$. The basis of this space associated with the basis $(e_i)_{1 \le i \le m}$ consists of the $\binom{m}{n}$ elements:

$$e_H = e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_n},$$

where the index sets H are defined as above, so that $H = \{i_1, ..., i_n\} \subset \{1, ..., m\}$, $i_1 < i_2 < ... < i_n$. Hence,

$$\dim\left(\bigwedge^n \mathbf{E}\right) = \begin{cases} \binom{m}{n} \text{ if } n \le m, \\ 0 \text{ if } n > m. \end{cases} \qquad [4.10]$$

In particular, $\dim\left(\bigwedge^n \mathbf{E}\right) = 1$ if $n = m$ and $\bigwedge^n \mathbf{E} = 0$ if $n > m$.

DEFINITION 4.7.– *The elements of* $\overset{p}{\bigwedge}\mathbf{E} := \mathbf{A}^p(\mathbf{E})$ *are said to be* p-vectors.

We write that $\det(\mathbf{E}) = \overset{m}{\bigwedge}\mathbf{E}$ (and so $\det(\mathbf{E}) \cong \mathbf{K}$).

(II) CALCULATING AN n-VECTOR IN A GIVEN BASIS Let $x_i = \sum_{j=1}^{m} \xi_i^j e_j$; explicitly writing out the expression from [P1], section 3.3.8**(VII)** gives:

$$x_1 \wedge ... \wedge x_n = \sum_{i_1 < i_2 < ... < i_n} \begin{vmatrix} \xi_1^{j_1} & \xi_1^{j_2} & \cdots & \xi_1^{j_n} \\ \xi_2^{j_1} & \xi_2^{j_2} & \cdots & \xi_2^{j_n} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_n^{j_1} & \xi_n^{j_2} & \cdots & \xi_n^{j_n} \end{vmatrix} .e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_n}. \quad [4.11]$$

(III) EXTERIOR ALGEBRA Let $\overset{0}{\bigwedge}\mathbf{E} = \mathbf{K}$ and consider the direct sum

$$\bigwedge \mathbf{E} = \bigoplus_{0 \le p \le m} \overset{p}{\bigwedge}\mathbf{E}.$$

This sum has an associative $\mathbf{K}$-algebra structure (for the exterior product) and is called the *exterior algebra* of $\mathbf{E}$. The exterior algebra $\bigwedge\mathbf{E}$ is therefore clearly an anticommutative graduated $\mathbf{K}$-algebra ([P1], sections 2.3.10**(I)** and 2.3.12), and the elements of $\overset{p}{\bigwedge}\mathbf{E}$ are the *homogeneous elements of degree* p.

4.2.4. *Duality in the exterior algebra*

We can repeat the above for covariant tensors and covectors. This gives

$$\overset{n}{\bigwedge}\mathbf{E}^\vee = \left(\overset{n}{\bigwedge}\mathbf{E}\right)^\vee$$

with the duality bracket ([P1], section 3.3.8**(VII)**)

$$\left\langle x_1 \wedge x_2 \wedge ... \wedge x_n, x^{\vee 1} \wedge x^{\vee 2} \wedge ... \wedge x^{\vee n} \right\rangle = \det\left(\left\langle x_j, x^{\vee i}\right\rangle\right) \quad [4.12]$$

and the following result:

LEMMA 4.8.– *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be finite-dimensional* $\mathbf{K}$*-vector spaces.*

1) The elements of $\overset{n}{\bigwedge}\mathbf{E}^\vee$ *can be identified with the alternating* n*-linear forms on* $\mathbf{E}$*, i.e. (with the same notation as above) the* n*-linear forms* $u : \mathbf{E}^n \to \mathbf{K}$ *such that, for every permutation* $\sigma \in \mathfrak{S}_n$ *and every* $(x_1, ..., x_n) \in \mathbf{E}^n$*,*

$$\sigma(u)(x_1, ..., x_n) = \varepsilon_\sigma . u(x_1, ..., x_n), \quad [4.13]$$

where $\sigma(u)(x_1,...,x_n) := u\left(x_{\sigma(1)},...,x_{\sigma(n)}\right).$

2) $\bigwedge^n \mathbf{E}$ *satisfies the following* universal property *([BOU 12], Chapter III, section 7.4, Proposition 7): For every alternating* n*-linear form* $u : \mathbf{E}^n \to \mathbf{F}$*, there exists a unique* $\mathbf{K}$*-linear mapping* $v : \bigwedge^n \mathbf{E} \to \mathbf{F}$ *such that*

$$u(x_1, x_2, ..., x_n) = v(x_1 \wedge x_2 \wedge ... \wedge x_n) = \langle x_1 \wedge x_2 \wedge ... \wedge x_n, v\rangle .$$

3) In particular, suppose that $\dim(\mathbf{E}) = m$*. Then,* $u : \mathbf{E}^n \to \mathbf{F}$ *is an alternating* n*-linear form if and only if it can be expressed in the form*

$$u(x_1, x_2, ..., x_n) = \sum_H a_H . \det\left(X^H\right),$$

where $H = \{j_1, ..., j_n\} \subset \{1, ..., m\}$*,* $1 \leq j_1 < ... < j_n$*,* X^H *is the square matrix* $\left(\eta_h^k\right)$ *of order* n *satisfying* $\eta_h^k = \xi_h^{j_k}$ *for* $1 \leq h, k \leq n$ *(see section 4.2.3(**II**)), and* $a_H \in \mathbf{F}$.

DEFINITION 4.9.– *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be finite-dimensional* $\mathbf{K}$*-vector spaces. Write* $\mathcal{A}_n(\mathbf{E}; \mathbf{F})$ *for the vector space of alternating* n*-linear mappings (i.e. which satisfy [4.13]) from* $\mathbf{E}^n$ *into* $\mathbf{F}$.

In particular, $\mathcal{A}_p(\mathbf{E}) := \mathcal{A}_p(\mathbf{E}; \mathbf{K}) = \bigwedge^p \mathbf{E}^\vee$ *is the space of* p-covectors *or* p-forms *([P1], section 3.3.8(**VII**)).*

LEMMA-DEFINITION 4.10.– *i) Let* u *(respectively* v*) be an alternating* p*-linear (respectively* q*-linear) form on* $\mathbf{E}^p$ *(respectively* $\mathbf{E}^q$*). The exterior product* $u \wedge v$ *is defined as the* antisymmetrization *of the* $(p+q)$*-linear form*

$$(x_1, ..., x_p, x_{p+1}, ..., x_{p+q}) \mapsto \frac{1}{p!q!} u(x_1, ..., x_p) . v(x_{p+1}, ..., x_{p+q}),$$

and $v \wedge u = (-1)^{pq}\, u \wedge v$*. Equivalently,*

$$(u \wedge v)(x_1, ..., x_{p+q}) = \sum_{\sigma \in Sh(p,q)} \varepsilon(\sigma)\, u\left(x_{\sigma(1)}, ..., x_{\sigma(p)}\right) . v\left(x_{\sigma(p+1)}, ..., x_{\sigma(p+q)}\right),$$

where $Sh(p,q)$ *is the set of permutations* $\sigma \in \mathfrak{S}_{p+q}$ *such that* $\sigma(1) < ... < \sigma(p)$ *and* $\sigma(p+1) < ... < \sigma(p+q)$*.*[4]

ii) The exterior product of alternating multilinear forms is an associative operation.

4 The notation Sh is short for *shuffle*.

iii) If $u_1, ..., u_n$ are linear forms, then $u_1 \wedge ... \wedge u_n = \mathbf{a}. \prod_{i=1}^n u_i$, or more explicitly:

$$(u_1 \wedge ... \wedge u_n)(x_1, ..., x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n u_i\left(x_{\sigma(i)}\right) = \det_{1 \leq i,j \leq n}(u_i(x_j)).$$

By Lemma 4.8(2),

$$(u_1 \wedge ... \wedge u_n)(x_1, ..., x_n) = \langle x_1 \wedge ... \wedge x_n, u_1 \wedge ... \wedge u_n \rangle.$$

If $\omega^i = \sum_{j=1}^m a_i^j e^{\vee i}$ $(i = 1, ..., n)$, then we have the following relation, similar to [4.11]:

$$\omega^1 \wedge ... \wedge \omega^n = \sum_{j_1 < j_2 < ... < j_n} \begin{vmatrix} a_1^{j_1} & a_1^{j_2} & \cdots & a_1^{j_n} \\ a_2^{j_1} & a_2^{j_2} & \cdots & a_2^{j_n} \\ \cdots & \cdots & \cdots & \cdots \\ a_n^{j_1} & a_n^{j_2} & \cdots & a_n^{j_n} \end{vmatrix} e^{\vee j_1} \wedge ... \wedge e^{\vee j_n}.$$

REMARK 4.11.– *Some authors ([KOB 69], Volume I, Chapter I, p. 35) define the exterior product of an alternating linear p-form u and an alternating linear q-form v as the alternating linear $(p+q)$-form*

$$(x_1, ..., x_p, x_{p+1}, ..., x_{p+q}) \mapsto \frac{1}{(p+q)!} \sum_{\sigma \in \mathfrak{S}_{p+q}} \varepsilon(\sigma) u\left(x_{j_1}, ..., x_{j_p}\right).v\left(x_{k_1}, ..., x_{k_q}\right).$$

We will instead write $u \mathbin{\bar{\wedge}} v$ for this exterior product, which is equal to $\frac{p!q!}{(p+q)!} u \wedge v$ in our notation. In particular, if $u_1, ..., u_n$ are linear forms, then $(u_1 \mathbin{\bar{\wedge}} ... \mathbin{\bar{\wedge}} u_n)(x_1, ..., x_n) = \frac{1}{n!} \det_{1 \leq i,j \leq n}(u_i(x_j))$ ([KOB 69], Volume I, Chapter I, p. 7).

4.2.5. *Interior products*

Let p, q be integers ≥ 0 and z_q a contravariant tensor in $\mathbf{T}_0^q(\mathbf{E})$, where $\mathbf{E}$ is an m-dimensional $\mathbf{K}$-vector space. The mapping $v_p \mapsto z_q \otimes v_p$ from $\mathbf{T}_0^p(\mathbf{E})$ into $\mathbf{T}_0^{p+q}(\mathbf{E})$ is $\mathbf{K}$-linear. Its transpose may therefore be identified with a $\mathbf{K}$-linear mapping from $\mathbf{T}_0^{p+q}(\mathbf{E}^\vee)$ into $\mathbf{T}_0^p(\mathbf{E}^\vee)$, written as (with $u_{p+q} \in \mathbf{T}_0^{p+q}(\mathbf{E}^\vee)$)

$$u_{p+q} \mapsto z_q \lrcorner u_{p+q},$$

which is said to be the *interior product* of z_q and u_{p+q}. Thus, by definition:

$$\langle v_p, z_q \lrcorner u_{p+q} \rangle = \langle v_p \otimes z_q, u_{p+q} \rangle.$$

For example, if ω is a p-covector and $v_1, ..., v_p$ are *vectors*, then

$$(v_1 \lrcorner \omega)(v_2, ..., v_p) = \omega(v_1, v_2, ..., v_p). \qquad [4.14]$$

If ψ is also a n-covector and v is a vector, this gives the "distributivity relation":

$$v \lrcorner (\omega \wedge \psi) = (v \lrcorner \omega) \wedge \psi + (-1)^p \omega \wedge (v \lrcorner \psi). \qquad [4.15]$$

If $\omega_1, ..., \omega_p$ are covectors,

$$v \lrcorner (\omega_1 \wedge ... \wedge \omega_p) = \sum_{i=1}^{p} (-1)^{i+1} \langle v, \omega_i \rangle \omega_1 \wedge ... \wedge \widehat{\omega_i} \wedge .... \wedge \omega_p. \qquad [4.16]$$

REMARK 4.12.– *If we write $i_v(u_p) := v \lrcorner u_p$, then [4.15] becomes:*

$$i_v(\omega \wedge \psi) = i_v(\omega) \wedge \psi + (-1)^p \omega \wedge i_v(\psi).$$

The interior product is an antiderivation of degree -1 of the exterior algebra $\bigwedge \mathbf{E}^\vee = \bigoplus_{1 \leq p \leq m} \bigwedge^p \mathbf{E}^\vee$ ([P1], section 2.3.12).

EXAMPLE 4.13.– *i) Suppose that $q = 1$, $p = 0$. Now, let $z \in \mathbf{E} = \mathbf{T}_0^1(\mathbf{E})$, $u \in \mathbf{E}^\vee = \mathbf{T}_1^0(\mathbf{E})$. Then, $z \lrcorner u \in \mathbf{K}$ and, for all $\lambda \in \mathbf{K}$, $\lambda . z \lrcorner u = \langle \lambda \otimes z, u \rangle = \lambda . \langle z, u \rangle$, so $z \lrcorner u = \langle z, u \rangle$. The interior product is therefore a generalization of the duality bracket.*

ii) In general, let

$$z_q = \sum_{i_1, \ldots, i_q} z^{i_1 \ldots i_q} e_{i_1} \otimes ... \otimes e_{i_q}, \quad u_{p+q} = \sum_{i_1, \ldots, i_{p+q}} u_{i_1 \ldots i_{p+q}} e^{\vee i_1} \otimes ... \otimes e^{\vee i_{p+q}}.$$

It can immediately be checked that

$$z_q \lrcorner u_{p+q} = \sum_{i_1, \ldots, i_{p+q}} z^{i_1 \ldots i_q} u_{i_1 \ldots i_{p+q}} e^{\vee i_{q+1}} \otimes ... \otimes e^{\vee i_{p+q}}.$$

iii) In the case of antisymmetric tensors*, let $z_q \in \bigwedge^q \mathbf{E}$, $u_{p+q} \in \bigwedge^{p+q}(\mathbf{E}^\vee)$. Define $z_q \lrcorner u_{p+q} \in \bigwedge^p (\mathbf{E}^\vee)$ by the equality*

$$\langle v_p, z_q \lrcorner u_{p+q} \rangle = \langle v_p \wedge z_q, u_{p+q} \rangle$$

for all $v_p \in \bigwedge^p \mathbf{E}$.

iv) In particular, suppose that $\dim(\mathbf{E}) = m \geq 3$, $q = m-1$, $p = 1$, $u_m = e^{\vee i_1} \otimes ... \otimes e^{\vee i_m}$, $z_{m-1} = x_1 \wedge ... \wedge x_{m-1}$, $v_p = x_0$, *where* $x_i \in \mathbf{E}$ $(0 \leq i \leq m-1)$. *Then:*

$$\langle x_0, (x_1 \wedge ... \wedge x_{m-1}) \lrcorner u_m \rangle = \det\left(\left\langle x_j, e^{\vee i}\right\rangle\right)_{\substack{1 \leq i \leq m \\ 0 \leq j \leq m-1}} \in \mathbf{K}$$

and $x_1 \wedge ... \wedge x_{m-1} \in \mathbf{E}^{\vee}$. *If* $\mathbf{E}$ *is a Euclidean space over the field of real numbers, the scalar product associates this linear form with a vector of* $\mathbf{E}$ *(this is a trivial special case of Riesz's theorem ([P2], section 3.10.2)). This vector is again written as* $x_1 \wedge ... \wedge x_{m-1}$ *and called the* vector product *of the vectors* $x_1, ..., x_{m-1}$. *The vector product does not just depend on the Euclidean structure of* $\mathbf{E}$, *but also on its* orientation *(see below, section 4.4.4).*

In the case where $m = 3$, *the vector product of two vectors can be identified with a vector. By [4.15], the equality [4.14] therefore implies the classical relation:*

$$x \wedge (y \wedge z) = (x.z)\, y - (x.y)\, z. \qquad [4.17]$$

4.2.6. *Tensors on Banach spaces*

(I) TENSORS Let $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ be subcategories of the category of Banach $\mathbb{K}$-spaces ([P2], section 3.4.1(**I**)) and let $\lambda : \mathfrak{A}^p \times \mathfrak{B}^q \to \mathfrak{C}$ be a functor that is covariant in $\mathfrak{A}^p$ and contravariant in $\mathfrak{B}^q$ ([P1], sections 1.1.1(**I**) and 1.2.1).

EXAMPLE 4.14.– *1) If* $\mathbf{E}$ *is a finite-dimensional* $\mathbb{K}$*-vector space, then* $\mu : \mathbf{E}^p \mapsto \mathbf{T}_0^p(\mathbf{E})$ *is a covariant functor,* $\nu : \mathbf{E}^q \mapsto \mathbf{T}_q^0(\mathbf{E})$ *is a contravariant functor and* $\lambda : \mathbf{E}^p \times \mathbf{E}^q \mapsto \mathbf{T}_q^p(\mathbf{E})$ *is a functor that is covariant in* $\mathbf{E}^p$ *and contravariant in* $\mathbf{E}^q$.

2) If $p = 1$, $\mathfrak{B}_1 = ... = \mathfrak{B}_q$, $\mathbf{E} \in \mathfrak{A}_1$, *and* $\mathbf{E}' \in \mathfrak{B}_1$,[5] *we can define a functor covariant in* $\mathbf{E}$ *and contravariant in* $\mathbf{E}'^q$ *by:*

$$\mathbf{T}_q : (\mathbf{E}', \mathbf{E}) \mapsto \mathcal{L}_q(\mathbf{E}'; \mathbf{E}),$$

where $\mathcal{L}_q(\mathbf{E}'; \mathbf{E})$ *denotes the Banach space of continuous q-linear mappings from* $\mathbf{E}'^q$ *into* $\mathbf{E}$ *([P2], section 3.8.9). This functor acts as follows on morphisms: let* $(f, f') : (\mathbf{E}_1, \mathbf{E}_1') \to (\mathbf{E}_2, \mathbf{E}_2')$ *(where* $f \in \mathcal{L}(\mathbf{E}_1; \mathbf{E}_2)$ *and* $f' \in \mathcal{L}(\mathbf{E}_1'; \mathbf{E}_2')$*); then*

$$\mathbf{T}_q(f, f') : \mathcal{L}_q(\mathbf{E}_2'; \mathbf{E}_1) \to \mathcal{L}_q(\mathbf{E}_1'; \mathbf{E}_2) : u \mapsto f \circ u \circ f'^q,$$

where $f'^q = (f', ..., f')$ *(q terms).*

3) With the same conditions as (2), let $Alt^q(\mathbf{E}'; \mathbf{E})$ *be the closed subspace of* $\mathcal{L}_q(\mathbf{E}'; \mathbf{E})$ *formed by the* antisymmetric *q-linear mappings, i.e. those which satisfy*

5 Recall that the dual of $\mathbf{E}$ is written as $\mathbf{E}^{\vee}$ rather than $\mathbf{E}'$, which for us denotes an arbitrary Banach space in $\mathfrak{B}_1$.

[4.13] for all $\sigma \in \mathfrak{S}_q$ *and every* $\left(x'_1, ..., x'_q\right) \in \mathbf{E}'^q$*, generalizing Definition 4.9. Write* $Alt^q\left(\mathbf{E}';\mathbb{K}\right) = Alt^q\left(\mathbf{E}'\right)$. *Now, define the following functor, covariant in* $\mathbf{E}$ *and contravariant in* $\mathbf{E}'$*:*

$$Alt^q : \left(\mathbf{E}', \mathbf{E}\right) \mapsto Alt^q\left(\mathbf{E}';\mathbf{E}\right).$$

This functor acts as follows on morphisms: let $(f, f') : \left(\mathbf{E}_1, \mathbf{E}'_1\right) \rightarrow \left(\mathbf{E}_2, \mathbf{E}'_2\right)$ *(where* $f \in \mathcal{L}\left(\mathbf{E}_1;\mathbf{E}_2\right)$ *and* $f' \in \mathcal{L}\left(\mathbf{E}'_1;\mathbf{E}'_2\right)$*); then*

$$Alt^q\left(f, f'\right) : Alt^q\left(\mathbf{E}'_2;\mathbf{E}_1\right) \rightarrow Alt^q\left(\mathbf{E}'_1;\mathbf{E}_2\right) : u \mapsto f \circ u \circ f'^q.$$

For every $\mathbf{E}, \mathbf{F} \in \mathfrak{A}$, $\mathbf{E}', \mathbf{F}' \in \mathfrak{B}$, with $\mathbf{E} = \mathbf{E}_1 \times ... \times \mathbf{E}_p$, $\mathbf{E}' = \mathbf{E}'_1 \times ... \times \mathbf{E}'_q$, and similarly for $\mathbf{F}$ and $\mathbf{F}'$, set

$$\operatorname{Hom}\left(\mathbf{E} \times \mathbf{E}';\mathbf{F} \times \mathbf{F}'\right) = \left(\prod\nolimits_{1 \leq i \leq p} \mathcal{L}\left(\mathbf{E}_i;\mathbf{F}_i\right)\right) \times \left(\prod\nolimits_{1 \leq i \leq p} \mathcal{L}\left(\mathbf{F}'_i;\mathbf{E}'_i\right)\right).$$

DEFINITION 4.15.– *The functor* λ *is said to be a* vector functor of class C^s*,* $s \in \mathbb{N}_{\mathbb{K}}$ *(section 1.2.1(**I**)) if the following conditions are satisfied:*

i) For every pair $\left(\mathbf{E}, \mathbf{E}'\right) \in \mathfrak{A} \times \mathfrak{B}$*,* $\lambda\left(\mathbf{E}, \mathbf{E}'\right) \in \mathfrak{C}$*.*

ii) For every $\mathbf{f} \in Hom\left(\mathbf{E} \times \mathbf{E}';\mathbf{F} \times \mathbf{F}'\right)$*,* $\lambda(\mathbf{f}) \in \mathcal{L}\left(\lambda\left(\mathbf{E}, \mathbf{E}'\right);\lambda\left(\mathbf{F}, \mathbf{F}'\right)\right)$*.*

iii) The mapping $\lambda : Hom\left(\mathbf{E} \times \mathbf{E}';\mathbf{F} \times \mathbf{F}'\right) \mapsto \mathcal{L}\left(\lambda\left(\mathbf{E}, \mathbf{E}'\right);\lambda\left(\mathbf{F}, \mathbf{F}'\right)\right)$ *is of class* C^s*.*

DEFINITION 4.16.– *Let* B *be a manifold (i.e. a Banach manifold of class* C^r*,* $r \geq \infty$*, by the conventions **(C1)**, **(C2)**),* $\mathbf{E} \in \mathfrak{A}$*,* $\mathbf{E}' \in \mathfrak{B}$*, and* $\lambda : \mathfrak{A}^p \times \mathfrak{B}^q \rightarrow \mathfrak{C}$ *a vector functor of class* C^r *that is covariant in the first variable and contravariant in the second. A* tensor *of type* (p, q) *is an element* $\lambda\left(\mathbf{E}^p, \mathbf{E}'^q\right) \in \mathfrak{C}$*.*

REMARK 4.17.– *1) The functor* $\lambda : \mathbf{E}^p \times \mathbf{E}^q \mapsto \mathbf{T}^p_q(\mathbf{E})$ *is* covariant in $\mathbf{E}^p$ and contravariant in $\mathbf{E}^q$*, but* $\mathbf{T}^p_q(\mathbf{E})$ *is the space of* q-times covariant and p-times contravariant tensors! *See Remark 4.3.*

2) We do not know whether the existence of λ *(Definition 4.15) can be established without restrictions for* $s = r$*, even for the case* $p = 1$*. However, it always holds for* $s = 0$.

(II) ANTISYMMETRIC CONTINUOUS MULTILINEAR MAPPINGS Let $\mathbf{E}, \mathbf{F}$ be two Banach spaces. The space $\operatorname{Alt}^q(\mathbf{E};\mathbf{F})$ of antisymmetric continuous q-linear mappings from $\mathbf{E}^q$ into $\mathbf{F}$ is defined as in Example 4.14(3). $\operatorname{Alt}^q(\mathbf{E};\mathbf{F})$ is a closed subspace of $\mathcal{L}_q(\mathbf{E};\mathbf{F})$ and hence a Banach space.

Let $\mathbf{G}, \mathbf{H}$ be two other Banach spaces and $\Phi : \mathbf{F} \times \mathbf{G} \to \mathbf{H}$ a continuous bilinear mapping; for every $(y, z) \in \mathbf{F} \times \mathbf{G}$, write $\Phi(y, z) = y.z$. Let $u \in \mathrm{Alt}^p(\mathbf{E}; \mathbf{F})$ and $v \in \mathrm{Alt}^q(\mathbf{E}; \mathbf{G})$. Then, the mapping

$$h : \mathbf{E}^{p+q} \to \mathbf{H} : (x_1, ..., x_p, x_{p+1}, ...x_{p+q}) \mapsto \frac{1}{p!q!} u(x_1, ..., x_p) .v(x_{p+1}, ...x_{p+q})$$

is continuous and $(p+q)$-linear. The statement of Lemma-Definition 4.10 remains valid in this more general context (note that the linear and multilinear mappings must be continuous), and $u \wedge v \in \mathrm{Alt}^{p+q}(\mathbf{E}; \mathbf{H})$.

REMARK 4.18.– *The exterior product $\wedge$ defined above depends on the continuous bilinear mapping Φ. We can use the notation $\wedge_\Phi$ to emphasize this more explicitly. In the case where $\mathbf{F} = \mathbf{G} = \mathbf{H} = \mathbf{K}$, Φ is the "canonical" coupling $(y, z) \mapsto y.z$ (usual product) and can be left implicit.*

The interior product $\mathbf{v}_1 \lrcorner \omega$ of $\mathbf{v}_1 \in \mathbf{E}$ and $\omega \in \mathrm{Alt}^p(\mathbf{E}; \mathbf{F})$ is defined by [4.14] and belongs to $\mathrm{Alt}^{p-1}(\mathbf{E}; \mathbf{F})$. The relations [4.15] and [4.16] still hold, as well as Remark 4.12.

4.3. Tensor fields

Let B be a manifold and U an open subset of B. Write $\mathcal{C}^r(U)$ for the $\mathbb{K}$-algebra of mappings of class C^r from U into $\mathbb{K}$.

4.3.1. *Vector fields*

DEFINITION 4.19.– *A* vector field *of class C^r on an open subset U of B is a morphic section of $T(B)$ over U (see Remark 3.24), i.e. an element of $\Gamma(U, T(B))$ (Corollary-Definition 3.21). The latter space is also written as $\mathcal{T}_0^1(U)$ and is a $\mathcal{C}^r(U)$-module.*

Let $(U, \xi, \mathbf{E})$ be a chart of B centered on some point $a \in B$, where $\mathbf{E}$ is a Banach space. The restriction to U of a vector field $X \in \mathcal{T}_0^1(B)$ is a mapping $X|_U : U \to T(B)$ such that $\pi \circ v = 1_U$, so

$$X|_U : b \mapsto (b, \mathbf{h}_b) \in T_b(U)$$

is a vector field on U. We already know that there exists an isomorphism [3.2]

$$\psi_c : U \times \mathbf{F} \to \pi^{-1}(U) : (b, \mathbf{h}) \mapsto \left(b, (d_b\xi)^{-1} .\mathbf{h}\right) = (b, \mathbf{h}_b).$$

4.3.2. *Covector field*

DEFINITION 4.20.– *A* covector field *of class C^r on U is a morphic section of $T^{\vee}(B)$ over U, i.e. an element of $\Gamma(U, T^{\vee}(B))$. The latter space can alternatively be written as $\mathcal{T}_1^0(U)$ or $\Omega^1(U)$. The elements of $\Omega^1(U)$ are also known as* Pfaff forms *on U.*

Let $(U, \xi, \mathbf{E})$ be a chart of B centered on a point $b \in B$. The restriction to U of any such covector field ω is a mapping $\omega|_U : U \to T^{\vee}(M)$ satisfying $\pi \circ \omega = 1_U$, so

$$\omega|_U : b \mapsto (b, \mathbf{h}_b^{\vee}) \in T_b^{\vee}(U),$$

is a covector field on U. (For the finite-dimensional case, see Theorem 2.71.)

4.3.3. *Tensor fields and scalar fields*

(I) TENSOR BUNDLE Let $\pi : M \to B$ be a vector bundle. Write $\mathbf{T}_q^p(M)$ for the vector bundle of class C^r whose fiber over an arbitrary point $b \in B$ is $\mathbf{T}_q^p(M_b)$ (sections 4.2.1 and 4.2.6). This is called the bundle of p-times contravariant and q-times covariant tensors (or tensors of type (p, q)) on M. If M is of finite rank, then $\mathbf{T}_q^p(M) = (M^{\vee})^{\otimes q} \otimes (M)^{\otimes p}$. By convention, $\mathbf{T}_0^1(M) = M$, $\mathbf{T}_1^0(M) = M^{\vee}$ (Definition 3.36) and $\mathbf{T}_0^0(M) = B \times \mathbb{K}$. The following result is clear:

LEMMA 4.21.– *Suppose that M is of rank n. Let $(\mathbf{s}_i)_{1 \le i \le n}$ be a frame of M over the open subset U of B (Lemma-Definition 3.23) and $\left(\mathbf{s}^{\vee i}\right)_{1 \le i \le n}$ its dual frame (Lemma-Definition 3.38). Then, the n^{p+q} tensor fields*

$$\mathbf{s}_{j_1} \otimes ... \otimes \mathbf{s}_{j_p} \otimes \mathbf{s}^{\vee i_1} \otimes ... \otimes \mathbf{s}^{\vee i_q}$$

form a frame of $\mathbf{T}_q^p(M)$.

(II) TENSOR PRODUCT, WHITNEY SUM AND EXTERIOR PRODUCT The constructions of section 3.4.4 can be extended to the tensor product and Whitney sum of an *arbitrary finite number* of vector bundles with the same base; in particular, we can define the tensor power $M^{n\otimes}$ of a vector bundle M with base B and *finite rank* (section 3.4.1). We can also define the tensor field $\mathbf{T}_0^n(M)$ whose fibers are $\mathbf{T}_0^n(M)_b$ $(b \in B)$.

Furthermore, the same construction enables us to define the exterior power

$$\bigwedge^p M$$

of a vector bundle M of finite rank (with $p > 0$). For every sequence $(\mathbf{s}_j)_{1 \le j \le p}$ of sections of M on an open subset U of B, write $\mathbf{s}_1 \wedge \mathbf{s}_2 \wedge ... \wedge \mathbf{s}_p$ for the mapping

$$U \ni b \mapsto \mathbf{s}_1(b) \wedge \mathbf{s}_2(b) \wedge ... \wedge \mathbf{s}_p(b) \in \bigwedge^p M_b.$$

REMARK 4.22.– *By convention,* $M^{\otimes 0} = \bigwedge^{0} M$ *is the trivial bundle* $B \times \mathbb{K}$*.*

Consider a morphism of vector bundles $u : M \to N$, where M and N are vector bundles with base B and finite rank. There exists a unique morphism

$$\bigwedge^{p} u : \bigwedge^{p} M \to \bigwedge^{p} N$$

such that, if $\mathbf{s}_1, ..., \mathbf{s}_p$ are arbitrary sections of M over the open set $U \subset B$, then

$$\bigwedge^{p} u \circ (\mathbf{s}_1 \wedge ... \wedge \mathbf{s}_p) = (u \circ \mathbf{s}_1) \wedge ... \wedge (u \circ \mathbf{s}_p) .$$

(III) TENSOR FIELDS Let B be a manifold, $T(B)$ its tangent bundle and $T^{\vee}(B)$ its cotangent bundle. Write $\mathbf{T}_q^p(B)$ for $\mathbf{T}_q^p(T(B))$ if this is not ambiguous. Thus, $\mathbf{T}_0^1(B) = T(B)$ and $\mathbf{T}_1^0(B) = T^{\vee}(B)$.

DEFINITION 4.23.– *A* tensor field *of type* (p, q) *and class* C^r *on an open subset* U *of* B *is a morphic section of* $\mathbf{T}_q^p(B)$ *over* U*, i.e. an element of* $\Gamma\left(U, \mathbf{T}_q^p(B)\right)$*. This is a mapping* $T : b \mapsto T(b)$*, where* $b \in U$*,* $T(b) \in \mathbf{T}_q^p(B)_b$*. The space* $\Gamma\left(U, \mathbf{T}_q^p(B)\right)$ *can also be written as* $\mathcal{T}_q^p(U)$ *(according to the notation of Definitions 4.19 and 4.20).*

Let (U, ξ, n) be a chart of B. Every tensor field $\mathbf{Z} \in \mathcal{T}_q^p(U)$ can be uniquely written in the form

$$\mathbf{Z} = \sum z_{i_1 \dots i_q}^{j_1 \dots j_p} \frac{\partial}{\partial \xi^{j_1}} \otimes ... \otimes \frac{\partial}{\partial \xi^{j_p}} \otimes d\xi^{i_1} \otimes ... \otimes d\xi^{i_q}, \qquad [4.18]$$

where the components $z_{i_1 \dots i_q}^{j_1 \dots j_p}$ are of class C^r. Hence, $\mathcal{T}_q^p(U)$ is a free $\mathcal{C}^r(U)$-module with basis

$$\left\{ \frac{\partial}{\partial \xi^{j_1}} \otimes ... \otimes \frac{\partial}{\partial \xi^{j_p}} \otimes d\xi^{i_1} \otimes ... \otimes d\xi^{i_q} : j_1 < ... < j_p, i_1 < ... < i_q \right\}.$$

More generally, if U is an arbitrary non-empty open subset of B, then the set $\mathcal{T}_q^p(U)$ is a $\mathcal{C}^r(U)$-module that is free whenever $T(U)$ is trivializable, and the mapping $U \mapsto \mathcal{T}_q^p(U)$ is a sheaf of $\mathcal{C}^r$-Modules, where $\mathcal{C}^r$ is the sheaf of rings $U \mapsto \mathcal{C}^r(U)$ (see Remark 3.34).

If B is m-dimensional, a section of class C^r of the bundle $\bigwedge^{p} T(B)$ of *tangent p-vectors* over U (section 4.2.3) is said to be a *p-vector field*. .

(IV) SCALAR FIELDS By Remark 4.22, $\mathcal{T}_0^0(U)$ is the set of mappings of class C^r $\boldsymbol{G} : U \to U \times \mathbb{K}$ and can be identified with $\mathcal{C}^r(U)$.

4.4. Differential forms

4.4.1. *Differential forms of degree* p

(I) Let B be a manifold. Write $\mathrm{Alt}^p\left(\left(T\left(B\right)\right);\mathbb{K}\right)$ for the vector bundle of class C^r of alternating continuous p-linear mappings from $T\left(B\right)$ into $\mathbb{K}$. The fiber of $\mathrm{Alt}^p\left(\left(T\left(B\right)\right);\mathbb{K}\right)$ over an arbitrary point $b\in U$ is given by $\mathrm{Alt}^p\left(\left(T\left(B\right)\right)_b;\mathbb{K}\right)$ (sections 4.2.1 and 4.2.6).

DEFINITION 4.24.– *A differential p-form (or a differential form of degree p) on an open subset U of B is a section of class C^r of $Alt^p\left(\left(T\left(B\right)\right);\mathbb{K}\right)$ on U, i.e. an element of $\Omega^p\left(U\right):=\Gamma\left(U,Alt^p\left(\left(T\left(B\right)\right);\mathbb{K}\right)\right)$ (Corollary-Definition 3.21).*

The set $\Omega^p\left(U\right)$ is a $\mathcal{C}^r\left(U\right)$-module. The mapping $U\mapsto\Omega^p\left(U\right)$, where U is an open subset of B, is a sheaf of $\mathcal{C}^r$-Modules. The set $\Omega^p\left(U\right)$ is in fact just the $\mathcal{C}^r\left(U\right)$-module of *antisymmetric p-times covariant tensor fields* of class C^r on U.

(II) FINITE-DIMENSIONAL CASE Suppose that the manifold B is locally finite-dimensional and let (U,ξ,m) be a chart of B. By Theorem 2.71, every differential form on U can be written in the form

$$b\mapsto\omega\left(b\right)=\sum_{i=1}^{m}a_i\left(b\right).d\xi^i\left(b\right), \qquad [4.19]$$

where the a_i are m functions $U\to\mathbb{K}$. For ω to be of class C^r, it is necessary and sufficient for the a_i to be of class C^r. Every p-form $\alpha\in\Omega^p\left(U\right)$ can be uniquely written as

$$\boxed{\alpha=\sum_{1\leq i_1<\ldots<i_p\leq m}a_{i_1,\ldots,i_p}.d\xi^{i_1}\wedge\ldots\wedge d\xi^{i_p},} \qquad [4.20]$$

where the $a_{i_1,\ldots,i_p}$ are of class C^r. Hence, the differential p-forms $d\xi^{i_1}\wedge\ldots\wedge d\xi^{i_p}$ stated above form a basis of the free $\mathcal{C}^r\left(U\right)$-module $\Omega^p\left(U\right)$. The $\mathcal{C}^r\left(U\right)$-module $\Omega^p\left(B\right)$ is free whenever $T\left(B\right)$ is trivializable.

Identifying $\bigwedge^p T^\vee\left(B\right)$ with the dual bundle of $\bigwedge^p T\left(B\right)$, if $Y_1,\ldots,Y_p$ are p vector fields, then

$$\left\langle\alpha,Y_1\wedge\ldots\wedge Y_p\right\rangle=\sum_{1\leq i_1<\ldots<i_p\leq m}a_{i_1,\ldots,i_p}.\det\left(\left\langle d\xi^{i_h},Y_k\right\rangle\right). \qquad [4.21]$$

The indices h and k range from 1 to p in each determinant by [4.12].

Let $B = \mathbb{K}^n$, $\alpha(x) = f(x)\, dx^1 \wedge ... \wedge dx^n$, and $Y(x) = \sum_{1 \leq i \leq n} y^i(x)\, \partial/\partial x_i$. Then, [4.14] implies that:

$$(Y \lrcorner \alpha)(x) = \sum_{i=1}^{n} (-1)^{i+1} f(x)\, y_i(x)\, dx^1 \wedge ... \wedge \widehat{dx^i} \wedge ... \wedge dx^n.$$

(III) De Rham algebra Set

$$\boxed{\Omega = \bigoplus_{p=0}^{\infty} \Omega^p,}$$

with $\Omega^0 = \mathcal{C}^r(B)$ (section 4.3.3(**IV**)) and $\Omega^p = \Omega^p(B)$. The notion of exterior algebra introduced in section 4.2.3 allows us to state the following result (see footnote 1(b), p. 132):

LEMMA-DEFINITION 4.25.– *1) Let (U, ξ, m) be a chart of B. The restriction $\Omega\,|U$ is the* exterior algebra *of the free $\Omega^0\,|U$-module $\Omega^1\,|U$.*

*2) The $\mathbb{K}$-algebra Ω equipped with the exterior product $\wedge$ of differential forms ([P1], section 3.3.8(**VII**)) is associative, graduated and anticommutative ([P1], section 2.3.5 (**IV**)). The exterior product satisfies:*

$$\alpha \wedge \beta = (-1)^{(\#\alpha)(\#\beta)}\, \beta \wedge \alpha,$$

where $\#$ denotes the degree. We say that Ω is the de Rham algebra *of differential forms on B.*

4.4.2. *Preimage of a differential p-form*

(I) Let B, B' be two manifolds, $\omega \in \Omega^p(B)$, and $f : B' \to B$ a morphism of manifolds.

LEMMA 4.26.– *There exists a uniquely determined differential p-form $f^*(\omega) \in \Omega^p(B')$ such that*

$$f^*(\omega)_{b'}\left(v'_1, ..., v'_p\right) = \omega_{(f(b'))}\left(T_{b'}(f).v'_1, ..., T_{b'}(f).v'_p\right) \qquad [4.22]$$

for every $b' \in B'$ and every family $\left\{v'_1, ..., v'_p\right\}$ of elements of the tangent space $T_{b'}(B')$.

DEFINITION 4.27.– *The differential p-form $f^*(\omega)$ specified above is said to be the* preimage *of ω under f (or the p-form induced by f from ω).*

In particular, if $p=1$, then $\langle f^*(\omega)_{b'}, v'_1\rangle = \left\langle \omega_{(f(b'))}, T_{b'}(f).v'_1\right\rangle = \langle {}^tT_{b'}(f).\omega_{(f(b'))}, v_1\rangle$, so

$$f^*(\omega)_{b'} = {}^tT_{b'}(f).\omega_{(f(b'))}.$$

If B'' is a manifold and $g : B'' \to B'$ is a morphism, then (**exercise**)

$$(f \circ g)^*(\omega) = g^*(f^*(\omega)).$$

(II) If B, B' have dimensions m, m', respectively, (U, ξ, m) is a chart of B with $\xi = (\xi^i)_{1\leq i\leq m}$, and ω is defined on U by [4.19], then the expression of $f^*(\omega)$ on $f^{-1}(U)$ is given by:

$$f^*(\omega) = \sum_{i=1}^{m} (a_i \circ f).d\left(\xi^i \circ f\right).$$

If the p-form $\alpha \in \Omega^p(B)$ can be written in the form [4.20] on U, i.e. $\alpha = \sum_{1\leq i_1<...<i_p\leq m} a_{i_1,...,i_p}.d\xi^{i_1} \wedge ... \wedge d\xi^{i_p}$, then $f^*(\alpha) \in \Omega^p(B')$ has the following expression on $f^{-1}(U)$, setting $f^*(a) = a \circ f$:

$$f^*(\alpha) = \sum_{1\leq i_1<...<i_p\leq m} f^*(a)_{i_1,...,i_p}.d\left(\xi^{i_1} \circ f\right) \wedge ... \wedge d\left(\xi^{i_p} \circ f\right).$$

Write $\xi = f(x)$, $x = \left(x^1, ..., x^{m'}\right)$. Then:

$$d\left(\xi^{i_k} \circ f\right) = \sum_{1\leq j_1<...<j_p\leq m'} \frac{\partial \xi^{i_k}}{\partial x^{j_p}} dx^{j_p}$$

and (**exercise**)

$$d\left(\xi^{i_1} \circ f\right) \wedge ... \wedge d\left(\xi^{i_p} \circ f\right) = \sum_{\substack{1\leq i_1<...<i_p\leq m\\ 1\leq j_1<...<j_p\leq m'}} \frac{\partial\left(\xi^{i_1}, ..., \xi^{i_p}\right)}{\partial\left(x^{j_1}, ..., x^{j_p}\right)} dx^{j_1} \wedge ... \wedge dx^{j_p},$$

where $\frac{\partial(\xi^{i_1},...,\xi^{i_p})}{\partial(x^{j_1},...,x^{j_p})}$ is the Jacobian of $\left(\xi^{i_1}, ..., \xi^{i_p}\right)$ relative to $\left(x^{j_1}, ..., x^{j_p}\right)$ (section 1.2.2(**IV**)). Hence:

$$\boxed{f^*(\alpha) = \sum_{\substack{1\leq i_1<...<i_p\leq m\\ 1\leq j_1<...<j_p\leq m'}} f^*(a)_{i_1,...,i_p} \cdot \frac{\partial(\xi^{i_1},...,\xi^{i_p})}{\partial(x^{j_1},...,x^{j_p})} dx^{j_1} \wedge ... \wedge dx^{j_p}.} \quad [4.23]$$

If $m = m' = p$, the sum [4.20] has a single term $a.d\xi^1 \wedge ... \wedge d\xi^p$, and [4.23] can be restated as follows:

$$f^* (\alpha) = f^* (a)_{i_1,...,i_p} \frac{\partial \left(\xi^1, ..., \xi^p\right)}{\partial \left(x^1, ..., x^p\right)} dx^1 \wedge ... \wedge dx^p. \qquad [4.24]$$

(III) Let $\alpha \in \Omega^q (B), \omega \in \Omega^p (B)$. Then (**exercise**: see [DIE 93], Volume 3, (16.20.9.5)):

$$f^* (\alpha \wedge \omega) = f^* (\alpha) \wedge f^* (\omega).$$

REMARK 4.28.– *Definition 4.27 (which is in fact just a* change of variables formula*) can be extended to the case of an arbitrary* covariant *tensor without requiring any modifications.*[6] *If* $\mathbf{Z} \in \mathcal{T}_p^0 (B)$ *is a field of* p*-times* covariant *tensors whose expression on* U *is given by*

$$\mathbf{Z} = \sum_{i_1...i_p} z_{i_1...i_p} d\xi^{i_1} \otimes ... \otimes d\xi^{i_p}$$

(see [4.18]) and $f : B' \to B$ *is a morphism of manifolds, then* $f^* (\mathbf{Z}) \in \mathcal{T}_p^0 (B')$ *has the following expression on* $f^{-1} (U)$*:*

$$f^* (\mathbf{Z}) = \sum_{i_1...i_p} \left(z_{i_1...i_p} \circ f\right) .d \left(\xi^{i_1} \circ f\right) \otimes ... \otimes d \left(\xi^{i_p} \circ f\right).$$

4.4.3. *Differential forms taking values in a fiber bundle. List of formulas*

The next section reproduces the presentation of [BOU 82a], sections 7.3 and 7.8, with slight changes to the order and a few simplifications.

(I) Let $\pi : N \to B$ be a vector bundle and consider the vector bundle $\mathrm{Alt}^p (M; N)$ $(M = T (B))$ whose fiber over an arbitrary point $b \in B$ is $\mathrm{Alt}^p (M; N)_b := \mathrm{Alt}^p (M_b; N_b)$ (section 4.2.6, Example 4.14(3)). If N is the trivial bundle $B \times \mathbf{F}$, where $\mathbf{F}$ is a Banach space, write $\mathrm{Alt}^p (M; \mathbf{F})$ for $\mathrm{Alt}^p (M; N)$.

DEFINITION 4.29.– *A differential* p*-form on an open subset* U *of* B *taking values in* N *is a morphic section of Alt*$^p (M; N)$*, i.e. an element of* $\Omega^p (U; N) := \Gamma (U, Alt^p (M; N))$.

Clearly, for any open subset U of B, $\Omega^p (U; N)$ is a $\mathcal{C}^r (U)$-module and the mapping $U \mapsto \Omega^p (U; N)$ is a sheaf of $\mathcal{C}^r$-Modules.

6 There is no equivalent definition for contravariant and mixed tensors, except for vector fields (see below, Definition 5.14).

REMARK 4.30.– *i) In the above, we can replace M by an arbitrary vector bundle of class C^r and base B. We can alternatively write $\Omega^p(U;N) := \Gamma(U, Alt^p(M;N))$, where the fiber bundle M is implicit on the left-hand side.*

ii) With $p = 0$, $\Omega^0(U;N)$ is the set of sections of class C^r of the form $\boldsymbol{G} : U \to N : b \mapsto (b, n_b)$ that are liftings of 1_U (thus, $n_b \in N_b$ and $\pi \circ \boldsymbol{G} = 1_U$).

iii) If N is the trivial bundle $B \times \mathbf{F}$, where $\mathbf{F}$ is a Banach space, then $\Omega^p(U;N)$ is written as $\Omega^p(U;\mathbf{F})$.

(II) EXTERIOR PRODUCT Let $\pi^i : N^i \to B$ $(i = 1, 2)$ and $\pi : N \to B$ be three vector bundles of class C^r and Φ a mapping from $N^1 \times_B N^2$ into N. Suppose that, for every $b_0 \in B$, there exist an open neighborhood U of b_0 in B and vector charts $t^i = (U, \varphi^i, \mathbf{F}^i)$ of π^i $(i = 1, 2)$ and $t = (U, \varphi, \mathbf{F})$ of π, together with a mapping λ of class C^r in $\mathcal{L}(\mathbf{F}^1 \times \mathbf{F}^2; \mathbf{F})$, satisfying the following condition: for every $b \in U$ and all $(x^1, x^2) \in \mathbf{F}^1 \times \mathbf{F}^2$ (with the notation of Definition 3.22(i), Condition **(V)**),

$$(t_b \circ \lambda(b))(x^1, x^2) = \Phi(t_b^1(x^1), t_b^2(x^2)).$$

DEFINITION 4.31.– *The mapping Φ defined above is said to be a* coupling *of the fiber product $N^1 \times_B N^2$ (section 2.3.9) in N.*

Suppose that one such coupling is given and let M be a fiber bundle with base B and of class C^r. For every $b \in B$, there is a continuous bilinear mapping

$$\Phi_b : \mathrm{Alt}^p(M;N^1)_b \times \mathrm{Alt}^q(M;N^2)_b \to \mathrm{Alt}^{p+q}(M;N)_b.$$

The collection of these continuous bilinear mappings determines a coupling

$$u : \mathrm{Alt}^p(M;N^1) \times \mathrm{Alt}^q(M;N^2) \to \mathrm{Alt}^{p+q}(M;N).$$

Given sections ω^1, ω^2 of $\mathrm{Alt}^p(M;N^1)$ and $\mathrm{Alt}^q(M;N^2)$ on U, the section $u(\omega^1, \omega^2)$ of $\mathrm{Alt}^{p+q}(M;N)$ on U is written as $\omega^1 \wedge_\Phi \omega^2$ (see Remark 4.18).

DEFINITION 4.32.– *We say that $\omega^1 \wedge_\Phi \omega^2$ is the* exterior product *of ω^1 and ω^2. This exterior product is written as $\omega^1 \wedge \omega^2$ whenever the coupling Φ is implicitly clear.*

If the $\mathbf{s}_i$ $(i = 1, ..., p+q)$ are morphic sections of M on U, then:

$$\begin{aligned}&(\omega^1 \wedge_\Phi \omega^2)(\mathbf{s}_1, ..., \mathbf{s}_{p+q})\\ &= \sum_\sigma \varepsilon(\sigma)\, \Phi\left(\omega^1(\mathbf{s}_{\sigma(1)}, ..., \mathbf{s}_{\sigma(p)}), \omega^2(\mathbf{s}_{\sigma(p+1)}, ..., \mathbf{s}_{\sigma(p+q)})\right),\end{aligned}$$

where the sum ranges over permutations of $\{1, ..., p+q\}$ such that $\sigma(1) < ... < \sigma(p)$ and $\sigma(p+1) < ... < \sigma(p+q)$ (see Lemma-Definition 4.10(i)).

(III) DE RHAM ALGEBRA: GENERALIZATION An *algebra bundle* of base B is a vector bundle A of base B equipped with a coupling from $A \times_B A$ into A. In the following, the fibers A_b $(b \in B)$ are assumed to be associative and commutative algebras with neutral element e_b ([P1], section 2.3.10(**I**)). Let $\pi : M \to B$ be a vector bundle with base B (e.g. $T(B)$). Write $\Omega^p(U; A)$ (where U is an open subset of B) for the $\mathcal{C}^r(U)$-module formed by the morphic sections of $\mathrm{Alt}^p(M; A)$ on U.

The exterior product allows the direct sum $\Omega(U; A) = \bigoplus_{p \geq 0} \Omega^p(U; A)$ to be equipped with the structure of an associative, anticommutative and graduated algebra, again called the *de Rham algebra* (section 4.4.1(**III**)).

If $\omega = \omega_1 \wedge ... \wedge \omega_p$, where $\omega_i \in \Omega^1(U; A)$, and $\mathbf{s}_i$ is a morphic section of M on U $(i = 1, ..., p)$, then (see [4.21]):

$$\boxed{\omega(\mathbf{s}_1, ..., \mathbf{s}_p) = \det(\langle \omega_i, \mathbf{s}_j \rangle).}$$

(IV) INTERIOR PRODUCT Let $p \geq 1$. There exists a coupling i from $M \times_B \mathrm{Alt}^p(M; A)$ into $\mathrm{Alt}^{p-1}(M; A)$ whose restriction to each fiber is the interior product (Remark 4.12). If $\mathbf{s}$ is a section of M on the open subset U of B and $\omega \in \Omega^p(U; A)$, write $i(\mathbf{s})\,\omega$ for the section $i(\mathbf{s}, \omega)$ of $\mathrm{Alt}^{p-1}(M; A)$ on U. Then, for every $b \in U$,

$$\boxed{(i(\mathbf{s})\,\omega)(b) = \mathbf{s}(b) \lrcorner \omega(b).}$$

We say that $i(\mathbf{s})\,\omega$ is the *interior product* of $\mathbf{s}$ and ω. From this definition and the formulas listed in section 4.2.5, we can deduce the following relations (where the $\mathbf{s}_j$ are sections of M):

$$\Big(i(\mathbf{s}) \underbrace{\omega}_{\text{degree } p}\Big)(\mathbf{s}_1, ..., \mathbf{s}_{p-1}) = \omega(\mathbf{s}, \mathbf{s}_1, ..., \mathbf{s}_{p-1}),$$

$$i(\mathbf{s}) \circ i(\mathbf{s}) = 0,$$

$$i(\mathbf{s}) \underbrace{\omega}_{\text{degree } 1} = \langle \omega, \mathbf{s} \rangle,$$

$$i(\mathbf{s}).\Big(\underbrace{\omega}_{\text{degree } p} \wedge \omega'\Big) = i(\mathbf{s})\,\omega \wedge \omega' + (-1)^p\,\omega \wedge i(\mathbf{s})\,\omega',$$

$$i(\mathbf{s})\Big(\underbrace{\omega_1 \wedge ... \wedge \omega_p}_{\text{forms of degree } 1}\Big) = \sum_{i=1}^{p} (-1)^{i+1} \langle \omega_i, \mathbf{s} \rangle\, \omega_1 \wedge ... \wedge \widehat{\omega_i} \wedge ... \wedge \omega_p.$$

The interior product is an antiderivation of degree -1 of the de Rham algebra ([P1], section 2.3.12).

(V) Preimage Let $f : B' \to B$ be a morphism of manifolds and $\omega \in \Omega^p(B; N)$. Let $f^*(N)$ be the preimage of N under f (section 3.4.6). There exists a uniquely determined differential p-form $f^*(\omega)$ such that [4.22] holds (*mutatis mutandis*). For every family of vector fields $X'_1, ..., X'_p$ of class C^r on B', the mapping

$$f^*(\omega)\left(X'_1, ..., X'_p\right) : b' \mapsto f^*(\omega)_{b'}\left(X'_1(b'), ..., X'_p(b')\right)$$

is a lifting of f into N (Definition 3.6, section 3.3.1).

In particular, a differential 0-form $\mathbf{G}$ on B' taking values in N is in fact just a lifting of class C^r of f into N, i.e. a mapping (of class C^r) $\mathbf{G} : B' \to N$ such that $\mathbf{G}(b') \in N_{f(b')}$ for all $b' \in B'$.

(VI) Vector-valued differential forms The fiber bundle $\Omega^p(B; N)$ is trivial if and only if N is trivial (section 3.4.1, Example 3.26(a)), or in other words $N = B \times \mathbf{F}$, where $\mathbf{F}$ is a Banach space (**exercise**). If so, the fibers N_b can be identified with $\mathbf{F}$. We write $\mathrm{Alt}^p(M; \mathbf{F})$ for $\mathrm{Alt}^p(M; N)$, $\Omega^p(B; \mathbf{F})$ for $\Omega^p(B; N)$, and we say that ω is a *vector-valued* differential p-form taking values in $\mathbf{F}$. In particular, if $\mathbb{K} = \mathbb{R}$, $N = T(B)$, and $\mathbf{F} = \mathbb{C}$, then ω is a *complex* differential p-form on the real manifold B.

The usual fiber bundle $\mathrm{Alt}^p((T(B)); \mathbb{K})$ can also be written as $\mathrm{Alt}^p((T(B)); \mathbb{K}_B) = \mathrm{Alt}^p((T(B)); B \times \mathbb{K})$, and

$$\Omega^p(B) = \Gamma\left(\mathrm{Alt}^p((T(B)); B \times \mathbb{K})\right). \qquad [4.25]$$

Remark 4.33.– *The space of differential p-forms on an $(\mathcal{FN})$ or $(\mathcal{SN})$ manifold can be defined using [4.25] ([KRI 97], Chapter VII, section 33).*

4.4.4. *Orientation*

For the rest of this chapter, excluding section 4.5.1, every manifold is a *locally finite-dimensional differential manifold* and $\mathbf{F}$ always denotes a Banach space.

(I) Orientation of a vector space Let $\mathbf{E}$ be a *real* m-dimensional vector space. We know that $\dim\left(\bigwedge^m \mathbf{E}\right) = 1$ (section 4.2.3(**I**)), so $\det(\mathbf{E}) := \bigwedge^m \mathbf{E}$ is the union of two closed half-lines on the real axis with opposite directions and origin 0. These half-lines are written as O and $-O$. The set $\{O, -O\}$ formed by these half-lines is written as $\mathrm{Or}(\mathbf{E})$.

(II) ORIENTATION OF A REAL MANIFOLD Let B be a manifold and $M = T(B)$. Write

$$\mathrm{Or}_M = \dot{\bigcup}_{b \in B} \mathrm{Or}(M_b).$$

LEMMA 4.34.– *Let $\tilde{\pi} : Or_M \to B$ be the mapping defined by $\tilde{\pi}((b, O)) = b$ for every $b \in B$. There exists a unique topological space structure on Or_M for which the following two conditions are satisfied:*

i) $\tilde{\pi}$ is continuous.

ii) If $\mathbf{s}$ is a continuous, everywhere non-zero section of the vector bundle $\det(M)$ *(whose fibers are the* $\det(M_b)$*, $b \in B$, with the notation of **(I)**, implying that $\mathbf{s}(b) \in$* $\det(M_b) = O(b) \cup (-O(b))$*) on an open subset U of B and $\mathbf{s}(b) \in O(\mathbf{s}(b))$ for every $b \in U$, then the mapping $B \to Or_M : b \mapsto O(\mathbf{s}(b))$ is continuous (**exercise**).*

Suppose that the topological space Or_M is equipped with the manifold structure determined by taking the preimage under $\tilde{\pi}$ of the manifold structure of B (Remark 2.44), and consider the fibration $\tilde{\pi} : \mathrm{Or}_M \to B$. The multiplicative group $\{\pm 1\}$ acts simply transitively (and hence freely) ([P1], section 2.2.8(**II**)) on Or_M by $O \mapsto -O$. The manifold of orbits of this action is $\mathrm{Or}_M \backslash \{\pm 1\} \cong B$.

COROLLARY-DEFINITION 4.35.– *1) The fibration $\pi : Or_M \to B$ is a principal bundle with structural group $\{\pm 1\}$. This principal bundle $\tilde{B}$ is a covering of B with fiber type $\{\pm 1\}$ (section 3.3.3), namely a covering of two leaves.*

2) The principal bundle $\tilde{B}$ is said to be the orientation covering. *An* orientation *of B is a continuous section $O : b \mapsto (b, O_b)$ of $\tilde{B}$ $(O_b \in \{\pm 1\})$. If any such section exists, the pair (B, O) is said to be an* oriented manifold. *The orientations O and $-O$ are said to be* opposite.

3) A manifold B is said to be orientable *if there exists an orientation on B.*

4) A pure m-dimensional manifold B is orientable if and only if one of the following equivalent conditions is satisfied:

i) The orientation covering $\tilde{B}$ is the trivial bundle $B \times \{\pm 1\}$.

ii) There exists a continuous *differential m-form v_0 such that $v_0(b) \neq 0$ for every $b \in B$; if so, v_0 is of class C^∞, i.e. $v_0 \in \Omega^m(B)$. Since* $\dim(\Omega^m(B))_b = 1$, *the sign of $v_0(b)$ must be constant on B, and v_0 determines an orientation $O : b \mapsto (b, sgn(v_0(b)))$ of B (where sgn denotes the sign).*

iii) There exists an atlas of B whose charts (U_i, φ_i, n_i) satisfy the property that, if $U_i \cap U_j \neq \varnothing$ (which implies $n_i = n_j$), then

$$\frac{\partial\left(\varphi_i^1, ..., \varphi_i^{n_i}\right)}{\partial\left(\varphi_j^1, ..., \varphi_j^{n_j}\right)} > 0$$

on $U_i \cap U_j$.

5) Hence, if B is a pure orientable m-dimensional manifold, then the relation $\sim$ on $\Omega^m(B)$ defined by $v \sim v'$ if $sgn(v(b)) = sgn(v'(b))$ (where b is an arbitrary point of B) is an equivalence relation. The orientation $O : b \longmapsto (b, sgn(v(b)))$ is the equivalence class of v, written as $\bar{v}$.

PROOF.– (4): See [DIE 93], Volume 3, (16.21.1), (16.21.16). (5): **exercise**. ■

COROLLARY 4.36.– *1) Let B be a manifold and b some point of B. By Corollary-Definition 4.35(4), there exists an open neighborhood U of b in B that is orientable.*

2) A manifold B is orientable if and only if each of its connected components is orientable.

3) Let (B, O) be a pure m-dimensional oriented manifold and $v_0 \in O$. Every differential m-form $\omega \in \Omega^m(B)$ can be uniquely written in the form $\omega = f.v_0$, where $f : B \to \mathbb{R}$ is of class C^∞. Given $b \in B$, write $\omega(b) \gtreqless 0$ if $f(b) \gtreqless 0$.

*4) Let (B, O) be a pure m-dimensional oriented manifold. If $\omega \in \Omega^m(B; \mathbf{F})$, there exists a unique mapping $\mathbf{f} : B \to \mathbf{F}$ of class C^∞ satisfying $\omega = \mathbf{f}.v_0$ (**exercise**).*

DEFINITION 4.37.– *Let (B, O) be an oriented manifold and $v_0 \in O$. A sequence $(Z_1, ..., Z_m)$ of vector fields is said to be* positive *or* direct *(respectively* negative *or* retrograde*) if, for every $b \in B$,*

$$\langle v_0(b), Z_1(b) \wedge ... \wedge Z_m(b) \rangle > 0 \quad \textit{(resp. } < 0\textit{).}$$

EXAMPLE 4.38.– *i) The space $\mathbb{R}^m$ is orientable and the canonical m-form $dx^1 \wedge ... \wedge dx^m$ (where x^i is the i-th coordinate function in the canonical basis) defines its* canonical orientation.

*ii) More generally, the underlying manifold of a finite-dimensional real Lie group $\mathbf{G}$ is always orientable. Indeed, suppose that $\mathbf{G}$ is m-dimensional, and let $z^\vee$ be an m-covector such that $z_e^\vee \neq 0$ (where e is the neutral element); then the differential m-form $g \mapsto \boldsymbol{\gamma}(g) z^\vee$ (of class C^ω), where $\boldsymbol{\gamma}$ is left translation (section 2.4.1(**I**)), is non-zero at every point.*

iii) Every simply connected manifold and every parallelizable manifold (Definition 3.28) is orientable ([NAR 73], Corollary 2.7.6; [LEE 02], Proposition 10.5).

iv) In particular, the sphere $\mathbb{S}^n$ (see footnote 2, p. 98) is orientable.

v) Every finite product of orientable manifolds is orientable. If B_1, B_2 are two manifolds with orientations O_1, O_2 respectively, the mapping $(b_1, b_2) \mapsto O_{b_1} O_{b_2}$ is an orientation of $B_1 \times B_2$ written as $O_1 \otimes O_2$.

vi) Let B be an oriented manifold with orientation O. Let U be a submanifold of B. The mapping $O\,|_U$ is an orientation of U. Hence, every submanifold of an orientable manifold is orientable.

vii) It can be shown that any finite-dimensional pure differential manifold B_0 underlying a holomorphic manifold B is orientable ([DIE 93], Volume 3, (16.21.13)).

viii) The Möbius strip (see Figure 3.2 in section 3.3.1 and footnote 3, p. 98) and the Klein bottle (see footnote 6, p. 70) are not orientable. The Möbius strip is a striking example of a non-orientable manifold: any reader who wishes to experiment with the concept of orientation can make a Möbius strip by gluing together the two ends of a strip of paper with a half-twist. Now, draw a pencil line along the middle of the strip – the line will almost magically reach "the other side" of the strip from the starting point.

Let (B, O) be an oriented manifold. If we write this oriented manifold as $\overset{\curvearrowright}{B}$, we can write $\overset{\curvearrowleft}{B}$ or $-\overset{\curvearrowright}{B}$ for the manifold equipped with the opposite orientation.

4.4.5. *Integral of a differential form of maximal degree*

(I) VOLUME INTEGRALS IN $\mathbb{R}^m$

LEMMA 4.39.– *Let U, U' be open subsets of $\mathbb{R}^m$ and $u : U \to U'$ a diffeomorphism. For every $x \in U$, let $J = \det\left(\frac{\partial u}{\partial x}\right)$ be the Jacobian of u (section 1.2.2**(IV)**). Let $\lambda_U^{\otimes m} := \lambda^{\otimes m}|_U$ and $\lambda_{U'}^{\otimes m}$ be the Radon measures induced on U and U', respectively, by the Lebesgue measure on $\mathbb{R}^m$ ([P2], section 4.1.5**(I)**). Then, the* image *of $|J|\,.\lambda_U^{\otimes m}$ under u is $\lambda_{U'}^{\otimes m}$ ([P2], section 4.1.5**(II)**); in other words, for every function $f \in \mathcal{K}(U')$, where $\mathcal{K}(U')$ denotes the space (of type $(\mathcal{L}_s\mathcal{F})$) of compactly supported continuous functions from U' into $\mathbb{R}$ ([P2], section 4.1.4**(IV)**), we have the following change of variable formula, by [4.24]:*

$$\int_{U'} f(x')\,.d\lambda^{\otimes m}(x') = \int_U f(u(x))\,.\,|J(x)|\,.d\lambda^{\otimes m}(x).$$

This relation still holds if we replace $f \in \mathcal{K}(U')$ by a $\lambda^{\otimes m}$-integrable mapping $\mathbf{f} : U' \mapsto \mathbf{F}$ ([P2], section 4.1.2).

EXAMPLE 4.40.– *Let U and U' be two open intervals on the real line and u a diffeomorphism from U onto U'. Let $\mathbf{f}$ be a continuous function on U taking values in the Banach space $\mathbf{F}$. The function $t \mapsto u(t)$ is monotone, so $\dot{u}(t)$ has constant sign ε on U, and:*

$$\int_U \mathbf{f}(t).dt = \int_{U'} \mathbf{f}(u(t)).\varepsilon.\dot{u}(t).dt = \int_{U'} \mathbf{f}(u(t)).|\dot{u}(t)|.dt.$$

EXAMPLE 4.41.– *Let us calculate the volume V of a sphere of radius R.*

*1) Cartesian coordinates. Pick the center of the sphere as the origin, and begin by calculating the volume of the hemisphere $z \geq 0$. We can use the Fubini-Tonelli theorem ([P2], section 4.1.3(**III**)) to do this by cutting the hemisphere into "slices" of infinitely small thickness dz. Each slice is a cylinder with a circular cross-section of radius $\sqrt{R^2 - z^2}$ and thickness dz. The volume of the hemisphere is therefore given by*

$$\int_0^R \pi\left(R^2 - z^2\right) dz = \frac{2}{3}\pi R^3,$$

which gives us the classical formula $V = \frac{4}{3}\pi R^3$.

2) Spherical coordinates. In mathematics, the radial, azimuthal, and zenithal coordinates $\{r, \phi, \theta\}$ are defined as shown in Figure 4.1, with $r \geq 0$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$ (in physics, the symbols θ and ϕ are often swapped). These coordinates satisfy the relations:

$$x = r\cos\theta\sin\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\phi. \qquad [4.26]$$

Hence, by Lemma 4.39,

$$dxdydz = \det\begin{bmatrix} \cos\theta\sin\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\sin\theta\cos\phi & r\cos\theta\sin\phi \\ \cos\phi & -r\sin\phi & 0 \end{bmatrix} drd\phi d\theta \qquad [4.27]$$
$$= r^2\sin\phi drd\phi d\theta, \text{ so}$$

$$V = \int_0^R dr \int_0^{2\pi} d\theta \int_0^{\pi} d\phi\left(r^2\sin\phi\right) = -\frac{R^3}{3}.2\pi.\left[-\cos\phi\right]_0^{\pi} = \frac{4}{3}\pi R^3.$$

Let C be the point with Cartesian coordinates (x, y, z). Then, $d\overrightarrow{C} = \overrightarrow{e_1}dx + \overrightarrow{e_2}dy + \overrightarrow{e_3}dz$, where $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$ are the vectors of the canonical basis. In spherical coordinates, write $\overrightarrow{u_r}, \overrightarrow{u_\phi}, \overrightarrow{u_\theta}$ for the unit vector along the direction of $\overrightarrow{OC}$, the unit vector tangent to the meridian in the direction of increasing ϕ, and the unit vector *parallel to* C *in the direction of increasing* θ*, respectively. Then,* $\overrightarrow{OC} = r\overrightarrow{u_r}$*,* $d\overrightarrow{OC} = \overrightarrow{u_r}.dr + r\frac{\partial \overrightarrow{u_r}}{\partial \phi}.d\phi + r\frac{\partial \overrightarrow{u_r}}{\partial \theta}.d\theta = \frac{\partial}{\partial r}dr + \frac{\partial}{\partial \varphi}d\varphi + \frac{\partial}{\partial \theta}d\theta$ *with* $\overrightarrow{u_r} = \cos\theta\sin\phi.\overrightarrow{e_1} + \sin\theta\sin\phi.\overrightarrow{e_2} + \cos\phi.\overrightarrow{e_3}$*, so*

$$d\overrightarrow{OC} = dr.\overrightarrow{u_r} + \frac{1}{r}.\overrightarrow{u_\phi} + \frac{1}{r\sin\phi}.\overrightarrow{u_\theta}, \qquad [4.28]$$

$$\frac{\partial}{\partial r} = \overrightarrow{u_r}, \quad \frac{\partial}{\partial \phi} = r\frac{\partial \overrightarrow{u_r}}{\partial \phi} = r\overrightarrow{u_\phi}, \quad \frac{\partial}{\partial \theta} = r\frac{\partial \overrightarrow{u_r}}{\partial \theta} = r\sin\phi.\overrightarrow{u_\theta}. \qquad [4.29]$$

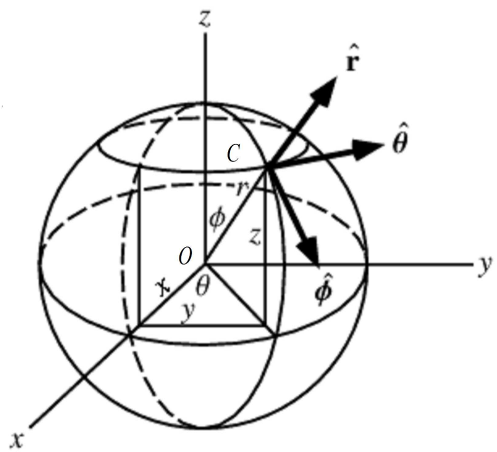


Figure 4.1. *Spherical coordinates*

The orthonormal frame $\{\overrightarrow{u_r}, \overrightarrow{u_\phi}, \overrightarrow{u_\theta}\}$ *is positively oriented, since the determinant [4.27] is positive.*

3) Calculation based on the radial coordinate, the longitude, and the latitude. The coordinates are now (r, φ, λ) *(see Example 2.12, section 2.2.1(**II**)). In radians, the latitude* φ *and the longitude* λ *may now be expressed as a function of the zenith* ϕ *and the azimuth* θ *by* $\varphi = \frac{\pi}{2} - \phi$ *and* $\lambda = \theta - \pi$*. Starting from the expressions [2.1], an analogous calculation gives* $dxdydz = r^2\cos\varphi dr d\varphi d\lambda$*, and*

$$V = \int_0^R dr \int_{-\pi}^{\pi} d\lambda \int_{-\pi/2}^{\pi/2} d\varphi \left(r^2\cos\varphi\right) = \frac{R^3}{3}.2\pi.\left[\sin\varphi\right]_{-\pi/2}^{\pi/2} = \frac{4}{3}\pi R^3.$$

(II) LEBESGUE MEASURES

COROLLARY-DEFINITION 4.42.– *Let B be a pure m-dimensional manifold. There exists a positive Radon measure μ on B ([P2], section 4.1.5(**V**)) with the following properties:*

*i) For every chart $c = (U, \varphi, m)$ of B, the image under φ ([P2], section 4.1.5(**II**)) of the induced measure μ_U is of the form $f_c.\lambda^{\otimes m}_{\varphi(U)}$, where f_c is a function of class C^∞ that does not vanish on $\varphi(U)$.*

ii) In other words, for every function $g \in \mathcal{K}(U)$,

$$\int_U g(b)\, d\mu(b) = \int_{\varphi(U)} g(\zeta)\, f_c(\zeta)\, d\lambda^{\otimes m}(\zeta).$$

iii) Any Radon measure μ with the property (i) is said to be Lebesgue *on B. Any two Lebesgue measures μ, μ' on B are equivalent in the sense that each is absolutely continuous with respect to the other ([P2], section 4.1.6(**III**)) and each has a density function of class C^∞ with respect to the other.*

(III) INTEGRAL OF A FORM OF MAXIMAL DEGREE OVER AN m-DIMENSIONAL ORIENTED MANIFOLD Let $\overset{\frown}{B}$ be a pure *oriented* manifold of dimension $m \geq 0$ and let ω be a differential m-form taking values in a Banach space $\mathbf{F}$. Our next task is to give meaning to the quantity

$$\boxed{\int_{\overset{\frown}{B}} \omega.}$$

By Corollary-Definition 4.35(5), there exists a differential m-form $v_0 \in \Omega^m(B)$ that belongs to the orientation of $\overset{\frown}{B}$. Let $c = (U, \varphi, m)$ be a chart of B such that the open set U is connected. For every point $\zeta = (\zeta^1, ..., \zeta^m) \in \varphi(U)$, we can write (with the same notation as above):

$$v_0\left(\varphi^{-1}(\zeta)\right) = f_c(\zeta)\, .d\zeta^1 \wedge ... \wedge d\zeta^m.$$

Since U is connected, f_c is of class C^∞ and has constant sign in $\varphi(U)$. Let $\mu_{v_0,U}$ be the positive Radon measure on U defined by

$$\mu_{v_0,U} = \varphi^{-1}\left(|f_c|\, .\lambda^{\otimes m}_{\varphi(U)}\right).$$

If we proceed in the same way for another arbitrary chart $c' = (U', \varphi', m)$ of B, where U' is also connected, we obtain a measure $\mu'_{v_0,U'}$ that is positive on U'. It is easy to show ([DIE 93], Volume 3, (16.24.1)) that, if $U \cap U' \neq \varnothing$, the restrictions of $\mu_{v_0,U}$ and $\mu'_{v_0,U'}$ to $U \cap U'$ are equal. The positive measures $\mu_{v_0,U}$ (as U ranges over the set of connected open subsets of B, which form a covering of B) are, therefore, the restrictions to U of a unique positive Radon measure μ_{v_0} (which is Lebesgue) defined on B. With the same hypotheses and notation as Corollary 4.36(4):

COROLLARY-DEFINITION 4.43.– *The differential m-form ω is said to be integrable (over $\overset{\curvearrowright}{B}$) if* $\mathbf{f}$ *is μ_{v_0}-integrable. If so, define:*

$$\int_{\overset{\curvearrowright}{B}} \omega := \int_{\overset{\curvearrowright}{B}} \mathbf{f}.d\mu_{v_0}.$$

This quantity only depends on the orientation of $\overset{\curvearrowright}{B}$ and not on the particular choice of v_0 made to specify this orientation.

Let $\omega \in \Omega^m(B;\mathbf{F})$ be an integrable differential m-form on $\overset{\curvearrowright}{B}$ taking values in $\mathbf{F}$. With the notation introduced at the end of section 4.4.4**(II)**:

$$\int_{-\overset{\curvearrowright}{B}} \omega = -\int_{\overset{\curvearrowright}{B}} \omega.$$

If $g \in \mathcal{K}(B)$ and $\omega \in \Omega^m(B;\mathbf{F})$, then $g.\omega$ is a *continuous* m-form and $g \mapsto \int_{\overset{\curvearrowright}{B}} g.\omega$ is a Radon measure $[\omega]$ taking values in $\mathbf{F}$ ([P2], section 4.1.5**(VIII)**).[7]

DEFINITION 4.44.– *The Radon measure $[\omega]$ defined on B is called the* volume form *on $\overset{\curvearrowright}{B}$ determined by the differential m-form ω.*

The above leads to the following result ([DIE 93], Volume 3, (16.24.2)):

COROLLARY 4.45.– *If $\omega \in \Omega^m(B)$* belongs to the orientation of $\overset{\curvearrowright}{B}$*, then the volume form*

$$[\omega] : \mathcal{K}(B) \to \mathbb{R} : g \mapsto \int_{\overset{\curvearrowright}{B}} g.\omega$$

is a positive *Lebesgue measure on $\overset{\curvearrowright}{B}$. Conversely, every positive Lebesgue measure on $\overset{\curvearrowright}{B}$ is of the form $[\omega]$, $\omega \in \Omega^m(B)$.*

7 The cited reference only considers measures taking values the dual $\mathbf{E}^\vee$ of a finite-dimensional space $\mathbf{E}$. However, extending this concept to measures taking values in an arbitrary Banach space $\mathbf{F}$ is entirely straightforward (see, for example, [SCH 93], Volume III, Definition 5.4.6).

(IV) Orientation of a morphism

Definition 4.46.– *Let B, B' be locally finite-dimensional manifolds and $f : B \to B'$ a morphism. An* orientation *of f is a morphism $\tilde{f} : \tilde{B} \to \tilde{B}' : (b, O_b) \to (b', O'_b)$ that makes the diagram*

$$\begin{array}{ccc} \tilde{B} & \xrightarrow{\tilde{f}} & \tilde{B}' \\ \downarrow \pi & & \downarrow \pi' \\ B & \xrightarrow{f} & B' \end{array}$$

commute and which is compatible with the group action $\{\pm 1\}$. In other words, $\tilde{f}(b, O_b)(\pm 1) = \left(f(b), \pm O'_{f(b)}\right)$. We write that $\tilde{f} = (f, \varepsilon)$, $\varepsilon \in \{-1, 1\}$.

Corollary-Definition 4.47.– *1) Let B, B' be orientable manifolds, $f : B' \to B$ a diffeomorphism, and $\tilde{f} = (f, \varepsilon)$ an orientation of f. Let O be an orientation of B and write $\overset{\frown}{B} = (B, O)$. There exists a unique orientation O' on B' such that, setting $\overset{\frown}{B}{}' = (B', O')$, every differential form $\omega \in \Omega^m(B; \mathbf{F})$ satisfies*

$$\int_{\overset{\frown}{B}} \omega = \varepsilon \int_{\overset{\frown}{B}{}'} f^*(\omega) \qquad [4.30]$$

*(**exercise**). The orientation O' is said to be* associated *with O by $\tilde{f}$.*

2) Conversely, let $\overset{\frown}{B}, \overset{\frown}{B}{}'$ be two oriented manifolds and $f : B' \to B$ a diffeomorphism. If the relation [4.30] is satisfied for every differential form $\omega \in \Omega^m(B; \mathbf{F})$ and $\varepsilon = +1$, we say that f preserves orientations*, or that f* reverses orientations *if $\varepsilon = -1$. The pair $\tilde{f} := (f, \varepsilon)$ is an orientation of f, where f is viewed as a diffeomorphism from B' onto B (and B and B' are the* orientable *manifolds underlying $\overset{\frown}{B}$ and $\overset{\frown}{B}{}'$, respectively).*

(V) Canonical orientation of the orientation covering Let B be a manifold and $\tilde{\pi} : \tilde{B} \to B$ the orientation covering (Corollary-Definition 4.35(1)). Let $\tilde{b} = (b, O_b) \in \tilde{B}$. The projection $\tilde{\pi} : \tilde{B} \twoheadrightarrow B$ is the linear mapping $\tilde{b} \mapsto b$. Its tangent linear mapping $T_{\tilde{b}}(\tilde{\pi}) : T_{\tilde{b}}\left(\tilde{B}\right) \to T_b(B)$ is an isomorphism, so $\tilde{\pi}$ is a local diffeomorphism (Theorem 2.61(2)), i.e. a diffeomorphism from an open subset U of $\tilde{b}$ onto an open subset V of b. We can choose U and V to be orientable (Corollary 4.36(1)); thus, there exists an orientation $O : b' \mapsto O_{b'}$ of V taking the value O_b at the point b. We have $\tilde{\pi}^{-1}(b) = \{(b, O_b), (b, -O_b)\}$. Let $\omega \in O$, $\tilde{\omega} = \tilde{\pi}^*(\omega)$, and $\tilde{O} = \overline{\tilde{\omega}}$ (see Corollary-Definition 4.35(5)); then $\tilde{b} \mapsto \tilde{O}_{\tilde{b}}$ is an orientation of $\tilde{B}$ ([DIE 93], Volume 3, (16.21.6)), which gives us the following result:

Theorem 4.48.– *The manifold $\tilde{B}$ is orientable.*

REMARK 4.49.– *(1) The space $\tilde{\pi} : \tilde{B} \twoheadrightarrow B$ is a covering of two leaves. It therefore has a canonical involution*[8] $\iota : \tilde{B} \to \tilde{B}$ *that permutes these leaves over each point $b \in B$. If $\tilde{O}$ is an orientation of $\tilde{B}$, then the orientation of $\tilde{B}$ associated with ι (Corollary-Definition 4.47) is $-\tilde{O}$. Therefore, both $\tilde{O}$ and $-\tilde{O}$ are orientations of $\tilde{B}$; nevertheless, we say that $\tilde{O}$ is the* canonical orientation. *This ambiguity is irrelevant in practice because $\tilde{O}$ is unique up to permutation of the two leaves of $\tilde{B}$.*

2) Conversely, if $\bar{\pi} : M \to B$ is an oriented covering of class C^∞ with two leaves and its canonical involution associates a given orientation of M with the opposite orientation, then $\bar{\pi} : M \to B$ is isomorphic to $\tilde{\pi} : \tilde{B} \to B$ ([LEB 82], Chapter I, section 5.C, Theorem 2).

4.4.6. *Differential forms of odd type*

(I) DEFINITION The fiber bundle $\tilde{\mathbb{R}} := \tilde{B} \times^{\{\pm 1\}} \mathbb{R}$ associated with $\tilde{B}$ of fiber type $\mathbb{R}$ (section 3.5.5) is called the bundle of scalars of odd type. Let N be a vector bundle of base B. A differential p-form $\underline{\omega} \in \Omega^p\left(B; \tilde{\mathbb{R}} \otimes N\right)$ is said to be a *differential p-form of odd type* on B taking values in N. Let $\tilde{\pi} : \tilde{B} \twoheadrightarrow B$ be the projection. There exists a bijection

$$\widetilde{\bullet} : \underline{\omega} \overset{\sim}{\longmapsto} \tilde{\omega} \qquad [4.31]$$

between the differential p-forms of odd type on B taking values in N and the differential p-forms on $\tilde{B}$ taking values in $\tilde{\pi}^*(N)$ such that $\tilde{\omega}(-O) = -\tilde{\omega}(O)$ for every orientation $O \in \tilde{B}$.

REMARK 4.50.– *a) When B is equipped with an orientation O, the bijection*

$$\Omega^p(B; N) \to \Omega^p\left(B; \tilde{\mathbb{R}} \otimes N\right) : \omega \mapsto \underline{\omega} := O \otimes \omega \qquad [4.32]$$

allows us to identify the usual differential p-forms ω on $\overset{\curvearrowright}{B}$ with the differential p-forms of odd type $\underline{\omega}$ on $\overset{\curvearrowright}{B}$. Thus, in most cases, we do not need to talk about differential forms of odd type on an oriented manifold. *However, the concept of differential p-form of odd type becomes crucial on manifolds without an orientation (especially when these manifolds are non-orientable). Whenever we talk about p-forms on these manifolds, we will always be referring to differential p-forms of odd type.*

b) We also say that an ordinary differential form (belonging to $\Omega^p(B; N)$) is even *and that a differential form of odd type (belonging to $\Omega^p\left(B; \tilde{\mathbb{R}} \otimes N\right)$) is* odd *([DER 84], Chapter II). This more simple terminology is adopted below.*

8 An involution is a bijection that coincides with its inverse bijection.

If N is the trivial bundle $B \times \mathbf{F}$, we write $\Omega^p\left(B; \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$ for the space $\Omega^p\left(B; \tilde{\mathbb{R}} \otimes N\right)$, and we say that the odd p-form $\underline{\omega} \in \Omega^p\left(B; \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$ takes values in $\mathbf{F}$.

(II) PREIMAGE Let $f : B \to B'$ be a morphism, $\tilde{f} : \tilde{B} \to \tilde{B}'$ an orientation of f (Definition 4.46), and $\underline{\omega} \in \Omega^p\left(B'; \tilde{\mathbb{R}} \otimes N\right)$.

DEFINITION 4.51.– *The preimage $f^*(\underline{\omega})$ is the odd differential form in $\Omega^p\left(B; \tilde{\mathbb{R}} \otimes f^*(N)\right)$ uniquely determined by the following property ([BOU 82a], 10.4.2): let $\tilde{\omega}$ (respectively $\widetilde{f^*(\underline{\omega})}$) be the differential p-form on $\tilde{B}'$ (respectively $\tilde{B}$) associated with $\underline{\omega}$ (respectively $f^*(\underline{\omega})$) by the bijection $\tilde{\bullet}$ from [4.31]; then the following relation holds:*

$$\boxed{\tilde{f}^*(\tilde{\omega}) = \widetilde{f^*(\underline{\omega})}.}$$

(III) EXTERIOR AND INTERIOR PRODUCTS Given three fiber bundles N^1, N^2, N with the same base B, a coupling Φ from $N^1 \times_B N^2$ into N (Definition 4.31), and two odd differential forms $\underline{\omega}^1 \in \Omega^p\left(B; \tilde{\mathbb{R}} \otimes N^1\right), \underline{\omega}^2 \in \Omega^q\left(B; \tilde{\mathbb{R}} \otimes N^2\right)$, we can define the exterior product $\underline{\omega}^1 \wedge_\Phi \underline{\omega}^2 \in \Omega^{p+q}\left(B; \tilde{\mathbb{R}} \otimes N\right)$ in the same way as in Definition 4.32. The coupling Φ uniquely determines a coupling $\tilde{\Phi}$ from $\tilde{N}^1 \times_B \tilde{N}^2$ into $\tilde{N}$, and (**exercise**)

$$\widetilde{\underline{\omega}^1 \wedge_\Phi \underline{\omega}^2} = \tilde{\omega}^1 \wedge_{\tilde{\Phi}} \tilde{\omega}^2.$$

If ω^1 and ω^2 are differential forms of same parity (respectively of opposite parity) (Remark 4.50(b)), then $\omega^1 \wedge_\Phi \omega^2$ is an even (respectively odd) differential form.

The formulas satisfied by the exterior product and the interior product (section 4.4.3**(II)**,**(IV)**) also hold for odd differential forms.

(IV) CHANGE OF VARIABLE In practical settings, an odd differential p-form $\underline{\alpha}$ can be expressed locally in the form [4.20]. Even differential p-forms satisfy the change-of-variable formula [4.23], whereas the corresponding formula satisfied by odd differential p-forms is as follows:

$$f^*(\underline{\alpha}) = \operatorname{sgn}(J) \sum_{\substack{1 \le i_1 < \ldots < i_p \le m \\ 1 \le j_1 < \ldots < j_p \le m'}} f^*(a)_{i_1,\ldots,i_p} .J.dx^{j_1} \wedge \ldots \wedge dx^{j_p}, \qquad [4.33]$$

where $J = \frac{\partial(\xi^1,\ldots,\xi^n)}{\partial(x^1,\ldots,x^n)}$ ([DER 84], Chapter II, section 5).

(V) Measure defined by an odd m-form Let B be an m-dimensional Hausdorff pure manifold, $\underline{\omega} \in \Omega^m\left(B; \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$, and O^m the canonical orientation of $\mathbb{R}^m$ (Example 4.38(i)). Let $c = (U, \varphi, m)$ be a chart of B.

Lemma 4.52.– *There exists a unique mapping* $\mathbf{f}_c : \varphi(U) \to \mathbf{F}$ *so that*

$$\underline{\omega}_U = \varphi^*\left(O^m \otimes \mathbf{f}_c . dx^1 \wedge ... \wedge dx^m\right).$$

We say that $\underline{\omega}$ is locally integrable if $\mathbf{f}_c$ is locally $\lambda_U^{\otimes m}$-integrable. If so, there exists a unique Radon measure $\alpha(\underline{\omega})$ taking values in $\mathbf{F}$ that satisfies the following property: for every chart $c = (U, \varphi, m)$ of B, $\varphi(\alpha(\underline{\omega})) = \mathbf{f}_c . \lambda_U^{\otimes m}$.

Definition 4.53.– *The Radon measure* $\alpha(\underline{\omega})$ *is said to be* defined *by the odd differential* m*-form* $\underline{\omega}$ *and is also written as* $\underline{\omega}$.

Remark 4.54.– *Suppose that* $\mathbf{F} = \mathbb{R}$; *let* $\overset{\curvearrowright}{B}$ *be an oriented* m*-dimensional manifold with orientation O. Let* $\omega \in \Omega^m(B)$, $\underline{\omega} = O \otimes \omega$ *(see [4.32]). Unlike the Radon measure* $[\omega]$ *from Definition 4.44, the Radon measure* $\underline{\omega}$ *from Definition 4.53 is not necessarily positive. We have the relation* $[\omega] = |\underline{\omega}|$ *([P2], section 4.1.5(**VI**)). In a certain sense, the notion of an odd differential form transfers the orientation of the manifold over to the form (see Example 4.55).*

Example 4.55.– *Let* B *be the cube* $0 < \xi^i < 1$ *in* $\mathbb{R}^3$*; this is an open subset of* $\mathbb{R}^3$ *and hence a submanifold (section 2.3.3). Write* $\overset{\curvearrowright}{B}$ *for the manifold* B *equipped with the orientation induced by the canonical orientation of* $\mathbb{R}^3$ *(Example 4.38(vi)).*

a) The "algebraic volume" of B *([SCH 93], Volume IV, Chapter VI, Remark 12) is (see Corollary 4.45)*

$$V = \int_{\overset{\curvearrowright}{B}} \omega = \int_B [\omega], \tag{4.34}$$

with $\omega := d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \in \Omega^3(B)$ *(*ordinary *differential 3-form) and* $[\omega] = \lambda_B^{\otimes 3}$.

Let $\varphi : \left(x^1, x^2, x^3\right) \mapsto \left(\xi^1, \xi^2, \xi^3\right)$, *where* $\xi^2 = x^1$, $\xi^1 = x^2$, $x^3 = \xi^3$. *By [4.24],*

$$\varphi^*(\omega) = dx^1 \wedge dx^2 \wedge dx^3 \Longrightarrow \varphi^*([\omega]) = [\omega] = \lambda_B^{\otimes 3}.$$

However, φ *reverses the orientation of* B *(Corollary-Definition 4.47(2)), so* $\varphi^*\left(\overset{\curvearrowright}{B}\right) = -\overset{\curvearrowright}{B}$, *and performing the change of variable* φ *transforms* V *into*

$$\varphi^*(V) := \int_{-\overset{\curvearrowright}{B}} \omega = -\int_{\overset{\curvearrowright}{B}} \omega = -\int_B [\omega] = -V.$$

b) Now consider the odd *differential 3-form* $\underline{\omega} := d\xi^1 \wedge d\xi^2 \wedge d\xi^3 \in \Omega^3\left(B; \tilde{\mathbb{R}}\right)$. *This time, we have:*

$$V = \int_B \underline{\omega}.$$

By [4.33], $\varphi^*(\underline{\omega}) = -\underline{\omega}$, *so*[9]

$$\varphi^*(V) = \int_B \varphi^*(\underline{\omega}) = \int_B (-\underline{\omega}) = -V. \qquad [4.35]$$

REMARK 4.56.– *Since the differential forms* ω *and* $\underline{\omega}$ *from Example 4.55 have the same expression* ($d\xi^1 \wedge d\xi^2 \wedge d\xi^3$)*, they are often written in the same way (even though the first belongs to* $\Omega^3(B)$ *and the second belongs to* $\Omega^3\left(B; \tilde{\mathbb{R}}\right)$*; strictly speaking, they are distinct objects).*

4.4.7. *Integration of a differential form over a chain*

(I) INTEGRATION OVER AN ODD SIMPLEX Recall that the standard m-simplex Δ^m in $\mathbb{R}^m$ is defined by

$$\Delta^m = \left\{ \sum_{i=0}^{m} t_i v_i : t_i \geq 0, \sum_{0 \leq i \leq m} t_i = 1 \right\},$$

where $v_0 = 0$ and $\{v_1, ..., v_m\}$ is the canonical basis of $\mathbb{R}^m$ ([P1], section 3.3.8(**V**)). Let U be an open neighborhood of Δ in $\mathbb{R}^m$, $\mathbf{F}$ a Banach space, and $\omega \in \Omega^m(U; \mathbf{F})$ an even m-form (Remark 4.32(b)). Since the space $\mathbb{R}^m$ is equipped with an orientation O, we can calculate $\int_{\overset{\frown}{U}} \chi_{\Delta^m}.\omega$, where χ_{Δ^m} is the characteristic function of Δ^m and $\overset{\frown}{U}$ is equipped with the orientation induced by O.

DEFINITION 4.57.– *Let* B *be a pure, metrizable,* m*-dimensional manifold. An* odd m*-simplex in* B *is a triple* $\underline{\tau} = (\Delta^m, \sigma, O)$*, where* Δ^m *is the standard* m*-simplex in* $\mathbb{R}^m$*,* U *is an open neighborhood of* Δ^m*,* σ *is a mapping of class* C^∞ *from* U *into* B *and* O *is an orientation of* $\mathbb{R}^m$.

In algebraic topology, B is a topological space, σ is only assumed to be continuous, O is the canonical orientation of $\mathbb{R}^m$ and $\underline{\tau}$ can be identified with σ. In differential geometry, where σ is of class C^∞, we sometimes specify that $\underline{\tau}$ is a *smooth simplex*.

9 In [4.34] and [4.35], we have avoided writing $\int_U d[\omega]$ and $\int_U d\breve{\omega}$ for the integrals relative to the measures $[\omega]$ and $\breve{\omega}$ respectively, differing from the convention adopted in ([P2], section 4.1.5(**II**)), as this would inevitably result in confusion with the exterior differential.

Let $\alpha \in \Omega^m(B;\mathbf{F})$ be an even m-form. The preimage $\sigma^*(\alpha)$ is an even m-form in $\Omega^m(U;\mathbf{F})$. Set:

$$\boxed{\int_{\underline{\tau}} \alpha := \int_{\widehat{U}} \chi_{\Delta^m}.\sigma^*(\alpha).}$$

(II) INTEGRATION OVER AN EVEN SIMPLEX Let $\Delta^m, U, \mathbf{F}, \sigma, B$ be as defined above.

DEFINITION 4.58.– *An* even *m-simplex in B is a pair $\tau = (\Delta^m, \tilde{\sigma})$, where $\tilde{\sigma} := (\sigma, \varepsilon)$ is an orientation of a mapping σ of class C^∞ from U into B (Definition 4.46).*

Consider an odd differential m-form $\underline{\alpha} \in \Omega^m\left(B;\tilde{\mathbb{R}} \otimes \mathbf{F}\right)$. Its preimage $\sigma^*(\underline{\alpha})$ is another odd differential m-form (Definition 4.51). Set:

$$\boxed{\int_\tau \underline{\alpha} = \int_U \chi_{\Delta^m}.\sigma^*(\underline{\alpha}).} \qquad [4.36]$$

REMARK 4.59.– *If $\alpha \in \Omega^p(B;\mathbf{F})$ (respectively $\underline{\alpha} \in \Omega^p\left(B;\tilde{\mathbb{R}} \otimes \mathbf{F}\right)$), where $p < m$, then $\int_{\widehat{\tau}} \alpha = 0$ (respectively $\int_\tau \underline{\alpha} = 0$), since we are integrating over a set of measure zero. If $\alpha \in \Omega(B;\mathbf{F})$ is an even non-homogeneous differential form $\sum_{0 \leq p \leq m} \alpha_p$, then $\int_\tau \alpha = \int_\tau \alpha_m$, and an analogous result holds when integrating an odd non-homogeneous differential form over an even simplex.*

(III) INTEGRATION OVER A CHAIN There are two cases to consider:

i) integration of an odd form over an even chain;

ii) integration of an even form over an odd chain.

The remarks up to and including part **(IV)** discuss the first case **(i)**. The second case **(ii)** is similar and is left to the reader.

Let $\{\tau_i : i \in I\}$ be a set of even m-simplexes in the metrizable, pure, m-dimensional manifold B. The free group with this set as a basis is the set $S_m(B)$ of linear combinations with integer coefficients

$$\tau = \sum_{i \in I} k_i.\tau_i, \qquad [4.37]$$

where all but finitely many of the k_i are zero ([P1], section 3.3.8**(VI)**). It is useful to consider the real vector space $S_m(B) \otimes_{\mathbb{Z}} \mathbb{R}$ formed by allowing these sums to have real coefficients; any such sum is called a *chain* of simplexes (*ibid.*).

EXAMPLE 4.60.– *A closed, convex polyhedron*[10] *in $\mathbb{R}^m$ is a finite intersection of closed half-spaces. Any closed polyhedron is a finite union of convex closed*

10 See the Wikipedia article *Polyhedron*.

*polyhedra ([BOU 82a], 11.3.1). A closed polyhedron is a chain of simplexes (see [P1], section 3.3.8(**VI**) for a demonstration of how to express a square as the sum of two simplexes), so a chain of closed polyhedra is a chain of simplexes.*

In the following, every chain is a chain of simplexes. If $\underline{\alpha} \in \Omega^m \left(B; \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$, set

$$\int_\tau \underline{\alpha} := \sum_{i \in I} k_i . \int_{\tau_i} \underline{\alpha}. \qquad [4.38]$$

This definition still makes sense in the obvious way if the chain $\tau = \sum_{i \in I} k_i \tau_i$ is allowed to be infinite (i.e. the sum is allowed to include infinitely many non-zero terms) and the set $\text{supp}\left(\underline{\alpha}\right) \cap \tau$ is compact.

(IV) CHANGE OF VARIABLE Let B' be another metrizable pure m-dimensional manifold, suppose that $f : B \to B'$ is a morphism, and let $\tilde{f} = (f, \varepsilon')$ be an orientation of f. If $\tau = (\Delta^m, \tilde{\pi})$ is an *even simplex* in B, then $f(\tau) := \left(\Delta^m, \tilde{f} \circ \tilde{\pi}\right)$ is an even simplex in B'. From [4.38], we can deduce the definition of $f(\tau)$ when τ is an *even chain* in B. By [4.36],

$$\boxed{\int_\tau \underline{\alpha} = \int_{f(\tau)} f^* \left(\underline{\alpha}\right).}$$

(V) BOUNDARY OF A CHAIN The i-th face of the standard m-simplex Δ^m in $\mathbb{R}^m$ is $\epsilon_i^m : \Delta^{m-1} \mapsto \Delta^m$, where ([P1], section 3.3.8(**VI**))

$$\epsilon_i^m : (t_0, ..., t_{m-1}) \mapsto \begin{cases} (0, t_0, ..., t_{m-1}) \text{ if } i = 0 \\ \left(t_0, ..., \widehat{t_i}, ..., t_{m-1}\right) \text{ if } 0 \leq i \leq m \end{cases}$$

Given an m-simplex $\tau = (\Delta^m, \tilde{\sigma})$, its *boundary* $\partial\tau$ is defined as follows for $m \geq 1$:

$$\partial\tau := \sum_{i=0}^{m} (-1)^i \sigma \circ \epsilon_i^m.$$

If $\tau \in S_m(B) \otimes_{\mathbb{Z}} \mathbb{R}$ is the chain defined by [4.37], where the τ_i are m-simplexes, write:

$$\partial\tau = \sum_{i \in I} k_i . \partial\tau_i. \qquad [4.39]$$

The boundary $\partial(\tau)$ is a chain with the same parity as the chain τ. If these chains are odd (*which will be assumed henceforth*), the orientation of $\partial(\tau)$ above is said to be *induced* by the orientation of τ. Consider, for example, the triangle Δ^2 in the plane, equipped with its canonical orientation, and write v_0, v_1, v_2 for its vertices; then ([P1], section 3.3.8(**V**)) $\partial\left(\Delta^2\right) = [v_0, v_1, \widehat{v_2}] + [\widehat{v_0}, v_1, v_2] + (-[v_0, \widehat{v_1}, v_2])$: see Figure 4.2.

REMARK 4.61.– *Consider one of the segments $[v_i, v_j]$. The interior $]v_i, v_j[$ of this segment is a submanifold of the plane (Oxy). Suppose that this submanifold is oriented as shown in Figure 4.2, writing ξ for this orientation and η for some point of $]v_i, v_j[$. The canonical orientation O of the plane (Oxy) is positive in the direction (Oz) and contains each of the elements $v \wedge u$, where v is a vector pointing strictly outward at η for Δ^2 (with the obvious meaning in this context) and u belongs to ξ. The orientation ξ is* induced *by O.*

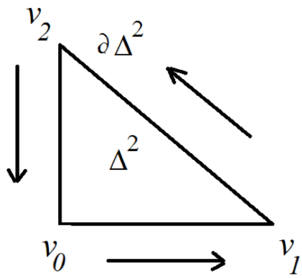


Figure 4.2. Δ^2 *and its boundary* $\partial\Delta^2$

LEMMA 4.62.– *Let $\partial_m : \Delta^m \to \Delta^{m-1}$. Then, $\partial_{m-1} \circ \partial_m = 0$.*

PROOF.– We have:

$$\partial_m [v_0, ..., v_m] = \sum_{i=0}^{m} (-1)^i [v_0, ..., \widehat{v_i}, ..., v_m],$$

$$\partial_{m-1} [v_0, ..., \widehat{v_i}, ..., v_m] = \sum_{j=0}^{i-1} (-1)^j [v_0, ..., \widehat{v_j}, ..., \widehat{v_i}, ..., v_m] + \sum_{k=i+1}^{m} (-1)^{k-1} [v_0, ..., \widehat{v_i}, ...\widehat{v_k}, ..., v_m],$$

and $\partial_{m-1} \circ \partial_m [v_0, ..., v_m] = 0$. ■

(VI) ORIENTATION OF A BOUNDARY Our next task is to specify the notion of the boundary of an odd chain, as well as the orientation of this boundary. This is straightforward if we proceed as described in Remark 4.61. Let τ be an odd chain in an m-dimensional manifold B $(m \geq 2)$ and write Fr (τ) for its frontier. The latter is the union of a *submanifold* $\partial\tau$ of dimension $m-1$, called the *regular boundary* (or simply the *boundary*) of τ (in the example from Remark 4.61, this is the union of the open segments $]v_i, v_j[$), and a set of points contained in a submanifold of dimension $m-2$ (the endpoints of the segments, in the example).

Let $b \in \partial\tau$ and suppose that ξ is an orientation of $T_b(\partial\tau)$. Write $\tilde{\imath}_b(\xi)$ for the orientation of $T_b(B)$ containing each of the vectors $v \wedge u$, where v is a vector pointing strictly outward from τ at b and u is a non-zero element of $\bigwedge^{m-1} T_b(\partial\tau)$ that belongs to the orientation ξ. The mapping $\tilde{\imath}_b$ is a bijection from Or $(T_x(\partial\tau))$ onto Or (B) and the $\tilde{\imath}_b$ $(b \in \partial\tau)$ determine a morphism $\tilde{\imath} : \widetilde{\partial\tau} \to \tilde{B}$ that is an orientation of the canonical injection $i : \partial\tau \hookrightarrow B$ (section 4.4.5**(IV)**).

DEFINITION 4.63.– *If O is an orientation of B, the orientation of $\partial\tau$ associated with O by $\tilde{\imath}$ (Corollary-Definition 4.47) is said to be* induced *by O.*

4.5. Pseudo-Riemannian manifolds

4.5.1. *Metric*

Let B be a Banach manifold of class C^r $(r \geq \infty)$ and $\mathbf{g} : (X, Y) \mapsto \mathbf{g}(X, Y) = \langle X|Y \rangle$ a twice covariant Hermitian tensor field of class C^r. For every $b \in B$,

$$\mathbf{g}(b) \in \mathbf{T}_2^0(T_b(B)) = \mathcal{L}_2(T_b(B), T_b(B); \mathbb{K}).$$

Consider the condition (**M**) and the weaker condition (**WM**) stated below:

(M) For every $b \in B$ and all $X_b \in T_b(B)$, the mapping

$$\mathbf{g}(Y_b, .)_b : T_b(B) \to \mathbb{K} : X_b \mapsto \langle Y_b|X_b \rangle_b$$

is an anti-linear bijection from $T_b(B)$ onto $T_b^{\vee}(B)$ ([P2], section 3.10.1(**I**)).

(WM) For every $b \in B$ and all $X_b \in T_b(B)$, the mapping

$$\mathbf{g}(Y_b, .)_b : T_b(B) \to \mathbb{K} : X_b \mapsto \langle Y_b|X_b \rangle_b$$

is non-degenerate (i.e. if $\mathbf{g}(Y_b, X_b)_b = 0$ for all $X_b \in T_b(B)$, then $Y_b = 0$).

If B is locally finite-dimensional (which is typically the case in practice), the conditions **(M)** and **(WM)** are equivalent.

DEFINITION 4.64.– *Let B be a manifold equipped with a twice covariant Hermitian tensor field* $\mathbf{g}$ *of class C^r. If* $\mathbf{g}$ *satisfies* **(M)** *(respectively* **(WM)***), the manifold B is said to be strongly (respectively weakly)* pseudo-Riemannian.[11] *The field* $\mathbf{g}$ *is called the* metric *or the* fundamental tensor field *of the manifold. This manifold is said to be strongly (respectively weakly)* Riemannian *if the Hermitian form* $\mathbf{g}(b)$ *is also positive definite for every $b \in B$.*

Suppose now that $\mathbb{K} = \mathbb{R}$ and let B be a pure m-dimensional manifold.

DEFINITION 4.65.– *A* frame $(\mathbf{h}_\alpha)_{1\leq\alpha\leq m}$ *of $T(B)$ (Definition 3.23) is said to be* orthonormal *for the metric* $\mathbf{g}$ *if*

$$\langle \mathbf{h}_\alpha (b) | \mathbf{h}_\beta (b) \rangle_b = \eta_b (\alpha, \beta) . \delta_\alpha^\beta \qquad [4.40]$$

for every $b \in B$, where $\eta_b(\alpha,\beta) \in \{1,-1\}$. We say that $\mathbf{g}(b)$ *has* signature $(p_b, m-p_b)$ *if, for every $\beta \in \{1, ..., m\}$, Card$\{\alpha \in \{1,...,m\} :$ sgn $(\eta_b(\alpha,\beta)) = 1\} = p_b$.*

If B is connected, the signature of $\mathbf{g}$ is constant. Let (U, ξ, m) be a chart; we can express $\mathbf{g}$ locally (over U) by

$$\boxed{\mathbf{g} = \sum_{i,j=1}^m g_{ij} . d\xi^i \otimes d\xi^j,} \qquad [4.41]$$

where the square matrix $G(b) = (g_{ij}(b))$ is real, symmetric, and nowhere singular. Any such chart is said to be a system of *normal pseudo-Riemannian coordinates* at the point $b \in B$ if the system of tangent vectors $\left\{ \left.\frac{\partial}{\partial\xi^1}\right|_b, ..., \left.\frac{\partial}{\partial\xi^m}\right|_b \right\}$ is orthonormal for the real symmetric form $\langle . | . \rangle_b$.

The real manifold B is *Riemannian* if the symmetric form $\mathbf{g}(b)$ has signature $(m, 0)$ at every point; if $m \geq 2$ and the signature is $(1, m-1)$ or $(m-1, 1)$ at every point, then the manifold B is said to be *Lorentzian* (the Lorentzian manifold of dimension 4 is encountered in general relativity; we will call it the *Einstein manifold*).

4.5.2. *Pseudo-Riemannian volume element*

Let $\{\sigma^\alpha : 1 \leq \alpha \leq m\}$ be the dual coframe of the orthonormal frame $(\mathbf{h}_\alpha)_{1\leq\alpha\leq m}$ (Lemma-Definition 3.38). The pseudo-Riemannian volume element for which the frame $(\mathbf{h}_\alpha)_{1\leq\alpha\leq m}$ is positively oriented is (Definition 4.37 and section 4.4.5**(III)**):

$$\boxed{\omega = \sigma^1 \wedge ... \wedge \sigma^m.}$$

11 In the case where B is locally finite-dimensional (the only case considered below), the adverbs *strongly* and *weakly* can be omitted.

Let $(\mathbf{h}_{\beta'})$ be another, arbitrary frame. Let $\left(\sigma^{\beta'}\right)$ be its dual, and suppose that $A = (A^{\alpha}_{\alpha'})$ is the *inverse* change-of-basis matrix satisfying $\mathbf{h}_{\alpha'} = \sum_{\alpha} \mathbf{h}_{\alpha} A^{\alpha}_{\alpha'}$ (see [1.2], section 1.2.1(**II**)). Then, by [4.40], using the fact that $\mathbf{g}$ is bilinear,

$$g_{\alpha'\beta'} = \sum_{\alpha,\beta} A^{\alpha}_{\alpha'} A^{\beta}_{\beta'} \mathbf{g}\left(\mathbf{h}_{\alpha}, \mathbf{h}_{\beta}\right) = \sum_{\alpha} A^{\alpha}_{\alpha'} A^{\alpha}_{\beta'} \eta,$$

where $g_{\alpha'\beta'} = \mathbf{g}\left(\mathbf{h}_{\alpha'}, \mathbf{h}_{\beta'}\right)$. Writing $(g_{\alpha'\beta'})$ for the matrix with entries $g_{\alpha'\beta'}$, we therefore have $\det(g_{\alpha'\beta'}) = \det(A)^2 \eta$, which implies that

$$\det(A)^2 = |\det(g_{\alpha'\beta'})| \Rightarrow \det(A) = \varepsilon\sqrt{|\det(g_{\alpha'\beta'})|},$$

where $\varepsilon = +1$ if the frame $(\mathbf{h}_{\beta'})$ is positively oriented and $\varepsilon = -1$ otherwise. Moreover, $\sigma^{\alpha} = \sum_{\alpha'} A^{\alpha}_{\alpha'} \sigma^{\alpha'}$, so

$$\omega = \sum_{\sigma \in \mathfrak{S}_m} \varepsilon_{\sigma} A^{1}_{\sigma(1)} ... A^{m}_{\sigma(m)} \sigma'^{1} \wedge ... \wedge \sigma'^{m} = \det(A) . \sigma'^{1} \wedge ... \wedge \sigma'^{m}$$

(see [P1], section 2.3.11(**V**)) and hence

$$\boxed{\omega = \varepsilon\sqrt{|\det(g_{\alpha'\beta'})|} . \sigma'^{1} \wedge ... \wedge \sigma'^{m}.} \qquad [4.42]$$

5

Differential and Integral Calculus on Manifolds

5.1. Introduction

The notion of a distribution, which we studied on open subsets of $\mathbb{R}^n$ in [P2], section 4.4.1, can be defined in the same way on a locally compact manifold B that is countable at infinity[1]. The notion of a current, due to de Rham ([DER 84], Chapter 3) and further specified by L. Schwartz ([SCH 66], Chapter 9), generalized the notion of a distribution: while the Dirac distribution represents the electric charge of a sphere as the radius becomes infinitely small, a current represents the electric current density in a conductive wire as the cross-section becomes infinitely small. Point distributions (such as the Dirac distribution $\delta_b : \varphi \mapsto \varphi(b)$ and its derivatives, where $\varphi \in \mathcal{E}(B)$) are defined on Banach manifolds and can be identified with scalar differential operators.

The Lie derivative of a function along a tangent vector was defined in Lemma-Definition 2.29(i). The notion of a Lie derivative along a vector field immediately follows; it can be extended to the Lie derivative of a tensor field, and in particular, a differential form.

The notion of an exterior differential, due to É. Cartan ([CAR 45], section 2.1)[2], provides the framework needed to establish one of the major results of differential geometry: Stokes' formula. There are myriad versions of this formula; in the words of L. Schwartz ([SCH 93], Volume IV, p. 352), "There are no good conditions for applying Stokes' formula". The history of the formula is complex [KAT 79]; the

1 In distribution theory and Radon measure theory, countability at infinity is not strictly necessary but allows simplifications to be made ([P2], Remark 4.1).

2 This generalizes the previous idea of the "bilinear covariant" of a 1-form (introduced by Lipschitz in 1870).

earliest version appeared in 1813 in a result stated by Gauss. Ostrogradsky gave a more sophisticated version in 1827, followed by Green in 1828, Riemann in 1851, then W. Thomson (Lord Kelvin) and Stokes around 1854. All of these mathematicians were also physicists. Their work, which led to very different versions of Stokes' formula, was motivated by problems from physics, especially electromagnetism. The unification of these formulas, plus another by Cauchy (1846), was achieved by Volterra (1889), Poincaré (1899) and most notably É. Cartan (1922) with the introduction of differential forms and exterior derivatives [CAR 22a]. Stokes' formula on a chain is due to de Rham [DER 84]; this is the version that we have chosen to consider here (drawing from the classification by [BOU 82a]) as a good compromise between simplicity and generality. Whitney gave a more general version, at the price of greater complexity ([WHI 57], Chapter 3, section 18). The de Rham homology and the dual notion of the cohomology of currents or chains, established from the work by Poincaré and de Rham – enable us to approach the global questions which follow naturally from Stokes' formula.

Unfortunately, it would be impossible for us to delve any deeper into the vast subject of partial differential equations; there are simply too many different perspectives and recent developments. For classical theory, readers can refer to ([COU 58], Volume 2), as well as [PET 91], which is more succinct. For our purposes, however, we will need to study differential equations on manifolds so that we can define geodesics in Chapter 7, as well as integral manifolds and the Frobenius theorem – a result that is regularly exploited by the theory of nonlinear dynamic systems. This theorem was discovered by H. Deahna in 1840 but expressed in a relatively obscure form; it was later clarified by Frobenius in 1875 [SAM 01].

Up to and including section 5.3, every manifold is a differential manifold.

5.2. Currents and differential operators

5.2.1. *Currents and distributions*

In this section, every manifold is locally compact and countable at infinity, hence paracompact ([P2], section 2.3.10), and furthermore C^∞-paracompact by Corollary 2.15.

(I) THE SPACES $\Omega^p(B)$ **AND** $\Omega_c^p(B)$ Let B be a pure m-dimensional manifold and p an integer such that $0 \leq p \leq m$. Recall that $\mathcal{E}(B)$ denotes the ring of C^∞ mappings from B into $\mathbb{C}$. The space $\Omega^p(B)$ of p-forms of class C^∞ on B is, like $\mathcal{E}(B)$ itself, a Fréchet nuclear space ([P2], sections 4.3.1**(I)** and 4.3.2**(III)**). The space $\Omega_c^p(B)$ of compactly supported p-forms of class C^∞ on B is of type $(\mathcal{L}_s\mathcal{F})$ (a strict inductive limit of Fréchet spaces) and is a Montel nuclear space (*ibid.*).

Write

$$\Omega(B) = \bigoplus_{0 \le p \le m} \Omega^p(B), \quad \Omega_c(B) = \bigoplus_{0 \le p \le m} \Omega_c^p(B)$$

(direct topological sums: see [P2], section 3.3.7).

(II) THE SPACES $\Omega^p\left(B;\tilde{\mathbb{R}}\right)$ **AND** $\Omega_c^p\left(B;\tilde{\mathbb{R}}\right)$ The space $\Omega^p\left(B;\tilde{\mathbb{R}}\right)$ of odd p-forms of class C^∞ on B and the space $\Omega_c^p\left(B;\tilde{\mathbb{R}}\right)$ of compactly supported odd p-forms of class C^∞ on B are locally convex spaces of same type as $\Omega^p(B)$ and $\Omega_c^p(B)$, respectively. When the manifold B is implicitly clear, write $\Omega^p, \Omega_c^p, \underline{\Omega^p}, \underline{\Omega_c^p}$ for $\Omega^p(B), \Omega_c^p(B), \Omega^p\left(B;\tilde{\mathbb{R}}\right)$, and $\Omega_c^p\left(B;\tilde{\mathbb{R}}\right)$, respectively, as well as

$$\underline{\Omega} = \bigoplus_{0 \le p \le m} \underline{\Omega^p}, \quad \underline{\Omega_c} = \bigoplus_{0 \le p \le m} \underline{\Omega_c^p}.$$

(III) CURRENTS An odd (respectively even) p-current on B (for $0 \le p \le m$) is a continuous linear form on $\Omega_c^p(B)$ (respectively $\Omega_c^p\left(B;\tilde{\mathbb{R}}\right)$)[3]. A compactly supported odd (respectively even) p-current on B (for $0 \le p \le m$) is a continuous linear form on $\Omega^p(B)$ (respectively $\Omega^p\left(B;\tilde{\mathbb{R}}\right)$). The support of a current is defined in the same way as the support of a Radon measure or a distribution, and a gluing principle can also be defined for currents ([P2], sections 4.1.5**(I)** and 4.4.1**(I)**).

If $\varphi \in \Omega_c^p$ (respectively $\underline{\varphi} \in \underline{\Omega_c^p}$) and $T \in \Omega_c^{p\vee}$ (respectively $\underline{T} \in \underline{\Omega_c^p}^\vee$), then the value $T\left(\underline{\varphi}\right)$ (respectively $\underline{T}(\varphi)$) is written as $\left\langle T, \underline{\varphi}\right\rangle$ (respectively $\langle \underline{T}, \varphi\rangle$). If $\varphi = \sum_{0 \le p \le m} \overset{p}{\varphi}$, where $\overset{p}{\varphi} \in \Omega_c^p$, and $\underline{T} = \sum_{0 \le p \le m} \overset{p}{\underline{T}}$, where $\overset{p}{\underline{T}} \in \Omega_c^{p\vee}$, set (see [P2], section 3.6.4, Theorem 3.105):

$$\langle \underline{T}, \varphi\rangle = \sum_{0 \le p \le m} \left\langle \overset{p}{\underline{T}}, \overset{p}{\varphi}\right\rangle.$$

Similarly, with the obvious adjustments to the notation, set:

$$\left\langle T, \underline{\varphi}\right\rangle = \sum_{0 \le p \le m} \left\langle \overset{p}{T}, \overset{p}{\underline{\varphi}}\right\rangle.$$

It can be shown in the same way as [P2], section 4.4.1 that these spaces of p-currents are all Montel nuclear spaces (and hence reflexive) and that the spaces of compactly supported p-currents are Silva nuclear spaces.

---

3 We will prefer the terminology of ([DIE 93], Volume 3, Chapter 17, section 3) over that of ([SCH 66], Chapter 9), which uses the term $(m-p)$-current for what we are calling a p-current.

EXAMPLE 5.1.– Dirac current*: Let $b \in B$ and suppose that $\mathbf{z}_b$ is a p-vector belonging to $\bigwedge^p T_b(B)$ (section 4.2.3(**I**), Definition 4.7). Furthermore, let $\omega \in \Omega^p(B)$. Then, $\omega(b) \in \bigwedge^p T_b(B)^\vee$, so we can form the duality bracket*

$$\underline{\delta_{\mathbf{z}_b}}(\omega) := \langle \mathbf{z}_b, \omega(b) \rangle .$$

The mapping $\omega \mapsto \underline{\delta_{\mathbf{z}_b}}(\omega)$ is an odd p-current $\underline{\delta_{\mathbf{z}_b}}$ such that supp $\left(\underline{\delta_{\mathbf{z}_b}}\right) = \{b\}$, which generalizes the Dirac distribution on the real line.

(IV) CHAINS OF SMOOTH SIMPLEXES Any even m-chain (respectively compactly supported even m-chain) τ of smooth simplexes on B uniquely determines an even m-current (respectively compactly supported even m-current) T by the relation

$$T : \underline{\omega} \mapsto \int_\tau \underline{\omega} \quad (\underline{\omega} \in \underline{\Omega_c^m}, \text{ resp. } \underline{\omega} \in \underline{\Omega^m}).$$

Similarly, any odd m-chain (respectively compactly supported odd m-chain) $\overset{\curvearrowright}{\tau}$ of smooth simplexes on an *oriented* manifold B uniquely determines an odd m-current (respectively compactly supported odd m-current) $\underline{T}$ by the relation

$$\underline{T} : \omega \mapsto \int_{\overset{\curvearrowright}{\tau}} \omega \quad (\omega \in \Omega_c^m, \text{ resp. } \omega \in \Omega^m). \qquad [5.1]$$

(V) EXTERIOR PRODUCT OF A CURRENT AND A FORM On a pure m-dimensional manifold B, let T be an even p-current $(0 \le p \le m)$ and $\underline{\omega}$ an odd q-form $(0 \le q \le p)$. For every odd $(p-q)$-form $\underline{\beta}$, $\underline{\omega} \wedge \underline{\beta}$ is an odd p-form. We can therefore form the quantity $\left\langle T, \underline{\omega} \wedge \underline{\beta} \right\rangle$. Thus, $\underline{\beta} \longmapsto \left\langle T, \underline{\omega} \wedge \underline{\beta} \right\rangle$ is a $(p-q)$-current, denoted $T \wedge \underline{\beta}$. Hence, by definition:

$$\boxed{\left\langle T \wedge \underline{\beta}, \underline{\omega} \right\rangle = \left\langle T, \underline{\omega} \wedge \underline{\beta} \right\rangle.}$$

In the special case where T is a distribution (i.e. when $p = 0$), $\underline{\beta}$ is a function of class C^∞ on B and $T \wedge \underline{\beta}$ is the product $\underline{\beta}.T$ of T and $\underline{\beta}$ ([P2], section 4.4.1(**IV**)).

The same reasoning is valid on an oriented manifold with an odd current and even forms.

(VI) Currents taking values in a fiber bundle Let $\pi : N \to B$ be a fiber bundle whose fibers N_b are of type $\mathbf{F}$, where $\mathbf{F}$ is a finite-dimensional vector space. The odd differential p-forms taking values in N are the elements of $\Omega^p\left(B, \tilde{\mathbb{R}} \otimes N\right)$ (section 4.4.6(**I**)).

Let $\underline{\omega} \in \Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes N\right)$ and $\underline{\varphi} \in \Omega_c^{m-p}\left(B, \tilde{\mathbb{R}} \otimes N^\vee\right)$, where $N^\vee$ is the dual of N (section 3.4.2, Definition 3.36). Then, $\underline{\varphi} \wedge \underline{\omega}$ is a compactly supported section of $\mathrm{Alt}^m\left(\tilde{\mathbb{R}} \otimes N^\vee \otimes N\right)$ of class C^∞. The universal property of the tensor product ([P1], section 3.1.5(**I**)) implies the existence of a "contraction operation" $\kappa : N^\vee \otimes N \to \mathbb{R}$ that is linear on fibers and which satisfies $\kappa\left(x \otimes x^\vee\right) = \left\langle x, x^\vee\right\rangle$ for every $\left(x^\vee, x\right) \in N^\vee \times N$. This operation κ uniquely determines a linear operation (whose linearity is induced by the linearity on fibers)

$$\hat{\kappa} : \Omega_c^{m-p}\left(B, \tilde{\mathbb{R}} \otimes N^\vee\right) \wedge \Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes N\right) \to \Omega_c^m\left(B, \tilde{\mathbb{R}}\right)$$

such that $\hat{\kappa}\left(\underline{\varphi} \wedge \underline{\omega}\right) \in \Omega_c^m\left(B, \tilde{\mathbb{R}}\right)$. Thus,

$$T_{\underline{\varphi}} : \underline{\omega} \longmapsto \int_B \hat{\kappa}\left(\underline{\varphi} \wedge \underline{\omega}\right)$$

is a continuous linear form on $\Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes N\right)$, said to be an $(m-p)$-current taking values in $N^\vee$.

The case where N is the trivial bundle $\mathbf{F}_B$ is completely straightforward: $\Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$ is a locally convex vector space of same type as $\Omega_c^p\left(B, \tilde{\mathbb{R}}\right)$ (see (**III**)) and the set of the $T_{\underline{\varphi}}$ specified above is the dual $\Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes \mathbf{F}\right)^\vee$; this dual is a complete Montel nuclear and ultrabornological space ([P2], section 4.4.1(**I**)). In particular, when $\mathbf{F} = \mathbb{C}$, the spaces $\Omega_c^p\left(B, \tilde{\mathbb{R}} \otimes \mathbb{C}\right)^\vee$ and $\Omega_c^0\left(B, \tilde{\mathbb{R}} \otimes \mathbb{C}\right)^\vee$ are the space of p-fields and the space of *complex* distributions on B, respectively.

The above cannot be extended to the case of vector bundles of infinite rank without introducing restrictions, even for distributions ([SCH 58], p. 140, Example 2).

(VII) Tensor product of currents Let A and B be pure n- and m-dimensional manifolds, respectively, and suppose that $\underline{\omega} \in \Omega^{p+q}\left(A \times B; \tilde{\mathbb{R}}\right)$ $(0 \leq p \leq n, 0 \leq q \leq m)$. Over the domain U of a chart of $A \times B$, we can decompose $\underline{\omega}$ into $\underline{\omega} = \sum_{i,j} \underline{\omega_i} \wedge d\xi^j$, where $\underline{\omega_i} \in \Omega^p\left(A; \tilde{\mathbb{R}}\right)$, $j = (j_1, ..., j_p)$ and $d\xi^j = d\xi^{j_1} \wedge ... \wedge d\xi^{j_p}$. Let $T \in \Omega^p\left(A; \tilde{\mathbb{R}}\right)^\vee$ be a p-current. Then, over U:

$$\langle T, \underline{\omega}\rangle = \sum_{i,j} \left\langle T, \underline{\omega_i}\right\rangle . d\xi^j \in \Omega^q\left(B; \tilde{\mathbb{R}}\right).$$

If $S \in \Omega^q\left(B;\tilde{\mathbb{R}}\right)^\vee$, we can form the quantity $\langle S, \langle T, \underline{\omega}\rangle\rangle$ and define the tensor product $S \otimes T \in \Omega^{p+q}\left(A \times B;\tilde{\mathbb{R}}\right)^\vee$ by the following the relation (which holds for any odd form $\underline{\omega} \in \Omega^{p+q}\left(A \times B;\tilde{\mathbb{R}}\right)$):

$$\boxed{\langle S \otimes T, \underline{\omega}\rangle = \langle S, \langle T, \underline{\omega}\rangle\rangle.}$$

This generalizes the tensor product of two measures ([P2], section 4.1.3(**II**)).

(VIII) Distributions on a manifold Write $\mathcal{D}$ for Ω_c^0 and $\mathcal{E}$ for Ω^0. In accordance with [P2], section 4.4.1, any distribution (respectively compactly supported distribution) on a pure m-dimensional *oriented* manifold B (e.g. an open subset U of $\mathbb{R}^m$ equipped with its canonical orientation) is an odd 0-current (respectively a compactly supported odd 0-current) that belongs to $\mathcal{D}'$ (respectively $\mathcal{E}'$)[4]. A distribution $T \in \mathcal{D}'(B)$ is said to be *real* if, for every *real* test function $\varphi \in \mathcal{D}(B)$, $\langle T, \varphi\rangle$ is a real number. Wherever necessary, to avoid ambiguity, we will say that a distribution takes values in $\mathbb{K}$, with $\mathbb{K} = \mathbb{C}$ (respectively $\mathbb{K} = \mathbb{R}$) when this distribution is complex (respectively real).

We need the following idea from general topology ([BOU 74], Chapter 1, section 10.3, Proposition 7):

Definition 5.2.– *Let X and Y be locally compact topological spaces and suppose that $f : X \to Y$ is a continuous mapping. We say that f is* proper *if, for every compact subset K of Y, $f^{-1}(K)$ is compact.*

Let B, B' be oriented manifolds, T a distribution (respectively a compactly supported distribution) on B and $u : B \to B'$ a proper mapping (respectively a mapping) of class C^∞. For every test function $\varphi \in \mathcal{D}(B')$ (respectively $\varphi \in \mathcal{E}(B')$), we have $\varphi \circ u \in \mathcal{D}(B)$ (respectively $\varphi \circ u \in \mathcal{E}(B)$), and $u(T) : \varphi \mapsto \langle T, \varphi \circ u\rangle$ is a distribution (respectively a compactly supported distribution) on B', called the *image* of the distribution T under u (compare with the image of a Radon measure in [P2], section 4.1.5(**II**)).

Let (U, ξ, n) be a chart of B, $T \in \mathcal{D}'(B)$, and $\alpha = (\alpha_1, ..., \alpha_n)$ a multi-index. Write $D_\xi^\alpha T = D^\alpha(\xi(T))$, where $\xi(T)$ is the image of the restriction of T to U under ξ and $D^\alpha(\xi(T))$ is defined as in [P2], section 4.4.1(**II**).

4 This convention is used by [DIE 93] (Volume 3, Chapter 17, section 3) but not by [SCH 66] (Chapter 9), where a distribution is an even m-current (with the definition of p-current adopted above); see footnote 3, p. 175. The space $\mathbb{R}^m$ is implicitly equipped with its canonical orientation in [P2], section 4.4.1.

Let B be an oriented manifold, $T \in \mathcal{D}'(B)$, and $\varphi \in \mathcal{E}(B)$. If $K := \operatorname{supp}(T) \cap \operatorname{supp}(\varphi)$ is compact, set $\langle T, \varphi \rangle = \langle \varphi.T, 1 \rangle$, which is well-defined because $\varphi.T \in \mathcal{E}'(B)$ and $1 \in \mathcal{E}(B)$ (where 1 denotes the constant function equal to 1).

The notion of the restriction of a distribution to an open subset of a manifold is defined in the usual way ([P2], section 4.4.1(**I**)). Let $\pi : B' \to B$ be a local diffeomorphism (Lemma-Definition 2.57). For every distribution $T \in \mathcal{D}'(B)$, there exists a distribution T' on B' uniquely determined by the following condition: for every open subset U' of B' such that the restriction $\pi|_{U'} : U' \to \pi(U')$ is a diffeomorphism, $\pi|_{U'}(T'|_{U'}) = T_{\pi(U')}$ ([DIE 93], Volume 3, (17.4.4)). We say that T' is the *preimage* of T under π, written as $\pi^*(T)$. (The same construction can be performed for currents.)

If $T \in \mathcal{D}'(B)$ and $\varphi \in \mathcal{D}(B)$, or if $T \in \mathcal{E}'(B)$ and $\varphi \in \mathcal{E}(B)$, we will adopt the following convention (strictly speaking an abuse of notation, but convenient and unambiguous nonetheless):

$$\langle T, \varphi \rangle = \int \varphi(x)\,.dT(x) = \langle T(x), \varphi(x) \rangle. \qquad [5.2]$$

Let $S \in \mathcal{D}'(A)$ and $T \in \mathcal{D}'(B)$. Then, $\operatorname{supp}(S \otimes T) = \operatorname{supp}(S) \times \operatorname{supp}(T)$ (**exercise**). If $\theta \in \mathcal{D}(A \times B)$,

$$\langle S \otimes T, \theta \rangle = \langle S(x), \langle T(y), \theta(x,y) \rangle \rangle = \langle T(y), \langle S(x), \theta(x,y) \rangle \rangle.$$

The quantity $\langle S \otimes T, \theta \rangle := \langle \theta.S \otimes T, 1 \rangle$ is well defined whenever $(\operatorname{supp}(S) \times \operatorname{supp}(T)) \cap \operatorname{supp}(\theta)$ is compact. It can be shown that the bilinear mapping $\mathcal{D}'(A) \times \mathcal{D}'(B) \to \mathcal{D}'(A \times B) : (S,T) \mapsto S \otimes T$ is continuous (for the strong dual topologies). The same is true if we replace $\mathcal{D}'$ by $\mathcal{E}'$ ([SCH 66], Chapter 4, section 4, Theorem VI). Furthermore, since the spaces $\mathcal{D}'$ and $\mathcal{E}'$ are nuclear ([P2], section 4.4.1(**I**)), their completions $\mathcal{D}'(A)\,\widehat{\otimes}\,\mathcal{D}'(B)$ and $\mathcal{E}'(A)\,\widehat{\otimes}\,\mathcal{E}'(B)$ are unambiguously defined ([P2], section 3.11.3(**I**)) and we have ([TRÈ 67], Chapter 51):

$$\boxed{\mathcal{D}'(A \times B) \cong \mathcal{D}'(A)\,\widehat{\otimes}\,\mathcal{D}'(B), \quad \mathcal{E}'(A \times B) \cong \mathcal{E}'(A)\,\widehat{\otimes}\,\mathcal{E}'(B).} \qquad [5.3]$$

A similar relation holds for $\mathcal{S}'$: see Theorem 6.149 in section 6.5.

(IX) CONVERGENCE OF SEQUENCES OF DISTRIBUTIONS Let B be an oriented manifold and (T_i) a sequence or bounded net of distributions that belong to $\mathcal{T}'(B)$, where $\mathcal{T} = \mathcal{D}$, $\mathcal{E}$, or $\mathcal{S}$. Write $\mathcal{T}'_b(B)$ for the strong dual of $\mathcal{T}(B)$ and $\mathcal{T}'_s(B)$ for its weak* dual ([P2], section 3.6).

THEOREM 5.3.– *The following conditions are equivalent: (i)* (T_i) *converges to the distribution* U *in* $\mathcal{T}'_b(B)$ *. (ii) For every function* $\varphi \in \mathcal{T}(B)$*,* $\langle T_i, \varphi\rangle \to \langle U, \varphi\rangle$ *.*

PROOF.– As recalled in **(I)**, $\mathcal{T}(B)$ is a Montel space, so (T_i) converges to U in $\mathcal{T}'_b(B)$ if and only if (T_i) converges to U in $\mathcal{T}'_s(B)$ ([P2], section 3.7.5, Corollary 3.124). This last property is equivalent to (ii) ([P2], section 3.6.2). ■

COROLLARY 5.4.– *Let* μ *be a Lebesgue measure on* B *(Corollary-Definition 4.42),* b *a point of* B*,* δ_b *the Dirac distribution at the point* b *and* (f_i) *a sequence of* μ*-measurable functions (respectively compactly supported* μ*-measurable functions) from* B *into* $\mathbb{R}_+$ *such that, for any compact neighborhood* V *of* b *in* B*,* $\int_V f_i.d\mu \to 1$ *as* $i \to +\infty$*. Then,* $f_i.\mu \to \delta_b$ *in* $\mathcal{D}'_b(B)$ *(respectively in* $\mathcal{E}'_b(B)$*) as* $i \to +\infty$ *(if* $B = \mathbb{R}^m$ *and* $\mu = \lambda^{\otimes m}$*, we simply write that* $f_i \to \delta_b$*, and, if each* f_i *is tempered, the convergence also takes place in* $\mathcal{S}'(\mathbb{R}^m)$*).*

PROOF.– We will give the reasoning when $\mathcal{T} = \mathcal{D}$. The reader is invited to adapt the proof to the other cases as an **exercise**.

First, note that $\int_K f_i.d\mu \to 0$ as $i \to +\infty$ for any compact set K that does not contain b, since, if V is a neighborhood of b with no points in common with K, then $\int_V f_i.d\mu \to 1$ and $\int_{V\cup K} f_i.d\mu \to 1$ as $i \to +\infty$. But $\int_{V\cup K} f_i.d\mu = \int_V f_i.d\mu + \int_K f_i.d\mu$.

Let $\varphi \in \mathcal{D}(B)$. Given $\varepsilon > 0$, there exists an open neighborhood Ω of b in B such that, for every $x \in \Omega$, $|\varphi(x) - \varphi(a)| \leq \varepsilon/4$. Let V be a compact neighborhood contained in Ω; we again have $|\varphi(x) - \varphi(a)| \leq \varepsilon/4$ for every $x \in V$. Furthermore,

$$\left(\int_B f_i\varphi.d\mu\right) - \varphi(b) = \int_{B-V} f_i\varphi.d\mu + \int_V f_i(\varphi - \varphi(b)).d\mu$$
$$+\varphi(b)\left[\left(\int_V f_i d\mu\right) - 1\right]. \quad [5.4]$$

Consider the first term on the right-hand side of [5.4]. If $\operatorname{supp}(\varphi) \subset K$, then, setting $K' = \overline{(B-V)\cap K}$,

$$\left|\int_{B-V} f_i\varphi.d\mu\right| = \left|\int_{(B-V)\cap K} f_i\varphi.d\mu\right| \leq \sup_{x\in B}|\varphi(x)|\,.\int_{K'} f_i.d\mu,$$

so this quantity tends to 0 as $i \to +\infty$. The same is true for the absolute value of the third term on the right-hand side of [5.4]. Hence, there exists an integer i_0 such that the sum of the absolute values of the first and third terms of [5.4] is $\leq \varepsilon/2$ whenever $i \geq i_0$. Furthermore,

$$\left|\int_V f_i(\varphi - \varphi(b)).d\mu\right| \leq \frac{\varepsilon}{4}\int_V f_i.d\mu$$

and $\int_V f_i.d\mu \to 0$ as $i \to +\infty$, so there exists an integer i'_0 such that $\left|\int_V f_i\left(\varphi - \varphi(b)\right).d\mu\right| \leq \varepsilon/2$ whenever $i \geq i'_0$. Finally, setting $i''_0 = \max(i_0, i'_0)$, we have $\left|\left(\int_B f_i\varphi.d\mu\right) - \varphi(b)\right| \leq \varepsilon$ whenever $i \geq i''_0$. Hence, $\int_B f_i\varphi.d\mu \to \varphi(b) = \langle\delta_b, \varphi\rangle$ as $i \to +\infty$. ■

5.2.2. *Differential operators and point distributions*

(I) DIFFERENTIAL OPERATORS Let B be a pure q-dimensional manifold that is locally compact and countable at infinity and $M \twoheadrightarrow B$, $N \twoheadrightarrow B$ two complex vector bundles of finite ranks m and n, respectively (section 3.4.1, Definition 3.22). The space $\Gamma(B, M)$ of sections of class C^∞ of M is a Fréchet nuclear space, like $\mathcal{E}(U)$ itself whenever U is an open subset of $\mathbb{R}^q$ ([P2], sections 4.3.1**(I)** and 4.3.2**(III)**). Hence, this space is separable ([P2], section 3.11.3**(I)**).

DEFINITION 5.5.– *A linear* differential operator *of class C^∞ from M into N is a continuous linear mapping $P : f \mapsto P.f$ from $\Gamma(B, M)$ into $\Gamma(B, N)$ that satisfies the following condition:*

(L) *For every open subset U of B and every morphic section $f \in \Gamma(B, M)$ such that $f|_U = 0$, we have $(P.f)|_U = 0$.*

The condition **(L)** expresses the *local* nature of the operator P. Write $\text{Diff}(B; M, N)$ for the set of these differential operators; this is an $\mathcal{E}(B)$-module.

The local trivialization condition **(V)** of the vector bundles M and N (Definition 3.22(i)) implies that, for every $b \in B$, there exists an open neighborhood U of b that is the domain of a chart $c = (U, \xi, q)$ of B over which these two fibers can be identified with the trivial bundles $U \times \mathbb{C}^m$ and $U \times \mathbb{C}^n$, respectively. Hence, for every section $f \in \Gamma(B, M)$, there exist a mapping $g_V \in \mathcal{E}(V)$, with $V = \xi(U)$, and a linear differential operator $Q : \mathcal{E}(V)^m \to \mathcal{E}(V)^n$ such that both squares of the following diagram commute (the rows of this diagram are not compositions):

$$\begin{array}{ccccc} U & \xrightarrow{f|_U} & U \times \mathbb{C}^m & \xrightarrow{(P.f)|_U} & U \times \mathbb{C}^n \\ \downarrow \xi & & \downarrow \xi \times 1_m & & \downarrow \xi \times 1_n \\ V & \xrightarrow{g_V} & V \times \mathbb{C}^m & \xrightarrow{Q.g_V} & V \times \mathbb{C}^n \end{array}$$

We say that Q is the local expression of P corresponding to the chart c (and the local trivializations specified above). Given the topology of $\mathcal{E}(V)$ ([P2], section 4.3.1**(I)**), with the notation of section 1.2.4**(IV)**, the operator Q is of the form

$$Q = \sum_{|\alpha| \leq p} A_\alpha.D^\alpha,$$

where $x \mapsto A_\alpha(x)$ $(x = \xi(b))$ is a mapping of class C^∞ from $V = \xi(U)$ into $\mathrm{Hom}(\mathbb{C}^m, \mathbb{C}^n) \cong \mathbb{C}^{m \times n}$ (**exercise***: see [DIE 93], Volume 3, (17.13.3)). The *order* of the differential operator P at the point b is defined as the greatest integer $|\alpha|$ such that $A_\alpha \neq 0$.

If M and N are the trivial bundles $B \times \mathbb{C}^m$ and $B \times \mathbb{C}^n$, respectively, then, for any mapping $f : B \to \mathbb{C}^m$ of class C^∞, we have $P.f = \sum_{|\alpha| \leq p} A_\alpha.D^\alpha(f \circ \xi)$. Hence, setting $D_\xi^\alpha(f) = D^\alpha(f \circ \xi)$,

$$P = \sum_{|\alpha| \leq p} A_\alpha.D_\xi^\alpha. \qquad [5.5]$$

If M and N are both equal to the trivial bundle $B \times \mathbb{C}$, then $\Gamma(B, M)$ and $\Gamma(B, N)$ can both be identified with $\mathcal{E}(B)$, in which case $\mathrm{Diff}(B; M, N)$ is simply written as $\mathrm{Diff}(B)$.

(II) Sheaf of differential operators For every $b \in B$ and every open neighborhood U of b, let $M|_U$ and $N|_U$ be the vector bundles induced by M and N, respectively, on U (section 3.3.1, Lemma-Definition 3.4(4)). Let $h \in \mathcal{E}(B)$ be a mapping such that $\mathrm{supp}(h) \subset U$ and h is equal to 1 in a neighborhood $W \subset U$ of b (the existence of such a function follows from Theorem 2.13 and Corollary 2.17). Let $f \in \Gamma(U; M)$ and $P \in \mathrm{Diff}(B; M, N)$; $h.f$, extended by 0 outside of $\mathrm{supp}(h)$, is an element of $\Gamma(B; M)$. Hence, we can form $P.(h.f)$. This quantity is independent of h, and $f \mapsto P.(h.f)$ is called the *restriction* $P|_U \in \mathrm{Diff}(U; M|_U, N|_U)$ of P to U.

Let $\mathcal{E}$ be the sheaf of rings $U \mapsto \mathcal{E}(U)$. The mapping $U \mapsto \mathrm{Diff}(U; M|_U, N|_U)$ is clearly a sheaf of $\mathcal{E}$-Modules ([P2], section 5.3.1).

(III) Point distributions Let $P \in \mathrm{Diff}(B)$. For every $b \in B$, $f \mapsto (P.f)(b)$ is a distribution with support in $\{b\}$, written as $P(b)$. We say that it is a *point distribution* at b, and so $\mathrm{Diff}(B)$ is said to be a field of point distributions. The local expression (see **(I)**) of a point distribution at b of order p is

$$\Delta_b = \sum_{|\alpha| \leq p} a_\alpha.D_{\xi,b}^\alpha, \quad D_{\xi,b}^\alpha : f \mapsto \left(D_\xi^\alpha f\right)(b), \quad a_\alpha \in \mathbb{C}. \qquad [5.6]$$

The set of point distributions at b is an $\mathcal{E}(B)$-module, written as $\mathcal{T}_b^\infty(B)$, and $\mathcal{T}^\infty(B) = \bigoplus_{b \in B} \mathcal{T}_b^\infty(B)$ is the $\mathcal{E}(B)$-module of distributions with finite support in B. The above shows that $\mathcal{T}^\infty(B) = \Gamma(B, \mathrm{Diff}(B))$. We have the following result ([SCH 66], Chapter 3, section 10, Theorem 35):

THEOREM 5.6.– *Any distribution on $\mathbb{R}^n$ whose support is contained in $\{0\}$ is a finite linear combination of the Dirac distribution and its derivatives.*

The Dirac distribution at the point b ([P2], section 4.4.1(**II**)) is $\delta_b : \mathcal{E}(B) \to \mathbb{C} : f \mapsto f(b)$. We have $D^{\alpha}_{\xi,b}.f = \left\langle \delta_b, D^{\alpha}_{\xi} f \right\rangle = (-1)^{|\alpha|} \left\langle D^{\alpha}_{\xi}\delta_b, f \right\rangle$, hence:

$$\boxed{D^{\alpha}_{\xi,b}(\bullet) = (-1)^{|\alpha|} \left\langle D^{\alpha}_{\xi}\delta_b, \bullet \right\rangle.} \qquad [5.7]$$

REMARK 5.7.– *We can extend the notion of a finitely supported distribution to the case where B is a Banach $\mathbb{K}$-manifold of class C^r ([BOU 82a], section 13). It might seem tempting to define a compactly supported distribution more generally as a continuous linear form on $\mathcal{E}(B)$; but this would require us to define a "good" locally convex topology on the latter space, which is surprisingly difficult (see [KRE 76]).*

5.3. Manifolds of mappings

Manifolds of mappings lie at the heart of the perspective of global analysis adopted in [EEL 66, PAL 68, KRI 97]. In this section, $\mathbb{K} = \mathbb{R}$.

5.3.1. *The Banach framework*

(I) BANACH SPACE OF MAPPINGS OF CLASS C^k ON A COMPACT MANIFOLD Let B be a compact manifold of class C^∞, $\mathbf{Y}$ a Banach space and k an integer ≥ 0. The space $\mathcal{C}^k(B; \mathbf{Y})$ of mappings of class C^k from B into $\mathbf{Y}$, equipped with the norm

$$\|\mathbf{f}\|_k := \sup_{\substack{b \in B \\ 0 \leq j \leq k}} \left\| D^j \mathbf{f}(b) \right\|,$$

is a Banach space (**exercise**).

The space $\mathcal{C}^\infty(B; \mathbf{Y})$ can be equipped with the increasing filtrant and countable family of norms $\|.\|_k$ defined above. The coarsest topology for which *all* of these norms are continuous is a non-normable Fréchet space topology (**exercise**).

(II) BANACH MANIFOLD OF MAPPINGS OF CLASS C^k ON A COMPACT MANIFOLD If Y is a Banach manifold of class C^∞ and k is an integer ≥ 0, we can still equip $\mathcal{C}^k(B; Y)$ with the uniform structure of uniform convergence of mappings from B into Y and their differentials (Remark 2.40(4)) up to order k. This uniform space is Hausdorff whenever Y is Hausdorff. We will equip $\mathcal{C}^k(B; Y)$ with a Banach manifold structure of class C^∞ and assume that B and Y are C^∞-paracompact[5].

5 The following results are due to R. Palais: see ([ABR 63], Section 11); however, we will present them as formulated by ([BOU 82a], Section 15), choosing stronger hypotheses to enable certain simplifications.

Let k be an integer > 0, $\pi_X : X \twoheadrightarrow B$ a Banach fiber bundle of class C^∞ (section 3.4.1, Definition 3.22(i)) and (with the notation of Corollary-Definition 3.21, section 3.3.4) $\sigma \in \Gamma^{(k)}(B; X)$.

DEFINITION 5.8.– *A* tubular neighborhood *of class C^∞ of σ is a triple (M, N, φ), where $\pi_M : M \twoheadrightarrow B$ is a Banach vector bundle of class C^∞, N is an open neighborhood of $\sigma(B)$ in X and φ is a B-isomorphism of class C^∞ from N onto an open subset of M (section 3.3.1, Lemma-Definition 3.7(iv)) that transforms the section σ into the zero section of M (namely $B \to M : b \longmapsto 0$).*

By the C^∞-paracompactness of B, σ admits tubular neighborhoods of class C^∞ ([LAN 99b], Chapter 4, Theorem 5.1). If N is an open subset of X, write $\Gamma^{(k)}(B; N)$ for the set of sections $\sigma \in \Gamma^{(k)}(B; X)$ such that $\sigma(B) \subset N$. With the notation of Definition 5.8, the space $\Gamma^{(k)}(B; M)$, equipped with the vector space structure induced by the fibers and the topology of uniform convergence of the mappings from B into M and their differentials of order up to k, is a Banach space ([ABR 63], Theorem 5.4) as in **(I)**. We have the following result ([ABR 63], Theorem 11.1; [BOU 82a], 15.3.2):

LEMMA 5.9.– *There exists a unique manifold structure of class C^∞ on $\Gamma^{(k)}(B; X)$ such that, for every section $\sigma \in \Gamma^{(k)}(B; X)$ and every tubular neighborhood (M, N, φ) of class C^k of σ, the triple $\left(\Gamma^{(k)}(B; N), \varphi^*, \Gamma^{(k)}(B; M)\right)$ is a chart of $\Gamma^{(k)}(B; X)$ for which the following diagram commutes, where $\varphi^* : \Gamma^{(k)}(B; N) \to \Gamma^{(k)}(B; M)$ is the mapping $f \longmapsto \varphi \circ f$:*

$$\begin{array}{ccc} B & \xrightarrow{f} & N \\ & \searrow & \downarrow \varphi \\ \varphi^*(f) & & \\ & & M \end{array}$$

The Banach manifold structure of class C^k of $\mathcal{C}^k(B; Y)$ can be obtained by identifying this space with $\Gamma^{(k)}(B; B \times Y)$.

(III) TANGENT SPACE Let $\pi : T(Y) \twoheadrightarrow Y$ be the tangent bundle of Y. If $f^0 \in \mathcal{C}^k(B; Y)$, the tangent space $T_{f^0}\left(\mathcal{C}^k(B; Y)\right)$ is given by

$$T_{f^0}\left(\mathcal{C}^k(B; Y)\right) = \left\{u \in \mathcal{C}^k(B; T(Y)) : \pi \circ u(b) = f^0(b), \forall b \in B\right\}.$$

Let $b \in B$, $f^0 \in \mathcal{C}^k(B; Y)$, and write $\delta_b : \mathcal{C}^k(B; Y) \to Y : g \mapsto g(b)$ for the Dirac measure at the point b; δ_b is a morphism of class C^∞. Let $\xi \in T_{f^0}\left(\mathcal{C}^k(B; Y)\right)$. The image of ξ under $T_{f^0}(\delta_b)$ is an element $\tilde{\xi}_b$ of $T_{f^0(b)}(Y)$ and the mapping $\tilde{\xi} : b \mapsto \xi_b$ (from B into $T_{f^0(b)}(Y)$) is a lifting of class C^k of f^0 into $T(Y)$ (section 3.3.1, Definition 3.6), which gives $\pi \circ \tilde{\xi} = f^0$. Hence, the following diagram commutes (see Lemma-Definition 3.11, section 3.3.2):

$$\begin{array}{ccc} B & \xrightarrow{f^0} & Y \\ \pi' \uparrow & \searrow \tilde{\xi} & \uparrow \pi \\ f^{0*}(T(Y)) & \xrightarrow{f'} & T(Y) \end{array}$$

and, by [3.7], $\tilde{\xi}$ can be identified with a section of class C^k of the fiber bundle $\pi' : f^{0*}(T(Y)) \to B$. We therefore have the isomorphism:

$$T_{f^0}\left(\mathcal{C}^k(B;Y)\right) \cong \Gamma^{(k)}\left(B; f^{0*}(T(Y))\right).$$

(IV) EXPONENTIAL LAW Let X, Y, Z be three sets; write Z^Y for the set of mappings from Y into Z ([P1], section 1.1.2**(IV)**). Then, we have the canonical bijection

$$(Y \times Z)^X \cong Y^X \times Z^X. \qquad [5.8]$$

Similarly, the mapping $Z^{Y \times X} \ni f \mapsto (x \mapsto f(.,x)) \in \left(Z^Y\right)^X$ is a bijection (**exercise**), and the canonical bijection

$$Z^{Y \times X} \cong \left(Z^Y\right)^X \qquad [5.9]$$

is called the exponential law. Finally, there exists a canonical mapping

$$Y^X \times Z^Y \to Z^X$$

given by $(u, v) \mapsto v \circ u$.

The two isomorphisms [5.8] and [5.9] also exist for uniform spaces. Furthermore, if X is compact, there exists a canonical isomorphism of uniform spaces $\mathcal{C}(X; Y \times Z) \cong \mathcal{C}(X;Y) \times \mathcal{C}(X;Z)$ (where these three spaces of continuous mappings are equipped with the uniform structure of uniform convergence); if X and Y are compact, there exists a canonical *homeomorphism* $\mathcal{C}(X \times Y; Z) \cong \mathcal{C}(X; \mathcal{C}(Y;Z))$, and the mapping $(u, v) \mapsto v \circ u$ from $\mathcal{C}(X;Y) \times \mathcal{C}(Y;Z)$ into $\mathcal{C}(X;Z)$ is continuous ([BOU 74], Chapter 10, section 3.4, Corollary 2 and Proposition 9).

Similar results can be established for manifolds. If X, Y, Z are manifolds of class C^∞, then, for every integer $k \geq 0$, we have a canonical bijection $\mathcal{C}^k(X; Y \times Z) \cong \mathcal{C}^k(X;Y) \times \mathcal{C}^k(X;Z)$:

$$(x \mapsto (g(x), h(x))) \mapsto ((x \mapsto g(x)), (x \mapsto h(x))).$$

Moreover, we have the following results ([ABR 63], Theorem 11.4; [BOU 82a], 15.3.7, 15.3.5):

THEOREM 5.10.– *Let X, Y, Z be C^{∞}-paracompact manifolds, where X, Y are compact, and let p, q be integers ≥ 1.*

*1) Write $\mathcal{C}^{p,q}(X \times Y; Z)$ for the manifold of continuous mappings $f : X \times Y \to Z$ satisfying the following property **(P)**:*

***(P)** For any choice of charts $c = (U, \varphi, \mathbf{E})$ of X, $d = (V, \psi, \mathbf{F})$ of Y and $e = (W, \theta, \mathbf{G})$ of Z, the partial differentials $D_{\mathbf{E}}^{p'} D_{\mathbf{F}}^{q'} \Phi$ of the expression Φ of f in the charts $c \times d$ and e (section 2.3.7**(I)** and Definition 2.19) exist and are continuous for every integer $p', q' \geq 0$ such that $p' \leq p$ and $q' \leq q$.*

Then, there exists a canonical bijection, called the exponential law

$$\mathcal{C}^{p,q}(X \times Y; Z) \cong \mathcal{C}^{p}(X; \mathcal{C}^{q}(Y; Z)).$$

2) The composition $(u, v) \mapsto v \circ u$ induces a mapping of class C^q

$$\mathcal{C}^{p}(X; Y) \times \mathcal{C}^{p+q}(Y; Z) \to \mathcal{C}^{p}(X; Z).$$

(V) EVALUATION Let X, Y be C^{∞}-paracompact manifolds, suppose that X is compact, and let $p \in \mathbb{N}^{\times}$. Then, the evaluation operator $\mathrm{ev} : \mathcal{C}^{p}(X; Y) \times X \to Y : (\varphi, x) \mapsto \varphi(x)$ (section 1.2.3**(IV)**) is of class C^p ([ABR 63], Theorem 11.7) and the following expression generalizes [1.11]:

$$T_{(\varphi,x)}(\mathrm{ev}).(h, \xi) = h(x) + T_x(\varphi).\xi. \qquad [5.10]$$

The generalization of [1.10] to the setting considered here is left to the reader.

5.3.2. *The "convenient" framework*

The Banach framework does not allow us to define $\mathcal{C}^{\infty}(X; Y)$ or $\mathcal{C}^{\omega}(X; Y)$ (see section 5.3.1**(I)**). However, if X, Y are metrizable, finite-dimensional manifolds of class $\mathfrak{c}^{\infty}$ and X is compact (section 2.2.1**(II)**), then the set $\mathfrak{c}^{\infty}(X; Y)$ can be canonically equipped with the structure of a metrizable $(\mathcal{FN})$ manifold of class $\mathfrak{c}^{\infty}$ ([KRI 97], Lemmas 30.3(2) and 42.5 and Theorem 42.1).

Let X, Y, Z be metrizable, finite-dimensional manifolds of class $\mathfrak{c}^{r}$ $(r \in \{\infty, \omega\})$ and suppose that Y is compact. Then, there exists a canonical bijection, called the *exponential law* ([KRI 97], Theorem 42.14):

$$\mathfrak{c}^{r}(X \times Y; Z) \cong \mathfrak{c}^{r}(X; \mathfrak{c}^{r}(Y; Z)).$$

The evaluation operator ev : $\mathfrak{c}^{\omega}(X;Y) \times X \to Y : (\varphi, x) \mapsto \varphi(x)$, where X is a compact manifold of class $\mathfrak{c}^{\omega}$ and Y is a finite-dimensional manifold of class C^{ω}, a Lie group of class $\mathfrak{c}^{\omega}$, or a $(\mathcal{KM})$ space (Section 1.3.3(**II**)), is of class $\mathfrak{c}^{\omega}$ ([KRI 97], Chapter 9, Corollary 42.15)[6] and the expression [5.10] remains valid.

Suppose that X, Y are as above and let M be a compact manifold of class $\mathfrak{c}^{\omega}$. The composition $(u, v) \mapsto v \circ u$ induces a mapping of class $\mathfrak{c}^{\omega} : \mathfrak{c}^{\omega}(M;X) \times \mathfrak{c}^{\omega}(X;Y) \to \mathfrak{c}^{\omega}(M;Y)$ (*ibid.*).

5.4. Lie derivatives

5.4.1. *Lie algebras*

(I) NOTION OF A LIE ALGEBRA Let $\mathbf{K}$ be a field. In the following, unless otherwise stated, every algebra is a $\mathbf{K}$-algebra[7] and every vector space is defined over $\mathbf{K}$. A *Lie algebra* $\mathfrak{g}$ is a $\mathbf{K}$-vector space equipped with a $\mathbf{K}$-bilinear mapping $(X, Y) \mapsto [X, Y]$ from $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$, called the *Lie bracket*, satisfying the identities

(Lie 1) $[X, X] = 0$

(Lie 2) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

for every $X, Y, Z \in \mathfrak{g}$. The relation (**Lie 2**) is called the *Jacobi identity*.

In general, the Lie algebra $\mathfrak{g}$ can be non-associative ([P1], section 2.3.10(**I**)). We have $[X + Y, X + Y] = 0$, so $[X, Y] = -[Y, X]$, which means that $\mathfrak{g}$ is anti-commutative.

Let $\mathfrak{g}$ be a Lie algebra and $(E_i)_{i \in I}$ a basis of $\mathfrak{g}$. The structure constants[8] of $\mathfrak{g}$ are the elements $c_{ij}^{k} \in \mathbf{K}$ $(i, j, k \in I)$, all but finitely many of which are zero, uniquely determined by the condition

$$[E_i, E_j] = \sum_{k \in I} c_{ij}^{k} E_k, \quad \forall i, j \in I. \qquad [5.11]$$

6 More generally, Y can be an analytic manifold with a "local addition" operation in the sense of ([KRI 97], Chapter 9, section 42.4).

7 The notions presented below are still valid when $\mathbf{K}$ is replaced by a commutative ring if we assume that the $\mathbf{K}$-algebras are free $\mathbf{K}$-modules.

8 The structure constants of an *arbitrary* $\mathbf{K}$-algebra $\mathbf{A}$ and its multiplication table are defined as in [P1], section 2.3.10(**III**); nevertheless, we can only substitute the elements of $\mathbf{A}$ into the indeterminates of a polynomial algebra with the additional hypotheses stated in *loc. cit.*

Conversely, let $\mathbf{A}$ be an algebra and $c_{ij}^{k} \in \mathbf{K}$ $(i, j, k \in I)$ its structure constants. Then, $\mathbf{A}$ is a Lie algebra if and only if, for $i, j, k, l \in I$ (**exercise**),

$$c_{ij}^{k} = -c_{ji}^{k}, \quad \sum_{r \in I} \left(c_{ij}^{r} c_{rk}^{l} + c_{jk}^{r} c_{ri}^{l} + c_{ki}^{r} c_{rj}^{l} \right) = 0.$$

(II) CATEGORY OF LIE ALGEBRAS A mapping $f : \mathfrak{g} \to \mathfrak{g}'$ is a morphism of Lie algebras if f is a morphism of algebras such that $f([X, Y]) = [(f(X), f(Y))]$ for every $X, Y \in \mathfrak{g}$. The Lie algebras form a category **LieAl**; this category is not additive, since the morphisms from $\mathfrak{g}$ into $\mathfrak{g}'$ do not form an abelian group. If $\mathfrak{a}, \mathfrak{b}$ are vector subspaces of the Lie algebra $\mathfrak{g}$, write $[\mathfrak{a}, \mathfrak{b}]$ for the vector space generated by all the elements $[A, B], A \in \mathfrak{a}, B \in \mathfrak{b}$. If $\mathfrak{g}$ is a Lie algebra, a *subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$ is a Lie algebra $\mathfrak{h} \subset \mathfrak{g}$.

A Lie algebra $\mathfrak{g}$ is said to be commutative if $[X, Y] = [Y, X]$ for all $X, Y \in \mathfrak{g}$. This condition is equivalent to having $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$, i.e. $[\mathfrak{g}, \mathfrak{g}] = 0$.

LEMMA 5.11.– *If* $\mathfrak{g}$ *is an associative algebra, the bilinear mapping* $(X, Y) \mapsto X.Y - Y.X$ *equips* $\mathfrak{g}$ *with a Lie algebra structure (**exercise**). We will therefore set* $[X, Y] = X.Y - Y.X$.

In particular, if $\mathbf{E}$ is a Banach space, then the Banach algebra $\mathcal{L}(\mathbf{E})$ ([P2], section 3.4.1**(II)**, Corollary 3.44) equipped with the above bilinear operation is a Lie algebra, written as $\mathfrak{gl}(\mathbf{E})$. Readers should be careful not to confuse multiplication $(f, g) \mapsto f \circ g$ in End $(\mathbf{E})$ with multiplication $(f, g) \mapsto [f, g]$ in $\mathfrak{gl}(\mathbf{E})$. A morphism of Lie algebras $\mathfrak{g} \to \mathfrak{gl}(\mathbf{E})$ is said to be a *representation* of $\mathfrak{g}$ in $\mathbf{E}$ (compare with the representation of a group in [P2], section 1.3.2). Any such representation is said to be *faithful* if it is injective.

Recall that the notions of left ideal, right ideal and two-sided ideal can be defined for any algebra ([P1], section 2.3.10**(II)**); in the special case of a Lie algebra, they all coincide, and a vector subspace $\mathfrak{a}$ of $\mathfrak{g}$ is an ideal if and only if $[\mathfrak{a}, \mathfrak{g}] = \mathfrak{a}$ (**exercise**). An ideal $\mathfrak{a}$ of $\mathfrak{g}$ is *proper* if $\mathfrak{a} \subsetneq \mathfrak{g}$ and *trivial* if $\mathfrak{a} = \{0\}$.

One example of a morphism of Lie algebras is the canonical decomposition, stated in [P1], section 2.3.10**(II)**, for arbitrary algebras, which gives the isomorphism of Lie algebras (Noether's first isomorphism) $f(\mathfrak{g}) \cong \mathfrak{g}/\ker(f)$, where $\ker(f) := \{X \in \mathfrak{g} : f(X) = 0\}$.

If $f : \mathfrak{g} \to \mathfrak{g}'$ is a morphism of Lie algebras, then $\ker(f)$ is an ideal of $\mathfrak{g}$; if $\mathfrak{a}$ is an ideal of the algebra $\mathfrak{g}$, then the quotient $\mathbf{K}$-vector space $\mathfrak{g}/\mathfrak{a}$ can be equipped with a canonical Lie algebra structure, since, if $X, Y \in \mathfrak{g}$, the Lie bracket $[X, Y]$ only depends on the classes $\dot{X}, \dot{Y}$ of X, Y $(\text{mod.}\mathfrak{a})$ and may therefore be written as $\left[\dot{X}, \dot{Y}\right]$ (**exercise**).

If $\mathfrak{a}, \mathfrak{b}$ are two ideals of a Lie algebra $\mathfrak{g}$, then $\mathfrak{a} + \mathfrak{b}$ and $\mathfrak{a} \cap \mathfrak{b}$ are ideals of $\mathfrak{g}$. Furthermore, if $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, then $\mathfrak{h} + \mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$ and we have Noether's second isomorphism (**exercise**) $(\mathfrak{h} + \mathfrak{a}) / \mathfrak{a} \cong \mathfrak{h} / (\mathfrak{h} \cap \mathfrak{a})$. The induced homomorphism theorem, Noether's third isomorphism and the correspondence theorem also hold ([P1], section 2.3.3); their statements and proofs are left to the reader.

If $\mathfrak{g}_1, \mathfrak{g}_2$ are two Lie algebras over $\mathbf{K}$, then the product $\mathbf{K}$-vector space $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$, equipped with the $\mathbf{K}$-bilinear mapping

$$((X_1, X_2), (Y_1, Y_2)) \mapsto ([X_1, Y_1], [X_2, Y_2]),$$

has the structure of a Lie $\mathbf{K}$-algebra. Writing $[.,.]$ for the bilinear mapping from $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$ thus obtained, we have $[(X_1, 0), (0, Y_2)] = 0$ for every $X_1 \in \mathfrak{g}_1$ and all $X_2 \in \mathfrak{g}_2$. The statement and proof of the product isomorphism is left to the reader ([P1], section 2.2.3(**IV**)).

If $\mathfrak{a}, \mathfrak{b}$ are Lie algebras, we say that a Lie algebra $\mathfrak{g}$ is an *extension* of $\mathfrak{b}$ by $\mathfrak{a}$ (similar to groups: see [P1], section 2.2.2(**I**)) if there exists a sequence

$$\mathfrak{a} \overset{\lambda}{\hookrightarrow} \mathfrak{g} \overset{\mu}{\twoheadrightarrow} \mathfrak{b}, \qquad [5.12]$$

where λ, μ are a monomorphism and an epimorphism of Lie algebras, respectively, with $\mathfrak{n} := \ker(\mu)$, $\lambda(\mathfrak{a}) = \mathfrak{n}$, so that $\mathfrak{b} \cong \mathfrak{g}/\mathfrak{n}$. In the category ${}_{\mathbf{K}}\mathbf{Vec}$ of $\mathbf{K}$-vector spaces, we therefore have the short exact sequence

$$\{0\} \longrightarrow \mathfrak{a} \overset{\lambda}{\longrightarrow} \mathfrak{g} \overset{\mu}{\longrightarrow} \mathfrak{b} \longrightarrow \{0\}. \qquad [5.13]$$

In this case, $\mathfrak{a} \cong \ker(\mu)$ is an ideal of $\mathfrak{g}$, and the short sequence [5.13] is said to be *exact* in $\mathbf{LieAl}$. The extension [5.12] is said to be *central* if $\mathfrak{n} \subset \mathfrak{z}(\mathfrak{g})$ where $\mathfrak{z}(\mathfrak{g})$ is the center of $\mathfrak{g}$ (see [P1], section 2.3.10(**III**) and (**III**) below) and *inessential* (respectively *trivial*) if there exists a subalgebra (respectively an ideal) of $\mathfrak{g}$ that is supplementary to $\mathfrak{n}$ in $\mathfrak{g}$ in ${}_{\mathbf{K}}\mathbf{Vec}$. Any central and inessential extension is trivial (**exercise**).

(III) Derivations; adjoint linear mapping Let $\mathbf{A}$ be a (not necessarily associative) algebra and $\mathrm{Der}(\mathbf{A})$ the algebra of derivations of $\mathbf{A}$ (i.e. derivations of $\mathbf{A}$ in the $\mathbf{A}$-module ${}_{\mathbf{A}}\mathbf{A}$, where $\mathbf{A}$ is equipped with the trivial graduation: see [P1], section 2.3.12). If $\boldsymbol{D}_1, \boldsymbol{D}_2 \in \mathrm{Der}(\mathbf{A})$, then $[\boldsymbol{D}_1, \boldsymbol{D}_2] := \boldsymbol{D}_1 \circ \boldsymbol{D}_2 - \boldsymbol{D}_2 \circ \boldsymbol{D}_1 \in \mathrm{Der}(\mathbf{A})$ and this bracket equips $\mathrm{Der}(\mathbf{A})$ with the structure of a Lie algebra (**exercise**).

Let $\mathfrak{g}$ be a Lie algebra. For every $X \in \mathfrak{g}$, the linear mapping $Y \longmapsto [X, Y]$ from $\mathfrak{g}$ to $\mathfrak{g}$ is a derivation of $\mathfrak{g}$, written as $\mathrm{ad}_{\mathfrak{g}} X$ or $\mathrm{ad} X$ when $\mathfrak{g}$ is implicitly clear. The mapping $X \mapsto \mathrm{ad} X$ is a morphism of Lie algebras from $\mathfrak{g}$ into $\mathrm{Der}(\mathfrak{g})$; in particular, $\mathrm{ad}[X, Y] = [\mathrm{ad} X, \mathrm{ad} Y]$ and, for every derivation $\boldsymbol{D} \in \mathrm{Der}(\mathfrak{g})$,

$[\boldsymbol{D}, \mathrm{ad}X] = \mathrm{ad}(\boldsymbol{D}X)$; hence, $\mathrm{ad}\mathfrak{g}$ is an ideal of $\mathrm{Der}(\mathfrak{g})$ (**exercise**). The mapping $\mathrm{ad}X$ is called the *inner derivation* defined by X; any derivations not of this form are said to be *outer* (compare with [P1], section 3.1.11(**I**)).

A classical procedure often effective in mathematics is to *linearize* the problem. With Lie algebras, this idea leads to the adjoint representation: the mapping $X \mapsto \mathrm{ad}X$ from $\mathfrak{g}$ into $\mathfrak{gl}(\mathfrak{g}) = \mathrm{End}(\mathfrak{g})$ is a representation of $\mathfrak{g}$ in the vector space $\mathfrak{g}$ (see (**II**)).

DEFINITION 5.12.– *ad* : $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ *is called the* adjoint linear representation*, and* adX *is called the* adjoint linear mapping *of* X.

Any ideal $\mathfrak{c}$ of $\mathfrak{g}$ is a submodule that is stable under inner derivations (**exercise**). This ideal is said to be *characteristic* if it is stable under *every* derivation $\boldsymbol{D} \in \mathrm{Der}(\mathfrak{g})$. The center of $\mathfrak{g}$, is the set of elements $Z \in \mathfrak{g}$ such that $[Z, X] = 0$ for every $X \in \mathfrak{g}$; this center $\mathfrak{z}(\mathfrak{g}) = \ker(\mathrm{ad})$ is a characteristic ideal (**exercise**). If $\mathfrak{a}, \mathfrak{b}$ are ideals of $\mathfrak{g}$, then $[\mathfrak{a}, \mathfrak{b}]$ is an ideal that is characteristic whenever $\mathfrak{a}$ and $\mathfrak{b}$ are characteristic (**exercise**).

Let $\mathfrak{g}$ be a p-dimensional Lie algebra and $(E_i)_{1 \leq i \leq p}$ a basis of $\mathfrak{g}$. The relation [5.11] shows that the matrix A_i representing $\mathrm{ad}E_i$ in this basis has entries $(A_i)_j^k = c_{ij}^k$ ([P1], section 3.1.3(**II**)); see Example 6.45 in section 6.3.6(**I**).

(IV) DIRECT SUMS AND SEMI-DIRECT SUMS OF LIE ALGEBRAS Let $\mathfrak{a}_1, \mathfrak{a}_2$ and be Lie subalgebras of the Lie algebra $\mathfrak{a}$; we say that $\mathfrak{a}$ is the *direct sum* of $\mathfrak{a}_1, \mathfrak{a}_2$, written as $\mathfrak{a} = \mathfrak{a}_1 \oplus \mathfrak{a}_2$, if $\mathfrak{a} \cong \mathfrak{a}_1 \times \mathfrak{a}_2$. This means that (i) the $\mathbf{K}$-vector spaces $\mathfrak{a}_1, \mathfrak{a}_2$ are supplementary in $\mathfrak{a}$ and (ii) $[\mathfrak{a}_1, \mathfrak{a}_2] = 0$. In other words, $\mathfrak{a} = \mathfrak{a}_1 \oplus \mathfrak{a}_2$ if and only if $\mathfrak{a}$ is a trivial extension of $\mathfrak{a}_2$ by $\mathfrak{a}_1$ (**exercise**). If so, we say that $\mathfrak{a}_1$ (and similarly $\mathfrak{a}_2$) is a *direct factor* of $\mathfrak{a}$.

Let $\mathfrak{h}, \mathfrak{k}$ be Lie subalgebras of the Lie algebra $\mathfrak{g}$ such that $\mathfrak{h}, \mathfrak{k}$ are supplementary $\mathbf{K}$-vector spaces. We say that $\mathfrak{g}$ is the *semi-direct sum* of $\mathfrak{h}$ and $\mathfrak{k}$ if $[\mathfrak{h}, \mathfrak{k}] \subset \mathfrak{k}$, i.e. if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. If so, $\mathfrak{g}/\mathfrak{h} \cong \mathfrak{k}$, and $\mathfrak{g}$ is an inessential extension of $\mathfrak{k}$ by $\mathfrak{h}$; if $K \in \mathfrak{k}$, the mapping $\sigma_K : H \mapsto [K, H]$ is a derivation of $\mathfrak{g}$ and $\sigma : K \mapsto \sigma_K$ is a representation of K in H. We write that $\mathfrak{g} = \mathfrak{h} \oplus_\sigma \mathfrak{k}$.

Lie algebras are discussed further in section 6.3.

5.4.2. *Lie derivative of a function*

Up to and including section 5.5, the field $\mathbf{K}$ from section 5.4.1 is taken to be equal to $\mathbb{K}$, and we will adopt the conventions **(C1)** (section 2.2.1(**II**), p. 55) and **(C2)** (section 2.2.2(**III**), p. 58).

Let B be a manifold. Recall that $\mathcal{C}^r(B), \mathcal{T}_0^1(B)$ and $\Omega^1(B)$ denote the space of functions of class C^r from B into $\mathbb{K}$, the space of vector fields of class C^r on B and the space of covector fields of class C^r on B, respectively (sections 4.3.1 and 4.3.2). The Lie derivative of a function $f \in \mathcal{C}^r(B)$ at a point along a tangent vector was introduced earlier (Lemma-Definition 2.29).

DEFINITION 5.13.– *Let $f \in \mathcal{C}^r(B)$ and $X \in \mathcal{T}_0^1(B)$. The mapping $\mathcal{L}_X.f := \langle df, X\rangle = b \mapsto \langle d_b f, X_b\rangle$ is called the* Lie derivative *of f along X.*

DEFINITION 5.14.– *Let B, B' be two manifolds, $\varphi : B \to B'$ a diffeomorphism and $Y \in \mathcal{T}_0^1(B')$. The preimage of Y under φ is the vector field $\varphi^*(Y) : f \mapsto \mathcal{L}_Y.(f \circ \varphi)$ $(f \in \mathcal{C}^r(B))$ (compare with the preimage of a distribution in section 5.2.1(**VIII**)).*

The Lie derivation $\mathcal{L}_X$ is a derivation of the algebra $\mathcal{C}^r(B)$ in the $\mathcal{C}^r(B)$-module $\mathcal{C}^{r-1}(B)$ ([P1], section 2.3.12); $\mathcal{C}^{r-1}(B) = \mathcal{C}^r(B)$, so $\mathcal{L}_X \in \mathrm{Der}(\mathcal{C}^r(B))$. The mapping $X \mapsto \mathcal{L}_X$ from $\mathcal{T}_0^1(B)$ into $\mathrm{Der}(\mathcal{C}^r(B))$ is injective by Lemma-Definition 2.29(i). Accordingly, the mapping $\mathcal{L}_X.f$ specified above is sometimes written as $X.f$, so that, for every $b \in B$,

$$X.f(b) = \langle d_b f, X_b\rangle .$$

Suppose that B is locally finite-dimensional. Let $(U, \xi, \mathbb{K}^m)$ be a chart of B such that $b \in U$, write $(\mathbf{e}_j)_{1\le j\le m}$ for the canonical basis of $\mathbb{K}^m$, and suppose that $X_j(b) = (d_b\xi)^{-1}.\mathbf{e}_j$. By section 2.2.4(**III**), $X_j(b) = \left(\frac{\partial}{\partial \xi^j}\right)_b$ and

$$X = \sum\nolimits_{1\le j\le m} \left(\mathcal{L}_X.\xi^j\right).\left(\frac{\partial}{\partial \xi^j}\right).$$

THEOREM 5.15.– *1) Let B be a manifold. The mapping $X \mapsto \mathcal{L}_X$ from $\mathcal{T}_0^1(B)$ into $\mathrm{Der}(\mathcal{C}^r(B))$ is a monomorphism of $\mathcal{C}^r(B)$-algebras.*

2) If B is locally finite-dimensional, this mapping is an isomorphism. If $(U, \xi, \mathbb{K}^m)$ is a chart of B, we have the following local expression over U:

$$\mathcal{L}_X = \sum\nolimits_{1\le j\le m} X^j.\left(\frac{\partial}{\partial \xi^j}\right), \quad X^j = \mathcal{L}_X.\xi^j \in \mathcal{C}^r(B).$$

PROOF.– (1) follows from the above. (2): Let $\boldsymbol{D} \in \mathrm{Der}(\mathcal{C}^r(B))$, $f \in \mathcal{C}^r(B)$, $b \in B$, and suppose that the chart (U, ξ, m) is centered on b; this is without loss of generality because $\boldsymbol{D}$ cancels constants, like any other derivation (**exercise**). By Taylor's formula with the Young's residual (Theorem 1.22(ii)), the following formula holds in the neighborhood of $\xi = 0$:

$$(f \circ \xi)(b') = \sum\nolimits_{1\le j\le m} \xi^j(b').\frac{\partial f}{\partial \xi^j}(0) + \varepsilon(\xi(b')),$$

where $\varepsilon \in \mathcal{C}^r(\xi(U))$ vanishes at $\xi = 0$. Hence,

$$(\boldsymbol{D}.f)(b) = \sum\nolimits_{1 \leq j \leq m} \left(\boldsymbol{D}.\xi^j\right)(0) \cdot \frac{\partial f}{\partial \xi^j}(0),$$

which gives $\boldsymbol{D} = \sum_{1 \leq j \leq m} \left(\mathbf{D}.\xi^j\right) \cdot \frac{\partial}{\partial \xi^j}$. ■

5.4.3. *Lie brackets*

THEOREM-DEFINITION 5.16.– *Let $X, Y \in \mathcal{T}_0^1(B)$ be vector fields of class C^r.*

1) There exists a unique vector field $[X, Y] \in \mathcal{T}_0^1(B)$ of class C^r, called the Lie bracket *of X, Y, such that*

$$\boxed{\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X.}$$

2) The $\mathcal{C}^r(B)$-module $\mathcal{T}_0^1(B)$ equipped with the Lie bracket is a Lie $\mathbb{K}$-algebra.

PROOF.– (1) If the derivation $\mathcal{L}_{[X,Y]}$ exists, it uniquely determines $[X, Y] \in \mathcal{T}_0^1(B)$ by Theorem 5.15(1). Moreover, if $\boldsymbol{D}_1, \boldsymbol{D}_2$ are derivations of an algebra $\mathbf{A}$, then $\boldsymbol{D} := \boldsymbol{D}_1 \circ \boldsymbol{D}_2 - \boldsymbol{D}_2 \circ \boldsymbol{D}_1$ is also a derivation of this algebra. We still need to show the existence of a vector field Z such that $Z = [X, Y]$. If B is locally finite-dimensional, this follows from Theorem 5.15(2). Let us go into slightly more detail. Let $c = (U, \xi, m)$ be a chart of B and $(\mathbf{s}_i)_{1 \leq i \leq m}$ a frame of class C^r over U (Lemma-Definition 3.23). Then, $(\mathbf{s}_i)_{1 \leq i \leq m}$ is a basis of the free $\mathcal{C}^r(U)$-module $\mathcal{T}_0^1(U)$ (Corollary 3.33(i)). Let x^j, y^j $(1 \leq j \leq m)$ be the components of the restrictions to U of X, Y, respectively, in this basis. Then, $\mathcal{L}_X : f \mapsto \langle df, X \rangle = \sum_i \frac{\partial f}{\partial \xi^i} x^i$, so

$$(\mathcal{L}_Y \circ \mathcal{L}_X)(f) = \sum_{i,j} \left(\frac{\partial f}{\partial \xi^i} \cdot \frac{\partial x^i}{\partial \xi^j} y^j + \frac{\partial^2 f}{\partial \xi^i \partial \xi^i} . x^i . y^j \right) \qquad [5.14]$$

and hence $(\mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X)(f) = \sum_i \frac{\partial f}{\partial \xi^i} . z^i$, with

$$z^i = \sum_j \left(\frac{\partial y^i}{\partial \xi^j} x^j - \frac{\partial x^i}{\partial \xi^j} y^j \right). \qquad [5.15]$$

In the general case, we can use charts to reduce to the case where B is an open subset of a Banach space $\mathbf{E}$ and X, Y are mappings $b \mapsto (b, \mathbf{x}(b))$ and $b \mapsto (b, \mathbf{y}(b))$, respectively. We then have:

$$[X, Y] : b \mapsto (b, \boldsymbol{D}\mathbf{y}(b) . \mathbf{x}(b) - \boldsymbol{D}\mathbf{x}(b) . \mathbf{y}(b)). \qquad [5.16]$$

2) The mapping $(X, Y) \mapsto [X, Y]$ is antisymmetric and we can easily check that it satisfies the Jacobi identity (section 5.4.1(**I**)). ■

REMARK 5.17.– *The differential operator* $\mathcal{L}_Y \circ \mathcal{L}_X$ *is* second order *by [5.14], whereas* $[X, Y]$ *is* first order.

5.4.4. *Lie derivative of vector, covector and tensor fields*

(I) LIE DERIVATIVE OF A VECTOR FIELD The mapping

$$\boxed{\mathcal{L}_X : \mathcal{T}_0^1(B) \to \mathcal{T}_0^1(B) : Y \mapsto [X, Y]} \qquad [5.17]$$

is a derivation of the $\mathcal{C}^r(B)$-algebra $\mathcal{T}_0^1(B)$ (equipped with the relevant bracket). More specifically, for every $X, Y \in \mathcal{T}_0^1(B)$ and $f \in \mathcal{C}^r(B)$ (**exercise**),

$$\mathcal{L}_X(f.Y) = [\mathcal{L}_X.f, Y] + f(\mathcal{L}_X.Y), \quad \mathcal{L}_X.[Y, Z] = [\mathcal{L}_X.Y, Z] + [Y, \mathcal{L}_X.Z].$$

The vector field $\mathcal{L}_X.Y = [X, Y]$ is the derivative of Y along the direction X. Let $X_1, X_2 \in \mathcal{T}_0^1(B)$ and write span $\{X_1, X_2\}$ for the $\mathcal{C}^r(B)$-module generated by X_1, X_2.

PROPOSITION 5.18.– *1) If* $[X_1, X_2] = 0$ *and* $X, Y \in$ *span* $\{X_1, X_2\}$*, then* $[X, Y] \in$ *span* $\{X_1, X_2\}$.

2) (functor property of the bracket) *Let* B, B' *be two manifolds,* $X, Y \in \mathcal{T}_0^1(B)$*, and* $\varphi : B \to B'$ *a diffeomorphism. Then (with the notation of Lemma 2.34, section 2.3.1),* $[\varphi_* X, \varphi_* Y] = \varphi_* [X, Y]$.

PROOF.– (1): It suffices to show that, if $f_1, f_2 \in \mathcal{C}^r(B)$, then $[f_1 X_1, f_2 X_2] \in$ span $\{X_1, X_2\}$. But

$$[f_1 X_1, f_2 X_2] = f_1.f_2.\underbrace{[X_1, X_2]}_{0} + (f_1.\mathcal{L}_{X_1}.f_2).X_2 - (f_2.\mathcal{L}_{X_2}.f_1).X_1.$$

(2): **exercise**. ■

(II) LIE DERIVATIVE OF A COVECTOR FIELD Let $\alpha \in \mathcal{T}_1^0(B) = \Omega^1(B)$ be a covector field of class C^r on B. For every vector field $X_1 \in \mathcal{T}_0^1(B)$, $\langle \alpha, X_1 \rangle$ belongs to $\mathcal{C}^r(B)$. If $X \in \mathcal{T}_0^1(B)$, we obtain $\mathcal{L}_X.\alpha \in \Omega^1(B)$ by writing $\mathcal{L}_X.\langle \alpha, X_1 \rangle = \langle \mathcal{L}_X.\alpha, X_1 \rangle + \langle \alpha, \mathcal{L}_X.X_1 \rangle$, which gives

$$\langle \mathcal{L}_X.\alpha, X_1 \rangle = \mathcal{L}_X.\langle \alpha, X_1 \rangle - \langle \alpha, [X, X_1] \rangle. \qquad [5.18]$$

(II) LIE DERIVATIVE OF A TENSOR FIELD Let $X \in \mathcal{T}_0^1(B)$. Earlier, we defined $\mathcal{L}_X.f$, $\mathcal{L}_X.X_1$ and $\mathcal{L}_X.\alpha$ for every function $f \in \mathcal{C}^r(B)$, every vector field $Y \in \mathcal{T}_0^1(B)$ and every covector field $\alpha \in \mathcal{T}_1^0(B)$. For all integers $p, q \geq 0$, we now inductively define an operator $\mathcal{L}_X : \mathcal{T}_q^p(B) \to \mathcal{T}_q^p(B)$ satisfying the following conditions: given any two arbitrary tensor fields $\mathbf{Z}' \in \mathcal{T}_{q'}^{p'}(B)$, $\mathbf{Z}'' \in \mathcal{T}_{q''}^{p''}(B)$,

$$\mathcal{L}_X.(\mathbf{Z}' \otimes \mathbf{Z}'') = (\mathcal{L}_X.\mathbf{Z}' \otimes \mathbf{Z}'') + (\mathbf{Z}' \otimes \mathcal{L}_X.\mathbf{Z}'').$$

If $\sigma \in \mathfrak{S}_p$ and $\mathbf{Z} \in \mathcal{T}_0^p(B)$ (respectively $\mathbf{Z} \in \mathcal{T}_p^0(B)$), then $\mathcal{L}_X.(\sigma.\mathbf{Z}) = \sigma.(\mathcal{L}_X.\mathbf{Z})$, and in particular, $\mathcal{L}_X$ commutes with the symmetrization and antisymmetrization operations (section 4.2.2). Hence, if $\mathbf{Z}'$ and $\mathbf{Z}''$ are a p-vector field and a q-vector field or a p-form and a q-form, respectively, of class C^r in either case, then

$$\mathcal{L}_X.(\mathbf{Z}' \wedge \mathbf{Z}'') = (\mathcal{L}_X.\mathbf{Z}' \wedge \mathbf{Z}'') + (\mathbf{Z}' \wedge \mathcal{L}_X.\mathbf{Z}''). \quad [5.19]$$

5.4.5. *Lie derivative of a p-form*

Let $\alpha \in \Omega^p(B)$. If $X \in \mathcal{T}_0^1(B)$, it follows from the above that there exists a unique differential p-form $\mathcal{L}_X.\alpha \in \Omega^p(B)$ such that, for any vector fields $X_1, ..., X_p \in \mathcal{T}_0^1(B)$,

$$\boxed{(\mathcal{L}_X.\alpha)(X_1, ..., X_p) = \mathcal{L}_X.\alpha(X_1, ..., X_p) - \sum_{i=1}^p \alpha(X_1, ..., [X, X_i], ..., X_p).} \quad [5.20]$$

This expression generalizes [5.18]. If $\mathbf{Z}'$ and $\mathbf{Z}''$ are a p-form and a q-form taking values in the Banach spaces $\mathbf{F}'$ and $\mathbf{F}''$, respectively, the expression [5.19] still holds, taking the exterior product relative to some continuous bilinear mapping Φ from $\mathbf{F}' \times \mathbf{F}''$ into a Banach space $\mathbf{F}$ (section 4.2.6(**II**)).

Moreover, for any vector fields $X, Y \in \mathcal{T}_0^1(B)$ and every 1-form $\omega \in \Omega^1(B)$, we have the following relation (**exercise**):

$$\mathcal{L}_X.(Y \lrcorner \omega) = [X, Y] \lrcorner \omega + Y \lrcorner (\mathcal{L}_X.\omega).$$

EXAMPLE 5.19.– *Let B be a locally finite-dimensional manifold, $\alpha \in \Omega^1(B)$, $X, Y \in \mathcal{T}_0^1(B)$ and $c = (U, \xi, m)$ a chart of B. Over U, setting $\partial_i = \partial/\partial\xi^i$: $\alpha = \sum_i \alpha_i.d\xi^i$, we can write $X = \sum_i x^i \partial_i$, $Y = \sum_i y^i \partial_i$. It follows that:*

$$(\mathcal{L}_X.\alpha)Y = \mathcal{L}_X.\langle\alpha, Y\rangle - \langle\alpha, [X, Y]\rangle = \sum_{i,j} \left((\partial_i \alpha_j) y^j x^i + \alpha_i \left(\partial_j x^i\right) y^j\right).$$

Hence,

$$\mathcal{L}_X.\alpha = \sum_{i,j} \left(\left(\partial_i \alpha_j \right) x^i + \alpha_i \left(\partial_j x^i \right) \right) d\xi^j. \qquad [5.21]$$

In particular,

$$\mathcal{L}_X.d\xi^i = \sum_j \left(\partial_j x^i \right) .d\xi^j, \quad \mathcal{L}_X.\frac{\partial}{\partial \xi^j} = -\sum_i \left(\partial_j x^i \right) .\frac{\partial}{\partial \xi^i},$$

where the second expression is a special case of [5.15].

5.5. Exterior differential

5.5.1. *É. Cartan's theorem*

(I) EXTERIOR DIFFERENTIAL OF A p-FORM Let B be a manifold, $\mathbf{F}$ a Banach space and $\mathbf{f} \in \mathcal{C}^r(B;\mathbf{F})$. The mapping $d : f \mapsto d\mathbf{f}$ is a differential operator of order one (section 5.2.2**(I)**) from the trivial bundle $B \times \mathbf{F}$ into the bundle $\mathcal{L}(T(B);\mathbf{F})$. We will now generalize this setting by defining a differential operator of order one from $\mathrm{Alt}^p(T(B);\mathbf{F})$ into $\mathrm{Alt}^{p+1}(T(B);\mathbf{F})$ for every $p \geq 0$ (section 4.4.3**(I)**). Recall that $\Omega^p(B;\mathbf{F})$ denotes the space of differential p-forms of class C^r taking values in $\mathbf{F}$ (Section 4.4.3**(VI)**).

LEMMA 5.20.– (É. Cartan) *For every integer $p \geq 1$, there exists a unique differential operator d from $\Omega^p(B;\mathbf{F})$ into $\Omega^{p+1}(B;\mathbf{F})$ satisfying the following properties:*

i) If $\mathbf{F}, \mathbf{F}', \mathbf{F}''$ are Banach spaces, $\Phi : \mathbf{F}' \times_B \mathbf{F}'' \to \mathbf{F}$ is a coupling (Definition 4.31), and $\alpha \in \Omega^p(B;\mathbf{F}')$, $\beta \in \Omega^q(B;\mathbf{F}')$, then $d(\alpha \wedge \beta) \in \Omega^{p+q}(\mathbf{F})$ and

$$d(\alpha \wedge \beta) = (d\alpha) \underset{\Phi}{\wedge} \beta + (-1)^p \alpha \underset{\Phi}{\wedge} d\beta, \qquad [5.22]$$

*so d is an anti-derivation ([P1], section 2.3.12) of the de Rham algebra $\Omega(B)$ (section 4.4.1**(III)**).*

ii) For $p = 0$, d is the mapping $\mathbf{f} \mapsto df$ specified above.

iii) For every function $f \in \mathcal{C}^r(B;\mathbf{F})$, $d(d\mathbf{f}) = 0$.

PROOF.– Since this is a local question, we can reduce to the case where B is an open subset U of a Banach space $\mathbf{E}$. To simplify the notation, we can restrict to the case where $\mathbf{F} = \mathbb{K}$. Let $\omega \in \Omega^p(B) : U \to \mathrm{Alt}^p(\mathbf{E};\mathbb{K})$ (section 4.4.1**(I)**). The (Fréchet) differential of ω is $D\omega : U \to \mathrm{Alt}^p(\mathbf{E};\mathbb{K})^\vee$. By antisymmetrization, we can associate any continuous linear form $\alpha \in \mathrm{Alt}^p(\mathbf{E};\mathbb{K})^\vee$ with a mapping $\mathbf{a}.\alpha \in \mathrm{Alt}^{p+1}(\mathbf{E};\mathbb{K})$ (section 4.2.2) that we will denote $\varphi_p(\alpha)$. We will show that $d\omega$ is the composition

$$U \xrightarrow{D\omega} \mathrm{Alt}^p(\mathbf{E};\mathbb{K})^\vee \xrightarrow{\varphi_p} \mathrm{Alt}^{p+1}(\mathbf{E};\mathbb{K}).$$

i) If $\alpha \in \Omega^p(U), \beta \in \Omega^q(U)$, then $D(\alpha \wedge \beta) = (D\alpha) \wedge \beta + \alpha \wedge (D\beta)$. Furthermore, $\varphi_{p+q}((D\alpha) \wedge \beta) = \varphi_p(D\omega) \wedge \beta = d\alpha \wedge \beta$ and $\varphi_{p+q}(\alpha \wedge (D\beta)) = (-1)^p \alpha \wedge \varphi_q(D\beta) = (-1)^p \alpha \wedge d\beta$, which shows (i).

ii) This is clear. (iii): See [P1], section 3.3.8(**VII**), Theorem 3.183. ■

DEFINITION 5.21.– *The differential operator* $d : \Omega^p(B;\mathbf{F}) \to \Omega^{p+1}(B;\mathbf{F})$ *is called the* exterior differential.

THEOREM 5.22.– (É. Cartan) *The exterior differential has the following properties:*

i) For every p*-form* $\alpha \in \Omega^p(B;\mathbf{F})$, $d(d\alpha) = 0$*; in other words,* $d^2 = 0$*, or alternatively* d *is a* differential *([P1], section 3.3.8(**III**)).*

Hence, if B *is pure and* n*-dimensional,* $c = (U, \xi, n)$ *is a chart of* B *and* $\alpha = \sum_{i_1,\dots,i_p} \alpha_{i_1,\dots,i_p} d\xi^1 \wedge \dots \wedge d\xi^p$*, then*

$$d\alpha = \sum_{i_1,\dots,i_p} d\alpha_{i_1,\dots,i_p} \wedge d\xi^1 \wedge \dots \wedge d\xi^p.$$

ii) For every morphism $u : B' \to B$ *and every* p*-form* $\alpha \in \Omega^p(B;\mathbf{F})$*, with the notation of section 4.4.2,* $d(u^*(\alpha)) = u^*(d\alpha)$*. In particular,* d *commutes with the operation of restriction to a submanifold.*

iii) For every vector field $X \in \mathcal{T}_0^1(B)$ *and every* p*-form* $\alpha \in \Omega^p(B;\mathbf{F})$*, Cartan's equations hold:*

$$\mathcal{L}_X.(d\alpha) = d(\mathcal{L}_X.\alpha), \qquad [5.23]$$

$$\mathcal{L}_X.\alpha = i_X.d\alpha + d(i_X.\alpha), \qquad [5.24]$$

where $i_X.\alpha$ *is the interior product* $X \lrcorner \alpha$ *(section 4.2.5).*

iv) For every p*-form* $\alpha \in \Omega^p(B)$ *and every vector field* $X_0, \dots, X_p \in \mathcal{T}_0^1(B)$*,*

$$\langle d\alpha, X_0 \wedge \dots \wedge X_p \rangle = \sum_{j=0}^{p} (-1)^j \mathcal{L}_{X_j}.\left\langle \alpha, X_0 \wedge \dots \wedge \widehat{X_j} \wedge \dots \wedge X_p \right\rangle$$

$$+\sum_{0\leq i<j\leq p}^{p}(-1)^{i+j}\left\langle \alpha,[X_i,X_j]\wedge X_0\wedge\ldots\wedge\widehat{X_i}\wedge\ldots\wedge\widehat{X_j}\wedge\ldots\wedge X_p\right\rangle. \quad [5.25]$$

In particular, for $\omega\in\Omega^1(B)$ *and* $X,Y\in\mathcal{T}_0^1(B)$*, we have the* Maurer–Cartan formula

$$\langle d\omega, X\wedge Y\rangle=\mathcal{L}_X.\langle\omega,Y\rangle-\mathcal{L}_Y.\langle\omega,X\rangle-\langle\omega,[X,Y]\rangle. \quad [5.26]$$

If $\omega\in\Omega^1(B;\mathbf{F})$*, this formula can be rewritten as:*

$$d\omega.(X\wedge Y)=\mathcal{L}_X.(\omega.Y)-\mathcal{L}_Y.(\omega.X)-\omega.[X,Y] \quad [5.27]$$

(the relation [5.25] can be modified analogously for a p*-form* $\alpha\in\Omega^p(B;\mathbf{F})$*).*

PROOF.– This theorem is shown inductively. First, for clarity, we will prove Cartan's equation [5.24] in the case where $\mathbf{F}=\mathbb{K}$ and B is locally finite-dimensional (for the general case, see ([ABR 83], section 6.4A)). By [4.14],

$$\langle i_X.d\alpha,Y\rangle=\langle d\alpha,X\wedge Y\rangle=\sum_{i,j}\partial_j\alpha_i\left(X^jY^i-X^iY^j\right).$$

Furthermore,

$$d(i_X.\alpha).Y=\sum_j\partial_j\left(\sum_i\alpha_iX^i\right).Y^j=\sum_{i,j}\left(\partial_j\alpha_iX^i+\alpha_i\partial_jX^i\right)Y^j,$$

which gives that

$$\langle i_X.d\alpha,Y\rangle+d(i_X.\alpha).Y=\sum_{i,j}\left(\partial_j\alpha_iX^iY^i+\alpha_i\partial_jX^iY^j\right)$$

$$\overset{(5.21)}{=}\langle\mathcal{L}_X.\alpha,Y\rangle.$$

■

REMARK 5.23.– *1) The Maurer–Cartan formula [5.27] allows us to calculate* $\langle d\omega(b),\mathbf{h}_b\wedge\mathbf{k}_b\rangle$ *by considering* $X,Y\in\mathcal{T}_0^1(B)$ *such that* $X(b)=\mathbf{h}_b$ *and* $Y(b)=\mathbf{k}_b$*. The bracket* $[X,Y]$ *on the right-hand side of [5.27], evaluated at* b*, depends on the values of* X *and* Y in a neighborhood of b *by [5.16], whereas the left-hand side only depends on the values of* X *and* Y at the point b*, since* $\omega(b)$ *is a bilinear mapping.*

2) Using the exterior product $\bar{\wedge}$ instead of $\wedge$ for p-forms (section 4.2.4, Remark 4.11) changes the exterior differential and hence the Maurer–Cartan formula [5.25] ([KOB 69], Volume I, p. 36), whose right-hand side needs to be multiplied by $\frac{1}{p+1}$ ***(exercise)****.*

(II) EXTERIOR DIFFERENTIAL OF AN ODD p-FORM Let B be a manifold, $\mathbf{F}$ a Banach space and $\underline{\omega} \in \Omega^p \left(B; \tilde{\mathbb{R}} \otimes \mathbf{F}\right)$ an odd differential p-form taking values in $\mathbf{F}$ (section 4.4.6(**I**)). Every sufficiently small non-empty open subset U of B is orientable, and, given any orientation O of U, there exists a bijection $\Omega^p (U; \mathbf{F}) \to \Omega^p \left(U; \tilde{\mathbb{R}} \otimes \mathbf{F}\right) : \omega \mapsto O \otimes \omega$ (Remark 4.50(a)). The exterior differential of $\underline{\omega} = O \otimes \omega$ is defined by the relation

$$\boxed{d (O \otimes \omega) = O \otimes d\omega.}$$

If $X \in \mathcal{T}_0^1 (B)$, we define $\mathcal{L}_X .\underline{\omega}$ in the same way as [5.20], and Cartan's equation [5.24] still holds, *mutatis mutandis*.

5.5.2. *Application to vector calculus*

Consider a real, Riemannian, pure, n-dimensional manifold B (section 4.5) and let $c = (U, \xi, n)$ be a chart such that $\left\{\frac{\partial}{\partial \xi^i}\right\}_{1 \le i \le n}$ is a field of frames.

(I) GRADIENT Suppose that $V \in \mathcal{C}^r (U)$ and let $dV \in \Omega^1 (U)$ be its (exterior) differential. The scalar product on $T_b (U)$ allows us to identify dV with the vector field grad$V = \overrightarrow{\nabla} V \in \mathcal{T}_0^1 (U)$ such that

$$dV.X = \left\langle \overrightarrow{\nabla} V | X \right\rangle$$

for every $X \in \Gamma^{(0)} (U)$ (Corollary-Definition 3.21), called the *gradient* of V. If $\left\{\frac{\partial}{\partial \xi^i}\right\}_{1 \le i \le n}$ is orthonormal, the $\frac{\partial V}{\partial \xi^i}$ $(1 \le i \le n)$ are the components of the gradient $\overrightarrow{\nabla} V$.

(II) DIVERGENCE Let $X \in \mathcal{T}_0^1(U)$. If $\left\{\frac{\partial}{\partial\xi^i}\right\}_{1\leq i\leq n}$ is orthonormal[9], the *divergence* $\operatorname{div}(X)$ of X is the scalar field $\sum_{1\leq i\leq n}\frac{\partial X^i}{\partial\xi^i} \in C^r(U)$. Writing $\overrightarrow{\nabla} = \left(\frac{\partial}{\partial\xi^1}, \ldots, \frac{\partial}{\partial\xi^n}\right)$, we therefore have:

$$\boxed{\operatorname{div} = \left\langle\overrightarrow{\nabla}|-\right\rangle.} \qquad [5.28]$$

(III) BELTRAMI LAPLACIAN Given a field $f \in \mathcal{C}^r(U;\mathbf{F})$, where $\mathbf{F}$ is a Banach space, the Beltrami Laplacian is defined by $\overrightarrow{\nabla}^2 V = \left\langle\overrightarrow{\nabla}|\overrightarrow{\nabla}\right\rangle V$. If $\left\{\frac{\partial}{\partial\xi^i}\right\}_{1\leq i\leq n}$ is orthonormal, it follows that $\overrightarrow{\nabla}^2 V = \sum_{i=1}^n \frac{\partial^2 f}{\partial\xi_i^2}$; thus, if $\mathbf{F} = \mathbb{R}$,

$$\boxed{\overrightarrow{\nabla}^2 = \sum_{i=1}^n \frac{\partial^2}{\partial\xi_i^2} = \operatorname{div}\circ\operatorname{grad}.} \qquad [5.29]$$

(IV) CURL Suppose that $n = 3$ and U is equipped with the orientation induced by the canonical orientation of $\mathbb{R}^3$ (section 4.4.4). The curl is the operator $\overrightarrow{\nabla} \wedge -$, where $\wedge$ denotes the usual vector product of two vectors in the Euclidean space $\mathbb{R}^3$ (Example 4.13(iii)).

(V) REMARKS

1) The gradient, divergence and Laplacian depend on the Riemannian structure of the manifold.

2) The curl only exists for $n = 3$ and depends on the orientation of the Riemannian manifold.

3) If U is an open subset of $\mathbb{R}^3$ equipped with the canonical orientation, let $\omega = A.dx + B.dy + C.dz$. Then, $d\omega = A_1.dx\wedge dy + B_1.dy\wedge dz + C_1.dz\wedge dx \in \Omega^1(U)$, where A_1, B_1, C_1 are the components of curl$X$ and X is the vector field with components A, B, C.

Let $\Omega = A_1.dx\wedge dy + B_1.dy\wedge dz + C_1.dz\wedge dx \in \Omega^2(U)$. Then, $d\Omega = A_2 dx\wedge dy\wedge dz$, where $A_2 = \operatorname{div}(Y)$ and Y is the vector field with components A_1, A_2, A_3.

9 For the case where $\left\{\frac{\partial}{\partial\xi^i}\right\}_{1\leq i\leq n}$ is not assumed to be orthonormal, see Theorem-Definition 5.42 below.

4) From the above, we can deduce the following relations on an open subset U of $\mathbb{R}^3$ equipped with the canonical orientation (the last equality can be deduced from [4.17] in Example 4.13 of section 4.2.5):

$$\text{curl} \circ \text{grad} = \overrightarrow{\nabla} \wedge \overrightarrow{\nabla} = 0, \qquad [5.30]$$

$$\text{div} \circ \text{curl} = \left\langle \overrightarrow{\nabla} | \overrightarrow{\nabla} \wedge - \right\rangle = 0, \qquad [5.31]$$

$$\text{curl} \circ \text{curl} = \text{grad} \circ \text{div} - \overrightarrow{\nabla}^2. \qquad [5.32]$$

5.6. Stokes' formula and applications

In this section, every manifold is real, pure, finite-dimensional and of class C^r $(2 \leq r \leq \infty)$.

5.6.1. *Stokes' formula on a chain*

Let B be an m-dimensional manifold. Let τ be a chain of simplexes in B (section 4.4.7(**II**)) and ω a differential m-form of class C^1 taking values in a Banach space $\mathbf{F}$. Suppose that $\text{supp}\,(\omega) \cap \tau$ is compact and furthermore assume either of the following hypotheses:

(H$_1$) τ is an odd chain whose boundary $\partial\tau$ is equipped with the induced orientation (section 4.4.7(**VI**)), and ω is even.

(H$_2$) τ is an even chain and ω is odd.

THEOREM 5.24.– (Stokes) *The following relation holds:*

$$\boxed{\int_{\partial\tau} \omega = \int_\tau d\omega.}$$

PROOF.– By linearity, we can reduce to the case where τ is an m-simplex. By change of variables, we can further reduce to the case where τ is the standard m-simplex in $\mathbb{R}^m$ (section 4.4.7(**I**)).

1) To illustrate, let us perform the calculation for $m = 2$ first. Thus, consider the triangle Δ^2 from Figure 4.2 (section 4.4.7(**V**)). The equation of the hypotenuse is $x + y = 1$. Consider the differential 1-form $\omega = \mathbf{P}.dx + \mathbf{Q}.dy$, where $\mathbf{P}, \mathbf{Q}$ are mappings of class C^1 from an open neighborhood U of $\tau := \Delta^2$ into $\mathbf{F}$. In the case of **(H$_1$)**, for example, we have:

$$d\omega = \left(\frac{\partial \mathbf{Q}}{\partial x} - \frac{\partial \mathbf{P}}{\partial y} \right) dx \wedge dy.$$

The first term can be calculated as follows by the Fubini–Tonelli theorem ([P2], section 4.1.3(**III**)).

$$\int_{\tau} \frac{\partial \mathbf{Q}}{\partial x} dx \wedge dy = \int_0^1 \left[\mathbf{Q}\left(x,y\right)\right]_{x=0}^{x=1-y} dy$$
$$= \underbrace{\int_0^1 \mathbf{Q}\left(1-y,y\right) dy}_{\text{integration over } [v_1, v_2]} + \underbrace{\int_1^0 \mathbf{Q}\left(0,y\right) dy}_{\text{integration over } [v_2, v_0]} = \int_{\partial\tau} \mathbf{Q}.dy.$$

Similarly, the second term gives:

$$-\int_{\tau} \frac{\partial \mathbf{P}}{\partial x} dx \wedge dy = -\int_0^1 \left[\mathbf{P}\left(x,y\right)\right]_{y=0}^{y=1-x} dx$$
$$= \underbrace{\int_0^1 \mathbf{P}\left(x,0\right) dx}_{\text{integration over } [v_0, v_1]} + \underbrace{\int_1^0 \mathbf{P}\left(x,1-x\right) dx}_{\text{integration over } [v_1, v_2]} = \int_{\partial\tau} \mathbf{P}.dx,$$

which proves the result.

2) Next, consider the general case, while remaining in the case $(\mathbf{H}_1)$; consider an oriented simplex Π with vertices $v_0, v_1, ..., v_m$, as well as the coordinate system $(t_1, ..., t_m) \in \mathbb{R}^m$ determined by writing each point P of this simplex in the form:

$$P = v_0 + \sum\nolimits_{i=1}^{m} t_i.\left(v_i - v_0\right) = \sum\nolimits_{i=0}^{m} t_i.v_i,$$

with $t_0 = 1 - \sum_{i=1}^{p} t_i$. Every m-form ω is a sum of m terms. We simply need to show Stokes' formula for one of these terms, for example $\omega = \mathbf{Q}.dt_2 \wedge ... \wedge dt_p$. Let Π_i be the face of the simplex opposite to v_i, P_i the hyperplane containing this face and $\iota_i : P_i \hookrightarrow \mathbb{R}^p$ the inclusion mapping. Equipping $\partial\Pi$ and $\partial\Pi_i$ with the induced orientation (section 4.4.7(**VI**), Definition 4.63), we have:

$$\int_{\partial\Pi} \omega = \sum_{i=0}^{m} \int_{\Pi_i} \iota_i^* \left(\omega\right),$$

where $\iota_i^*\left(\omega\right)$ is the preimage of ω under ι_i (section 4.4.2). For every $i > 1$, $\iota_i^*\left(\omega\right) = 0$ (in (1) above, $\mathbf{Q}.dy = 0$ on $\Pi_2 = [v_0, v_1]$). In the two hyperplanes P_0 and $P_1 = 0$, $(t_2, ..., t_m)$ is a coordinate system. In P_0, this system is positively oriented and $t_1 = 1 - \sum_{i=2}^{p} t_i$, so

$$\iota_0^*\left(\omega\right) = \mathbf{Q}\left(1 - \sum_{i=2}^{m} t_i, t_2, ..., t_m\right) dt_2 \wedge ... \wedge dt_m.$$

In P_1, the coordinate system $(t_2, ..., t_m)$ is negatively oriented (in (1) above, the integration is performed from v_2 to v_0, i.e. with respect to $-dy$), and hence

$$\iota_1^* (\omega) = -\mathbf{Q} (0, t_2, ..., t_m) \, dt_2 \wedge ... \wedge dt_m.$$

The integration is performed over the region of $\mathbb{R}^{m-1}$ determined by its characteristic function $\chi = \chi (0, t_2, ..., t_m)$. Therefore,

$$\int_{\partial \Pi} \omega = \int_{\Pi_0} \iota_0^* (\omega) + \int_{\Pi_1} \iota_1^* (\omega)$$
$$= \int_{\mathbb{R}^{m-1}} \chi . \underbrace{\left(\mathbf{Q} \left(1 - \sum_{i=2}^{m} t_i, t_2, ..., t_m \right) - \mathbf{Q} (0, t_2, ..., t_m) \right)}_{\mathbf{P}} dt_2 ... dt_m.$$

Furthermore, $d\omega = \frac{\partial \mathbf{Q}}{\partial t_1} dt_1 \wedge ... \wedge dt_m$. The characteristic function of Π is the product $\chi . \chi'$, where $\chi' (t_1, ... t_m)$ is the characteristic function of the domain defined by $0 < t_1 < 1 - \sum_{i=2}^{m} t_i$, so

$$\int_{\Pi} d\omega = \int_{\mathbb{R}^{m-1}} dt_2 ... dt_m . \chi (0, t_2, ..., t_m) \underbrace{\int_{\mathbb{R}} dt_1 . \chi' (t_1, ... t_m) \frac{\partial \mathbf{Q}}{\partial t_1}}_{\mathbf{P}},$$

which is identical to the expression found earlier for $\int_{\partial \Pi} \omega$. ■

Part (1) proves the Green–Riemann formula. We can go further by specifying orientations when the hypothesis $(\mathbf{H}_1)$ is assumed:

COROLLARY 5.25.– (Green–Riemann formula) *Let B be an open subset of the plane equipped with its canonical orientation, τ an odd chain and $\omega = \mathbf{P}.dx + \mathbf{Q}.dy$ an even differential 1-form taking values in a Banach space $\mathbf{F}$. Assume that $\mathbf{P}$ and $\mathbf{Q}$ are of class C^1 from B into $\mathbf{F}$ and that supp $(\omega) \cap \tau$ is compact. Then, equipping $\partial \tau$ with the induced orientation (see Figure 5.1),*

$$\boxed{\iint_{\tau} \left(\frac{\partial \mathbf{Q}}{\partial x} - \frac{\partial \mathbf{P}}{\partial y} \right) dx \wedge dy = \int_{\partial \tau} (\mathbf{P}.dx + \mathbf{Q}.dy).}$$

Another special case of Stokes' formula is obtained by considering an oriented path $\tau = a \to b$ that is piecewise of class C^1 in a manifold B, together with a mapping $\mathbf{f}$ of class C^1 from an open neighborhood of τ into a Banach space $\mathbf{F}$. This recovers the fundamental theorem of integration:

$$\int_{a \to b} d\mathbf{f} = \mathbf{f} (b) - \mathbf{f} (a) .$$

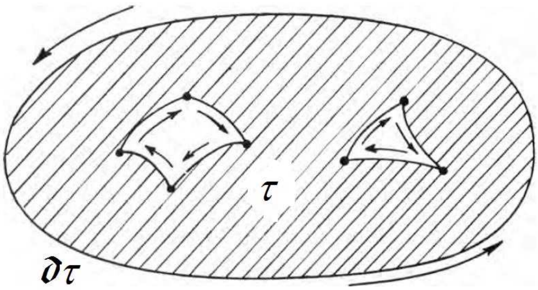


Figure 5.1. *Orientation of the boundary (Green–Riemann formula)*

5.6.2. *Ostrogradsky and Green formulas*

(I) OSTROGRADSKY'S FORMULA Let τ be a chain of simplexes in an m-dimensional manifold B, α an m-form of class C^1 taking values in a Banach space $\mathbf{F}$ and X a vector field of class C^1 in an open neighborhood of τ. Then, $d\alpha = 0$, so $\mathcal{L}_X.\alpha = d\,(i_X.\alpha)$ by Cartan's equation [5.24]. Stokes' formula (with one of the hypotheses $(\mathbf{H}_1)$ or $(\mathbf{H}_2)$) implies that:

THEOREM 5.26.– (Gauss) *The following relation holds:*

$$\boxed{\int_\tau \mathcal{L}_X.\alpha = \int_{\partial\tau} i_X.\alpha.} \qquad [5.33]$$

In the case where B is a Riemannian manifold, let $\upsilon = \upsilon^1 \wedge ... \wedge \upsilon^m$ be its volume form (section 4.5(**II**)) and write σ for the volume form on $\partial\tau$. Then, $\mathcal{L}_X.\upsilon = \operatorname{div}(X).\upsilon$ and there exists a 1-form ϖ such that $\upsilon = \varpi \wedge \sigma$. Using the scalar product, we can identify ϖ with a unit vector field $\mathbf{n}$ such that, for all $b \in \partial\tau$, $\mathbf{n}_b \in T_b(\tau)$ is orthogonal to $T_b(\partial\tau)$. Thus, by [4.15] (section 4.2.4),

$$i_X.\upsilon = \langle \mathbf{n} | X \rangle \sigma - \varpi \wedge (i_X.\sigma)$$

and $\varpi \wedge (i_X.\sigma) = 0$ on $\partial\tau$.

EXAMPLE 5.27.– *Suppose, for example, that τ is the square $ABCD$ in the plane $(0, x^1, x^2)$, with $A = (0,0)$, $B = (1,0)$, $C = (1,1)$, $D = (0,1)$. First, integrate over AB. We have $\sigma = dx^1, \varpi = -dx^2$. If $X = (X^1, X^2)$, then, by [4.16], $i_X\sigma = X^1$, so $\varpi \wedge (i_X.\sigma) = -X^1 dx^2$, hence $\varpi \wedge (i_X.\sigma) = 0$ on AB, where $x^2 = 0$. Similarly, $\varpi \wedge (i_X.\sigma) = 0$ on BC, CD and DA.*

By [5.33], we obtain:

THEOREM 5.28.– (Ostrogradsky) *The following relation holds (sometimes called the* flux-divergence theorem*):*

$$\boxed{\int_\tau div\,(X)\,.\upsilon = \int_{\partial\tau} \langle \mathbf{n}|X\rangle\,.\sigma.} \qquad [5.34]$$

The quantity on the right-hand side can be interpreted as the *flux* of the vector field X through $\partial\tau$ (if B is three-dimensional Euclidean space, then τ is a volume, $\partial\tau$ is the surface enclosing this volume and $\mathbf{n}$ is the outward normal to $\partial\tau$: see Figure 5.2).

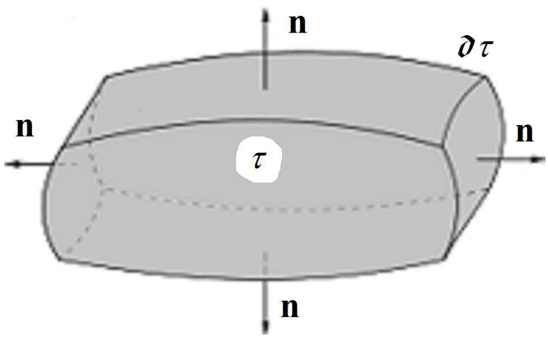


Figure 5.2. *Flux through* $\partial\tau$

REMARK 5.29.– *In particular, if* X_b *denotes some physical quantity at the point* b *(e.g. momentum, electric charge), then the relation* $div\,(X) = 0$ *means that, given any volume* τ*, nothing can escape from* τ*; the relation* $div\,(X) = 0$ *is therefore a "conservation relation".*

(II) GREEN'S FORMULAS Once again, let B be a Riemannian manifold and τ a chain in B, with the same notation as before. Let φ be a scalar field of class C^1 in an open neighborhood of $\bar{\tau}$. The *normal derivative* of φ is the function defined on $\partial\tau$ by the relation:

$$\partial_{\mathbf{n}}\varphi = \langle \mathbf{n}, \nabla\varphi\rangle\,.$$

THEOREM 5.30.– (Green) *Consider scalar fields* φ *and* ψ *of class* C^2 *in an open neighborhood of* $\bar{\tau}$ *whose support intersects compactly with* $\bar{\tau}$. Green's first formula *can be stated as follows:*

$$\int_\tau \langle\nabla\varphi|\nabla\psi\rangle\,.\upsilon + \int_\tau \varphi.\nabla^2\psi.\upsilon = \int_{\partial\tau} \varphi.\partial_{\mathbf{n}}\psi.\sigma. \qquad [5.35]$$

Green's second formula *is given by:*

$$\boxed{\int_\tau \left(\varphi\nabla^2\psi - \psi\nabla^2\varphi\right).\upsilon = \int_{\partial\tau}\left(\varphi\partial_{\mathbf{n}}\psi - \psi\partial_{\mathbf{n}}\varphi\right).\sigma.} \qquad [5.36]$$

PROOF.– i) We have $\langle\varphi.\nabla\psi|\mathbf{n}\rangle = \varphi.\langle\nabla\psi|\mathbf{n}\rangle = \varphi.\partial_{\mathbf{n}}\psi$. Furthermore, in an orthonormal frame,

$$\text{div}\,(\varphi.\nabla\psi) = \sum_i \frac{\partial}{\partial\xi^i}\left(\varphi\frac{\partial\psi}{\partial\xi^i}\right) = \langle\nabla\varphi|\nabla\psi\rangle + \varphi.\nabla^2\psi.$$

Ostrogradsky's theorem (Theorem 5.28) gives

$$\int_\tau \text{div}\,(\varphi.\nabla\psi).\upsilon = \int_{\partial\tau}\langle\varphi.\nabla\psi|\mathbf{n}\rangle.\sigma = \int_{\partial\tau}\varphi.\partial_{\mathbf{n}}\psi.\sigma,$$

and we therefore obtain [5.35].

ii) By switching the roles of φ and ψ in [5.35] and subtracting the result from the original formula, we obtain [5.36]. ■

Green's formulas have many applications, such as the following example with harmonic functions:

COROLLARY 5.31.– (Hopf's theorem) *Let B be a Riemannian manifold without a boundary and f a* compactly supported *function of class C^2 in B such that $\nabla^2 f \leq 0$. Then, f is* harmonic, *i.e.* $\nabla^2 f = 0$. *It follows that f is constant: every compactly supported harmonic function is constant.*

PROOF.– i) Let us show that f is harmonic. To do this, we can apply Green's second formula with $\varphi = f$, $\psi = 1$ and $\tau = B$, i.e. $\partial\tau = \varnothing$. This gives $\int_B \left(\nabla^2 f\right).\upsilon = 0$, and, since $\nabla^2 f \leq 0$, $\nabla^2 f = 0$.

ii) Next, show that f is constant. We can apply Green's first formula with $\varphi = f^2$, $\psi = 1$ and $\tau = B$. We can immediately check that

$$\nabla^2 f^2 = 2f\nabla^2 f + 2\left\|\nabla f\right\|^2,$$

so $\nabla^2 f^2 = 2\left\|\nabla f\right\|^2$, since $\nabla^2 f = 0$, and hence $\int_B \left\|\nabla f\right\|^2 dv = 0$, giving $\nabla f = 0$ and $f = const.$ ■

Let τ be an odd chain of a surface Σ in the oriented Euclidean space E^3 and X a vector field of class C^1 in an open neighborhood of $\bar{\tau}$. Let $\mathbf{t}$ be the unit tangent vector of $\partial\tau$. Assuming that supp $(X)\cap\bar{\tau}$ is compact, the work performed by X along $\partial\tau$ is $\int_{\partial\tau}\langle X|\mathbf{t}\rangle .ds$, where s denotes the arc length. Writing σ for the "surface form", Theorem 5.24 then gives us the original form of Stokes' theorem, also known as the Kelvin–Stokes theorem (see Figure 5.3):

COROLLARY 5.32.– (Kelvin–Stokes theorem) *The work performed by X along $\partial\tau$ is equal to the flux of the curl of X through τ :*

$$\boxed{\int_\tau \langle curlX|\mathbf{n}\rangle .\sigma = \int_{\partial\tau} \langle X|\mathbf{t}\rangle .ds.}$$

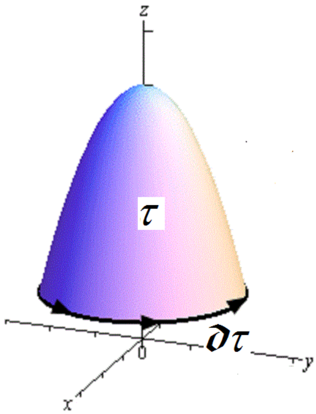


Figure 5.3. *Kelvin–Stokes theorem*

5.6.3. *Hodge duality and codifferentials*

(I) Let B be a real oriented n-dimensional pseudo-Riemannian manifold with metric $\mathbf{g}$ (section 4.5).

LEMMA 5.33.– *1) The metric* $\mathbf{g}$ *induces a symmetric non-degenerate tensor field* $\mathbf{g}^{(k)}$ *on* $\Omega^k(B)$ $(k = 0, 1, ..., n)$ *as follows: let* $(\mathbf{h}_i)_{1\leq i\leq n}$ *be an orthonormal frame of* $T(B)$, $(\theta^i)_{1\leq i\leq n}$ *the dual coframe (section 3.4.2, Lemma-Definition 3.38), and*

$$\alpha = \sum_{i_1,...,i_k} \alpha_{i_1,...,i_k} \theta^{i_1} \wedge ...\theta^{i_k},$$

$$\beta = \sum_{i_1,...,i_k} \beta_{i_1,...,i_k} \theta^{i_1} \wedge ...\theta^{i_k}$$

two arbitrary elements of $\Omega^k(B)$ *. Let* $g_{ij} = \mathbf{g}(\mathbf{h}_i, \mathbf{h}_j)$, (g^{ij}) *be the inverse matrix of* (g_{ij}) *and*

$$\beta^{i_1,...,i_k} = \sum_{j_1,...,j_k} g^{i_1,j_1,...,i_k,j_k} \beta_{j_1,...,j_k}$$

Then:

$$\mathbf{g}^{(k)}(\alpha,\beta)=\sum_{i_1<\ldots<i_k}\alpha_{i_1,\ldots,i_k}\beta^{i_1,\ldots,i_k}.$$

2) This symmetric non-degenerate tensor field is independent of the choice of orthonormal frame $(\mathbf{h}_i)_{1\leq i\leq n}$*. We write that* $\mathbf{g}^{(k)}(\alpha,\beta)=\langle\alpha|\beta\rangle$*.*

3) In particular, if the matrix of $\mathbf{g}$ *in the frame* $(\mathbf{h}_i)_{1\leq i\leq m}$ *is diag* (η_i)*, where* $\eta_i=\pm1$*, i.e. if* $\mathbf{g}=\sum_i\eta_i\theta^i\wedge\theta^j$*, then:*

$$\left\langle\theta^{i_1}\wedge\ldots\wedge\theta^{i_k}|\theta^{i_1}\wedge\ldots\wedge\theta^{i_k}\right\rangle=\eta_{i_1}\ldots\eta_{i_k}.$$

PROOF.– Let us show that $\mathbf{g}^{(k)}$ is non-degenerate. If $\mathbf{g}^{(k)}(\alpha,\beta)=0$ for all $\beta\in\Omega^k(B)$, then in particular $\mathbf{g}^{(k)}\left(\alpha,\theta^{i_1}\wedge\ldots\wedge\theta^{i_k}\right)=0$, so $\alpha_{i_1,\ldots,i_k}=0$. This is true for every k-tuple of indices $\{i_1,\ldots,i_k\}$, so $\alpha=0$. ■

LEMMA 5.34.– *1) There exists a uniquely determined isomorphism* $*:\Omega^k(B)\to\Omega^{n-k}(B)$ *such that*

$$\alpha\wedge*\beta=\langle\alpha|\beta\rangle\,\theta\quad\text{for every }\alpha,\beta\in\Omega^k(B),\qquad[5.37]$$

where θ *is the pseudo-Riemannian volume element (section 4.5).*

2) If $(\mathbf{h}_i)_{1\leq i\leq n}$ *is a positively-oriented orthonormal frame (section 4.4.4(**II**), Definition 4.37) and* $\left(\theta^i\right)_{1\leq i\leq n}$ *is the dual coframe, then, for every permutation* $\sigma\in\mathfrak{S}_n$*,*

$$*\left(\theta^{\sigma(1)}\wedge\ldots\wedge\theta^{\sigma(k)}\right)=\eta_{\sigma(1)}\ldots\eta_{\sigma(k)}.\varepsilon_\sigma.\left(\theta^{\sigma(k+1)}\wedge\ldots\wedge\theta^{\sigma(n)}\right)\qquad[5.38]$$

for $\sigma(1)<\ldots<\sigma(k)$ *and* $\sigma(k+1)<\ldots<\sigma(n)$*.*

PROOF.– i) Let us first show uniqueness by assuming that $*$ satisfies [5.37]. Let $\beta=\theta^{\sigma(1)}\wedge\ldots\wedge\theta^{\sigma(k)}$ and suppose that α is one of the k-vectors $\theta^{i_1}\wedge\ldots\wedge\theta^{i_k}$. Then, by Lemma 5.33, $\alpha\wedge*\beta=0$ except if $(i_1,\ldots,i_k)=(\sigma(1),\ldots,\sigma(k))$. Hence, there exists a scalar field a such that $*\beta=a.\theta^{\sigma(k+1)}\wedge\ldots\wedge\theta^{\sigma(n)}$. It follows that $\beta\wedge*\beta=a.\varepsilon_\sigma.\theta$ by Lemma 4.8 (section 4.2.4). But Lemma 5.33(3) implies that $\langle\beta|\beta\rangle=\eta_{\sigma(1)}\ldots\eta_{\sigma(k)}$, so $a=\varepsilon_\sigma\eta_{\sigma(1)}\ldots\eta_{\sigma(k)}$, and $*$ satisfies [5.38], which shows uniqueness.

ii) Define $*$ as in [5.38], noting that

$$\left\{\theta^{\sigma(1)}\wedge\ldots\wedge\theta^{\sigma(k)}:\sigma(1)<\ldots<\sigma(k)\right\}$$

is an orthonormal basis of $\Omega^k(B)$. Then, $*$ sends an orthonormal basis of $\Omega^k(B)$ onto an orthonormal basis of $\Omega^{n-k}(B)$; it is therefore an isomorphism and [5.37] follows from [5.38]. ■

DEFINITION 5.35.– *The isomorphism* $*:\Omega^k(B)\to\Omega^{n-k}(B)$ *is called the* Hodge operator.

Let $\eta=\eta_1...\eta_n=(-1)^{\mathrm{ind}(\mathbf{g})}$, where $\mathrm{ind}(\mathbf{g})$ is the number of $-$ signs in the canonical form of $\mathbf{g}$, so that $\mathbf{g}$ has signature $(n-\mathrm{ind}(\mathbf{g}),\mathrm{ind}(\mathbf{g}))$ (section 4.5.1, Definition 171). It is easy to check the following properties:

PROPOSITION 5.36.– *Let* $\alpha,\beta\in\Omega^k(B)$.

1) $\alpha\wedge*\beta=\beta\wedge*\alpha=\langle\alpha|\beta\rangle\,\theta$*;*

2) $*1=\theta$*;*

3) $*\theta=(-1)^{ind(\mathbf{g})}$*;*

4) $**\alpha=(-1)^{ind(\mathbf{g})+k(n-k)}\,\alpha$*;*

5) $\langle\alpha|\beta\rangle=(-1)^{ind(\mathbf{g})}\langle*\alpha|*\beta\rangle$*;*

6) $\langle*\alpha|\beta\rangle=\langle\alpha\wedge\beta|\theta\rangle$*.*

EXAMPLE 5.37.– *The Hodge operator on* $\Omega^1(\mathbb{R}^3)$ *is determined by the following properties:* $*\theta^1=\theta^2\wedge\theta^3$*,* $*\theta^2=-\theta^1\wedge\theta^3$ *and* $*\theta^3=\theta^1\wedge\theta^2$*.*

Note the analogy with the usual vector product. To clarify this analogy, we can introduce the "musical operators" flat $^\flat$ and sharp $^\sharp$ as follows: given two vector fields $X,Y\in\mathcal{T}^1(B)$, write $\langle X|Y\rangle=\mathbf{g}(X,Y)$. We then have two isomorphisms:

$$^\flat:\mathcal{T}^1(B)\to\mathcal{T}_1(B):X\mapsto\langle X,.\rangle,$$
$$^\sharp:\mathcal{T}_1(B)\to\mathcal{T}^1(B),\ \text{the inverse isomorphism.}$$

Hence, if $X=\sum_i X^i\mathbf{e}_i\in\mathcal{T}^1(B)$, where $\{\mathbf{e}_1,...,\mathbf{e}_n\}$ is a frame, then $X^\flat=\sum_i X_i\mathbf{e}^{\vee i}$, where $\{\mathbf{e}^{\vee 1},...,\mathbf{e}^{\vee n}\}$ is the dual coframe and $X_i=\sum_j g_{ij}X^j$, with $g_{ij}=\langle\mathbf{e}_i|\mathbf{e}_j\rangle$. Conversely, if $\alpha=\sum_i\alpha_i\mathbf{e}^{\vee i}\in\mathcal{T}_1(B)$, then $\alpha^\sharp=\sum_i\alpha^i\mathbf{e}_i$, where $\alpha^i=\sum_j g^{ij}\alpha_i$ and (g^{ij}) is the inverse matrix of (g_{ij}).

(II) The above shows that the metric $\mathbf{g}$ induces the following symmetric non-degenerate bilinear forms on $\mathcal{T}^1(B)$ and $\mathcal{T}_1(B)$:

$$\langle X|Y\rangle=\sum_i X^iY_j=\sum_{i,j}g_{ij}X^iY^j,\quad\langle\alpha|\beta\rangle=\sum_i\alpha_i\beta^i=\sum_{i,j}g^{ij}\alpha_i\beta_j.$$

More generally, we have the following rule ([TON 65], p. 22):

(R) If $\mathbf{Z}$ is a tensor field of type (p,q) with components $Z^{i_1,\dots,i_p}_{j_1,\dots,j_q}$, we *lower* the index i_k $(1 \le k \le p)$ (from contravariant to covariant) by writing

$$Z^{i_1,\dots,\widehat{i_k},\dots,i_p}_{i'_k,i_1,\dots,i_q} = \sum_{i_k} g_{i'_k i_k} Z^{i_1,\dots,i_k,\dots,i_p}_{j_1,\dots,j_q}.$$

We *raise* the index j_k $(1 \le k \le q)$ (from covariant to contravariant) by writing

$$Z^{j'_k,i_1,\dots,i_p}_{j_1,\dots,\widehat{j_k},\dots,j_q} = \sum_{j_k} g^{j'_k j_k} Z^{i_1,\dots,i_p}_{j_1,\dots,j_k,\dots,j_q}.$$

In particular,

$$g^{i}_{j'} = \sum_{j} g_{j'j} g^{ij} = \sum_{i'} g^{ii'} g_{j'i'} = \delta^{i}_{j'}.$$

COROLLARY 5.38.– *Let $f \in \mathcal{C}^r(B)$. Then,*

$$\boxed{\mathit{grad} f = (df)^{\sharp}.}$$

Given positively oriented local coordinates, we have:

$$\operatorname{grad} f = \sum_{i,j} g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}. \qquad [5.39]$$

PROOF.– We have $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$ and $\left(dx^i\right)^{\sharp} = \sum_j g^{ij} \frac{\partial}{\partial x^j}$. ■

EXAMPLE 5.39.– *1) In the Euclidean space $\mathbb{R}^3$, consider the vector field $F = F_1 \frac{\partial}{\partial x} + F_2 \frac{\partial}{\partial y} + F_3 \frac{\partial}{\partial z}$, where $F_i = F_i(x,y,z)$ $(i = 1,2,3)$. Then, $F^{\flat} = F_1 dx + F_2 dy + F_3 dz$, so:*

$$dF^{\flat} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx \wedge dy - \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) dx \wedge dz + \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) dy \wedge dz.$$

Furthermore, if $G = G_1 \frac{\partial}{\partial x} + G_2 \frac{\partial}{\partial y} + G_3 \frac{\partial}{\partial z}$, then, by Example 5.37,

$$*G^{\flat} = G_3 dx \wedge dy - G_2 dx \wedge dz + G_1 dy \wedge dz. \qquad [5.40]$$

Hence, $dF^\flat = *(curlF)^\flat$, *or alternatively*

$$\boxed{curl = *d.}$$

Moreover, by [5.40], $d * G^\flat = (divG)\, dx \wedge dy \wedge dz$. *Hence,* $divG = *d * G^\flat$.

2) Let us find the gradient in spherical coordinates [4.26] (section 4.4.5, Example 4.41(2)). We have $\begin{bmatrix} dx\ dy\ dz \end{bmatrix}^T = J \begin{bmatrix} dr\ d\theta\ d\phi \end{bmatrix}^T$, *where* J *is the Jacobian matrix from the equality [4.27]. The usual Euclidean metric of* $\mathbb{R}^3$ *is given by:*

$$\mathbf{g} = \begin{bmatrix} dx\ dy\ dz \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} dr\ d\theta\ d\phi \end{bmatrix} G \begin{bmatrix} dr \\ d\theta \\ d\phi \end{bmatrix},$$

where

$$G = J^T J = diag\left(1, r^2, r^2 \sin^2 \phi\right). \qquad [5.41]$$

Hence, by (5.39), $gradf = \frac{\partial f}{\partial r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial f}{\partial \phi}\frac{\partial}{\partial \phi} + \frac{1}{r^2 \sin^2 \phi}\frac{\partial f}{\partial \theta}\frac{\partial}{\partial \theta}$. *In the right-handed orthonormal frame* $\{\overrightarrow{u_r}, \overrightarrow{u_\phi}, \overrightarrow{u_\theta}\}$, *by [4.29], we therefore have:*

$$gradf = \frac{\partial f}{\partial r}\overrightarrow{u_r} + \frac{1}{r}\frac{\partial f}{\partial \phi}\overrightarrow{u_\phi} + \frac{1}{r \sin \phi}\frac{\partial f}{\partial \theta}\overrightarrow{u_\theta}. \qquad [5.42]$$

DEFINITION 5.40.– *The* codifferential $\boldsymbol{\delta} : \Omega^k(B) \to \Omega^{k-1}(B)$ *is defined by* $\boldsymbol{\delta}\left(\Omega^0(B)\right) = 0$ *and, on* $\Omega^k(B)$ $(k \geq 1)$*:*

$$\boldsymbol{\delta}\alpha = (-1)^{n(k+1)+1+ind(\mathbf{g})} * d * \alpha.$$

This definition is justified by the following property:

PROPOSITION 5.41.– *Let* B *be a real compact orientable Riemannian manifold without a boundary.*

i) Let $L^2\left(B; Alt^k(B)\right)$ *be, the subspace of* $\Gamma^{(0)}\left(B, Alt^k(B)\right)$ *formed by the continuous* k*-forms* α *satisfying*

$$\int_B \alpha \wedge *\alpha = \int_B \langle \alpha | \alpha \rangle\, \theta < +\infty.$$

Then, $\langle . | . \rangle_2 : L^2\left(B; Alt^k(B)\right) \times L^2\left(B; Alt^k(B)\right) \to \mathbb{R} : (\alpha, \beta) \mapsto \int_B \alpha \wedge *\beta$ *is a symmetric positive definite bilinear form, and* $L^2\left(B; Alt^k(B)\right)$, *equipped with this form, is therefore a Hausdorff pre-Hilbert space ([P2], section 3.10.1)*[10].

10 The bilinear form $\langle . | . \rangle_2$ should not be confused with $\langle . | . \rangle$ introduced in Lemma 5.33.

ii) The codifferential $\boldsymbol{\delta}$ is the adjoint operator of the exterior differential d for this form. In other words, for any differential form $\alpha \in \Omega^k(B)$, $\beta \in \Omega^{k+1}(B)$,

$$\langle d\alpha|\beta\rangle_2 = \langle \alpha|\boldsymbol{\delta}\beta\rangle_2 .$$

PROOF.– i) If $\langle\alpha|\beta\rangle_2 = 0$ for all $\beta \in L^2\left(B; \mathrm{Alt}^k(B)\right)$, then, in particular, $0 = \langle\alpha|\alpha\rangle_2 = \int_B \langle\alpha|\alpha\rangle\, \theta$, and, since $\langle\alpha(x)|\alpha(x)\rangle \geq 0$ (the manifold is Riemannian), $\alpha = 0$. The form $\langle .|.\rangle_2$ is symmetric by Proposition 5.36(i).

ii) Let us calculate $I = \langle d\alpha|\beta\rangle_2 - \langle\alpha|\boldsymbol{\delta}\beta\rangle_2 = \int_B (d\alpha \wedge *\beta - \alpha \wedge *\boldsymbol{\delta}\beta)$. We have:

$$\begin{aligned} d\alpha \wedge *\beta - \alpha \wedge *\boldsymbol{\delta}\beta &= d\alpha \wedge *\beta + (-1)^{n(k+2)+\mathrm{ind}(\mathbf{g})}\, \alpha \wedge * * d * \beta \\ &= d\alpha \wedge *\beta + (-1)^{n(k+2)+\mathrm{ind}(\mathbf{g})}\, (-1)^{\mathrm{ind}(\mathbf{g})+k(n-k)}\, \alpha \wedge d * \beta \\ &= d\alpha \wedge *\beta + (-1)^{k^2}\, \alpha \wedge d * \beta = d(\alpha \wedge *\beta) . \end{aligned}$$

By Stokes' theorem, $I = 0$, so $d\alpha \wedge *\beta = \alpha \wedge *\boldsymbol{\delta}\beta$. ■

We can now give an intrinsic definition of the divergence:

THEOREM-DEFINITION 5.42.– *A vector field $X \in \mathcal{T}^1(B)$ uniquely determines a scalar field $\mathfrak{D}$ such that $\mathcal{L}_X.\theta = \mathfrak{D}.\theta$. This scalar field is the* divergence *of X and is given by [5.28], i.e.*

$$\boxed{div(X) = -\boldsymbol{\delta} X^\flat.} \qquad [5.43]$$

Given positively-oriented local coordinates ξ^i $(1 \leq i \leq n)$, with $g = \det(g_{ij})$, we have:

$$div(X) = \frac{1}{\sqrt{|g|}} \sum_k \frac{\partial}{\partial \xi^k}\left(\sqrt{|g|}.X^k\right). \qquad [5.44]$$

PROOF.– 1) We have the relation $i_X.\theta = *X^\flat$. Indeed, if $X(b) \neq 0$, choose $v_2, ..., v_n \in T_b(B)$ such that $\left\{\frac{X(b)}{\mathbf{g}(X,X)(b)}, v_2, ..., v_n\right\}$ is a positively oriented orthonormal basis of $T_b(B)$. Then, by [4.14],

$$\begin{aligned} (i_X.\lambda)(v_2, ..., v_n) &= \theta(X(b), v_2, ..., v_n) = \mathbf{g}(X, X)(b) \\ &= X^\flat(X)(b)(v_2, ..., v_n) . \end{aligned}$$

2) Hence, by Theorem 5.22(iii), $\mathcal{L}_X.\theta = d\left(i_X.\theta\right) = d * X^\flat$.

3) Since $\boldsymbol{\delta} X^\flat = (-1)^{1+\mathrm{ind}(\mathbf{g})} * d * X^\flat$, we have $*\boldsymbol{\delta} X^\flat = (-1)^{1+\mathrm{ind}(\mathbf{g})} * *d * X^\flat$, where $d{*}X^\flat$ is a scalar field, so (by Proposition 5.36(iii)) $**d{*}X^\flat = (-1)^{\mathrm{ind}(\mathbf{g})}\, d{*}X^\flat$, and hence $*\boldsymbol{\delta} X^\flat = -d * X^\flat$. By Proposition 5.36(ii), $*\boldsymbol{\delta} X^\flat = \boldsymbol{\delta} X^\flat * 1 = \boldsymbol{\delta} X^\flat.\theta$, which finally gives that $\mathcal{L}_X.\theta = -\boldsymbol{\delta} X^\flat.\theta$. The equality $\mathcal{L}_X.\theta = \mathfrak{D}.\theta$ is satisfied if and only if $\mathfrak{D} = -\,\boldsymbol{\delta} X^\flat$.

4) By ([4.42], section 4.5), $\theta = \sqrt{|g|}.d\xi^1 \wedge \ldots \wedge d\xi^n$, so ([4.16], section 4.2.5) $i_X.\theta = \sqrt{|g|}\sum_k (-1)^k X^k.d\xi^1 \wedge \ldots \wedge \widehat{d\xi^k} \wedge \ldots \wedge d\xi^n$. Hence:

$$di_X.\theta = \sum_k \left(\frac{\partial}{\partial \xi^k}\left(\sqrt{|g|}X^k\right)\right) d\xi^1 \wedge \ldots \wedge d\xi^n = \frac{1}{\sqrt{|g|}}\sum_k \left(\frac{\partial}{\partial \xi^k}\left(\sqrt{|g|}X^k\right)\right).\theta.$$

■

EXAMPLE 5.43.– *Let us find the divergence in spherical coordinates. Consider the vector field* $\overrightarrow{X} = A.\overrightarrow{u_r} + B.\overrightarrow{u_\phi} + C.\overrightarrow{u_\theta}$. *By* (4.29), $\overrightarrow{X} = A.\frac{\partial}{\partial r} + \frac{B}{r}.\frac{\partial}{\partial \phi} + \frac{C}{r\sin\phi}.\frac{\partial}{\partial \theta}$ *and, by [5.41],* $g = r^4 \sin^2\phi$. *Therefore:*

$$div_{\mathbf{g}}\left(X\right) = \frac{2A}{r} + \frac{\partial A}{\partial r} + \frac{B}{r\tan\phi} + \frac{1}{r}\frac{\partial B}{\partial \phi} + \frac{1}{r\sin\phi}\frac{\partial C}{\partial \theta}. \qquad [5.45]$$

The intrinsic definition of the Laplacian can be established as follows:

COROLLARY-DEFINITION 5.44.– *The* Beltrami Laplacian $\nabla^2 : \Omega^0(B) \to \Omega^0(B)$ *is given by:*

$$\nabla^2 = div \circ grad = -\Delta\left|_{\Omega^0(B)}\right.,$$

where $\Delta : \Omega^k(B) \to \Omega^k(B)$ $(k \geq 0)$ *is the* de Rham Laplacian

$$\boxed{\Delta := d\boldsymbol{\delta} + \boldsymbol{\delta} d = (d + \boldsymbol{\delta})^2.}$$

Given positively-oriented local coordinates,

$$\nabla^2 f = \frac{1}{\sqrt{|g|}}\sum_{k,l}\frac{\partial}{\partial x^k}\left(g^{lk}.\sqrt{|g|}.\frac{\partial f}{\partial x^l}\right). \qquad [5.46]$$

PROOF.– 1) If f is a scalar field, $\boldsymbol{\delta} f = 0$ by Definition 5.40, so

$$\left(d\boldsymbol{\delta} + \boldsymbol{\delta} d\right) f = \boldsymbol{\delta} df = \boldsymbol{\delta}\left(\mathrm{grad} f\right)^\flat = -\mathrm{div} \circ \mathrm{grad}\left(f\right).$$

2) The expression [5.46] follows from [5.44] and [5.39]. ■

REMARK 5.45.– *The de Rham Laplacian is symmetric relative to the bilinear form* $\langle.|.\rangle_2$*, whereas the (more common) Beltrami Laplacian is antisymmetric; this is the motivation behind the definition of the de Rham Laplacian. Indeed, let* $\alpha, \beta \in L_c^2\left(Alt_k\left(B\right)\right)$. *Then:*

$$\begin{aligned}\langle \Delta\alpha|\beta\rangle_2 &= \langle d\boldsymbol{\delta}\alpha|\beta\rangle_2 + \langle \boldsymbol{\delta} d\alpha|\beta\rangle_2 = \langle \boldsymbol{\delta}\alpha|\boldsymbol{\delta}\beta\rangle_2 + \langle d\alpha|d\beta\rangle_2 \\ &= \langle \alpha|d\boldsymbol{\delta}\beta\rangle_2 + \langle \alpha|\boldsymbol{\delta} d\beta\rangle_2 = \langle \alpha|\Delta\beta\rangle_2 .\end{aligned}$$

EXAMPLE 5.46.– *Let us find the Laplacian in spherical coordinates. We can simply apply the formula [5.45] for the divergence with* A, B, C *as the components of the gradient [5.42]. This gives:*

$$-\Delta f = \nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\tan\phi}\frac{\partial f}{\partial \phi} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{1}{r^2\sin^2\phi}\frac{\partial^2 f}{\partial \theta^2}.$$

For more details about Hodge duality on an Einstein manifold (section 4.5.1), see [MIS 73], section 3.5.

5.6.4. *Gauss' theorem and Poisson's formula*

Consider a Riemannian submanifold $\overset{\frown}{B}$ of the Euclidean space $\mathbb{R}^3$ equipped with its canonical orientation. On this manifold, suppose that the density $\rho(P) \geq 0$ is defined at every point P and assume that the function $\rho : P \mapsto \rho(P)$ is continuous. In physics, ρ might represent the density of matter or an electric charge. Suppose that an infinitely small volume δV around P generates a field $\delta\overrightarrow{E} = -K.\rho.\delta V.\frac{\overrightarrow{r}}{r^3} = -\frac{K.\rho.\delta V}{r^2}.\overrightarrow{u_r}$ at every point P', where $\overrightarrow{r} = \overrightarrow{PP'}$, $r = |\overrightarrow{r}|$ and $\overrightarrow{u_r} = \frac{\overrightarrow{r}}{r}$. For a gravitational field, $K = G$, where G is the universal constant of gravitation (Newton's law); for an electric field, $K = -\frac{1}{4\pi\varepsilon_0}$, where ε_0 is the permittivity of vacuum (Coulomb's law).

Let $\overset{\frown}{V}$ be a chain of closed polyhedra in $\overset{\frown}{B}$ (section 4.4.7**(III)**, Example 4.60) and $\overset{\frown}{S} = \partial\overset{\frown}{V}$ its boundary, equipped with the induced orientation. Consider a surface element dS. The flux of $\delta\overrightarrow{E}$ through the small oriented surface $\overset{\frown}{dS}$ is $d\delta\Phi = \left\langle \delta\overrightarrow{E}|\overrightarrow{n}\right\rangle.dS$, where $\overrightarrow{n} = \mathbf{n}$ is the outward normal unit vector field of V (section 5.6.2**(I)**), so let $d\delta\Phi = -K.\rho.dV.\frac{\langle\overrightarrow{r}|\overrightarrow{n}\rangle}{r^3}.dS$. But $\frac{\langle\overrightarrow{r}|\overrightarrow{n}\rangle}{r^3}.dS$ is by definition the solid angle[11] $d\Omega$ subtended by the surface element dS from the point P whenever

11 See the Wikipedia article on the *Solid angle*.

the latter is inside $\overset{\frown}{V}$. Hence, $d\delta\Phi = -K.\rho.\delta V.d\Omega$, and so the flux of $\delta\overrightarrow{E}$ through $\overset{\frown}{S}$ is

$$\delta\Phi = -\int_{\overset{\frown}{S}} K\rho.\delta V.d\Omega = -K\rho.\delta V.\Omega = -4\pi K\rho.\delta V$$

because $\Omega = 4\pi$. Thus, the flux through $\overset{\frown}{S}$ of the field created by the entire volume V is $\Phi = -\int_{\overset{\frown}{V}} 4\pi K\rho.dV$. For gravitation, the integral $M = \int_{\overset{\frown}{V}} \rho.dV$ is interpreted as the total mass within the volume V; in electrostatics, it is interpreted as the total electric charge. We therefore have:

$$\boxed{\Phi = -4\pi KM} \qquad [5.47]$$

(Gauss' theorem). Ostrogradksy's theorem (Theorem 5.28) implies that $\Phi = \int_{\overset{\frown}{V}} \text{div}\left(\overrightarrow{E}\right).dV$. Since $\overset{\frown}{V}$ is arbitrary, we can take it to be infinitely small around P. With the assumption that $\overrightarrow{E}$ is of class C^1, we deduce that $\text{div}\overrightarrow{E} = -4\pi K\rho$. Now suppose that $\overrightarrow{E}$ derives from the potential $U : \overrightarrow{E} = -\text{grad}U$. From [5.29], we obtain *Poisson's formula*:

$$\boxed{\overrightarrow{\nabla}^2 U = 4\pi K\rho.} \qquad [5.48]$$

Conversely, suppose that Poisson's formula [5.48] is given and consider a point mass or charge M located at the point P. The rest of the space is assumed to be empty. Let P' be another point, referenced by its spherical coordinates relative to the origin P. By symmetry, the potential U only depends on $r = \left|\overrightarrow{PP'}\right|$. Poisson's equation $\nabla^2 U = 0$ therefore simplifies to $\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial U}{\partial r}\right) = 0$, so $r^2\frac{\partial U}{\partial r} = C$ and $U = -\frac{C}{r}$ for $r \neq 0$, where C is a constant. By [5.42], this gives a field $\overrightarrow{E} = -\nabla U = -\frac{C}{r^2}\overrightarrow{u_r}$ whose flux through the positively oriented sphere $\overset{\frown}{V}$ of center P and radius $r > 0$ is $\Phi = -\frac{C}{r^2}.4\pi r^2 = -4\pi C$. This flux is also given by [5.47], so $C = -KM$, where $M = \int_{\overset{\frown}{V}} \rho.dV$. Since $r > 0$ is arbitrary, $\rho = M.\delta$, where δ is the Dirac distribution (this is the origin of the term "distribution"). Finally,

$$\boxed{\overrightarrow{E} = -\frac{KM}{r^2}\overrightarrow{u_r}}$$

(Newton's law when M is a mass, Coulomb's law when M is an electric charge).

5.6.5. *Homology, cohomology and duality*

The results from algebraic topology presented below complement the results already outlined in [P1], section 3.3.8, and illustrate Stokes' theorem[12]. In this section, $\overset{\frown}{B}$ is an oriented, pure, m-dimensional differential manifold that is countably at infinity and of class C^{∞}. This determines an orientable manifold B.

(I) HOMOLOGY AND COHOMOLOGY SPACES In section 5.2.1, we saw that the space $\underline{\Omega_c^p}(B)$ of compactly supported odd differential p-forms on B can be equipped with the structure of a strict inductive limit of Fréchet spaces and a Montel nuclear space (if B is compact, $\underline{\Omega_c^p}(B) = \underline{\Omega^p}(B)$ is a Fréchet nuclear space, in which case, the index c can be omitted). The exterior differential $d : \underline{\Omega_c^{p-1}}(B) \to \underline{\Omega_c^p}(B)$ $(p \geq 1)$ is a continuous linear mapping (this is a straightforward generalization of [P2], section 4.3.1**(III)**, Theorem 4.73). We have $d^2 = 0$ (section 5.5.1**(I)**), so d is a coboundary ([P1], section 3.3.8**(VII)**) and its transpose ([P2], section 3.5.3) is a continuous linear mapping ∂ (that can be written as ∂_p wherever necessary) from the space $\underline{\Omega_c^p}(B)^{\vee}$ of even p-currents into the space $\underline{\Omega_c^{p-1}}(B)^{\vee}$ of even $(p-1)$-currents[13].

DEFINITION 5.47.– *The continuous linear mapping* $\partial_p : \underline{\Omega_c^p}(B)^{\vee} \to \underline{\Omega_c^{p-1}}(B)^{\vee}$ *is called the* boundary *mapping.*

By Lemma 5.20(i), d is an anti-derivation of degree 1 in the graduated algebra $\underline{\Omega_c}(B)$. Set:

$$\mathbf{Z}_c^p(B;\mathbb{R}) = \ker\left(d : \underline{\Omega_c^p}(B) \to \underline{\Omega_c^{p+1}}(B)\right),$$

$$\mathbf{B}_c^p(B;\mathbb{R}) = \operatorname{im}\left(d : \underline{\Omega_c^{p-1}}(B) \to \underline{\Omega_c^p}(B)\right),$$

$$\mathbf{H}_c^p(B;\mathbb{R}) = \mathbf{Z}_c^p(B;\mathbb{R}) / \mathbf{B}_c^p(B;\mathbb{R});$$

then $\mathbf{H}_c^p(B;\mathbb{R})$ is the p-th compactly supported de Rham cohomology space ([P1], section 3.3.8**(VII)**). The elements of $\mathbf{Z}_c^p(B;\mathbb{R})$ are the compactly supported p-cocycles (closed odd differential p-forms), the elements of $\mathbf{B}_c^p(B;\mathbb{R})$ are the compactly supported p-coboundaries (exact odd differential p-forms) and every coboundary is a cocycle. Any two closed odd differential forms $\omega, \omega' \in \mathbf{Z}_c^p(B;\mathbb{R})$ with the same canonical image in $\mathbf{H}_c^p(B;\mathbb{R})$ are said to be *cohomologous*.

---

12 More details on these questions of *Analysis Situs*, admittedly lengthy reading but nonetheless highly enjoyable to browse, can be found on the website [DES 19] (in French), written by a group of mathematicians and completely open access (featuring the elegant illustrations of the Poincaré duality theorem originally published in [POP 12]).

13 The reader is invited to reformulate the next results for the case where B is oriented, the differential p-forms are even and the p-currents are odd.

Dually, we have $\partial^2 = {}^t d^2 = 0$, so the boundary mapping $\partial_\bullet = (\partial_p)_{p\geq 1}$ is an anti-derivation of degree -1 in the graduated vector space

$$\underline{\Omega_c}^\vee := \bigoplus_{p\geq 0} \underline{\Omega_c^p}^\vee$$

that is also a codifferential ([P1], section 3.3.8(**II**)). Set (*ibid.*):

$$\mathbf{Z}'_p(B) = \ker(\partial_p), \mathbf{B}'_p(B) = \operatorname{im}(\partial_{p+1}),$$
$$\mathbf{H}'_p(B) = \mathbf{Z}'_p(B)/\mathbf{B}'_p(B).$$

We say that $T \in \underline{\Omega_c^p}(B)^\vee$ is a *p-cycle* (respectively a *p-boundary*) if $T \in \mathbf{Z}'_p(B)$ (respectively $T \in \mathbf{B}'_p(B)$). Any two p-currents are said to be *homologous* if their difference is a boundary. We also say that an element $T \in \mathbf{Z}'_p(B)$ is a *closed* p-current; thus, any two closed p-currents are homologous if and only if they have the same canonical image in $\mathbf{H}'_p(B)$.

DEFINITION 5.48.– *The space* $\mathbf{H}'_p(B)$ *is called the* p*-th homology space of* B *and* $\mathbf{H}'_\bullet(B) = (\mathbf{H}'_p(B))_{p\geq 0}$ *is called the (total) homology space of* B *(for currents).*

By section 5.2.1(**IV**), we know that an even p-chain uniquely determines an even p-current T. Conversely, a closed even p-current uniquely determines an even p-chain τ by the relation [5.1] ([DER 84], Chapter 4, section 23, Theorem 18; [DIE 93], Volume 9, (24.26.4)). This gives us the result:

THEOREM 5.49.– (de Rham) *The homology spaces* $\mathbf{H}'_p(B)$ *of closed even currents coincide with the real homology spaces* $\mathbf{H}_p(B;\mathbb{R}) = \mathbf{H}_p(B) \otimes_{\mathbb{Z}} \mathbb{R}$ *of complexes of even chains of smooth simplexes (section 4.4.7(**I**)) on* B *([P1], section 3.3.8(**VI**)).*

REMARK 5.50.– *This theorem begins by proving the case where* B *is a chain of smooth simplexes, i.e. when* B *admits a triangulation. But a theorem by Whitehead from 1940 shows that every differential manifold is triangulable ([DER 84], Chapter 4, section 23, footnote 8). The homology space* $\mathbf{H}_p(B) \otimes_{\mathbb{Z}} \mathbb{R}$ *is sometimes written as* $\mathbf{H}_p^\infty(B;\mathbb{R})$ *to emphasize that we are considering* smooth *simplexes ([DIE 93], Volume 9, (24.20.6)).*

(II) MAYER–VIETORIS THEOREM The Mayer–Vietoris theorem is an effective way to calculate the homology of a manifold. Set $\mathbf{H}_p(\varnothing;\mathbb{R}) = 0$ for every $p \geq 0$.

THEOREM 5.51.– (Mayer–Vietoris) *Consider two non-empty open sets U, V that give an open covering of B. We have the following long exact sequences, called* Mayer–Vietoris *sequences: the first [5.49] is a homology sequence ending with $\mathbf{H}_0(B;\mathbb{R}) \to 0$, the second [5.50] is a cohomology sequence ending with $\mathbf{H}^m(B;\mathbb{R})$:*

$$... \longrightarrow \mathbf{H}_p(U \cap V;\mathbb{R}) \longrightarrow \mathbf{H}_p(U;\mathbb{R}) \oplus \mathbf{H}_p(V;\mathbb{R}) \longrightarrow \mathbf{H}_p(B;\mathbb{R})$$

$$\xrightarrow{\delta} \mathbf{H}_{p-1}(U \cap V;\mathbb{R}) \longrightarrow ... \qquad [5.49]$$

$$\longrightarrow \mathbf{H}^{p-1}(U;\mathbb{R}) \oplus \mathbf{H}^{p-1}(V;\mathbb{R}) \longrightarrow \mathbf{H}^{p-1}(U \cap V;\mathbb{R}) \xrightarrow{\delta^*} \mathbf{H}^p(B;\mathbb{R})$$

$$\longrightarrow \mathbf{H}^p(U;\mathbb{R}) \oplus \mathbf{H}^p(V;\mathbb{R}) \longrightarrow ... \qquad [5.50]$$

PROOF.– Consider the following commutative diagram and canonical injections:

$$\begin{array}{ccccc} & & U \cap V & & \\ & {}^{i_U}\swarrow & & \searrow^{i_V} & \\ U & & & & V \\ & {}_{j_U}\searrow & & \swarrow_{j_V} & \\ & & U \cup V & & \end{array}$$

The Mayer–Vietoris sequence [5.49] is the long exact homology sequence ([P1], section 3.3.8**(III)**) deduced from the short sequence

$$0 \longrightarrow \underline{\Omega}(U \cup V) \overset{{}^t j_U \oplus {}^t j_V}{\longrightarrow} \underline{\Omega}(U) \oplus \underline{\Omega}(V) \overset{{}^t i_U - {}^t i_V}{\longrightarrow} \underline{\Omega}(U \cap V) \longrightarrow 0. \qquad [5.51]$$

The result therefore hinges upon showing that the latter sequence is exact. For ${}^t j_U \oplus {}^t j_V$ to be injective, a differential form must be zero on $U \cup V$ if and only if it is zero on U and V, which is clear. For exactness at $\underline{\Omega}(U) \oplus \underline{\Omega}(V)$, a pair $(\alpha, \beta) \in \underline{\Omega}(U) \oplus \underline{\Omega}(V)$ must originate from a form on $U \cup V$ if and only if α and β coincide on $U \cap V$, which is also clear. It remains to be shown that ${}^t i_U - {}^t i_V$ is surjective, i.e. an odd form $\underline{\omega}$ on $U \cap V$ is the difference of an odd form $\underline{\alpha}$ that can be extended to U and an odd form $\underline{\beta}$ that can be extended to V. To do this, we can simply use a partition of unity of class C^∞ subordinate to the covering $\{U, V\}$; such a partition of unity must exist by Corollary 2.16.

By changing p to $-p$, we can also deduce the long cohomology sequence [5.50] from the short exact sequence [5.51]. ■

(III) POINCARÉ DUALITY THEOREM

LEMMA 5.52.– (Poincaré theorem for compactly supported sets) *If the oriented manifold* $\overset{\frown}{B}$ *is the space* $\mathbb{R}^m$ *equipped with its canonical orientation*[14]*, the mapping* $\int_{\overset{\frown}{B}} . : \mathbf{Z}_c^m(B;\mathbb{R}) \to \mathbb{R} : \omega \to \int_{\overset{\frown}{B}} \omega$ *induces an isomorphism* $\mathbf{H}_c^m(B;\mathbb{R}) \cong \mathbb{R}$*, and* $\mathbf{H}_c^p(B;\mathbb{R}) = \{0\}$ *if* $p \neq m$.

PROOF.– It is clear that $\int_{\overset{\frown}{B}} .$ is surjective. Since $\mathbb{R}^m$ is star-shaped, we have the following exact sequence, which can be proved in the same way as the version of Poincaré's lemma established earlier ([P1], section 3.3.8**(VII)**, Theorem 3.184) (**exercise***: see [MAD 97], Lemma 10.15):

$$\Omega_c^{m-1} \xrightarrow{d} \Omega_c^m \xrightarrow{\int_{\overset{\frown}{B}} .} \mathbb{R} \longrightarrow 0$$

and $\Omega_c^m = \mathbf{Z}_c^m(B;\mathbb{R})$, $\mathbf{B}_c^m(B;\mathbb{R}) := \operatorname{im}(d) = \ker\left(\int_{\overset{\frown}{B}} .\right)$, so the following diagram commutes:

$$\begin{array}{ccc} \mathbf{Z}_c^m(B;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ \downarrow & \nearrow\cong & \\ \mathbf{Z}_c^m(B;\mathbb{R})/\mathbf{B}_c^m(B;\mathbb{R}) & & \end{array}$$

which implies that $\mathbf{H}_c^m(B;\mathbb{R}) \cong \mathbb{R}$. We have $\mathbf{H}_c^0(B;\mathbb{R}) = \mathbf{Z}_c^0(B;\mathbb{R}) = \{f \in \mathcal{D}(B) : df = 0\} = \{0\}$ and similarly for $\mathbf{H}_c^p(B;\mathbb{R})$ if $0 < p < m$; this last property can be shown inductively using the relation $\mathbf{H}_c^p(\mathbb{R}^k) \cong \mathbf{H}_c^{p+1}(\mathbb{R}^{k+1})$ (which follows from an integral and holds for every integer $p, k \geq 0$) ([MAD 97], Lemma 13.2). ■

EXAMPLE 5.53.– *The vector or scalar fields and p-forms of class* C^∞ *considered below are assumed to be* compactly supported *in* $\overset{\frown}{B}$*, which is taken to be the Euclidean space* $\mathbb{R}^3$ *equipped with its canonical orientation.*

i) $\mathbf{H}_c^0(B;\mathbb{R}) = 0$ *means that the only zero function is the function* f *such that* $df = 0$ *(since* f *is compactly supported).*

ii) $\mathbf{H}_c^1(B;\mathbb{R}) = 0$ *means that a vector field* $\overrightarrow{E}$ *is of the form* $-\nabla U$ *if and only if it has zero curl. If so,* U *is uniquely determined (because it is compactly supported).*

---

14 The framework chosen here is the hypothesis **($\mathbf{H_1}$)** from section 5.6.1. This avoids needing to consider the bundle of scalars of odd type in the proofs and is also more common in physics. Nevertheless, the results still hold with the hypothesis **($\mathbf{H_2}$)**. Despite the change of setting, we will keep the same notation for cycles, cocycles, boundaries, coboundaries, homology spaces and cohomology spaces (see footnote 13, p. 215).

Indeed, let $\omega = E_x.dx + E_y.dy + E_z.dz$ *be the even differential 1-form determined by* $\overrightarrow{E}$*. Then,* $d\omega = \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) dx \wedge dy + ...$*, so* $\overrightarrow{E}$ *is of the form* $-\nabla U$ *if and only if* $\omega \in \mathbf{B}_c^1(B;\mathbb{R})$*, and* $\omega \in \mathbf{Z}_c^1(B;\mathbb{R})$ *if and only if* $d\omega = 0$*, i.e. curl* $\left(\overrightarrow{E}\right) = 0$*. Furthermore,* U *is uniquely determined by (i).*

iii) $\mathbf{H}_c^2(B;\mathbb{R}) = 0$ *means that a vector field* $\overset{\hookrightarrow}{B}$ *is of the form curl*$\overrightarrow{A}$ *if and only if div*$\overset{\hookrightarrow}{B} = 0$*; if so,* $\overrightarrow{A}$ *is uniquely determined up to addition by a term of the form* $-\nabla U$*. Indeed, let* $\omega = B_x.d_y \wedge dz + B_y.dz \wedge dx + B_z.dx \wedge dy$ *be the even differential 2-form determined by* $\overset{\hookrightarrow}{B}$*. Then,* $d\omega =$ *div* $\left(\overset{\hookrightarrow}{B}\right).dx \wedge dy \wedge dz$*. Hence,* $\overset{\hookrightarrow}{B}$ *is of the form curl*$\overrightarrow{A}$ *if and only if* $\omega \in \mathbf{B}_c^2(B;\mathbb{R})$*, and* $\omega \in \mathbf{Z}_c^2(B;\mathbb{R})$ *if and only if div*$\overset{\hookrightarrow}{B} = 0$*. Furthermore,* $\ker\left(d : \left(\Omega_c^1 B\right) \to \Omega_c^2(B)\right) = \mathbf{Z}_c^1(B;\mathbb{R}) = \mathbf{B}_c^1(B;\mathbb{R})$ *can be identified with the fields* $\overrightarrow{E}$ *of the form* $-\nabla U$ *by (ii).*

iv) $\mathbf{H}_c^3(B;\mathbb{R}) = \mathbb{R}$ *means that a scalar field* Φ *is of the form div*$\overrightarrow{B}$ *if and only if* $\int_{\overset{\frown}{B}} \Phi.dx \wedge dy \wedge dz = 0$*. If so,* $\overrightarrow{B}$ *is uniquely determined up to addition by a term of the form curl*$\overrightarrow{A}$*. Indeed, every even 3-form* $\omega \in \Omega_c^3(B)$ *is of the form* $\Phi.dx \wedge dy \wedge dz$*, every odd 2-form* $\alpha \in \Omega_c^2(B)$ *is of the form* $B_x.dy \wedge dz + B_y.dz \wedge dx + B_z.dx \wedge dy$*, and im* $\left\{d : \Omega_c^2(B) \to \Omega_c^3(B)\right\} = \left\{\omega \in \Omega_c^3(B) : \int_{\overset{\frown}{B}} \omega = 0\right\}$*. But* $d\alpha =$ *div* $\left(\overrightarrow{B}\right).dx \wedge dy \wedge dz$ *and* $\ker\left(d : \Omega_c^2(B) \to \Omega_c^3(B)\right)$ *is the space of even 2-forms* $\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dx \wedge dy + ...$

DEFINITION 5.54.– *The manifold* B *is said to be* of finite type *if there exists a covering (said to be admissible) of* B *by finitely many open sets* $U_1, ..., U_k$ *such that every non-empty intersection* $U_{i_1} \cap ... \cap U_{i_l}$ *is isomorphic to* $\mathbb{R}^m$ $(m = \dim(B))$*. The smallest cardinal of an admissible open covering is said to be the* minimal type *of* B*.*

If B is compact, it is of finite type by the Borel–Lebesgue property ([P2], section 2.3.7) (by extracting a finite subcovering from a covering by charts).

COROLLARY 5.55.– *The homology and cohomology spaces of a manifold* B *of finite type are finite-dimensional.*

PROOF.– We will give the reasoning for the cohomology spaces. Proceed by induction on the minimal type k of the manifold. The claim is true for $k = 1$ by Lemma 5.52.

If B is of finite type but is not diffeomorphic to $\mathbb{R}^m$ $(m = \dim(B))$, there exists a covering of B by two open sets U, V such that U, V and $U \cap V$ have minimal type strictly lower than B (**exercise**). The long exact sequence [5.50] gives the exact sequence

$$\mathbf{H}^{p-1}(U \cap V; \mathbb{R}) \xrightarrow{\partial^*} \mathbf{H}^p(B; \mathbb{R}) \longrightarrow \mathbf{H}^p(U; \mathbb{R}) \oplus \mathbf{H}^p(V; \mathbb{R}),$$

where $\mathbf{H}^{p-1}(U \cap V; \mathbb{R})$ and $\mathbf{H}^p(U; \mathbb{R}) \oplus \mathbf{H}^p(V; \mathbb{R})$ are finite-dimensional, so $\mathbf{H}^p(B; \mathbb{R})$ is finite-dimensional. ■

The proof of the Poincaré duality theorem is based on the five lemma, which is proved by "diagram chasing" in the same way as the nine lemma ([P1], section 3.3.7(**IV**), Lemma 3.167) or by applying the snake lemma ([P1], section 3.3.7(**IV**), Lemma 3.169) (**exercise**):

LEMMA 5.56.– (five lemma) *In an abelian category, consider the commutative diagram:*

$$\begin{array}{ccccccccc} A_1 & \longrightarrow & A_2 & \longrightarrow & A_3 & \longrightarrow & A_4 & \longrightarrow & A_5 \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ B_1 & \longrightarrow & B_2 & \longrightarrow & B_3 & \longrightarrow & B_4 & \longrightarrow & B_5 \end{array}$$

where the two rows are exact and each vertical morphism $A_i \to B_i$ is written as f_i $(i = 1, ..., 5)$. If f_1 is an epimorphism and f_2, f_4 are monomorphisms, then f_3 is a monomorphism; if f_5 is a monomorphism and f_2, f_4 are epimorphisms, then f_3 is an epimorphism.

THEOREM 5.57.– (Poincaré duality) *The linear mapping $\Lambda : \mathbf{Z}^p(B; \mathbb{R}) \longrightarrow \mathbf{Z}_c^{m-p}(B; \mathbb{R})^*$ defined by*

$$\langle \Lambda.\alpha, \beta \rangle = \int_{\overset{\frown}{B}} \alpha \wedge \beta$$

($\alpha \in \mathbf{Z}^p(B; \mathbb{R}), \beta \in \mathbf{Z}_c^{m-p}(B; \mathbb{R})$) induces an isomorphism $\mathbf{H}^p(B; \mathbb{R}) \cong \mathbf{H}_c^{m-p}(B; \mathbb{R})^*$[15].

PROOF.– 1) The quantity $\langle \Lambda.\alpha, \beta \rangle$, called the "cup product" $\alpha \bigcup \beta$ of α, β, only depends on the canonical images $[\alpha], [\beta]$ in the cohomology spaces $\mathbf{H}^p(B; \mathbb{R})$ and $\mathbf{H}_c^{m-p}(B; \mathbb{R})$, respectively. Indeed, if $\alpha \in \mathbf{B}^p(B; \mathbb{R}), \beta \in \mathbf{Z}_c^{m-p}(B; \mathbb{R})$, then, by [5.22] and the relation $d^2 = 0$, $\alpha \wedge \beta = (-1)^p d\beta$ and $\int_{\overset{\frown}{B}} \alpha \wedge \beta = (-1)^p \int_{\overset{\frown}{B}} d\beta = 0$ by Stokes' theorem, since $\partial \overset{\frown}{B} = \varnothing$. Similarly, if $\alpha \in \mathbf{Z}^p(B; \mathbb{R}), \beta \in \mathbf{B}_c^{m-p}(B; \mathbb{R})$, then $\alpha \wedge \beta = 0$.

15 Here, $(.)^*$ denotes the algebraic dual of $(.)$.

2) We will prove the isomorphism by induction on the minimal type k of B. For $k=1$, $\mathbf{H}^0(B;\mathbb{R})=\mathbf{Z}^0(B;\mathbb{R})=\ker\left(d:\Omega^0(B)\to\Omega^1(B)\right)=\{f\in\mathcal{E}(B):df=0\}\cong\mathbb{R}$, since B is connected. Thus, $\mathbf{H}^p(B;\mathbb{R})\cong\mathbf{H}_c^{m-p}(B;\mathbb{R})^*$ for $p=0$ by Lemma 5.52.

Consider now the exact Mayer–Vietoris cohomology sequence [5.50], where U and V are the open sets used earlier in the proof of Corollary 5.55, and the transposed exact sequence, since any exact sequence $E_1\longrightarrow E_2\longrightarrow E_3$ of finite-dimensional yields the transposed exact sequence $E_1^*\longleftarrow E_2^*\longleftarrow E_3^*$ ([P2], section 3.8.1, Corollary 3.126). This gives a diagram identical to the diagram in the five lemma, with:

$$f_1:\mathbf{H}^{p-1}(U;\mathbb{R})\oplus\mathbf{H}^{p-1}(V;\mathbb{R})\tilde{\to}\mathbf{H}_c^{m-(p-1)}(U;\mathbb{R})^*\oplus\mathbf{H}_c^{m-(p-1)}(V;\mathbb{R})^*$$
$$f_2:\mathbf{H}^{p-1}(U\cap V;\mathbb{R})\tilde{\to}\mathbf{H}_c^{m-(p-1)}(U\cap V;\mathbb{R})^*$$
$$f_3:\mathbf{H}^{p}(B;\mathbb{R})\to\mathbf{H}_c^{m-p}(B;\mathbb{R})^*$$
$$f_4:\mathbf{H}^{p}(U;\mathbb{R})\oplus\mathbf{H}^{p}(V;\mathbb{R})\tilde{\to}\mathbf{H}_c^{m-p}(U;\mathbb{R})^*\oplus\mathbf{H}_c^{m-p}(V;\mathbb{R})^*$$
$$f_5:\mathbf{H}^{p}(U\cap V;\mathbb{R})\tilde{\to}\mathbf{H}_c^{m-p}(U\cap V;\mathbb{R})^*$$

The mappings f_i are isomorphisms for $i\neq 3$ by the hypotheses. Hence, the five lemma implies that f_3 is also an isomorphism. ■

If the manifold B is compact, the integer $b_p(B):=\dim(\mathbf{H}^p(B;\mathbb{R}))$ is called the p-th *Betti number* of B.

COROLLARY 5.58.– (Poincaré)[16] *For all $p\geq 0$, $b_p(B)=b_{m-p}(B)$ (**exercise**).*

(IV) DE RHAM DUALITY THEOREM A boundary is a cycle, but the converse is not true in general. As an illustration, assume the hypothesis $(\mathbf{H}_2)$ from the beginning of section 5.6.1 (see footnote 13, p. 215).

LEMMA 5.59.– *i) The integral of a p-cocycle over a p-boundary is zero.*

ii) The integral of a p-coboundary over a p-cycle is zero.

16 This is one of the key results of the *Analysis Situs* published by Poincaré in 1895 [POI 00a], but a similar result had already been formulated earlier by É. Picard in 1889. Poincaré himself had cited this result in 1893 (see [POP 12] and [DIE 89] for more details). For an intuitive interpretation of the Betti numbers and their relation with the *genus* of a surface, see [POP 16] and the Wikipedia article on *Betti numbers*.

PROOF.– i) Let $\underline{\omega} \in \Omega_c^p(B)$ be a compactly supported p-cocycle and $\partial\tau$ a p-boundary, i.e. the boundary of a $(p+1)$-chain of simplexes τ. By Stokes' theorem,

$$\int_{\partial\tau} \underline{\omega} = \int_{\tau} d\underline{\omega} = 0.$$

ii) If $\underline{\omega} = d\underline{\alpha}$ and τ is a cycle, then, by Stokes' theorem:

$$\int_{\tau} \underline{\omega} = \int_{\tau} d\underline{\alpha} = \int_{\partial\tau} \underline{\alpha} = 0,$$

since $\partial\tau = \varnothing$ ([P1], section 3.3.8(**I**)). ■

REMARK 5.60.– *The integral of a p-coboundary over a p-boundary is zero by Lemma 4.62. However, the integral of a p-cocycle over a p-cycle is not necessarily zero.*

Consider the bilinear form:

$$B : \mathbf{Z}_p(B;\mathbb{R}) \times \mathbf{Z}_c^p(B;\mathbb{R}) : (\tau, \underline{\omega}) \longmapsto \int_{\tau} \underline{\omega}.$$

Lemma 5.59 implies the following result.

COROLLARY 5.61.– *The bilinear form B induces a bilinear form*

$$\bar{B} : \mathbf{H}_p(B;\mathbb{R}) \times \mathbf{H}_c^p(B;\mathbb{R}) : ([\tau], [\underline{\omega}]) \longmapsto \int_{\tau} \underline{\omega},$$

where $[\tau]$ (respectively $[\underline{\omega}]$) is the canonical image of $\tau \in \mathbf{Z}_p(B;\mathbb{R})$ (respectively $\underline{\omega} \in \mathbf{Z}_c^p(B;\mathbb{R})$) in $\mathbf{H}_p(B;\mathbb{R})$ (respectively $\mathbf{H}_c^p(B;\mathbb{R})$).

The next theorem was originally formulated as a conjecture by É. Cartan in 1928 ([DIE 89], section I.3, p. 63). It is a very deep result from algebraic topology established by de Rham in 1931; the proof (as well as the proof of Theorem 5.49) exceeds the scope of this book ([DER 84], Chapter 4, Theorem 17; [DIE 93], (21.11.3); [SCH 66], Chapter 9, section 3, Theorem I):

THEOREM 5.62.– (de Rham) *Let B be a manifold of finite type. The bilinear form $\bar{B}$ is non-degenerate (both on the left and the right), which implies the isomorphism*

$$\boxed{\mathbf{H}_c^p(B;\mathbb{R}) \cong \mathbf{H}_p(B;\mathbb{R})^*.}$$

REMARK 5.63.– *If B is not of finite type but is locally compact, m-dimensional, and countable at infinity, then $B = \bigcup_{j\geq 1} K_j$, where $(K_j)_{j\geq 1}$ is a sequence of compact submanifolds such that $K_j \Subset K_{j+1}$. Thus, $\mathbf{H}_c^p(B;\mathbb{R}) = \varinjlim_j \mathbf{H}^p(K_j;\mathbb{R})$ and*

$\mathbf{H}_p(B;\mathbb{R})^* = \varinjlim_j \mathbf{H}_p(K_j;\mathbb{R})^*$ *([HAT 02], section 3.3, Proposition 3.33); the isomorphisms* $\mathbf{H}^p(K_j) \cong \mathbf{H}_c^{m-p}(K_j)^*$ *(Theorem 5.57) and* $\mathbf{H}^p(K_j;\mathbb{R}) \cong \mathbf{H}_p(K_j;\mathbb{R})^*$ *(Theorem 5.62) imply the isomorphisms of vector spaces* $\mathbf{H}^p(B) \cong \mathbf{H}_c^{m-p}(B)^*$ *and* $\mathbf{H}_c^p(B;\mathbb{R}) \cong \mathbf{H}_p(B;\mathbb{R})^*$.

(V) EXAMPLES

(i) If B is a connected Hausdorff manifold, then $\mathbf{H}_c^0(B;\mathbb{R}) \cong \mathbb{R}$ if B is compact and $\mathbf{H}_c^0(B;\mathbb{R}) = \{0\}$ otherwise.

Indeed, $\mathbf{B}_c^0(B;\mathbb{R}) = \{0\}$ by definition, and $\mathbf{Z}_c^0(B;\mathbb{R}) = \ker(d : \Omega_c^0(B) \to \Omega_c^1(B)) = \{f \in \mathcal{D}(B) : df = 0\}$. If B is compact, $\mathbf{Z}_c^0(B;\mathbb{R}) \cong \mathbb{R}$, otherwise $\mathbf{Z}_c^0(B;\mathbb{R}) = \{0\}$. For an interpretation when B is not compact, see Example 5.53(i); when B is compact, the vector space of all real functions $f \in \mathcal{D}(B)$ such that $df = 0$ is isomorphic to $\mathbb{R}$.

(ii) If B is a contractible Hausdorff manifold ([P1], section 3.3.8**(VI)**), then $\mathbf{H}_c^p(B;\mathbb{R}) \cong \mathbb{R}$ if $p = m$ ($m = \dim(B)$) and $\mathbf{H}_c^p(B;\mathbb{R}) = \{0\}$ if $p \neq m$.

Any such manifold has the same compactly supported cohomology as $\mathbb{R}^m$; simply apply Lemma 5.52.

(iii) If B is the union of a family $(B_i)_{i\in I}$ of pairwise disjoint submanifolds, then $\mathbf{H}_c^p(B;\mathbb{R}) \cong \bigoplus_{i\in I} \mathbf{H}_c^p(B_i;\mathbb{R})$ (**exercise**).

(iv) For the sphere $\mathbb{S}^m$ (footnote 2, p. 98), $\mathbf{H}^0(\mathbb{S}^0;\mathbb{R}) \cong \mathbb{R}^2$, $\mathbf{H}^p(\mathbb{S}^0;\mathbb{R}) = \{0\}$ if $p \neq 0$, and, for $m \geq 1$, $\mathbf{H}^0(\mathbb{S}^m;\mathbb{R}) \cong \mathbf{H}^m(\mathbb{S}^m;\mathbb{R}) \cong \mathbb{R}$ and $\mathbf{H}^p(\mathbb{S}^m;\mathbb{R}) = \{0\}$ if $p \neq 0, m$ ([P1], section 3.3.8**(VI)**).

Indeed, since $\mathbb{S}^m$ is compact, $\mathbf{H}^p(\mathbb{S}^m;\mathbb{R}) \cong \mathbf{H}_p(\mathbb{S}^m;\mathbb{R})$ by Theorem 5.62. We have $\mathbb{S}^0 = \{-1,1\}$, $\mathbf{H}^0(\{1\};\mathbb{R}) \cong \mathbf{H}^0(\{-1\};\mathbb{R}) \cong \mathbb{R}$ by **(i)**, so $\mathbf{H}^0(\mathbb{S}^0;\mathbb{R}) \cong \mathbf{H}^0(\{1\};\mathbb{R}) \oplus \mathbf{H}^0(\{-1\};\mathbb{R}) \cong \mathbb{R}^2$. For $p \geq 1$, $\mathbf{H}^p(\{1\};\mathbb{R}) = \mathbf{H}^p(\{-1\};\mathbb{R}) = \{0\}$, so $\mathbf{H}^p(\mathbb{S}^0;\mathbb{R}) = \{0\}$ by **(iii)**.

For $m \geq 1$, we can proceed inductively. The circle $\mathbb{S}^1$ has a covering formed by two connected arcs $U, V \subsetneq \mathbb{S}^1$ with non-empty intersection. Each arc is contractible and therefore has the same homology space as a pair of points, namely $\mathbf{H}_p(U \cap V;\mathbb{R}) = \mathbf{H}_p(U;\mathbb{R}) \oplus \mathbf{H}_p(V;\mathbb{R}) = 0$ for $p \geq 1$ and $\mathbf{H}_0(U \cap V;\mathbb{R}) \cong \mathbf{H}_0(U;\mathbb{R}) \oplus \mathbf{H}_0(V;\mathbb{R}) \cong \mathbb{R}^2$. Furthermore, $\mathbf{H}_0(\mathbb{S}^1;\mathbb{R}) \cong \mathbf{H}^0(\mathbb{S}^1;\mathbb{R}) \cong \mathbb{R}$ by

Theorem 5.62 and **(i)**, since $\mathbb{S}^1$ is connected and compact. We therefore have the exact Mayer–Vietoris homology sequence [5.49]:

$$0 \longrightarrow 0 \longrightarrow \mathbf{H}_1\left(\mathbb{S}^1;\mathbb{R}\right) \stackrel{f}{\longrightarrow} \underbrace{\mathbf{H}_0\left(U \cap V;\mathbb{R}\right)}_{\mathbb{R}^2} \stackrel{g}{\longrightarrow} \underbrace{\mathbf{H}_0\left(U;\mathbb{R}\right) \oplus \mathbf{H}_0\left(V;\mathbb{R}\right)}_{\mathbb{R}^2}$$

$$\longrightarrow \underbrace{\mathbf{H}_0\left(\mathbb{S}^1;\mathbb{R}\right)}_{\mathbb{R}} \longrightarrow 0.$$

Since f is injective, $\mathbf{H}_1\left(\mathbb{S}^1;\mathbb{R}\right) \cong \operatorname{im}(f) = \ker(g)$ and $\operatorname{coker}(g) \cong \mathbb{R}$, so $\ker(g) \cong \mathbb{R}$ and $\mathbf{H}_1\left(\mathbb{S}^1;\mathbb{R}\right) \cong \mathbb{R}$.

The sphere $\mathbb{S}^2$ is a connected, compact, two-dimensional manifold with a north pole N and a south pole S. It has a covering formed by the two open sets $U = \mathbb{S}^2 - \{N\}$ and $V = \mathbb{S}^2 - \{S\}$. The open sets U and V are contractible and therefore have the same homology space as a point. The intersection $U \cap V$ retracts to the equator $\mathbb{S}^1$. Moreover, $\mathbf{H}_2\left(\mathbb{S}^2;\mathbb{R}\right) \cong \mathbb{R}$ by **(i)**. We therefore have the exact Mayer–Vietoris homology sequence

$$\underbrace{\mathbf{H}_1\left(U;\mathbb{R}\right) \oplus \mathbf{H}_1\left(V;\mathbb{R}\right)}_{0} \stackrel{f}{\longrightarrow} \mathbf{H}_1\left(\mathbb{S}^2;\mathbb{R}\right)$$

$$\stackrel{g}{\longrightarrow} \underbrace{\mathbf{H}_0\left(U \cap V;\mathbb{R}\right)}_{\mathbb{R}} \stackrel{h}{\longrightarrow} \underbrace{\mathbf{H}_0\left(U;\mathbb{R}\right) \oplus \mathbf{H}_0\left(V;\mathbb{R}\right)}_{\mathbb{R}^2} \longrightarrow \underbrace{\mathbf{H}_0\left(\mathbb{S}^2;\mathbb{R}\right)}_{\mathbb{R}} \longrightarrow 0,$$

so $\mathbf{H}_1\left(\mathbb{S}^2;\mathbb{R}\right) \cong \operatorname{im}(f) = \ker(g)$ and $\operatorname{im}(g) = \ker(h)$, $\operatorname{coker}(h) \cong \mathbb{R}$ and hence $\operatorname{im}(h) \cong \mathbb{R}$. Thus, $\mathbb{R}/\ker(h) \cong \mathbb{R}$, which implies that $\ker(h) = \{0\}$, and the exact sequence $0 \longrightarrow \mathbf{H}_1\left(\mathbb{S}^2;\mathbb{R}\right) \longrightarrow 0$ gives us that $\mathbf{H}_1\left(\mathbb{S}^2;\mathbb{R}\right) = \{0\}$.

If $m \geq 3$, we can cover $\mathbb{S}^m$ by two open sets U, V that are diffeomorphic to $\mathbb{R}^m$ and whose intersection retracts to $\mathbb{S}^{m-1}$, allowing us to repeat the same reasoning as above (**exercise**).

For an interpretation of the case $m = 1$, $p = 0$, see (**i**); when $m = 1$, $p = 1$, a function $f \in \mathcal{D}\left(\mathbb{S}^1\right)$ is a derivative if and only if $\int_{\mathbb{S}^1} f(t)\,.dt = 0$.

5.7. Integral curves and manifolds

5.7.1. *First-order differential equations*

Consider a Hausdorff real Banach manifold B of class C^r $(2 \leq r \leq \infty)$ and a vector field X of class C^{r-1} on B. Let J be a non-empty open interval of $\mathbb{R}$. The tangent bundle of J is the trivial bundle $J_{\mathbb{R}} = J \times \mathbb{R}$ (section 3.4.1, Example 3.26(a)) and we have a canonical section $\iota : J \to J_{\mathbb{R}}$ such that, for every $t \in J$, $\iota(t) \in (J_{\mathbb{R}})_t = \mathbb{R}$ and $\iota(t) = 1$.

A curve of class C^r in B is a mapping $c : J \to B$, where J is a non-empty open interval of $\mathbb{R}$ (Definition 2.20). Given a curve c, the following diagram commutes (section 2.3.1):

$$\begin{array}{ccc} J_{\mathbb{R}} & \xrightarrow{c_*} & T(B) \\ \downarrow \pi_J & & \downarrow \pi_B \\ J & \xrightarrow{c} & B \end{array}$$

We write that $dc/dt = c_* \circ \iota : J \to T(B)$. Hence, dc/dt is a curve in $T(B)$ of class C^{r-1}. If $c(t)$ is the position of a point at time t, $dc/dt = v$ is the velocity of this point.

A curve of class C^r, $c : J \to B$, where J is an open interval of $\mathbb{R}$, is said to be an integral of the *first-order differential equation*

$$\frac{du}{dt} = X(u) \qquad [5.52]$$

if, for all $t \in J$, $\left.\frac{dc}{dt}\right|_t = X(c(t))$. By translation of the origin, we can assume that J contains 0; the curve c is said to satisfy the initial condition c_0 if $c(0) = c_0$ $(c_0 \in B)$.

Note that, if B' is a manifold and $B = \mathbb{R} \times B'$, then u is written as (τ, w), where $w \in N$ and [5.52] takes the form

$$\frac{dw}{dt} = X(\tau, w), \quad \frac{d\tau}{dt} = 1.$$

The *local study* of a differential equation on a manifold can be reduced to the study of a differential equation on an open subset of a Banach space (section 1.5). Let $c_0 \in B$. By the Cauchy–Lipschitz theorem (Corollary 1.72 and Remark 1.73), there exist a real number $\varepsilon > 0$, an open neighborhood V of c_0 in B and a mapping $\phi :]-\varepsilon, +\varepsilon[\times V \to B$ such that $\phi(0, q) = q$ and $X_q = \left.\frac{\partial(\phi(t,q))}{\partial t}\right|_{t=0}$ for all $q \in V$. Setting $\phi_t(q) = \phi(t, q)$, we have $\phi_{s+t} = \phi_s \circ \phi_t$ for any sufficiently small t, s. For every function $f : B \to \mathbb{R}$ of class C^r,

$$(\mathcal{L}_X . f)(q) = \lim_{t \to 0} \frac{f(\phi_t(q)) - f(q)}{t}.$$

DEFINITION 5.64.– *The mapping* ϕ *is called the* local flow *defined by* X *in* $]-\varepsilon,+\varepsilon[\times V$.

Manifolds provide an appropriate framework for the *global study* of first-order differential equations.

LEMMA 5.65.– *The union* $J(c_0)$ *of all intervals containing* 0 *and on which there exists an integral curve* c *of [5.52] satisfying the initial condition* c_0 *is an open interval* $]t^-(c_0), t^+(c_0)[$ *of* $\mathbb{R}$*. The curve* $c : J(c_0) \to B$ *is a maximal integral of [5.52] (section 1.5.1, Definition 1.75).*

PROOF.– (A) Let $c_1 : J_1 \to B$ and $c_2 : J_2 \to B$ be two integral curves of [5.52] satisfying the same initial condition c_0 and let J^* be the subset of $J_1 \cap J_2$ defined by

$$J^* = \{t \in J_1 \cap J_2 : c_2(t) = c_1(t)\}.$$

Since B is Hausdorff, J^* is closed in $J_1 \cap J_2$ ([P2], section 2.3.3(**II**), Lemma 2.30). If $t_0 \in J^*$, the Cauchy–Lipschitz theorem shows that there exists an open interval $]t_0-\varepsilon, t_0+\varepsilon[$ $(\varepsilon > 0)$ contained in J^*. Hence, since $J_1 \cap J_2$ is connected, we must have $J^* = J_1 \cap J_2$.

(B) We can therefore define a uniquely determined mapping $c_3 : J_1 \cup J_2 \to B$ that coincides with c_1 on J_1 and with c_2 on J_2 so that c_3 is an integral curve of [5.52] on $J_1 \cup J_2$ satisfying the initial condition c_0. Proceeding in the same way for every integral curve of [5.52] satisfying the initial condition c_0 gives the desired integral curve $c : J(c_0) \to B$. The set $J(c_0)$ is an interval of $\mathbb{R}$ by (A). It is open as a union of open sets. ■

Let $\mathfrak{D}(X) \subset B$ be the subset of $\mathbb{R} \times B$ formed by the points (t,q) such that $t^-(q) < t < t^+(q)$.

DEFINITION 5.66.– *A* flow *(or* global flow*) on* X *is a mapping* $\phi : \mathfrak{D}(X) \to B : (t,q) \to \phi(t,q) = \phi_t(q)$*, where* $\phi(.,q) : t \mapsto \phi(t,q)$ *is a* maximal *integral curve of [5.52] satisfying the initial condition* q *(compare with Definition 5.64).*

By translation, it is clear that, for every $t_0 \in J(q)$,

$$J(\phi_{t_0}, q) = J(q) - t_0 \qquad [5.53]$$

(translation of $J(q)$ by $-t_0$).

THEOREM 5.67.– *If* $t^+(q) < +\infty$ *(respectively* $t^-(q) > -\infty$*), then, for every compact subset* K *of* B*, there exists* $\varepsilon > 0$ *such that, for all* $t > t^+(q) - \varepsilon$ *(respectively* $t < t^-(q) + \varepsilon$*),* $\phi_t(q) \notin K$ *(i.e. the maximal integral curve "ends outside" every compact subset of* B*).*

PROOF.– Suppose that $t^+(q) < +\infty$ and that there does not exist a real number $\varepsilon > 0$ with the stated property. Then, we can find a sequence (t_n) of points > 0 in $J(q)$ tending to $t^+(q)$ from below such that $\phi_{t_n}(q) \in K$ for every n. Since K is compact, there exists a subsequence (t_{n_j}) such that $\left(\phi_{t_{n_j}}(q)\right)$ converges to some point z of K. By the Cauchy–Lipschitz theorem, there exist an open neighborhood U of z and a number $\delta > 0$ such that $t^+(y) > \delta$ for every $y \in U$. By choosing j to be sufficiently large, $t^+(q) < t_{n_j} + \delta$. But $\phi_{t_{n_j}}(q) \in U$, so $t^+\left(\phi_{t_{n_j}}(q)\right) > \delta$, and, by [5.53], $t^+\left(\phi_{t_{n_j}}(q)\right) = t^+(q) - t_{n_j} > \delta$; hence, $t^+(q) > t_{n_j} + \delta$, which gives a contradiction. ■

COROLLARY 5.68.– *If X is compactly supported (in particular, if the manifold B is compact), then $J(q) = \mathbb{R}$.*

Let ϕ be the flow determined by the differential equation [5.52] and the initial condition $\phi_0(q) = q$. Each ϕ_t is a diffeomorphism of class C^r, and, for all sufficiently small t, t', we have $\phi_{t'} \circ \phi_t = \phi_{t+t'}$. If X is compactly supported, this flow is therefore a *group* of diffeomorphisms of class C^r by Corollary 5.68.

Conversely, let ϕ be a flow on B. The *infinitesimal generator* A of the group ϕ is given by

$$A.q = \lim_{t \to 0^+} \frac{\phi_t - 1_B}{t}.q = \left.\frac{\partial \phi_t}{\partial t}\right|_{t=0}.q = X(q).$$

Theorem 5.15 implies the following result:

COROLLARY 5.69.– *If B is locally finite-dimensional, every derivation on B induces a flow on B, and vice versa.*

If X is a vector field on B of class C^{r-1}, ϕ its flow, M a Banach manifold of class C^r and $f : B \to M$ is of class C^r, write $\phi_t^* f := f \circ \phi_t$.

THEOREM 5.70.– *1) Let $\mathbf{F}$ be a Banach space and $\mathbf{f} \in C^r(B; \mathbf{F})$. Then $\frac{d}{dt}(\phi_t^* \mathbf{f}) = \phi_t^*(\mathcal{L}_X.\mathbf{f})$.*

2) Let Y be a field of vectors of class C^{r-1} on B. Then:

$$\boxed{\left.\frac{d}{dt}\right|_{t=0} (\phi_t^* Y) = [X, Y].} \qquad [5.54]$$

PROOF.– 1): Let $b \in B$. Then:

$$\begin{aligned} \frac{d}{dt}(\phi_t^* \mathbf{f})(b) &= \frac{d}{dt}(\mathbf{f} \circ \phi_t(b)) = d\mathbf{f}(\phi_t(b)).\frac{d\phi_t(b)}{dt} \\ &= d\mathbf{f}(\phi_t(b)).X(\phi_t(b)) = \mathcal{L}_X.\mathbf{f}(\phi_t(b)) = (\phi_t^*(\mathcal{L}_X.\mathbf{f}))(b). \end{aligned}$$

2) Similarly, $\frac{d}{dt}\left(\phi_t^* Y\right) = \phi_t^*\left(\mathcal{L}_X.Y\right) = \phi_t^*\left([X,Y]\right)$ by [5.17], and $\phi_0^* = 1_B$. ■

The equality [5.54] can be taken as an alternative definition of the Lie bracket.

5.7.2. *Second-order differential equations*

Let B be a real Hausdorff Banach manifold of class C^r $(3 \leq r \leq \infty)$ and $c : J \to B$ a curve of class C^r. In section 5.7.1, we defined the velocity $v(t)$ of a point $c(t)$ of the manifold B by $v = dc/dt : J \to T(B)$, where $T(B)$ is the tangent bundle of B, which is a manifold of class C^{r-1}. The second derivative, or acceleration, $\gamma = d^2c/dt^2$ is therefore a mapping $J \to T(T(B))$, where $T(T(B))$ is now a manifold of class C^{r-2}. To express a second-order differential equation, we need to be given a vector field Z of class C^{r-2} on $T(B)$. Any such vector field is a mapping $Z : T(B) \to T(T(B)) : v \mapsto Z(v)$. We also need to assume that there exists a lifting β of class C^{r-1} of c into $T(B)$ satisfying $\pi \circ \beta = c$, where $\pi : T(B) \twoheadrightarrow B$ is the projection. This leads us to the following commutative diagram:

$$\begin{array}{ccccc} & & T(B) & \underset{\pi_*}{\overset{Z}{\rightleftarrows}} & T(T(B)) \\ & \beta \nearrow & \downarrow \pi & & \\ J & \xrightarrow{c} & B & & \end{array}$$

It must be the case that $\pi_*(Z(v)) = v$ for every $v \in T(B)$, so

$$\pi_* \circ Z = 1_{T(B)}. \qquad [5.55]$$

Since $\pi \circ \beta = c$, $\beta = v$ and $\beta(t) = dc/dt = v(t)$, this implies that

$$\frac{d(\pi \circ \beta)}{dt} = \beta. \qquad [5.56]$$

DEFINITION 5.71.– *Given a vector field Z on $T(B)$ of class C^{r-2} satisfying [5.55], a solution of the* second-order differential equation

$$\frac{d^2u}{dt^2} = Z\left(\frac{du}{dt}\right) \qquad [5.57]$$

is an integral curve of the field Z, i.e. a function $c : J \to B$ satisfying this equation.

By iterating this method, we can similarly define differential equations of order $n \geq 3$.

Consider a chart $(U, \varphi, \mathbf{E})$ of B. Then, $T(B)$ can be identified with $U \times \mathbf{E}$, and any field Z satisfying [5.55] has the local expression $(x, \mathbf{v}) \mapsto ((x, \mathbf{v}), \mathbf{f}(x, \mathbf{v}))$, where $\mathbf{f}$

is a mapping of class C^{r-2} from $\varphi(U) \times \mathbf{E}$ into $\mathbf{E}$. Given any solution $c : J \to B$ of [5.57], $y = \varphi \circ c$ therefore satisfies the differential equation

$$\frac{d^2 y}{dt^2} = \mathbf{f}\left(y, \frac{dy}{dt}\right). \qquad [5.58]$$

Given the differential equation [5.57] and $v \in T(B)$, there exists a unique integral curve β_v on $T(B)$ such that

$$\frac{d\beta_v}{dt} = Z(\beta_v), \quad \beta_v(0) = v. \qquad [5.59]$$

Write $\mathfrak{O}$ for the set of vectors of $T(X)$ such that β_v is defined in an interval J containing $[0,1]$; we say that 1 *belongs to the domain of* β_v. It can be shown that $\mathfrak{O}$ is an open subset of $T(B)$ and that the mapping

$$\exp_Z : \mathfrak{O} \to B : v \mapsto \pi \circ \beta_v(1),$$

called the *exponential mapping* defined by Z, is of class C^{r-2} (see, for example, [DIE 93], Volume 4, section 18.4).

5.7.3. *Sprays*

(I) An important special case of a second-order differential equation is when the equations are *isochronous*, i.e. independent of the choice of unit of time. In other words, an isochronous differential equation remains unchanged whenever t is replaced by λt $(\lambda \neq 0)$. Set $z(t) = y(\lambda t) = y(\tau)$ $(\tau = \lambda t)$. Then, $\frac{dz}{dt}(t) = \lambda \frac{dy}{d\tau}(\lambda t)$, $\frac{d^2 z}{dt^2}(t) = \lambda^2 \frac{d^2 y}{d\tau^2}(\lambda t)$. If y is a solution of [5.58], then

$$\lambda^2 \frac{d^2 y}{d\tau^2}(\lambda t) = \lambda^2 f\left(y(\lambda t), \frac{dy}{d\tau}(\lambda t)\right).$$

Hence, z is a solution of [5.58] if and only if the function $f(b, .) : \mathbf{v} \mapsto f(b, \mathbf{v})$ is homogeneous of degree 2, i.e. satisfies the relation $f(b, \lambda \mathbf{v}) = \lambda^2 f(b, \mathbf{v})$ for every $b \in U$, $\mathbf{v} \in \mathbf{E}$, $\lambda \in \mathbb{R}$. If so, the vector field Z (or the differential equation [5.57] that it determines) is said to be a *spray*.

Consider the integral curve $\beta_{\mathbf{v}} : J \to T(B)$ defined by [5.59]. Then, $\beta_{\mathbf{v}}(t) = \frac{dy}{dt}(t)$, $\beta_{\mathbf{v}}(0) = v$, where y is a solution of [5.58]. Thus, $\beta_{\mathbf{v}}(\lambda t) = \frac{dy}{d\tau}(\lambda t)$ $(\tau = \lambda t)$ and $\lambda \frac{dy}{d\tau}(\lambda t) = \frac{dz}{dt}(t)$, where $z(t) = y(\lambda t)$. If we assume that the differential equation is isochronous, then z is also a solution of [5.58] and $\frac{dz}{dt}(0) = \lambda \frac{dy}{d\tau}(0) = \lambda \mathbf{v}$. Thus, $\beta_{\lambda \mathbf{v}}(t) = \beta_{\mathbf{v}}(\lambda t)$. By switching the roles of λ and t, we further have $\beta_{\mathbf{v}}(\lambda t) = \beta_{t\mathbf{v}}(\lambda)$. Moreover, from the relation $\lambda \frac{dy}{d\tau}(\lambda t) = \frac{dz}{dt}(t)$, we deduce that $\lambda \beta_{\mathbf{v}}(\lambda t) = \beta_{\lambda \mathbf{v}}(t)$.

(II) Let us return to the global formulation [5.57] of the second-order differential equation. The next lemma follows from the above discussion:

LEMMA 5.72.– *The following conditions are equivalent:*

i) A time $t > 0$ belongs to the domain of β_v if and only if 1 belongs to the domain of β_{tv}, in which case $\pi.\beta_v(t) = \pi.\beta_{tv}(1)$.

ii) Given times $t, \lambda > 0$, λt belongs to the domain of β_v if and only if λ belongs to the domain of β_{tv}, in which case $\pi.\beta_{tv}(\lambda) = \pi.\beta_v(\lambda t)$.

iii) A time $t > 0$ belongs to the domain of $\beta_{\lambda v}$ $(\lambda > 0)$ if and only if λt belongs to the domain of β_v, in which case $\beta_{\lambda v}(t) = \lambda.\beta_v(\lambda t)$.

iv) For every $\lambda \in \mathbb{R}$ and every $v \in T(B)$, $Z(\lambda v) = \lambda^2.Z(v)$.

DEFINITION 5.73.– *Let Z be a vector field on $T(B)$ satisfying [5.55]. If the equivalent conditions from Lemma 5.72 are satisfied, the field Z is said to be a* spray.

Let $b \in B$ and write 0_b for the zero vector in the fiber $T_b(B)$. If Z is a spray, then $Z(0_b) = 0$ by Lemma 5.72(iv), so

$$\exp_Z(0_b) = \pi(0_b) = b. \qquad [5.60]$$

Thus, let $\exp_Z|_b : T_b(B) \to B$ be the restriction of $\exp_Z$ to the tangent space $T_b(B)$ and $(\exp_Z|_b)_* : T(T_b(B)) \to T_{\exp_Z|_b}(B)$ its tangent linear mapping. Since $T_b(B)$ is a Banach space, $T_{0_b}(T_b(B))$ can be identified with $T_b(B)$ and $T_{\exp_Z|_b(0_b)}(B)) = T_b(B)$ by [5.60]. Hence, with these identifications, $(\exp_Z|_b)_*(0_b) : T_b(B) \to T_b(B)$.

THEOREM 5.74.– *Let B be a real manifold of class C^r, where $r \geq 3$, and let Z be a spray on B. Then, $\exp_Z|_b : T_b(B) \to B$ induces a local isomorphism of class C^{r-1} in the neighborhood of 0_b. More precisely, $(\exp_Z|_b)_*(0_b) = 1_{T_b(B)}$ (see Figure 5.4).*

PROOF.– Consider the straight line $\alpha_b(t) = t.v$ in $T_b(B)$, where $v \in T_b(B)$. We have $T_{0_b}(T_b(B)) \ni \frac{d\alpha_b}{dt}(0) = v$. This makes sense because $T_{0_b}(T_b(B))$ and $T_b(B)$ can be identified. Thus, $(\exp_Z|_b)(\alpha_b(t)) = (\exp_Z|_b)(tv)$. Hence:

$$(\exp_Z|_b)_*(0_b).v = \frac{d}{dt}((\exp_Z|_b)(tv))\bigg|_{t=0} = \frac{d}{dt}(\pi \circ \beta_{tv})\bigg|_{t=0}.$$

By [5.56], $\frac{d}{dt}(\pi \circ \beta_v) = \beta_v$ and $\beta_v(0) = v$, so $(\exp_Z|_b)_*(0_b).v = v$. By the inverse mapping theorem (section 1.2.6(**II**), Theorem 1.29), $\exp_Z|_b$ is a local isomorphism at 0_b. ■

EXAMPLE 5.75.– *1) Let $B =]0, +\infty[$ and identify $T(B)$ with $\mathbb{R}$. The field $Z : (b, \mathbf{v}) \mapsto ((b, \mathbf{v}), (\mathbf{v}, \mathbf{v}^2/b))$ corresponding to the differential equation $\ddot{x} = \dot{x}^2/x$ is a spray. We will calculate the associated exponential mapping. We have $\ddot{x}/\dot{x} = \dot{x}/x$, so $\dot{x} = c.x$, where c is a constant, and $x(t) = e^{ct}b$, where $b = x(0)$. Hence, $\dot{x}(t) = e^{ct}bc$. Writing $\dot{x}(0) = v$ gives $v = bc$, so $c = v/b$. Thus, $x(1) = e^c b = e^{v/b}b$. The exponential mapping is, therefore, the diffeomorphism $(b, v) \mapsto e^{v/b}b$ from $T_b(B) = \{b\} \times \mathbb{R}$ onto $\mathbb{R}_+^\times$.*

2) Now, let $B = \mathbb{R}$ and consider the spray $Z : (b, \mathbf{v}) \mapsto ((b, \mathbf{v}), (\mathbf{v}, \mathbf{v}^2))$ corresponding to the differential equation $\ddot{x} = \dot{x}^2$. Then, $\ddot{x}/\dot{x}^2 = 1$, so $-1/\dot{x} = t + c$, or alternatively $\dot{x} = -\frac{1}{t+c}$, and $x(t) = -\ln(t + c) + c_1$. Thus, $x(0) = -\ln(c) + c_1 = b$ and $\dot{x}(0) = -1/c = v$, giving $c = -1/v$. Hence, $x(1) = -\ln(1 + c) + c_1 = -\ln(1 + c) + b + \ln(c) = b - \ln(1 + 1/c) = b - \ln(1 - v)$. The exponential mapping is therefore the diffeomorphism $(b, v) \mapsto b - \ln(1 - v)$ from $\Omega \cap T_b(\mathbb{R})$ onto $\mathbb{R}$, where $\Omega = \{(b, v) \in \mathbb{R}^2 : v < 1\}$. Hence, $\Omega \cap T_b(\mathbb{R}) = \{b\} \times]-\infty, 1[$.

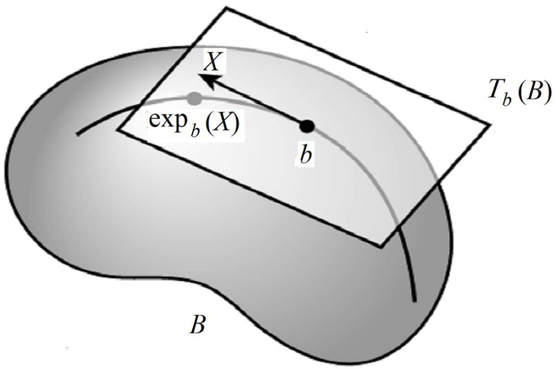


Figure 5.4. *Exponential mapping*

5.7.4. *Straightening of vector fields and frames*

In this section, B is a real, pure, n-dimensional manifold of class C^r $(2 \leq r \leq \infty)$.

THEOREM 5.76.– (straightening theorem for vector fields) *Let X be a vector field of class C^{r-1} on B. If there exists $q_0 \in B$ such that $X(q_0) \neq 0$, then there exists a chart*

(U, ξ, n) centered on q_0 such that (with the notation of Definition 2.33, section 2.3.1) $\xi_* (X|_U) = \frac{\partial}{\partial \xi^1}$, where $\xi = (\xi^1, ..., \xi^n)$.

PROOF.– Since this is a local question, we can reduce to the case where $B = \mathbb{R}^n$, $q_0 = 0$ and $X(0) = (1, 0, ..., 0)$. Let $\phi : (t, x) \mapsto \phi_t(x) = \phi(t, x)$ be the local flow of the vector field X in the neighborhood of 0. Then, the mapping

$$\theta : (x^1, x^2, ..., x^n) \mapsto \phi_{x^1}(0, x^2, ..., x^n)$$

is of class C^r in the neighborhood of $0 \in \mathbb{R}^n$. Its Jacobian matrix at 0 is $1_{\mathbb{R}^n}$, so θ is a local diffeomorphism of class C^r by the inverse mapping theorem (section 1.2.6, Theorem 1.29). Furthermore, writing $\{e_1, ..., e_n\}$ for the canonical basis,

$$\begin{aligned} D\theta(x).e_1 &= \frac{\partial \theta}{\partial x^1}(x) = \left.\frac{d}{dt}\right|_{t=x_1} \phi_t(0, x^2, ..., x^n) \\ &= X(\phi_{x^1}(0, x^2, ..., x^n)) = X(\theta(x)), \end{aligned}$$

so θ_* sends the constant field $e_1 = \frac{\partial}{\partial x^1}$ to the field X in the neighborhood of 0. Hence, $\xi = \theta^{-1}$ satisfies the stated condition. ■

THEOREM 5.77.– (straightening theorem for fields of frames). *Let $X_1, ..., X_p$ be vector fields of class C^{r-1} on B. Suppose that there exists $q \in B$ such that $X_1(q), ..., X_p(q)$ $(q \leq p)$ are linearly independent. The following conditions are equivalent:*

1) The brackets $[X_i, X_j]$ are all zero.

2) There exists a chart (U, ξ, n) centered on x such that $\xi_* (X_i|_U) = \frac{\partial}{\partial \xi^i}$ $(1 \leq i \leq p)$.

PROOF.– This condition is necessary because $\xi_*([X_i, X_j]) = [\xi_*(X_i), \xi_*(X_j)]$ by the functor property of the bracket (Proposition 5.18(2)), and $\left[\frac{\partial}{\partial \xi^i}, \frac{\partial}{\partial \xi^j}\right] = 0$ (Theorem-Definition 5.16(1)). We will show that it is also sufficient. As before in the proof of Theorem 5.76, suppose that $B = \mathbb{R}^n$, $q_0 = 0$, and $X_i(0) = e_i$. Let ϕ^i be the local flow of X_i. The Jacobian matrix of the mapping

$$\Theta : (x_1, ..., x_n) \mapsto \phi^1_{x_1} \circ \phi^2_{x_2} \circ ... \circ \phi^p_{x_p}(0, ..., 0, x_{p+1}, ..., x_n)$$

is $1_{\mathbb{R}^n}$ at 0, so this mapping is a local diffeomorphism of class C^r. As before, we have

$$\begin{aligned} D\Theta(x).e_1 &= \left.\frac{d}{dt}\right|_{t=x_1} \phi^1_t\left(\phi^2_{x_2} \circ ... \circ \phi^p_{x_p}(0, ..., 0, x_{p+1}, ..., x_n)\right) \\ &= X_1(\Theta(x)), \end{aligned}$$

so Θ_* sends the constant field $e_1 = \frac{\partial}{\partial x^1}$ to the field X in the neighborhood of 0. Since $[X_i, X_j] = 0$, the fields X_i and X_j commute pairwise, so their respective flows ϕ^i and ϕ^j also commute. Hence, Θ_* sends $\frac{\partial}{\partial x^i}$ to X_i for $1 \leq i \leq p$. ■

5.7.5. *Integral manifolds, foliations*

(I) In the plane, consider the circle of center 0 and radius $R > 0$. In Cartesian coordinates, it has the equation $x^2 + y^2 = R^2$. Differentiating this expression gives $\omega = 0$, where $\omega = x.dx + y.dy$. Conversely, the "integral curves" of the equation $\omega = 0$, with $\omega = x.dx + y.dy$, are the circles of center 0. As a set, the union of the family of these circles is equal to $\mathbb{R}^2 - \{0\}$. We say that this family forms a *foliation* of the manifold $\mathbb{R}^2 - \{0\}$. More generally, if $\omega = P(x, y).dx + Q(x, y).dy$, where P and Q are functions of class C^1 and Q is not identically zero, the equation $\omega = 0$ is equivalent to the differential equation $\frac{dy}{dx} = -\frac{P(x,y)}{Q(x,y)}$. Theorem 5.67 implies that this equation has maximal integral curves defined on non-empty intervals of the real line. These integral curves do not intersect and form a foliation of the so-called *phase space* of the differential equation.

If we add one or more dimensions, the equation $\omega = 0$ (known as a *Pfaff equation*) is no longer equivalent to a differential equation but a partial differential equation. In this case, the above reasoning no longer applies: the integrability conditions of the Pfaff equation are now non-trivial. These conditions are the object of the Frobenius theorem studied in this section.

In this section, M is a $\mathbb{K}$-manifold of class C^r $(r \in \mathbb{N}_{\mathbb{K}}, r \geq 2)$. Every diffeomorphism is of class C^r. Write $\mathcal{T}_0^1(M)$ for the set of vector fields of class C^{r-1} on M.

(II) CONTACT DISTRIBUTIONS

DEFINITION 5.78.– *Any subbundle Δ of class C^{r-1} of the tangent bundle $T(M)$ (section 3.4.3) is said to be a* contact distribution[17] *of class C^{r-1}.*

For every $x \in M$, we have $\Delta_x \subset T_x(M)$ (where Δ_x is the fiber of Δ over $\{x\}$). Let U be an open subset of M and $p \leq \infty$ the rank of Δ (section 3.4.1, Definition 3.22(iii)). A section of Δ over U (i.e. a mapping $U \rightarrow \Delta : x \mapsto \Delta_x$) is said to be a *p-field* over U.

For every $x \in M$, let Δ_x^0 be the *polar set* of Δ_x ([P2], section 3.5.2), i.e. the subspace of the cotangent space $T_x^\vee(M)$ formed by the continuous linear forms that vanish on Δ_x.

17 This notion of distribution is due to Chevalley [CHE 46]. It should not be confused with the notion of distribution due to Schwartz that is studied in [P2], section 4.4.1, and above, in section 5.2.1(**IV**).

DEFINITION 5.79.– *The subbundle* Δ^0 *of* $T^{\vee}(M)$ *is said to be a* codistribution.

Since Δ_x is closed in $T_x(M)$ (section 3.4.3, Lemma-Definition 3.40), the bipolar set theorem ([P2], section 3.5.2, Theorem 3.78) implies that $\Delta_x^{00} = \Delta_x$ for every $x \in M$. This can be expressed as follows:

$$\Delta^{00} = \Delta. \qquad [5.61]$$

Note the following result, which uses the notation of Corollary-Definition 3.21 (section 3.3.4), and whose proof is an **exercise**:

LEMMA 5.80.– *i) If* p *is finite, a* p*-field* $M \to \Delta : x \mapsto \Delta_x$ *is of class* C^{r-1} *if and only if it is possible to choose vector fields* $X_1, ..., X_p$ *in* $\mathcal{T}_0^1(M)$ *satisfying the following property: for every* $x \in M$*, there exists an open neighborhood* U *of* x *such that* $X_1(y), ..., X_p(y)$ *form a basis of* Δ_y *for every* $y \in U$*. If so, the set of vector fields* Y *of class* C^{r-1} *such that* $Y(x) \in \Delta_x$ *for every* $x \in M$ *is the free* $\mathcal{C}^r(M)$*-module* $\Gamma^{(r-1)}(M, \Delta)$ *with basis* $\{X_1, ..., X_p\}$.

ii) If the vector fields $Y_1, ..., Y_m$ *in* $\mathcal{T}_0^1(M)$ *are such that, for every* $x \in M$*,* $Y_1(x), ..., Y_m(x)$ *generate* Δ_x *(i.e. if* Δ_x *is the vector space span* $\{Y_1(x), ..., Y_m(x)\}$ *generated by* $Y_1(x), ..., Y_m(x)$*), we write that* $\Gamma^{(r-1)}(M, \Delta) = \text{span}\{Y_1, ..., Y_m\}$*, which means that* $\Gamma^{(r-1)}(M, \Delta)$ *is the* $\mathcal{C}^r(M)$*-module generated by* $Y_1, ..., Y_m$.

iii) In the situation of (i), for every $x \in M$*,* $\Delta_x^0 \subset T_x^{\vee}(M)$ *has a basis* $\{\omega^j(x) : 1 \leq j \leq n-p\}$*, where the* ω^j *are Pfaff forms (section 4.3.2, Definition 4.20) of class* C^{r-1} *that vanish on the* X_i*, i.e. which satisfy*

$$\langle \omega^j, X_i \rangle = 0, \quad 1 \leq j \leq n-p, \quad 1 \leq i \leq p.$$

The set of Pfaff forms ϖ *so that* $\varpi_x \in \Delta_x^0$ *is thus the free* $\mathcal{C}^r(M)$*-module* $\Gamma^{(r-1)}(M, \Delta^0)$ *with basis* $\{\omega^1, ..., \omega^{n-p}\}$.

Let N be a submanifold of M and $\iota : N \hookrightarrow M$ the inclusion mapping.

DEFINITION 5.81.– *1) The submanifold* N *is said to be an* integral manifold *of the contact distribution* Δ *if* $\iota_*(T_x(N)) = \Delta_x$ *for every* $x \in N$ *(i.e. the tangent space* $T_x(N)$ *can be identified with the subspace* $\Delta_x \subset T_x(M)$*). It is said to be a* maximal *integral manifold of* Δ *if every integral manifold containing* N *is equal to* N.

2) The contact distribution Δ *is said to be* integrable *(or* completely integrable*) if it admits an integrable manifold.*

3) The contact distribution Δ *is said to be* involutive *if* $[Y, Z] \in \Gamma^{(r-2)}(U, \Delta)$ *for every* $Y, Z \in \Gamma^{(r-1)}(U, \Delta)$ *(where* U *is a sufficiently small open subset of* M*).*

Note that the notion of the integrability of a manifold is local and invariant under diffeomorphism. Furthermore, a maximal integral manifold of a contact distribution Δ of rank p, if it exists, necessarily has dimension p.

EXAMPLE 5.82.– *If N is one-dimensional, we locally recover the notion of an integral curve $c : I \rightarrow M$ of a first-order differential equation [5.52] (where $X(x)$, assumed to be $\neq 0$, forms a basis of Δ_x for every $x \in \{c(t) : t \in I\}$). In this case, the contact distribution $\Delta = span\{X\}$ is clearly involutive.*

(III) PFAFF SYSTEMS Consider a contact distribution Δ and $n-k$ Pfaff forms ω^j such that, for every $x \in M$, the $\omega^j(x)$ $(1 \leq j \leq n-k)$ generate $\Delta_x^0 \subset T_x^\vee(M)$ (without necessarily being linearly independent). Then, N is an integral manifold of Δ if and only if, for every $x \in N$, $\iota_*(T_x(N)) = \text{span}\{X_1(x), ..., X_p(x)\}$, where

$$\left\langle \omega^j, X_i \right\rangle = 0, \quad 1 \leq j \leq n-k, \quad 1 \leq i \leq p. \qquad [5.62]$$

This can also be expressed by saying that the restrictions to N of the $n-k$ Pfaff forms ω^j are zero, or that N is an integral manifold of the following *Pfaff system*, said to be *associated with* Δ :

$$\boxed{\omega^j = 0 : 1 \leq j \leq n-k.} \qquad [5.63]$$

EXAMPLE 5.83.– *Consider the case where M is an open subset of $\mathbb{R}^n$. Let (U, ξ, k) be a chart of N, $z^i = pr_i \circ \iota \circ \xi^{-1} : M \rightarrow N$ $(1 \leq i \leq n)$ and $\Delta = span\left\{\frac{\partial}{\partial \xi^1}, ..., \frac{\partial}{\partial \xi^k}\right\}$. Then, $\omega^j = d\xi^{k+j}$ $(1 \leq j \leq n-k)$ and the manifold N is the set*

$$N = \left\{\left(z^1, ..., z^n\right) \in M : z^{k+1} = c^1, ..., z^n = c^{n-k}\right\},$$

where the $c^1, ..., c^{n-k}$ are constants: N is the manifold whose tangent bundle $T(N)$ is generated by the $X_i = \frac{\partial}{\partial \xi^i}$ $(1 \leq i \leq k)$, i.e. $T(N) = \Delta$. Note that for every pair (i,j), $[X_i, X_j] = \left[\frac{\partial}{\partial \xi^i}, \frac{\partial}{\partial \xi^j}\right] = 0$, because $\xi_*\left(\left[\frac{\partial}{\partial \xi^i}, \frac{\partial}{\partial \xi^j}\right]\right) = [\mathbf{e}_i, \mathbf{e}_j]$ by the functor property of the bracket (Proposition 5.18(2)).

(IV) SYSTEM OF PARTIAL DIFFERENTIAL EQUATIONS ASSOCIATED WITH A PFAFF SYSTEM Let us consider a somewhat more general situation than Example 5.83. As before, let M be an open subset of $\mathbb{K}^n$, and consider the Pfaff system [5.63], where $k = p$ and the ω^j $(1 \leq j \leq p)$ are linearly independent at every point. Let $\left\{x^1, ..., x^n\right\}$ be the coordinates of a point of $\mathbb{K}^n$ in the canonical basis $\{e_1, ..., e_n\}$. Since $\left\{dx^1, ..., dx^n\right\}$ is a basis of $(\mathbb{K}^n)^\vee$, we can write:

$$\omega^j(x) = \sum_{i=1}^{n} a_i^j(x) . dx^i \ \ (1 \leq j \leq n-p) .$$

If N is an integral p-dimensional manifold of $\Delta = \Delta^{00}$ (see [5.61]) and (U, ξ, p) is a chart of N, let $\iota : U \to M$ be inclusion and $z^i = \mathrm{pr}_i \circ \iota \circ \xi^{-1} : U \to M$ $(1 \leq i \leq n)$. Then, U is given by the parametric representation (section 2.3.6, Lemma 2.58) $U = \{z(u) : u \in \xi(U)\}$. At the point $z = \left(z^1, ..., z^n\right)$, its tangent vectors are $X_i = \sum_{k=1}^{p} \frac{\partial z^i}{\partial u^k}(z).e_k$. The relations [5.62] are translated by the following system of partial differential equations:

$$\sum_{i=1}^{n} a_i^j (z^1, ..., z^n). \frac{\partial z^i}{\partial u^k} = 0 \quad (1 \leq k \leq p, \quad 1 \leq j \leq n-p).$$

Let $x_0 \in M$. The $\omega^j(x_0)$ $(1 \leq j \leq n-p)$ are linearly independent by the hypotheses. Furthermore, these $\omega^j(x_0)$ and the dx^i $(1 \leq i \leq n)$ generate the vector space $T_{x_0}^{\vee}(M)$. Hence, there exists a subset $\{i_1, ..., i_p\}$ of $\{1, ..., n\}$ so that the $\omega^j(x_0)$ $(1 \leq j \leq n-p)$ and the dx^{i_h} $(1 \leq h \leq p)$ form a basis of $T_{x_0}^{\vee}(M)$. Relabeling the x_h if necessary, we may assume that $i_h = h$ $(1 \leq h \leq p)$. By continuity of ω, the $\omega^j(x)$ $(1 \leq j \leq n-p)$ and the dx^h $(1 \leq h \leq p)$ form a basis of $T_x^{\vee}(M)$ for every x in a neighborhood W of x_0 in M; thus, there exist uniquely determined $A_k^j(x)$ $(1 \leq k \leq n-p)$ and $B_h^j(x)$ $(1 \leq h \leq p)$ such that, for every $k \in \{1, ..., n-p\}$,

$$dx^{p+k} = \sum_{j=1}^{n-p} A_j^k(x).\omega^j(x) + \sum_{h=1}^{p} B_h^k(x).dx^h. \quad [5.64]$$

Since the variables $x_1, ..., x_p$ are independent, the manifold $N \cap W$ is defined by the $n-p$ equations (Lemma 2.58):

$$x^{p+k} = y^k\left(x^1, ..., x^p\right), \quad (1 \leq k \leq n-p). \quad [5.65]$$

The tangent space of this manifold at the point x is therefore the set of vectors $\left(\delta x^1, ..., \delta x^p, \delta x^{p+1}, ..., \delta x^n\right)$ such that

$$\delta x^{p+k} = \sum_{h=1}^{p} \frac{\partial y^k}{\partial x^h}\left(x^1, ..., x^p, y^1, ..., y^{n-p}\right).\delta x^h,$$

where the ω^k vanish on $N \cap W$. Hence, by [5.64],

$$\boxed{\frac{\partial y^k}{\partial x^h} = B_h^k\left(x^1, ..., x^p, y^1, ..., y^{n-p}\right), \quad (1 \leq k \leq n-p, \quad 1 \leq h \leq p).} \quad [5.66]$$

Conversely, if the y^k satisfy this system of partial differential equations on W, then the ω^k satisfying [5.64] vanish on N.

This gives the following result:

LEMMA 5.84.– *The contact distribution Δ is completely integrable if and only if the system of partial differential equations [5.66] has a solution (we also say that this system "is completely integrable"). The integral manifold N is then given by [5.65].*

(V) ANALYTIC VERSION OF THE FROBENIUS THEOREM Now let $\mathbf{E}, \mathbf{F}$ be two Banach $\mathbb{K}$-spaces, U (respectively V) an open subset of $\mathbf{E}$ (respectively $\mathbf{F}$) and B a mapping of class C^{r-1} $(r \geq 2)$ from $U \times V$ into the Banach space $\mathcal{L}(\mathbf{E}; \mathbf{F})$. Then, following differential equation is said to be *total*:

$$\boxed{Dy = B(x, y)} \qquad [5.67]$$

This equation generalizes differential equations such as [1.22] (case where $\mathbf{E} = \mathbb{K} = \mathbb{R}$) and partial differential equations such as [5.66] (case where $\mathbf{E}$ and $\mathbf{F}$ are finite-dimensional). In the latter case, we can set $\mathbf{E} = \mathbb{K}^p$ and $\mathbf{F} = \mathbb{K}^{n-p}$. Then, Dy and B are represented by the Jacobian matrix $\left(\frac{\partial y^k}{\partial x^h}\right)$ and the matrix $\left(B_h^k\right)$, respectively, in the canonical bases, with $k = 1, ..., n-p$ and $h = 1, ..., p$; the equation [5.67] is now of the form [5.66].

DEFINITION 5.85.– *A differentiable mapping $\eta : U \to V$ is said to be a solution of the total differential equation [5.67] if, for every $x \in U$,*

$$D\eta(x) = B(x, \eta(x)). \qquad [5.68]$$

The equation [5.67] is said to be completely integrable *in $U \times V$ if, for every point $(x_0, y_0) \in U \times V$, there is a neighborhood U' of x_0 in U so that there exists a unique solution η of [5.67] defined in U' and taking values in V so that $\eta(x_0) = y_0$.*

THEOREM 5.86.– (analytic version of the Frobenius theorem).

1) The total differential equation [5.67] is completely integrable if and only if, for every $(x, y) \in U \times V$, the bilinear mapping

$$D_1 B(x, y) + D_2 B(x, y) \circ B(x, y) \in \mathcal{L}_2(E; F) \qquad [5.69]$$

*is symmetric (section 1.2.1**(III)**).*

2) Whenever [5.67] is of the form [5.66], the necessary and sufficient condition (1) is equivalent to the "Frobenius integrability condition"

$$\frac{\partial B_h^k}{\partial x^l} + \sum_{j=1}^{n-p} \frac{\partial B_h^k}{\partial v^j} B_l^j = \frac{\partial B_l^k}{\partial x^h} + \sum_{j=1}^{n-p} \frac{\partial B_l^k}{\partial v^j} B_h^j,$$
$$(1 \leq k \leq n-p, 1 \leq l \leq p, 1 \leq h \leq p).$$

PROOF.– 2) is clear. We will show the necessary condition of (1) here; the sufficient condition is shown later (Theorem 5.87). If $\eta : U \to V$ is differentiable and satisfies [5.68], then $D\eta$ is of class C^{r-1}, so η is of class C^r. Furthermore, for every $x \in U$,

$$D\eta(x) = D_1 B(x, \eta(x)) + D_2 B(x, \eta(x)) \circ D\eta(x).$$

But Schwarz's theorem (Theorem 1.16(i)) implies that $D\eta(x)$ is symmetric. Hence, for every $(x_0, y_0) \in U \times V$, $D_1 B(x_0, y_0) + D_2 B(x_0, y_0) \circ B(x_0, y_0)$ is symmetric. ■

(VI) GEOMETRIC VERSION OF THE FROBENIUS THEOREM Let Δ be a contact distribution on the manifold M, $\Omega(M) = \bigoplus_{p=0}^{\infty} \Omega^p(M)$ the graduated de Rham $\mathbb{K}$-algebra (section 4.4.1**(III)**) and $\mathfrak{D}(\Delta)$ the graduated ideal of $\Omega(M)$ generated by $\Gamma^{(r-1)}(M, \Delta^0)$ ([P1], section 2.3.12). Write $d(\mathfrak{D}(\Delta))$ for the $\mathbb{K}$-vector space consisting of all $d\alpha$ $(\alpha \in \mathfrak{D}(\Delta))$ and, given an open subset W of M, write $\mathfrak{D}(\Delta)|_W$ for the graduated $\Omega(W)$-ideal generated by $\Gamma^{(r-1)}(W, \Delta^0)$.

We say that $\mathfrak{D}(\Delta)$ is a *differential ideal* if $d(\mathfrak{D}(\Delta)) \subset \mathfrak{D}(\Delta)$.

Suppose that the 1-forms $\omega^1, ..., \omega^{n-p} \in \mathfrak{D}(\Delta)$ are linearly independent at every point. We say that they *locally generate* the $\Omega(M)$-ideal $\mathfrak{D}(\Delta)$ if, for every $x \in M$, there exists an open neighborhood W of x so that the restrictions $\omega^1|_W, ..., \omega^{n-p}|_W$ generate the graduated $\Omega(W)$-ideal $\mathfrak{D}(\Delta)|_W$; in other words, for every form $\alpha \in \mathfrak{D}(\Delta)|_W$, there exist forms $\theta^i \in \Omega(W)$ such that $\alpha = \sum_{i=1}^{n-p} \theta^i \wedge \omega^i|_W$.

The connection between the analytic version of the Frobenius theorem (Theorem 5.86) and the geometric version stated below is made explicit by Lemma 5.84.

THEOREM 5.87.– (geometric version of the Frobenius theorem). *Let Δ be a contact distribution of class C^{r-1}. The conditions (i) and (ii) below are equivalent. Furthermore, if M is pure and n-dimensional and [5.63] is the Pfaff system associated with Δ, where $k = p$ and the forms ω^j $(j = 1, ..., n-p)$, assumed to be linearly independent at every point, locally generate $\mathfrak{D}(\Delta)$, then the conditions (i) and (ii) are equivalent to each of the conditions (iii) to (vii), where $\omega := \omega^1 \wedge ... \wedge \omega^{n-p}$:*

i) Δ is completely integrable.

ii) Δ is involutive.

iii) $\mathfrak{D}(\Delta)$ is a differential ideal.

iv) For every $x \in M$, there exist an open neighborhood W of x and 1-forms α_k^j of class C^{r-2} over W such that

$$d\omega^j \left|_W\right. = \sum_{k=1}^{n-p} \alpha_k^j \wedge \omega^k \left|_W\right. , \quad (1 \leq j \leq n-p) .$$

v) For every $x \in M$, there exist an open neighborhood W of x and a 1-form θ of class C^{r-2} over W such that $d\omega\left|_W\right. = \theta \wedge \omega\left|_W\right.$.

vi) For every $j \in \{1, ..., n-p\}$, $d\omega^j \wedge \omega = 0$.

vii) For every $x \in M$, there exist an open neighborhood W of x and functions $f_k^j \in \mathcal{C}^{r-1}(W)$, $g^k \in \mathcal{C}^r(W)$ $(j, k \in \{1, ..., n-p\})$ such that

$$\omega^j \left|_W\right. = \sum_k f_k^j . dg^k. \qquad [5.70]$$

PROOF.– The following diagram summarizes the proof:

$$\begin{array}{ccccc} \text{(ii)} \Leftrightarrow & \text{(i)} & \Rightarrow & \text{(vii)} \\ & \Uparrow & & \Downarrow \\ & \text{(iii)} & \Longleftarrow & \text{(vi)} \\ & \Downarrow & & \Uparrow \\ & \text{(iv)} & \Longrightarrow & \text{(v)} \end{array}$$

(i)⇒(ii): This is a local condition, so we can reduce to the case $M = U \times V$, where U and V are open neighborhoods of 0 in the Banach spaces $\mathbf{E}$ and $\mathbf{F}$, respectively. Let the mapping $B : U \times V \to \mathcal{L}(\mathbf{E}; \mathbf{F})$ be of class C^{r-1} and such that the fiber of Δ over (x, y) is

$$\Delta_{(x,y)} = \{(\mathbf{u}, B(x, y) . \mathbf{u}) : \mathbf{u} \in \mathbf{E}\} \subset \mathbf{E} \times \mathbf{F}.$$

Consider two arbitrary vector fields $X_{\mathbf{u}} = (\mathbf{u}, B.\mathbf{u})$, $X_{\mathbf{v}} = (\mathbf{v}, B.\mathbf{v})$ belonging to $\Gamma^{(r-1)}(\Delta)$. Then, $DX_{\mathbf{u}}.X_{\mathbf{v}} = (0, D_1B.\mathbf{u}.\mathbf{v} + D_2B.\mathbf{v}.B.\mathbf{u})$. By [5.16], $[X_{\mathbf{u}}, X_{\mathbf{v}}] = 0$ if and only if this expression is symmetric in $\mathbf{u}$ and $\mathbf{v}$; in other words, if and only if the bilinear mapping [5.69] is symmetric. Hence, the Frobenius integrability condition (i) implies (ii), namely that Δ is involutive.

(ii) ⇒(i): Our proof argues the case where M is pure and n-dimensional (for the general case, see [LAN 99b], Chapter 6, Theorem 1.1); since this is a local question, we can assume that M is an open subset of $\mathbb{R}^n$ and that Δ is generated by (and therefore has a basis given by) the first p vectors $\mathbf{e}_i = \frac{\partial}{\partial x^i}$ $(1 \leq i \leq p)$ of the canonical basis of $\mathbb{R}^n$. Suppose that $[Y, Z] \in \Gamma^{(r-1)}(M, \Delta)$ for every $Y, Z \in \Gamma^{(r-1)}(M, \Delta)$. Let $\pi : \mathbb{R}^n \to \mathbb{R}^p$ be the canonical projection. Then, the mapping $\pi_* |_{\Delta_x}$ is an isomorphism from Δ_x onto $T_x(\mathbb{R}^p) \cong \mathbb{R}^p$ for every x in a neighborhood of 0 in M (once again, we can assume that this neighborhood is equal to M). On M, we can, therefore, find vector fields $X_1, ..., X_p$ such that $X_1(x), ..., X_p(x) \in \Delta_x$ for every $x \in M$ and $\pi_*(X_i) = \frac{\partial}{\partial x^i}\ (1 \leq i \leq p)$. By the functor property of the Lie bracket (Proposition 5.18(2)), we deduce that:

$$\pi_*\left([X_i, X_j]_x\right) = [\mathbf{e}_i, \mathbf{e}_j]_{\pi(x)} = 0,$$

so $[X_i, X_j] = 0$. Hence, by the straightening theorem for fields of frames (Theorem 5.77), there exists an open subset $U \subset M$ that we can once again assume is equal to M, as well as a diffeomorphism $\varphi : M \to N$, with $N = \varphi(M) \subset \mathbb{R}^n$, such that the differential $D\varphi$ sends X_i to $\frac{\partial}{\partial x^i}\ (1 \leq i \leq p)$. The manifold N is integral, since it is defined by the relations $x^i = const.\ (p+1 \leq i \leq n)$, or alternatively by its tangent bundle $T(N) = \left\{X \in \mathcal{T}_0^1(\mathbb{R}^n) : \langle \omega, X \rangle = 0\right\}$ with $\omega \in \Delta^0 =$ span $\left\{dx^{p+1}, ..., dx^n\right\}$.

(iii)⇒(i): We only need to consider 1-forms, for which the Maurer–Cartan formula holds (Theorem 5.22(iv)). Consider a 1-form $\omega \in \mathfrak{O}(\Delta)$. Then, $\langle d\omega, X \wedge Y \rangle = 0$ for every $X, Y \in \Gamma^{(r-1)}(M, \Delta)$ if and only if $\langle \omega, [X, Y] \rangle = 0$, i.e. $[X, Y] \in \Gamma^{(r-2)}(M, \Delta)$.

(iii)⇒(iv): The exterior products $\omega^j \wedge dx^h, \omega^j \wedge \omega^k, dx^l \wedge dx^h$ $(1 \leq j \leq k \leq n-p, 1 \leq l < h \leq p)$ form a basis of differential 2-forms in W. For $1 \leq j \leq n-p$, we can, therefore, write that:

$$d\omega^j = \sum_{k,h} C^j_{kh} \omega^k \wedge dx^h + \sum_{i,k} D^j_{ik} \omega^i \wedge \omega^k + \sum_{h,l} E^j_{hl} dx^l \wedge dx^h,$$

where the coefficients are of class C^{r-2} in W. Hence, (iii) is equivalent to saying that the E^j_{hl} are all zero, which implies (and is even equivalent to) (iv).

(iv)⇒(v): If β is a 1-form and γ is a 2-form, then $\beta \wedge \gamma = \gamma \wedge \beta$. Hence, by [5.22], $d\omega = \sum_{j=1}^{n-p} d\omega^j \wedge \omega^1 \wedge ... \wedge \widehat{\omega^j} \wedge ... \wedge \omega^{n-p}$ and (iv) implies (v) with $\theta = \sum_{j=1}^{n-p} (-1)^j \alpha^j_j$.

(v)⇒(vi) is clear.

(vii)⇒(vi): If (vii) is satisfied, then, for every $i \in \{1, ..., n-p\}$, $d\omega^j = \sum_k df^j_k \wedge dg^k$, so $d\omega^j \wedge \omega = 0$.

(i)⇒(vii): Suppose that (i) holds. Then, there exists (Corollary and Definition 2.42) a chart (U, ξ, n) of M for which the integral submanifold N is such that $N \cap U$ has local coordinates $\xi^{p+1} = ... = \xi^n = 0$. The 1-forms $d\xi^{p+1}, ..., d\xi^n$ therefore form a basis of the $\mathcal{C}^r(U)$-module $\Gamma^{(r-1)}(U, \Delta^0)$. Hence, the 1-forms ω^j $(1 \leq j \leq n-p)$ can be expressed in the form [5.70] with $g^k = \xi^{p+k}$.

(vi)⇒(iii): This is a consequence of the following algebraic lemma:

If E is a vector space of dimension n and $\omega_1, ..., \omega_r$ are r linearly independent vectors of E, then an element $\alpha \in \bigwedge E$ of the exterior algebra of E (section 4.2.3(**III**)) belongs to the ideal $\mathfrak{I}$ generated by $\omega_1, ..., \omega_r$ if and only if $\omega_1 \wedge ... \wedge \omega_r \wedge \alpha = 0$.

Indeed, let $\omega_{r+1}, ..., \omega_n \in E$ be such that $\omega_1, ..., \omega_n$ form a basis of E. Since $\mathfrak{I}$ is the direct sum of its homogeneous components, it is sufficient to consider the case where ω is homogeneous of degree k, i.e.

$$\omega = \sum_{1 \leq i_1 \leq ... \leq i_k \leq n} a_{i_1,...,i_k} \omega_{i_1} \wedge ... \wedge \omega_{i_k}$$

Thus, α belongs to $\mathfrak{I}$ if and only if $a_{i_1,...,i_k} = 0$ for $i_1 > r$, which is equivalent to $\omega_1 \wedge ... \wedge \omega_r \wedge \alpha = 0$. ■

EXAMPLE 5.88.– *Let M be a non-empty open subset of $\mathbb{R}^3$ and $\omega = P(x,y,z).dx + Q(x,y,z).dy + R(x,y,z).dz$, where $P, Q, R : M \to \mathbb{R}$ are of class C^1. Then:*

$$d\omega = \frac{\partial P}{\partial y} dy \wedge dx + \frac{\partial P}{\partial z} dz \wedge dx + \frac{\partial Q}{\partial x} dx \wedge dy + \frac{\partial Q}{\partial z} dz \wedge dy + \frac{\partial R}{\partial x} dx \wedge dz + \frac{\partial R}{\partial y} dy \wedge dz,$$

so

$$\omega \wedge d\omega = \left[P\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + Q\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + R\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right] dx \wedge dy \wedge dz,$$

and hence the Pfaff equation $\omega = 0$ is completely integrable if and only if

$$P\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + Q\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + R\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = 0. \quad [5.71]$$

This result can be interpreted as follows: by the condition (vii) of Theorem 5.87, the Pfaff equation is completely integrable if and only if, for every $x \in M$, there exist an open neighborhood W of x and functions $f \in \mathcal{C}^1(W), g \in \mathcal{C}^2(W)$ so that $\omega = f.dg$. The case where $f = 0$ is trivial. Setting this case aside, we can assume that

$f \neq 0$ in W, since f is continuous; this function f is said to be an "integrating factor"*. Thus, $\omega = 0 \Leftrightarrow dg = 0$. The desired contact distribution, if it exists, is therefore an "equipotential" $dg = 0 \Leftrightarrow g = const$. Its existence is conditional on the fact that the vector field X with components $\varphi P, \varphi Q, \varphi R$, where $\varphi = 1/f$, "derives from a potential", i.e. is a gradient. But this holds if and only if $curl X = 0$ by Example 5.53 and Poincaré's lemma ([P1], section 3.3.8**(VII)**, Theorem 3.184), choosing W to be simply connected. Hence, there exists a solution to $\omega = 0$ if and only if there exists a function $\varphi \neq 0$ de classe C^1 for which curl $(\varphi Y) = 0$, where Y is the field of vectors with components P, Q, R. But curl $(\varphi Y) = \nabla \wedge (\varphi Y) = (\nabla \varphi) \wedge Y + \varphi . (\nabla \wedge Y)$. Therefore, it is necessary to have $(\nabla \varphi) \wedge Y = -\varphi . (\nabla \wedge Y)$; given that $(\nabla \varphi) \wedge Y$ is orthogonal to Y, it must be the case that Y is orthogonal to $\nabla \wedge Y =$ curl (Y), i.e. satisfies [5.71]. The Frobenius theorem shows that this necessary condition is also sufficient.*

Suppose that W is simply connected. Then, by Poincaré's lemma, ω is an exact total differential, i.e. $\omega = dg$, where $g \in \mathcal{C}^2(W)$, if and only if $d\omega = 0$, i.e. $\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0$, $\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = 0$, $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0$. This condition is more restrictive than [5.71] due to the absence of any integrating factor.

(VII) FOLIATIONS

DEFINITION 5.89.– *Let M be a locally finite-dimensional manifold and Φ a submanifold of M. We say that Φ is a* foliation *of M if (i) every point of M belongs to Φ and (ii) for every $x \in M$, there exists a chart (U, ξ, n) centered on x, with $\xi(U) =]-\varepsilon, \varepsilon[^n$ $(\varepsilon > 0)$, so that the connected components of $\Phi \cap U$ are sets of the form*

$$\left\{y \in U : \xi^{k+1}(x) = a^{k+1}, ..., \xi^n(x) = a^n\right\}, \quad \left(a^i = const., \quad i = k+1, ..., n\right),$$

said to be slices *of U. A connected component of Φ_a $(a \in \mathcal{A})$ is said to be a* leaf *of Φ, and $\Phi = \dot{\bigcap}_{a \in \mathcal{A}} \Phi_a$ (where $\dot{\bigcap}$ denotes the disjoint union). The subbundle $T(M, \Phi)$ of $T(M)$ defined by $T(M, \Phi) = \bigcup_{a \in \mathcal{A}} T(\Phi_a)$ is said to be the* tangent bundle of the foliation Φ*, and the quotient bundle $T(M)/T(M, \Phi)$ (Section 3.4.3) is said to be the* normal bundle *of this foliation.*

Each leaf is an immersed submanifold of M (Lemma-Definition 2.43, section 2.3.4), but this immersion is not an embedding in general. A very simple example of foliations is as follows: if $M = \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$, we obtain a foliation of M by considering the leaves Φ_a defined by $x^{k+1} = a^{k+1}, ..., x^n = a^n$, each isomorphic to $\mathbb{R}^k$, with $a = \left(a^{k+1}, ..., a^n\right) \in \mathbb{R}^{n-k}$, and setting $\Phi = \dot{\bigcap}_{a \in \mathbb{R}^{n-k}} \Phi_a$ (Figure 5.5).

The notion of a foliation can be extended to the case where M is a Banach manifold ([ABR 83], Definition 4.4.4):

DEFINITION 5.90.– *Let M be a $\mathbb{K}$-manifold of class C^r $(r \in \mathbb{N}_{\mathbb{K}}, r \geq 2)$ and $\Phi = (\Phi_a)_{a \in \mathcal{A}}$ a partition of M into connected subsets, called leaves. The partition Φ is said to be a* foliation *of M if, for each point $x \in M$, there exists a chart $(U, \xi, \mathbf{E})$ centered on x, where $\mathbf{E} = \mathbf{E}_1 \oplus \mathbf{E}_2$, the subspaces $\mathbf{E}_1$, $\mathbf{E}_2$ split in the Banach space $\mathbf{E}$ ([P1], section 3.2.2(**IV**)), $\varphi(U) \cap \mathbf{E}_1 = U_1$, $\varphi(U \cap \mathbf{E}_2) = U_2$, and, for every $a \in \mathcal{A}$, the connected components $(U \cap \Phi_a)_\gamma$ of $U \cap \Phi_a$ are given by $\varphi\left((U \cap \Phi_a)_\gamma\right) = U_1 \times \{c_{a,\gamma}\}$, where the $c_{a,\gamma} \in \mathbf{E}_2$ are constants. Any such chart is said to be* foliated.

THEOREM 5.91.– (global version of the Frobenius theorem) *Let Δ be a contact distribution on a $\mathbb{K}$-manifold M of class C^r. The following conditions are equivalent:*

i) Δ is integrable.

ii) There exists a foliation Φ of M such that $\Delta = T(M, \Phi)$.

iii) For any choice of open subset U of M and the vector fields $X, Y \in \Gamma^{(r-1)}(U, \Delta)$, we have $[X, Y] \in \Gamma^{(r-2)}(U, \Delta)$.

PROOF.– If M is finite-dimensional, we can cover it by a family of charts (U_i, z_i, n_i) in such a way that, on each U_i, we have the situation from Example 5.83. The same reasoning can be extended to the case where M is a Banach manifold ([ABR 83], Theorem 4.4.7). ■

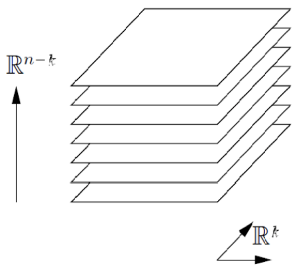


Figure 5.5. *Foliation of $\mathbb{R}^n$*

This page intentionally left blank

6

Analysis on Lie Groups

6.1. Introduction

Lie groups generally have a richer structure than manifolds, which allows us to perform additional operations. The first such operation is the convolution product (historically also known as the composition product). This idea is omnipresent in various engineering sciences, especially signal theory and automatic systems. As an example, the output of a stationary linear filter is the convolution of the input of this filter and its impulse response.

Given a locally compact topological group $\mathbf{G}$, there exists a left invariant measure and a right invariant measure on $\mathbf{G}$, distinct except when $\mathbf{G}$ is "unimodular" and in particular when $\mathbf{G}$ is commutative, but unique up to multiplication by a real number > 0. These measures, called the left and right Haar measures, play an analogous role to the Lebesgue measure on $\mathbb{R}$ or $\mathbb{R}^n$ (which they generalize). They allow us to define the convolution product of two functions. The existence of these invariant measures was proved by Hurwitz for a Lie group $\mathbf{G}$ in 1897; Haar generalized the result in 1933 by showing that these invariant measures still exist when $\mathbf{G}$ is a locally compact separable topological group. In 1938, A. Weil proved that we can omit the separability condition [WEI 38]. However, by solving Hilbert's fifth problem[1], Gleason, Montgomery and Zippin established in 1952 that every locally compact, metrizable and locally connected topological group is a Lie group[2]. Hence, retrospectively, Hurwitz's result was very close to optimal. This is the result that we will present below, with a simple and concise proof.

The Lie algebras can be classified very exhaustively (section 6.3). Thanks to a "dictionary" that was primarily established by S. Lie (section 6.4.3), so can the Lie

1 See the Wikipedia article on *Hilbert's fifth problem*.

2 There is also a theory of distributions on locally compact groups [BRU 61].

groups. We will present a modern reformulation of the three “fundamental theorems” of S. Lie (the original versions can be found in the historical notes in [BOU 82b], Chapter 3).

Harmonic analysis on Lie groups (section 6.5) is one of the major fields of mathematics, derived from analysis; it has also crept into other areas, especially the “p-adic Lie groups” ([BOU 82b], Chapter 3), [BRU 61, SCH 11] in number theory [WEI 74]; however, these topics exceed the scope of this book, so our presentation will be limited to harmonic analysis on real Lie groups, primarily in the commutative case. An overview and a few historical notes are given in section 6.5.1.

6.2. Convolution

In this section, unless otherwise stated, every Lie group $\mathbf{G}$ is real and locally compact; the neutral element is written as e and $\mathbf{G}$ is assumed to be generated by any given neighborhood of e, so $\mathbf{G}$ is also countable at infinity (Lemma 2.80).

6.2.1. *Convolution of distributions*

(I) Convolution of functions on the real line Let f, g be two locally Lebesgue-integrable functions on the real line ([P2], section 4.1.5(**III**)) and consider

$$h(x) = \int_{\mathbb{R}} f(y)\, g(x-y)\, .dy.$$

If this integral converges, we write that $h = f \star g$. The change of variable $x - y = y'$ gives $h(x) = \int_{\mathbb{R}} g(y')\, .f(x-y')\, dy'$, so $f \star g = g \star f$. We can give a *sufficient* condition for h to be a locally integrable function defined almost everywhere. Let $\varphi \in \mathcal{K}(\mathbb{R})$, i.e. φ is a continuous function with compact support K, and suppose that:

$$\begin{aligned}\mu_h(\varphi) &= \int_{\mathbb{R}} h(x)\, .\varphi(x)\, .dx = \iint f(y)\, .g(x-y)\, .dy.dx \\ &= \iint f(y)\, .g(z)\, .\varphi(y+z)\, .dy.dz.\end{aligned}$$

This integral converges if, setting $\varphi^{\Delta}(y,z) = \varphi(y+z)$ and $K^{\Delta} = \operatorname{supp}(\varphi^{\Delta})$, the condition (**S**) is satisfied:

(S) For every compact set $K \subset \mathbb{R}$, $(\operatorname{supp}(f) \times \operatorname{supp}(g)) \cap K^{\Delta}$ is compact.

If so, $h.\varphi$ is integrable and $h(x)$ is defined for every x such that $\varphi(x) \neq 0$. This is true for every function $\varphi \in \mathcal{K}(\mathbb{R})$, so h is a locally integrable function defined almost everywhere. We say that the functions f, g are *strictly convolvable* and that $f \star g$ is their convolution product (or simply their convolution).

If either f or g is compactly supported, then the condition **(S)** is satisfied. Suppose that f and g have left bounded support, i.e. there exist two real numbers a, b such that $f(x) = 0$ if $x < a$ and $g(x) = 0$ if $x < b$. Then:

$$h(x) = \int_a^{x-b} f(y) . g(x-y) . dy,$$

so h is a locally integrable function defined almost everywhere with left bounded support, because $h(x) = 0$ for $x < a + b$. If f and g are both compactly supported, then h is compactly supported.

(II) GENERALIZATION Let us now extend the situation described above to the case where the real line is replaced by a Lie group $\mathbf{G}$ and the functions f, g are distributions on $\mathbf{G}$.

Let $\varphi \in \mathcal{E}(\mathbf{G})$ and $\boldsymbol{m} : \mathbf{G} \times \mathbf{G} \to \mathbf{G} : (x, y) \mapsto x.y$. In general, even if φ is compactly supported, $\varphi^{\Delta} =: \varphi \circ \boldsymbol{m} : \mathbf{G} \times \mathbf{G} \to \mathbb{C}$ is not, because, if $K = \operatorname{supp}(\varphi)$, then $\operatorname{supp}(\varphi^{\Delta}) = K^{\Delta}$, where $K^{\Delta} := \{(x, y) \in \mathbf{G} \times \mathbf{G} : x.y \in K\}$. Let $T, S \in \mathcal{D}'(\mathbf{G})$. By section 5.2.1**(VII)**, we can form the quantity

$$\boxed{\langle T \star S, \varphi \rangle = \langle T \otimes S, \varphi^{\Delta} \rangle} \qquad [6.1]$$

whenever $(\operatorname{supp}(T) \times \operatorname{supp}(S)) \cap K^{\Delta}$ is compact, which generalizes the condition **(S)**. We say that $T \star S$ is the convolution of T and S.

DEFINITION 6.1.– *We say that a finite family $(A_i)_{1 \leq i \leq h}$ of closed subsets of $\mathbf{G}$ is* admissible for convolution *if, for every compact subset K of $\mathbf{G}$, the set $\left(\prod_{i=1}^{h} A_i\right) \cap K^{\Delta}$ is compact. A finite family of distributions $(T_i)_{1 \leq i \leq h}$ is said to be* strictly convolvable *if the family $(\operatorname{supp}(T_i))_{1 \leq i \leq h}$ is admissible for convolution. The convolution product $T_1 \star ... \star T_h$ is then recursively defined by $T_1 \star ... \star T_{j+1} = (T_1 \star ... \star T_j) \star T_{j+1}$ $(1 \leq j \leq h-1)$.*

THEOREM 6.2.– *i) If $(T_i)_{1 \leq i \leq h}$ is a family of h distributions, all compactly supported except for at most one, then this family is strictly convolvable (**exercise***: see [DIE 93], Volume 2, (14.6.4)). If these distributions are all compactly supported, then $T_1 \star ... \star T_h$ is compactly supported.*

*ii) Suppose that $\mathbf{G} = \mathbb{R}$. If $(T_i)_{1 \leq i \leq h}$ is a family of h distributions, all with left bounded support, then this family is strictly convolvable and $T_1 \star ... \star T_h$ has left bounded support (**exercise**).*

COROLLARY 6.3.– *i) The space $\mathcal{E}'(\mathbf{G})$ of compactly supported distributions equipped with the convolution product is an associative and unitary algebra (said to be a* convolution algebra*) ([P1], section 2.3.10(**I**)) whose neutral element is the Dirac distribution $\delta_e : \varphi \mapsto \varphi(e)$ at the point e (section 5.2.2(**III**)). This algebra is commutative if and only if $\mathbf{G}$ is commutative.*

ii) The space $\mathcal{D}'_+(\mathbb{R})$ of positively supported distributions on $\mathbb{R}$ (i.e. whose support is contained in $\mathbb{R}_+$) equipped with the convolution product is an associative, unitary and commutative algebra.

PROOF.– We only need to show (i). It is immediately clear that the mapping $(S,T) \mapsto S \star T$ is bilinear. Let $S \in \mathcal{E}'(G)$ and $\varphi \in \mathcal{E}(G)$. With the notation of section 5.2.1(**VIII**), $\langle S(x) \otimes \delta_e(y), \varphi(x.y)\rangle = \langle S(x), \varphi(x.e)\rangle = \langle S, \varphi\rangle$, so $S \otimes \delta_e = S$. Similarly, $\delta_e \otimes S = S$. Take $S = \delta_s, T = \delta_t$; then, $\langle S \star T, \varphi\rangle = \varphi(s.t)$ and $\langle T \star S, \varphi\rangle = \varphi(t.s)$. If $(\mathcal{E}'(\mathbf{G}), \star)$ is a commutative algebra, then $\varphi(s.t) = \varphi(t.s)$ for every function $\varphi \in \mathcal{E}(\mathbf{G})$, so $s.t = t.s$ by Theorem 2.13 and Corollary 2.17. ■

REMARK 6.4.– *1) The bilinear mapping $(T,S) \mapsto T \star S$ is separately continuous from $\mathcal{D}'(\mathbf{G}) \times \mathcal{E}'(\mathbf{G})$ into $\mathcal{D}'(\mathbf{G})$, from $\mathcal{E}'(\mathbf{G}) \times \mathcal{E}'(\mathbf{G})$ into $\mathcal{E}'(\mathbf{G})$ and from $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$. By [P2], section 3.9.1(**II**), Theorem 3.137, it is therefore continuous from $\mathcal{E}'(\mathbf{G}) \times \mathcal{E}'(\mathbf{G})$ into $\mathcal{E}'(\mathbf{G})$ and from $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$, since $\mathcal{E}'(\mathbf{G})$ is a Silva space and hence a barreled $(\mathcal{DF})$ space ([P2], section 4.4.1(**I**), Theorem 4.89 and Remark 4.90) and $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space ([P2], section 4.3.1(**III**), Theorem 4.72).*

*2) We can embed the space of Radon measures on $\mathbf{G}$ into the space of distributions on $\mathbf{G}$ ([P2], section 4.4.1(**II**)). Hence, the above enables us to define the convolution product of two measures. It then suffices to assume that $\mathbf{G}$ is a locally compact* topological *group that is countable at infinity.*

(III) RELATION WITH THE DERIVATION With the notation from section 5.2.1(**VIII**), we have:

$$\boxed{D_\xi^\alpha \delta_e \star T = D_\xi^\alpha T.} \qquad [6.2]$$

PROOF.– For every test function $\varphi \in \mathcal{D}(\mathbf{G})$,

$$\begin{aligned}\left\langle D_\xi^\alpha \delta_e \star T, \varphi\right\rangle &= \left\langle D_\xi^\alpha \delta_e(x) \otimes T(y), \varphi(x.y)\right\rangle = \left\langle T(y), \left\langle D_\xi^\alpha \delta_e(x), \varphi(x.y)\right\rangle\right\rangle \\ &= \left\langle T(y), \left\langle (-1)^{|\alpha|} D_\xi^\alpha \varphi(y)\right\rangle\right\rangle = \left\langle D_\xi^\alpha T, \varphi\right\rangle.\end{aligned}$$ ■

(IV) RELATION WITH TRANSLATIONS Consider the left and right translations $\boldsymbol{\lambda}(s)$ and $\boldsymbol{\rho}(s)$, respectively (section 2.4.1(**I**)). For every function $\varphi : \mathbf{G} \to \mathbb{C}$, set $\boldsymbol{\lambda}(s)\varphi : \boldsymbol{\lambda}(s).x \mapsto \varphi(x)$ and $\boldsymbol{\rho}(s)\varphi : \boldsymbol{\rho}(s).x \mapsto \varphi(x)$, which gives:

$$\boldsymbol{\lambda}(s)\varphi : x \mapsto \varphi\left(s^{-1}.x\right), \quad \boldsymbol{\rho}(s)\varphi : x \mapsto \varphi(x.s).$$

LEMMA 6.5.– *Let T be a distribution. For every test function $\varphi \in \mathcal{D}(G)$, set $\langle\boldsymbol{\lambda}(s)T, \varphi\rangle = \left\langle T, \boldsymbol{\lambda}\left(s^{-1}\right)\varphi\right\rangle$, $\langle\boldsymbol{\rho}(s)T, \varphi\rangle = \left\langle T, \boldsymbol{\rho}\left(s^{-1}\right)\varphi\right\rangle$. Then:*

$$\boxed{\boldsymbol{\lambda}(s)T = \delta_s \star T, \quad \boldsymbol{\rho}(s)T = T \star \delta_{s^{-1}}.} \qquad [6.3]$$

PROOF.– For every test function $\varphi \in \mathcal{D}(\mathbf{G})$,

$$\begin{aligned}\langle\boldsymbol{\lambda}(s)T, \varphi\rangle &= \left\langle T, \boldsymbol{\lambda}\left(s^{-1}\right)\varphi\right\rangle = \langle T(x), \varphi(s.x)\rangle \\ &= \langle\delta(s)\otimes T(x), \varphi(y.x)\rangle = \langle\delta_s \star T, \varphi\rangle.\end{aligned}$$

A similar reasoning works for the second equality. ■

REMARK 6.6.– *1) The convolution algebra $(\mathcal{E}'(\mathbf{G}), \star)$ is not integral in general. For example, if $\mathbf{G}$ is the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and C is a non-zero constant, then $0 \neq C \in \mathcal{E}'(\mathbb{T})$, $\dot{\delta}_0 \neq 0$ and $\delta_0 \star C = 0$. However, the convolution algebras $(\mathcal{E}'(\mathbb{R}^n), \star)$ and $\left(\mathcal{D}'_+(\mathbb{R}), \star\right)$ are integral; for the first, this follows from the Paley–Wiener–Schwartz theorem ([DIE 93], (22.18.7)), and for the second, this follows from a theorem by Titchmarsh ([KHO 72], Volume 2, p. 127, Exercise 7).*

2) If $\mathbf{G}$ is a locally compact topological group that is countable at infinity and T is a Radon measure on $\mathbf{G}$, the equalities [6.3] remain valid.

(V) CASE OF POINT DISTRIBUTIONS The $\mathcal{E}(\mathbf{G})$-modules $\mathcal{T}_g^{(\infty)}(\mathbf{G})$ and $\mathcal{T}^{(\infty)}(\mathbf{G})$ formed by the distributions with support contained in $\{g\}$ and the point distributions, respectively (section 5.2.2(**III**)), are convolution subalgebras of $(\mathcal{E}'(\mathbf{G}), \star)$. The first is clearly commutative; the second is commutative if and only if $\mathbf{G}$ is commutative, and its neutral element is δ_e.

In the case where $\mathbf{G}$ is a Banach or non-Banach Lie $\mathbb{K}$-group, the convolution product $T \otimes S$ of two-point distributions $T, S \in \mathcal{T}^{(\infty)}(\mathbf{G})$ (section 5.2.2, Remark 5.7) can no longer be defined using [6.1]. However, the following definition still works, since $\boldsymbol{m}$ is a proper mapping:

DEFINITION 6.7.– *Let $T, S \in \mathcal{T}^{(\infty)}(\mathbf{G})$. Their convolution $S \star T$ is $\boldsymbol{m}(S \otimes T)$ (section 5.2.1(**VIII**)).*

If $\mathbf{G}$ is locally compact and countable at infinity and $\varphi \in \mathcal{E}(\mathbf{G})$, then $\langle \boldsymbol{m}(S \otimes T), \varphi \rangle = \langle S \otimes T, \varphi \circ \boldsymbol{m} \rangle$ and we recover the definition from [6.1].

(VI) ALGEBRAIC PARENTHESES Let $\mathbf{G}$ be a group, $\mathbf{K}$ a commutative ring, and consider the $\mathbf{K}$-module $\mathbf{K}^{(\mathbf{G})}$ consisting of the finitely supported families $(\alpha_g)_{g \in \mathbf{G}}$ of elements of $\mathbf{K}$ ([P1], sections 1.6.2**(II)**, 2.2.1**(IV)**). This module has the structure of a $\mathbf{K}$-algebra equipped with the multiplication $(\alpha_g) \star (\beta_h) = (\gamma_s)$, where $\gamma_s = \sum_{g \in \mathbf{G}} \alpha_g . \beta_{g^{-1}.s}$; we say that it is *the algebra of the group* $\mathbf{G}$ *over* $\mathbf{K}$.

Let $\mathcal{T}^{(0)}(\mathbf{G})$ be the set of point distributions of order 0 on a Lie group $\mathbf{G}$ (section 5.2.2**(III)**), i.e. the set of Radon point measures. Then, $\mathcal{T}^{(0)}(\mathbf{G}) = \bigoplus_{g \in \mathbf{G}} \mathcal{T}_g^{(0)}(\mathbf{G})$ is a convolution subalgebra of $\mathcal{T}^{(\infty)}(\mathbf{G})$. An element $T \in \mathcal{T}^{(0)}(\mathbf{G})$ can be uniquely written in the form $T = \sum_{g \in \mathbf{G}} \alpha_g . \delta_g$, where $(\alpha_g) \in \mathbb{K}^{(\mathbf{G})}$ is a finitely supported sequence of elements of $\mathbb{K}$, whenever this Radon measure takes values in $\mathbb{K}$. We have $\delta_g \star \delta_s = \delta_{gs}$, so

$$\left(\sum_{g \in \mathbf{G}} \alpha_g . \delta_g \right) \left(\sum_{h \in \mathbf{G}} \beta_h . \delta_h \right) = \sum_{s \in \mathbf{G}} \left(\sum_{g \in \mathbf{G}} \alpha_g . \beta_{g^{-1}.s} \right),$$

and hence $\mathcal{T}^{(0)}(\mathbf{G})$ is isomorphic to (and can be identified with) the algebra of the group $\mathbf{G}$ over $\mathbb{K}$.

6.2.2. *Haar measure and convolution of functions*

(I) HAAR MEASURE Let $\mathbf{G}$ be an n-dimensional Lie group.

LEMMA 6.8.– *There exists a unique positive non-zero measure μ_l (respectively μ_r) on $\mathbf{G}$ that is left invariant (respectively right invariant), i.e. which satisfies $\boldsymbol{\lambda}(s)\mu_l = \mu_l$ (respectively $\boldsymbol{\rho}(s)\mu_r = \mu_r$) for every $s \in \mathbf{G}$, and every other positive non-zero left invariant (respectively right invariant) measure is of the form $a.\mu_l$ (respectively $a.\mu_r$), $a > 0$.*

PROOF.– We will give the proof for μ_l. Let $\omega_e \in \bigwedge^n T_e^{\vee}(\mathbf{G})$ be a non-zero alternating n-linear form (section 4.2.4) and, for every $g \in \mathbf{G}$, define the preimage $\omega_g = \boldsymbol{\lambda}^*(g^{-1}).\omega_e$ in the same way as section 4.4.2, i.e. for every n-tuple of tangent vectors $(v_1, ..., v_n) \in T_g(\mathbf{G})^n$,

$$\omega_g.(v_1, ..., v_n) = \omega_e.\left(T_g(\boldsymbol{\lambda}_{g^{-1}}).v_1, ..., T_g(\boldsymbol{\lambda}_{g^{-1}}).v_n\right),$$

writing $\boldsymbol{\lambda}_{g^{-1}}$ for $\boldsymbol{\lambda}\left(g^{-1}\right)$. Then, the mapping $\omega : \mathbf{G} \to \bigwedge^n T^\vee(\mathbf{G}) : g \mapsto \omega_g$ is a complex differential n-form that belongs to $\Omega^n(\mathbf{G};\mathbb{C})$ (section 4.4.3(**VI**)). For every $g, h \in \mathbf{G}$,

$$\left(\boldsymbol{\lambda}_g^*.\omega\right)_h = \boldsymbol{\lambda}_g^*.\omega_{gh} = \boldsymbol{\lambda}_g^*\left(\boldsymbol{\lambda}_{(gh)^{-1}}^*\right).\omega_e = \boldsymbol{\lambda}_g^*.\boldsymbol{\lambda}_{h^{-1}g^{-1}}^*.\omega_e$$
$$= \left(\boldsymbol{\lambda}_{h^{-1}g^{-1}}.\boldsymbol{\lambda}_g\right)^*.\omega_e = \boldsymbol{\lambda}_{h^{-1}}^*.\omega_e = \omega_h,$$

so ω is a left invariant measure that is positive if a suitable orientation is chosen on $\mathbf{G}$ (section 4.4.4(**II**), Example 4.38(ii)). Since $\dim_{\mathbb{R}}\left(\bigwedge^n T_e^*(\mathbf{G})\right) = 1$, every other non-zero element of $\bigwedge^n T_e^*(\mathbf{G})$ is of the form $C.\omega_e$, where C is a non-zero real number. The measure $C.\omega$ is positive if and only if $C > 0$. ■

DEFINITION 6.9.– *The measure μ_l (respectively μ_r), which is unique up to multiplication by a strictly positive real number, is called a left (respectively right)* Haar measure *of the group* $\mathbf{G}$.

COROLLARY 6.10.– *We have supp (μ_l) = supp (μ_r) = $\mathbf{G}$. If V is a compact neighborhood of e, then $\mu_l(V) > 0$ and $\mu_r(V) > 0$.*

PROOF.– Since μ_l is non-zero and left invariant, supp $(\boldsymbol{\lambda}(s).\mu_l) = s.$supp (μ_l) and supp $(\mu_l) \neq \varnothing$, so supp $(\mu_l) = \mathbf{G}$. Let V be a compact neighborhood of e and let $U = V \cap V^{-1}$. Then, U is another compact neighborhood of e, $U = U^{-1}$ and $\mathbf{G} = U^\infty$ (see the start of section 6.2); therefore, if we had $\mu_l(V) = 0$, we would also have $\mu_l(\mathbf{G}) = 0$ ([P2], section 4.1.1(**II**)), which is impossible because $\mu_l \neq 0$. ■

LEMMA-DEFINITION 6.11.– *There exists a morphism of groups $\Delta_{\mathbf{G}} : \mathbf{G} \to \mathbb{R}_+^* : s \mapsto \Delta_{\mathbf{G}}(s)$ such that $\boldsymbol{\rho}(s)\mu_l = \Delta_{\mathbf{G}}(s)\mu_l$ for every $s \in \mathbf{G}$. This morphism is independent of the choice of left Haar measure μ_l and is analytic. It is called the* modulus function *of* $\mathbf{G}$. *If $\Delta_{\mathbf{G}}(s) = 1$ for every $s \in \mathbf{G}$, the Lie group $\mathbf{G}$ is said to be* unimodular.

PROOF.– For every $s, t \in \mathbf{G}$,

$$\boldsymbol{\lambda}(t)\left(\boldsymbol{\rho}(s).\mu_l\right) = \boldsymbol{\rho}(s)\left(\boldsymbol{\lambda}(t).\mu_l\right) = \boldsymbol{\rho}(s).\mu_l,$$

so $\boldsymbol{\rho}(s)\mu_l$ is another left invariant positive measure. Hence, by Lemma 6.8, there exists a number $\Delta_{\mathbf{G}}(s) > 0$ independent of μ_l such that $\boldsymbol{\rho}(s).\mu_l = \Delta_{\mathbf{G}}(s).\mu_l$. It is clear that $\Delta_{\mathbf{G}}$ is a morphism of groups from $\mathbf{G}$ into $\mathbb{R}_+^*$; this morphism is analytic by Corollary-Definition 4.42(iii), since $r = \omega$. ■

NOTATION 6.12.– *Given a function $\varphi \in \mathcal{K}_{\mathbb{C}}(\mathbf{G})$ and a Radon measure μ, write $\check{\varphi}$ and $\check{\mu}$ for their respective images under the diffeomorphism $s \mapsto s^{-1}$, so that $\check{\varphi}(s) = \varphi\left(s^{-1}\right)$ and $\langle\check{\mu},\varphi\rangle = \langle\mu,\check{\varphi}\rangle$. The same convention is adopted for arbitrary functions or distributions.*

THEOREM 6.13.– *1)* $\check{\mu}_l = \Delta_{\mathbf{G}}^{-1}.\mu_l$ *is a right Haar measure* μ_r.

2) If there exists a compact neighborhood V *of* e *that is invariant under every inner automorphism of* $\mathbf{G}$ *([P1], section 2.2.2(**III**)), then* $\mathbf{G}$ *is unimodular. For this to be the case, it suffices for* $\mathbf{G}$ *to be commutative, compact or discrete (section 2.4.1, Example 2.74(i)).*

PROOF.– 1) For every $s \in \mathbf{G}$,

$$\boldsymbol{\rho}(s)\left(\Delta_{\mathbf{G}}^{-1}.\mu_l\right) = \left(\boldsymbol{\rho}(s)\,\Delta_{\mathbf{G}}^{-1}\right).\boldsymbol{\rho}(s).\mu_l$$
$$= \left(\Delta_{\mathbf{G}}^{-1}.\Delta_{\mathbf{G}}^{-1}\right).\left(\Delta_{\mathbf{G}}.\mu_l\right) = \Delta_{\mathbf{G}}^{-1}.\mu_l,$$

so $\Delta_{\mathbf{G}}^{-1}.\mu_l$ is a right Haar measure. Furthermore, for every function $\varphi \in \mathcal{K}_{\mathbb{C}}(\mathbf{G})$,

$$\langle \boldsymbol{\rho}(s).\check{\mu}_l, \varphi\rangle = \langle \check{\mu}_l, \boldsymbol{\rho}(s).\varphi\rangle = \left\langle \mu_l, (\boldsymbol{\rho}(s).\varphi)^{\vee}\right\rangle,$$

where $(\boldsymbol{\rho}(s).\varphi)^{\vee}(g) := (\boldsymbol{\rho}(s).\varphi)\left(g^{-1}\right) = \varphi\left(g^{-1}s\right) = \check{\varphi}\left(s^{-1}g\right) = \boldsymbol{\lambda}(s).\check{\varphi}(g)$, so $(\boldsymbol{\rho}(s).\varphi)^{\vee} = \boldsymbol{\lambda}(s).\check{\varphi}$. Hence,

$$\left\langle \mu_l, (\boldsymbol{\rho}(s).\varphi)^{\vee}\right\rangle = \langle \mu_l, \boldsymbol{\lambda}(s).\check{\varphi}\rangle = \langle \boldsymbol{\lambda}(s).\mu_l, \check{\varphi}\rangle = \langle \mu_l, \check{\varphi}\rangle = \langle \check{\mu}_l, \varphi\rangle,$$

so $\check{\mu}_l$ is a right Haar measure. Thus, there exists a real number $C > 0$ such that $\check{\mu}_l = C.\Delta_{\mathbf{G}}^{-1}.\mu_l$. But $\mu_l = (\check{\mu}_l)^{\vee}$, so $C = 1$.

2) For every μ_l-integrable function f and every $s \in \mathbf{G}$,

$$\int f(gs).d\mu_l(g) = \int (\boldsymbol{\rho}(s).f)(g).d\mu_l(g) = \int f(g).d\left(\boldsymbol{\rho}\left(s^{-1}\right).\mu_l\right)(g)$$
$$= \Delta_{\mathbf{G}}\left(s^{-1}\right).\int f(g).d\mu_l(g).$$

In particular, for every μ_l-integrable set A (i.e. every set A whose characteristic function χ_A is μ_l-integrable: see [P2], section 4.1.6(**II**)),

$$\mu_l(A.s) = \int \chi_{A.s}(g).d\mu_l(g) = \int \chi_A\left(g.s^{-1}\right).d\mu_l(g) = \Delta_{\mathbf{G}}(s).\mu_l(A).$$

Thus, if V is invariant under the inner automorphisms of $\mathbf{G}$, then:

$$\mu_l(V) = \mu_l\left(s^{-1}.V.s\right) = \Delta_{\mathbf{G}}(s).\mu_l\left(s^{-1}.V\right) = \Delta_{\mathbf{G}}(s).\mu_l(V).$$

But $\mu_l(V) \neq 0$ by (2), so $\Delta_{\mathbf{G}}(s) = 1$. ■

The Haar measure μ'' on a quotient group $\mathbf{G}''$ is determined as follows ([BOU 69], Chapter 7, section 2.7):

THEOREM 6.14.– *Let* $\mathbf{G}$ *be a Lie group,* $\mathbf{G}'$ *a closed normal subgroup of* $\mathbf{G}$*,* $\mathbf{G}''$ *the quotient group* $\mathbf{G}/\mathbf{G}'$*,* π *the canonical surjection from* $\mathbf{G}$ *onto* $\mathbf{G}''$ *and* μ, μ', μ'' *three left Haar measures on* $\mathbf{G}, \mathbf{G}', \mathbf{G}''$*, respectively. Then, multiplying* μ *by a constant factor if necessary, we have* $\mu = \mu' \otimes \mu''$*. This can be rewritten as* $\mu'' = \mu/\mu'$*, and, for every function* $f \in \mathcal{K}_{\mathbb{C}}(\mathbf{G})$*,*

$$\int_{\mathbf{G}} f(x) \, .d\mu(x) = \int_{\mathbf{G}''} d\mu''(\dot{x}) \, . \int_{\mathbf{G}'} f(x\xi) \, .d\mu'(\xi), \quad \dot{x} = \pi(x).$$

REMARK 6.15.– *Lemma 6.8, Definition 6.9 and Theorem 6.14 still hold when* $\mathbf{G}$ *is a locally compact topological group, as we noted earlier in section 6.1 ([BOU 69], Chapter 7). Lemma-Definition 6.11 also holds, except that the homomorphism* $\Delta_{\mathbf{G}}$ *is merely continuous (rather than analytic). Theorem 6.13 holds without change.*

(II) CONVOLUTION OF A DISTRIBUTION AND A FUNCTION We will adopt the following convention:

(C3) In the following, unless otherwise stated, $\mathbf{G}$ will always be equipped with a left Haar measure denoted $\mathfrak{m}_{\mathbf{G}}$, or simply $\mathfrak{m}$ whenever $\mathbf{G}$ is implicitly clear. Furthermore, if $\mathbf{G} = \mathbb{R}^n$, then $\mathfrak{m}_{\mathbf{G}}$ is taken to be the Lebesgue measure; if $\mathbf{G}$ is discrete, then $\mathfrak{m}_{\mathbf{G}}$ is the measure defined by $\mathfrak{m}_{\mathbf{G}}(\{x\}) = 1$ for every $x \in \mathbb{Z}^n$; if $\mathbf{G}$ is compact, then $\mathfrak{m}_{\mathbf{G}}$ is the Haar measure with total mass equal to 1.

Let $f \in \mathcal{E}(\mathbf{G})$ (respectively $f \in \mathcal{D}(\mathbf{G})$) and $T_f = f.\mathfrak{m}$. Thus, $T_f \in \mathcal{D}'(\mathbf{G})$ (respectively $T_f \in \mathcal{E}'(\mathbf{G})$) and, for every distribution $S \in \mathcal{E}'(\mathbf{G})$ (respectively $S \in \mathcal{D}'(\mathbf{G})$), the convolution products $S \star T_f$ and $T_f \star S$ are well defined. The proof from ([DIE 93], Volume 3, section 17.12) can be adapted to show the following result ([BRU 56], Proposition 1; 2):

LEMMA 6.16.– *The convolutions* $S \star T_f$ *and* $T_f \star S$*, written as* $S \star f$ *and* $f \star S$*, respectively, are measures with base* $\mathfrak{m}$ *and densities* $\mathfrak{m}$*-almost everywhere equal to functions* g *and* h *in* $\mathcal{E}(\mathbf{G})$*. Writing* f, g *and* h *for* $f.\mathfrak{m}$*,* $g.\mathfrak{m}$ *and* $h.\mathfrak{m}$*, respectively, the regularizing bilinear mappings (or* regularizations*)* $(f, S) \mapsto S \star f$ *and* $(f, S) \mapsto f \star S$ *are separately continuous from* $\mathcal{E}(\mathbf{G}) \times \mathcal{E}'(\mathbf{G})$ *(respectively* $\mathcal{D}(\mathbf{G}) \times \mathcal{D}'(\mathbf{G})$*) into* $\mathcal{E}(\mathbf{G})$*.*

With the notation from [5.2] and $s, x \in \mathbf{G}$, Theorem 6.13(1) implies that:

$$(S \star f)(x) = \int_{\mathbf{G}} dS(s).f\left(s^{-1}.x\right), (f \star S)(x) = \int_{\mathbf{G}} f\left(x.s^{-1}\right) \Delta_{\mathbf{G}}\left(s^{-1}\right) .dS(s).$$

In particular,

$$\delta_s \star f = f\left(s^{-1}\bullet\right) = \boldsymbol{\lambda}(s) f,$$
$$f \star \delta_s = \Delta_{\mathbf{G}}\left(s^{-1}\right) f\left(\bullet s^{-1}\right) = \Delta_{\mathbf{G}}\left(s^{-1}\right) \boldsymbol{\delta}\left(s^{-1}\right) f.$$

Let (U, ξ, n) be a chart centered on s and α a multi-index. For every $x \in U$,

$$\boxed{\left(\left(D_\xi^\alpha \delta_s\right) \star f\right)(x) = D_\xi^\alpha f\left(s^{-1}x\right).}$$

(III) CONVOLUTION OF FUNCTIONS We have the following result ([DIE 93], Volume 2, (14.10.1)):

LEMMA 6.17.– *Let f, g be complex functions on $\mathbf{G}$ that are locally $\mathfrak{m}$-integrable. The measures $f.\mathfrak{m}$ and $g.\mathfrak{m}$ are convolvable (Remark 6.4(2)) if and only if there exists an $\mathfrak{m}$-negligible subset N such that, for every $x \in N$, the function $s \mapsto f(s)\, g\left(s^{-1}.x\right)$ (defined $\mathfrak{m}$-almost everywhere) is locally $\mathfrak{m}$-integrable. If so, the following equality holds $\mathfrak{m}$-almost everywhere:*

$$h(x) := \int_{\mathbf{G}} f(s) \, .g\left(s^{-1}.x\right).d\mathfrak{m}(s) = \int_{\mathbf{G}} f\left(x.s^{-1}\right).g(s)\, .\Delta_{\mathbf{G}}\left(s^{-1}\right).d\mathfrak{m}(s)\, .$$

The function $h : x \mapsto h(x)$ (defined $\mathfrak{m}$-almost everywhere) is locally $\mathfrak{m}$-integrable, and $(f.\mathfrak{m}) \star (g.\mathfrak{m}) = h.\mathfrak{m}$.

DEFINITION 6.18.– *The function h from Lemma 6.17 is called the* convolution *of f and g, written as $f \star g$.*

Let $\mathcal{C}_0(\mathbf{G}; \mathbb{C})$ be the space of continuous complex functions satisfying the following property: for every $\varepsilon > 0$, there exists a compact set $K \subset \mathbf{G}$ such that $|f(x)| \leq \varepsilon$ for every $x \in X - K$.

DEFINITION 6.19.– *We say that $\mathcal{C}_0(\mathbf{G}; \mathbb{C})$ is the space of* continuous functions *that are* zero at infinity.

THEOREM 6.20.– *1) If $f \in \mathcal{L}^p(\mathbf{G}, \mathfrak{m})$, $g \in \mathcal{L}^q(\mathbf{G}, \mathfrak{m})$, $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ ($p, q, r \in [1, \infty]$) and $p = 1$ (so $r = q$) or $\mathbf{G}$ is unimodular, then $f \star g$ is $\mathfrak{m}$-almost everywhere equal to a function $h \in \mathcal{L}^r(\mathbf{G}, \mathfrak{m})$. Furthermore, with the notation of [P2], section 4.1.2(**III**), Theorem 4.12, $N_r(f \star g) \leq N_p(f).N_q(g)$ (*Young's inequality*).*

In particular, setting $p = 1$ gives $r = q$; if we set $q = 1$, then $\mathcal{L}^1(\mathbf{G}, \mathfrak{m})$ is a convolution algebra. For $p = 1$, $q = \infty$, the function $f \star g \in \mathcal{L}^\infty(\mathbf{G}, \mathfrak{m})$ is $\mathfrak{m}$-almost everywhere equal to a continuous function h, written as $h = f \star g$.

*2) If $p, q \in]1, +\infty[$ are a pair of conjugate exponents ([P2], section 2.1.1(**II**), Definition 2.1), i.e. $\frac{1}{p} + \frac{1}{q} = 1$, $f \in \mathcal{L}^p(\mathbf{G}, \mathfrak{m})$ and $g \in \mathcal{L}^q(\mathbf{G}, \check{\mathfrak{m}})$, then $f \star g \in \mathcal{C}_0(\mathbf{G}; \mathbb{C})$ and $N_\infty(f \star g) \leq N_p(f).N_q(\check{g})$ (Notation 6.12).*

PROOF.– 1) (i) First, consider the case where $p = 1$. Then (the integrals might *a priori* be infinite):

$$\int |(f \star g)(x)| \, d\mathfrak{m} \leq \int d\mathfrak{m}(x) \int \left| f(s) \, g\left(s^{-1}x\right) \right| d\mathfrak{m}(s).$$

By performing the change of variable $x \mapsto t = s^{-1}x$,

$$\int |(f \star g)(x)| \, d\mathfrak{m} \leq \int d\mathfrak{m}(t) \int |f(s)| \, |g(t)| \, d\mathfrak{m}(s) \leq N_1(f) \, N_1(g).$$

ii) If we allow the semi-norms to take the value $+\infty$, then, setting $\check{g}_x(s) = g\left(s^{-1}.x\right) = (\boldsymbol{\lambda}(x) \, g)^{\vee}$,

$$|(f \star g)(x)| \leq \int \left(|f|^p \, |\check{g}_x|^q\right)^{1/r} . \, |f|^{p(1/p-1/r)} . \left|(\boldsymbol{\lambda}(x) \, g)^{\vee}\right|^{q(1/q-1/r)} . d\mathfrak{m},$$

where $1 = 1/r + (1/p - 1/r) + (1/q - 1/r)$, so, by the generalized Hölder inequality ([P2], section 4.1.2**(III)**, Theorem 4.14),

$$|(f \star g)(x)| \leq N_r\left(\left(|f|^p \, |\check{g}_x|^q\right)^{1/r}\right) . N_{\frac{1}{1/p-1/r}}(|f|)^{p(1/p-1/r)}$$
$$. N_{\frac{1}{1/q-1/r}}\left(\left|(\boldsymbol{\lambda}(x) \, g)^{\vee}\right|^{q(1/q-1/r)}\right).$$

But

$$\left(N_{\frac{1}{1/q-1/r}}\left(\left|(\boldsymbol{\lambda}(x) \, g)^{\vee}\right|^{q(1/q-1/r)}\right)\right)^r = \left(\int \left|(\boldsymbol{\lambda}(x) \, g)^{\vee}(s)\right|^q d\mathfrak{m}(s)\right)^{\frac{1}{q}(r-q)}$$
$$= N_q\left((\boldsymbol{\lambda}(x) \, g)^{\vee}\right)^{r-q}.$$

Hence, $N_r(f \star g) \leq N_1\left(|f|^p \star |g|^q\right) . N_p(|f|)^{r-p} . N_q\left((\boldsymbol{\lambda}(x) \, g)^{\vee}\right)^{r-q}$, and, by (i), $N_r(f \star g) \leq N_p(f) \, N_q(g)^{q/r} \, N_q\left((\lambda(x) \, g)^{\vee}\right)^{1-q/r}$. If $\mathbf{G}$ is unimodular, then $N_q\left((\boldsymbol{\lambda}(x) \, g)^{\vee}\right) = N_q(g)$ by Theorem 6.13(1).

iii) If $p = 1$ and $q = \infty$, let $x \in \mathbf{G}$ and suppose that (x_n) is a sequence of elements of $\mathbf{G}$ converging to x. Then:

$$\begin{aligned}|h(x) - h(x_n)| &\leq \int \left|f\left(xs^{-1}\right) - f\left(x_n s^{-1}\right)\right| |g(s)| \Delta_{\mathbf{G}}\left(s^{-1}\right) d\mathfrak{m}(s) \\ &\leq N_\infty(g) \int \left|f\left(xs^{-1}\right) - f\left(x_n s^{-1}\right)\right| \Delta_{\mathbf{G}}\left(s^{-1}\right) d\mathfrak{m}(s) \\ &= N_\infty(f) \int \left|f\left(t^{-1}\right) - f\left(x_n x^{-1} t\right)\right| \Delta_{\mathbf{G}}\left(x^{-1} t\right) d\mathfrak{m}(t).\end{aligned}$$

This tends to 0 by the dominated convergence theorem ([P2], section 4.1.2**(II)**).

2) The function $\breve{g}_x : s \mapsto g\left(s^{-1}x\right)$ belongs to $\mathcal{L}^q(\mathbf{G}, \check{\mathfrak{m}})$ by the left invariance of $\mathfrak{m}$, and $N_q(\breve{g}) = N_q(\breve{g}_x)$. Hence, the function $s \mapsto f.\breve{g}_x$ is $\mathfrak{m}$-integrable and $|h(x)| = N_1(f.\breve{g}_x) \leq N_p(f).N_q(\breve{g})$ by Hölder's inequality ([P2], section 4.1.2**(III)**, Lemma 4.13). For the proof that h belongs to $\mathcal{C}_0(\mathbf{G}; \mathbb{C})$, see [BOU 69], Chapter 8, section 4.5, Proposition 15, where $\mathcal{C}_0(\mathbf{G}; \mathbb{C})$ is written as $\overline{\mathcal{K}(\mathbf{G})}$. ■

Setting $p = q = 2$ gives the following result:

COROLLARY 6.21.– *If $f, g \in \mathcal{L}^2(\mathbf{G}, \mathfrak{m})$, then f and $\tilde{g} : x \mapsto \overline{g(x^{-1})}$ are convolvable, $f \star \tilde{g}$ belongs to $\mathcal{C}_0(\mathbf{G}; \mathbb{C})$ and $(f \star \tilde{g})(e) = \int_{\mathbf{G}} f(x).\overline{g(x)}.d\mathfrak{m}(x) := \langle f, g \rangle_2$. In particular, we have Schwarz's inequality:*

$$\boxed{|(f \star \tilde{g})(e)| \leq \|f\|_2 \cdot \|g\|_2.}$$

6.3. Classification of Lie algebras

6.3.1. *Additional notions from algebra*

(I) TRACE OF AN ENDOMORPHISM; SEMI-SIMPLE ENDOMORPHISMS Let $\mathbf{K}$ be a field. If $\mathbf{E}$ is an n-dimensional $\mathbf{K}$-vector space, then any element $\mathbf{u} \in \mathrm{End}(\mathbf{E})$ can be represented in a given basis of $\mathbf{E}$ by a matrix $A = \mathrm{Mat}(\mathbf{u}) \in \mathfrak{M}_n(\mathbf{K})$, where $\mathfrak{M}_n(\mathbf{K})$ is the algebra of square matrices of order n with entries in $\mathbf{K}$ ([P1], sections 2.3.11**(I)** and 3.1.3**(I)**). If P is a change-of-basis matrix, then $\det\left(PAP^{-1}\right) = \det(A)$ ([P1], section 2.3.11**(V)**), i.e. $\det(A)$ is invariant under similarity ([P1], section 3.4.3**(III)**); the same is true for the characteristic polynomial $P_A(X) = \det(X.I_n - A) = X^n - \mathrm{Tr}(A) X^{n-1} + \ldots + (-1)^n \det(A)$ and hence also for the trace $\mathrm{Tr}(A)$ ([P1], section 2.3.11**(VII)**). We can therefore write $\det(\mathbf{u}) = \det(A)$, $P_{\mathbf{u}}(X) = P_A(X)$ and $\mathrm{Tr}(\mathbf{u}) = \mathrm{Tr}(A)$.

The elementary divisors of A ([P1], section 3.4.3**(IV)**) are invariant under similarity and may therefore be called the elementary divisors of $\mathbf{u}$; similarly, the minimal polynomial of A is said to be the minimal polynomial of $\mathbf{u}$. The matrix A, or any matrix similar to A, is diagonalizable over an algebraic closure of $\mathbf{K}$ if and only if its elementary divisors are all of degree one, or alternatively if the minimal polynomial does not have repeated factors. If so, we say that the endomorphism $\mathbf{u}$ is *semi-simple*.

(II) CHANGE OF BASE FIELD Let $\mathbf{K}, \mathbf{L}$ be two fields, $\rho : \mathbf{K} \to \mathbf{L}$ a morphism of fields, $\rho^* = - \otimes_{\mathbf{K}} \mathbf{L}$ the functor of "extension of the field of scalars" ([P1], section 3.1.5**(VI)**), and $\mathbf{A}$ a $\mathbf{K}$-algebra ([P1], section 2.3.10**(I)**). The $\mathbf{L}$-vector space $\rho^* (\mathbf{A})$ can be equipped with a canonical $\mathbf{L}$-algebra structure as follows: the canonical isomorphism $(\mathbf{A} \otimes_{\mathbf{K}} \mathbf{L}) \otimes_{\mathbf{L}} (\mathbf{A} \otimes_{\mathbf{K}} \mathbf{L}) \cong (\mathbf{A} \otimes_{\mathbf{K}} \mathbf{L}) \otimes_{\mathbf{K}} \mathbf{L}$ implies that, if $\boldsymbol{m} : (\mathbf{A} \otimes_{\mathbf{K}} \mathbf{A}) \to \mathbf{A}$ is the $\mathbf{K}$-linear mapping defining multiplication in $\mathbf{A}$ (by $a.a' = \boldsymbol{m}(a \otimes a')$), then the $\mathbf{L}$-linear mapping $\boldsymbol{m} \otimes 1_{\mathbf{L}}$ can be canonically identified with an $\mathbf{L}$-linear mapping $\boldsymbol{m}' : \rho^* (\mathbf{A}) \otimes_{\mathbf{L}} \rho^* (\mathbf{A}) \to \rho^* (\mathbf{A})$, which defines multiplication in $\rho^* (\mathbf{A})$, and $(a \otimes l) . (a' \otimes l') = (a.a') \otimes (l.l')$ for every $a, a' \in \mathbf{A}, l, l' \in \mathbf{L}$ ([BOU 12], Chapter 3, section 1.5).

In the special case where $\mathbf{K} = \mathbb{R}$ and $\mathbf{L} = \mathbb{C}$ (the most important case below), the *complexification* of the real algebra $\mathbf{A}$ is $\mathbf{A} \otimes_{\mathbb{R}} \mathbb{C} = \mathbf{A} \oplus \mathbf{A}i$, with $(\lambda + \lambda' i)(a + a' i) = (\lambda a - \lambda' a') + (\lambda' a + \lambda a') i$, where $a, a' \in \mathbf{A}$ and $\lambda, \lambda' \in \mathbb{R}$ or $\lambda, \lambda' \in \mathbf{A}$. The real algebra $\mathbf{A}$ is called the *real form* of $\mathbf{A} \otimes_{\mathbb{R}} \mathbb{C}$.

If $\mathbf{B}$ is an $\mathbf{L}$-algebra, then it is clear that the $\mathbf{K}$-module $\rho_* (\mathbf{B}) = \mathbf{B}_{[\mathbf{L}]}$ obtained from $\mathbf{B}$ by restriction of the field of scalars ([P1], section 3.1.5**(VI)**) has a canonical $\mathbf{K}$-algebra structure.

Let $\mathbf{E}, \mathbf{F}$ be two $\mathbf{K}$-vector spaces and $\Phi : \mathbf{E} \times \mathbf{F} \to \mathbf{K}$ a bilinear form. There exists a unique bilinear form $\check{\Phi} : \rho^* (\mathbf{E}) \times \rho^* (\mathbf{F}) \to \mathbf{L}$ such that $\check{\Phi}(x \otimes l, y \otimes l') = \Phi(x, y) \otimes l.l'$ for every $x \in \mathbf{E}, y \in \mathbf{F}, l, l' \in \mathbf{L}$. The bilinear mapping $\check{\Phi}$ is said to be *obtained from* Φ *by extension of the base field*. If $\mathbf{M}$ is a vector subspace of $\mathbf{E}$, then its polar set is $\mathbf{M}^0 := \{y \in \mathbf{F} : \Phi(x, y) = 0, \forall x \in \mathbf{M}\}$, and we can similarly define the polar set $\mathbf{N}^0$ of a vector subspace $\mathbf{N}$ of $\mathbf{F}$. The bilinear mapping Φ is non-degenerate if and only if $\mathbf{E}^0 = \{0\}$ and $\mathbf{F}^0 = \{0\}$ ([P2], section 3.5.1). We have $\rho^* \left(\mathbf{M}^0\right) = (\rho^* (\mathbf{M}))^0$; in particular, Φ is non-degenerate if and only if $\check{\Phi}$ is non-degenerate ([BOU 12], Chapter 9, section 1.4).

If $\mathbf{K} = \mathbb{R}$ and $\mathbf{L} = \mathbb{C}$, the *complexification* $\check{\Phi}$ of a real bilinear form Φ is given by $\check{\Phi}(x + yi, x' + y'i) = \Phi(x, x') - \Phi(y, y') + (\Phi(x, y') + \Phi(y, x')) i$ for every $x, y \in E, y, y' \in F$.

(III) LINEAR REPRESENTATION OF AN ALGEBRA Let $\mathbf{A}$ be a $\mathbf{K}$-algebra and $\mathbf{E}$ a $\mathbf{K}$-vector space. A *linear representation* of the algebra $\mathbf{A}$ in $\mathbf{E}$ is a $\mathbf{K}$-linear mapping from $\mathbf{A}$ into the algebra End $(\mathbf{E})$. Any representation of a Lie algebra (section 5.4.1(**II**)) is clearly linear.

(IV) QUADRATIC FORMS

(a) Let $\mathbf{E}$ be a real m-dimensional vector space and $\Phi : \mathbf{E} \times \mathbf{E} \to \mathbb{R}$ a symmetric non-degenerate bilinear form ([P2], section 3.5.1). Equipped with this form, $\mathbf{E}$ is a Hausdorff pre-Hilbert space whenever Φ is positive definite ([P2], section 3.10.1); furthermore, $\mathbf{E}$ is complete because it is finite-dimensional ([P2], section 3.2.2(**IV**)), and hence it is a Hilbert space. If we write x_j, y_j $(1 \leq j \leq m)$ for the components of two elements $x, y \in \mathbf{E}$ in a Hilbert basis, then $\Phi(x,y) = \sum_{i=1}^{m} x_j.y_j = \mathbf{x}^T.\mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ are the columns of the x_j and y_j, respectively; $Q : \mathbf{E} \to \mathbb{R} : x \longmapsto \Phi(x,x)$ is the quadratic form $Q(x) = \sum_{j=1}^{m} x_j^2$. Knowing Q is equivalent to knowing Φ, since $2\Phi(x,y) = Q(x+y) - Q(x) - Q(y)$.

(b) The complexification $\mathbf{E}_{(\mathbb{C})} := \mathbf{E} \oplus i\mathbf{E}$ of $\mathbf{E}$ is equipped with the complexification of Φ (also written as Φ) and the quadratic form $Q(x) = \sum_{j=1}^{m} x_j^2$. If $m = 2n$, the unitary transformation ([P2], section 3.10.3(**I**)) $(x_1, ..., x_m) \mapsto (w_1, ..., w_m)$, where $w_j = \frac{1}{\sqrt{2}}(x_j + ix_{m-j+1})$, $w_{m-j+1} = \frac{1}{\sqrt{2}}(x_j - ix_{m-j+1})$ for $j = 1, ..., m$, transforms Q into the equivalent quadratic form $\sum_{j=1}^{m} w_j w_{m-j+1}$. When $m = 2n+1$, the same is true for the unitary transformation $(x_1, ..., x_m) \mapsto (w_1, ..., w_m)$, where w_j is defined as above for $j \neq n$ and $w_n = x_n$.

(c) Let $\mathbf{K}$ be a field whose characteristic is not 2 and $\mathbf{E}$ an m-dimensional $\mathbf{K}$-vector space over $\mathbf{K}$. A mapping $Q : \mathbf{E} \to \mathbf{K}$ is said to be a *quadratic form* if $Q(\lambda.x) = \lambda^2.Q(x)$ for every $\lambda \in \mathbf{K}, x \in \mathbf{E}$ and the mapping $\Phi : (x,y) \mapsto \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$ from $\mathbf{E} \times \mathbf{E}$ into $\mathbf{K}$ is a (necessarily symmetric) bilinear form. If Φ is non-degenerate, we say that Q is non-degenerate. This condition is assumed to hold in the following. A subspace $\mathbf{F}$ of $\mathbf{E}$ is said to be *totally isotropic* if $\mathbf{F} \subset \mathbf{F}^0$.

(α) If $\mathbf{K}$ is the field of real numbers, there exists an integer $p \leq m$ and a basis of $\mathbf{E}$ in which $Q(x) = \mathbf{x}^T.\Lambda_{p,m}.\mathbf{x}$, where $\Lambda_{p,m} = \operatorname{diag}\left(\underbrace{1, ..., 1}_{p}, \underbrace{-1, ... -1}_{m-p}\right)$; the pair $(p, m-p)$ is the signature of Q, and $\nu = \inf(p, m-p)$ is said to be its *Witt index*. This index is the maximal dimension of any totally isotropic subspace of $\mathbf{E}$; it satisfies $\nu \leq n$ if $m = 2n$ or $m = 2n+1$. Hence, $\nu = n$ if and only if there exists a totally isotropic subspace of $\mathbf{E}$ that has dimension n, and this condition is satisfied if and only if there exists a basis of $\mathbf{E}$ in which $Q(w) = \sum_{j=1}^{n} w_j w_{m-j+1}$, where the w_j are the components of the vector w.

(β) For an arbitrary field $\mathbf{K}$ whose characteristic is not 2, the signature of Q is not defined. However, we can still define the Witt index of Q using (α); Q is a quadratic form with maximal Witt index if and only if there exists a totally isotropic subspace of $\mathbf{E}$ with dimension $n = \frac{m}{2}$ or $\frac{m-1}{2}$, according to whether m is even or odd. If so, there exists a basis of $\mathbf{E}$ in which $Q(w) = \sum_{j=1}^{n} w_j w_{m-j+1}$ for every $w \in \mathbf{E}$. This quadratic form can be written as $\mathbf{w}^T.\tilde{I}_m.\mathbf{w}$ (with the conventions adopted earlier), where $\tilde{I}_m$ is the second unit diagonal in $\mathfrak{M}_m(\mathbf{K})$[3].

6.3.2. *Classical Lie algebras*

(I) Let $\mathbf{K}$ be a field. The algebras of matrices defined below (equipped with the bilinear mapping specified in Lemma 5.11) are Lie algebras over a field $\mathbf{K}$ (section 5.4.1). The proofs of the stated results are left to the reader as an **exercise**. Write A^* for the conjugate transpose of a matrix $A \in \mathfrak{M}_n(\mathbb{C})$. Recall that a *subalgebra* of a Lie algebra $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{g}$ (section 5.4.1**(II)**).

1) The associative $\mathbf{K}$-algebra $\mathfrak{M}_n(\mathbf{K})$, equipped with the bracket specified in Lemma 5.11 (section 5.4.1), is a Lie $\mathbf{K}$-algebra written as $\mathfrak{gl}_n(\mathbf{K})$ and a $\mathbf{K}$-vector space of dimension n^2. The real algebra $\mathfrak{gl}_n(\mathbb{R})$ is simply written as $\mathfrak{gl}_n$. Similar conventions are adopted for each of the Lie algebras considered below.

2) The set $\mathfrak{sl}_n(\mathbf{K})$ of matrices $A \in \mathfrak{M}_n(\mathbf{K})$ such that $\operatorname{Tr}(A) = 0$ is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension $n^2 - 1$.

3) The set $\mathfrak{u}_n(\mathbb{C})$ of matrices $A \in \mathfrak{M}_n(\mathbb{C})$ such that $A + A^* = 0$ is a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ of dimension $n(n-1)/2$ over $\mathbb{C}$. Let E_r^s be the matrix $\left(a_i^j\right)$ with zeros everywhere except a_r^s, which is equal to 1. Then, the n matrices iE_{rr} $(i = 1, ..., n)$ and the $n^2 - n$ matrices $E_r^s - E_s^r$, $i(E_r^s + E_s^r)$ $(1 \le r < s \le n)$ form a basis of the real vector space $\mathfrak{u}_n := \mathfrak{u}_n(\mathbb{C})_{[\mathbb{R}]}$, so the complexification $(\mathfrak{u}_n)_{\mathbb{C}}$ of $\mathfrak{u}_n$ is $\mathfrak{gl}_n(\mathbb{C})$ (the reader is invited to fill in the details for the case $n = 2$ as an illustration). Both Lie algebras $\mathfrak{u}_n$ and $\mathfrak{gl}_n$ are real forms of $\mathfrak{gl}_n(\mathbb{C})$ (section 6.3.1**(II)**) (but $\mathfrak{u}_n \ncong \mathfrak{gl}_n(\mathbb{R})$ except when $n = 1$).

4) Similarly, $\mathfrak{su}_n(\mathbb{C}) := \mathfrak{u}_n(\mathbb{C}) \cap \mathfrak{sl}_n(\mathbb{C})$ is a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ of dimension $n^2 - 1$. The complexification $(\mathfrak{su}_n)_{(\mathbb{C})}$ of $\mathfrak{su}_n$ is $\mathfrak{sl}_n(\mathbb{C})$ by the above. The Lie algebras $\mathfrak{su}_2$ and $\mathfrak{sl}_2$ are not isomorphic (**exercise**), but they both share the same complexification $\mathfrak{sl}_2(\mathbb{C})$.

5) The set $\mathfrak{sp}_{2n}(\mathbf{K})$ of matrices $E \in \mathfrak{gl}_{2n}(\mathbf{K})$ such that $J.E + E^T.J = 0$, where J is defined in the same way as in section 2.4.1**(V)**, is a subalgebra of $\mathfrak{gl}_{2n}(\mathbf{K})$ of dimension $n(2n+1)$, called the *symplectic algebra*.

6) If $P \in \mathrm{GL}_n(\mathbf{K})$ is the matrix of a non-degenerate quadratic form $Q : x \mapsto x^T P x$ (where $x \in \mathbf{K}^n$ is identified with a column), then the set $\mathfrak{o}_n(\mathbf{K}; P)$ of matrices

3 The second unit diagonal in $\mathfrak{M}_m(\mathbf{K})$ is the matrix $\tilde{I}_m = \left(a_i^j\right)$ such that $a_i^j = 1$ when $i + j = m + 1$ $(i = 1, ..., m)$ and $a_i^j = 0$ otherwise.

$A \in \mathfrak{M}_n(\mathbf{K})$ such that $A^T P + PA = 0$ is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension $n(n-1)/2$. This Lie algebra is written as $\mathfrak{o}_n(\mathbf{K})$ when $P = \tilde{I}_n$; in the case where $\mathbf{K}$ is the field of complex numbers, it is equivalent and in practice more convenient to choose $P = I_n$ by section 6.3.1**(IV)(b)**; Corollaries 6.42, 6.56 (section 6.3.6 below) prompt us to do the same when $\mathbf{K}$ is the field of real numbers. Thus, $\mathfrak{o}_n(\mathbb{C})$ is the set of matrices $A \in \mathfrak{M}_n(\mathbb{C})$ such that $A + A^T = 0$, and $\mathfrak{o}_n$ is the set of matrices $A \in \mathfrak{M}_n(\mathbb{R})$ such that $A + A^T = 0$ (antisymmetric matrices); its complexification $(\mathfrak{o}_n)_{(\mathbb{C})}$ is clearly $\mathfrak{o}_n(\mathbb{C})$.

7) The set $\mathfrak{d}_n(\mathbf{K})$ of diagonal matrices $A \in \mathfrak{M}_n(\mathbf{K})$ is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension n.

8) The set $\mathfrak{t}_n(\mathbf{K})$ of upper triangular matrices $A \in \mathfrak{M}_n(\mathbf{K})$ is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension $n(n+1)/2$.

9) The set $\mathfrak{st}_n(\mathbf{K})$ of upper triangular matrices $A \in \mathfrak{M}_n(\mathbf{K})$ with zero trace is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension $n(n+1)/2 - 1$.

10) The set $\mathfrak{n}_n(\mathbf{K})$ of upper triangular matrices $A \in \mathfrak{M}_n(\mathbf{K})$ with zeros along the diagonal is a subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ of dimension $n(n-1)/2$.

11) The set $\mathfrak{a}_n$ (respectively $\mathfrak{se}_n$) of matrices of the form $\begin{bmatrix} \Omega & \nu \\ 0 & 0 \end{bmatrix}$, where $\Omega \in \mathfrak{gl}_n$ (respectively $\Omega \in \mathfrak{o}_n$) and $\nu \in \mathbb{R}^n$, is a subalgebra of $\mathfrak{gl}_{2n}$ of dimension $n(n+1)$ satisfying $\mathfrak{a}_n \cong \mathfrak{d}_n \oplus_s \mathfrak{gl}_n$ and $\mathfrak{se}_n \cong \mathfrak{d}_n \oplus_s \mathfrak{o}_n$; furthermore, $\mathfrak{d}_n$ is an ideal of both $\mathfrak{a}_n$ and $\mathfrak{se}_n$ (section 5.4.1**(II)**).

12) The set $\mathfrak{c}_n$ of matrices of the form $\lambda i I_n$ $(\lambda \in \mathbb{R})$ is a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$.

(II) The Lie algebra $\mathfrak{sl}_n(\mathbf{K})$ is an ideal of $\mathfrak{gl}_n(\mathbf{K})$. The Lie algebra $\mathfrak{o}_n(\mathbf{K})$ is a subalgebra of $\mathfrak{sl}_n(\mathbf{K})$ and can be written as $\mathfrak{so}_n(\mathbf{K})$. The Lie algebras $\mathfrak{sl}_n(\mathbf{K})$ $(n \geq 2)$, $\mathfrak{o}_{2n+1}(\mathbf{K})$, $\mathfrak{sp}_{2n}(\mathbf{K})$ $(n \geq 1)$ and $\mathfrak{o}_{2n}(\mathbf{K})$ $(n \geq 2)$ are written as A_{n-1}, B_n, C_n and D_n, respectively, ever since S. Lie. They satisfy $C_1 \cong A_1$, $C_2 \cong B_2$, $D_2 \cong A_1 \oplus A_1$, $D_3 \cong A_3$ ([BOU 82b], Chapter 8, section 13). See section 6.3.6**(III)**.

We also have $\mathfrak{c}_n = \mathfrak{z}(\mathfrak{u}_n)$ and $\mathfrak{u}_n = \mathfrak{c}_n \oplus \mathfrak{su}_n$ (**exercise**).

6.3.3. *General notions about Lie algebras*

(I) Killing forms Let $\mathfrak{g}$ be a Lie algebra over the field $\mathbf{K}$ and $X, Y \in \mathfrak{g}$. Then, $\mathrm{ad}X \circ \mathrm{ad}Y \in \mathfrak{gl}(\mathfrak{g})$ (section 5.4.1**(III)**). Hence, we can consider the trace $B_{\mathfrak{g}}(X,Y) := \mathrm{Tr}(\mathrm{ad}X \circ \mathrm{ad}Y)$ (section 6.3.1**(I)**). The bilinear form $(X,Y) \mapsto B_{\mathfrak{g}}(X,Y)$ from $\mathfrak{g} \times \mathfrak{g}$ into $\mathbf{K}$ is called the *Killing form* of $\mathfrak{g}$.

(II) DESCENDING CENTRAL SERIES AND DERIVED SERIES The characteristic ideal (section 5.4.1(**II**)) $[\mathfrak{g}, \mathfrak{g}]$ is said to be the *derived ideal* of $\mathfrak{g}$ and is written as $\mathfrak{g}'$ or $\mathcal{D}\mathfrak{g}$; this and the notions of descending central series and derived series are comparable to the corresponding concepts introduced for groups in [P1], sections 2.2.6 and 2.2.7. Given an ideal $\mathfrak{a}$ of $\mathfrak{g}$ (section 5.4.1(**II**))), the quotient algebra $\mathfrak{g}/\mathfrak{a}$ commutes if and only if $\mathfrak{a} \supset \mathcal{D}\mathfrak{g}$ (**exercise**), and $\mathfrak{g}^{ab} := \mathfrak{g}/\mathcal{D}\mathfrak{g}$ is called the abelianization of $\mathfrak{g}$.

DEFINITION 6.22.– *1) The* descending central series *of* $\mathfrak{g}$ *is the descending sequence of characteristic ideals* $\mathcal{C}^1\mathfrak{g}, \mathcal{C}^2\mathfrak{g}, \ldots$ *determined inductively as follows:*

$$\mathcal{C}^1\mathfrak{g} = \mathfrak{g}, \quad \mathcal{C}^{p+1}\mathfrak{g} = [\mathfrak{g}, \mathcal{C}^p\mathfrak{g}] .$$

2) The derived series *of* $\mathfrak{g}$ *is the descending sequence of characteristic ideals* $\mathcal{D}^0\mathfrak{g}, \mathcal{D}^1\mathfrak{g}, \ldots$ *determined inductively as follows:*

$$\mathcal{D}^0\mathfrak{g} = \mathfrak{g}, \quad \mathcal{D}^{p+1}\mathfrak{g} = [\mathcal{D}^p\mathfrak{g}, \mathcal{D}^p\mathfrak{g}] .$$

Note that $\mathcal{C}^2\mathfrak{g} = \mathcal{D}^1\mathfrak{g} = \mathcal{D}\mathfrak{g}$, and $\mathcal{C}^{p+1}\mathfrak{g} \supset \mathcal{D}^p\mathfrak{g}$ for every $p \geq 1$.

(III) CHANGE OF BASE FIELD Let $\mathbf{K}, \mathbf{L}$ be fields and $\rho : \mathbf{K} \to \mathbf{L}$ a morphism of fields. Let $\mathfrak{g}$ be a Lie $\mathbf{L}$-algebra and $\rho_* (\mathfrak{g})$ the algebra obtained from $\mathfrak{g}$ by restriction of the field of scalars (see (**I**)(b)). Then, $\rho_* (\mathfrak{g})$ is a Lie $\mathbf{K}$-algebra, the subalgebras (respectively ideals) of $\rho_* (\mathfrak{g})$ are the $\rho_* (\mathfrak{a})$, where the $\mathfrak{a}$ are the subalgebras (respectively ideals) of $\mathfrak{g}$, and $\rho_* ([\mathfrak{a}, \mathfrak{b}]) = [\rho_* (\mathfrak{a}), \rho_* (\mathfrak{b})]$ for every pair of $\mathbf{L}$-vector subspaces $\mathfrak{a}, \mathfrak{b}$ of $\mathfrak{g}$. We have $\mathcal{C}^p\rho_* (\mathfrak{g}) = \rho_* (\mathcal{C}^p\mathfrak{g})$ and $\mathcal{D}^p\rho_* (\mathfrak{g}) = \rho_* (\mathcal{D}^p\mathfrak{g})$ for every p.

If $\mathfrak{g}$ is a Lie $\mathbf{K}$-algebra, the above still holds when ρ_* is replaced by ρ^* (section 5.4.1(**I**)(b)). The Killing form of $\rho^* (\mathfrak{g})$ can be deduced from the Killing form of $\mathfrak{g}$ by extension of the base field to $\mathbf{L}$ (see (**I**)(**b**)). If V is a finite-dimensional vector space over $\mathbf{K}$, $\mathfrak{gl} (\rho^* (V))$ can be canonically identified with $\rho^* (\mathfrak{gl} (V))$.

(IV) ENVELOPING ALGEBRA Henceforth, any $\mathbf{K}$-vector spaces are always finite-dimensional. Let $\mathfrak{g}$ be a Lie $\mathbf{K}$-algebra and consider the tensor algebra $\mathbf{T}$ of the $\mathbf{K}$-vector space $\mathfrak{g}$ (Lemma-Definition 4.1). The algebra $\mathbf{T}$ is associative and unitary ([P1], section 2.3.10(**I**)). Let J be the two-sided ideal of $\mathbf{T}$ generated by the tensors $X \otimes Y - Y \otimes X - [X, Y]$ $(X, Y \in \mathfrak{g})$.

DEFINITION 6.23.– *The associative unitary* $\mathbf{K}$*-algebra* $\boldsymbol{U} (\mathfrak{g}) = \mathbf{T}/J$ *is said to be the enveloping algebra of* $\mathfrak{g}$*. The restriction* σ_0 *of the canonical surjection* $\mathbf{T} \twoheadrightarrow \boldsymbol{U} (\mathfrak{g})$ *to* $\mathfrak{g}$ *is said to be the canonical mapping from* $\mathfrak{g}$ *into* $\boldsymbol{U} (\mathfrak{g})$.

It can be shown that $\sigma_0 : \mathfrak{g} \to \boldsymbol{U}(\mathfrak{g})$ is injective ([BOU 82b], Chapter 1, section 2.7, Corollary 2).

The algebra $\boldsymbol{U}(\mathfrak{g})$ satisfies the following universal property: for every associative unitary $\mathbf{K}$-algebra $\boldsymbol{A}$ and every morphism $\sigma : \mathfrak{g} \to \boldsymbol{A}$ such that $\sigma([X,Y]) = \sigma(X)\sigma(Y) - \sigma(Y)\sigma(X)$, there exists a unique morphism τ of unitary algebras ([P1], section 2.3.10(**I**)) from $\boldsymbol{U}(\mathfrak{g})$ into $\boldsymbol{A}$ such that $\sigma = \tau \circ \sigma_0$, thus making the following diagram commute:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\sigma_0} & \boldsymbol{U}(\mathfrak{g}) \\ & {}_{\sigma}\searrow & \uparrow \tau \\ & & \boldsymbol{A} \end{array}$$

If $\mathfrak{g}'$ is another Lie algebra over $\mathbf{K}$ and $\varphi : \mathfrak{g} \to \mathfrak{g}'$ is a morphism of Lie algebras, there exists a unique morphism $\tilde{\varphi}$ of unitary algebras for which the following diagram commutes (**exercise**):

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\varphi} & \mathfrak{g}' \\ \sigma_0 \downarrow & & \downarrow \sigma_0' \\ \boldsymbol{U}(\mathfrak{g}) & \xrightarrow{\tilde{\varphi}} & \boldsymbol{U}(\mathfrak{g}') \end{array}$$

We have $\boldsymbol{U}(\mathfrak{g} \times \mathfrak{g}') \cong \boldsymbol{U}(\mathfrak{g}) \otimes_{\mathbf{K}} \boldsymbol{U}(\mathfrak{g}')$ ([BOU 82b], Chapter 1, section 2.2, Proposition 2). If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, there exists a canonical morphism of unitary algebras $\boldsymbol{U}(\mathfrak{h}) \to \boldsymbol{U}(\mathfrak{g})$; this morphism is injective ([BOU 82b], Chapter 1, section 2.7, Corollary 5).

Let $\{E_i : 1 \leq i \leq n\}$ be a basis of the $\mathbf{K}$-vector space $\mathfrak{g}$. The algebra $\boldsymbol{U}(\mathfrak{g})$ is generated by 1 and $\sigma_0(\mathfrak{g})$. Hence, identifying $\mathfrak{g}$ with its canonical image under σ_0, $\boldsymbol{U}(\mathfrak{g})$ is generated by the "standard monomials" $E_{i_1}.E_{i_2}...E_{i_k}$, $1 \leq i_1 \leq i_2 \leq ... \leq i_k \leq n$. We have the following result ([MCC 01], 1.7.5; [BOU 82b], Chapter 1, section 2.7, Theorem 1):

THEOREM 6.24.– (Poincaré–Birkhoff–Witt) *The standard monomials form a basis of the* $\mathbf{K}$*-vector space* $\boldsymbol{U}(\mathfrak{g})$.

The $\mathbf{K}$-vector space $\boldsymbol{U}(\mathfrak{g})$ has the canonical structure of a ring of skew polynomials (in several indeterminates), which is Noetherian by a generalized version of the *Basissatz* ([P1], section 3.1.11(**I**)).

By the definition of $\boldsymbol{U}(\mathfrak{g})$, there is a one-to-one correspondence between the set of representations of $\mathfrak{g}$ in a $\mathbf{K}$-vector space V and the set of linear representations of $\boldsymbol{U}(\mathfrak{g})$ in V; furthermore, there is an equivalence between the linear representations of the associative algebra $\boldsymbol{U}(\mathfrak{g})$ and the left $\boldsymbol{U}(\mathfrak{g})$-modules ([BOU 12], Chapter 8, p. 365).

DEFINITION 6.25.– *Let* $\mathfrak{g}$ *be a Lie algebra over* $\mathbf{K}$ *and* $\boldsymbol{U}(\mathfrak{g})$ *its enveloping algebra.*

i) A left $\boldsymbol{U}(\mathfrak{g})$*-module* M *is said to be a* $\mathfrak{g}$*-module.*

ii) A representation ρ *of* $\mathfrak{g}$ *in* M *is said to be* completely reducible *if the corresponding* $\mathfrak{g}$*-module is semi-simple ([P1], section 2.3.5(**III**)).*

(V) ADO'S THEOREM Let $\mathbf{K}$ be a field of characteristic zero. The importance of the subalgebras of $\mathfrak{gl}_n(\mathbf{K})$ resides in the following fundamental result ([BOU 82b], Chapter 1, section 7, Theorem 2).

THEOREM 6.26.– (Ado) *Every finite-dimensional Lie* $\mathbf{K}$*-algebra* $\mathfrak{g}$ *has a faithful finite-dimensional representation. In other words, given any Lie algebra* $\mathfrak{g}$*, there exist an integer* n *and a monomorphism* $\rho : \mathfrak{g} \hookrightarrow \mathfrak{gl}_n(\mathbf{K})$ *that allow* $\mathfrak{g}$ *to be identified with a subalgebra of* $\mathfrak{gl}_n(\mathbf{K})$.

6.3.4. *Nilpotent Lie algebras*

In the following, $\mathbf{K}$ is a field (of characteristic zero from section 6.3.5 onward), and every Lie algebra is finite-dimensional over $\mathbf{K}$.

(I) DEFINITION OF A NILPOTENT LIE ALGEBRA; EXAMPLES

DEFINITION 6.27.– *A Lie algebra* $\mathfrak{g}$ *is said to be* nilpotent *if there exists a finite sequence of ideals* $(\mathfrak{a}_i)_{0\leq i\leq p}$ *of* $\mathfrak{g}$*, where* $\mathfrak{a}_0 = \mathfrak{g}$*,* $\mathfrak{a}_p = \{0\}$*, such that* $[\mathfrak{g}, \mathfrak{a}_i] \subset \mathfrak{a}_{i+1}$ $(0 \leq i \leq p-1)$.

Any commutative Lie algebra is nilpotent. It is possible to show the following result ([BOU 82b], Chapter 1, section 4.1, Proposition 1):

THEOREM 6.28.– *Let* $\mathfrak{g}$ *be a Lie algebra. The following conditions are equivalent:*

i) $\mathfrak{g}$ *is nilpotent.*

ii) $\mathcal{C}^k\mathfrak{g} = \{0\}$ *for* k *sufficiently large.*

iii) There exists a decreasing sequence of ideals $(\mathfrak{b}_i)_{0\leq i\leq n}$ *of* $\mathfrak{g}$*, where* $\mathfrak{b}_0 = \mathfrak{g}$*,* $\mathfrak{b}_n = \{0\}$*, such that* $[\mathfrak{g}, \mathfrak{b}_i] \subset \mathfrak{b}_{i+1}$ *and* $\dim(\mathfrak{b}_i/\mathfrak{b}_{i+1}) = 1$ $(0 \leq i \leq n-1)$.

iv) There exists an integer k *such that* $adX_1 \circ adX_2 \circ ... \circ adX_k = 0$ *for any choice of elements* $X_1, ..., X_k$ *of* $\mathfrak{g}$.

PROOF.– (i)⇔(ii)⇔(iii): **exercise*** (see *loc. cit.*). (ii)⇔(iv): $\mathcal{C}^i\mathfrak{g}$ is the set of linear combinations of elements of the form $[X_1, [X_2, ..., [X_{i-2}, [X_{i-1}, X_i]]]]$, where the $X_1, X_2, ..., X_i$ range over $\mathfrak{g}$. ■

COROLLARY 6.29.– *1) The center $\mathfrak{z}(\mathfrak{g})$ of a non-zero nilpotent Lie algebra $\mathfrak{g}$ is non-zero.*

2) The Killing form $B_{\mathfrak{g}}$ of a nilpotent Lie algebra $\mathfrak{g}$ is zero.

PROOF.– 1): If $\mathfrak{z}(\mathfrak{g}) = 0$, then $\ker(\mathrm{ad}) = 0$, so if $\mathrm{ad}X_1 \circ \mathrm{ad}X_2 \circ ... \circ \mathrm{ad}X_k = 0$, then $X_1 = X_2 = ... = X_k = 0$.

2) If $\mathfrak{g}$ is nilpotent, then $A := \mathrm{ad}X \circ \mathrm{ad}Y$ is nilpotent for every $X, Y \in \mathfrak{g}$ by Theorem 6.28. Since $\mathfrak{g}$ is finite-dimensional, ad$\mathfrak{g}$ can be identified with an algebra of square matrices (section 5.4.1(**IV**)), so A can be identified with a nilpotent matrix. The canonical Jordan form J of A over an algebraic closure of $\mathbf{K}$ ([P1], section 3.4.3(**IV**)) has zeros along the diagonal, so $\mathrm{Tr}(A) = \mathrm{Tr}(J) = 0$. ■

Any subalgebra, quotient algebra, central extension of a nilpotent Lie algebra or finite product of nilpotent algebras is a nilpotent Lie algebra (**exercise***: see [BOU 82b], Chapter 1, section 4.1, Proposition 2). The nilpotent Lie algebras are precisely the Lie algebras obtained from commutative Lie algebras by a sequence of central extensions (**exercise**).

EXAMPLE 6.30.– *Let $\mathfrak{l}_n(\mathbf{K})$ be the subalgebra of $\mathfrak{gl}_n(\mathbf{K})$ formed by the matrices of the form $\lambda.I_n + A$ $(\lambda \in \mathbf{K}, A \in \mathfrak{n}_n(\mathbf{K}))$. Writing $\mathfrak{m}_n(\mathbf{K})$ for the commutative (and hence nilpotent) Lie algebra formed by the matrices of the form $\lambda.I_n$ $(\lambda \in \mathbf{K})$, we have $\mathfrak{l}_n(\mathbf{K}) = \mathfrak{m}_n(\mathbf{K}) \oplus_\sigma \mathfrak{n}_n(\mathbf{K})$ (section 5.4.1(**IV**)) and $\mathfrak{l}_n(\mathbf{K})$ is a central extension of $\mathfrak{n}_n(\mathbf{K})$ by $\mathfrak{m}_n(\mathbf{K})$. Therefore, the Lie algebra $\mathfrak{l}_n(\mathbf{K})$ is nilpotent.*

(II) ADO'S THEOREM AND ENGEL'S THEOREM The following theorem is assumed without proof ([BOU 82b], Chapter 1, section 7, Theorem 3 and Exercise 1(a); section 4.2, Corollary 1). Part (1) extends Theorem 6.26:

THEOREM 6.31.– *1)* (Ado) *Suppose that $\mathbf{K}$ is of characteristic zero, let $\mathfrak{g}$ be a Lie algebra, and let $\mathfrak{n}$ (also written as $\mathfrak{N}(\mathfrak{g})$) be the largest nilpotent ideal of $\mathfrak{g}$. There exists a faithful finite-dimensional representation ρ of $\mathfrak{g}$ such that every element of $\rho(\mathfrak{n})$ is nilpotent.*

*2) (*Engel*) Let $\mathfrak{g}$ be a Lie algebra. The conditions stated below satisfy the implications (iii)⇒(i)⇔(ii); furthermore, (ii)⇒(iii) whenever $\mathbf{K}$ is of characteristic zero:*

i) $\mathfrak{g}$ is nilpotent;

ii) for every $X \in \mathfrak{g}$, adX is nilpotent; in other words, there exists an integer k such that $\mathfrak{g}^k = 0$, where $\mathfrak{g}^k := \underbrace{[\mathfrak{g}[\mathfrak{g}...[\mathfrak{g},\mathfrak{g}]]]}_{k \text{ terms}}$;

iii) $\mathfrak{g}$ is isomorphic to a subalgebra of the algebra $\mathfrak{n}_n(\mathbf{K})$ (section 6.3.2.10).

EXAMPLE 6.32.– *With the notation of Example 6.30,* $\mathfrak{m}_1(\mathbf{K}) \cong \mathfrak{n}_2(\mathbf{K}) : \lambda \mapsto \begin{bmatrix} 0 & \lambda \\ 0 & 0 \end{bmatrix}$.

The smallest nilpotent ideal $\mathfrak{n}$ of a Lie algebra $\mathfrak{g}$ is orthogonal to $\mathfrak{g}$ for the Killing form (**exercise**).

(III) EXTENSION OF THE BASE FIELD If the field extension $\mathbf{L}/\mathbf{K}$ is separable (necessarily true whenever $\mathbf{K}$ is of characteristic zero: see [P2], section 1.1.5**(III)**) and $\rho : \mathbf{K} \hookrightarrow \mathbf{L}$ is the canonical injection, then $\mathfrak{g}$ is nilpotent if and only if $\rho^*(\mathfrak{g})$ is nilpotent. Furthermore, $\rho^*(\mathfrak{N}(\mathfrak{g})) = \mathfrak{N}(\rho^*(\mathfrak{g}))$ ([BOU 82b], Chapter 1, section 4.5).

6.3.5. *Solvable Lie algebras*

Recall that $\mathbf{K}$ is assumed to be of characteristic zero.

(I) NOTION OF A SOLVABLE ALGEBRA

DEFINITION 6.33.– *A Lie algebra* $\mathfrak{g}$ *is said to be* solvable *if* $\mathcal{D}^k\mathfrak{g} = \{0\}$ *for sufficiently large* k.

In the words of M. Hausner and J.T. Schwartz ([HAU 68], p. viii):

> Unfortunately, the structure theory of solvable algebras, like most algebraic theories of structures which are neither commutative nor semi-simple, remains, in spite of some successful partial analyses, largely a mystery, perhaps even an impenetrable mystery.

Nevertheless, there are a few things that we can say. Every nilpotent Lie algebra is solvable. Any subalgebra or quotient algebra of a solvable Lie algebra, finite product of solvable Lie algebras or extension of a solvable Lie algebra by a solvable Lie algebra is solvable (**exercise***: see [BOU 82b], Chapter 1, section 5.1, Proposition 1).

(II) LEVI SUBALGEBRA Given a Lie algebra $\mathfrak{g}$, let $B_{\mathfrak{g}}$ be its Killing form, and let $\check{\mathfrak{t}} = \check{\mathfrak{T}}(\mathfrak{g})$ be the orthogonal complement of $\mathfrak{g}$ for $B_{\mathfrak{g}}$. If $\mathfrak{a}, \mathfrak{b}$ are solvable ideals of $\mathfrak{g}$, then $\mathfrak{a} + \mathfrak{b}$ is a solvable ideal as an extension of $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$ by $\mathfrak{b}$, and $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$ is solvable. The sum of every solvable ideal is therefore a solvable ideal, which is the largest solvable ideal.

DEFINITION 6.34.– *The* radical *(respectively* nilpotent radical*) of a Lie algebra* $\mathfrak{g}$ *is its largest solvable ideal* $\mathfrak{r}$*, also written as* $\mathfrak{R}(\mathfrak{g})$ *(respectively the ideal* $\mathfrak{s}$*, also written as* $\mathfrak{S}(\mathfrak{g})$*, which is equal to* $\mathfrak{r} \cap \mathcal{D}\mathfrak{g}$*). A* Levi subalgebra *of* $\mathfrak{g}$ *is any subalgebra that is supplementary to* $\mathfrak{r}$ *in* $\mathfrak{g}$ *as a* $\mathbf{K}$*-vector space.*

PROPOSITION 6.35.– *1) The ideals* $\mathfrak{s}, \mathfrak{n}, \check{\mathfrak{k}}, \mathfrak{r}$ *are all characteristic (section 5.4.1**(II)**), and* $\mathfrak{s} \subset \mathfrak{n} \subset \check{\mathfrak{k}} \subset \mathfrak{r}$ *(**exercise***: see [BOU 82b], Chapter 1, p. 70).*

2) $\mathfrak{r}$ *is the orthogonal of* $\mathcal{D}\mathfrak{g}$ *for the Killing form, and* $\mathfrak{s} = [\mathfrak{r}, \mathfrak{g}]$ *([BOU 82b], Chapter 1, section 5.5, Proposition 5; section 6.4, Proposition 6).*

*3) In the situation of section 6.3.4**(III)**), we have* $\mathfrak{S}(\rho^*(\mathfrak{g})) = \rho^*(\mathfrak{S}(\mathfrak{g}))$, $\check{\mathfrak{T}}(\rho^*(\mathfrak{g})) = \rho^*\left(\check{\mathfrak{T}}(\mathfrak{g})\right)$, $\mathfrak{R}(\rho^*(\mathfrak{g})) = \rho^*(\mathfrak{R}(\mathfrak{g}))$. *The Killing form* $B_{\rho^*}(\mathfrak{g})$ *is non-degenerate if and only if the Killing form* $B_{\mathfrak{g}}$ *is non-degenerate.*

PROOF.– (3) follows from section 6.3.3(**V**). ■

If $\mathfrak{a}$ is an ideal of the Lie algebra $\mathfrak{g}$, then $\mathfrak{R}(\mathfrak{a}) = \mathfrak{R}(\mathfrak{g}) \cap \mathfrak{a}$; $\mathfrak{R}(\mathfrak{g})$ is the smallest ideal $\mathfrak{r}$ of $\mathfrak{g}$ such that $\mathfrak{R}(\mathfrak{g}/\mathfrak{r}) = \{0\}$ (**exercise***: see [BOU 82b], Chapter 1, section 5.2, Proposition 3 and section 5.5, Corollary 3).

It is possible to show the following result ([BOU 82b], Chapter 1, section 5.8, Theorem 5):

THEOREM 6.36.– (Levi–Malcev) *Every Lie algebra has a Levi subalgebra (not unique in general)* $\mathfrak{l}$ *that is isomorphic to* $\mathfrak{g}/\mathfrak{r}$. *Up to isomorphism,* $\mathfrak{g}$ *is the semi-direct sum of* $\mathfrak{r}$ *and* $\mathfrak{l}$: $\mathfrak{g} \cong \mathfrak{r} \oplus_\sigma \mathfrak{l}$ *("Levi–Malcev decomposition").*

(III) LIE'S THEOREM AND CARTAN'S CRITERION

THEOREM 6.37.– *Let* $\mathfrak{g}$ *be a Lie algebra. The following conditions are equivalent:*

i) $\mathfrak{g}$ *is solvable;*

*ii) (*S. Lie's theorem*)* $\mathcal{D}\mathfrak{g}$ *is nilpotent;*

*iii) (*Cartan's criterion*)* $\mathcal{D}\mathfrak{g} \subset \check{\mathfrak{k}}$ *(i.e.* $B_{\mathfrak{g}}(X, Y) = 0$ *for every* $X \in \mathcal{D}\mathfrak{g}$ *and every* $Y \in \mathfrak{g}$*).*

PROOF.– (i)⇔(ii): see [BOU 82b], Chapter 1, section 5.3, Corollary 5. (i)⇔(iii): this immediately follows from Proposition 6.35(2). ■

We have $\mathcal{D}\mathfrak{t}_n(\mathbf{K}) = \mathfrak{n}_n(\mathbf{K})$ (**exercise**), so $\mathfrak{t}_n(\mathbf{K})$ is solvable. Hence, $\mathfrak{st}_n(\mathbf{K})$ is solvable.

We have $\mathfrak{d}_n(\mathbf{K}) = \mathfrak{R}(\mathfrak{se}_n(\mathbf{K}))$, so $\mathfrak{d}_n(\mathbf{K}) \oplus_\sigma \mathfrak{o}_n(\mathbf{K})$ is a Levi–Malcev decomposition of $\mathfrak{se}_n(\mathbf{K})$.

COROLLARY 6.38.– *Suppose that* $\mathbf{K}$ *is algebraically closed. For a Lie* $\mathbf{K}$*-algebra to be solvable, it is necessary and sufficient for it to be isomorphic to a subalgebra of* $\mathfrak{t}_n(\mathbf{K})$ *(**exercise**).*

(IV) EXTENSION OF THE BASE FIELD In the situation of section 6.3.4(**III**)), we have $\mathcal{D}^k \rho^* (\mathfrak{g}) = \rho^* \left(\mathcal{D}^k \mathfrak{g}\right)$, so $\mathfrak{g}$ is solvable if and only if $\rho^* (\mathfrak{g})$ is solvable.

6.3.6. *Simple and semi-simple Lie algebras*

(I) NOTION OF A SEMI-SIMPLE LIE ALGEBRA In the following, the field $\mathbf{K}$ is of characteristic zero. Unless otherwise stated, every algebra is a $\mathbf{K}$-algebra and every vector space is a $\mathbf{K}$-vector space.

DEFINITION 6.39.– *A Lie algebra* $\mathfrak{g}$ *is said to be*

– simple *if it is not commutative and its unique proper ideal is trivial;*

– semi-simple *if its unique commutative ideal is trivial.*

The algebra $\{0\}$ is semi-simple but not simple. It is possible to show the following result ([BOU 82b], Chapter 1, section 6, Theorem 1 and Proposition 2):

THEOREM 6.40.– *Let* $\mathfrak{g}$ *be a Lie algebra. The following conditions are equivalent:*

i) $\mathfrak{g}$ *is semi-simple;*

ii) $\mathfrak{g}$ *is a finite product of simple Lie algebras;*

*iii) (*Cartan's criterion*) the Killing form of* $\mathfrak{g}$ *is non-degenerate;*

iv) the radical $\mathfrak{r}$ *of* $\mathfrak{g}$ *is zero.*

COROLLARY 6.41.– *Any Levi subalgebra of a Lie algebra is semi-simple. Conversely, any semi-simple Lie algebra coincides with its unique Levi algebra.*

PROOF.– Let $\mathfrak{l}$ be a Levi subalgebra of $\mathfrak{g}$. Then, $\mathfrak{l} \cong \mathfrak{g}/\mathfrak{r}$, where $\mathfrak{r} = \mathfrak{R}(\mathfrak{g})$, so $\mathfrak{R}(\mathfrak{l}) = \{0\}$. ■

The Lie algebras $\mathfrak{o}_n(\mathbf{K}), \mathfrak{sl}_n(\mathbf{K}), \mathfrak{sp}_{2n}(\mathbf{K})$ are semi-simple for $n \geq 2$ because their Killing forms are non-degenerate (**exercise**). It can be shown that $\mathfrak{sl}_n(\mathbf{K})$ is simple for $n \geq 2$ ([BOU 82b], Chapter 1, section 6, Exercise 21).

If $\mathfrak{g}$ is a semi-simple Lie algebra, then every ideal and every quotient algebra of $\mathfrak{g}$ is semi-simple; every extension of $\mathfrak{g}$ by a semi-simple Lie algebra is semi-simple and trivial; every derivation of $\mathfrak{g}$ is an inner derivation ([BOU 82b], Chapter 1, section 6); furthermore, $\mathfrak{g} = \mathcal{D}\mathfrak{g}$ and $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

COROLLARY 6.42.– *In the situation of section 6.3.4(**III**), the Lie algebra* $\mathfrak{g}$ *is semi-simple if and only if* $\rho^*(\mathfrak{g})$ *is semi-simple (**exercise**). If* $\rho^*(\mathfrak{g})$ *is simple, then* $\mathfrak{g}$ *is* *simple (**exercise**), but the converse does not hold, even when $\mathbf{K} = \mathbb{R}$ and $\mathbf{L} = \mathbb{C}$ ([BOU 82b], Chapter 1, section 6, Exercise 17).*

Let $\mathfrak{g}$ be a Lie algebra. The *normalizer* of a subspace $\mathfrak{a}$ of $\mathfrak{g}$ is defined as the set of $X \in \mathfrak{g}$ such that $(\mathrm{ad}X) .\mathfrak{a} \subset \mathfrak{a}$.

DEFINITION 6.43.– *We say that a subalgebra $\mathfrak{h}$ of the Lie algebra $\mathfrak{g}$ is a* Cartan subalgebra *of $\mathfrak{g}$ if $\mathfrak{h}$ is nilpotent and the normalizer of $\mathfrak{h}$ coincides with $\mathfrak{h}$.*

Every Lie algebra $\mathfrak{g}$ over $\mathbf{K}$ has a Cartan subalgebra $\mathfrak{h}$, and all Cartan subalgebras of $\mathfrak{g}$ have the same dimension ([BOU 82b], Chapter 7, section 2.3, Corollary 1). We say that $\sigma \in \mathrm{Aut}(\mathfrak{g})$ is an *elementary automorphism* if it is of the form $\exp(\mathrm{ad}X_1) ... \exp(\mathrm{ad}X_r)$, where $X_i \in \mathfrak{g}$ and $\mathrm{ad}X_i$ is nilpotent for $i = 1, ..., r$; the elementary automorphisms form a group[4] $\mathcal{E}(\mathfrak{g})$, and Chevalley showed that, if $\mathbf{K}$ is algebraically closed, $\mathcal{E}(\mathfrak{g})$ is normal in $\mathrm{Aut}(\mathfrak{g})$ and acts transitively on the set of Cartan subalgebras of $\mathfrak{g}$ ([BOU 82b], Chapter 7, section 3.2, Theorem 1); hence, any Cartan subalgebra is unique up to conjugation by an elementary automorphism σ; in other words, if $\mathfrak{h}_1, \mathfrak{h}_2$ are two Cartan subalgebras of $\mathfrak{g}$, there exists $\sigma \in \mathcal{E}(\mathfrak{g})$ such that $\mathfrak{h}_2 = \sigma(\mathfrak{h}_1)$.

It is possible to show the following result ([VAR 84], Theorem 4.1.5):

THEOREM 6.44.– *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h}$ a subalgebra of $\mathfrak{g}$. Then, $\mathfrak{h}$ is a Cartan subalgebra if and only if (i) $\mathfrak{h}$ is a maximal commutative subalgebra of $\mathfrak{g}$ and (ii) $\mathrm{ad}_{\mathfrak{g}}X$ is semi-simple (section 6.3.3(**I**)(a)) for every $X \in \mathfrak{h}$. If so, the restriction $B_{\mathfrak{h}}$ of the Killing form $B_{\mathfrak{g}}$ to $\mathfrak{h} \times \mathfrak{h}$ is non-degenerate.*

EXAMPLE 6.45.– *Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbf{K})$. The set $\{E_1, E_2, E_3\}$ is a basis of $\mathfrak{g}$, where $E_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. In this basis, the endomorphisms $\mathrm{ad}E_1$ and $\mathrm{ad}E_2$ are represented by the matrices $A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, respectively (section 5.4.1(**III**)). The subalgebra $\mathfrak{h} = \mathbf{K}.E_1$ is maximal commutative and $\mathrm{ad}E_1$ is semi-simple, so $\mathfrak{h}$ is a Cartan subalgebra. However, $\mathbf{K}.E_2$ is a maximal commutative subalgebra but is not a Cartan subalgebra, because $\mathrm{ad}E_2$ is not semi-simple: this endomorphism has the unique elementary divisor $\pi(\lambda) = \lambda^3$ (**exercise**).*

If $\mathfrak{g}$ is solvable, its only Cartan subalgebra is $\mathfrak{g}$ (**exercise**). We have the following result ([BOU 82b], Chapter 1, section 6.2, Theorem 2):

4 The exponential of a nilpotent endomorphism $\mathbf{u}$ from a finite-dimensional vector space V over a field of characteristic zero is defined by the Taylor expansion $e^{\mathbf{u}} = 1_V + \frac{\mathbf{u}}{1!} + ... + \frac{\mathbf{u}^k}{k!}$, where k is the smallest integer such that $\mathbf{u}^{k+1} = 0$.

THEOREM 6.46.– (H. Weyl's complete reducibility theorem) *Every finite-dimensional representation of a semi-simple Lie algebra is completely reducible (Definition 6.25).*

Conversely, if every finite-dimensional representation of $\mathfrak{g}$ is completely reducible, then $\mathfrak{g}$ is semi-simple (*ibid.*, Remark, p. 75).

(II) ROOT SYSTEMS Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra.

DEFINITION 6.47.– *We say that* $(\mathfrak{g}, \mathfrak{h})$ *is* split *if, for every* $H \in \mathfrak{h}$, $ad_{\mathfrak{g}}H$ *is diagonalizable. We say that* $\mathfrak{g}$ *is* splitting *if there exists a Cartan subalgebra* $\mathfrak{h}$ *for which* $(\mathfrak{g}, \mathfrak{h})$ *is split.*

We know that $\mathrm{ad}_{\mathfrak{g}}H$ is semi-simple (Theorem 6.44). Hence, if $\mathbf{K}$ is algebraically closed, then $(\mathfrak{g}, \mathfrak{h})$ is split.

DEFINITION 6.48.– *Suppose that* $(\mathfrak{g}, \mathfrak{h})$ *is split. The* roots $\alpha(H) \in \mathbf{K}$ *of* $(\mathfrak{g}, \mathfrak{h})$ *are the eigenvalues of* $ad_{\mathfrak{g}}H$ $(H \in \mathfrak{h})$. *The corresponding eigenvectors are called the* root vectors.

Suppose that $(\mathfrak{g}, \mathfrak{h})$ is split. If $\alpha(H)$ is a root of $\mathfrak{g}$ and Y_α is a corresponding root vector, then $[H, Y_\alpha] = \alpha(H).Y_\alpha$, and the mappings $\alpha : \mathfrak{h} \to \mathbf{K} : X \mapsto \alpha(X)$ are therefore linear forms on $\mathfrak{h}$, i.e. elements of the dual $\mathfrak{h}^\vee$ of $\mathfrak{h}$. If $\alpha \in \mathfrak{h}^\vee$, let

$$\mathfrak{g}_\alpha = \left\{ X \in \mathfrak{g} : \underbrace{(\mathrm{ad}_{\mathfrak{g}}H).X}_{[H,X]} = \langle \alpha, H \rangle .X \right\}$$

be the eigenspace corresponding to the eigenvalue $\langle \alpha, H \rangle$ $(\alpha \in \mathfrak{h}^\vee, H \in \mathfrak{h} - \{0\})$, and let Δ be the set of non-zero roots. Then, $\mathfrak{h} = \mathfrak{g}_0$, and $\mathfrak{g}$ can be expressed as follows:

$$\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus\nolimits_{\alpha \in \Delta} \mathfrak{g}_\alpha \right). \qquad [6.4]$$

The vector space $\mathfrak{g}_\alpha$ is said to be the *root subspace* associated with α and the direct sum [6.4] is known as the decomposition of $\mathfrak{g}$ into root subspaces relative to $\mathfrak{h}$. Each $\mathfrak{g}_\alpha$ $(\alpha \in \Delta)$ has dimension 1.

Let us give some hints about root systems, a notion which goes back to the works on crystallography of the years 1830–1840, and later used by Killing, then Weyl and Cartan about 1925 ([BOU 82b], Chapter 6, section 1).

DEFINITION 6.49.– *Let* V *be a finite-dimensional vector space over a field* $\mathbf{K}$ *of characteristic zero (hence,* $\mathbf{K} \supset \mathbb{Q}$*) and* R *a part of* V. *We say that* R *is a* root system *in* V *if the following conditions are satisfied:*

i) R is finite, does not contain 0 and generates V;

ii) for every $\alpha \in R$, there exists $\alpha^\vee \in V^\vee$ such that $\langle \alpha, \alpha^\vee \rangle = 2$ and R is stable under the "Weyl reflection" $s_{\alpha,\alpha^\vee} : \beta - \langle \beta, \alpha^\vee \rangle .\alpha$;

iii) for every $\alpha \in R$, $\alpha^\vee (R) \subset \mathbb{Z}$.

It is easy to see that, if the conditions of Definition 6.49 are satisfied, the linear form $\alpha^\vee$ is uniquely determined by α; thus, the Weyl reflection s_α can be denoted $s_{\alpha,\alpha^\vee}$. The group $A(R)$ of automorphisms of V is finite, since it can be identified with $\mathfrak{S}_B$ for any basis B of V ([P1], section 2.2.1(**I**)), and the subgroup $W(R)$ generated by the s_α is called the *Weyl group* of R. The symmetric non-degenerate bilinear form $(x|y)$ that is invariant under $W(R)$ is given by

$$(x|y) = \sum_{\alpha \in R} \langle \alpha^\vee, x \rangle \langle \alpha^\vee, y \rangle .$$

Using this form, V can be identified with $V^\vee$; in particular, if $\alpha \in R$, $\alpha^\vee = \frac{2\alpha}{(\alpha|\alpha)}$.

DEFINITION 6.50.– *A root system $R \neq \varnothing$ is* reducible *if $W(R)$ is of the form $\prod_{i=1}^{r} W(R_i)$, where $r \geq 2$ and $R_i \neq \varnothing$; it is* irreducible *otherwise. It is* reduced *if, for every root $\alpha \in R$, $\alpha/2 \notin R$.*

When $\mathbf{K} = \mathbb{R}$, the following result holds ([BOU 82b], Chapter 8, sections 2.2 and 3.2):

THEOREM 6.51.– *The set Δ in [6.4] is a reduced root system of $\mathfrak{h}^\vee$ that determines $\mathfrak{g}$ up to isomorphism, and every reduced root system is isomorphic to a root system obtained in this way. The Lie algebra $\mathfrak{g}$ is simple if and only if Δ is irreducible.*

(III) CARTAN'S CLASSIFICATION In a thesis published in 1894, following a profound but incomplete and occasionally incorrect work by Killing (from 1888 to 1890), É. Cartan showed that the classification of complex semi-simple Lie algebras reduces to the study of their reduced root systems, which in turn reduces to the study of their *Cartan matrices*, whose elements a_i^j are integers (equal to 2 for $i = j$ and to $0, -1, -2$ or -3 for $i \neq j$) that can be expressed as a function of the reduced roots α_i according to $a_i^j = 2\frac{(\alpha_i|\alpha_j)}{(\alpha_j|\alpha_j)}$

Cartan matrices are non-symmetric in general but satisfy $a_j^i = 0$ whenever $a_i^j = 0$ ([BOU 82b], Chapter 6, sections 1.3 and 1.5).

Cartan's classification was extended and simplified by Weyl (1926), van der Waerden (1933), Coxeter (1934) and Witt (1941), as well as Dynkin with the introduction of Dynkin diagrams in 1950, which (like Coxeter's graphs) enable us to represent reduced root systems graphically[5]. Connected Dynkin diagrams correspond to simple Lie algebras. In essence, Weyl and van der Waerden showed that each Dynkin diagram can only correspond to one semi-simple algebra up to isomorphism; Harish-Chandra, based on a contribution by Chevalley in 1948, showed the converse in 1951 [HAR 51]: each Dynkin diagram corresponds to a semi-simple Lie algebra. Between 1962 and 1964, Tits completed an algebraic formulation of these methods that we do not have the space to present here: see [BOU 82b], Chapter 6.

However, Cartan had already obtained a remarkable result as early as 1894, a true *tour de force* – the enumeration of every possible complex simple Lie algebra (a result that also holds for every splitting Lie algebra over any field $\mathbf{K}$ of characteristic zero): up to isomorphism, any such algebra must either be one of the infinite sequences of "principal" algebras A_{n-1}, B_n $(n \geq 2)$, C_n $(n \geq 3)$ and D_n $(n \geq 4)$ (section 6.3.2), of dimensions $n^2 - 1$, $n(2n+1)$, $n(2n+1)$ and $n(2n-1)$, respectively, or one of the "exceptional" algebras G_2, F_4, E_6, E_7 and E_8, of dimensions 14, 52, 78, 133 and 248, respectively ([JAC 62], Chapter 4). A Lie algebra of any one of these types is not isomorphic to a Lie algebra of any other of these types. Since $D_2 \cong A_1 \times A_1$, this Lie algebra is semi-simple but not simple. For each of the algebras A_{n-1}, B_n, C_n $(n \geq 2)$, D_n $(n \geq 3)$, the diagonal elements form a split Cartan subalgebra ([BOU 82b], Chapter 8, sections 13.1, 13.3 and section 13, Exercise 4).

It should be noted that fundamental applications in physics have been found for *every single one* of these Lie algebras [SAT 86, BIN 13].

Recall that a real Lie algebra is semi-simple if and only if its complexification is semi-simple (Corollary 6.42). The classification of every simple Lie algebra over a non-algebraically closed field $\mathbf{K}$ of characteristic zero exceeds the scope of this book. It was established in full by É. Cartan in 1914 in the case where $\mathbf{K} = \mathbb{R}$ (see [HAU 68], Part 3), and the general case was established by N. Jacobson (1938) [JAC 38].

6.3.7. *Reductive Lie algebras*

As before, $\mathbf{K}$ is a field of characteristic zero, and every Lie algebra is assumed to be finite-dimensional over $\mathbf{K}$.

5 For more on the history of semi-simple Lie algebras, see the historical notes in [BOU 82b], Chapters 4–6. For the period between Lie and Weyl, see [HAW 00]. Coxeter graphs are equivalent to Dynkin diagrams but additionally present the types $H_3, H_4, I_2(p)$ $(p = 5$ or $p \geq 7)$, which do not correspond to any of the "crystallographic" Coxeter groups, i.e. the symmetries of the groups of roots of semi-simple Lie algebras.

DEFINITION 6.52.– *A Lie algebra* $\mathfrak{g}$ *is said to be reductive if its adjoint representation is semi-simple.*

Any semi-simple Lie algebra is therefore reductive. We have the following result ([BOU 82b], Chapter 1, section 6.4, Proposition 5):

THEOREM 6.53.– *Let* $\mathfrak{g}$ *be a Lie algebra. The following conditions are equivalent:*

i) $\mathfrak{g}$ *is reductive;*

ii) $\mathcal{D}\mathfrak{g}$ *is semi-simple;*

iii) $\mathfrak{g}$ *is the product of a semi-simple algebra and a commutative algebra;*

iv) the largest solvable ideal $\mathfrak{r}$ *is equal to* $\mathfrak{z}(\mathfrak{g})$.

For example, $\mathfrak{gl}_n(\mathbf{K}) \cong \mathfrak{sl}_n(\mathbf{K}) \times \mathfrak{d}_n(\mathbf{K})$, $\mathfrak{sl}_n(\mathbf{K})$ is semi-simple, and $\mathfrak{d}_n(\mathbf{K})$ is commutative, so $\mathfrak{gl}_n(\mathbf{K})$ is reductive. The center $\mathfrak{z}(\mathfrak{gl}_\mathfrak{n}(\mathbf{K}))$ is the set of homotheties $\mathfrak{gl}_\mathfrak{n}(\mathbf{K}) \to \mathfrak{gl}_\mathfrak{n}(\mathbf{K}) : X \mapsto \lambda.X$ $(\lambda \in \mathbf{K})$.

Any product or quotient of reductive Lie algebras or center of a reductive Lie algebra is a reductive Lie algebra. If $\mathfrak{g}$ is a reductive Lie algebra and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{a}$ is a direct factor and is a reductive Lie algebra.

In the situation of section 6.3.4(**III**)), a Lie algebra $\mathfrak{g}$ is reductive if and only if $\rho^*(\mathfrak{g})$ is reductive (**exercise**).

6.3.8. *Real compact Lie algebras*

Ever since H. Weyl, a *real* Lie algebra is said to be *compact* if it is reductive and its Killing form is semi-definite negative ([BOU 82b], Chapter 9, section 1.3). A reductive Lie algebra $\mathfrak{g}$ is compact if and only if, for every $X \in \mathfrak{g}$, the endomorphism ad$X$ is semi-simple and its eigenvalues are purely imaginary. The real compact Lie algebras are the product algebras of a commutative algebra and a compact semi-simple algebra.

REMARK 6.54.– *Some authors define a Lie algebra as compact if its Killing form is negative definite. Any such algebra is semi-simple by Cartan's criterion (Theorem 6.40). There exist non-commutative Lie algebras for which the Killing form is zero. One example is Heisenberg's algebra ([SAT 86], Chapter 4, section 13).*

The following result is due to Weyl ([BOU 82b], Chapter 9, section 3.3, Corollary 2):

THEOREM 6.55.– *Let $\mathfrak{a}$ be a complex Lie algebra. The following conditions are equivalent:*

i) $\mathfrak{a}$ is reductive;

ii) there exists a real compact Lie algebra $\mathfrak{g}$ whose complexification $\mathfrak{g}_{(\mathbb{C})}$ is isomorphic to $\mathfrak{a}$.

Furthermore (*ibid.*, Corollary 3), if $\mathfrak{a}_1, \mathfrak{a}_2$ are semi-simple complex Lie algebras, then the real compact forms of $\mathfrak{a}_1 \times \mathfrak{a}_2$ (section 6.3.1(**II**)) are the products $\mathfrak{g}_1 \times \mathfrak{g}_2$ such that $\mathfrak{g}_i$ is a real compact form of $\mathfrak{a}_i$ for $i = 1, 2$. Hence:

COROLLARY 6.56.– *A real compact Lie algebra $\mathfrak{g}$ is simple if and only if its complexification $\mathfrak{g}_{(\mathbb{C})}$ is simple.*

Thus, it is not ambiguous to speak of a simple real compact Lie algebra of type A_{n-1}, B_n $(n \geq 2)$, C_n $(n \geq 3)$, D_n $(n \geq 4)$ $, G_2, F_4, E_6, E_7$ or E_8. These are the real compact forms of the corresponding complex Lie algebras. In particular (see section 6.3.2), $\mathfrak{su}_{n-1}, \mathfrak{o}_{2\mathfrak{n}+1}$ $(n \geq 2)$ and $\mathfrak{o}_{2n}$ $(n \geq 4)$ are real compact algebras whose complexifications are $\mathfrak{sl}_{n-1}(\mathbb{C}), \mathfrak{o}_{2n+1}(\mathbb{C})$ and $\mathfrak{o}_{2n}(\mathbb{C})$, of types A_{n-1}, B_n and D_n, respectively.

6.4. Relation between Lie groups and Lie algebras

In this section, every Lie group is Banach, real or complex, and of class C^ω, unless otherwise stated.

6.4.1. *Lie algebra of a Lie group*

(I) INFINITESIMAL ALGEBRA OF A LIE GROUP Let $\mathbf{G}$ be a Lie group. The convolution algebra $\mathcal{T}_e^{(\infty)}(\mathbf{G})$ (section 6.2.1(**V**)), equipped with the Lie bracket $[T, S] = T \star S - S \star T$ is a Lie algebra (Lemma 5.11).

DEFINITION 6.57.– *Let $P \in$ Diff$(\mathbf{G})$ be a differential operator (section 5.2.2(**I**)). For every $s \in \mathbf{G}$ and every function $f \in \mathcal{E}(\mathbf{G})$, set $(\boldsymbol{\lambda}(s) P).f = \boldsymbol{\lambda}(s)$ $\left(P.\left(\boldsymbol{\lambda}\left(s^{-1}\right) f\right)\right)$. The differential operator P is said to be* left invariant *if $\boldsymbol{\lambda}(s) P = P$ for every $s \in \mathbf{G}$. Right invariance is defined similarly, replacing $\boldsymbol{\lambda}$ by $\boldsymbol{\rho}$.*

LEMMA 6.58.– *If $s \in \mathbf{G}$ and $P, Q \in$ Diff$(\mathbf{G})$, then $\boldsymbol{\lambda}(s)(P \circ Q) = \boldsymbol{\lambda}(s)(P) \circ \boldsymbol{\lambda}(s)(Q)$ (**exercise**).*

LEMMA-DEFINITION 6.59.– *The subset of Diff* $(\mathbf{G})$ *formed by the left invariant differential operators is a subalgebra* $\mathfrak{G}$ *of Diff* $(\mathbf{G})$*, called the* infinitesimal algebra *of* $\mathbf{G}$.

PROOF.– It is clear that $\mathfrak{G}$ is a vector subspace of Diff $(\mathbf{G})$. It is also a subalgebra by Lemma 6.58. ■

COROLLARY 6.60.– *The mapping* $\mathfrak{E} : P \to P(e)$ *(section 5.2.2(**III**)) is an isomorphism (of associative algebras) from* $\mathfrak{G}$ *onto* $\mathcal{T}_e^{(\infty)}(\mathbf{G})$.

PROOF.– By [5.5], $\mathfrak{E}$ is linear and surjective. If $P, Q \in \mathfrak{G}$, then $(P \circ Q)(e) = P(e) \star Q(e)$, so $\mathfrak{E}$ is an epimorphism of algebras. Any operator $P \in \mathfrak{G}$ is uniquely determined by $P(e)$, since, for every $s \in \mathbf{G}$, $(P.f)(s) = s.\left(P.\left(\boldsymbol{\lambda}\left(s^{-1}\right)f\right)(e)\right)$, so $\mathfrak{E}$ is injective. ■

(II) LIE ALGEBRA Let $\mathbf{G}$ be a Lie group and $X \in \mathcal{T}_0^1(\mathbf{G})$. For every $s \in \mathbf{G}$ and every function $f \in \mathcal{E}(\mathbf{G})$, set $X.f = \mathcal{L}_X.f$ (section 5.4.2) and $(\boldsymbol{\lambda}(s)X).f = \boldsymbol{\lambda}(s)\left(X.\left(\boldsymbol{\lambda}\left(s^{-1}\right)f\right)\right)$ (similar to Definition 6.57, *mutatis mutandis*).

DEFINITION 6.61.– *A vector field* $X \in \mathcal{T}_0^1(\mathbf{G})$ *is said to be left invariant if* $(\boldsymbol{\lambda}(s)X)(x) = X(s.x)$ *for every* $s, x \in \mathbf{G}$.

Let $X, Y \in \mathcal{T}_0^1(\mathbf{G})$ be left invariant vector fields. Then, $[X, Y]$ is left invariant (**exercise**). The Lie group $\mathbf{G}$ acts analytically on itself by left translation, so, if $\mathbf{u} \in T_e(\mathbf{G})$, this tangent vector uniquely determines a vector field $X_{\mathbf{u}} : s \mapsto s.\mathbf{u}$ (with the notation from [2.12] in section 2.4.2(**I**)) that is left invariant. Conversely, any left invariant vector field $X \in \mathcal{T}_0^1(\mathbf{G})$ uniquely determines the element $X(e) \in T_e(\mathbf{G})$. Hence, the mapping $\mathcal{T}_0^1(\mathbf{G}) \tilde{\to} T_e(\mathbf{G}) : X \mapsto X(e)$ is an isomorphism of vector spaces, and so we obtain the following result:

THEOREM 6.62.– (Lie's third fundamental theorem) $T_e(\mathbf{G})$*, equipped with the bracket* $[X(e), Y(e)] := [X, Y](e)$ *for all* $X, Y \in \mathcal{T}_0^1(\mathbf{G})$*, is a Lie algebra.*

DEFINITION 6.63.– *The tangent space* $T_e(\mathbf{G})$*, equipped with the above bracket, is called the* Lie algebra *of* $\mathbf{G}$ *and is written as Lie* $(\mathbf{G})$.

Suppose that the associative algebras $\mathfrak{G}$ and $\mathcal{T}_e^{(\infty)}(\mathbf{G})$ are identified using the isomorphism $\mathfrak{E}$. The subset of $\mathcal{T}_e^{(\infty)}(\mathbf{G})$ formed by the differential operators of order ≤ 1 that vanish on constants is a subspace $\mathfrak{g}$ of the $\mathbf{K}$-vector space $\mathcal{T}_e^{(\infty)}(\mathbf{G})$. Furthermore, $\mathfrak{g}$ is a Lie subalgebra of $\mathcal{T}_e^{(\infty)}(\mathbf{G})$, since, if $T, S \in \mathfrak{g}$, the mappings $f \longmapsto T \star f$ and $f \longmapsto S \star f$ from $\mathcal{E}(\mathbf{G})$ into $\mathcal{E}(\mathbf{G})$ are derivations of $\mathcal{E}(\mathbf{G})$, so $[T, S]$ is a derivation of $\mathcal{E}(\mathbf{G})$ (section 5.4.1(**III**)). If $\mathbf{u} \in T_e(\mathbf{G})$, then $X_{\mathbf{u}}(e) \in \mathfrak{g}$. Conversely, if $S \in \mathfrak{g}$, then $X_S : f \mapsto S \star f$ $(f \in \mathcal{E}(\mathbf{G}))$ is a derivation of $\mathcal{E}(\mathbf{G})$ that can be identified with an element of $T_e(\mathbf{G})$ by Theorem 5.15 whenever $\mathbf{G}$ is finite-dimensional. This gives the following result:

THEOREM 6.64.– *There exists a monomorphism of Lie algebras Lie* $(\mathbf{G}) \hookrightarrow \mathfrak{g}$ *that is an isomorphism whenever* $\mathbf{G}$ *is finite-dimensional.*

Lie $(\mathbf{G})$ can therefore be identified with a Lie subalgebra of $\mathfrak{g}$ and with $\mathfrak{g}$ itself if $\mathbf{G}$ is finite-dimensional. For a Banach Lie group, it is possible to show the following result ([BOU 82b], Chapter 3, section 3.7, Proposition 25):

THEOREM 6.65.– *The Lie algebra* $\mathcal{T}_e^{(\infty)}(\mathbf{G})$ *is isomorphic to (and can be identified with) the enveloping algebra* $\boldsymbol{U}$ *(Lie* $(\mathbf{G})$*) of Lie* $(\mathbf{G})$ *(section 6.3.3(**IV**)).*

(III) ADJOINT REPRESENTATION AND ADJOINT LINEAR MAPPING Consider the adjoint representation Ad : $\mathbf{G} \to \mathrm{GL}(\mathfrak{g})$ of $\mathbf{G}$ (section 2.4.1, Definition 2.88) and the adjoint linear mapping ad : $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ (section 5.4.1, Definition 5.12).

LEMMA 6.66.– *We have* $ad = T_e(Ad)$ *(**exercise**).*

COROLLARY 6.67.– *Let* $\mathbf{G}$ *be a Lie group,* $\mathfrak{g} = Lie(\mathbf{G})$ *and* $s \in \mathbf{G}$*. Then,* $T_s(\mathbf{G}) = s.\mathfrak{g}$ *and* $T(\mathbf{G}) = \mathfrak{g} \ltimes \mathbf{G}$ *(semi-direct product of* $\mathbf{G}$ *and* $\mathfrak{g}$*), so* $T(\mathbf{G})$ *is a Lie group: the* tangent Lie group.

PROOF.– This follows from Lemma 2.89 and Definition 2.79. ■

REMARK 6.68.– *The tangent bundle* $(T(\mathbf{G}), \mathbf{G}, \pi)$ *of a Lie group* $\mathbf{G}$*, where* π *is the canonical surjection* $s.A \mapsto s$ $(s \in \mathbf{G}, A \in \mathfrak{g})$*, is a principal bundle with structural group* $\mathbf{G}$ *and base* $\mathbf{G}$*.*

(IV) FUNDAMENTAL VECTOR FIELDS ON A PRINCIPAL BUNDLE Let P be a principal bundle with structural group $\mathbf{G}$ and base $B \cong \mathbf{G}\backslash P$. For every point $q \in P$, let $q\bullet : \mathbf{G} \to P : g \mapsto q.g$ (section 3.5.1). The tangent linear mapping at the point e (neutral element of $\mathbf{G}$) is therefore $T_e(q\bullet) : T_e(\mathbf{G}) \to T_q(P)$. But $T_e(\mathbf{G}) = \mathrm{Lie}(\mathbf{G}) = \mathfrak{g}$; thus, $T_e(q\bullet)$ sends each element X of $\mathfrak{g}$ to a tangent vector $T_e(q\bullet)(X) \in T_q(P)$, denoted $Z_X(q) = q.X$. Each vector $X \in \mathfrak{g}$ is therefore associated with a vector field $Z_X : q \mapsto q.X$ on P.

DEFINITION 6.69.– $Z_X \in \mathcal{T}_0^1(P)$ $(Z_X : q \mapsto q.X)$ *is the* fundamental vector field *(or* Killing field*) associated with* $X \in \mathfrak{g}$*.*

The vector $Z_X(q) = q.X$ is tangent to the fiber $\mathbf{G}_{\pi(q)}$ and hence is vertical. More precisely:

LEMMA 6.70.– *The set of fundamental vectors* $Z_X(q)$ $(X \in \mathfrak{g})$ *on a principal bundle coincides with the set* $V_q(P)$ *of vertical vectors at the point* $q \in P$*.*

(V) INDUCED MORPHISMS OF LIE ALGEBRAS Let $\mathbf{G}, \mathbf{G}'$ be two Lie groups with neutral elements e, e' and left translations $\boldsymbol{\lambda}, \boldsymbol{\lambda}'$, respectively, and let $\varphi : \mathbf{G} \to \mathbf{G}'$ be a morphism of Lie groups. Write $\varphi_{*e} : T_e(\mathbf{G}) \to T_{e'}(\mathbf{G}')$ for the tangent linear mapping $T_e(\varphi)$ of φ at the point e (Definition 2.33). We will show that φ_{*e} is a morphism of Lie algebras.

Let $\mathbf{u} \in T_e(\mathbf{G})$ and $\mathbf{u}' = \varphi_{*e}(\mathbf{u})$. By the above, $\mathbf{u}$ and $\mathbf{u}'$ uniquely determine the left invariant vector fields $X_{\mathbf{u}} : s \mapsto s.\mathbf{u}$ and $X'_{\mathbf{u}'} : s' \mapsto s'.\mathbf{u}'$ in $\mathcal{T}_0^1(\mathbf{G})$ and $\mathcal{T}_0^1(\mathbf{G}')$, respectively.

LEMMA 6.71.– *For every $s \in \mathbf{G}$, we have $\varphi_{*s}.X_{\mathbf{u}}(s) = X'_{\mathbf{u}'}(\varphi(s))$.*

PROOF.– For every $s, t \in \mathbf{G}$, $\varphi(s.t) = \varphi(s).\varphi(t)$, so $\varphi \circ \boldsymbol{\lambda}(s) = \boldsymbol{\lambda}'(\varphi(s)) \circ \varphi$. Furthermore, by definition, $X_{\mathbf{u}}(s) = \boldsymbol{\lambda}(s)_{*e}.\mathbf{u}$ and $X'_{\mathbf{u}'}(s') = \boldsymbol{\lambda}(s')_{*e'}.\mathbf{u}'$. Hence, $\varphi_{*s}.X_{\mathbf{u}}(s) = \varphi_{*s} \circ \boldsymbol{\lambda}(s)_{*e}.\mathbf{u} = (\varphi \circ \boldsymbol{\lambda}(s))_{*e}.\mathbf{u} = \left(\boldsymbol{\lambda}'(\varphi(s)) \circ \varphi\right)_{*e}.\mathbf{u} = X'_{\mathbf{u}'}(\varphi(s))$. ■

THEOREM 6.72.– *The mapping φ_{*e} is a morphism of Lie algebras.*

PROOF.– This follows from Lemma 6.71 and Remark 5.18. ■

COROLLARY 6.73.– (Lie algebra of a Lie subgroup) *i) Let $\mathbf{G}, \mathbf{H}$ be Lie groups such that $\mathbf{H} \subset \mathbf{G}$. Then, Lie $(\mathbf{H})$ can be identified with a subalgebra of Lie $(\mathbf{G})$.*

ii) If $\mathbf{H}_1, \mathbf{H}_2$ are Lie subgroups of a finite-dimensional Lie group $\mathbf{G}$, then so is $\mathbf{H}_1 \cap \mathbf{H}_2$, and Lie $(\mathbf{H}_1 \cap \mathbf{H}_2) =$ Lie $(\mathbf{H}_1) \cap$ Lie $(\mathbf{H}_2)$.

PROOF.– (i): Let $\iota : \mathbf{H} \hookrightarrow \mathbf{G}$ be the canonical injection. By Theorem 6.72, $\iota_{*e}(\text{Lie}(\mathbf{H}))$ is a subalgebra of Lie $(\mathbf{G})$. Furthermore, $\mathbf{H}$ is a submanifold of $\mathbf{G}$, so $T_e(\mathbf{H})$ can be identified with the subspace $\iota_{*e}(T_e(\mathbf{H}))$ of $T_e(\mathbf{G})$ (Section 2.3.3). (ii): See ([BOU 82b], Chapter 3, Section 3.8, Corollary 2). ■

Note that the hypotheses of Corollary 6.73(ii) are very different from those of Theorem 2.64, even though they lead to an analogous conclusion.

THEOREM 6.74.– (Lie) *If two Lie groups are locally isomorphic, their Lie algebras are isomorphic.*

PROOF.– This follows from Lemma 2.83 and Theorem 6.72. ■

(VI) LIE ALGEBRAS OF CLASSICAL GROUPS We will show that the Lie algebras of the classical groups (in the sense specified in section 2.4.1**(VI)**) coincide with the classical Lie algebras.

Let $\mathbf{E}$ be a Banach space. Then, $1_{\mathbf{E}}$ is the neutral element of GL $(\mathbf{E})$. Since GL $(\mathbf{E})$ is open in the Banach algebra $\mathcal{L}(\mathbf{E})$ ([P2], section 3.4.1**(II)**, Corollary 3.49), for every $A \in \mathcal{L}(\mathbf{E})$, there exists $\varepsilon > 0$ such that $1_{\mathbf{E}} + \varepsilon A \in \text{GL}(\mathbf{E})$, and hence $T_e(\text{GL}(\mathbf{E})) = \mathcal{L}(\mathbf{E})$. By Definition 6.63 and section 5.4.1**(II)**, it follows that:

THEOREM 6.75.– *The Lie algebra of* $GL(\mathbf{E})$ *is* $\mathfrak{gl}(\mathbf{E})$.

If $\mathbf{E} = \mathbb{K}^n$, then $\mathrm{Lie}(\mathrm{GL}_n(\mathbb{K})) = \mathfrak{gl}_n(\mathbb{K})$. By [P1], section 2.3.11(**VII**):

$$\det(1_{\mathbf{E}} + A) = 1 + \mathrm{Tr}(A) + o(\|A\|). \qquad [6.5]$$

The group $\mathrm{SL}_n(\mathbb{K})$ is the subgroup of $\mathrm{GL}_n(\mathbb{K})$ formed by the matrices X such that $\det(X) = 1$. Therefore, its tangent space at $1_{\mathbf{E}}$ is the set of matrices $A \in \mathfrak{M}_n(\mathbf{K})$ with zero trace, so $\mathrm{Lie}(\mathrm{SL}_n(\mathbb{K})) = \mathfrak{sl}_n(\mathbb{K})$.

The group $\mathrm{O}_n(\mathbb{K})$ is formed by the matrices $X \in \mathfrak{M}_n(\mathbf{K})$ such that $X.X^T = 1_n$. Write X in the form $I_n + \varepsilon.A$. Then:

$$\frac{(I_n + \varepsilon.A).(I_n + \varepsilon.A)^T - I_n}{\varepsilon} = A + A^T + O(|\varepsilon|)$$

so $\mathrm{Lie}(\mathrm{O}_n(\mathbb{K})) = \mathfrak{o}_n(\mathbb{K})$. Similarly, $\mathrm{Lie}(\mathrm{Sp}_{2n}(\mathbb{K})) = \mathfrak{sp}_{2n}(\mathbb{K})$, $\mathrm{Lie}(\mathrm{A}_n) = \mathfrak{a}_n$ and $\mathrm{Lie}(\mathrm{E}_n) = \mathrm{Lie}(\mathrm{SE}_n) = \mathfrak{se}_n$.

We have $\mathrm{SO}_n(\mathbb{K}) = \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K})$, so by Corollary 6.73(2), setting $\mathfrak{so}_n(\mathbb{K}) = \mathrm{Lie}(\mathrm{SO}_n(\mathbb{K}))$, we have $\mathfrak{so}_n(\mathbb{K}) = \mathfrak{sl}_n(\mathbb{K}) \cap \mathfrak{o}_n(\mathbb{K}) = \mathfrak{o}_n(\mathbb{K})$. Similarly, $\mathrm{Lie}(\mathrm{U}(\mathbb{C})) = \mathfrak{u}_n(\mathbb{C})$ and $\mathrm{SU}_n(\mathbb{C}) = \mathrm{U}_n(\mathbb{C}) \cap \mathrm{SL}_n(\mathbb{C})$, so $\mathrm{Lie}(\mathrm{SU}_n(\mathbb{C})) = \mathfrak{u}_n(\mathbb{C}) \cap \mathfrak{sl}_n(\mathbb{C}) = \mathfrak{su}_n(\mathbb{C})$.

The Lie algebras of $\mathrm{D}_n(\mathbb{K})$, $\mathrm{T}_n(\mathbb{K})$, $\mathrm{ST}_n(\mathbb{K})$ and $\mathrm{N}_n(\mathbb{K})$ are $\mathfrak{d}_n(\mathbb{K})$, $\mathfrak{t}_n(\mathbb{K})$, $\mathfrak{st}_n(\mathbb{K})$ and $\mathfrak{n}_n(\mathbb{K})$, respectively (**exercise**).

Since $\mathrm{Spin}_n(\mathbb{K})$ and $\mathrm{SO}_n(\mathbb{K})$ are locally isomorphic (section 3.3.3(**IV**)), $\mathrm{Lie}(\mathrm{Spin}_n(\mathbb{K})) = \mathfrak{o}_n(\mathbb{K})$. Similarly, $\mathrm{PGL}_n(\mathbb{K})$, $\mathrm{PSL}_n(\mathbb{K})$ and $\mathrm{SL}_n(\mathbb{K})$ (respectively $\mathrm{PO}_n(\mathbb{K})$, $\mathrm{PSO}_n(\mathbb{K})$ and $\mathrm{O}_n(\mathbb{K})$, respectively $\mathrm{PU}_n(\mathbb{C})$ and $\mathrm{SU}_n(\mathbb{C})$) are locally isomorphic (section 2.4.1(**VI**)), so $\mathrm{Lie}(\mathrm{PGL}_n(\mathbb{K})) = \mathrm{Lie}(\mathrm{PSL}_n(\mathbb{K})) = \mathfrak{sl}_n(\mathbb{K})$ (respectively $\mathrm{Lie}(\mathrm{PO}_n(\mathbb{K})) = \mathrm{Lie}(\mathrm{PSO}_n(\mathbb{K})) = \mathfrak{o}_n(\mathbb{K})$, respectively $\mathrm{Lie}(\mathrm{PU}_n) = \mathfrak{su}_n(\mathbb{C})$). We have therefore established the following result:

THEOREM 6.76.– (Lie–Killing–Cartan classification) *The simple complex projective groups*[6] *(respectively the simply connected classical Lie groups*[7]*) are as follows:*

PSL$_n(\mathbb{C})$ *(respectively SU*$_n(\mathbb{C})$*), of type* A_{n-1} *(*$n \geq 2$*);*

PSO$_{2n+1}(\mathbb{C})$ *(respectively Spin*$_{2n+1}(\mathbb{R})$*), of type* B_n *(*$n \geq 1$*;* $n \geq 2$[8]*);*

6 Lie's classification.

7 Killing–Cartan classification

8 Here and below, the first value of n is the minimal index. The second value should be considered if we wish to eliminate any isomorphisms between Lie algebras of different types: see section 6.3.2(**II**).

$PSp_{2n}(\mathbb{C})$ (respectively USp_{2n}), of type C_n ($n \geq 1$; $n \geq 3$);

$PSO_{2n}(\mathbb{C})$ (respectively $Spin_{2n}(\mathbb{R})$), of type D_n ($n \geq 3$; $n \geq 4$).

6.4.2. *Passing from a Lie algebra to a Lie group*

(I) CAMPBELL–BAKER–HAUSDORFF FORMULA Let $\mathbf{E}$ be a Banach space, $f : z \mapsto f(z) = \sum_{k\geq 0} a_k.z^k$ a power series with radius of convergence $\rho > 0$ and $A \in \mathcal{L}(\mathbf{E})$. Then, $f(A) = \sum_{k\geq 0} a_k.A^k$ converges in $\mathcal{L}(\mathbf{E})$ for every $A \in \mathcal{L}(\mathbf{E})$ such that $\|A\| < \rho$ (section 1.2.5(**I**)). We know that, if $|z-1| < 1$, then $\ln(z) = \ln(1-(1-z)) = \sum_{k\geq 1} \frac{(1-z)^k}{k}$ and $e^{\ln(z)} = z$, which leads us to the following result, according to [P2], section 3.4.1(**II**), Theorem 3.46, and section 1.2.5(**II**):

LEMMA 6.77.– *Let $U = \{B \in \mathcal{L}(\mathbf{E}) : \|B - 1_{\mathbf{E}}\| < 1\}$. Then, U is a neighborhood of $1_{\mathbf{E}}$ in $GL(\mathbf{E})$ and the mapping $\ln : B \mapsto \ln(B)$ is an analytic diffeomorphism from U onto a neighborhood V of 0 in $\mathfrak{gl}(\mathbf{E})$ whose inverse diffeomorphism is the* exponential mapping $\exp : A \mapsto \exp(A)$.

LEMMA 6.78.– *Let $A : \mathbb{R} \to \mathcal{L}(\mathbf{E})$ be a mapping of class C^1. Then, for every $t \in \mathbb{R}$, $e^{A(t)}.\frac{d}{dt}\left(e^{-A(t)}\right) = -f(adA(t)).\dot{A}$, where $f(z) = (e^z - 1)/z$.*

PROOF.– Let $B(s,t) = e^{s.A(t)}.\frac{d}{dt}\left(e^{-s.A(t)}\right)$. By section 5.4.1(**III**),

$$\frac{\partial B}{\partial s}(s,t) = \mathrm{ad}A(t)(B(s,t)) - \dot{A}(t), \quad B(0,t) = 0.$$

This differential equation is of the form $\dot{u}(s) = z.u(s) - w$ with $z = \mathrm{ad}A(t)$ and $w = -\dot{A}(t)$, and $u(0) = 0$. It follows that:

$$u(s) = \frac{1}{z}(e^{s.z} - 1).w \Longrightarrow u(1) = f(z).w. \qquad \blacksquare$$

THEOREM 6.79.– (Campbell–Baker–Hausdorff) *There exists a neighborhood V of 0 in $\mathcal{L}(\mathbf{E})$ such that, if $A, B \in V$, then $C = \ln\left(e^A.e^B\right)$ is the element of V uniquely determined by the Campbell–Baker–Hausdorff formula:*

$$\boxed{C = B + \int_0^t g\left(e^{t.adA}.e^{adB}\right).A.dt, \quad g(z) = \frac{\ln(z)}{z-1}.} \qquad [6.6]$$

PROOF.– Let $C(t) = \ln\left(e^{t.A}.e^B\right)$. Then:

$$\exp(C(t)).\frac{d}{dt}(\exp(-C(t))) = -A,$$

since A and $e^{t.A}$ commute, and, by Lemma 6.78,

$$A = f\left(\mathrm{ad}C\left(t\right)\right).\dot{C}\left(t\right). \qquad [6.7]$$

It can be checked that $\exp\left(\mathrm{ad}C\left(t\right)\right) = \exp\left(t.\mathrm{ad}A\right).\exp\left(\mathrm{ad}B\right)$ by differentiating both terms of this equality. Therefore, there exists a neighborhood V of 0 in $\mathcal{L}\left(\mathbf{E}\right)$ such that the right-hand side of this expression is defined for $A, B \in V$ and $\mathrm{ad}C\left(t\right) = \ln\left(\exp\left(t.\mathrm{ad}A\right).\exp\left(\mathrm{ad}B\right)\right)$. Since $f\left(\ln\left(z\right)\right).g\left(z\right) = 1$, we deduce that $f\left(\mathrm{ad}C\left(t\right)\right) = g\left(\exp\left(t.\mathrm{ad}A\right).\exp\left(\mathrm{ad}B\right)\right)^{-1}$ and, by [6.7], $\dot{C}\left(t\right) = g\left(\exp\left(t.\mathrm{ad}A\right).\exp\left(\mathrm{ad}B\right)\right).A$ with $C\left(0\right) = B$. Integrating then shows [6.6]. ■

By developing the power series of the analytic function g, we obtain the Hausdorff series expansion (**exercise**):

$$\ln\left(e^{A}.e^{B}\right) = A + B + \frac{1}{2}\left[A, B\right] + \frac{1}{12}\left[A, \left[A, B\right]\right] + \frac{1}{12}\left[B, \left[B, A\right]\right] + ...$$

The higher "degree" terms quickly become difficult to calculate (see [GOD 17], section 6.6, (6.12)).

COROLLARY 6.80.– *If $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}\left(\mathbf{E}\right)$ and $A, B \in \mathfrak{g}$, then $\ln\left(e^{A}.e^{B}\right) \in \mathfrak{g}$.*

PROOF.– The vector subspace $\mathfrak{g}$ of $\mathcal{L}\left(\mathbf{E}\right)$ is stable under the operators $\mathrm{ad}A$ and $\mathrm{ad}B$ (section 5.4.1(**III**)), and hence also under the operator $e^{t.\mathrm{ad}A}.e^{\mathrm{ad}B}$, and so also under $g\left(e^{t.\mathrm{ad}A}.e^{\mathrm{ad}B}\right)$. Thus, on the right-hand side of the first equality of [6.6], we are integrating in $\mathfrak{g}$. ■

(II) FUNDAMENTAL THEOREMS

LEMMA 6.81.– *Let $\mathfrak{g}$ and $\mathfrak{g}'$ be the Lie algebras of the Lie groups $\mathbf{G}$ and $\mathbf{G}'$, respectively, assumed to be finite-dimensional over $\mathbb{K}$, and let $\sigma : \mathfrak{g} \to \mathfrak{g}'$ be a morphism of Lie algebras. There exists a local morphism of Lie groups φ from $\mathbf{G}$ to $\mathbf{G}'$ (Definition 2.81) such that (with the notation of section 6.4.1(**III**)) $\varphi_{*e} = \sigma$.*

PROOF.– By Ado's theorem (Theorem 6.26), there exists an integer n such that $\mathfrak{g}, \mathfrak{g}'$ can be embedded in $\mathfrak{M}_n\left(\mathbb{K}\right)$. We will show that $\varphi\left(e^{A}\right) = e^{\sigma(A)}$ in an open neighborhood U of 0 in $\mathfrak{M}_n\left(\mathbb{K}\right)$. To do this, we must show that, if $A, B \in U$ and $C = A.B$, then $\varphi\left(e^{A}.e^{B}\right) = \varphi\left(e^{A}\right).\varphi\left(e^{B}\right)$. But $\varphi\left(e^{A}.e^{B}\right) = \sigma\left(C\right)$, where $C = \ln\left(e^{A}.e^{B}\right)$. Since σ preserves brackets, $\sigma\left(C\right) = \sigma\left(A\right).\sigma\left(B\right)$ by the same reasoning as the proof of Corollary 6.80. ■

THEOREM 6.82.– (Lie) *Let $\mathfrak{g}$ and $\mathfrak{g}'$ be the Lie algebras of the Lie groups $\mathbf{G}$ and $\mathbf{G}'$, respectively, assumed to be finite-dimensional over $\mathbb{K}$. If $\mathfrak{g}$ and $\mathfrak{g}'$ are isomorphic, then $\mathbf{G}$ and $\mathbf{G}'$ are locally isomorphic. If $\mathbf{G}$ and $\mathbf{G}'$ are also simply connected, then they are isomorphic.*

PROOF.– This follows from Lemma 6.81 and Theorem 2.85. ■

COROLLARY 6.83.– *Let* $\mathbf{G}$ *be a finite-dimensional Lie group. There exists an open neighborhood* U *of* 0 *in Lie* $(\mathbf{G})$ *such that the* exponential mapping $\exp_{\mathbf{G}} : X \mapsto e^X$ *of* $\mathbf{G}$ *is a diffeomorphism from* U *onto an open neighborhood of* e *in* $\mathbf{G}$.

PROOF.– This follows from Theorem 2.61(2) and Lemma 6.81. ■

THEOREM 6.84.– (converse of Lie's third fundamental theorem) *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra over* $\mathbb{K}$. *There exists a local Lie group (Definition 2.81) whose Lie algebra is isomorphic to* $\mathfrak{g}$.

PROOF.– Again, this follows from Ado's theorem. ■

The connection to the global setting was established by É. Cartan in 1930 ([CAR 52], Chapter 2, section I.22; [BOU 82b], Chapter 3, section 6.3, Theorem 3):

THEOREM 6.85.– (Cartan) *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra over* $\mathbb{K}$. *There exists a simply connected Lie group whose Lie algebra is isomorphic to* $\mathfrak{g}$.

REMARK 6.86.– *1) The exponential mapping is not surjective in general for a connected Lie group of matrices. In other words, if* $\mathfrak{g}$ *is a subalgebra of* $\mathfrak{gl}_n(\mathbb{K})$ *and* $\mathfrak{g} = Lie(\mathbf{G})$, *where* $\mathbf{G} \subset GL_n(\mathbb{K})$ *is a connected Lie group, it is not true that every element* $g \in \mathbf{G}$ *must be of the form* $\exp(A)$ *with* $A \in \mathfrak{g}$. *For example, diag* $(\lambda, 1/\lambda) \in SL_2(\mathbb{R})$ *is not of the form* $\exp(A)$ *with* $A \in \mathfrak{sl}_2(\mathbb{R})$ *whenever* $\lambda < 0$ *and* $\lambda \neq -1$ *(**exercise**).*

2) Similarly, the exponential mapping is not injective in general for a connected Lie group of matrices: if $A = \begin{bmatrix} 0 & 2\pi \\ 2\pi & 0 \end{bmatrix} \in \mathfrak{sl}_2(\mathbb{R})$, *then* $\exp(A) = I_2 = e^0$.

3) Every finite-dimensional Lie algebra $\mathfrak{g}$ *can be identified with a subalgebra of* $\mathfrak{gl}_n(\mathbb{K})$ *by Ado's theorem (Theorem 6.26) and is the Lie algebra of a Lie group* $\mathbf{G}$ *by Cartan's theorem (Theorem 6.85). By the proof of Lemma 6.81, there exist a neighborhood* U *of* 0 *in* $\mathfrak{g}$, *a Lie subgroup* $\mathbf{G}'$ *of* $GL_n(\mathbb{K})$ *and a neighborhood* V *of* e' *in* $\mathbf{G}'$ *(where* e' *is the neutral element of* $\mathbf{G}'$*) such that* $\exp$ *is an analytic diffeomorphism from* U *onto* V. *Hence,* $\mathbf{G}$ *is* locally *isomorphic to the Lie group of matrices* $\mathbf{G}'$ *by Theorem 6.82; however, G. Birkhoff showed in 1936 [BIR 36] that* $\mathbf{G}$ *is not necessarily (globally) isomorphic to a Lie group of matrices.*

Nevertheless, we have the following results:

LEMMA 6.87.– (É. Cartan) *If* $\mathbf{G} \subset GL_n(\mathbb{K})$ *is* compact *and connected, then* $\exp_{\mathbf{G}}$ *is* surjective *from Lie* $(\mathbf{G})$ *onto* $\mathbf{G}$.

PROOF.– See [GOD 17], section 3.11, Theorem 9. ■

The exponential mapping is also surjective from $\mathfrak{gl}_n(\mathbb{C})$ onto $\mathrm{GL}_n(\mathbb{C})$ (*ibid.*, Theorem 10).

LEMMA 6.88.– *If* $\mathbf{G}$ *is connected and* nilpotent *([P1], section 2.2.7(**II**)), then* $\exp_{\mathbf{G}}$ *is* étale *(Lemma-Definition 2.57) and surjective; furthermore, if* $\mathbf{G}$ *is* simply connected*, then* $\exp_{\mathbf{G}}$ *is* bijective*.*

PROOF.– See [BOU 82b], Chapter 3, section 9.5, Propositions 13 and 14. ■

The statement of Lemma 6.88 fails if $\mathbf{G}$ is assumed to be solvable instead of nilpotent ([HOC 65], Chapter 12, Exercise 1).

Let $\mathbf{G}$ be a finite-dimensional Lie group and $\mathfrak{h}$ a subalgebra of $\mathrm{Lie}(\mathbf{G})$. In general, there does not necessarily exist a Lie subgroup of $\mathbf{G}$ such that $\mathrm{Lie}(\mathbf{H}) = \mathfrak{h}$. However, it can be shown ([DIE 93], Volume 4, (19.7.4)) that there exist a subgroup $\mathbf{H}$ of $\mathbf{G}$ and a *connected* Lie group structure on $\mathbf{H}$ such that the monomorphism $\iota : \mathbf{H} \hookrightarrow \mathbf{G}$ is an immersion and $\iota_{*e} : \mathrm{Lie}(\mathbf{H}) \to \mathfrak{h}$ (where e is the neutral element) is an isomorphism; furthermore, $\mathbf{H}$ and ι are determined up to isomorphism.

DEFINITION 6.89.– $\mathbf{H}$ *is said to be the* integral subgroup of $\mathbf{G}$ *corresponding to the Lie algebra* $\mathfrak{h} \subset$ *Lie* $(\mathbf{G})$.

The integral subgroup $\mathbf{H}$ of $\mathbf{G}$ is not necessarily closed in $\mathbf{G}$, even if $\mathbf{G}$ is simply connected ([DIE 93], Volume 4, Chapter 19, section 7, Exercise 2), in which case $\mathbf{H}$ is not a Lie subgroup of $\mathbf{G}$ (Theorem 2.76(1)).

If $\rho : \mathbf{G} \to \mathrm{GL}(\mathbf{F})$ is a linear representation of $\mathbf{G}$ in a Banach space $\mathbf{F}$ (Definition 2.87), then $\rho_{*e} : \mathrm{Lie}(\mathbf{G}) \to \mathfrak{gl}(\mathbf{F})$ is a representation of $\mathrm{Lie}(\mathbf{G})$ (section 5.4.1(**II**)).

6.4.3. *Dictionary*

In the following, every Lie group and every Lie algebra is finite-dimensional, unless otherwise stated.

(I) THE LIE FUNCTOR FOR CONNECTED GROUPS The functor $\mathrm{Lie} : \mathbf{LieGrp} \to \mathbf{LieAl}$ is defined as follows: $\mathrm{Lie}(\mathbf{G})$ is the Lie algebra of the Lie group $\mathbf{G}$; if $\varphi : \mathbf{G} \to \mathbf{G}'$ is a Lie group morphism, then $\mathrm{Lie}(\varphi) = \varphi_{*e} : \mathrm{Lie}(\mathbf{G}) \to \mathrm{Lie}(\mathbf{G}')$.

If $\mathbf{G}$ is a Lie group, its neutral component is a normal Lie subgroup (Corollary 2.77). Furthermore, if $\mathbf{G}, \mathbf{H}$ are Lie groups, $f : \mathbf{G} \to \mathbf{H}$ is a morphism of $\mathbf{LieGrp}$ and $\mathbf{G}$ is connected, then $f(\mathbf{G})$ is a connected subgroup of $\mathbf{H}$ ([P2], section 2.3.8, Theorem 2.46), and hence a Lie subgroup (Theorem 2.76(2)). Write $\mathbf{cLieGrp}$ for the full subcategory of $\mathbf{LieGrp}$ whose objects are the connected Lie groups.

Consider the diagram:

$$\begin{array}{ccccccccc} \{1\} & \longrightarrow & \mathbf{H} & \xrightarrow{\iota} & \mathbf{G} & \xrightarrow{\varphi} & \mathbf{G}/\mathbf{H} & \longrightarrow & \{1\} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \{0\} & \longrightarrow & \mathfrak{h} & \xrightarrow{\iota_{*e}} & \mathfrak{g} & \xrightarrow{\varphi_{*e}} & \mathfrak{g}/\mathfrak{h} & \longrightarrow & \{0\} \end{array} \qquad [6.8]$$

where e denotes the neutral element in each case, the top sequence is exact in $\mathbf{LieGrp}$, the vertical arrows denote the Lie functor and the bottom sequence is in the category $\mathbf{LieAl}$. If $\mathbf{G}$ is connected, then $\mathbf{G}/\mathbf{H} = \varphi(\mathbf{G})$ is connected; furthermore, if $\mathbf{H}$ is connected, we say that the top sequence is exact in $\mathbf{cLieGrp}$. The Lie functor determines a subfunctor from $\mathbf{cLieGrp}$ into $\mathbf{LieAl}$ ([P1], section 1.2.1(**V**)) that we will also write Lie.

THEOREM 6.90.– *The functor Lie is faithful and exact from* $\mathbf{cLieGrp}$ *into* $\mathbf{LieAl}$ *and, for this functor, the diagram [6.8] commutes.*

PROOF.– (1) If $\mathbf{G}$ is a connected Lie group and $\mathbf{K}$ is a Lie group, the mapping $f \mapsto f_{*e}$ from the set of morphisms from $\mathbf{G}$ into $\mathbf{K}$ into the set of morphisms from Lie $(\mathbf{G})$ into Lie $(\mathbf{K})$ is injective ([DIE 93], Volume 4, (19.7.6)), so the functor Lie : $\mathbf{cLieGrp} \to \mathbf{LieAl}$ is faithful ([P1], section 1.2.1(**III**)). Furthermore, Lie $(f(\mathbf{G})) = f_{*e}$ (Lie $(\mathbf{K})$) ([DIE 93], Volume 4, (19.7.5)), and the Lie subgroup $\mathbf{H}$ of $\mathbf{G}$ is normal in $\mathbf{G}$ if and only if Lie $(\mathbf{H})$ is an ideal of Lie $(\mathbf{G})$ ([BOU 82b], Chapter 3, section 6.6, Proposition 14), so, if the top sequence of the diagram [6.8] is exact in $\mathbf{cLieGrp}$, then the bottom sequence is exact in $\mathbf{LieAl}$. Hence, the Lie functor is exact ([P1], section 1.2.9(**I**)). ■

It is also possible to show the following result ([VAR 84], section 3.15): let $\mathbf{G}$ be the semi-direct product $\mathbf{H} \times_{\tau} \mathbf{K}$, where $\mathbf{H}$ is a normal subgroup of $\mathbf{G}$ and $\mathbf{K}$ is a Lie subgroup of $\mathbf{H}$ (section 2.4.1(**III**)). Then, Lie $(\mathbf{G}) \cong$ Lie $(\mathbf{H}) \oplus_{\sigma}$ Lie $(\mathbf{K})$, where $\sigma = \tau_{*e}$ (section 5.4.1(**IV**)). Furthermore, ([BOU 82b], Chapter 3, section 9.2, Corollary of Proposition 4):

THEOREM 6.91.– *Let* $\mathbf{G}$ *be a connected Lie group and* $\mathfrak{g}$ *its Lie algebra. The derived subgroups* $\mathfrak{D}^i\mathbf{G}$ *(respectively the central subgroups* $\mathfrak{C}^i\mathbf{G}$*) ([P1], section 2.2.7) are the integral subgroups* $\mathbf{G}$ *corresponding to the characteristic ideals* $\mathcal{D}^i\mathfrak{g}$ *(respectively* $\mathcal{C}^i\mathfrak{g}$*) (Definition 6.22); they are Lie subgroups if* $\mathbf{G}$ *is simply connected. Hence,* $\mathbf{G}$ *is solvable (respectively nilpotent) if and only if its Lie algebra is solvable (respectively nilpotent).*

(II) THE LIE FUNCTOR FOR SIMPLY CONNECTED GROUPS In section 3.3.3(**IV**), we saw that every connected Lie group $\mathbf{G}$ is isomorphic to the quotient $\tilde{\mathbf{G}}/\mathbf{D}$ of a simply connected group $\tilde{\mathbf{G}}$, the universal covering of $\mathbf{G}$, by a discrete subgroup $\mathbf{D} \cong \pi_1(\mathbf{G})$ contained in the center $\mathfrak{Z}(\mathbf{G})$ of $\mathbf{G}$.

LEMMA 6.92.– *Let* $\mathbf{G}$ *be a simply connected Lie group and* $\mathbf{H}$ *a Lie group. The mapping* $(.)_{*e}$: *Hom* $(\mathbf{G},\mathbf{H}) \to$ *Hom* $(\mathfrak{g},\mathfrak{h})$ *is bijective, where* $\mathfrak{g} =$ *Lie* $(\mathbf{G})$ *and* $\mathfrak{h} =$ *Lie* $(\mathbf{G})$.

PROOF.– Since $\mathbf{G}$ is connected, it is generated by every neighborhood of the neutral element (Lemma 2.80(i)), and hence by $\exp(\mathfrak{g})$ by Corollary 6.83. In other words, any element g of $\mathbf{G}$ can be written in the form $\exp(A_1)\ldots\exp(A_n)$. We need to show that, if $f \in \mathrm{Hom}(\mathfrak{g},\mathfrak{h})$, then there exists $\varphi \in \mathrm{Hom}(\mathbf{G},\mathbf{H})$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathbf{G} & \xrightarrow{\varphi} & \mathbf{H} \\ \exp_{\mathbf{G}} \uparrow & & \uparrow \exp_{\mathbf{H}} \\ \mathfrak{g} & \xrightarrow{f} & \mathfrak{h} \end{array}$$

i.e. such that $\varphi(\exp_{\mathbf{G}}(A)) = \exp_{\mathbf{H}}(f(A))$ for every $A \in \mathfrak{g}$, or alternatively $\varphi(g) = \varphi(\exp_{\mathbf{G}}(A_1))\ldots\varphi(\exp_{\mathbf{G}}(A_n))$. The graph $\mathfrak{l} = \mathrm{Gr}(f)$ is a Lie subalgebra of $\mathfrak{g}\times\mathfrak{h}$; let $\mathbf{L}$ be the integral subgroup of $\mathbf{G}\times\mathbf{H}$ corresponding to the Lie algebra $\mathfrak{l}$ (Definition 6.89). The projection $\psi : \mathbf{L} \to \mathbf{G}$ is a local isomorphism (Theorem 6.82) and an epimorphism of Lie groups, so $\ker(\psi)$ is discrete, or alternatively ψ is a covering (section 3.3.3(**I**)); since $\mathbf{G}$ is simply connected, ψ is an isomorphism (Lemma 3.18) and $\mathbf{L}$ is the graph of the desired mapping φ. ■

Let **scLieGrp** be the full subcategory of **LieGrp** whose objects are the simply connected Lie groups.

THEOREM 6.93.– *The functor Lie* : **scLieGrp** $\to$ **LieAl** *is an equivalence of categories ([P1], section 1.2.2(**II**)).*

PROOF.– By Theorem 6.90 and Lemma 6.92, the functor Lie : **scLieGrp** $\to$ **LieAl** is fully faithful. It is also injective, and, by Theorem 6.85, essentially surjective. ■

(III) ALMOST SIMPLE, SEMI-SIMPLE, REDUCTIVE, COMMUTATIVE GROUPS A connected Lie group is *almost simple* (respectively *semi-simple*, respectively *reductive*) if its Lie algebra is simple (respectively semi-simple, respectively reductive) (sections 6.3.6(**I**) and 6.3.7). By **(I)**, **(II)** above and Theorem 6.40, we have the following result:

THEOREM 6.94.– *1) Let* $\mathbf{G}$ *be a* connected *Lie group. Then,* $\mathbf{G}$ *is commutative if and only if Lie* $(\mathbf{G})$ *is a commutative Lie algebra.*

Furthermore, the following conditions are equivalent:

i) $\mathbf{G}$ *is almost simple.*

ii) The only immersed Lie groups in $\mathbf{G}$ *are* $\{e\}$ *and* $\mathbf{G}$*, and* $\mathbf{G}$ *is not commutative.*

2) Let $\mathbf{G}$ *be a* simply connected *Lie group. The following conditions are equivalent:*

i) $\mathbf{G}$ *is semi-simple (respectively reductive).*

ii) $\mathbf{G}$ *is isomorphic to the direct product of finitely many almost simple Lie groups (respectively the direct product of a semi-simple Lie group and a commutative Lie group).*

PROOF.– If $\mathbf{G}$ is commutative, it is clear that $\text{Lie}(\mathbf{G})$ is commutative. Conversely, if $\text{Lie}(\mathbf{G})$ is commutative, then, by Theorem 6.82, $\mathbf{G}$ is locally isomorphic to the additive Lie group $\text{Lie}(\mathbf{G})$, since $\text{Lie}(\text{Lie}(\mathbf{G})) \cong \text{Lie}(\mathbf{G})$ (**exercise**). Hence, $\mathbf{G}$ is commutative by Lemma 2.80(i). The other claims follow from the results of **(I)**, **(II)**.■

Any connected Lie group that is a non-commutative simple group ([P1], section 2.2.4(**I**)) is almost simple. If $\mathbf{G}$ is an almost simple Lie group and $\mathbf{C}$ is its center, then $\mathbf{G}/\mathbf{C}$ is simple [TIT 62]. The Lie group $\text{GL}_n(\mathbb{C})$ is reductive.

6.5. Harmonic analysis

6.5.1. *Introduction*

The harmonic analysis of a time signal studies the "frequency components" of this signal. We can pass from a time-based representation to a frequency-based representation using Fourier transforms (Remarks 6.99(1) and 6.108). However, it would be extremely reductive to suggest that the applications of Fourier transforms are restricted to this specific area.

As early as 1755, D. Bernoulli gave a solution to the problem of vibrating strings, presented in terms of sums of trigonometric series we now call *Fourier series*. But it was Joseph Fourier who made leaps and bounds in harmonic analysis in 1804 by studying the propagation of heat in solids. A pragmatically minded man (the prefect of the department of Isère under Napoleon I), he was relatively uninterested in the convergence of the series which bear his name; this convergence was later studied by others including Dirichlet (1829), Riemann (1854), Dini (1880), Jordan (1881) and Fejèr (1900) drawing from work by Cesàro on series (1897)[9]. The fine convergence of Fourier series and Fourier integrals is highly complex and preoccupied numerous mathematicians for the best part of a century and a half (section 6.5.3(**VII**)). Two approaches can be adopted to simplify the problem considerably: Schwartz distribution theory (1947), which forms the basis of any modern presentation of the subject, and the Hilbertian formulation, due to Plancherel (1910), which studies mean square convergence (sections 6.5.2(**VII**) and 6.5.3(**VI**)). These two formulations are presented in sections 6.5.2 and 6.5.3, which are dedicated to "classical" harmonic analysis.

9 See Theorem 6.122, as well as the Wikipedia articles on the *Dini criterion* and *Fejèr's theorem*.

Harmonic analysis was developed on locally compact commutative groups by A. Weil in 1938 [WEI 38]. This generalization was made possible by the Pontryagin–van Kampen duality theorem (established by Pontryagin in 1934 for separable groups, then extended to arbitrary locally compact commutative groups by van Kampen the following year). The theorem revolves around the fact that, on a locally compact *commutative* group, every irreducible unitary component of the regular representation is of degree 1 (see Corollary 6.148). This makes "abstract" commutative harmonic analysis a relatively direct extension of classical harmonic analysis; nevertheless, it plays an invaluable role in certain unexpected places, such as number theory ([WEI 74], Chapter 7). A few details are given below to present this field (without a proof of the duality theorem).

Non-commutative harmonic analysis, in the words of J. Dieudonné, is a theory "hedged [...] with difficulties both conceptual and technical" ([DIE 93], Volume 6, Chapter 22, p. 2). Despite its complexity, this field has found applications not only in mathematics (number theory) and theoretical physics but also more recently in engineering (signal and image process, robotics, automation, solid physics, etc. [CHI 01]). We will also give a brief overview of this field in section 6.5.5. Frobenius himself, as well as Molien independently, observed in 1897 that finite non-commutative groups have irreducible unitary representations of degree higher than 1, where the Fourier coefficients are therefore unitary matrices; Frobenius' work was continued by Schur and Burnside between 1904 and 1905 ([BOU 12], Chapter 8). Weyl and his doctoral student Peter generalized these results to the case of arbitrary compact groups in 1927, replacing the finite sums of unitary matrices by convergent sums of such matrices ([WEI 38], Chapter 5; [DIE 93], Volume 5, (21.2.3)). In the case of non-compact groups (e.g. non-compact connected semi-simple groups), the theory is deduced from the structure of the von Neumann algebras (section 6.5.5(**I**)), now working with irreducible unitary representations that are no longer matrices but Hilbert–Schmidt operators (Example 6.151(2)). Nevertheless, everything remains relatively straightforward (and a brief presentation will suffice in section 6.5.5) provided that we remain within the formal perspective developed around 1950. The true difficulties lie in explicitly performing the calculations. This daunting task was completed by Harish-Chandra for semi-simple groups between 1965 and 1966 [HAR 66]. A representative special case – highly relevant from a pedagogical perspective – is when the group is $\mathrm{SL}_2(\mathbb{R})$ [LAN 85], [VAR 84]; sadly, we cannot explore it here[10].

10 Readers can find some details in the Wikipedia article on *Non-commutative harmonic analysis*.

6.5.2. *Harmonic analysis on* $\mathbb{R}^n$

(I) NOTATION Henceforth, write $\mathcal{L}^p(\mathbb{R}^n)$ for $\mathcal{L}^p(\mathbb{R}^n, \lambda^{n\otimes})$, where λ is the Lebesgue measure on $\mathbb{R}$ ([P2], 4.1.3**(II)**), also written as dt when $f \in \mathcal{L}^p(\mathbb{R}^n)$ is written as $t \mapsto f(t)$. If $\nu, t \in \mathbb{R}^n$, write $\nu.t$ for the usual scalar product $\sum_{1\leq i\leq n} \nu_i.t_i$. The Euclidean norm of t is written as $|t|$. The norm in $L^p(\mathbb{R}^n) := \mathcal{L}^p(\mathbb{R}^n)/\mathcal{N}$, where $\mathcal{N}$ is the subspace of $\mathcal{L}^p(\mathbb{R}^n)$ formed by the Lebesgue-negligible functions ([P2], section 4.1.2**(III)**), is written as N_p $(1 \leq p \leq \infty)$. We do not need to distinguish between a Lebesgue-measurable function f and its Lebesgue class $\mathring{f}$ ([P2], section 4.1.1**(III)**) unless there is a risk of ambiguity. Write $\mathcal{C}_0(\mathbb{R}^n)$ for the space of continuous functions $f : \mathbb{R}^n \to \mathbb{C}$ that are zero at infinity (Definition 6.19) equipped with the norm $N_\infty(f) = \sup_{t\in\mathbb{R}^n} |f(t)|$; the space $\mathcal{C}_0(\mathbb{R}^n)$ is a Banach space ([P2], section 4.1.4**(II)**).

(II) FOURIER TRANSFORM OF AN INTEGRABLE FUNCTION Let $f \in \mathcal{L}^1(\mathbb{R}^n)$. The integral

$$\boxed{\mathcal{F}[f](\nu) = \int_{\mathbb{R}^n} f(t).e^{-2\pi i\nu.t}.dt}$$

converges for every $\nu \in \mathbb{R}^n$. If f, g are equal $\lambda^{n\otimes}$-almost everywhere, then $\mathcal{F}[f](\nu) = \mathcal{F}[g](\nu)$ for every $\nu \in \mathbb{R}^n$, so the above integral can be written as $\mathcal{F}[f]$, where $f \in L^1(\mathbb{R}^n)$.

LEMMA 6.95.– (Gaussian integral) *Let* $\Sigma = diag(\sigma_1, ..., \sigma_n)$, $\sigma_j > 0$ $(j = 1, ..., n)$. *Then:*

$$\frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}x.\Sigma^{-1}x}.dx = \int_{\mathbb{R}^n} e^{-\pi.|t|^2}.dt = 1.$$

PROOF.– We can deduce the second integral from the first by performing the change of variable $x = \sqrt{2\pi}.\sqrt{\Sigma}.t$, where $\sqrt{\Sigma} = \text{diag}\left(\sqrt{\sigma_1}, ..., \sqrt{\sigma_n}\right)$. The second integral is equal to I^n, where $I = \int_{\mathbb{R}} e^{-\pi.t^2}.dt$. We have $I^2 = \int_{\mathbb{R}} e^{-\pi x^2}.dx. \int_{\mathbb{R}} e^{-\pi y^2}.dy = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)}.dx.dy$. Setting $x = \rho\cos\theta$, $y = \rho\sin\theta$ gives $I^2 = \int_0^{2\pi} d\theta \int_0^{+\infty} e^{-\pi\rho^2}\rho.d\rho = 1$ (Lemma 4.39). ■

THEOREM 6.96.– *If* $f \in L^1(\mathbb{R}^n)$, *then* $\mathcal{F}[f]$ *is uniformly continuous, bounded and zero at infinity, and* $\mathcal{F} : L^1(\mathbb{R}^n) \to \mathcal{C}_0(\mathbb{R}^n)$ *is a continuous linear mapping of norm 1.*

PROOF.– 1) By [P2], section 4.1.2**(I)**, Lemma 4.8,

$$|\mathcal{F}[f](\nu + \eta) - \mathcal{F}[f](\nu)| \leq \int_{\mathbb{R}^n} |f(t)|.\left|e^{-2\pi i\eta.t} - 1\right|.dt := \varphi(\eta).$$

By the dominated convergence theorem ([P2], section 4.1.2**(II)**), if (η_k) is a sequence of elements of $\mathbb{R}^n$ that converges to 0, then the sequence $\varphi(\eta_k)$ converges to 0, so $\mathcal{F}[f]$ is uniformly continuous.

2) We have $N_\infty(\mathcal{F}[f]) = \sup_{\nu\in\mathbb{R}^n} \left|\int_{\mathbb{R}^n} f(t).e^{-2\pi i\nu.t}\right| \leq N_1(f)$ and $\mathcal{F}$ is clearly linear. As a linear operator from $L^1(\mathbb{R}^n)$ into $L^\infty(\mathbb{R}^n)$, the norm $\|\mathcal{F}\|$ of $\mathcal{F}$ is therefore ≤ 1.

3) Let d be the function such that $d(t) = e^{-\pi.t^2}$; then, $N_1(d) = 1$ by Lemma 6.95 and, if $n = 1$, then $\mathcal{F}[d](\nu) = e^{-\pi\nu^2}\int_{\mathbb{R}} e^{-\pi z^2}.dt$, where $z = t + i\nu$, since $-\pi t^2 - 2\pi i\nu t = -\pi(t+i\nu)^2 - \pi\nu^2$. Replacing $\int_{\mathbb{R}}$ by $\int_\gamma$, where γ is a closed path commonly used to compute integrals along the real axis (see the end of [P2], section 4.2.5), we deduce that $\mathcal{F}[d] = d$ (**exercise**). If $n > 1$, $\mathcal{F}[d](\nu)$ is the product of n integrals of this type, and we again have $\mathcal{F}[d] = d$. Hence, $N_\infty(\mathcal{F}[d]) = 1$ and $\|\mathcal{F}\| = 1$.

4) We will show that $\lim_{|\nu|\to\infty}\mathcal{F}[f](\nu) = 0$. Together with (1), this is the *Riemann–Lebesgue theorem*. To simplify the notation, consider the case $n = 1$ (extending to the general case is left to the reader); then, for $|\nu| \to \infty$,

$$\left|\int_{\mathbb{R}} f(t).e^{-2\pi i\nu.t}\right| = \frac{1}{2}\left|\int_{\mathbb{R}}\left(f(t) - f\left(t+\frac{1}{2\nu}\right)\right).e^{-2\pi i\nu.t}.dt\right|$$
$$\leq \frac{1}{2}\int_{\mathbb{R}}\left|f(t) - f\left(t+\frac{1}{2\nu}\right)\right|.dt \to 0. \qquad \blacksquare$$

DEFINITION 6.97.– *The operator $\mathcal{F}$ is called the* Fourier transform. *The operator $\overline{\mathcal{F}}$ from $L^1(\mathbb{R}^n)$ into $\mathcal{C}_0(\mathbb{R}^n)$ defined by*

$$\boxed{\overline{\mathcal{F}}[\varphi](t) = \int_{\mathbb{R}^n}\varphi(\nu).e^{2\pi i\nu.t}.d\nu}$$

is called the Fourier cotransform.

Later, we will see (Theorem 6.105) that the Fourier transform (and similarly the Fourier cotransform) is injective from $L^1(\mathbb{R}^n)$ into $\mathcal{C}_0(\mathbb{R}^n)$ and that, given certain hypotheses, $\overline{\mathcal{F}}$ is the inverse transformation of $\mathcal{F}$.

If $a \in \mathbb{R}^n$ and f is a mapping from $\mathbb{R}^n$ into $\mathbb{C}$, the left translation of vector a is $\boldsymbol{\lambda}(a) : f \mapsto \boldsymbol{\lambda}(a) f$, where $\boldsymbol{\lambda}(a) f : t \mapsto f(t-a)$ (section 6.2.1**(IV)**). Write χ_a for the mapping $t \mapsto \exp(2\pi i a.t)$, so that $\chi_{-a} = \overline{\chi_a}$. Recall that $\check{f}$ denotes the mapping $t \mapsto f(-t)$ (Notation 6.12). Then:

LEMMA 6.98.– *1) The following three equalities are satisfied:*

$$\overline{\mathcal{F}}\left[\overline{f}\right]=\overline{\mathcal{F}\left[f\right]},\quad \check{\mathcal{F}}\left[f\right]=\overline{\mathcal{F}}\left[f\right],\quad \overline{\mathcal{F}}\left[\check{f}\right]=\mathcal{F}\left[f\right].$$

2) The Fourier transform exchanges translation for multiplication by an exponential. Specifically, for every $\tau\in\mathbb{R}^n$,

$$\mathcal{F}\left[\boldsymbol{\lambda}\left(\tau\right)f\right]=\overline{\chi_\tau}.\mathcal{F}\left[f\right],\quad \mathcal{F}\left[\chi_\tau.f\right]=\boldsymbol{\lambda}\left(\tau\right).\mathcal{F}\left[f\right].$$

3) If $f,\varphi\in L^1\left(\mathbb{R}^n\right)$, then we have the transfer theorem*:*

$$\int_{\mathbb{R}^n}\mathcal{F}\left[f\right]\left(\nu\right).\varphi\left(\nu\right).d\nu=\int_{\mathbb{R}^n}f\left(t\right).\mathcal{F}\left[\varphi\right]\left(t\right).dt.$$

4) The Riesz formula *holds for $f,\varphi\in L^1\left(\mathbb{R}^n\right)$:*

$$\int_{\mathbb{R}^n}\mathcal{F}\left[f\right]\left(\nu\right).\varphi\left(\nu\right).\chi_\tau\left(\nu\right).d\nu=\int_{\mathbb{R}^n}f\left(\tau+t\right).\mathcal{F}\left[\varphi\right]\left(t\right).dt.$$

PROOF.– The proof of (1) and (2) is a very straightforward **exercise**. (3): According to [P2], section 4.1.2, Lemma 4.13, and Theorem 6.96, $\mathcal{F}\left[f\right].\varphi$ and $f.\mathcal{F}\left[\varphi\right]$ belong to $\mathcal{L}^1\left(\mathbb{R}^n\right)$; we now simply need to apply the Fubini–Tonelli theorem to $f\otimes\varphi\in\mathcal{L}^1\left(\mathbb{R}^n\times\mathbb{R}^n\right)$ ([P2], section 4.1.3(**III**), Theorem 4.21). The details of the calculations are left to the reader. (4) can be deduced from (2) and (3). ■

REMARK 6.99.– *1) In signal theory, with $n=1$, the variable t denotes the time, ν denotes the frequency and $\omega=2\pi\nu$ denotes the angular frequency.*

*2) Consider a compactly supported distribution $T\in\mathcal{E}'\left(\mathbb{R}^n\right)$ ([P2], section 4.4.1(**I**)). Since $\overline{\chi_\nu}\in\mathcal{E}\left(\mathbb{R}^n\right)$, we can set*

$$\mathcal{F}\left[T\right]\left(\nu\right)=\left\langle T,\overline{\chi_\nu}\right\rangle, \qquad [6.9]$$

and $\nu\mapsto\mathcal{F}\left[T\right]\left(\nu\right)$ is a function of class C^∞. In particular, $\mathcal{F}\left[\delta\right]\left(\nu\right)=1$ and $\delta_\tau=\boldsymbol{\lambda}\left(\tau\right)\delta$, so $\mathcal{F}\left[\delta_\tau\right]\left(\nu\right)=e^{-2\pi i\nu.\tau}=\overline{\chi_\tau}\left(\nu\right)$.

THEOREM 6.100.– (exchange theorem) *If $f,g\in\mathcal{L}^1\left(\mathbb{R}^n\right)$, then $f\star g\in\mathcal{L}^1\left(\mathbb{R}^n\right)$ and $\mathcal{F}\left[f\star g\right]=\mathcal{F}\left[f\right].\mathcal{F}\left[g\right]$.*

PROOF.– We know that $\mathcal{L}^1\left(\mathbb{R}^n\right)$ is a convolution algebra (Theorem 6.20) and

$$\begin{aligned}\mathcal{F}\left[f\star g\right]\left(\nu\right)&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}f\left(t-\tau\right).g\left(\tau\right).e^{-2\pi i\nu.(t-\tau)}.e^{-2\pi i\nu.\tau}.dt.d\tau\\&=\mathcal{F}\left[f\right]\left(\nu\right).\mathcal{F}\left[g\right]\left(\nu\right).\end{aligned}$$

■

(III) SOBOLEV SPACES For every $p \in [1, +\infty]$, write $W^{k,p}(\mathbb{R}^n)$ for the space of functions $f : \mathbb{R}^n \to \mathbb{R}$ that belong to $\mathcal{L}^p$ together with their partial derivatives $\partial^\alpha f$ in the sense of distribution for the multi-indices $\alpha = (\alpha_1, ..., \alpha_n)$ such that $|\alpha| \leq k$ ([P2], section 4.4.1(**II**)). The space $W^{k,p}(\mathbb{R}^n)$ is equipped with the norm

$$N_{k,p}(f) = \left(\sum_{0 \leq |\alpha| \leq k} (N_p(\partial^\alpha f))^p \right)^{1/p} \quad \text{if } p < \infty,$$

$$N_{k,\infty}(f) = \sup_{0 \leq |\alpha| \leq k} N_\infty(\partial^\alpha f).$$

The space $W^{k,p}(\mathbb{R}^n)$ is called the Sobolev space of indices k, p. This is a Banach space that coincides with the completion of $\mathcal{E}(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$ for the norm $N_{k,p}$ (*Meyers–Serrin theorem* [MEY 64]); $W^{k,2}(\mathbb{R}^n)$ is a Hilbert space (**exercise**).

Let $\mathcal{C}_b^k(\mathbb{R}^n)$ (respectively $\mathcal{C}_0^k(\mathbb{R}^n)$) be the space of bounded (respectively zero at infinity) functions of class C^k in $\mathbb{R}^n$ together with their partial derivatives of order $\leq k$, equipped with the norm obtained by restriction of the norm of $W^{k,\infty}(\mathbb{R}^n)$. In 1938 and 1950, Sobolev showed the following result from Hölder's inequality ([P2], section 4.1.2, Lemma 4.13) (**exercise***: see ([ADA 03], Theorem 4.12; [KHO 72], Volume 2, Chapter DC, section III.3)); the argument $\mathbb{R}^n$ is omitted to lighten the notation.

THEOREM 6.101.– (Sobolev embedding theorem) *Let k be an integer ≥ 0. If $mp > n$, there exist canonical injections $W^{k+m,p} \hookrightarrow \mathcal{C}_b^k$ and $W^{k+m,p} \hookrightarrow W^{k,q}$ whenever $p \leq q \leq \infty$. For $p = 2$, there exists a continuous injection $W^{k+m,2} \hookrightarrow \mathcal{C}_0^k$. If $mp = n$, there exist continuous injections $W^{k+m,p} \hookrightarrow W^{k,p} \hookrightarrow \mathcal{L}^r$ whenever $p \leq q \leq r < \infty$.*

The completion of $\mathcal{D} \cap W^{k,p}$ for the norm $N_{k,p}$ is a closed subspace $W_0^{k,p}$ of $W^{k,p}$; the spaces $W^{k,p}$ and $W_0^{k,p}$ are separable for $1 \leq p < \infty$ and reflexive for $1 < p < \infty$ ([GIL 98], section 7.5).

Given any multi-index $\alpha = (\alpha_1, ..., \alpha_n)$, write $\boldsymbol{m}^\alpha$ for the mapping $(t_1, ..., t_n) \longmapsto \prod_{1 \leq i \leq n} t_i^{\alpha_i}$.

THEOREM 6.102.– (exchange theorem between differentiation and monomial multiplication) *Let $k \geq 0$.*

i) Let f be a function such that $\boldsymbol{m}^\alpha . f \in \mathcal{L}^1$ for every multi-index α such that $|\alpha| \leq k$. Then, $\mathcal{F}[f] \in W^{k,\infty} \cap \mathcal{C}^k$ and

$$\partial^\alpha (\mathcal{F}[f]) = \mathcal{F}[(-2\pi i \boldsymbol{m})^\alpha . f]. \qquad [6.10]$$

ii) Let $f \in W^{k,1} \cap \mathcal{C}^k$. Then, $\boldsymbol{m}^{\beta}.(\mathcal{F}.f) \in \mathcal{L}^{\infty}$ for every multi-index β such that $|\beta| \leq k$, and

$$(2\pi i\boldsymbol{m})^{\beta}.\mathcal{F}[f] = \mathcal{F}\left[\partial^{\beta} f\right]. \qquad [6.11]$$

PROOF.– To simplify the notation, consider the case $n = 1$. i) The function $\nu \mapsto -2\pi i\nu.t.e^{-2\pi i\nu.t}$ is bounded above by the integrable function $t \mapsto 2\pi\,|t.f(t)|$. We can therefore differentiate under the $\int$ sign with respect to ν according to [P2], section 4.1.2(**II**), Theorem 4.11, which establishes the result for $\alpha = 1$ by Theorem 6.96. The result follows by induction.

ii) For $\beta = 1$, the right-hand side of [6.11] is $\int_{\mathbb{R}} \dot{f}(t).e^{-2\pi i\nu.t}.dt$, which we integrate by parts, then proceed by induction. ■

Theorem 6.102 enables us to establish the relation $\overline{\mathcal{F}} = \mathcal{F}^{-1}$ for a function $f \in \mathcal{L}^1(\mathbb{R})$ on which a few additional conditions have been imposed (see Theorem 6.122(i) and Corollary 6.123, as well as **(IV)** below).

(IV) FOURIER TRANSFORM IN $\mathcal{S}(\mathbb{R}^n)$ The Fourier transform [6.9] has remarkable properties in the space $\mathcal{S}(\mathbb{R}^n)$ of declining functions ([P2], section 4.3.1). Recall that $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet–Schwartz space[11] (*ibid.*, Theorem 4.72), and hence a Montel space ([P2], section 3.4.6(**III**), Theorem 3.76). Its dual, the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions, is therefore also a Montel space ([P2], section 3.7.5, Theorem 3.123). The topology of $\mathcal{S}(\mathbb{R}^n)$ can be defined using the countable family of seminorms $(q_{\alpha,\beta})$ from ([P2], section 4.3.1(**III**)), or alternatively using the family $(\hat{q}_{s,m})$, where $\hat{q}_{s,m}(f) = \sup_{|\beta|\leq s}\left(\left(1+|\boldsymbol{m}|^2\right)^m.\partial^{\beta} f\right)$ (**exercise**).

LEMMA 6.103.– *Let $d_{\sigma}(x) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^n} e^{-|x|^2/\left(2\sigma^2\right)}$ and suppose that (σ_k) is a sequence of real numbers > 0 converging to 0. Then, the sequence of functions (d_{σ_k}) converges to the Dirac distribution δ in the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$.*

PROOF.– It suffices that to show that, for every test function $\varphi \in \mathcal{S}(\mathbb{R}^n)$, the sequence $(\langle d_{\sigma_k}, \varphi\rangle)$ converges to $\varphi(0)$ (Theorem 5.3). To simplify the notation, consider the case $n = 1$. In this case:

$$\langle d_{\sigma_k}, \varphi\rangle = \frac{1}{\sqrt{2\pi}\sigma_k}\int_{\mathbb{R}} e^{-\pi x^2/\sigma_k^2}.\varphi(x).dx = \int_{\mathbb{R}} e^{-\pi.t^2}.\varphi\left(\sqrt{2\pi}\sigma_k.t\right).dt.$$

The sequence of functions g_k that appear under the $\int$ sign converges pointwise to $t \mapsto e^{-\pi.t^2}.\varphi(0)$ for $k \to +\infty$, and $|g_k(t)| \leq g(t)$, where $g(t) = e^{-\pi t^2} N_{\infty}(\varphi)$

11 Some authors call $\mathcal{S}(\mathbb{R}^n)$ the "Schwartz space", which could potentially lead to ambiguity with the *category* of Schwartz spaces ([P2], section 3.4.6(**II**)).

and g is integrable, so the dominated convergence theorem ([P2], section 4.1.2(**II**)) implies that $\int_{\mathbb{R}} g_k(t).dt \to \int_{\mathbb{R}} e^{-\pi t^2}.\varphi(0).dt = \varphi(0)$. ■

THEOREM 6.104.– (automorphism theorem) *The Fourier transform $\mathcal{F}$ is an automorphism (of locally convex spaces) of $\mathcal{S}(\mathbb{R}^n)$ with inverse automorphism $\overline{\mathcal{F}}$ (reciprocity formula).*

PROOF.– 1) Let $f \in \mathcal{S}$. Since $\boldsymbol{m}^\alpha.f \in S$ for every multi-index β, $\mathcal{F}[f]$ is of class C^∞ by [6.10]. Similarly, $\mathcal{F}[f]$ is rapidly decreasing by [6.11]. It can then be shown by induction that $\mathcal{F}[f] \in \mathcal{S}$.

2) By Theorem 6.102,

$$(2\pi i\boldsymbol{m}^\alpha)\,\partial^\beta(\mathcal{F}.f) = (2\pi i\boldsymbol{m}^\alpha).\mathcal{F}\left((-2\pi i\boldsymbol{m})^\beta f\right) = \mathcal{F}\left(\partial^\alpha(-2\pi i\boldsymbol{m})^\beta f\right).$$

Now, observe that $\left(1+|\boldsymbol{m}|^2\right)^{-n} \in \mathcal{L}^1$. It is possible to find a constant $c_{s,m}$ such that $\hat{q}_{2m,s+n+1}(f) \leq 1$ implies $\hat{q}_{s,m}(\mathcal{F}.f) \leq c_{s,m}$ (**exercise***: see ([DIE 93], Volume 6, (22.16.10))), which proves the continuity of $\mathcal{F} : \mathcal{S} \to \mathcal{S}$.

3) We can show that $\overline{\mathcal{F}} \circ \mathcal{F} = \mathrm{id}_{\mathcal{S}(\mathbb{R}^n)}$ by calculating the following integral for $f \in \mathcal{S}$:

$$\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)[f](\tau) = \int_{\mathbb{R}^n} \mathcal{F}[f](\nu).\varphi(\nu).\chi_\tau(\nu).d\nu,$$

where $\varphi(\nu) = 1 = \mathcal{F}[\delta]$. By Riesz's formula (Lemma 6.98(4)),

$$\begin{aligned}\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)[f](\tau) &= \int_{\mathbb{R}^n\times\mathbb{R}^n} f(t).e^{2\pi i\nu.(\tau-t)}.dt \otimes d\nu \\ &= \int_{\mathbb{R}^n} \mathcal{F}[f](\nu).\varphi(\nu).e^{2\pi i\nu.\tau}.d\nu.\end{aligned}$$

Replacing φ by $\overline{\mathcal{F}}[d_{\sigma_k}]$ and taking the limit gives $\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)[f](\tau) = f(\tau)$. ■

The same calculation as in part (3) of the above proof shows the following result:

THEOREM 6.105.– *If $f \in \mathcal{L}^1(\mathbb{R}^n)$ and $\mathcal{F}[f] \in \mathcal{L}^1(\mathbb{R}^n)$, then $\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)[f](\tau) = f(\tau)$ almost everywhere. In particular, if $f \in \mathcal{L}^1(\mathbb{R}^n)$ and $\mathcal{F}[f] = 0$, then $f(\tau) = 0$ almost everywhere.*

The above calculations also lead to the following result (**exercise***: see ([DIE 93], Volume 6, (22.18.2))), which is analogous to Theorem 6.100:

THEOREM 6.106.– (exchange theorem) *If $f, g \in \mathcal{S}(\mathbb{R}^n)$, then $f \star g \in \mathcal{S}(\mathbb{R}^n)$ and $\mathcal{F}[f \star g] = \mathcal{F}[f].\mathcal{F}[g]$. The bilinear mapping $(f,g) \mapsto f \star g$ is continuous from $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$.*

We also have the following result:

COROLLARY 6.107.– (Plancherel–Parseval) *Let $g, h \in \mathcal{S}(\mathbb{R}^n)$. Then:*

$$\int_{\mathbb{R}^n} \overline{h(t)}.g(t).dt = \int_{\mathbb{R}^n} \overline{\mathcal{F}[h](\nu)}.\mathcal{F}[g](\nu).d\nu.$$

PROOF.– Simply apply Riesz's formula (Lemma 6.98(4)) with $f = \overline{\mathcal{F}[h]}, g = \varphi$ and $\tau = 0$. ■

REMARK 6.108.– *Instead of $\mathcal{F}$, we often consider the "normalized Fourier transform" $\mathcal{F}_{\mathbf{n}}[f](\omega) = \int_{\mathbb{R}^n} f(t).e^{-i\omega.t}.dt$, where $\omega = 2\pi\nu$ can be interpreted as an angular frequency. The reader is invited to check that $\mathcal{F}_{\mathbf{n}}^{-1} = \frac{1}{(2\pi)^n}\overline{\mathcal{F}_{\mathbf{n}}}$.*

(V) CONVOLUTION BY A TEMPERED DISTRIBUTION The convolution product of two tempered distributions is not defined in general. For example, the constant function $\mathbf{1} : t \mapsto 1$ belongs to $\mathcal{S}'(\mathbb{R}^n)$, but there is no non-trivial definition for its convolution product with itself, since

$$\int_{-a}^{a} \mathbf{1}(t-\tau).\mathbf{1}(\tau).d\tau = 2a$$

tends to $+\infty$ as $a \to +\infty$. If $\psi \in \mathcal{S}(\mathbb{R}^n)$ and $T \in \mathcal{S}'(\mathbb{R}^n)$, then, for every test function $\varphi \in \mathcal{S}(\mathbb{R}^n)$, the convolution product $\check{\psi} \star \varphi$ belongs to $\mathcal{S}(\mathbb{R}^n)$. Thus, we can set:

$$\boxed{\langle T \star \psi, \varphi\rangle = \left\langle T, \varphi \star \check{\psi}\right\rangle.}$$

We have the following relations (**exercise**):

$$\check{T} \star \check{\psi} = (T \star \psi)^{\vee}, \quad \boldsymbol{\lambda}(a)(T \star \psi) = (\boldsymbol{\lambda}(a)T) \star \psi, \quad \partial^{\alpha}(T \star \psi) = (\partial^{\alpha}T) \star \psi.$$

The above still holds when ψ is replaced by a distribution U such that $U \star \varphi \in \mathcal{S}(\mathbb{R}^n)$ whenever $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

DEFINITION 6.109.– *We say that a distribution U is a convolutor (for $\mathcal{S}(\mathbb{R}^n)$) if the mapping $\varphi \mapsto U \star \varphi$ sends $\mathcal{S}(\mathbb{R}^n)$ continuously into $\mathcal{S}(\mathbb{R}^n)$.*

A convolutor function $t \mapsto \psi(t)$ satisfies the property that, for any integer $k \geq 0$, the function $t \mapsto \left(1 + |t|^2\right)^{k/2} \psi(t)$ is bounded. It is clear that the set of convolutors for $\mathcal{S}(\mathbb{R}^n)$ forms a complex vector space. By Theorem 6.2 and Remark 6.4(1), this vector space contains $\mathcal{E}'(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$.

DEFINITION 6.110.– *The space of convolutors is written as $\mathcal{O}'_c(\mathbb{R}^n)$ and is called the space of rapidly decreasing distributions.*

By the above,

$$\mathcal{E}'(\mathbb{R}^n) + \mathcal{S}(\mathbb{R}^n) \subset \mathcal{O}'_c(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n).$$

If $U \in \mathcal{O}'_c(\mathbb{R}^n)$, then $\check{U}, \boldsymbol{\lambda}(a).U, \partial^\beta U, \boldsymbol{m}^\alpha U$ belong to $\mathcal{O}'_c(\mathbb{R}^n)$ (**exercise**). In Theorem 6.112(v) below, we will see that the space of convolutors $\mathcal{O}'_c(\mathbb{R}^n)$ is isomorphic to the space $\mathcal{O}_M(\mathbb{R}^n)$ of multipliers, i.e. the space of slowly increasing infinitely differentiable functions ([P2], section 4.4.1**(V)**).

Recall that the following conditions are equivalent: (i) the linear mapping $T \mapsto g.T$ is a continuous endomorphism of $\mathcal{S}(\mathbb{R}^n)$; (ii) the linear mapping $T \mapsto g.T$ is a continuous endomorphism of $\mathcal{S}'(\mathbb{R}^n)$; (iii) $g \in \mathcal{O}_M(\mathbb{R}^n)$, i.e. $g \in \mathcal{E}(\mathbb{R}^n)$ and, for every multi-index α, $|\partial^\alpha g|$ is bounded above by a polynomial. The space $\mathcal{O}_M(\mathbb{R}^n)$ can be equipped with a locally convex topology presented in ([SCH 66], Chapter 7, section 5) for which the multiplication $\times : \mathcal{O}_M(\mathbb{R}^n) \times \mathcal{O}_M(\mathbb{R}^n) \to \mathcal{O}_M(\mathbb{R}^n)$ is continuous and $\mathcal{O}_M(\mathbb{R}^n)$ is complete, thus giving a complete topological $\mathbb{C}$-algebra ([P2], section 2.8.3).

Similarly, $\mathcal{O}'_c(\mathbb{R}^n)$ can be equipped with a locally convex topology presented in [SCH 66], Chapter 7, section 5) for which the convolution product $\star : \mathcal{O}'_c(\mathbb{R}^n) \times \mathcal{O}'_c(\mathbb{R}^n) \to \mathcal{O}'_c(\mathbb{R}^n)$ is continuous.

(VI) FOURIER TRANSFORMS IN $\mathcal{S}'(\mathbb{R}^n)$

DEFINITION 6.111.– *The Fourier transform from $\mathcal{S}'(\mathbb{R}^n)$ into $\mathcal{S}'(\mathbb{R}^n)$ is the transpose of $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, again written as $\mathcal{F}$.*

If $T \in \mathcal{S}'(\mathbb{R}^n)$, then, for any test function $\varphi \in \mathcal{S}(\mathbb{R}^n)$, write:

$$\langle \mathcal{F}[T], \varphi \rangle = \langle T, \mathcal{F}[\varphi] \rangle, \quad \langle \overline{\mathcal{F}}[T], \varphi \rangle = \langle T, \overline{\mathcal{F}}[\varphi] \rangle.$$

THEOREM 6.112.– *i) The Fourier transform is an automorphism of $\mathcal{S}'(\mathbb{R}^n)$ whose inverse automorphism $\overline{\mathcal{F}}$ is the transpose of $\overline{\mathcal{F}} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$.*

ii) If $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and $T \in \mathcal{S}'(\mathbb{R}^n)$, then $\langle \mathcal{F}[T \star \psi], \varphi \rangle = \langle T \star \psi, \mathcal{F}[\varphi] \rangle$.

iii) $\mathcal{O}'_c(\mathbb{R}^n)$ *is the space of convolutors for* $\mathcal{S}'(\mathbb{R}^n)$*, i.e. the linear mapping* $T \mapsto T \star U$ *is a continuous endomorphism of* $\mathcal{S}'(\mathbb{R}^n)$ *if and only if* $U \in \mathcal{O}'_c(\mathbb{R}^n)$.

iv) $\mathcal{O}_M(\mathbb{R}^n)$ *is a* $\mathbb{C}$*-algebra and* $\mathcal{O}'_c(\mathbb{R}^n)$ *is a convolution algebra.*

v) The Fourier transform $\mathcal{F}$ *is an isomorphism of locally convex spaces* $\mathcal{O}_M(\mathbb{R}^n)\tilde{\rightarrow}\mathcal{O}'_c(\mathbb{R}^n)$ *whose inverse isomorphism is* $\overline{\mathcal{F}}$*. If* $T \in \mathcal{S}'(\mathbb{R}^n)$*,* $U \in \mathcal{O}'_c(\mathbb{R}^n)$*, then the* exchange theorem *holds:*

$$\boxed{\mathcal{F}\left[\underbrace{U}_{\substack{\in\\ \mathcal{O}'_c(\mathbb{R}^n)}} \star \underbrace{T}_{\substack{\in\\ \mathcal{S}'(\mathbb{R}^n)}}\right] = \underbrace{\mathcal{F}[U]}_{\substack{\in\\ \mathcal{O}_M(\mathbb{R}^n)}} \cdot \underbrace{\mathcal{F}[T]}_{\substack{\in\\ \mathcal{S}'(\mathbb{R}^n)}} .}$$

PROOF.– (i) and (ii): See [P2], section 3.5.3. (iii), (iv): **exercise**. (v): Let $T \in \mathcal{S}'$ and suppose that U is a distribution. Then:

$$T \star U \in \mathcal{S}' \Leftrightarrow \mathcal{F}[T \star U] \in \mathcal{S}' \iff \mathcal{F}[T].\mathcal{F}[U] \in \mathcal{S}'.$$

But, for any test function $\varphi \in \mathcal{S}$,

$$\langle \mathcal{F}[T].\mathcal{F}[U], \varphi\rangle = \langle \mathcal{F}[T], \mathcal{F}[U].\varphi\rangle$$

and $[\forall\varphi \in \mathcal{S}, \mathcal{F}[U].\varphi \in \mathcal{S}] \iff \mathcal{F}[U] \in \mathcal{O}_M$. Similarly,

$$\langle T \star S, \varphi\rangle = \langle T, \varphi \star \check{S}\rangle$$

and $\left[\forall\varphi \in \mathcal{S}, \varphi \star \check{U} \in \mathcal{S}'\right] \iff \check{U} \in \mathcal{O}'_c \iff U \in \mathcal{O}'_c$. Finally, a net (U_α) converges to 0 in $\mathcal{O}'_c$ if and only if $(\mathcal{F}[U_\alpha])$ converges to 0 in $\mathcal{O}_M$. The rest is clear. ■

EXAMPLE 6.113.– *1) By Remark 6.99(2),* $\mathcal{F}[\delta] = 1$*, so* $\mathcal{F}[\partial^\alpha\delta_\tau] = (2\pi i\boldsymbol{m})^\alpha\,\overline{\chi_\tau} : \nu \mapsto (2\pi i\nu)^\alpha e^{-2\pi i\nu.\tau}$.

2) We have $\mathcal{F}[\chi_\lambda] = \delta_\lambda$*, since* $\overline{\mathcal{F}}[\delta_\lambda] = \langle\delta_\lambda, \chi_t\rangle = \chi_\lambda$*, so* $\delta_\lambda = \mathcal{F}\left[\overline{\mathcal{F}}[\delta_\lambda]\right] = \mathcal{F}[\chi_\lambda]$*. In "abusive notation", this equality can be expressed as follows:*

$$\int_{\mathbb{R}^n} e^{2\pi i(\lambda-\nu).t}.dt = \delta(\nu - \lambda).$$

3) With $n = 1$*, let* $\mathbf{e}_\lambda : t \mapsto e^{-\lambda.t}$ *and write* Υ *for the Heaviside function ([P2], section 4.1.1(**II**), Example 4.4(4)). If* $\lambda > 0$*,* $\mathcal{F}[\mathbf{e}_\lambda.\Upsilon] = \frac{1}{\lambda+2\pi i\bullet} : \nu \mapsto \frac{1}{\lambda+2\pi i\nu}$.

If $\lambda = 1$*, then* $\mathbf{e}_\lambda.\Upsilon = \Upsilon$ *is a tempered distribution whose Fourier transform is* $p.v.\left(\frac{1}{2\pi i \bullet}\right)$*, where* $p.v.$ *denotes the* Cauchy principal value*, namely the tempered distribution such that, for every test function* $\varphi \in \mathcal{S}(\mathbb{R})$*,*

$$\left\langle p.v.\left(\frac{1}{2\pi i \bullet}\right), \varphi\right\rangle = \lim_{\varepsilon \to 0^+} \left\{\int_{-\infty}^{-\varepsilon} \frac{\varphi(\nu)}{2\pi i \nu}.d\nu + \int_{\varepsilon}^{+\infty} \frac{\varphi(\nu)}{2\pi i \nu}.d\nu\right\}.$$

(VII) FOURIER TRANSFORMS IN $L^2(\mathbb{R}^n)$ It is possible to show the following result ([TRÈ 67], Chapter 15, Corollary 3 of Theorem 15.3):

LEMMA 6.114.– *Let* Ω *be a non-empty open subset of* $\mathbb{R}^n$*. If* $p \in [1, \infty[$*, then* $\mathcal{D}(\Omega)$ *is dense in* $L^p(\Omega)$*, and in particular,* $\mathcal{S}(\mathbb{R}^n)$ *is dense in* $L^p(\mathbb{R}^n)$*.*

THEOREM 6.115.– (Plancherel–Parseval) *The Fourier transform* $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ *can be extended to an isometry from* $L^2(\mathbb{R}^n)$ *onto* $L^2(\mathbb{R}^n)$ *whose inverse isometry is* $\overline{\mathcal{F}}$*. In other words, if* $f, g \in L^2(\mathbb{R}^n)$*, then*

$$\boxed{\langle \mathcal{F}[f], \mathcal{F}[g]\rangle_2 = \langle f, g\rangle_2 := \int_{\mathbb{R}^n} \overline{f(t)}.g(t).dt.}$$

PROOF.– This theorem follows from Corollary 6.107 and Lemma 6.114. ■

REMARK 6.116.– *Let* $f \in \mathcal{L}^2(\mathbb{R}^n)$ *and suppose that* (B_k) *is a sequence of bounded measurable subsets of* $\mathbb{R}^n$ *satisfying* $B_k \Subset B_{k+1}$ *and* $\bigcup_k B_k = \mathbb{R}^n$*. Let* $g_k(\nu) = \int_{\mathbb{R}^n} f_k(t).\overline{\chi_\nu(t)}.dt$*, where* $f_k = f.1_k$*,* 1_k *denotes the characteristic function of* B_k *([P2], section 2.3.3(**III**)). Since* $f_k \in \mathcal{L}^1(\mathbb{R}^n)$*,* g_k *is uniformly continuous, bounded and zero at infinity (Theorem 6.96). Thus,* $\mathcal{F}[f]$ *is the limit* in $L^2(\mathbb{R}^n)$ *of the sequence of functions* (g_k)*. This reasoning shows that* $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ *is dense in* $L^2(\mathbb{R}^n)$ *and that the Fourier transform defined in* $L^2(\mathbb{R}^n)$ *extends the Fourier transform defined in* $L^1(\mathbb{R}^n)$*.*

(VII) FOURIER TRANSFORMS IN $\mathcal{E}'(\mathbb{R}^n)$ The next result presents recent developments in “time-frequency” approaches and the importance of wavelets [FLA 98]:

THEOREM 6.117.– $\mathcal{E}'(\mathbb{R}^n) \cap \mathcal{F}[\mathcal{E}'(\mathbb{R}^n)] = \{0\}$.

PROOF.– If $T \in \mathcal{E}'(\mathbb{R}^n)$, $f := \mathcal{F}[T]$ can be extended to an entire function in $\mathbb{C}^n$ by Remark 6.99. Hence, f is analytic in $\mathbb{R}^n$ ([P2], sections 4.3.2(**I**), (**II**)). If $f \in \mathcal{E}'(\mathbb{R}^n)$, then f is compactly supported and hence zero by the principle of analytic continuation[12]. ■

12 We can further specify $\mathcal{F}[\mathcal{E}'(\mathbb{R}^n)]$ using the Paley–Wiener–Schwartz theorem (see the Wikipedia article on the *Paley–Wiener theorem*).

6.5.3. *Fourier series and Fourier transforms on the torus*

(I) FOURIER SERIES EXPANSION OF A FUNCTION Write $l^p(\mathbb{Z}^n)$ for the space $\mathcal{L}^p(\mathbb{Z}^n, \mathfrak{m}_{\mathbb{Z}^n})$ (with Convention **(C3)**, section 6.2.2**(II)**). Thus, $l^p(\mathbb{Z}^n)$ is a Banach space equipped with the norm $N_p(a) = \left(\sum_{k\in\mathbb{Z}^n} |a_k|^p\right)^{1/p}$ if $p \in [1,\infty[$, $N_\infty(a) = \sup_{k\in\mathbb{Z}^n} |a_k|$ if $p = \infty$, and $l^2(\mathbb{Z}^n)$ is a separable Hilbert space ([P2], section 3.10.2**(VII)**). Write $c_0(\mathbb{Z}^n)$ for the space of sequences of complex numbers $a = (a_k)_{k\in\mathbb{Z}^n}$ such that $\lim_{|k|\to\infty} a_k = 0$; this is a closed subspace of the Banach space $l^\infty(\mathbb{Z}^n)$ (**exercise**) and $c_0(\mathbb{Z}^n)^\vee \cong l^1(\mathbb{Z}^n)$ ([YOS 80], Chapter 4, section 9, Example 1). There are continuous canonical injections $l^p(\mathbb{Z}^n) \hookrightarrow l^q(\mathbb{Z}^n)$ $(1 \le p \le q \le \infty)$ and $l^p(\mathbb{Z}^n) \hookrightarrow c_0(\mathbb{Z}^n) \hookrightarrow l^\infty(\mathbb{Z}^n)$ $(1 \le p < \infty)$ (**exercise**).

Let $a \in l^1(\mathbb{Z}^n)$. The trigonometric series

$$\boxed{\mathcal{F}[a](t) = \sum\nolimits_{k\in\mathbb{Z}^n} a_k.e^{-2\pi i k.t} = \sum\nolimits_{k\in\mathbb{Z}^n} a_k.\overline{\chi_k}(t)}$$

converges uniformly and is therefore a continuous function in $\mathbb{R}^n$.

Let $T = (T_1, ..., T_n) \in \left(\mathbb{R}_+^\times\right)^n$. A function $f : \mathbb{R}^n \to \mathbb{C}$ is said to be T-periodic if $\boldsymbol{\lambda}(T).f = f$. After normalizing, the study of T-periodic functions can be reduced to the study of $(1, ..., 1)$-periodic functions; the same is true for distributions. To simplify the notation and the terminology, simply write 1 for the unit period $(1, ..., 1)$ and say that a 1-periodic function is periodic.

If $a \in l^1(\mathbb{Z}^n)$, then $\mathcal{F}[a]$ is clearly a periodic function. Writing $\mathbb{I}$ for the *unit hypercube* $[-1/2, 1/2]^n$ and $\mathcal{C}_{\mathbb{I}}(\mathbb{R}^n)$ (respectively $\mathcal{C}_{\mathbb{I}}^m(\mathbb{R}^n)$) for the space of continuous periodic functions (respectively periodic functions of class C^m), we therefore have $\mathcal{F}[a] \in \mathcal{C}_{\mathbb{I}}(\mathbb{R}^n)$.

Write $\mathcal{L}_{\mathbb{I}}^1(\mathbb{R}^n)$ for the space of periodic functions $f : \mathbb{R}^n \to \mathbb{C}$ that are integrable on $\mathbb{I}$. For every $k \in \mathbb{Z}^n$, we can calculate the integral

$$\boxed{\overline{\mathcal{F}}[f]_k = \int_{\mathbb{I}} f(t).e^{2\pi i k.t}.dt = \int_{\mathbb{I}} f(t).\chi_k(t).dt,} \qquad [6.12]$$

and $\left|\overline{\mathcal{F}}[f]_k\right| \le N_{1,\mathbb{I}}(f)$, where $N_{1,\mathbb{I}}(f) = \int_{\mathbb{I}} |f(t)|.dt$. The same reasoning as the proof of the Riemann–Lebesgue theorem (part (4) of the proof of Theorem 6.96) shows that $\lim_{|k|\to\infty} \overline{\mathcal{F}}[f]_k = 0$ (*Riemann summability theorem*). Clearly, $\overline{\mathcal{F}}[f]_k$ only depends on the restriction $f|_{\mathbb{I}} \in \mathcal{L}^1(\mathbb{I})$. We have therefore obtained the following result:

LEMMA 6.118.– *Let $f \in \mathcal{L}_{\mathbb{I}}^1(\mathbb{R}^n)$ or $f \in \mathcal{L}^1(\mathbb{I})$. Then, $\overline{\mathcal{F}}[f] \in c_0(\mathbb{Z}^n)$ and $N_\infty\left(\overline{\mathcal{F}}[f]\right) \le N_{1,\mathbb{I}}(f)$.*

In [6.12], we can of course replace the hypercube $\mathbb{I}$ by any other unit hypercube of the form $[-1/2+\alpha, 1/2+\alpha]^n$, $\alpha \in \mathbb{R}$. Analogously to Definition 6.97, consider the following definition:

DEFINITION 6.119.– *The operator* $\mathcal{F} : l^1(\mathbb{Z}^n) \to \mathcal{C}_{\mathbb{I}}(\mathbb{R}^n) : a \mapsto \mathcal{F}[a]$ *is called the* Fourier transform *in* $l^1(\mathbb{Z}^n)$. *The operator* $\overline{\mathcal{F}} : \mathcal{L}^1_{\mathbb{I}}(\mathbb{R}^n) \to c_0(\mathbb{Z}^n)$ *(respectively* $\mathcal{L}^1(\mathbb{I}) \to c_0(\mathbb{Z}^n)$*) is called the* Fourier cotransform *in* $\mathcal{L}^1_{\mathbb{I}}(\mathbb{R}^n)$ *(respectively* $\mathcal{L}^1(\mathbb{I})$*). The* $\overline{\mathcal{F}}[f]_k$ $(k \in \mathbb{Z}^n)$ *are called the* Fourier coefficients *of* $f \in \mathcal{L}^1_{\mathbb{I}}(\mathbb{R}^n)$ *(respectively* $f \in \mathcal{L}^1(\mathbb{I})$*).*

By Theorem 6.20, $l^1(\mathbb{Z}^n)$ is a convolution algebra whose convolution product $c = a \star b$ is given by

$$c_k = \sum_{l \in \mathbb{Z}^n} a_{k-l}.b_l.$$

THEOREM 6.120.– *Parts (1) and (2) of Lemma 6.98 remain valid in this new setting,* mutatis mutandis*, as well as the statement of Theorem 6.100; furthermore, if* $f \in W^{m,1}(\mathbb{I})$[13]*, then, for any multi-index* β *such that* $|\beta| \leq m$*, the following relation holds, analogous to [6.11]:*

$$(-2\pi i \boldsymbol{m})^{\beta}.\overline{\mathcal{F}}[f] = \overline{\mathcal{F}}\left[\partial^{\beta} f\right],$$

where $(-2\pi i \boldsymbol{m})^{\beta}.\overline{\mathcal{F}}[f]$ *denotes the sequence* $k \mapsto (-2\pi i k)^{\beta}.\overline{\mathcal{F}}[f]_k$.

DEFINITION 6.121.– *Let* $f \in \mathcal{L}^1_{\mathbb{I}}(\mathbb{R}^n)$. *The (possibly non-convergent) trigonometric series* $\left(\mathcal{F} \circ \overline{\mathcal{F}}\right)[f]$ *is called the* Fourier series expansion *(or* Fourier series*) of* f.

In the case $n = 1$, $\mathbb{I} = \left[-\frac{1}{2}, \frac{1}{2}\right]$, two of the most important results concerning pointwise convergence are as follows ([KHO 72], Volume 2, Chapter CB, section IV.2; Chapter CC, section IV.6):

THEOREM 6.122.– *i)* (Fourier–Dirichlet) *Let* $f \in \mathcal{L}^1(\mathbb{R})$ *be a function of locally bounded variation ([P2], section 4.1.7(**II**)). Then, for every* $t \in \mathbb{R}$*, the principal value (Example 6.113(2))*

$$p.v. \int_{-\infty}^{+\infty} \mathcal{F}[f](\nu).\chi_{\nu}(t).d\nu = \frac{1}{2}[f(t_+) + f(t_-)]$$

converges uniformly (with respect to T*) on any compact interval where* f *is continuous.*

13 The space $W^{m,p}(\mathbb{I})$ is the Sobolev space defined in section 6.5.2(**II**), replacing $\mathbb{R}^n$ by $\mathbb{I}$.

*ii) (*Jordan–Dirichlet*) Let f be a periodic function of locally bounded variation and write $\mathcal{F}[f]_k$ $(k \in \mathbb{Z})$ for its Fourier coefficients. For every $t \in \mathbb{R}$, the principal value*

$$p.v. \sum_{k=-\infty}^{+\infty} \mathcal{F}[f]_k . \chi_k(t) = \frac{1}{2}[f(t_+) + f(t_-)]$$

converges uniformly (with respect to t) on any compact interval where f is continuous (where $p.v. \sum_{k=-\infty}^{+\infty} = \lim_{k \to +\infty} \sum_{-k}^{k}$).

COROLLARY 6.123.– *i) If $f \in \mathcal{L}^1(\mathbb{R})$ (respectively $\mathcal{L}^1_{\mathbb{I}}(\mathbb{R})$) is continuous and locally of bounded variation, and if $\overline{\mathcal{F}}[f] \in \mathcal{L}^1(\mathbb{R})$ (respectively $\overline{\mathcal{F}}[f] \in l^1(\mathbb{Z})$), then $\left(\mathcal{F} \circ \overline{\mathcal{F}}\right)[f] = f$.*

ii) In particular, if $f \in \mathcal{C}^2(\mathbb{R})$ (respectively $f \in \mathcal{C}^2_{\mathbb{I}}(\mathbb{R})$), then $\left(\mathcal{F} \circ \overline{\mathcal{F}}\right)[f] = f$.

PROOF.– We will give the reasoning in the case of a Fourier series expansion. (i): Since $\overline{\mathcal{F}}[f] \in l^1(\mathbb{Z})$, the series $s = \sum_{k=-\infty}^{+\infty} \mathcal{F}[f]_k . \chi_k$ is normally convergent, and, by the Jordan–Dirichlet theorem, $s(t) = t$ for every $t \in \mathbb{I}$. (ii) We have $f \in W^{2,1}(\mathbb{I})$ and the second derivative is continuous and hence bounded. Therefore, $\overline{\mathcal{F}}[f] = O\left(\frac{1}{|k|^2}\right)$ for $|k| \to \infty$, and $\overline{\mathcal{F}}[f] \in l^1(\mathbb{Z})$. ■

REMARK 6.124.– *Let $f : \mathbb{R} \to \mathbb{C}$ be a T-periodic function ($T > 0$) that is integrable on $[-T/2, T/2]$. Its Fourier coefficients and Fourier series are, respectively, given by:*

$$\overline{\mathcal{F}}[f]_k = \frac{1}{T}\int_{-T/2}^{T/2} f(t) . e^{\frac{2\pi}{T} i.k.t} dt, \quad \sum_{k \in \mathbb{Z}} \overline{\mathcal{F}}[f]_k . e^{-\frac{2\pi}{T} i.k.t}.$$

The reader is invited to generalize these expressions to the case of a T-periodic function on $\mathbb{R}^n$, where $T = (T_1, ..., T_n) \in \left(\mathbb{R}_+^\times\right)^n$.

(II) PERIODIC DISTRIBUTIONS AND DISTRIBUTIONS ON THE TORUS Let $U \in \mathcal{D}'(\mathbb{R}^n)$ and $T = (T_1, ..., T_n) \in \left(\mathbb{R}_+^\times\right)^n$.

DEFINITION 6.125.– *The distribution U is said to be T-periodic if $\boldsymbol{\lambda}(T).U = U$.*

As above in **(I)**, we normalize first to ensure that $T = (1, ..., 1)$; this period is denoted 1, and we say that U is *periodic*. Write $\mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$ for the space of periodic distributions on $\mathbb{R}^n$ and $\mathcal{E}_{\mathbb{I}}(\mathbb{R}^n)$ (respectively $\mathcal{E}(\mathbb{I})$) for the space of functions that are infinitely differentiable and periodic in $\mathbb{R}^n$ (respectively the space of functions that are infinitely differentiable in the unit hypercube $\mathbb{I}$ (see **(I)**)).

It can be shown that $\mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$ is contained in the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions ([VLA 79], Chapter 2, section 7.1). Let $\mathbb{T}^n$ be the torus $(\mathbb{R}/\mathbb{Z})^n$ and

$p:\mathbb{R}^n \to \mathbb{T}^n$ the canonical surjection. Every periodic distribution $U \in \mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$ is the preimage (section 5.2.1(**VIII**)) $p^*\left(\dot{U}\right)$ of a uniquely determined distribution $\dot{U} \in \mathcal{D}'(\mathbb{T}^n)$ on the torus $\mathbb{T}^n$, said to be associated with U. We have $\mathcal{E}'(\mathbb{T}^n) = \mathcal{D}'(\mathbb{T}^n)$, and, since every compactly supported distribution is tempered, $\mathcal{D}'(\mathbb{T}^n)$ is once again the space $\mathcal{S}'(\mathbb{T}^n)$ of tempered distributions on $\mathbb{T}^n$; similarly, $\mathcal{S}(\mathbb{T}^n) = \mathcal{D}(\mathbb{T}^n)$. Hence, the mapping $\dot{U} \mapsto p^*\left(\dot{U}\right)$ is an isomorphism from $\mathcal{S}'(\mathbb{T}^n)$ onto $\mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$, and $(p^*)^{-1} : U \mapsto \dot{U}$ denotes its inverse isomorphism. Likewise, there exists an isomorphism $(p^*)^{-1} : \mathcal{E}_{\mathbb{I}}(\mathbb{R}^n) \to \mathcal{E}(\mathbb{T}^n) : f \mapsto \dot{f}$, where $\dot{f}\left(\dot{t}\right) = f(t)$ for $\dot{t}$ the canonical image of $t \in \mathbb{R}^n$ in $\mathbb{T}^n$.

(III) The spaces $\mathcal{S}(\mathbb{Z}^n)$ and $\mathcal{S}'(\mathbb{Z}^n)$ Write $\mathcal{S}(\mathbb{Z}^n)$ for the space of "rapidly decreasing" complex sequences $a = (a_k)_{k\in\mathbb{Z}^n}$, i.e. such that, for every $j > 0$, $|k|^j |a_k| \to 0$ for $|k| \to \infty$. This space, equipped with the increasing filtrant sequence of seminorms

$$q_m(a) = \sum_{k\in\mathbb{Z}^n} (1+|k|)^m |a_k|,$$

is a nuclear Fréchet space ([TRÈ 67], Chapter 51, Theorem 51.5). In light of Remark 2.40(5), $\mathcal{S}(\mathbb{Z}^n)$ is precisely the space of declining functions on $\mathbb{Z}^n$ ([P2], section 4.3.1(**III**)).

Write $\mathcal{S}'(\mathbb{Z}^n)$ for the space of "slowly increasing" sequences $a = (a_k)_{k\in\mathbb{Z}^n}$, i.e. for which there exists $j > 0$ such that $|a_k| = O\left(|k|^j\right)$. The space $\mathcal{S}'(\mathbb{Z}^n)$ is the dual of $\mathcal{S}(\mathbb{Z}^n)$ (**exercise**); it is therefore the space of *tempered distributions* on $\mathbb{Z}^n$ ([P2], section 4.4.1(**I**)); this is a Silva nuclear space equipped with the strong topology ([P2], section 3.11.3(**II**)). In particular, $\mathcal{S}(\mathbb{Z}^n)$ and $\mathcal{S}'(\mathbb{Z}^n)$ are reflexive. The duality bracket $\langle -,- \rangle : \mathcal{S}'(\mathbb{Z}^n) \times \mathcal{S}(\mathbb{Z}^n) \to \mathbb{C}$ is given by

$$\langle b, a \rangle = \sum_{|k|\leq\infty} b_k.a_k.$$

(IV) Fourier series expansion of a distribution

LEMMA 6.126.– *The series $U = \sum_{k\in\mathbb{Z}^n} c_k.\overline{\chi_k}$ converges in $\mathcal{S}'(\mathbb{R}^n)$ if and only if $c := (c_k)_{k\in\mathbb{Z}^n}$ belongs to $\mathcal{S}'(\mathbb{Z}^n)$, in which case U is a periodic distribution (belonging to $\mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$), which can be uniquely associated with the distribution $\dot{U} = (p^*)^{-1}(U)$ on the torus $\mathbb{T}^n$ given by $\dot{U} = \sum_{k\in\mathbb{Z}} c_k.\overline{\chi_k}$.*

PROOF.– Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then, $\langle \overline{\chi_k}, \varphi \rangle = \mathcal{F}[\varphi](k)$ by [6.9]. We know that $\mathcal{F}[\varphi] \in \mathcal{S}(\mathbb{R}^n)$. The series $\sum c_k.\mathcal{F}[\varphi](k)$ is therefore absolutely convergent if and only if $c \in \mathcal{S}'(\mathbb{Z}^n)$ (**exercise**). The sum U is therefore the tempered distribution $\varphi \longmapsto \sum c_k.\mathcal{F}[\varphi](k)$. We have $\boldsymbol{\lambda}(1).\overline{\chi_k} = \overline{\chi_k}$, so $U \in \mathcal{D}'_{\mathbb{I}}(\mathbb{R}^n)$. ■

THEOREM 6.127.– *1) Let* $\dot{U} \in \mathcal{S}'(\mathbb{T}^n)$ *and*

$$\boxed{\overline{\mathcal{F}}\left[\dot{U}\right]_k = \left\langle \dot{U}, \chi_k \right\rangle.} \qquad [6.13]$$

Then

$$\boxed{\dot{U} = \sum\nolimits_{k \in \mathbb{Z}} \overline{\mathcal{F}}\left[\dot{U}\right]_k . \overline{\chi_k}.} \qquad [6.14]$$

2) The correspondence $\overline{\mathcal{F}} : \dot{U} \mapsto \left(\overline{\mathcal{F}}\left[\dot{U}\right]_k\right)$ *in the expression [6.13] is an isomorphism from* $\mathcal{S}'(\mathbb{T}^n)$ *onto* $\mathcal{S}'(\mathbb{Z}^n)$ *whose inverse isomorphism is*

$$\boxed{\mathcal{F} : (c_k) \mapsto \sum_{k \in \mathbb{Z}} c_k . \overline{\chi_k}.} \qquad [6.15]$$

PROOF.– By Lemma 6.126, if $c \in \mathcal{S}'(\mathbb{Z}^n)$, then $\sum_{k \in \mathbb{Z}} c_k . \overline{\chi_k}$ converges to a distribution $\dot{U} \in \mathcal{S}'(\mathbb{T}^n)$. Conversely, assume that every distribution $\dot{U} \in \mathcal{S}'(\mathbb{T}^n)$ is of this form (see ([SCH 66], Chapter 7, section 1, Theorem I)). Now, let $\mathcal{F}\left[\dot{U}\right]_k$ be given by 6.13, where $\dot{U} = \sum_{k \in \mathbb{Z}} c_k . \overline{\chi_k}$. Then:

$$\mathcal{F}\left[\dot{U}\right]_k = \sum\nolimits_{l \in \mathbb{Z}} c_l \int_{\mathbb{I}} e^{2\pi i (k-l).t} . dt = c_k. \qquad \blacksquare$$

DEFINITION 6.128.– *The expression on the right-hand side of* 6.14*, namely* $\left(\mathcal{F} \circ \overline{\mathcal{F}}\right)\left[\dot{U}\right]$*, is called the* Fourier series expansion *(or the* Fourier series*) of* $\dot{U}$*, and the* $\overline{\mathcal{F}}\left[\dot{U}\right]_k$ *are called the* Fourier coefficients *of* $\dot{U}$*. The correspondence* $\mathcal{F} : \mathcal{S}'(\mathbb{Z}^n) \tilde{\to} \mathcal{S}'(\mathbb{T}^n)$ *is called the* Fourier transform *in* $\mathcal{S}'(\mathbb{Z}^n)$*, and* $\overline{\mathcal{F}} : \mathcal{S}'(\mathbb{T}^n) \tilde{\to} \mathcal{S}'(\mathbb{Z}^n)$ *is called the* Fourier cotransform *in* $\mathcal{S}'(\mathbb{T}^n)$.

EXAMPLE 6.129.– *Consider the* 1*-periodic signal* $U : t \mapsto A.e^{-i2\pi t}$*,* $A \in \mathbb{C}$*. Its Fourier coefficients* c_k $(k \in \mathbb{Z})$ *are given by:*

$$c_k = \int_{-1/2}^{1/2} A.e^{i2\pi(k-1)t}.dt = \begin{cases} A \text{ if } k = 0 \\ 0 \text{ if } k \neq 0 \end{cases}$$

(V) POISSON SUMMATION FORMULA Consider the Dirac comb $\varpi : \mathcal{S}(\mathbb{R}^n) \ni \varphi \mapsto \langle \varpi, \varphi \rangle = \sum_{k \in \mathbb{Z}^n} \varphi(k)$ ([P2], section 4.1.5(**VII**)). This is a periodic distribution, and

$$\mathcal{F}[\varpi](\nu) = \sum_{k \in \mathbb{Z}^n} \langle \delta_k, \overline{\chi_\nu} \rangle = \sum_{k \in \mathbb{Z}^n} e^{-i2\pi\nu.k} = \sum_{k \in \mathbb{Z}^n} \overline{\chi_k}(\nu),$$

which gives the following equality in $\mathcal{S}'(\mathbb{R}^n)$:

$$\boxed{\mathcal{F}[\varpi] = \sum_{k \in \mathbb{Z}^n} \overline{\chi_k}.} \qquad [6.16]$$

Furthermore, $\mathcal{F}[\dot{\varpi}]_k = \langle \dot{\varpi}, \overline{\chi_k} \rangle = \langle \delta, \overline{\chi_k} \rangle = 1$, so

$$\dot{\varpi} = \sum_{k \in \mathbb{Z}^n} 1.\overline{\chi_k} \Longrightarrow \varpi = \sum_{k \in \mathbb{Z}^n} \overline{\chi_k},$$

which gives us:

THEOREM 6.130.– (Poisson summation formula) *We have $\mathcal{F}[\varpi] = \varpi$. Equivalently, for every function $\varphi \in \mathcal{S}(\mathbb{R}^n)$,*

$$\boxed{\sum_{k \in \mathbb{Z}^n} \mathcal{F}[\varphi](k) = \sum_{k \in \mathbb{Z}^n} \varphi(k).}$$

The Poisson summation formula still holds if $\varphi \in W^{1,1}(\mathbb{R})$ ([BOC 59], Chapter 2, section 10, Theorem 10a). Other sufficient conditions for this formula to hold are given by Theorem 6.150 below in a more general setting.

(VI) HILBERTIAN INTERPRETATION If $a = (a_k)_{k \in \mathbb{Z}^n}$, $b = (b_k)_{k \in \mathbb{Z}^n}$ belong to the Hilbert space $l^2(\mathbb{Z}^n)$ of complex square-summable sequences on $\mathbb{Z}^n$, consider their scalar product ([P2], section 3.10.2(**VII**)) $\langle a|b \rangle = \sum_{k \in \mathbb{Z}^n} \overline{a_k}.b_k$. According to Theorem 6.127(2), with $\alpha = \mathcal{F}[a], \beta = \mathcal{F}[b]$, by abuse of notation,

$$\langle a|b \rangle = \sum_{k \in \mathbb{Z}^n} \int_{\mathbb{I} \times \mathbb{I}} \overline{\alpha(t)}.\beta(\tau).e^{2\pi i k.(\tau - t - k)}.dt.d\tau,$$

and, by [6.16], $\sum_{k \in \mathbb{Z}^n} e^{2\pi i k.(\tau - t - k)} = \sum_{k \in \mathbb{Z}^n} \delta(t + k - \tau)$. Hence:

$$\boxed{\langle a|b \rangle = \langle \mathcal{F}[a] | \mathcal{F}[b] \rangle = \int_{\mathbb{I}} \overline{\mathcal{F}[a](t)}.\mathcal{F}[b](t).dt.} \qquad [6.17]$$

The Stone–Weierstrass theorem implies that the family $\{\chi_k\}_{k \in \mathbb{Z}^n}$ is total ([P2], section 3.2.2(**III**)) in $L^2(\mathbb{I})$ ([DIE 93], Volume 1, (7.4.3) and Volume 2, (13.11.6)). It immediately follows that $\langle \chi_k | \chi_l \rangle = 1$ if $k = l$ and 0 otherwise, which gives the following result:

THEOREM 6.131.– *1) The Hilbert space $L^2(\mathbb{I})$ of Lebesgue classes of square-summable functions on $\mathbb{I}$ is separable and admits the Hilbert basis $\{\chi_k\}_{k\in\mathbb{Z}^n}$; furthermore, [6.17] is the* Bessel–Parseval equality *([P2], section 3.10.2(**VI**)).*

2) The Fourier transform $\mathcal{F} : l^2(\mathbb{Z}^n) \to L^2(\mathbb{I})$ is an isomorphism of Hilbert spaces whose inverse isomorphism is $\overline{\mathcal{F}} : L^2(\mathbb{I}) \to l^2(\mathbb{Z}^n)$ (Theorem 6.127(2)).

The reader is invited to formulate an analogous result for Remark 6.116.

(VII) REMARKS ON THE CONVERGENCE AND RECIPROCITY THEOREMS We have seen that the Fourier transform is an isomorphism from $\mathcal{S}(\mathbb{R}^n)$ onto $\mathcal{S}(\mathbb{R}^n)$, from $\mathcal{S}'(\mathbb{R}^n)$ onto $\mathcal{S}'(\mathbb{R}^n)$, from $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$, from $\mathcal{S}'(\mathbb{Z}^n)$ onto $\mathcal{S}'(\mathbb{T}^n)$ and from $l^2(\mathbb{Z}^n)$ onto $L^2(\mathbb{I})$ whose inverse isomorphism is $\overline{\mathcal{F}}$ in each case. We also saw that the Fourier transform of a function $f \in \mathcal{L}^1(\mathbb{R}^n)$ is a uniformly continuous function that is zero at infinity (Riemann–Lebesgue theorem).

Consider again $\mathbb{I} = [-1/2, 1/2]$. In 1926, A. Kolmogorov showed that there exists a function $f \in \mathcal{L}^1(\mathbb{I})$ whose Fourier series diverges at every point ([KOL 77], Chapter 8, section 1). L. Carleson showed in 1966 that, if $f \in \mathcal{L}^2(\mathbb{I})$, then the Fourier series $\left(\mathcal{F} \circ \overline{\mathcal{F}}\right)[f]$ of f converges to f almost everywhere [CAR 66]; this result was extended by R. Hunt in 1967 to any function $f \in \mathcal{L}^p(\mathbb{I})$ for $p \in]1, +\infty[$ (*Carleson–Hunt theorem*) [HUN 67]. Conversely, J.P. Kahane and Y. Katznelson showed in 1966 that, for every set E with zero measure in $\mathbb{I}$, there exists a continuous function f on $\mathbb{I}$ whose Fourier series diverges at every point of E (*Kahane–Katznelson theorem*) [KAH 66].

6.5.4. *Fourier transform on a locally compact commutative group*

(I) Let $\mathbf{G}$ be a locally compact topological group that is countable at infinity, equipped with a Haar measure $\mathfrak{m}_{\mathbf{G}}$ (Remark 6.15 and Convention **(C3)**, section 6.2.2(**II**)). Write $L^p(\mathbf{G})$ for $L^p(\mathbf{G}, \mathfrak{m}_{\mathbf{G}}; \mathbb{C})$ ($p \in [1, \infty]$).

(II) REGULAR REPRESENTATION OF A LOCALLY COMPACT GROUP Let $\mathbf{G}$ be a (not necessarily commutative) locally compact unimodular group (Lemma-Definition 6.11) and $\mathfrak{H}$ a Hilbert space. Write $\mathrm{U}(\mathfrak{H})$ for the group of unitary endomorphisms of $\mathfrak{H}$ ([P2], section 3.10.3(**I**)).

DEFINITION 6.132.– *A* unitary representation *of* $\mathbf{G}$ *in* $\mathfrak{H}$ *is a homomorphism of groups* $U : \mathbf{G} \to U(\mathfrak{H})$ *([P2], section 3.10.3(**I**)) such that, for every* $x \in \mathfrak{H}$*, the mapping* $g \mapsto U(g).x$ *from* $\mathbf{G}$ *into* $\mathfrak{H}$ *is continuous. The (finite or infinite) dimension of* $\mathfrak{H}$ *is said to be the* degree *of* U.

DEFINITION 6.133.– *A unitary representation* $U : \mathbf{G} \to U(\mathfrak{H})$ *is said to be* irreducible *if* $\mathfrak{H}$ *is the only non-trivial closed subspace of* $\mathfrak{H}$ *that is stable under* U.

Two unitary representations $U_1 : \mathbf{G} \to \mathrm{U}(\mathfrak{H}_1)$ and $U_2 : \mathbf{G} \to \mathrm{U}(\mathfrak{H}_2)$ are said to be *equivalent* if there exists an isomorphism of Hilbert spaces $T : \mathfrak{H}_1 \tilde{\to} \mathfrak{H}_2$ such that $U_2(g) = T.U_1(g).T^{-1}$ for every $g \in \mathbf{G}$. The *trivial representation* of $\mathbf{G}$ in $\mathfrak{H}$ is the constant mapping $g \mapsto 1_{\mathfrak{H}}$, and a *scalar representation* is a scalar multiple of the trivial representation, i.e. a representation of the form $t \mapsto \mathbf{u}(g).1_{\mathfrak{H}}$, where $\mathbf{u}(g) \in \mathbb{C}$ and $|\mathbf{u}(g)| = 1$.

DEFINITION 6.134.– *The regular (left) representation of* $\mathbf{G}$ *is the representation* $\boldsymbol{\lambda} : x \mapsto \boldsymbol{\lambda}(x)$ *of* $\mathbf{G}$ *in* $L^2(\mathbf{G})$*, where* $\boldsymbol{\lambda}(x) f : y \mapsto f(x^{-1}.y)$.

The regular right representation $\boldsymbol{\rho}$ is defined in the same way; in the following, every regular representation is a left representation.

Let (Z, π) be a measure space ([P2], section 4.1.1**(II)**) and, for every $z \in Z$, let U_z be a unitary representation of $\mathbf{G}$ in a Hilbert space $\mathfrak{H}_z$. The family $(U_z)_{z \in Z}$ is called a *field of unitary representations*. Suppose that the field $(\mathfrak{H}_z)_{z \in Z}$ is π-measurable, and consider the continuous sum of Hilbert spaces $\mathfrak{H} = \int_Z \mathfrak{H}_z.d\pi(z)$. With the notation from [P2], section 4.1.8, if the functions $z \mapsto \mathfrak{H}_z.\varepsilon_z^i$ are all π-measurable, then there exists a unique unitary representation of $\mathbf{G}$ in $\mathfrak{H}$ such that the relation $g = U.f$ is equivalent to $[g_z = U_z.f_z, \forall z \in Z]$. We say that U is the continuous sum of the unitary representations U_z, writing that

$$U = \int_Z U_z.d\pi(z).$$

The central problem of harmonic analysis on a locally compact group $\mathbf{G}$ *is to study the continuous sum decomposition of the irreducible unitary components of the regular representation of* $\mathbf{G}$.

This allows us to interpret the Fourier transform and Fourier series. Indeed:

1) If $\mathbf{G} = \mathbb{R}^n$ (case of a Fourier transform), then $Z = \mathbb{R}^n$, $\pi = \lambda^{\otimes n}$, $\mathfrak{H}_\nu = \mathbb{C}.\chi_\nu \cong \mathbb{C}$, so $\boldsymbol{\lambda}(\tau).f = \overline{\chi_\nu}(\tau).f$ if $f \in \mathfrak{H}_\nu$. If $f \in \mathcal{L}^2(\mathbb{R}^n)$ satisfies $\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)(f) = f$, then

$$\boldsymbol{\lambda}(\tau) f(t) = \int_{\mathbb{R}^n} \overline{\chi_\nu}(\tau).\underbrace{\chi_\nu(t).\mathcal{F}[f](\nu)}_{\in \mathfrak{H}_\nu}.d\nu,$$

so $\boldsymbol{\lambda} = \int_{\mathbb{R}^n} \overline{\chi_\nu}.d\nu$, where the $\overline{\chi_\nu}$ are unitary irreducible representations of $\mathbb{R}^n$.

2) If $\mathbf{G} = \mathbb{T}^n$ (case of a Fourier series), then $Z = \mathbb{Z}^n$, π is the Haar measure on $\mathbb{Z}^n$ such that $\pi(\{k\}) = 1$ for every $k \in \mathbb{Z}^n$ (Convention **(C3)**, section 6.2.2**(II)**), $\mathfrak{H}_k = \mathbb{C}.\chi_k \cong \mathbb{C}$, so $\boldsymbol{\lambda}(\tau).f = \overline{\chi_k}(\tau).f$ if $f \in \mathfrak{H}_k$. If $f \in \mathcal{L}^2(\mathbb{T}^n)$ is equal to its Fourier series, then

$$\boldsymbol{\lambda}(\tau) f(t) = \sum_{k \in \mathbb{Z}^n} \overline{\chi_k}(\tau).\underbrace{\chi_k(t).\mathcal{F}[f](k)}_{\in \mathfrak{H}_k},$$

so $\boldsymbol{\lambda} = \sum_{k \in \mathbb{Z}^n} \overline{\chi_k}$, where the $\overline{\chi_k}$ are unitary irreducible representations of $\mathbb{T}^n$.

(III) Characters of a locally compact commutative group

DEFINITION 6.135.– *A* unitary character *of a* commutative *locally compact group* $\mathbf{G}$ *is a morphism of groups* $\chi : \mathbf{G} \to \mathbf{U}$*, where* $\mathbf{U}$ *is the multiplicative group of complex numbers with unit modulus.*

In the following, the term *character* always denotes a *unitary character*. The set $\widehat{\mathbf{G}}$ of characters of $\mathbf{G}$ is the commutative group $\mathbf{U}^{\mathbf{G}}$; equipped with the topology of compact convergence, $\widehat{\mathbf{G}}$ is a locally compact group ([BOU 67], Chapter 2, section 1.1, Corollary 2)[14] that is metrizable, separable and countable at infinity if $\mathbf{G}$ is metrizable and separable ([DIE 93], Volume 6, (22.10.3); [BOU 67], Chapter 2, section 2, Exercise 1).

DEFINITION 6.136.– *The locally compact commutative group* $\widehat{\mathbf{G}}$ *is called the* dual group *of* $\mathbf{G}$.

The locally compact group $\widehat{\mathbf{G}}$ can be equipped with a Haar measure $\mathfrak{m}_{\widehat{\mathbf{G}}}$. It is possible to show the following result ([BOU 67], Chapter 2, section 2.1, Corollary 2).

THEOREM 6.137.– *Let* $\mathbf{G}$ *be a locally compact commutative group. The following conditions are equivalent:*

i) $\mathbf{G}$ *and* $\widehat{\mathbf{G}}$ *are generated by compact neighborhoods of their neutral elements (see Lemma 2.80 and the beginning of section 6.2).*

ii) $\mathbf{G}$ *is isomorphic to an* "elementary group"*, i.e. a product* $\mathbb{R}^p \times \mathbb{Z}^q \times \mathbb{T}^r \times \Phi$*, where* Φ *is a finite commutative group (*p, q, r *are necessarily unique and* Φ *is unique up to isomorphism).*

iii) $\widehat{\mathbf{G}}$ *is isomorphic to an elementary group.*

If the group $\mathbf{G}$ above is connected, then it is isomorphic to $\mathbb{R}^p \times \mathbb{T}^r$. If it is simply connected, then it is isomorphic to $\mathbb{R}^p$.

LEMMA 6.138.– *For any character* $\chi \in \widehat{\mathbf{G}}$ *and every* $x \in \mathbf{G}$*, we have* $\chi\left(x^{-1}\right) = \overline{\chi(x)}$.

PROOF.– For every $x \in \mathbf{G}$, we have $1 = \chi(e) = \chi\left(x^{-1}.x\right) = \chi\left(x^{-1}\right).\chi(x)$, so $\chi(x)^{-1} = \overline{\chi(x)}$. ■

14 This group should not be confused with the completion $\hat{\mathbf{G}}$ of $\mathbf{G}$ ([P2], section 2.8.1(**I**)).

In the following, for every $x \in \mathbf{G}$ and every character $\widehat{x} \in \widehat{\mathbf{G}}$, write $\langle \widehat{x}, x \rangle$ for the complex number $\widehat{x}(x)$; we say that $\langle -, - \rangle : \widehat{\mathbf{G}} \times \mathbf{G} \ni (\widehat{x}, x) \mapsto \langle \widehat{x}, x \rangle \in \mathbb{C}^{\times}$ is the *duality bracket*.

(IV) Fourier transform and cotransform

DEFINITION 6.139.– *The* Fourier transform *of a function* $f \in \mathcal{L}^1(\mathbf{G})$ *is*

$$\mathcal{F}[f] : \widehat{\mathbf{G}} \ni \widehat{x} \mapsto \int_{\mathbf{G}} \overline{\langle \widehat{x}, x \rangle}.f(x).d\mathfrak{m}_{\mathbf{G}}(x),$$

and the Fourier cotransform *of* $\varphi \in \mathcal{L}^1\left(\widehat{\mathbf{G}}\right)$ *is*

$$\overline{\mathcal{F}}[\varphi] : \mathbf{G} \ni x \mapsto \int_{\widehat{\mathbf{G}}} \langle \widehat{x}, x \rangle .\varphi(\widehat{x}).d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}).$$

This definition generalizes:

1) Definition 6.97 when $\mathbf{G} = \widehat{\mathbf{G}} = \mathbb{R}^n$ and $\langle \widehat{x}, x \rangle = e^{2\pi i \widehat{x}.x}$;

2) Definition 6.119 when $\mathbf{G} = \mathbb{Z}^n$, $\widehat{\mathbf{G}} = \mathbb{T}^n$ and $\langle \widehat{x}, x \rangle = e^{2\pi i \widehat{x}.x}$.

Theorems 6.100, 6.120, 6.105 and Lemmas 6.98, 6.118 can be (partially) generalized as follows:

THEOREM 6.140.– *1) If* $f, g \in \mathcal{L}^1(\mathbf{G})$*, then*

$$\mathcal{F}[f \star g] = \mathcal{F}[f].\mathcal{F}[g], \qquad [6.18]$$

$$\mathcal{F}[\boldsymbol{\lambda}(x).f] = \overline{\langle \widehat{x}, x \rangle}.\mathcal{F}[f], \quad x \in \mathbf{G}, \qquad [6.19]$$

$$\overline{\mathcal{F}}\left[\check{f}\right] = \mathcal{F}[f]. \qquad [6.20]$$

2) The Fourier transform is injective from $L^1(\mathbf{G})$ *into* $\mathcal{C}_0\left(\widehat{\mathbf{G}}\right)$*, where* $\mathcal{C}_0\left(\widehat{\mathbf{G}}\right)$ *is the algebra of continuous functions that are zero at infinity on* $\widehat{\mathbf{G}}$ *(Definition 6.19).*

Writing $\tilde{f}$ for $\overline{\check{f}} : x \longmapsto \overline{f(x^{-1})}$, [6.20] can be rewritten as:

$$\mathcal{F}\left[\tilde{f}\right] = \overline{\mathcal{F}[f]}. \qquad [6.21]$$

DEFINITION 6.141.– *Let $\mathcal{A}$ be a $\mathbb{C}$-algebra. An* involution *of $\mathcal{A}$ is a bijection $x \mapsto x^*$ from $\mathcal{A}$ onto itself satisfying the following conditions:*

$$(x^*)^* = x, \quad (x+y)^* = x^* + y^*, \quad (\lambda.x)^* = \overline{\lambda}.x^*, \quad (x.y)^* = y^*.x^*.$$

An algebra equipped with an involution is said to be an involutive algebra.

The bijection $f \mapsto \tilde{f}$ from $L^1(\mathbf{G})$ onto $L^1(\mathbf{G})$ is an involution of the convolution algebra $L^1(\mathbf{G})$ (**exercise**).

Plancherel's theorem ([BOU 67], Chapter 2, section 1.3, Theorem 1) shows that, if $f \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$, then $\mathcal{F}[f] \in L^2\left(\widehat{\mathbf{G}}\right)$. The Haar measure $\mathfrak{m}_{\widehat{\mathbf{G}}}$ can be uniquely chosen in such a way that the Fourier transform can be extended to an isometry from $L^2(\mathbf{G})$ onto $L^2\left(\widehat{\mathbf{G}}\right)$ (see Theorem 6.115 and Remark 6.116). If the measure $\mathfrak{m}_{\widehat{\mathbf{G}}}$ is chosen in this way, we then have the following result, analogous to Theorem 6.115 and the Bessel–Parseval equality [6.22]:

THEOREM 6.142.– (Plancherel) *If $f, g \in L^2(\mathbf{G})$, then Parseval's equality holds:*

$$\int_{\mathbf{G}} f(x) . \overline{g(x)} . d\mathfrak{m}_{\mathbf{G}}(x) = \int_{\widehat{\mathbf{G}}} \mathcal{F}[f](\widehat{x}) . \overline{\mathcal{F}[g](\widehat{x})} . d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}). \qquad [6.22]$$

We will see that Plancherel's theorem can be used to determine the inverse Fourier transform.

LEMMA 6.143.– *The vector space $A(\mathbf{G})$ generated by the $f \star \tilde{g}$ $\left(f, g \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})\right)$ is an ideal of the convolution algebra $L^1(\mathbf{G})$ that is dense in $L^2(\mathbf{G})$.*

PROOF.– If $f, g \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$, then $\tilde{g} \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$ and $f \star \tilde{g} \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$ by Theorem 6.20(1). Let (V_k) be a sequence of compact neighborhoods of e in $\mathbf{G}$ such that $V_k \Subset V_{k+1}$, $\bigcap_k V_k = \{e\}$; since the topological space $\mathbf{G}$ is locally compact and countable at infinity, it is paracompact, and hence normal ([P2], section 2.3.11)), and there exists a sequence of continuous real functions (γ_k) such that $\gamma_k(x) \geq 0$, $\operatorname{supp}(\gamma_k) \subset V_k$, $\gamma_k(x) = 1$ for every $x \in V_{k+1}$; let $\varepsilon_k = \gamma_k / \left(\int_{\mathbf{G}} \gamma_k(x) . d\mathfrak{m}_{\mathbf{G}}\right)$. Then, $\int_{\mathbf{G}} \varepsilon_k(x) . d\mathfrak{m}_{\mathbf{G}} = 1$, $\varepsilon_k \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$, $(\varepsilon_k) \to \delta$ in the vague topology ([P2], section 4.1.5(**IV**)) (**exercise**) and, if $h \in L^2(\mathbf{G})$, then $(\varepsilon_k \star h) \to h$ ([DIE 93], Volume 2, (14.11.1)). But $\varepsilon_k \star h \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$, and hence $A(\mathbf{G})$ is dense in $L^2(\mathbf{G})$. ■

DEFINITION 6.144.– *The mapping $\eta : x \mapsto [\langle -, x \rangle : \widehat{x} \mapsto \langle \widehat{x}, x \rangle]$ is called the* canonical mapping *from $\mathbf{G}$ into its bidual $\widehat{\widehat{\mathbf{G}}}$.*

THEOREM 6.145.– *Let $f \in A(\mathbf{G})$; then $f \in \mathcal{C}_0(\mathbf{G})$ (Lemma 6.20(2)), $\mathcal{F}[f] \in L^1\left(\hat{\mathbf{G}}\right)$, and the Fourier inversion formula $f(x) = \overline{\mathcal{F}} \circ \mathcal{F}[f](x), \forall x \in \mathbf{G}$, holds, so*

$$f = \left(\overline{\mathcal{F}}[\mathcal{F}[f]]\right) \circ \eta. \qquad [6.23]$$

This formula also holds for every function $f \in L^2(\mathbf{G})$ such that $\mathcal{F}[f] \in L^1(\mathbf{G})$, in which case, the function f is $\mathfrak{m}_{\mathbf{G}}$-almost everywhere equal to the continuous and bounded function on the right-hand side of [6.23].

PROOF.– Let $f, g \in L^1(\mathbf{G}) \cap L^2(\mathbf{G})$ and suppose that $h = \boldsymbol{\lambda}(x)(f \star \tilde{g})$, so that $h(x) = (f \star \tilde{g})(e)$. By [6.19],

$$\begin{aligned}\mathcal{F}[h](\widehat{x}) = \overline{\langle \widehat{x}, x \rangle}.\mathcal{F}[f \star \tilde{g}](\widehat{x}) \Rightarrow \mathcal{F}[f \star \tilde{g}](\widehat{x}) = \overline{\langle \widehat{x}, x \rangle}^{-1}.\mathcal{F}[h](\widehat{x}) \\ = \langle \widehat{x}, x \rangle .\mathcal{F}[h](\widehat{x}) \Longrightarrow \mathcal{F}[f \star \tilde{g}](\widehat{x}) = \langle \widehat{x}, x \rangle .\mathcal{F}[h](\widehat{x}).\end{aligned}$$

The left-hand side of Parseval's equality [6.22] is equal to $(f \star \tilde{g})(e) = h(x)$. By [6.18] and [6.21],

$$\mathcal{F}[f \star \tilde{g}](\widehat{x}) = \mathcal{F}[f](\widehat{x}).\mathcal{F}[\tilde{g}](\widehat{x}) = \mathcal{F}[f](\widehat{x}).\overline{\mathcal{F}[g](\widehat{x})}.$$

Hence,

$$\begin{aligned}h(x) &= \int_{\widehat{\mathbf{G}}} \mathcal{F}[f](\widehat{x}).\overline{\mathcal{F}[g](\widehat{x})}.d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}) = \int_{\widehat{\mathbf{G}}} \mathcal{F}[f \star \tilde{g}](\widehat{x}).d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}) \\ &= \int_{\widehat{\mathbf{G}}} \langle \widehat{x}, x \rangle .\mathcal{F}[h](\widehat{x}).d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}). \qquad [6.24]\end{aligned}$$

To conclude, it suffices to use the denseness of $A(\mathbf{G})$ in $L^2(\mathbf{G})$ (Lemma 6.143).■

REMARK 6.146.– *In the above, we deduced the Fourier inversion formula [6.23] from Parseval's equality [6.22]. Conversely, starting from [6.23] with $h = f \star \tilde{g}$, we can deduce [6.22] by explicitly finding $h(e)$.*

(V) THE DUALITY THEOREM It is possible to show the following fundamental result ([BOU 67], Chapter 2, section 1.5):

THEOREM 6.147.– (Pontryagin–van Kampen) *Let $\mathfrak{m}_{\widehat{\widehat{\mathbf{G}}}}$ be the measure on $\widehat{\widehat{\mathbf{G}}}$ associated with $\mathfrak{m}_{\widehat{\mathbf{G}}}$. The canonical mapping η from $\mathbf{G}$ into $\widehat{\widehat{\mathbf{G}}}$ (Definition 6.144) is* *an isomorphism of topological groups that transforms* $\mathfrak{m}_{\mathbf{G}}$ *into* $\mathfrak{m}_{\widehat{\widehat{\mathbf{G}}}}$*. After identifying* $\mathbf{G}$ *and* $\widehat{\widehat{\mathbf{G}}}$ *with this isomorphism, the Fourier cotransform from* $L^2\left(\widehat{\mathbf{G}}\right)$ *onto* $L^2(\mathbf{G})$ *and the Fourier transform from* $L^2(\mathbf{G})$ *onto* $L^2\left(\widehat{\mathbf{G}}\right)$ *are inverses of each other.*

Examples

1) The group $\mathbb{R}^n$ is dual to itself under the "duality bracket" $\langle -,-\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C} : (\widehat{x}, x) \mapsto e^{2\pi i \widehat{x}.x}$.

2) The dual of a discrete group is compact and vice versa.

3) In particular, the groups $\mathbb{Z}^n$ and $\mathbb{T}^n$ are dual to one another under the duality bracket $\langle -,-\rangle : \mathbb{Z}^n \times \mathbb{T}^n \to \mathbb{C} : (\widehat{x}, \dot{x}) \mapsto e^{2\pi i \widehat{x}.x}$, where $\dot{x}$ is the canonical image of $x \in \mathbb{R}^n$ in $\mathbb{T}^n$.

4) Any finite commutative group Φ is a finitely generated $\mathbb{Z}$-module that is isomorphic to a direct sum of cyclic groups ([P1], section 3.4.2(**III**)). This group is compact when equipped with the finite topology ([P2], section 1.1.12(**II**)).

If Φ is cyclic of order N, let $g \in \Phi$ be an element of order N and $\zeta_N = e^{2\pi i/N}$; if $\chi \in \widehat{\Phi}$, then $\chi(g) = \zeta_N^j$ for some unique integer j such that $0 \leq j < N$, and $\chi(g^m) = \chi(g)^m$, so χ is determined by the value $\chi(g)$. Conversely, if $0 \leq j < N$, we can define χ by $\chi(g^m) = \zeta_N^{j.m}$, and $\chi(g) = \zeta_N^j$. Hence, $j \mapsto \chi$ is a bijection from $\{0, ..., N-1\}$ onto $\widehat{\Phi}$, a group that must therefore be of order N and isomorphic to Φ. Thus, $\langle \widehat{x}, x\rangle^N = 1$.

In the general case, if $\Phi = \bigoplus_{\lambda\in\Lambda} \Phi_\lambda$, where Λ is a finite set and each group Φ_λ is cyclic, the direct sum (or coproduct) is an inductive limit, so $\mathrm{Hom}_{\mathbb{Z}}(\Phi, \mathbb{C}^\times) \cong \prod_{\lambda\in\Lambda} \mathrm{Hom}_{\mathbb{Z}}\left(\Phi_\lambda, \mathbb{C}^\times\right)$ (where $\mathbb{C}^\times$ is the multiplicative abelian group $\mathrm{GL}_1(\mathbb{C})$ of non-zero elements of $\mathbb{C}$), since the functor $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{C}^\times)$ is left exact ([P1], sections 1.2.8(**I**), (**II**)). Hence, $\widehat{\Phi} \cong \Phi$, again with $\langle \widehat{x}, x\rangle^N = 1$, where N is the order of Φ. The group $\widehat{\Phi}$ is equipped with the finite topology.

Let $f : \Phi \to \mathbb{C}$; then

$$\begin{aligned}\mathcal{F}[f](\widehat{x}) &= \sum_{x\in\Phi} f(x).e^{-\frac{2\pi i}{N}\widehat{x}.x} = \int_\Phi f(x).\overline{\langle \widehat{x}, x\rangle}.d\mathfrak{m}_\Phi(x) \\ &\Longrightarrow \sum_{\widehat{x}\in\widehat{\Phi}} \mathcal{F}[f](\widehat{x}).e^{\frac{2\pi i}{N}\widehat{x}.x} = N.f(x),\end{aligned}$$

where $\mathfrak{m}_{\Phi}(\{x\}) = 1$, since Φ is discrete (Convention **(C3)**, section 6.2.2**(II)**) and hence, given a function $\varphi : \widehat{\Phi} \to \mathbb{C}$,

$$\overline{\mathcal{F}}[\varphi](x) = \int_{\widehat{\Phi}} \langle \widehat{x}, x \rangle . \varphi(\widehat{x}) . d\mathfrak{m}_{\widehat{\Phi}}(\widehat{x}),$$

where $\mathfrak{m}_{\widehat{\Phi}}\left(\widehat{\Phi}\right) = 1$, since $\widehat{\Phi}$ is compact (Convention **(C3)**), so $\mathfrak{m}_{\widehat{\Phi}}(\{\widehat{x}\}) = 1/N$[15]. The characters of the group $\mathbf{U}(\mathbb{Z}/m\mathbb{Z})$ (group of units of $\mathbb{Z}/m\mathbb{Z}$) are called the *Dirichlet characters* modulo m; they play an important role in number theory ([IRE 90], Chapter 16).

Returning to the reasoning from section 6.5.3**(III)**, we see that $\mathcal{S}(\Phi) \cong \mathcal{S}'(\Phi) \cong \mathbb{C}^N$. Since this space is finite-dimensional, it is clearly nuclear.

Let us see how the above results can be interpreted according to **(II)**:

Let $Z = \widehat{\mathbf{G}}$, $\pi = \mathfrak{m}_{\widehat{\mathbf{G}}}$, $\mathfrak{H}_{\widehat{x}} = \mathbb{C}.\langle \widehat{x}, . \rangle$, so $\boldsymbol{\lambda}(x).f = \overline{\langle \widehat{x}, x \rangle}.f$ if $f \in \mathfrak{H}_{\widehat{x}}$. If $f \in \mathcal{L}^2(\mathbf{G})$ satisfies $\left(\overline{\mathcal{F}} \circ \mathcal{F}\right)(f) = f$, then

$$\boxed{\boldsymbol{\lambda}(x).f(y) = \int_{\widehat{\mathbf{G}}} \overline{\langle \widehat{x}, x \rangle}.\underbrace{\langle \widehat{x}, y \rangle . \mathcal{F}[f](\widehat{x})}_{\in \mathfrak{H}_{\widehat{x}}}.d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x}),}$$

so $\boldsymbol{\lambda} = \int_{\widehat{\mathbf{G}}} \overline{\langle \widehat{x}, . \rangle}.d\mathfrak{m}_{\widehat{\mathbf{G}}}(\widehat{x})$, where the $\overline{\langle \widehat{x}, . \rangle}$ are irreducible unitary representations of $\mathbf{G}$.

This establishes the next result, explaining the simplicity of *commutative* harmonic analysis:

COROLLARY 6.148.– *On a locally compact commutative group,* every unitary irreducible component of the regular representation is of degree 1 *(Definition 6.132).*

THEOREM 6.149.– *Let $\mathbf{G}$ be an elementary group $\prod_{1 \le i \le 4} \mathbf{G}_i$, where $\mathbf{G}_1 = \mathbb{R}^p$, $\mathbf{G}_2 = \mathbb{Z}^q$, $\mathbf{G}_3 = \mathbb{T}^r$ and $\mathbf{G}_4 = \Phi$, and let $\widehat{\mathbf{G}} = \prod_{1 \le i \le 4} \widehat{\mathbf{G}_i}$ be its dual group. Setting $\mathcal{S}'(\mathbf{G}) = \widehat{\bigotimes}_{1 \le i \le 4} \mathcal{S}'(\mathbf{G}_i)$, the Fourier transform $\mathcal{F}$ is an isomorphism from $\mathcal{S}'(\mathbf{G})$ onto $\mathcal{S}'\left(\widehat{\mathbf{G}}\right)$ determined by the four isomorphisms $\mathcal{F} : \mathcal{S}'(\mathbf{G}_i) \to \mathcal{S}'\left(\widehat{\mathbf{G}_i}\right)$ presented in sections 6.5.2**(VI)**, 6.5.3**(IV)** and Example 3 above. Its inverse isomorphism is $\overline{\mathcal{F}}$, which is determined by the four isomorphisms $\overline{\mathcal{F}} : \mathcal{S}'\left(\widehat{\mathbf{G}_i}\right) \to \mathcal{S}'(\mathbf{G}_i)$.*

15 There is an ambiguity here: we assumed that Φ is discrete and $\widehat{\Phi}$ is compact, but the reverse is also possible. The key observation is that, for a finite group, duality works according to compact $\longleftrightarrow$ discrete, and never compact $\longleftrightarrow$ compact, nor discrete $\longleftrightarrow$ discrete.

PROOF.– The expression $\mathcal{S}'(\mathbf{G}) = \widehat{\bigotimes_{1\leq i\leq 4}}\mathcal{S}'(\mathbf{G}_i)$ is similar to [5.3] (section 5.2.1(**VIII**)). An arbitrary element of $\bigotimes_{1\leq i\leq 4}\mathcal{S}'(\mathbf{G}_i)$ is of the form $\sum_{j\in J}\bigotimes_{i=1}^{4}T_i^j$, where J is finite and $T_i^j \in \mathcal{S}'(\mathbf{G}_i)$. Thus, $\mathcal{F}\left(\bigotimes_{i=1}^{4}T_i^j\right) = \bigotimes_{i=1}^{4}\mathcal{F}\left(T_i^j\right)$ ("Fubini–Tonelli theorem"), so $\mathcal{F}\left(\sum_{j\in J}\bigotimes_{i=1}^{4}T_i^j\right) = \sum_{j\in J}\bigotimes_{i=1}^{4}\mathcal{F}\left(T_i^j\right)$. Hence, $\mathcal{F}(\mathcal{S}'(\mathbf{G})) = \widehat{\bigotimes_{1\leq i\leq 4}}\mathcal{F}(\mathcal{S}'(\mathbf{G}_i))$. But $\mathcal{F}(\mathcal{S}'(\mathbf{G}_i)) \cong \mathcal{S}'\left(\widehat{\mathbf{G}_i}\right)$, so $\mathcal{F}(\mathcal{S}'(\mathbf{G})) \cong \widehat{\bigotimes_{1\leq i\leq 4}}\mathcal{S}'\left(\widehat{\mathbf{G}_i}\right) = \mathcal{S}'\left(\widehat{\mathbf{G}}\right)$. ■

(VI) GENERALIZED POISSON SUMMATION FORMULA Let $\mathbf{G}$ be a locally compact commutative group and write $\widehat{\mathbf{G}}$ for its dual. We say that two elements $x \in \mathbf{G}$, $\widehat{x} \in \widehat{\mathbf{G}}$ are orthogonal if $\langle \widehat{x}, x\rangle = 1$. If $\mathbf{H}$ is a subgroup of $\mathbf{G}$, then the set $\mathbf{H}^{\perp}$ of elements of $\widehat{\mathbf{G}}$ that are orthogonal to every element of $\mathbf{H}$ is a subgroup of $\widehat{\mathbf{G}}$, and $\left(\mathbf{H}^{\perp}\right)^{\perp} = \overline{\mathbf{H}}$ (closure of $\mathbf{H}$ in $\mathbf{G}$). The orthogonal complement of $\{e\}$ is $\widehat{\mathbf{G}}$ and the orthogonal complement of $\{\widehat{e}\}$ is $\mathbf{G}$ (**exercise**).

Let $\mathbf{H}$ be a closed subgroup of $\mathbf{G}$, write $\mathbf{L} = \mathbf{G}/\mathbf{H}$, suppose that $\mathfrak{m}_{\mathbf{G}}$ and $\mathfrak{m}_{\mathbf{H}}$ are Haar measures on $\mathbf{G}$ and $\mathbf{H}$, respectively, and let $\mathfrak{m}_{\mathbf{L}} = \mathfrak{m}_{\mathbf{G}}/\mathfrak{m}_{\mathbf{H}}$ (Theorem 6.14). Then, $\widehat{\mathbf{G}/\mathbf{H}}$ can be identified with $\mathbf{H}^{\perp}$ and $\mathfrak{m}_{\mathbf{L}}$ can be identified with a Haar measure on $\mathbf{H}^{\perp}$; we also have the following result ([BOU 67], Chapter 2, section 1.8, Corollary):

THEOREM 6.150.– *Let $f \in L^1(\mathbf{G})$. Suppose that (i) the restriction of $\mathcal{F}[f]$ to $\mathbf{H}^{\perp}$ is integrable, (ii) for every $g \in \mathbf{G}$, the function $h \mapsto f(g.h)$ is integrable on $\mathbf{H}$ and (iii) the integral $\int_{\mathbf{H}} f(g.h).d\mathfrak{m}_{\mathbf{H}}(h)$ is a continuous function of g. Then, the* Poisson summation formula *holds:*

$$\boxed{\int_{\mathbf{H}} f(h).d\mathfrak{m}_{\mathbf{H}}(h) = \int_{\mathbf{H}^{\perp}} \mathcal{F}[f](l).d\mathfrak{m}_{\mathbf{L}}(l).}$$

6.5.5. *Overview of non-commutative harmonic analysis*

In this section, every group is a real, unimodular, metrizable, separable, locally compact Lie group.

(I) VON NEUMANN ALGEBRAS Let $\mathfrak{H}$ be a complex Hilbert space and, if $u \in \operatorname{End}(\mathfrak{H})$, write u^* for its adjoint ([P2], section 3.10.3(**I**)). The mapping $u \mapsto u^*$ is an involution of $\operatorname{End}(\mathfrak{H})$ (Definition 6.141). An algebra $\mathcal{A} \subset \operatorname{End}(\mathfrak{H})$ is a *von Neumann algebra* (or $\mathbf{W}^*$-algebra) if it is stable under the involution $u \mapsto u^{**}$ and $\mathcal{A}^{**} = \mathcal{A}$ ([SCH 99], Chapter 6, section 4.3); we know that $\|u\|^2 = \|u.u^*\| = \|u^*.u\|$ ([P2], section 3.10.3(**II**)); we say that $\mathcal{A}$ is a *star* algebra (or $\mathbf{C}^*$-algebra). If $u \in \mathcal{A}$, write $u \geq 0$ if u is positive self-adjoint (*ibid.*), and write $\mathcal{A}_+$ for the set of such endomorphisms. Furthermore, write $v \leq u$ if $u - v \geq 0$; this determines an order relation in $\mathcal{A}_+$.

A *trace* of a von Neumann algebra $\mathcal{A}$ is a mapping $\text{Tr} : \mathcal{A}_+ \to [0, +\infty]$ (with the convention $0.(+\infty) = 0$) such that

$$\text{Tr}\{u+v\} = \text{Tr}\{u\}+\text{Tr}\{v\}, \quad \text{Tr}\{\lambda.u\} = \lambda.\text{Tr}\{u\}, \quad \text{Tr}\{u.u^*\} = \text{Tr}\{u^*.u\}.$$

A trace is said to be *faithful* if $\text{Tr}\{u\} = 0 \Rightarrow u = 0$ $(u \in \mathcal{A}_+)$, *finite* if it only takes finite values in $\mathcal{A}_+$ and semi-finite if, for every $u \in \mathcal{A}_+$,

$$\text{Tr}\{u\} = \sup\{\text{Tr}\{v\} : 0 \le v \le u, \text{Tr}\{v\} < +\infty\}.$$

A von Neumann algebra $\mathcal{A}$ is said to be *finite* (respectively *semi-finite*) if there exists a finite (respectively semi-finite) trace that is not identically zero on $\mathcal{A}_+$. It is said to be a *factor* if its center $\mathfrak{Z}(\mathcal{A})$ reduces to the scalar operations ([SCH 99], Chapter 6, section 8), i.e. $\mathfrak{Z}(\mathcal{A}) = \mathbb{C}.1_{\mathfrak{H}}$. On a factor, the trace, if it exists, is uniquely determined up to multiplication by a real number > 0 ([DIX 81], Chapter 1, section 6, Corollary of Theorem 3).

EXAMPLE 6.151.– *1) In the algebra $\mathfrak{M}_n(\mathbb{C})$, every matrix $v \in \mathfrak{M}_n(\mathbb{C})_+$ is of the form $u^*.u$. The usual trace Tr on $\mathfrak{M}_n(\mathbb{C})$) is faithful (*Schur's theorem*), and $\mathfrak{M}_n(\mathbb{C})$ is a finite factor said to be of type I_n.*

2) If the Hilbert space $\mathfrak{H}$ is separable, we say that $u \in End(\mathfrak{H})$ is a Hilbert–Schmidt operator *if, for a Hilbert basis (a_n), the series with terms $\|u(a_n)\|^2$ converges. If (b_n) is another Hilbert basis, then the series with terms $\|u(b_n)\|^2$ converges. Write $\mathcal{L}_2(\mathfrak{H})$ for the set of Hilbert–Schmidt operators; this is an algebra with the norm*

$$\|u\|_2 := \left(\sum_n \|u(a_n)\|^2\right)^{1/2},$$

and $\mathcal{L}_2(\mathfrak{H})$ is an involutive Banach algebra for this norm. If u, v are Hilbert–Schmidt operators, then $u.v$ is a nuclear *operator ([P2], section 3.11.2(**II**)); conversely, every nuclear operator is the product of two Hilbert–Schmidt operators, and every nuclear operator $T \ge 0$ can be written in the form $u^*.u$, where $u \in \mathcal{L}_2(\mathfrak{H})$. Given such an operator, the series*

$$Tr(T) = \sum_n \langle T.a_n | a_n \rangle$$

converges; it is called the trace of T ([KÖT 79], Volume 2, section 42.6; [DIE 93], Volume 2, section 15.11, Exercise 7). If $\mathfrak{H}$ is infinite-dimensional, then the algebra $\mathcal{L}_2(\mathfrak{H})$, equipped with this trace, is a semi-finite factor, said to be of type I_∞.

Factors of type I_n or I_∞ are said to be of type I. There are other, more complicated factors of type II or III.

(II) Non-commutative groups of type I – Plancherel measure The Fourier cotransform on a locally compact group $\mathbf{G}$ is related to the irreducible components of its regular representation $\boldsymbol{\lambda}$. These components are no longer always of degree 1. It is possible to show the following result ([GOD 15], Chapter 11, sections 6.23 and 8.28):

LEMMA 6.152.– (Schur) *Let U be an irreducible unitary representation of $\mathbf{G}$ in a Hilbert space $\mathfrak{H}$. If $T \in End(\mathfrak{H})$ commutes with every $U(g)$ $(g \in \mathbf{G})$, then T is a scalar multiple of $1_{\mathfrak{H}}$. Let U, U' be two irreducible unitary representations of $\mathbf{G}$ in the Hilbert spaces $\mathfrak{H}, \mathfrak{H}'$. If $T : \mathfrak{H} \to \mathfrak{H}'$ is an operator such that $T.U(g) = V(g).T, \forall g \in \mathbf{G}$, then T is an isomorphism of Hilbert spaces, and U, U' are therefore equivalent.*

A theorem established by Gelfand and Raïkov in 1943 also shows that every locally compact group has a complete system of irreducible unitary representations in the sense that, for every element $g \neq e$, there exist a Hilbert space $\mathfrak{H}$ and an irreducible unitary representation $U : \mathbf{G} \mapsto \mathrm{Aut}(\mathfrak{H})$ such that $U(g) \neq 1_{\mathfrak{H}}$ [HIS 49].

Let $\mathbf{G}$ be a locally compact group, $\mathfrak{H} = L^2(\mathbf{G})$, $\boldsymbol{\lambda}$ the regular representation of $\mathbf{G}$ in $\mathfrak{H}$ and $\mathcal{A}$ the von Neumann algebra generated by the $\boldsymbol{\lambda}(g).\mathfrak{H}$, $g \in \mathbf{G}$. Schur's lemma shows that $\mathcal{A}$ is a factor, and Mautner showed in 1950 that we can decompose $\boldsymbol{\lambda}$ into the continuous sum

$$\boldsymbol{\lambda} = \int_Z U_z.d\pi(z), \quad \mathfrak{H} = \int_Z \mathfrak{H}_z.d\pi(z), \qquad [6.25]$$

where each U_z is an irreducible unitary representation of $\mathbf{G}$ in $\mathfrak{H}_z$ and each $\mathcal{A}_z = \mathrm{End}(\mathfrak{H}_z)$ is a factor [MAU 50]. We have $\boldsymbol{\lambda}(x.y) = \boldsymbol{\lambda}(x).\boldsymbol{\lambda}(y)$, so $U_z(x.y) = U_z(x).U_z(y)$ and $U_z(x^{-1}) = U_z(x)^{-1} = U_z^*(x)$, since $U_z(x)$ is unitary.

DEFINITION 6.153.– *A locally compact unimodular group is said to be of type I (or "tame") if each factor $\mathcal{A}_z$ $(z \in Z)$ is of type I.*

The class of locally compact groups of type I contains the finite groups, the commutative groups, the compact groups, the semi-simple connected groups, the nilpotent connected groups (section 6.4.3(**I**)) and every algebraic linear group (section 2.4.1(**VI**)). However, there also exist "wild" discrete groups and solvable connected groups [KIR 76]; these groups are not considered below.

For a group of type I, each $\mathcal{A}_z$ $(z \in Z)$ is therefore equipped with a trace Tr_z (unique up to multiplication by a real number > 0), and the representations U_z are said to be factorial of type I. Furthermore, the measure space (Z, π) can be identified with the set $\widehat{\mathbf{G}}$ of equivalence classes of the irreducible unitary representations of $\mathbf{G}$; a "natural" topology can be defined on this set, in which case, π is a canonical Borel measure ([P2], section 4.1.1(**II**)), called the *Plancherel measure*. The topological space $\widehat{\mathbf{G}}$ is called the *dual space* of $\mathbf{G}$ (it is the *dual group* if $\mathbf{G}$ is commutative, in which case the Plancherel measure is the Haar measure on $\widehat{\mathbf{G}}$).

(III) FOURIER COTRANSFORM It is relatively straightforward to define the Fourier transform and cotransform abstractly on a tame group; this was accomplished by Segal and Mautner in the early 1950s [SEG 50, MAU 55].

Let $f \in L^2(\mathbf{G}) \cap L^1(\mathbf{G})$; from [6.25], writing $\widehat{x}(x)$ for $U^*_{\widehat{x}}(x)$, we deduce that:

$$\boxed{\overline{\mathcal{F}}[f](\widehat{x}) = \int_{\mathbf{G}} \widehat{x}(x).f(x).d\mathfrak{m}_{\mathbf{G}}(x).}$$

Setting $\mathfrak{F}\left(\widehat{\mathbf{G}}\right) = \prod_{\widehat{x}\in\widehat{\mathbf{G}}} \mathcal{A}_{\widehat{x}}$ leads us to the following definition:

DEFINITION 6.154.– *The* Fourier cotransform *is the operator* $\overline{\mathcal{F}} : L^2(\mathbf{G}) \cap L^1(\mathbf{G}) \to \mathfrak{F}\left(\widehat{\mathbf{G}}\right) : f \mapsto \overline{\mathcal{F}}[f]$.

The Hilbert space $L^2\left(\widehat{\mathbf{G}}\right) = \int_{\widehat{\mathbf{G}}} \mathcal{A}_{\widehat{x}}.d\pi(\widehat{x})$ is the space of π-measurable families $\Phi = \left(\Phi_{\widehat{x}}\right)_{\widehat{x}\in\widehat{\mathbf{G}}}$ such that $\int_{\widehat{\mathbf{G}}} \left\|\Phi_{\widehat{x}}\right\|_2^2.d\pi(\widehat{x}) < \infty$, where $\|.\|_2$ is the Hilbert-Schmidt norm on $\mathcal{A}_{\widehat{x}}$ (Example 6.151(2)) and, if $\mathrm{Tr}_{\widehat{x}}$ is the associated trace (see (**II**)), the scalar product on $L^2\left(\widehat{\mathbf{G}}\right)$ is

$$\langle \Phi, \Psi \rangle_{L^2\left(\widehat{\mathbf{G}}\right)} = \int_{\widehat{\mathbf{G}}} \mathrm{Tr}_{\widehat{x}} \left\{\Phi^*_{\widehat{x}}.\Psi_{\widehat{x}}\right\}.d\pi(\widehat{x}).$$

Parseval's equality establishes that the Fourier cotransform can be extended to an isometry from $L^2(\mathbf{G})$ onto $L^2\left(\widehat{\mathbf{G}}\right)$; this allows us to invert the Fourier cotransform, and hence define the *Fourier transform* by repeating the steps of the proof of Theorem 6.145. Thus:

THEOREM 6.155.– *If* $f \in A(\mathbf{G})$ *(see Lemma 6.143), then* $f \in \mathcal{C}_0(\mathbf{G})$*, and, for every* $x \in \mathbf{G}$*, the Fourier inversion formula (or* Plancherel's formula*)* $f(x) = \left(\mathcal{F}\circ\overline{\mathcal{F}}\right)[f](x)$ *holds, where, for every function* $\varphi \in \mathfrak{F}\left(\widehat{\mathbf{G}}\right)$*, we set:*

$$\boxed{\mathcal{F}[\varphi](x) = \int_{\widehat{\mathbf{G}}} Tr_{\widehat{x}}\left\{\widehat{x}^*(x).\varphi(\widehat{x})\right\}.d\pi(\widehat{x}).} \qquad [6.26]$$

By the denseness of $A(\mathbf{G})$ *in* $L^2(\mathbf{G})$, *the Fourier cotransform can be extended to an isometry from* $L^2(\mathbf{G})$ *onto* $L^2\left(\widehat{\mathbf{G}}\right)$ *("Plancherel's theorem").*

The difficulties (resolved by Harish-Chandra in the case of semi-simple Lie groups) lie in explicitly determining the Plancherel measure.

(IV) CASE OF COMPACT GROUPS When $\mathbf{G}$ is compact, the theory is more straightforward, since the dual space $\widehat{\mathbf{G}}$ is discrete and the irreducible unitary representations $\widehat{x}$ all have finite degree $d(\widehat{x})$, so are unitary matrices in $\mathfrak{M}_{d(\widehat{x})}(\mathbb{C})$ ([DIE 93], Volume 5, (21.2.3))[16]. The measure π is given by $\pi(\widehat{x}) = d(\widehat{x})$, and the relation [6.26] therefore becomes:

$$\mathcal{F}[\varphi](x) = \sum\nolimits_{\widehat{x}\in\widehat{\mathbf{G}}} d(\widehat{x}) . \mathrm{Tr}\{\widehat{x}^*(x) . \varphi(\widehat{x})\},$$

where Tr denotes the usual trace of a matrix. The *character* of $\mathbf{G}$ associated with the irreducible unitary representation $\widehat{x}$ is the continuous function $\chi_{\widehat{x}} = \mathrm{Tr}\{\widehat{x}\} : \mathbf{G} \to \mathbf{U}$. If $f \in A(\mathbf{G})$, then its *Weyl character* is $\Theta_{\widehat{x}}(f) = \mathrm{Tr}\left\{\overline{\mathcal{F}}[f](\widehat{x})\right\}$, which gives us Plancherel's formula for any function $f \in A(\mathbf{G})$:

$$f(x) = \sum\nolimits_{\widehat{x}\in\widehat{\mathbf{G}}} d(\widehat{x}) . \Theta_{\widehat{x}}\left(\boldsymbol{\lambda}\left(x^{-1}\right) . f\right). \tag{6.27}$$

When $\mathbf{G}$ is finite, the above sum is finite.

REMARK 6.156.– *1) In the case of a semi-simple Lie group, we can similarly define the* Harish-Chandra character $\Theta_{\widehat{x}}(f) = Tr\left\{\overline{\mathcal{F}}[f](\widehat{x})\right\}$ *of a function* $f \in \mathcal{D}(\mathbf{G})$, *which gives us another Plancherel formula, generalizing [6.27]:*

$$f(x) = \int_{\widehat{\mathbf{G}}} \Theta_{\widehat{x}}\left(\boldsymbol{\lambda}\left(x^{-1}\right) . f\right) . d\pi(\widehat{x}).$$

2) The Fourier cotransform is "covariant", since the exchange theorem states that $\overline{\mathcal{F}}[f \star g] = \overline{\mathcal{F}}[f] . \overline{\mathcal{F}}[g]$. *By contrast, the Fourier transform is "contravariant", since* $\mathcal{F}[f \star g] = \mathcal{F}[g] . \mathcal{F}[f]$ ***(exercise)****.*

16 See also the Wikipedia article on the *Peter–Weyl theorem*.

7

Connections

I must admit that I found the book[1], like most of Cartan's papers, hard reading. Does the reason lie only in the great French geometric tradition on which Cartan draws, and the style and contents of which he takes more or less for granted as a common ground for all geometers, while we, born and educated in other countries, do not share it?

H. WEYL [WEY 38]

The distinguished services that Ricci and Levi-Cività's absolute differential calculus have provided and will continue to provide must not prevent us from avoiding excessively formal calculations whose superabundance of indices obscures an often very simple geometry reality.

É. CARTAN [CAR 51] (Preface to the first edition)

7.1. Introduction

A connection **C** on a manifold M is a rule that connects the tangent spaces of two infinitely close points. More precisely, according to É. Cartan, an *affinely connected manifold* is a manifold whose tangent spaces are affine spaces (section 3.5.7(**I**)) together with a collection of transformation formulas (similar to those of the affine space itself) that allow us to pass from a coordinate system in the tangent space $T_P(M)$ at a point P of M to a coordinate system in $T_{P+dP}(M)$, where $P+dP$ is infinitely close to P. The term "affine connection" is due to H. Weyl ([WEY 52],

1 The book in question is [CAR 51].

section 12). It was adopted by É. Cartan in an article that generalized the idea, while also introducing the concept of torsion [CAR 25]. If we view $T_P(M)$ as the space of translations of our affine space (section 1.3.1(**I**)), we could instead say that $\mathbf{C}$ is a *linear connection*[2].

The opening quotes of this chapter suggest that there are at least two possible formalisms, and even two possible approaches, to the study of connections. The "superabundance of indices" decried by Cartan can be traced back to Riemann's *Pariserarbeit* (1861) ([SPI 99], Volume 2, Chapter 4, Part C) and a later article by Christoffel (1869) [CHR 69] but perhaps most notably the foundational article of tensor calculus by Ricci and Levi-Cività in 1900 [RIC 00] (mentioned earlier in section 4.1); geometrically, this article was supplemented in 1920 by L. Bianchi's rediscovery of the crucial equalities that bear his name[3]. It was further extended by Levi-Cività in 1917 with the introduction of the notion of parallel transport [LEV 17]; the intrinsic formulation of this idea was established by H. Weyl the next year ([WEY 52], Chapter 2, footnotes 9 and 10). Of the manifolds equipped with an affine connection, the most important to date for physics are the pseudo-Riemannian manifolds equipped with their Levi-Cività connection, which is determined by the metric of the manifold and is torsion-free. In Einstein's theory of general relativity (1915), massive objects acted upon by gravity alone follow a geodesic in space-time, which is a four-dimensional pseudo-Riemannian space deformed by the gravitational field; this field is in fact identified with the spatial curvature.

After Einstein's revolution, which elevated gravity to a distinguished role as the only force to be geometrized, there were numerous attempts to unify gravity and electromagnetism in a way that could reduce both phenomena to spatial geometry. This inspired people to consider connections other than the Levi-Cività connection: as early as 1918, H. Weyl imagined a connection with a "homothety curvature" where the "length" of a connection changes under parallel transport ([WEY 52], section 35). Five-dimensional manifolds with a projective (or quasi-projective) connection were considered by Kaluza and O. Klein (1921, 1926), Veblen (1930) [VEB 33], Pauli (1933), P. Jordan (1947), and many other physicists ([LIC 55], Book 2), ([TON 65], Book 2). Around 1950, Einstein ([EIN 54], Appendix 2), after a wide and diverse range of attempts spanning more than 30 years, like Schrödinger before him, considered a four-dimensional space-time equipped with a linear connection with torsion similar to the one considered by Cartan in 1923–1925 [CAR 25]. While highly interesting mathematically, these theories all failed in physics, partly because the orders of magnitude of the two phenomena differ too

2 Affine connections and linear connections (whose fundamental groups are $\mathrm{A}_n(\mathbb{K})$ and $\mathrm{GL}_n(\mathbb{K})$, respectively) are closely related ([KOB 69], Volume 1, Theorem 3.3).

3 These equalities were originally stated by Ricci in 1889, communicated to E. Padova, published by the latter without proof the same year, then forgotten until rediscovered by Bianchi ([LEV 27], footnote 1, p. 182).

strongly (e.g. for an electron, gravity is much weaker than the electromagnetic force), as well as, and most importantly, because the advent of quantum mechanics revealed that the desired unification was a false problem. The real problem – today more than ever as research into quantum gravity continues – is to unify the theories of general relativity and quantum mechanics, so that there are now four fundamental forces instead of just two (adding the weak and strong interactions to the ranks of gravity and electromagnetism), which requires the smooth space of relativity to be reconciled with the discrete space implied by Heisenberg's quanta and uncertainty principle. Today (as of 2019), we still do not know whether this unification is possible. Pauli himself, toward the end of his life, expressed his full skepticism with an ironic twist on the words of the gospel: "What God has split apart, let no man unite". If any such unification ever should prove possible, it will no doubt heavily draw from the geometric theories described above, which, according to the latest developments in analysis, can all be viewed as variations on the theme of "principal connections" (section 7.3.6) and, more specifically, "Cartan connections" (section 7.3.10).

Both types of connection were originally proposed in the work by É. Cartan in 1922–1926 [CAR 23, CAR 24, CAR 25, CAR 26]; they unify the notion of parallel transport and the idea expressed by F. Klein in his *Erlangen program*[4] [SHA 97]. They offer a very general framework that includes linear connections (section 7.2.3), Euclidean connections such as the Levi-Cività connection (section 7.4.1), projective connections (section 7.3.10(**II**)), conformal connections and others as special cases. A precise formulation was proposed by Ehresmann in 1950 [EHR 51], later extended by a wide range of mathematicians, including Chern (1951–1953), Ambrose-Singer (1953) [AMB 53], Lichnerowicz (1955), then Kobayashi-Nomizu in 1956–1963 ([NOM 56], Chapter 2), [KOB 57], ([KOB 69], Volume 1, Chapter 2). Going any further would require us to explore non-commutative geometry, which is beyond the scope of this book [CON 95].

Throughout this chapter, every manifold is real, of class C^∞, pure and finite-dimensional; every vector bundle is of finite rank and class C^∞; every vector field, tensor field, differential p-form, etc., is of class C^∞. The reader is invited to refine these hypotheses in each section.

7.2. Linear connections

7.2.1. *Curvilinear coordinates*

(I) NOTION OF CURVILINEAR COORDINATES In the affine space $\mathbb{A}^n_{\mathbb{R}}$ ([P1], section 3.2.7(**I**); section 1.3.1(**I**)), let $(O;\mathbf{f})$, $\mathbf{f} = (\mathbf{f}_1, ..., \mathbf{f}_n)$ be an affine frame and consider n functions $y^1, ..., y^n$ of the coordinates $\mathbf{b} = \left(b^1, ..., b^n\right)$, each of class C^2, such that

4 See the Wikipedia article on the *Erlangen program*.

Jacobian $\frac{\partial\left(y^1,\ldots y^n\right)}{\partial\left(b^1,\ldots,b^n\right)}$ does not vanish on any connected non-empty open subset Ω of $\mathbb{R}^n$ (section 1.2.2(**IV**)). Then, $\mathbf{y} = \left(y^1, \ldots y^n\right)$ is said to be a curvilinear coordinate system (or a Gaussian coordinate system) in the domain $B = \left\{A : \overrightarrow{OA} \in \Omega\right\} \subset \mathbb{R}^n$. Consider a curve $a : t \mapsto a\left(t\right)$ of class C^2 in B and set $A\left(t\right) = \overrightarrow{a\left(t_0\right) a\left(t\right)}$.

(II) DIFFERENTIATION IN AN AFFINE FRAME In the frame $\left(a\left(t_0\right), \mathbf{f}\right)$, we have $A\left(t\right) = \sum_{1\leq i\leq n} a^i.\mathbf{f}_i$, $\mathbf{f}_i = \partial/\partial b^i$ (section 2.2.4(**III**)), so

$$A\left(t\right) = \sum_{1\leq i\leq n} a^i.\frac{\partial}{\partial b^i}. \qquad [7.1]$$

Since the basis vectors $\partial/\partial b^i$ are constant, this implies that:

$$\frac{dA}{dt}\left|_{t_0}\right. = \lim_{\Delta t\to 0}\frac{A\left(t_0+\Delta t\right)-A\left(t_0\right)}{\Delta t} = \sum_{1\leq i\leq n}\frac{da^i}{dt}.\frac{\partial}{\partial b^i}. \qquad [7.2]$$

In particular, in B, consider the j-th coordinate curve $b^j = t$, $b^i = const.$ $\left(i \neq j\right)$. Then:

$$\frac{dA}{dt} = \frac{\partial A}{\partial b^j}. \qquad [7.3]$$

(III) DIFFERENTIATION IN CURVILINEAR COORDINATES. COVARIANT DERIVATIVE Suppose now that the a^i are the coordinates of the vector field A in the coordinate system $\mathbf{y} = \left(y^1, \ldots, y^n\right)$. This gives us a similar relation to [7.1], namely $A\left(t\right) = \sum_{1\leq j\leq n} a^j.\frac{\partial}{\partial y^j}$, where the tangent vectors $\mathbf{e}_j = \partial/\partial y^j$ vary as a function of t. Since $\frac{d}{dt}\left(\frac{\partial}{\partial y^j}\right)$ is a tangent vector, there exists a matrix $\left(M_j^k\right) \in \mathrm{GL}_n\left(\mathbb{R}\right)$ such that

$$\frac{d}{dt}\left(\frac{\partial}{\partial y^j}\right) = \sum_{1\leq k\leq n} M_j^k.\frac{\partial}{\partial y^k}. \qquad [7.4]$$

Hence, instead of [7.2], we have:

$$\frac{dA}{dt} = \sum_{1\leq k\leq n}\left(\frac{da^k}{dt} + \sum_{1\leq j\leq n} a^j.M_j^k\right).\frac{\partial}{\partial y^k}.$$

By [7.4],

$$\sum_{1\le k\le n} M_j^k.\underbrace{\frac{\partial}{\partial y^k}}_{\mathbf{e}_k} = \sum_l \frac{d}{dt}\left(\frac{\partial b^l}{\partial y^j}\right).\frac{\partial}{\partial b^l} = \sum_k \underbrace{\sum_{i,l}\left(\frac{\partial^2 b^l}{\partial y^i \partial y^j}.\frac{dy^i}{dt}\right).\frac{\partial y^k}{\partial b^l}}_{M_j^k}.\underbrace{\frac{\partial}{\partial y^k}}_{\mathbf{e}_k},$$

which gives $M_j^k = \sum_i \Gamma_{ij}^k.\frac{dy^i}{dt}$, where

$$\Gamma_{ij}^k = \sum_l \frac{\partial^2 b^l}{\partial y^i \partial y^j}.\frac{\partial y^k}{\partial b^l}. \qquad [7.5]$$

The expression [7.4] can be written in the form

$$d\mathbf{e}_i = \sum_k \underbrace{\sum_i \Gamma_{ij}^k.dy^i}_{\omega_j^k}.\mathbf{e}_k, \qquad [7.6]$$

where

$$\omega_j^k = \sum_i \Gamma_{ij}^k.dy^i. \qquad [7.7]$$

Finally:

$$\frac{dA}{dt} = \sum_k \left(\frac{da^k}{dt} + \sum_{i,j}\Gamma_{ij}^k.a^j.\frac{dy^i}{dt}\right).\mathbf{e}_k. \qquad [7.8]$$

DEFINITION 7.1.– Γ_{ij}^k, *also denoted* $\left\{ \begin{matrix} k \\ ij \end{matrix} \right\}$, *is called the* Christoffel symbol of the second kind.

REMARK 7.2.– *1)* $\left\{ \begin{matrix} k \\ ij \end{matrix} \right\}$ *is symmetric in* i, j : $\left\{ \begin{matrix} k \\ ij \end{matrix} \right\} = \left\{ \begin{matrix} k \\ ji \end{matrix} \right\}$.

2) $\left\{ \begin{matrix} k \\ ij \end{matrix} \right\}$ *can also be defined in the case of a pseudo-Riemannian manifold equipped with the Levi-Cività connection (section 7.4.1). The symmetry in* i, j *still holds in this context.*

3) Γ_{ij}^k *can also be defined on a fiber bundle equipped with a linear connection (section 7.2.2); this term, which is no longer symmetric in* i, j *in general, is then said to be the* coefficient of the connection. *It should not be written as* $\left\{ \begin{matrix} k \\ ij \end{matrix} \right\}$.

As before, consider the i-th (curvilinear) coordinate $y^i = t$, $y^j = const.$ $(j \neq i)$, and write:

$$\frac{dA}{dt} = \nabla_i A.$$

This expression replaces [7.3], leading us to define the *covariant derivative*[5] as follows:

$$\nabla_i A = \sum_k \left(\frac{\partial a^k}{\partial y^i} + \sum_j \Gamma^k_{ij}.a^j \right).\mathbf{e}_k,$$

where the k-th component is

$$(\nabla_i A)^k = \frac{\partial a^k}{\partial y^i} + \sum_j \Gamma^k_{ij}.a^j. \qquad [7.9]$$

Let $Z = \sum_i z^i.\frac{\partial}{\partial y^i}$ be a vector field. The derivative of A in the direction of Z (section 5.17(**I**)) is given by:

$$\nabla_Z A = \sum_i \nabla_i A.z^i = \sum_{i,j,k} \left(\frac{\partial a^k}{\partial y^i} + \sum_j \Gamma^k_{ij}.a^j \right).z^i.\mathbf{e}_k, \qquad [7.10]$$

where the k-th component is

$$\boxed{(\nabla_Z A)^k = \sum_i \left(\frac{\partial a^k}{\partial y^i} + \sum_j \Gamma^k_{ij}.a^j \right).z^i.} \qquad [7.11]$$

Here, A and Z are vector fields (elements of $\mathcal{T}^1_0(B)$). Later, we will see (Definition 7.7) that $\nabla_Z A \in \mathcal{T}^1_0(T(B))$; in our case, $T(B)$ can be identified with $\mathbb{R}^n$.

(IV) CASE OF EUCLIDEAN SPACE. CHRISTOFFEL SYMBOLS OF THE FIRST KIND In Euclidean space, there exists a scalar product $\langle - | - \rangle$. As we did in section 4.5.1, set

$$\boxed{g_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle = \mathbf{e}_i.\mathbf{e}_j,} \qquad [7.12]$$

5 Especially in physics, the term "covariant" has two entirely distinct meanings: one which refers to the *covariant components* of a tensor, as opposed to its contravariant components, and another meaning that is synonymous with *intrinsic*, or in other words *invariant under change of coordinates*. For example, Maxwell's equations are covariant relative to the Lorentz group but not the Galilean group.

and suppose that the frame $(O;\mathbf{f})$ is orthonormal, i.e. $\langle \mathbf{f}_k|\mathbf{f}_l\rangle = \delta_k^k$. Then:

$$g_{ij} = \left\langle \sum_k \frac{\partial b^k}{\partial y^i}.\frac{\partial}{\partial b^k} | \sum_l \frac{\partial b^l}{\partial y^j}.\frac{\partial}{\partial b^l} \right\rangle = \sum_k \left\langle \frac{\partial b^k}{\partial y^i} | \frac{\partial b^k}{\partial y^j} \right\rangle.$$

Furthermore,

$$\begin{aligned} dg_{ij} &= \mathbf{e}_i.d\mathbf{e}_j + d\mathbf{e}_i.\mathbf{e}_j \underset{(7.6)}{=} \mathbf{e}_i.\sum_k \omega_j^k.\mathbf{e}_k + \sum_{k'} \omega_i^{k'}.\mathbf{e}_{k'}.\mathbf{e}_j \\ &= \underbrace{\sum_k g_{ik}.\omega_j^k}_{\omega_{ij}} + \underbrace{\sum_k g_{kj}.\omega_i^k}_{\omega_{ji}}. \end{aligned} \qquad [7.13]$$

Explicitly expanding ω_{ij} gives:

$$\omega_{ij} = \sum_k g_{ik}.\underbrace{\sum_l \Gamma_{lj}^k.dy^l}_{\omega_j^k} = \sum_l \underbrace{\sum_k g_{ik}.\Gamma_{lj}^k}_{\Gamma_{lj,i}}.dy^l.$$

DEFINITION 7.3.– $\Gamma_{pq,j} = \sum_i g_{ij}.\Gamma_{pq}^i$ *is called the* Christoffel symbol of the first kind *(also written as* $[pq,j]$*).*

REMARK 7.4.– $\Gamma_{pq,j}$ *is symmetric in* p, q*.*

Let g^{ij} be the component of the inverse of the matrix formed by the g_{ij} :

$$\boxed{\left(g^{ij}\right) = \left(g_{ij}\right)^{-1}.}$$

The dual basis of $(\mathbf{e}_i)$ consists of the $\mathbf{e}^j$ $(j = 1, ..., n)$ such that $\left\langle \mathbf{e}^j, \mathbf{e}_i \right\rangle = \delta_i^j$, so:

$$\mathbf{e}^j = \sum_k g^{jk}.\mathbf{e}_k. \qquad [7.14]$$

Thus, we indeed have:

$$\left\langle \mathbf{e}^j, \mathbf{e}_i \right\rangle = \sum_k g^{jk}.\mathbf{e}_k.\mathbf{e}_i = \sum_k g^{jk}.g_{kj} = \delta_i^j.$$

The next result summarizes the above calculations, with $g = \det\left(g_{ij}\right)$.

THEOREM 7.5.– *The following relations are satisfied:*

$$[pq,j]=\sum_i g_{ij}.\left\{\begin{matrix} i \\ pq \end{matrix}\right\}, \quad \left\{\begin{matrix} i \\ pq \end{matrix}\right\}=\sum_j g^{ij}.\,[pq,j] \tag{7.15}$$

$$\frac{\partial \mathbf{e}_i}{\partial y^j}=\sum_k \left\{\begin{matrix} k \\ ij \end{matrix}\right\}.\mathbf{e}_k=\sum_k [ij,k]\,.\mathbf{e}^k \tag{7.16}$$

$$[pq,i]=\frac{1}{2}\left(\frac{\partial g_{pi}}{\partial y^q}+\frac{\partial g_{qi}}{\partial y^p}-\frac{\partial g_{pq}}{\partial y^i}\right) \tag{7.17}$$

$$\left\{\begin{matrix} i \\ pq \end{matrix}\right\}=\frac{1}{2}\sum_j g^{ij}.\left(\frac{\partial g_{pj}}{\partial y^q}+\frac{\partial g_{qj}}{\partial y^p}-\frac{\partial g_{pq}}{\partial y^j}\right) \tag{7.18}$$

$$\sum_i \left\{\begin{matrix} i \\ ih \end{matrix}\right\}=\frac{\partial\left(\ln\left(\sqrt{|g|}\right)\right)}{\partial y^h} \tag{7.19}$$

PROOF.– We have already shown the relations in [7.15], as well as the first equality in [7.16]. Next, observe that $\sum_k \left\{\begin{matrix} k \\ ij \end{matrix}\right\}.\mathbf{e}_k = \sum_k\sum_l g^{kl}.\,[ij,l]\,.\mathbf{e}_k = \sum_l [ij,l].\sum_k g^{kl}.\mathbf{e}_k = \sum_l [ij,l]\,.\mathbf{e}^l$, which shows the second equality of [7.16] by [7.14]. By [7.13],

$$dg_{ij}=\omega_{ij}+\omega_{ji}=\sum_q\left([jq,i]+[iq,j]\right).dy^q \Rightarrow [jq,i]+[iq,j]=\frac{\partial g_{ij}}{\partial y^q}.$$

Applying a circular permutation to the three indices and combining the relations thus obtained give us [7.17]. The equality [7.18] immediately follows by the second equality of [7.15]. It also implies that $\sum_i \left\{\begin{matrix} i \\ ih \end{matrix}\right\} = \frac{1}{2}\sum_{i,j} g^{ij}\partial_h g_{ij}$ and $\partial_h g = \sum_{i,j}\alpha^{ij}\partial_h g_{ij}$, where α^{ij} is the cofactor of indices (i,j) of the matrix (g_{ij}) ([P1], section 2.3.11(**II**)). This is a symmetric matrix, so $g^{ij}=\alpha^{ij}/g$, and hence:

$$\sum_i \left\{\begin{matrix} i \\ ih \end{matrix}\right\}=\frac{1}{2}\frac{\partial_h g}{g}=\partial_h\left(\ln\left(\sqrt{|g|}\right)\right). \qquad \blacksquare$$

REMARK 7.6.– *1) The components a^k of A in the coordinate system* $\mathbf{y}$ *appear on the right-hand side of [7.8], but the left-hand side of this equality is intrinsic (independent of the choice of coordinate system). Therefore, the right-hand side is also intrinsic, and so is the covariant derivative* $(\nabla_i A)^k=\partial_i a^k+\sum_j \Gamma^k_{ij}.a^j$, *unlike*

$\partial_i a^k$ (where $\partial_i = \partial/\partial y^i$). In physics, the quantities $(\nabla_i A)^k$ and $\partial_i a^k$ are often written as $a^k_{:i}$ (or $a^k_{;i}$) and $a^k_{,i}$, respectively (we will not use this notation). The connection is characterized by the bilinear operator $\nabla : T(B) \times T(B) \to T(T(B)) : (\mathbf{Z}, \mathbf{A}) \mapsto \nabla_{\mathbf{Z}}\mathbf{A}$. The mapping $\mathbf{Z} \mapsto \nabla_{\mathbf{Z}}\mathbf{A}$ is a tensor of type $(1,1)$. The Γ^k_{ij} are again called the coefficients of the connection*; they depend on the choice of coordinate system, because, in the affine coordinates $\mathbf{b}$, they are all zero. Thus, they are not the components of a tensor field.*

2) Unlike the Γ^k_{ij}, the Christoffel symbols of the first kind $\Gamma_{pq,j} = [pq, j]$ are defined using the scalar product. Therefore, they only have meaning in the Euclidean space, not the affine space. Later, we will see that they can also be defined in a pseudo-Riemannian manifold (section 4.5), for analogous reasons.

7.2.2. *Linear connection on a vector bundle*

(I) Notion of a linear connection Let B be a manifold, E a vector bundle with base B and $\mathbf{A} \in \Gamma(B; E)$ (section 3.3.4, Corollary-Definition 3.21).

DEFINITION 7.7.– *Specifying a* linear connection $\mathbf{C}$ *on the vector bundle E means specifying a* bilinear mapping

$$\nabla : T(B) \times E \to T(E) : (\mathbf{h}, \mathbf{a}) \mapsto \nabla_{\mathbf{h}}.\mathbf{a}$$

such that, for every function $\sigma \in \mathcal{T}^0_0(B) = \mathcal{E}(B)$, every vector field $Z \in \mathcal{T}^1_0(M)$ and every section $\mathbf{A} \in \Gamma(B; E)$:

$$\nabla_{\sigma.Z}.\mathbf{A} = \sigma(\nabla_Z.\mathbf{A}), \qquad [7.20]$$

$$\nabla_Z.(\sigma.\mathbf{A}) = (\mathcal{L}_Z.\sigma)\mathbf{A} + \sigma(\nabla_Z.\mathbf{A}), \qquad [7.21]$$

*where $\mathcal{L}_Z$ denotes the Lie derivative along Z (section 5.17**(I)**).*

The operator ∇ is the covariant derivative determined by the linear connection $\mathbf{C}$. Conversely, ∇ determines $\mathbf{C}$[6]. The relation [7.20] expresses the fact that $(\nabla_Z.\mathbf{A})_b$ only depends on Z_b (and not on Z in a whole neighborhood of b). The relation [7.21] can be interpreted as the *Leibniz rule*. The element $\nabla_Z.\mathbf{A} \in \mathcal{T}^1_0(T(E))$ is said to be the covariant derivative of $\mathbf{A}$ along the vector field Z.

Using suitable charts, we can reduce to the case where B is an open subset of $\mathbb{R}^n$ and $E = B \times \mathbb{R}^p$; $\mathbf{A}$ is of the form $B \ni b \mapsto (b, \mathbf{a}(b)) \in B \times \mathbb{R}^p$. Then, with $\mathbf{h}_b = (b, \mathbf{h})$, we set:

$$\nabla_{\mathbf{h}_b}.\mathbf{A} = (b, D\mathbf{a}(b).\mathbf{h} + \Gamma(b)(\mathbf{h}, \mathbf{a}(b))),$$

6 Readers should take care to distinguish between the covariant differentiation operator ∇ and the gradient operator $\overrightarrow{\nabla}$ from section 5.5.2.

where $\Gamma(b) : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$ is bilinear and $\Gamma : b \mapsto \Gamma(b)$ is of class C^∞. This expression can be written explicitly as follows:

$$(\nabla_{\mathbf{h}_b}.\mathbf{A})^k = \sum_i \left(\partial_i a^k(b) + \sum_j \Gamma^k_{ij}(b).a^j(b) \right).\mathbf{h}^i,$$

where $\partial_i = \partial/\partial b^i$. Note that the k, j range from 1 to $\dim(E_b)$, and i ranges from 1 to $\dim(B)$. Replacing $\mathbf{h}_b$ by a vector field $Z \in \mathcal{T}^1_0(B)$, we also obtain an expression similar to [7.11]:

$$(\nabla_Z.\mathbf{A})^k = \sum_i \left(\partial_i a^k + \sum_j \Gamma^k_{ij}.a^j \right).z^i. \qquad [7.22]$$

The Γ^k_{ij} are called the *coefficients of the linear connection* $\mathbf{C}$ on the vector bundle $\pi : E \to B$ relative to a given chart of B (they are no longer called Christoffel symbols of the second kind; in this more general context, there is no meaningful equivalent for the Christoffel symbols of the first kind). Note that Γ^k_{ij} is not symmetric in i, j in general.

COROLLARY 7.8.– *1) When $Z = X_i$, the k-th component of the covariant derivative $\nabla_i\mathbf{A} = \nabla_{X_i}.\mathbf{A}$ satisfies the expression [7.9].*

2) If additionally $E = T(B)$, then taking $\mathbf{A} = X_j$ implies that $a^k = \delta^k_j$, so $(\nabla_{X_i}.X_j)^k = \Gamma^k_{ij}.\delta^k_j$, and hence $\nabla_{X_i}.X_j = \sum_k \Gamma^k_{ij}.X_k$. In other words, with $X_i = \frac{\partial}{\partial \xi^i}$:

$$\nabla_{\frac{\partial}{\partial \xi^i}}.\frac{\partial}{\partial \xi^j} = \sum_k \Gamma^k_{ij}.\frac{\partial}{\partial \xi^k}. \qquad [7.23]$$

EXAMPLE 7.9.– *The sections of class C^∞ of the trivial bundle $B \times \mathbb{R}^p$ are the mappings $\mathbf{A} : B \to \mathbb{R}^p$ of class C^∞. For every $b \in B : d_b\mathbf{A} : T_b(B) \to \mathbb{R}^p$, $X_b \in T_b(B)$, $d\mathbf{A}.X = \nabla_X\mathbf{A} : b \mapsto d_b\mathbf{A}.X_b : B \to \mathbb{R}^p$. The connection of this bundle is $\nabla_0 : (X, \mathbf{A}) \mapsto d\mathbf{A}.X$, called the* trivial connection.

(II) PARAMETRIZATION OF LINEAR CONNECTIONS ON A VECTOR BUNDLE It can be shown that any given vector bundle has a linear connection ([DIE 93], Volume 3, (17.16.8)). Below, we will see that this linear connection is not unique; however, the linear connections can easily be parametrized.

Let $\mathbf{A} \in \Gamma(B; E)$, where B is pure and finite-dimensional. Then, ([P1], section 3.1.5**(I)**),

$$\mathrm{Hom}_{\mathbb{R}}\left(\Gamma\left(B;T\left(B\right)\right),\Gamma\left(B,E\right)\right) \overset{\text{identification}}{=} \Gamma\left(B;T^{\vee}\left(B\right)\otimes E\right),$$

so the mapping $\nabla.\mathbf{A} : Z \mapsto \nabla_Z.\mathbf{A}$ belongs to $\Gamma\left(B;T^{\vee}\left(B\right)\otimes E\right)$.

COROLLARY 7.10.– *Any linear connection on the vector bundle $\pi : E \to B$ is an $\mathbb{R}$-linear operator $\nabla : \Gamma\left(B,E\right) \to \Gamma\left(B,T^{\vee}\left(B\right)\otimes E\right)$ satisfying the Leibniz rule* $\nabla\left(\sigma\mathbf{A}\right) = d\sigma \otimes \mathbf{A} + \sigma.\nabla\mathbf{A}$.

From this result, it is easy to deduce a parametrization of the set of all linear connections on a vector bundle. Indeed, suppose that $\overset{1}{\nabla}, \overset{2}{\nabla}$ are two connections on the same vector bundle $\pi : E \twoheadrightarrow B$. Let $\alpha = \overset{2}{\nabla} - \overset{1}{\nabla}$. By the Leibniz rule, it immediately follows that $\alpha\left(\sigma.\mathbf{A}\right) = \sigma.\alpha\left(\mathbf{A}\right)$, which gives the following result:

COROLLARY 7.11.– *1) If $\overset{1}{\nabla}, \overset{2}{\nabla}$ are two linear connections on the same vector bundle $\pi : E \to B$, then $\alpha = \overset{2}{\nabla} - \overset{1}{\nabla} : \Gamma\left(B,E\right) \to \Gamma\left(B,T^{\vee}\left(B\right)\otimes E\right)$ is an $\mathcal{E}\left(B\right)$-linear operator.*

2) Conversely, let $\pi : E \twoheadrightarrow B$ be a vector bundle and $\overset{1}{\nabla}$ a linear connection on E. Then, the set of all linear connections on E is the set of $\overset{1}{\nabla} + \alpha$, where α ranges over the space of all $\mathcal{E}\left(B\right)$-linear operators from $\Gamma\left(B,E\right)$ into $\Gamma\left(B,T^{\vee}\left(B\right)\otimes E\right)$.

In terms of coordinates, as before in [7.22], we have:

$$\left(\overset{2}{\nabla}_Z.\mathbf{A}\right)^k - \left(\overset{1}{\nabla}_Z.\mathbf{A}\right)^k = \sum_{i,j}\left(\overset{2}{\Gamma}{}^k_{ij} - \overset{1}{\Gamma}{}^k_{ij}\right).a^j.z^i.$$

7.2.3. *Linear connection on a manifold*

(I) COVARIANT DERIVATIVE OF A VECTOR FIELD Consider the situation where $E = T\left(B\right)$. A *linear connection on B* is a linear connection on the tangent bundle $T\left(B\right)$, i.e. a B-morphism (section 3.3.1, Lemma-Definition 3.7) from $T\left(B\right) \times T\left(B\right)$ into $T\left(T\left(B\right)\right)$ satisfying the conditions [7.20, 7.21].

The connection, or any related tensors, are said to be expressed "relative to the *natural frames*" if, for every chart (U, ξ, n), we choose the frame consisting of the n vector fields $\frac{\partial}{\partial\xi^1}, ..., \frac{\partial}{\partial\xi^n}$, the coframe consisting of the n differential forms $d\xi^1, ..., d\xi^n$ with local coordinates $\xi^1, ..., \xi^n$ (section 2.2.4, Definition 2.27), and if we characterize the connection or these tensors by their coordinates in these bases.

Let $X, Y \in \mathcal{T}_0^1(U)$; the $(\nabla_X.Y)^i$ are the components of $\nabla_X.Y$ in the basis $\left(\frac{\partial}{\partial \xi^i}\right)$ of $T(U)$, so, by [7.24],

$$\nabla_X.Y = \sum_{i,j} \left(\partial_j y^i + \sum_k \Gamma^i_{jk}.y^k\right).X^j \otimes \frac{\partial}{\partial \xi^i},$$

and hence

$$\nabla Y = \sum_{i,j} \left(\partial_j y^i + \sum_k \Gamma^i_{jk}.y^k\right).d\xi^j \otimes \frac{\partial}{\partial \xi^i}, \qquad [7.24]$$

or in other words

$$\boxed{(\nabla_i y)^k = \partial_i y^k + \sum_j \Gamma^k_{ij}.y^j.} \qquad [7.25]$$

By Theorem-Definition 5.42, div $(Y) = \sum_i (\nabla_i y)^i$.

(II) COVARIANT DERIVATIVE OF A COVECTOR FIELD To define and calculate the covariant derivative of a covector field $\alpha \in \mathcal{T}_1^0(B)$, it suffices to note that, for any vector field $Y \in \mathcal{T}_0^1(B)$, the product $\langle \alpha, Y \rangle = \alpha.Y$ is an element of $\mathcal{E}(B)$, so $\partial_j(\alpha.Y)$ can be written in two equivalent ways:

$$\partial_i(\alpha.Y) = \sum_j \left((\partial_i \alpha_j).y^j + \alpha_j.\partial_i y^j\right),$$

$$\partial_i(\alpha.Y) = \sum_j \left((\nabla_i \alpha)_j.y^j + \alpha_j.(\nabla_i y)^j\right).$$

By [7.25], noting that Y was arbitrarily chosen, we deduce (**exercise**):

$$\boxed{(\nabla_i \alpha)_k = \partial_i \alpha_j - \sum_k \Gamma^k_{ij}.\alpha_k.} \qquad [7.26]$$

(III) COVARIANT DERIVATIVE OF A TENSOR FIELD More generally, we can consider the covariant derivative of a *tensor field* $\mathbf{U} \in \mathcal{T}_s^r(B)$ along a vector field $X \in \mathcal{T}_0^1(B)$: the space $\mathbf{T}_s^r(B)$ (with the convention from section 4.3.3(**III**)) is a vector bundle with base B, and so the remarks of section 7.2.2 are applicable; the connection determines an operator

$$\nabla : T(B) \times \mathbf{T}_s^r(B) \to \mathbf{T}_{s+1}^r(B).$$

The mapping $(\mathbf{V}^*, X) \to \langle \nabla_X.\mathbf{U}, \mathbf{V}^* \rangle$ from $\mathcal{T}_r^s(B) \times \mathcal{T}_0^1(B)$ into $\mathcal{E}(B)$ is $\mathcal{E}(B)$-bilinear. It therefore determines a tensor field $\nabla\mathbf{U} \in \mathcal{T}_{s+1}^r(B)$, called the *covariant differential* of $\mathbf{U}$ (for the connection $\mathbf{C}$) according to the equality $\langle \nabla\mathbf{U}, \mathbf{V}^* \otimes X \rangle = \langle \nabla_X.\mathbf{U}, \mathbf{V}^* \rangle$. Consider a tensor field $\mathbf{T} = \mathbf{U} \otimes \mathbf{V}$. Then, $\nabla\mathbf{T} = (\nabla\mathbf{U}) \otimes \mathbf{V} + \mathbf{U} \otimes \nabla\mathbf{V}$. We can use linearity to calculate the covariant exterior differential of any tensor field from the relations [7.24] and [7.26]. The general rule is that, for each contravariant (respectively covariant) index, we need to add (respectively remove) an appropriate term $\Gamma^{\bullet}_{\bullet}$. This gives the following result:

THEOREM 7.12.– *Consider the field of tensors of type (q, p) with components $A^{j_1 \dots j_q}_{i_1 \dots i_p}$. Then, the covariant derivative $\nabla_h A$ of this tensor is the tensor of type $(q, p+1)$ given by*

$$\nabla_h A^{j_1 \dots j_q}_{i_1 \dots i_p} = \partial_h A^{j_1 \dots j_q}_{i_1 \dots i_p} + \sum_{s=1}^{q} \sum_{\nu=1}^{n} \Gamma^{j_s}_{h\nu}.A^{j_1 \dots j_{s-1} \nu j_{s+1} \dots j_q}_{i_1 \dots i_p}$$
$$- \sum_{r=1}^{p} \sum_{\nu=1}^{n} \Gamma^{\nu}_{h i_r}.A^{j_1 \dots j_q}_{i_1 \dots i_{r-1} \nu i_{r+1} \dots i_p}. \qquad [7.27]$$

Note that, if $\sigma \in \mathcal{E}(B)$ and $\mathbf{U} \in \mathcal{T}_r^s(B)$, then $\nabla\sigma = d\sigma$, and, by [7.21], the *Leibniz rule* holds:

$$\nabla(\sigma.\mathbf{U}) = \sigma(\nabla\mathbf{U}) + \mathbf{U} \otimes d\sigma.$$

Furthermore, the operator ∇ commutes with the index contraction c_i^j (section 4.2.1(**II**)), i.e. if $\mathbf{U} \in \mathcal{T}_r^s(B)$ with $r, s \geq 1$, then $\nabla\left(c_i^j \mathbf{U}\right) = c_i^j \nabla(\mathbf{U})$ (**exercise**).

7.2.4. *Parallel transport and geodesics*

(I) NOTION OF PARALLEL TRANSPORT Suppose that a Chinese explorer who lived three millennia ago (according to legend) is located at the North Pole. The explorer decides to move directly forward, indicating the direction taken (necessarily south) by leaving an arrow attached to a post planted in the ground. Now, suppose that the explorer is sitting in a *south-pointing chariot*[7] (whose pointer is always directed toward the celestial south). After reaching a certain latitude, such as the equator, the explorer follows it around to the east for around 45° of longitude, then returns north. Surprise: the direction of the chariot's pointer is no longer the same as the arrow left earlier; there is an angle θ between them (Figure 7.1). Hence, the Earth cannot be flat. The pointer of the chariot can be viewed as a vector field that was transported parallel to itself throughout the entire journey.

7 See the Wikipedia article on the *South-pointing chariot*.

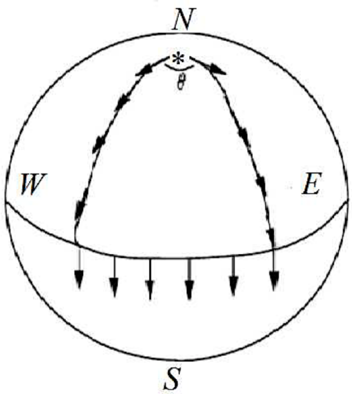


Figure 7.1. *Parallel transport on Earth*

(II) FORMALIZATION

DEFINITION 7.13.– *Let c be a curve of class C^2 on the manifold B equipped with the linear connection ∇. A vector field $Y \in T(B)$ is said to be* transported parallel (to itself) *along c if, setting $v = dc/dt \in T(B)$, Y satisfies the differential equation:*

$$\boxed{\nabla_v Y = 0.} \qquad [7.28]$$

See Definition 7.36 below for a more general setting. We sometimes write $\frac{\nabla Y}{dt} := \nabla_{\frac{dc}{dt}} Y$. In a system of local coordinates (U, ξ, n), the equation [7.28] can be written in the form $\sum_j \frac{dc^j}{dt} \frac{\partial y^k}{\partial \xi^j} + \sum_{i,j} \Gamma_{ij}^k y^j \frac{dc^i}{dt} = 0$, setting $c^k(t) = \xi^k(c(t))$. This expression is equivalent to:

$$\boxed{\frac{dy^k}{dt} + \sum_{i,j} \Gamma_{ij}^k . y^j . \frac{dc^i}{dt} = 0.}$$

Thus, $\frac{dy^k}{dt} = -\sum_{i,j} \Gamma_{ij}^k . y^j . \frac{dc^i}{dt}$. We sometimes define:

$$\left(dy^k\right)_{\|} := -\sum_{i,j} \Gamma_{ij}^k . y^j . d\xi^i .$$

By section 1.5.2**(II)**, we have the following result:

LEMMA 7.14.– *Given a curve* $c : [a,b] \to B$ *and a tangent vector* $\tau \in T_{c(a)}(B)$, *there exists a unique vector field* V *such that* $V(a) = \tau$ *and* V *is parallel to itself along* c.

Passing from the point $M = b$ to an infinitely close point $M' = b' = b + \overrightarrow{db}$ makes the vector field Y change from $\mathbf{A} = Y(b) \in T_b(B)$ to $\mathbf{A}' = Y(b') \in T_{b'}(B)$. A frame $(\mathbf{e}_\mu)$ (respectively $(\mathbf{e}'_\mu)$) is attached to the point M (respectively M'). This connection associates the point M' (respectively the vector $\mathbf{A}'$, respectively the vector $\mathbf{e}'_\mu$) with the point $m' = M + \mathbf{d}M = b + \overrightarrow{\mathbf{d}b}$ (respectively the vector $\mathbf{A} + \mathbf{dA}$, respectively $\mathbf{e}_\mu + \mathbf{de}_\mu$) in $T_b(B)$ (see Figure 7.2). If the transport of b to $b + \overrightarrow{db}$ is parallel along a curve c, then $\nabla_v Y = 0$, where $v = dc/dt$.

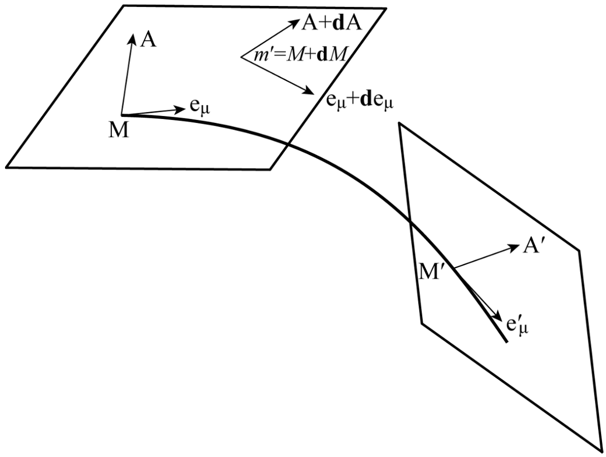


Figure 7.2. *Correspondence between* $M' \longleftrightarrow m'$

(III) GEODESICS The curve c is said to be a *geodesic* if the "speed vector" $v = dc/dt$ is transported parallel to itself along c, or in other words if

$$\boxed{\nabla_v v = 0,} \qquad [7.29]$$

or, alternatively, in the system of local coordinates considered above,

$$\boxed{\frac{d^2 c^k}{dt^2} + \sum_{i,j} \Gamma^k_{ij} \cdot \frac{dc^i}{dt} \cdot \frac{dc^j}{dt} = 0.} \qquad [7.30]$$

The equation [7.29] (or [7.30]) is known as the *geodesic equation*.

It follows from the very definition that the geodesics of the plane are the straight lines.

7.2.5. *Covariant exterior differential*

Consider the space $\Omega^p(B;\mathbf{F})$ of p-forms on a manifold B taking values in a Banach space $\mathbf{F}$ (section 4.4.3(**VI**)); any element $\varpi \in \Omega^p(B;\mathbf{F})$ is a mapping of class C^∞

$$\bigwedge^p T(B) \to \mathbf{F}$$

such that $\varpi(X(b) \wedge ... \wedge X_p(b)) \in \mathbf{F}$ for every vector field $X_1, ..., X_p \in T(B)$ and every $b \in B$.

First, let $\varpi \in \Omega^1(B;\mathbf{F})$, $b \in B$, $\mathbf{h}_b, \mathbf{k}_b \in T_b(B)$, and let $d\varpi \in \Omega^2(B;\mathbf{F})$ be the exterior differential of ϖ (section 5.5.1(**I**)). Remark 5.23 still holds if we replace the Lie derivative by the covariant derivative: the quantity

$$\boxed{\langle \nabla\varpi, X \wedge Y \rangle := \nabla_X.\langle \varpi, Y \rangle - \nabla_Y.\langle \varpi, X \rangle - \langle \varpi, [X,Y] \rangle} \qquad [7.31]$$

evaluated at b only depends on the values of X and Y *at the point* b; calculated in this way, $\nabla\varpi \in \Omega^2(B;\mathbf{F})$ is said to be the *covariant exterior differential* of ϖ; it is sometimes written as $\mathbf{d}\varpi$ (as in section 7.2.4(**II**))[8]. Unlike $d\varpi$, it is a *tensor field* (see section 7.3.2).

This can be generalized as follows:

LEMMA-DEFINITION 7.15.– *Let B be a manifold, E a vector bundle with base B and $\varpi \in \Omega^p(B;E)$ a differential p-form of class C^∞ taking values in E (section 4.4.3(**I**)). Let $X_0, X_1, ..., X_p$ be a set of p vector fields $\in \mathcal{T}_0^1(B)$ and consider*

$$\nabla\varpi.(X_0 \wedge X_1 \wedge ... \wedge X_p) = \sum_{j=0}^{p} (-1)^j \nabla_{X_j}.\left(\varpi.\left(X_0 \wedge ... \wedge \hat{X}_j \wedge ... \wedge X_p\right)\right)$$

$$+ \sum_{0 \le i < j \le p} (-1)^{i+j} \left(\varpi.\left([X_i, X_j] \wedge X_0 \wedge ... \wedge \hat{X}_i \wedge ... \wedge \hat{X}_j \wedge ... \wedge X_p\right)\right) \qquad [7.32]$$

(see Theorem 5.22(iv)). At every point $b \in B$, the left-hand side takes a value that only depends on the values $X_j(b)$ $(0 \le j \le p)$; $\nabla\varpi$ is called the covariant exterior differential *(or* absolute differential*) of ϖ; this is a differential $(p+1)$-form of class C^∞ on B taking values in E. Thus, $\nabla : \Omega^p(B;E) \to \Omega^{p+1}(B;E)$.*

8 Here, the reader should take care to distinguish between the symbols d and $\mathbf{d}$.

Writing $A_{i_1 \dots i_p}$ for the components of ϖ, in accordance with [7.27], we have:

$$\nabla_h A_{i_1 \dots i_p} = \partial_h A_{i_1 \dots i_p} - \sum_{r=1}^{p} \sum_{\nu=1}^{n} \Gamma^{\nu}_{hi_r} . A_{i_1 \dots i_{r-1} \nu i_{r+1} \dots i_p} . \qquad [7.33]$$

If we set $p = 0$, $\Omega^0 (B; E)$ is the set of sections $\boldsymbol{G}$ of class C^∞ of E (section 4.4.3**(I)**, Remark 4.30(ii)); if we set $\omega = \boldsymbol{G}$ and $X_0 = X$, we obtain the covariant differential $\nabla \boldsymbol{G} \in \Omega^1 (B; E)$ defined by

$$\boxed{\langle \nabla \boldsymbol{G}, X \rangle := \nabla \boldsymbol{G} . X = \nabla_X . \boldsymbol{G}.} \qquad [7.34]$$

Next, we can calculate the covariant differential of $\nabla \boldsymbol{G}$. In general, $\nabla^2 := \nabla \circ \nabla \neq 0$ (unlike d^2). By [7.34] and [7.31],

$$\boxed{(\nabla \circ \nabla) \boldsymbol{G} (X \wedge Y) = \nabla_X . (\nabla_Y . \boldsymbol{G}) - \nabla_Y . (\nabla_X . \boldsymbol{G}) - \nabla_{[X,Y]} . \boldsymbol{G},} \qquad [7.35]$$

and $\nabla (\nabla \boldsymbol{G})$ is a differential 2-form of class C^∞ taking values in E. For every $\mathbf{h}_b, \mathbf{k}_b \in T_b (B)$, $(\nabla \circ \nabla) \boldsymbol{G} (\mathbf{h}_b \wedge \mathbf{k}_b) \in E_b$ only depends on the value $\boldsymbol{G} (b)$ and not on the values of $\boldsymbol{G}$ in the neighborhood of b ([DIE 93], Volume 3, (17.20.1)).

The mapping $\boldsymbol{G} \mapsto (\nabla \circ \nabla) \boldsymbol{G} (\mathbf{h}_z \wedge \mathbf{k}_z)$ is linear, so there exists an endomorphism $R_b (\mathbf{h}_b \wedge \mathbf{k}_b)$ of E_b such that $(\nabla \circ \nabla) \boldsymbol{G} (\mathbf{h}_b \wedge \mathbf{k}_b) = R_b (\mathbf{h}_b \wedge \mathbf{k}_b) . \boldsymbol{G} (b)$. Furthermore, the mapping

$$\mathbf{r} : \mathbf{h}_b \wedge \mathbf{k}_b \mapsto R_b (\mathbf{h}_b \wedge \mathbf{k}_b)$$

is a morphism of vector bundles from $\bigwedge^2 T (B)$ into $\mathrm{End} (E) = E^\vee \otimes E$ ([P1], section 3.1.5**(I)**); hence, $\mathbf{r}$ is a differential 2-form on B taking values in $E^\vee \otimes E$.

7.2.6. *Curvature and torsion of a linear connection*

Consider the case where $E = T (B)$.

(I) CURVATURE

DEFINITION 7.16.– *The mapping* $\mathbf{r} : \bigwedge^2 T (B) \to T (B)^\vee \otimes T (B)$ *is called the* curvature B-morphism *(or simply the* curvature*) of the linear connection* $\mathbf{C}$ *(section 3.3.1, Lemma-Definition 3.7(iv)).*

This morphism $\mathbf{r}$ determines a bilinear B-morphism

$$(\mathbf{h}_b, \mathbf{k}_b) \mapsto \mathbf{r} . (\mathbf{h}_b \wedge \mathbf{k}_b)$$

from $T (B) \times T (B)$ into $\mathbf{T}^1_1 (B)$ that can be identified with a tensor field $\mathbf{r} \in \mathcal{T}^1_3 (B)$, called the *curvature tensor* (or the *Riemann–Christoffel tensor*) by abuse of language.

THEOREM 7.17.– *i) Given any vector fields $X, Y, Z \in \mathcal{T}_0^1(B)$ and any 1-form $\varpi \in \mathcal{T}_1^0(B)$, the curvature tensor $\mathbf{r}$ is defined by:*

$$\boxed{\mathbf{r}(X, Y, Z, \varpi) = \varpi\left(\nabla_X.\nabla_Y.Z - \nabla_Y.\nabla_X.Z - \nabla_{[X,Y]}.Z\right).}$$

ii) Let (U, ξ, m) be a chart of B and $(X_i)_{1 \leq i \leq m}$ the basis of $\mathcal{T}_0^1(U)$ on $\mathcal{E}(U)$ associated with this chart, namely $X_i = \frac{\partial}{\partial \xi^i}$ (Corollary-Definition 2.30). Then:

$$(\mathbf{r}.(X_j \wedge X_k)).X_i = \sum_l R^l_{i,jk}.X_l,$$

where the components $R^l_{i,jk}$ are given by

$$\boxed{R^l_{i,jk} = \partial_j \Gamma^l_{ki} - \partial_k \Gamma^l_{ji} + \sum_h \left(\Gamma^h_{ki}.\Gamma^l_{jh} - \Gamma^h_{ji}.\Gamma^l_{kh}\right).} \qquad [7.36]$$

Thus, $\mathbf{r} = \left(\mathbf{r}^l_i\right)_{1 \leq i,l \leq n}$, where, by [4.12] (section 4.2.4),

$$\mathbf{r}^l_i = \sum_{j,k} R^l_{i,jk}.d\xi^j \wedge d\xi^k$$

(where $\mathbf{r}$ is viewed as a twice covariant antisymmetric tensor field taking values in $T(B)^\vee \otimes T(B)$)[9]*.*

PROOF.– This immediately follows from Definition 7.16 and [7.35]. ■

(II) TORSION In the same situation, the identity mapping $1_{T(B)}$ can be viewed as a differential 1-form on B taking values in $T(B)$. Its covariant exterior differential $\mathbf{t} = \nabla\left(1_{T(B)}\right)$ is therefore a B-morphism from $\bigwedge^2 T(B)$ into $T(B)$, called the *torsion B-morphism* (or simply the *torsion*) of the linear connection $\mathbf{C}$. Hence:

DEFINITION 7.18.– *Given any vector fields $X, Y \in \mathcal{T}_0^1(B)$, the torsion B-morphism $\mathbf{t}$ is defined by:*

$$\boxed{\mathbf{t}.(X \wedge Y) = \nabla_X.Y - \nabla_Y.X - [X, Y].}$$

This morphism $\mathbf{t}$ defines a bilinear B-morphism $(\mathbf{h}_b, \mathbf{k}_b) \mapsto \mathbf{t}.(\mathbf{h}_b \wedge \mathbf{k}_b)$ from $T(B) \times T(B)$ into $T(B)$ and may therefore be identified with a tensor field $\mathbf{t} \in \mathcal{T}_2^1(B)$, called the torsion tensor field (or the *torsion tensor*, by abuse of language) of the linear connection $\mathbf{C}$.

9 The purpose of the comma in $R^l_{i,jk}$ is to indicate that $\mathbf{r}$ is a twice contravariant tensor (with indices j, k) taking matrix values (matrix with indices i, l).

THEOREM 7.19.– *Let* (U, ξ, m) *be a chart of* B *and* $(X_i)_{1 \leq i \leq m}$ *the basis of* $\mathcal{T}_0^1(U)$ *on* $\mathcal{E}(U)$ *associated with this chart, namely* $X_i = \frac{\partial}{\partial \xi^i}$*. Then:*

$$\mathbf{t}.(X_j \wedge X_k) = \sum_i T^i_{jk} X_i,$$

where the components T^i_{jk} *are given by*

$$\boxed{T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}} \quad [7.37]$$

(a quantity that is antisymmetric in the indices j, k*). Hence,* $\mathbf{t} = \left(\mathbf{t}^i\right)_{1 \leq i \leq n}$*, where, by [4.12] (section 4.2.4),*

$$\boxed{\mathbf{t}^i = \sum_{j,k} T^i_{jk}.d\xi^j \wedge d\xi^k} \quad [7.38]$$

(where $\mathbf{t}$ *is viewed as a twice covariant antisymmetric tensor field taking values in* $T(B)$*).*

PROOF.– Let $(\xi_i)_{1 \leq i \leq m}$ be a system of local coordinates and write x^i, y^i for the coordinates of two arbitrary vector fields X, Y in the basis $(\partial/\partial \xi_i)_{1 \leq i \leq m}$. By [7.22], $(\nabla_X.Y)^i = \sum_k \left(\partial_k y^i\right) x^k + \sum_{j,k} \Gamma^i_{jk}.x^j.y^k$ and, by [5.15], $[X,Y]^i = \sum_j \left(\partial_j y^i\right).$ $x^j - \sum_j \left(\partial_j x^i\right).y^j$. It therefore follows that:

$$\begin{aligned}\nabla_X.Y - \nabla_Y.X - [X,Y] &= \sum_{i,j,k} \left(\Gamma^i_{jk} - \Gamma^i_{kj}\right).x^j.y^k.\frac{\partial}{\partial \xi^i} \\ &= \sum_{i,j,k} T^i_{jk}.x^j.y^k.\frac{\partial}{\partial \xi^i}.\end{aligned}$$

■

7.3. Method of moving frames

The codomain of a frame mapping r (section 3.1**(II)**) is a principal bundle $R(B)$ with base B and structural group $\mathrm{GL}_n(\mathbb{R})$ (section 3.5.1**(I)**, Lemma-Definition 3.51) on which we know how to define a principal connection since Ehresmann. This is the modern perspective adopted by differential geometry. It has the advantage of being extremely general and the disadvantage of being extremely abstract; rather than tackling it directly, it might be more insightful to retrace the path from Cartan's original ideas to those of Ehresmann, starting with the method of moving frames [CAR 35]. According to this method, we keep the same base B, and any given frame mapping determines a linear connection **C** on B.

7.3.1. *Moving frame and gauge potential*

(I) BUNDLE OF FRAMES A frame $r_b \in R(B)$ is an n-tuple $\left[\mathbf{h}_{1,b} \cdots \mathbf{h}_{n,b}\right]$ (the *row* formed by the vectors $\mathbf{h}_{i,b}$ of a basis of the tangent space $T_b(B)$). The mapping $r : b \mapsto r_b$ was called a *moving frame* by É. Cartan, generalizing the notion introduced by Darboux for surfaces [DAR 13]. The group $\mathrm{GL}_n(\mathbb{R})$ acts on r on the right, since any matrix $\boldsymbol{g} \in \mathrm{GL}_n(\mathbb{R})$ sends the frame r_b to the frame

$$\boxed{r'_b = r_b.\boldsymbol{g} := \left[\mathbf{h}_{1,b}.\boldsymbol{g} \cdots \mathbf{h}_{n,b}.\boldsymbol{g}\right]} \qquad [7.39]$$

(the *row* formed by the $\mathbf{h}_{i,b}.\boldsymbol{g}$). Write $A^j_{i'}$ for the elements of $\boldsymbol{g}$ and $(\mathbf{h}_{i',b})_{1 \leq i' \leq n}$ for the new frame. Then, we have the relations $\mathbf{h}_{i',b} = \sum_i \mathbf{h}_{i,b}.A^i_{i'}$, $\mathbf{h}_{i,b} = \sum_{i'} \mathbf{h}_{i',b}.A^{i'}_i$, where $\left(A^{i'}_i\right)$ denotes the inverse matrix $\boldsymbol{g}^{-1}$ of $\boldsymbol{g} = \left(A^i_{i'}\right)$ (section 1.2.1(**II**)).

(II) BUNDLE OF G-FRAMES We can add constraints to the change-of-frame matrices $\boldsymbol{g}$. This is equivalent to replacing the Lie group $\mathrm{GL}_n(\mathbb{R})$ by a Lie *subgroup* $\mathbf{G}$ of $\mathrm{GL}_n(\mathbb{R})$ (section 2.4.1(**VI**)) and replacing the Lie algebra $\mathrm{Mat}_n(\mathbb{R})$ of $\mathrm{GL}_n(\mathbb{R})$ by the Lie algebra $\mathfrak{g} \subset \mathrm{Mat}_n(\mathbb{R})$ of $\mathbf{G}$.

For example, in the case of a differential manifold:

– $\mathbf{G} = \mathrm{SL}_n(\mathbb{R})$ to define an orientation, again giving $\mathfrak{g} = \mathrm{Mat}_n(\mathbb{R})$;

– $\mathbf{G} = \mathrm{O}_n(\mathbb{R})$ to define a Riemannian structure (section 4.5.1), giving $\mathfrak{g} = \mathfrak{o}_n(\mathbb{R})$;

– $\mathbf{G} = \mathrm{SO}_n(\mathbb{R})$ to define an oriented Riemannian structure, again giving $\mathfrak{g} = \mathfrak{o}_n(\mathbb{R})$, etc.

We say that the manifold B is equipped with a $\mathbf{G}$-structure. The corresponding connection is called a $\mathbf{G}$-connection (see **(V)** below); the corresponding moving frame $r : b \mapsto r_b$ is said to be a moving $\mathbf{G}$-frame, and each r_b is said to be a $\mathbf{G}$-frame. The bundle of moving $\mathbf{G}$-frames is denoted $R_{\mathbf{G}}(B)$, or just $R_{\mathbf{G}}$ when the base B is implicitly clear. This informal definition of these concepts is made precise in section 7.3.8 (using the notion of principal connection), where we also show that arbitrary manifolds cannot be equipped with arbitrary $\mathbf{G}$-structures.

In the following, the manifold B is equipped with a $\mathbf{G}$-structure.

(III) SOLDERING 1-FORM ON B Now, consider the cotangent space $T^\vee_b(B)$ of the base B at the point b and let $\left(\sigma^i_b\right)$ (respectively $\left(\sigma^{i'}_b\right)$) be the dual basis of $(\mathbf{h}_{i,b})$ (respectively $(\mathbf{h}_{i',b})$), which is therefore a coframe with origin b. Then,

$\sigma_b^i = \sum_{i'} A_{i'}^i . \sigma_b^{i'}$, $\sigma_b^{i'} = \sum_i A_i^{i'} . \sigma_b^i$. Alternatively, writing $\underline{\sigma}_b$ and $\underline{\sigma}'_b$ for the *columns* of the σ_b^i and the $\sigma_b^{j'}$ respectively,

$$\boxed{\underline{\sigma}_b = \boldsymbol{g}.\underline{\sigma}'_b,} \qquad [7.40]$$

where $\boldsymbol{g} = \left(A_{i'}^i\right)$. Each $\sigma^i : b \mapsto \sigma_b^i$ is a field of covectors, i.e. a 1-form, so $\underline{\sigma}$ is a vector-valued 1-form on B taking values in $\mathbb{R}^n$ (section 4.4.3(**VI**)): $\underline{\sigma} \in \Omega^1\left(B; \mathbb{R}^n\right)$[10].

DEFINITION 7.20.– *The vector-valued differential 1-form* $\underline{\sigma}$ *is called the* soldering form *of* **C**.

(IV) GAUGE POTENTIAL Let (U, ξ, n) be a chart of B and, for every $b \in U$, let $\underline{\sigma}_b^U$ be a coframe with origin b over U; suppose that the mapping $b \mapsto \underline{\sigma}_b^U$ is of class C^∞. By duality, from each $\underline{\sigma}_b^U$, we deduce a frame $r_{U,b}$ with origin b. Given a chart (V, ζ, n) such that $U \cap V \neq \varnothing$, we do the same. For every $b \in U \cap V$, there exists a matrix $A_V^U(b) \in \mathbf{G}$ such that

$$\underline{\sigma}_b^U = A_V^U(b) . \underline{\sigma}_b^V, \quad r_{V,b} = r_{U,b} . A_V^U(b) . \qquad [7.41]$$

Repeating the above for a chart (W, υ, n) such that $U \cap V \cap W \neq \varnothing$ yields the relation $A_W^U = A_V^U . A_W^V$.

LEMMA 7.21.– *Let*

$$\Lambda_{UV} = A_U^V . dA_V^U, \qquad [7.42]$$

where $A_U^V := \left(A_V^U\right)^{-1}$. *Then,* Λ_{UV} *is a differential 1-form on* B *taking values in* $\mathfrak{g}$ *and, setting* $\underline{\omega}' = \Lambda_{UW}$, $\underline{\omega} = \Lambda_{UV}$, *and* $\boldsymbol{g} = \left(A_W^V\right)$, *we obtain:*

$$\underline{\omega}' = \boldsymbol{g}^{-1} . \underline{\omega} . \boldsymbol{g} + \boldsymbol{g}^{-1} . d\boldsymbol{g}. \qquad [7.43]$$

PROOF.– We have $A_W^U = A_V^U . A_W^V$, so $dA_W^U = dA_V^U . A_W^V + A_V^U . dA_W^V$ by the Leibniz rule. Furthermore, $dA_V^U = A_V^U . \Lambda_{UV}$ by [7.42]. Hence, $dA_W^U = A_V^U . \Lambda_{UV} . A_W^V + A_V^U . dA_W^V$, and applying [7.42] once again gives

$$\Lambda_{UW} = \underbrace{A_U^W . A_V^U}_{A_V^W} . \Lambda_{UV} . A_W^V + \underbrace{A_U^W . A_V^U}_{A_V^W} . dA_W^V,$$

showing [7.43]. ■

The following definition is commonly adopted in physics ([COQ 02], section 4.1.5):

10 The underlined symbols (here $\underline{\sigma}$) denote forms on the base B (similarly, we will introduce the forms $\underline{\omega}$ and $\underline{\boldsymbol{\Theta}}$ on B). The non-underlined symbols $\sigma, \omega, \boldsymbol{\Theta}$ will later denote the corresponding forms "lifted" onto the bundle of **G**-frames.

DEFINITION 7.22.– *The differential 1-form (or* Pfaff form*: see section 4.3.2, Definition 4.20)* $\underline{\omega}$ *on* B *taking values in* $\mathfrak{g}$ *is called the* gauge potential *with structural group* $\mathbf{G}$.

The most classical example of a gauge potential is the vector potential $\overrightarrow{A}$ of the electromagnetic field $\overset{\hookrightarrow}{B}$, which satisfies $\overset{\hookrightarrow}{B} = \text{curl}\,\overrightarrow{A}$ (Example 5.53(iii)) with $\mathbf{G} = \mathrm{U}_1(\mathbb{C})$[11].

REMARK 7.23.– *Writing* $\mathbf{g}$ *and* $\underline{\omega}$*, respectively, for the matrices* $\left(A^i_{i'}\right)$ *and* $\left(\omega^k_i\right)$*, the relation [7.43] becomes*

$$\omega^{k'}_{i'} = \sum_{i,k} A^{k'}_k.\omega^k_i.A^i_{i'} + \sum_i A^{k'}_i.dA^i_{i'}. \tag{7.44}$$

(V) COEFFICIENTS OF THE CONNECTION C The Pfaff form $\underline{\omega}$ taking values in $\mathfrak{g}$ is a matrix of Pfaff forms $\left(\omega^k_j\right)$ taking values in $\mathbb{R}$. Set

$$\omega^k_j = \sum_i \gamma^k_{ij}.\sigma^i; \tag{7.45}$$

this expression replaces [7.7], and the γ^k_{ij} are the connection coefficients of the $\mathbf{G}$-connection $\mathbf{C}$ in the coframe $\underline{\sigma}^U$[12]. By duality,

$$\gamma^k_{ij} = \omega^k_j.\mathbf{h}_i. \tag{7.46}$$

These coefficients transform straightforwardly under change of frame. By [7.44] and [7.45],

$$\sum_{i'} \gamma^{k'}_{i'j'}\,\sigma^{i'} = \sum_{j,k} A^{k'}_k\,\gamma^k_{ij}\,\sigma^i\,A^j_{j'} + \sum_i A^{k'}_i\,dA^i_{j'},$$
$$dA^i_{j'} = \sum_{i'} \partial_{i'}A^i_{j'}\,\sigma^{i'}, \quad \sigma^i = \sum_{i'} A^i_{i'}\,\sigma^{i'}.$$

Hence,

$$\gamma^{k'}_{i'j'} = \sum_{i,j,k} A^{k'}_k A^i_{i'} A^j_{j'} \gamma^k_{ij} + \sum_i A^{k'}_i.\partial_{i'}A^i_{j'}, \tag{7.47}$$

11 See the Wikipedia article on *Gauge theory*.

12 The notation Γ^k_{ij} is reserved for the case where σ^U is a natural coframe (section 7.2.3(**I**)). In this case, $\omega^k_j = \sum_i \Gamma^k_{ij}.d\xi^i$, which shows that the ω^k_j determine the connection $\mathbf{C}$ by [7.27].

showing that the γ_{ij}^k are not the components of a tensor (field) (compare with Remark 7.6(1)).

By contrast, consider two $\mathbf{G}$-connections $\hat{\mathbf{C}}$, $\check{\mathbf{C}}$ on the manifold B with coefficients $\hat{\gamma}_{ij}^k$ and $\check{\gamma}_{ij}^j$, respectively, in the coframe $\underline{\sigma}^U$. By [7.47], setting $\Delta_{ij}^k = \hat{\gamma}_{ij}^k - \check{\gamma}_{ij}^k$,

$$\Delta_{i'j'}^{k'} = \sum_{i,j,k} A_k^{k'} A_{i'}^i A_{j'}^j . \Delta_{ij}^k .$$

Thus, the Δ_{ij}^k are the components of a once contravariant and twice covariant tensor (field).

7.3.2. *Curvature, torsion and covariant exterior differential of a* $\mathbf{G}$*-connection*

(I) TORSION By [7.41],

$$d\underline{\sigma}^V = A_U^V . d\underline{\sigma}^U + dA_U^V \wedge \underline{\sigma}^U = A_U^V . d\underline{\sigma}^U + dA_U^V \wedge A_V^U . \underline{\sigma}^V .$$

But $I = A_U^V . A_V^U$, so $A_U^V . dA_V^U + dA_U^V . A_V^U = 0$ and hence, by [7.42],

$$\Lambda_{UV} = A_U^V . dA_V^U = -dA_U^V . A_V^U \qquad [7.48]$$

and also

$$d\underline{\sigma}^V = A_U^V . d\underline{\sigma}^U - A_U^V . dA_V^U . \underline{\sigma}^V = A_U^V . d\underline{\sigma}^U - \Lambda_{UV} \wedge \underline{\sigma}^V . \qquad [7.49]$$

Set $\underline{\omega}_V = \underline{\omega}'$ and $\underline{\omega}_U = \underline{\omega}$; by [7.43] and [7.42], it follows that

$$\underline{\omega}_V = A_U^V . \underline{\omega}_U . A_V^U + \Lambda_{UV} , \qquad [7.50]$$

which gives

$$\underline{\omega}_V \wedge \underline{\sigma}^V = A_U^V . \underline{\omega}_U \wedge A_V^U . \underline{\sigma}^V + \Lambda_{UV} \wedge \underline{\sigma}^V = A_U^V . \underline{\omega}_U \wedge \underline{\sigma}^U + \Lambda_{UV} \wedge \underline{\sigma}^V .$$

Furthermore, by [7.49], $d\underline{\sigma}^V + \underline{\omega}_V \wedge \underline{\sigma}^V = A_U^V \left(d\underline{\sigma}^U + \underline{\omega}_U \wedge \underline{\sigma}^U \right)$.

$$\boxed{\underline{\mathbf{\Theta}}^U := \nabla \underline{\sigma}^U := d\underline{\sigma}^U + \underline{\omega}_U \wedge \underline{\sigma}^U} \qquad [7.51]$$

is a twice covariant and once contravariant antisymmetric tensor (field); $\underline{\mathbf{\Theta}}^U = \left[\Theta^1 \cdots \Theta^n \right]$, where

$$\boxed{\Theta^k = \sum_{i,j} S_{ij}^k . \sigma^i \wedge \sigma^j .}$$

Note that $\underline{\omega}_U \wedge \underline{\sigma}^U$ is the column of 2-forms $\left(\underline{\omega}_U \wedge \underline{\sigma}^U \right)^k = \sum_j \omega_j^k \wedge \sigma^j$.

THEOREM 7.24.– *If* $\mathbf{C}$ *is a linear connection, then, over* U, $\underline{\mathbf{\Theta}}^U = \frac{1}{2}\mathbf{t}$, *where* $\mathbf{t}$ *is the torsion tensor of* $\mathbf{C}$.

PROOF.– It suffices to show the result when $\underline{\sigma}^U$ is a natural frame. Then, $\sigma^i = d\xi^i$, so $d\sigma^i = 0$, and $\gamma_{ij}^k = \Gamma_{ij}^k$. Hence, by [7.51] and [7.45],

$$\Theta^k = \sum_{i,j} \Gamma_{ij}^k.\sigma^i \wedge \sigma^j = \frac{1}{2}\sum_{i,j}\left(\Gamma_{ij}^k - \Gamma_{ji}^k\right).\sigma^i \wedge \sigma^j,$$

and therefore (see [7.37])

$$S_{ij}^k = \frac{1}{2}\left(\Gamma_{ij}^k - \Gamma_{ji}^k\right) = \frac{1}{2}T_{ij}^k. \quad \blacksquare \qquad [7.52]$$

(II) CURVATURE By differentiating [7.50], noting that [5.22] holds (section 5.5.1(**I**)) and [7.42], we obtain:

$$d\underline{\omega}_V = A_U^V.d\underline{\omega}_U.A_V^U + dA_U^V \wedge \underline{\omega}_U.A_V^U - A_U^V.\underline{\omega}_U \wedge dA_V^U + \underbrace{d\Lambda_{UV}}_{dA_V^U \wedge dA_V^U}. \qquad [7.53]$$

Similarly, setting $[\underline{\omega}_V, \underline{\omega}_V] = \underline{\omega}_V \wedge \underline{\omega}_V$,

$$[\underline{\omega}_V, \underline{\omega}_V] = A_U^V.\underline{\omega}_U \wedge \underline{\omega}_U.A_V^U + \underbrace{A_U^V\left(\underline{\omega}_U.A_V^U \wedge \Lambda_{UV}\right)}_{\underline{\omega}_U.A_V^U \wedge A_U^V.\Lambda_{UV}}$$
$$+\underbrace{\Lambda_{UV} \wedge \left(A_U^V.\underline{\omega}_U.A_V^U\right)}_{\Lambda_{UV}.A_U^V \wedge \underline{\omega}_U.A_V^U} + \Lambda_{UV} \wedge \Lambda_{UV}.$$

Thus, by [7.48],

$$[\underline{\omega}_V, \underline{\omega}_V] = A_U^V.\underline{\omega}_U \wedge \underline{\omega}_U.A_V^U + A_U^V.\underline{\omega}_U \wedge dA_V^U$$
$$-dA_U^V \wedge \underline{\omega}_U.A_V^U - dA_U^V \wedge dA_U^V. \qquad [7.54]$$

Adding both sides of [7.53] and [7.54] together gives that $\underline{\mathbf{\Omega}}_V = A_U^V.\underline{\mathbf{\Omega}}_U.A_V^U$, where

$$\boxed{\underline{\mathbf{\Omega}}_U := \nabla\underline{\omega}_U := d\underline{\omega}_U + [\underline{\omega}_V, \underline{\omega}_V].} \qquad [7.55]$$

THEOREM 7.25.– *If* $\mathbf{C}$ *is a linear connection, then, over* U, $\underline{\mathbf{\Omega}}_U = \frac{1}{2}\mathbf{r}$, *where* $\mathbf{r}$ *is the curvature tensor of the connection* $\mathbf{C}$.

PROOF.– We proceed in the same way as the proof of Theorem 7.24 after choosing a natural frame. Then, by [7.45], $\underline{\omega}_U = \left(\omega_i^k\right)$, where $\omega_i^l = \sum_k \Gamma_{ji}^k . d\xi^k$. Hence, by [7.36] (**exercise**), $\underline{\mathbf{\Omega}}_U = \left(\Omega_i^l\right)$, where

$$\Omega_i^l = \frac{1}{2} \sum_{j,k} R_{i,jk}^l . d\xi^j \wedge d\xi^k. \qquad \blacksquare \quad [7.56]$$

DEFINITION 7.26.– *1) The 2-forms $\underline{\mathbf{\Theta}}^U$ and $\underline{\mathbf{\Omega}}_U$ (taking values in $\mathbb{R}^n$ and $\mathfrak{g}$) are, respectively, called the* torsion form *and the* curvature form *of the $\mathbf{G}$-connection $\mathbf{C}$ (over U).*

2) Let $\mathbf{C}$ be a $\mathbf{G}$-connection. Its curvature tensor $\mathbf{r}$ and its torsion tensor $\mathbf{t}$ are defined over U by the relations $\mathbf{r} = 2\underline{\mathbf{\Omega}}_U$ and $\mathbf{t} = 2\underline{\mathbf{\Theta}}^U$.

(III) COVARIANT EXTERIOR DIFFERENTIAL Let B be a pure n-dimensional manifold and

$$\mathbf{B}_q^p(B) = \left(\bigwedge^q (T(B))^\vee\right) \otimes \left(\bigwedge^p (T(B))\right).$$

There exists a multiplication B-morphism $\mathbf{B}_q^p(B) \otimes \mathbf{B}_s^r(B) \to \mathbf{B}_{q+s}^{p+r}(B)$ that sends each section $(\alpha \otimes \mathbf{u}) \otimes (\beta \otimes \mathbf{v})$ to the section $(\alpha \wedge \beta) \otimes (\mathbf{u} \wedge \mathbf{v})$. The direct sum $\mathbf{B}(B) = \bigoplus_{q,p} \mathbf{B}_q^p(B)$ is an algebra bundle on B (section 4.4.3**(III)**). Write $\mathfrak{B}_q^p(B)$ (respectively $\mathfrak{B}(B)$) for the $\mathcal{C}^\infty(B)$-module (respectively the $\mathcal{C}^\infty(B)$-algebra) of sections of $\mathbf{B}_q^p(B)$ (respectively $\mathbf{B}(B)$) of class C^∞. The elements of $\mathfrak{B}_q^0(B)$ and the elements of $\mathfrak{B}_q^1(B)$ can be identified with the differential q-forms on B taking values in $\mathbb{R}$ and $T(B)$, respectively (section 4.4.3**(I)**, Definition 4.29).

We have the following result (**exercise***: see [DIE 93], Volume 4, Chapter 20, section 6, Exercise 2):

LEMMA 7.27.– *There exists a unique differential operator (section 5.2.2**(I)**) $\mathbf{d}$ from $\mathfrak{B}(B)$ into itself that satisfies $\mathbf{d}\left(\mathfrak{B}_q^p(B)\right) \subset \mathfrak{B}_{q+1}^p(B)$, coincides with the exterior differentiation d in each of the $\mathfrak{B}_q^0(B) = \Omega^q(B)$, coincides with covariant exterior differentiation $\bigtriangledown$ in each of the $\mathfrak{B}_q^1(B) = \Omega^q(B; T(B))$ (section 7.2.5) and satisfies the following relation for $\boldsymbol{v} \in \mathfrak{B}_q^p(B), \boldsymbol{w} \in \mathfrak{B}_s^r(B)$:*

$$\mathbf{d}(\boldsymbol{v}.\boldsymbol{w}) = (\mathbf{d}\boldsymbol{v}).\boldsymbol{w} + (-1)^q \boldsymbol{v}.(\mathbf{d}\boldsymbol{w}). \qquad [7.57]$$

This result is straightforward to extend to the case where $\boldsymbol{v}$ and $\boldsymbol{w}$ take matrix values, in which case the number of columns of $\boldsymbol{v}$ must be equal to the number of rows of $\boldsymbol{w}$.

In the following, the differential operator from Lemma 7.27 is denoted $\bigtriangledown$.

7.3.3. *Quasi-parallelogram method*

(I) THE QUASI-PARALLELOGRAM On the manifold B, consider a coordinate system $\left(y^i\right)_{1\leq i\leq n}$ (section 7.2.1), where each $y^i:\mathbb{R}\to B$ is of class C^∞, and, for every i, consider a curve of class C^∞

$$c_\alpha^i : t\mapsto \left(y^1\left(\alpha\right),...,y^{i-1}\left(\alpha\right),y^i\left(t\right),y^{i+1}\left(\alpha\right),...,y^n\left(\alpha\right)\right).$$

If $M\left(t\right)=\left(y^1\left(t\right),...,y^n\left(t\right)\right)$, set

$$\delta M\left(t\right)=\left(0,...,0,\underbrace{\dot{y}^i\left(t\right).dt}_{i-\text{th place}},0,...,0\right),\widehat{\delta}M\left(t\right)=\left(0,...,0,\underbrace{\dot{y}^j\left(t\right).dt}_{j-\text{th place}},0,...,0\right).$$

This defines two differentiation symbols δ and $\widehat{\delta}$. Below, the quantities δM and $\widehat{\delta}M$ are assimilated with infinitely small increments, as is common practice in physics.

DEFINITION 7.28.– *The differentiation symbols δ and $\hat{\delta}$ are said to be* interchangeable *if $\delta\circ\widehat{\delta}=\widehat{\delta}\circ\delta$.*

The first of the next two lemmas is due to É. Cartan ([CAR 45], section 2.1(27)), and the second is due to H. Weyl ([WEY 52], section 13, (25)):

LEMMA 7.29.– *The interchangeability of the two differentiation symbols is a "covariant" property, i.e. a property that is invariant under change of variable (see footnote 5, p. 320).*

PROOF.– Let $\left(x^i\right)_{1\leq i\leq n}$ be the old variables and $\left(y^i\right)_{1\leq i\leq n}$ the new ones. Then:

$$\delta y^i=\sum_k\frac{\partial y^i}{\partial x^k}.\delta x^k,\quad \widehat{\delta}\circ\delta y^i=\sum_{k,h}\frac{\partial^2 y^i}{\partial x^h\partial x^k}.\widehat{\delta}x^h.\delta x^k+\sum_k\frac{\partial y^i}{\partial x^k}\widehat{\delta}\circ\delta x^k,$$

$$\widehat{\delta}y^i=\sum_k\frac{\partial y^i}{\partial x^k}.\delta x^k,\quad \delta\circ\widehat{\delta}y^i=\sum_{k,h}\frac{\partial^2 y^i}{\partial x^k\partial x^h}.\delta x^k.\widehat{\delta}x^h+\sum_k\frac{\partial y^i}{\partial x^k}\delta\circ\widehat{\delta}x^k,$$

which gives the stated result by Schwarz's theorem (section 1.2.3, Theorem 1.16(i)).■

LEMMA 7.30.– *Given a coordinate system $\left(y^i\right)_{1\leq i\leq n}$, let M_0 be the point with coordinates $\left(y^1\left(\alpha_0\right),...,y^n\left(\alpha_0\right)\right)$, $M_1=M_0+\delta M_0$, $M_1'=M_1+\delta M_1$, $M_2=M_0+\widehat{\delta}M_0$, $M_2'=M_2+\delta M_2$ (see Figure 7.3). Assuming that the increments are infinitely small, $M_1'=M_2'$ if and only if the two differentiation symbols $\delta,\widehat{\delta}$ are interchangeable.*

PROOF.– We have $M_1' = (M_0 + \delta M_0) + \widehat{\delta}(M_0 + \delta M_0) = M_0 + \delta M_0 + \widehat{\delta} M_0 + \widehat{\delta} \circ \delta M_0$. Similarly, $M_2' = M_0 + \delta M_0 + \widehat{\delta} M_0 + \delta \circ \widehat{\delta} M_0$. ■

If the two differentiation symbols $\delta, \widehat{\delta}$ are interchangeable, the curve in Figure 7.3 is closed, defining a *quasi-parallelogram* P.

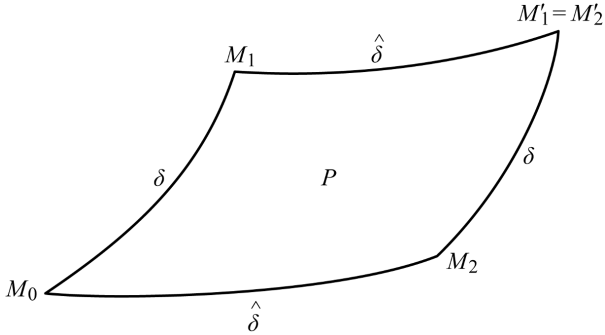


Figure 7.3. *Quasi-parallelogram on the manifold* B

(II) TORSION INTERPRETED AS TRANSLATION Every point M of P is located at the intersection of two curves c_α^i, c_β^j $(i \neq j)$. Identifying M with the element $(M, 1)$ of $\Omega^0(B; T(B)) = B \times \mathbb{R}$ (section 4.4.3(I), Remark 4.30(ii)), write $\triangledown M \in \Omega^1(B; T(B))$ for the differential element δM along c_α^i and $\widehat{\delta} M$ for the differential element along c_β^j; $\triangledown M$ is therefore a differential 1-form on B. If $\delta, \widehat{\delta}$ are interchangeable, as is assumed to be the case below, then Lemmas 7.29 and 7.30, respectively, show that $\triangledown$ is a covariant exterior differential (which justifies the notation) and that the quasi-parallelogram P and its boundary ∂P are well defined; suppose that ∂P is equipped with the orientation induced by the orientation of P (section 4.4.7(**VI**)). The covariant exterior differential $\triangledown$ determines a *linear connection* **C** on B (section 7.2.3).

Let us evaluate $\int_{\partial P} \triangledown M$. By Stokes' formula (section 5.6.1, Theorem 5.24),

$$\int_{\partial P} \triangledown M = \int_P d(\triangledown M),$$

and this quantity is non-zero if $\triangledown M$ is not an exact differential. To evaluate $\int_{\partial P} \triangledown M$, we can (provided that the quasi-parallelogram P is infinitely small and we may neglect infinitely small quantities of higher order than the desired quantity) associate M_1' (respectively M_2') with the point m_1' (respectively m_2') of the tangent space $T_{M_0}(B)$, as we did in section 7.2.4(**II**) (see Figure 7.2). Similarly, we associate m_1 with M_1, m_2 with M_2, set $m_0 = M_0$ (see Figure 7.4) and integrate over $T_{M_0}(B)$.

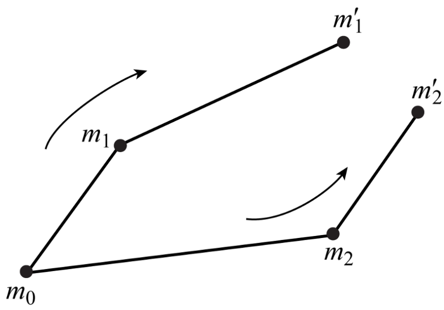


Figure 7.4. *Image of the quasi-parallelogram P on $T_{M_0}(B)$*

We now need to account for the fact that m_1' and m_2' do not necessarily coincide. Indeed, the vertices of Figure 7.4 satisfy the following relations:

$$\begin{array}{ccc} m_0 & = & m_0 \\ \downarrow & & \downarrow \\ m_1 = m_0 + \overrightarrow{\delta m_1} & & m_2 = m_0 + \overrightarrow{\widehat{\delta} m_0} \\ \downarrow & & \downarrow \\ m_1' = m_0 + \overrightarrow{\delta m_0} + \widehat{\delta}\left(m_0 + \overrightarrow{\delta m_0}\right) & & m_2' = m_0 + \overrightarrow{\widehat{\delta} m_0} + \delta\left(m_0 + \overrightarrow{\widehat{\delta} m_0}\right) \end{array}$$

Write $\overrightarrow{\delta m_0} = \sum_i d\xi^i \overrightarrow{\mathbf{e}_i}$, $\overrightarrow{\widehat{\delta} m_0} = \sum_i \widehat{\delta}\xi^i.\overrightarrow{\mathbf{e}_i}$ and $\overrightarrow{\delta \mathbf{e}_i} = \sum_h \omega_i^h(\delta).\overrightarrow{\mathbf{e}_h}$, $\overrightarrow{\widehat{\delta}\mathbf{e}_i} = \sum_h \omega_i^h\left(\widehat{\delta}\right).\overrightarrow{\mathbf{e}_h}$, as well as (in accordance with the right-hand side of [7.6]):

$$\omega_j^h(\delta) = \sum_j \Gamma_{ij}^h(\delta).\delta\xi^i, \quad \omega_j^h\left(\widehat{\delta}\right) = \sum_j \Gamma_{ij}^h\left(\widehat{\delta}\right).\widehat{\delta}\xi^i. \qquad [7.58]$$

With the hypothesis $\Gamma_{ij}^h(\delta) = \Gamma_{ij}^h\left(\widehat{\delta}\right)$, it follows that:

$$\begin{aligned} \overrightarrow{m_1' m_2'} &= \widehat{\delta}\left(\overrightarrow{\delta m_0}\right) - \delta\left(\overrightarrow{\widehat{\delta} m_0}\right) = \widehat{\delta}\left(\sum_i \delta\xi^i.\overrightarrow{\mathbf{e}_i}\right) - \delta\left(\sum_j \widehat{\delta}\xi^j.\overrightarrow{\mathbf{e}_j}\right) \\ &= \sum_i \widehat{\delta}\circ\delta\xi^i.\overrightarrow{\mathbf{e}_i} + \sum_{i,j,h} \delta\xi^i \otimes \Gamma_{ji}^h.\widehat{\delta}\xi^j.\overrightarrow{\mathbf{e}_h} \\ &\quad - \sum_j \delta\circ\widehat{\delta}\xi^j.\overrightarrow{\mathbf{e}_j} - \sum_{i,j,h} \widehat{\delta}\xi^j \otimes \Gamma_{ij}^h.\delta\xi^i.\overrightarrow{\mathbf{e}_h} \end{aligned}$$

$$= \sum_{i,j,h} \left(\Gamma_{ij}^h - \Gamma_{ji}^h\right) .\delta\xi^i \otimes \widehat{\delta}\xi^j.\overrightarrow{\mathbf{e}_h}$$

$$\underset{(4.8)}{=} \frac{1}{2} \sum_{i,j,h} \left(\Gamma_{ij}^h - \Gamma_{ji}^h\right) . \left(\delta\xi^i \otimes \widehat{\delta}\xi^j - \delta\xi^j \otimes \widehat{\delta}\xi^i\right) .\overrightarrow{\mathbf{e}_h}$$

$$= \sum_h \Theta^h.\mathbf{e}_h.$$

The infinitely small higher-order quantities that we neglected integrate to zero. Since $\overrightarrow{m_1' m_2'} = d\left(\bigtriangledown M\right)$, by [7.38], we have the relation

$$\boxed{\int_{\partial P} \bigtriangledown M = \int_P \sum_h \Theta^h.\overrightarrow{\mathbf{e}_h},} \qquad [7.59]$$

which is an infinitesimal translation.

(III) CURVATURE INTERPRETED AS A ROTATION We can proceed in the same way for the tangent vectors $\overrightarrow{\mathbf{e}_i}$. Thus:

$$\begin{aligned}
\widehat{\delta}\left(\delta\overrightarrow{\mathbf{e}_i}\right) - \delta\left(\widehat{\delta}\overrightarrow{\mathbf{e}_i}\right) &= \widehat{\delta}\left(\sum_l \omega_i^l\left(\delta\right).\overrightarrow{\mathbf{e}_l}\right) - \delta\left(\sum_l \omega_i^l\left(\widehat{\delta}\right).\overrightarrow{\mathbf{e}_l}\right)\\
&= \sum_l \widehat{\delta}\omega_i^l\left(\delta\right).\overrightarrow{\mathbf{e}_l} + \sum_k \omega_i^k\left(\delta\right) \otimes \widehat{\delta}\overrightarrow{\mathbf{e}_k}\\
&\quad - \sum_l \delta\omega_i^l\left(\widehat{\delta}\right).\overrightarrow{\mathbf{e}_l} - \sum_k \omega_i^k\left(\widehat{\delta}\right) \otimes \delta\overrightarrow{\mathbf{e}_k}\\
&= \sum_l \left[\widehat{\delta}\omega_i^l\left(\delta\right) - \delta\omega_i^l\left(\widehat{\delta}\right)\right].\overrightarrow{\mathbf{e}_l}\\
&\quad + \sum_{k,l} \left(\omega_i^k\left(\widehat{\delta}\right) \otimes \omega_k^l\left(\delta\right) - \omega_i^k\left(\delta\right) \otimes \omega_k^l\left(\widehat{\delta}\right)\right).\overrightarrow{\mathbf{e}_l}\\
&= \sum_l \sum_{j,k} \underbrace{\left(\partial_j\Gamma_{ki}^l - \partial_k\Gamma_{ji}^l + \sum_h \left(\Gamma_{ki}^h.\Gamma_{jh}^l - \Gamma_{ji}^h.\Gamma_{kh}^l\right)\right)}_{R_{i,jk}^l}\\
&\qquad\qquad .\widehat{\delta}\xi^j \otimes \delta\xi^k.\overrightarrow{\mathbf{e}_l}\\
&= \sum_h \mathbf{\Omega}_i^h.\overrightarrow{e_h},
\end{aligned}$$

which gives that $d\left(\nabla \overrightarrow{e_i}\right) = \sum_h \mathbf{\Omega}_i^h.\overrightarrow{e_h}$. By Stokes' formula,

$$\boxed{\int_{\partial P} \nabla \overrightarrow{e_i} = \int_P \sum_h \mathbf{\Omega}_i^h.\overrightarrow{e_h}.}$$

The $\int_P \mathbf{\Omega}_i^h$ $(i, j = 1, ..., n)$ are the components of an infinitesimal change-of-basis matrix (and more precisely an infinitesimal rotation if $\mathbf{G} = \mathrm{O}_n(\mathbb{R})$) and $\sum_i \mathbf{\Omega}_i^i$ is said to be the *homothety curvature* ([CAR 25](3), p. 394).

7.3.4. *Fundamental equalities*

(I) Structure equations By [7.55] and [7.51], the covariant exterior differential of a 1-form $\underline{\alpha}$ is $\nabla\underline{\alpha} = d\underline{\alpha} + \underline{\omega} \wedge \underline{\alpha}$. The first and second structure equations are as follows:

$$\boxed{\underline{\mathbf{\Omega}} = \nabla\underline{\omega}} \qquad [7.60]$$

and

$$\boxed{\underline{\mathbf{\Theta}} = \nabla\underline{\sigma}.} \qquad [7.61]$$

(II) Bianchi identities

First Bianchi identity By [7.61],

$$\begin{aligned} d\underline{\mathbf{\Theta}} &= d\left(d\underline{\sigma} + \underline{\omega} \wedge \underline{\sigma}\right) = d\left(\underline{\omega} \wedge \underline{\sigma}\right) \\ &= \left(d\underline{\omega}\right) \wedge \underline{\sigma} - \underline{\omega} \wedge \left(d\underline{\sigma}\right) \\ &= \left(\underline{\mathbf{\Omega}} - \underline{\omega} \wedge \underline{\omega}\right) \wedge \underline{\sigma} - \underline{\omega} \wedge \left(\underline{\mathbf{\Theta}} - \underline{\omega} \wedge \underline{\sigma}\right) \\ &= \underline{\mathbf{\Omega}} \wedge \underline{\sigma} - \underline{\omega} \wedge \underline{\mathbf{\Theta}}, \end{aligned}$$

which gives the first Bianchi identity:

$$\boxed{\nabla\underline{\mathbf{\Theta}} = \underline{\mathbf{\Omega}} \wedge \underline{\sigma}.} \qquad [7.62]$$

Second Bianchi identity By [7.55],

$$d\underline{\mathbf{\Omega}} = d\left[\underline{\omega}, \underline{\omega}\right] = d\underline{\omega} \wedge \underline{\omega} - \underline{\omega} \wedge d\underline{\omega} = \left(\underline{\mathbf{\Omega}} - \underline{\omega} \wedge \underline{\omega}\right) \wedge \underline{\omega} - \underline{\omega} \wedge \left(\underline{\mathbf{\Omega}} - \underline{\omega} \wedge \underline{\omega}\right).$$

We thus obtain the second Bianchi identity:

$$d\underline{\mathbf{\Omega}} = \underline{\mathbf{\Omega}} \wedge \underline{\omega} - \underline{\omega} \wedge \underline{\mathbf{\Omega}}.$$

We can express this identity in terms of the covariant exterior differential. Observe that:

$$\triangledown \underline{\mathbf{\Omega}} = \triangledown [\underline{\omega}, \underline{\omega}] \underset{(7.57)}{=} \triangledown \underline{\omega} \wedge \underline{\omega} - \underline{\omega} \wedge \triangledown \underline{\omega}.$$

Hence,

$$\boxed{\triangledown \underline{\mathbf{\Omega}} = \underline{\mathbf{\Omega}} \wedge \underline{\omega}.} \qquad [7.63]$$

Expressions in terms of the components We can express these identities as a function of the components $R^l_{i,jk}$ and T^k_{ij} of the curvature and torsion tensors (**exercise***: see [SPI 99], Volume 2, Chapter 5, Proposition 9) as follows:

$$\begin{aligned} R^i_{j,kl} + R^i_{k,lj} + R^i_{l,jk} &= \left(\nabla_j T^i_{kl} + \nabla_k T^i_{lj} + \nabla_l T^i_{jk}\right) \\ &\quad + \sum_\mu \left(T^\mu_{jk} T^i_{\mu l} + T^\mu_{kl} T^i_{\mu j} + T^\mu_{lj} T^i_{\mu k}\right), \\ \nabla_l R^h_{i,jk} + \nabla_j R^h_{i,kl} + \nabla_k R^h_{i,lj} &= -\sum_\mu \left(T^\mu_{jk} R^h_{i\mu l} + T^\mu_{kl} R^h_{i\mu j} + T^\mu_{lj} R^h_{i\mu k}\right). \end{aligned}$$

Note that these identities become much simpler in the case where the torsion is zero. In this case, they can be rewritten as follows for any vector fields X, Y, Z of class C^∞ on B:

$$\begin{aligned} &(\mathbf{r}.(X \wedge Y)).Z + (\mathbf{r}.(Y \wedge Z)).X + (\mathbf{r}.(Z \wedge X)).Y = 0, \\ &(\nabla_X.\mathbf{r})(Y \wedge Z) + (\nabla_X.\mathbf{r})(Z \wedge X) + (\nabla_Z.\mathbf{r})(X \wedge Y) = 0 \end{aligned}$$

(for the case of non-zero torsion, see ([DIE 93], Volume 3, section 20.6, Exercise 3)).

7.3.5. *Connection form on the bundle of* G*-frames*

The gauge potential is a 1-form on B. Below, we will find an expression for a 1-form ω *on the bundle of* $\mathbf{G}$*-frames* $R_{\mathbf{G}}(B)$ taking values in $\mathfrak{g}$ that is related to this potential. To do this, consider a local section $s : U \to R_{\mathbf{G}}$, where U is a non-empty open subset of B. If $v \in T(B)$ and $b \in U$, then $T_b(s).v \in T_{s(b)}(R)$. Thus,

$s^*(\omega) : (b, v) \mapsto \omega_{s(b)} (T_b (s) .v)$ is a 1-form on U taking values in $\mathfrak{g}$, which leads us to set

$$s^*(\omega) = \underline{\omega}. \qquad [7.64]$$

By [7.43], the 1-form ω must satisfy the relation

$$(s.\boldsymbol{g})^* (\omega) = \boldsymbol{g}^{-1}.s^* (\omega) .\boldsymbol{g} + \boldsymbol{g}^{-1}.d\boldsymbol{g}, \qquad [7.65]$$

where $s.\boldsymbol{g}$ is the section $b \mapsto s(b).\boldsymbol{g}$ and $\boldsymbol{g}^{-1}.s^*(\omega).\boldsymbol{g} = \mathrm{Ad}\left(\boldsymbol{g}^{-1}\right).s^*(\omega)$ (Definition 2.88).

DEFINITION 7.31.– *The form of a connection on the bundle of* $\mathbf{G}$*-frames* $R_{\mathbf{G}}(B)$ *is a 1-form of class* C^∞ *on* $R_{\mathbf{G}}$ *taking values in* $\mathfrak{g} = Lie(\mathbf{G})$ *that satisfies [7.65] for every local section* $s : U \to R_{\mathbf{G}}$ *(where* U *is an open subset of* B*).*

In section 3.5.2, we saw that the vertical tangent vectors (tangent to the fibers $\mathbf{G}_q, q \in R_{\mathbf{G}}$) are independent of the connection. Hence, if Z_X is the Killing field associated with $X \in \mathfrak{g}$ (section 6.4.1, Definition 6.69), then, for every $q \in R_{\mathbf{G}}$,

$$\boxed{\omega_q. (Z_X (q)) = X,} \qquad [7.66]$$

which implies that $\omega_q \big|_{V_q(R_{\mathbf{G}})}$ is an isomorphism from the space of vertical tangent vectors $V_q (R_{\mathbf{G}})$ onto $\mathfrak{g}$. In particular, $\omega_q : T_q (R_{\mathbf{G}}) \to \mathfrak{g}$ is surjective.

If $\boldsymbol{g}$ is a constant matrix, then, by [7.65],

$$\omega (q.\boldsymbol{g}) . (\mathbf{h}.\boldsymbol{g}) = \mathrm{Ad}\left(\boldsymbol{g}^{-1}\right) . (\omega (q) .\mathbf{h}) \qquad [7.67]$$

for every $\mathbf{h} \in T_q (R_{\mathbf{G}})$. The tangent linear mapping of $\rho_{\boldsymbol{g}} : q \longmapsto q.\boldsymbol{g}$ is $T_q (\rho_{\boldsymbol{g}}) : T_q (R_{\mathbf{G}}) \to T_{q.\boldsymbol{g}} (R_{\mathbf{G}}) : \mathbf{h} \mapsto \mathbf{h}.\boldsymbol{g}$[13], so [7.67] can alternatively be written as follows, setting $\rho_{\boldsymbol{g}*} = T_q (\rho_{\boldsymbol{g}})$:

$$\boxed{\omega (\rho_{\boldsymbol{g}*}.\mathbf{h}) = \mathrm{Ad}\left(\boldsymbol{g}^{-1}\right) .\omega (\mathbf{h}).} \qquad [7.68]$$

Since ω_q is surjective, $T_q (P) = H_q \oplus V_q (P)$, where $H_q := \ker (\omega_q)$ and, if $q = s (b)$, then $H_q \cong T_b (B)$ by Definition 3.54.

By [7.68], we have the so-called *equivariance* relation

$$\boxed{H_{q.\boldsymbol{g}} = \rho_{\boldsymbol{g}*}.H_q.} \qquad [7.69]$$

13 The mapping $\rho_{\boldsymbol{g}} : g \longmapsto q.\boldsymbol{g}$ acting on $R_{\mathbf{G}}$ is not the same as the right translation $\boldsymbol{\rho}(\boldsymbol{g})$ (section 2.4.1(**V**)), which acts on $\mathbf{G}$.

Finally, suppose that H is a *subbundle* (or a *contact distribution*) *of class* C^∞ of $T(R_{\mathbf{G}})$ (section 5.7.5**(II)**, Definition 5.78).

The elements of the tangent space $T_q(R)$ are the infinitesimal displacements that are possible for a $\mathbf{G}$-frame q (section 3.1**(II)**). The equivariance condition means that a change-of-basis matrix $\boldsymbol{g} \in \mathbf{G}$ acts globally on the whole fiber H_q over the origin $b = \pi(q)$, or in other words that the subbundle (of class C^∞) $H : q \mapsto H_q$ is stable under the action of $\mathbf{G}$.

7.3.6. *Principal connections and parallel transport*

(I) NOTION OF A PRINCIPAL CONNECTION We are now faced with the task of summarizing which of the above results remain valid when the bundle of $\mathbf{G}$-frames is replaced by an *arbitrary* principal bundle $(P, B, \mathbf{G}, \pi)$. The vertical tangent vectors are parallel to the fibers $\mathbf{G}_q$; ideally, we would like to define the vertical tangent vectors as parallel to the base B (section 3.1**(II)**). However, this is impossible because B is not a submanifold of P. Ehresmann's idea was to "lift everything up to the principal bundle", which leads to Lemma-Definition 7.37.

DEFINITION 7.32.– *1) We say that* $\mathbf{P}$ *is a* principal connection *on a principal bundle* $(P, B, \mathbf{G}, \pi)$ *if, for every* $q \in P$*, we have a subspace* H_q *of* $T_q(P)$ *that satisfies the following three conditions:*

i) $T_q(P) = V_q(P) \oplus H_q$*;*

ii) $\rho_{\boldsymbol{g}*}.H_q = H_{q.\boldsymbol{g}}$, $\forall \boldsymbol{g} \in \mathbf{G}$ *(equivariance condition of* H_q *under the action of* $\mathbf{G}$*);*

iii) H *is a* subbundle *of class* C^∞ *of* $T(P)$*. The elements of* H_q *are called* horizontal tangent vectors.

2) Given a principal connection $\mathbf{P}$ *on the principal bundle* $(P, B, \mathbf{G}, \pi)$*, the form of this connection is the 1-form* ω *on* P *taking values in* $\mathfrak{g}$ *that is uniquely determined by the following two conditions:*

i') $\omega_q((Z_X(q))) = X$ $(X \in \mathfrak{g}, q \in P)$*;*

ii') $\omega_q(Y_q) = 0$ $(Y_q \in H_q)$.

COROLLARY 7.33.– *Specifying a principal connection* $\mathbf{P}$ *on* P is equivalent to *specifying a connection 1-form* $\omega : P \to \mathfrak{g}$.

PROOF.– If $\mathbf{P}$ is given, we can determine ω by part (2) of Definition 7.32, and the equivariance condition [7.69] is automatically satisfied by (i). Conversely, if ω is given, we can determine $\mathbf{P}$ by setting $H_q = \ker(\omega_q)$, and ω is the connection 1-form determined by $\mathbf{P}$. ■

We can complete Figure 3.6 as shown in Figure 7.5. It is possible to show the following result (**exercise***: (see [DIE 93], Volume 4, (20.2.5))):

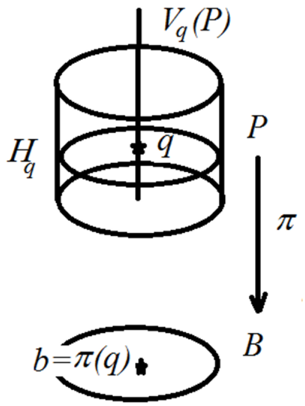


Figure 7.5. *Principal bundle (P), vertical tangent spaces $(V_q(P))$, and horizontal tangent spaces (H_q)*

THEOREM 7.34.– *A 1-form of class C^∞ on P taking values in $\mathfrak{g}$ is the 1-form of a connection on P if and only if it satisfies the following two conditions:*

1) ω, viewed as a mapping from $T(P)$ into $\mathfrak{g}$, is invariant for the right action of $\mathbf{G}$ on $T(P)$ induced by the action of $\mathbf{G}$ on P and for the right action $(\mathbf{u}, s) \mapsto \mathbf{u}.Ad(s)$ of $\mathbf{G}$ on $\mathfrak{g}$ (section 2.4.1, Definition 2.88). In other words, for every $\mathbf{h} \in T_q(P)$ and $s \in \mathbf{G}$, we have:

$$\omega(q.s).(\mathbf{h}.s) = Ad\left(s^{-1}\right).(\omega(q).\mathbf{h}). \qquad [7.70]$$

2) The equality [7.66] is satisfied (replacing $R_{\mathbf{G}}$ by P).

REMARK 7.35.– *The condition [7.70] can also be written as follows (see Definition 4.27):*

$$\rho_s^*(\omega) = Ad\left(s^{-1}.\omega\right), \qquad \forall s \in \mathbf{G}. \qquad [7.71]$$

(II) PARALLEL TRANSPORT Let $(P, B, \mathbf{G}, \pi)$ be a principal bundle on which a principal connection $\mathbf{P}$ is given. If $c : t \mapsto c(t)$ is a curve that is piecewise of class

C^1 in B, write b_1, b_2 for its first and last points; we can define a mapping $\varphi : \mathbf{G}_{b_1} \twoheadrightarrow \mathbf{G}_{b_2}$ as follows: given $q \in \mathbf{G}_{b_1}$, there exists a unique curve $\widehat{c_q}$ in P with initial point q satisfying the following two conditions ([KOB 69], Volume 1, Chapter 2, section 3):

(a) $\pi(\widehat{c_q}) = c$;

(b) every tangent vector of $\widehat{c_q}$ is horizontal.

The curve $\widehat{c_q}$ is the *horizontal lifting* of c (Definition 3.6) with initial point q. If p is the final point of $\widehat{c_q}$, set $\varphi(q) = p$. Then, $\widehat{c_{q.\boldsymbol{g}}} = \rho_{\boldsymbol{g}}(\widehat{c_q})$, so c_q commutes with the transformations induced by the elements of $\mathbf{G}$.

DEFINITION 7.36.– *The mapping* φ *is the* parallel transport *along* c.

This situation is represented graphically in Figure 7.6. Given a tangent vector $X_b \in T_b(B)$, where $b = \pi(q)$, let $\lambda_q(X_b)$ be the element of H_q uniquely determined by the condition $T_q(\pi)(\lambda_q(X_b)) = X_b$.

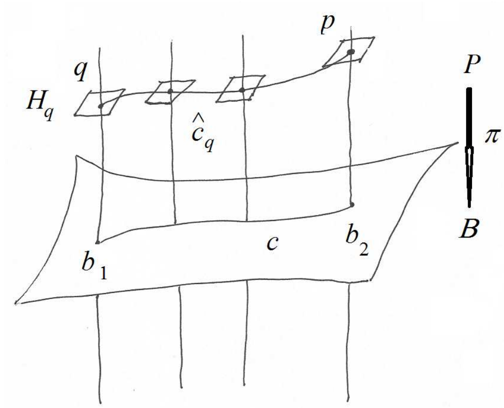


Figure 7.6. *Parallel transport in a principal bundle*

LEMMA-DEFINITION 7.37.– *1) The mapping* $\lambda_q : T_{\pi(q)}(B) \to H_q$ *is a linear bijection.*

2) This bijection is called the horizontal lifting *(considered at the point* q *in this case).*

PROOF.– Let $q \in P$. Then, $T_q(\pi) : T_q(P) \to T_{\pi(q)}(B)$ is a linear surjection and $V_q(P) = \ker(T_q(\pi))$ (section 3.5.2, Definition 3.54). Hence, $\mu_q = T_q(\pi)\big|_{H_q} : H_q \to T_{\pi(q)}(B)$ is a linear bijection and $\lambda_q = \mu_q^{-1}$. ■

REMARK 7.38.– *1) The* horizontal lifting *of a vector field $X \in \mathcal{T}_0^1(B)$ is the unique vector field $X^* \in \mathcal{T}_0^1(P)$ obtained by setting $X_q^* = \lambda_q\left(X_{\pi(q)}\right)$ for every $q \in P$. The mapping $X \longmapsto X^*$ is an $\mathbb{R}$-linear isomorphism that meets the objective described above at the start of* ***(I)****.*

2) The horizontal lifting of $f \in \mathcal{E}(B)$ is $f^* \in \mathcal{E}(P)$ such that $f^* = f \circ \pi$.

3) The horizontal component of $[X^*, Y^*]$ is the horizontal lifting of $[X, Y]$ ***(exercise)****.*

4) Consider the bundle of $\mathbf{G}$-frames $R_{\mathbf{G}}(B)$, where the base B is a pure n-dimensional manifold. Each frame $q = r_b \in R_{\mathbf{G}}(B)$ determines an isomorphism from $\mathbb{R}^n$ onto $T_{\pi(q)}(B)$, namely $e_i \mapsto q_i$, where $(e_i)_{1 \leq i \leq n}$ is the canonical basis of $\mathbb{R}^n$ and $q_i \in T_{\pi(q)}(B)$ is the i-th vector of q. We will write $q : \mathbb{R}^n \tilde{\rightarrow} T_{\pi(q)}(B)$ for this isomorphism.

7.3.7. *Covariant exterior differentiation on a principal bundle*

Consider again the remarks of sections 7.2.5 and 7.2.6**(I)**, this time working on the principal bundle $(P, B, \mathbf{G}, \pi)$ rather than the base B. The principal connection is denoted $\mathbf{P}$.

(I) HORIZONTAL AND VERTICAL FORMS Let α be a vector-valued differential n-form of class C^∞ taking values in a Banach space $\mathbf{F}$ (section 4.4.3**(VI)**): $\alpha \in \Omega^n(P; \mathbf{F})$.

DEFINITION 7.39.– *We say that α is* vertical *(respectively* horizontal*) if $\alpha(q).(\mathbf{h}_1 \wedge ... \wedge \mathbf{h}_n) = 0$ whenever one of the tangent vectors $\mathbf{h}_j \in T_q(P)$ is horizontal (respectively vertical).*

The 1-form ω is therefore vertical by definition. The same is true for $[\omega, \omega] := \omega \wedge \omega$.

(II) COVARIANT EXTERIOR DIFFERENTIATION For every $q \in P$, let η_q (respectively ν_q) be the projection $V_q(P) \oplus H_q \twoheadrightarrow H_q$ parallel to $V_q(P)$ (respectively the projection $V_q(P) \oplus H_q \twoheadrightarrow V_q(P)$ parallel to H_q). The mapping

$$(\mathbf{h}_1, ..., \mathbf{h}_n) \mapsto \alpha(q)(\eta_q.\mathbf{h}_1 \wedge ... \wedge \eta_q.\mathbf{h}_n)$$

belongs to $\mathrm{Alt}^n(P; \mathbf{F})$, so can be written as

$$(\mathbf{h}_1, ..., \mathbf{h}_n) \mapsto \alpha_{\mathbf{P}}(q)(\mathbf{h}_1 \wedge ... \wedge \mathbf{h}_n),$$

where $\alpha_{\mathbf{P}} : q \mapsto \alpha_{\mathbf{P}}(q)$ is a differential n-form on P of class C^∞ taking values in $\mathbf{F}$ that is clearly *horizontal*.

If α is *invariant* for the right action of $\mathbf{G}$ on $\bigwedge^n T(P)$ and a given right action of $\mathbf{G}$ on $\mathbf{F}$, i.e. for every $\boldsymbol{g} \in \mathbf{G}$,

$$\alpha\left(q.\boldsymbol{g}\right).\left(\mathbf{h}_1.\boldsymbol{g} \wedge ... \wedge \mathbf{h}_n.\boldsymbol{g}\right) = \alpha\left(q\right)\left(\mathbf{h}_1 \wedge ... \wedge \mathbf{h}_n\right).\boldsymbol{g},$$

then the same is true for $\alpha_{\mathbf{P}}$ (**exercise**).

DEFINITION 7.40.– *The covariant exterior differential of the differential n-form α relative to the principal connection $\mathbf{P}$ is the* horizontal *differential $(n+1)$-form (see section 5.5.1(**I**))*

$$\boxed{\mathbf{D}\alpha = (d\alpha)_{\mathbf{P}}.}$$

REMARK 7.41.– *If α is invariant in the sense specified above, then so is $\mathbf{D}\alpha$. If α is vertical and $n \geq 2$, then $\mathbf{D}\alpha = 0$.*

7.3.8. *Characterization of a* G*-connection*

Let B be a pure n-dimensional manifold and $R\left(B\right)$ its bundle of frames. Suppose that $\mathbf{G}$ is a closed subgroup of $\mathrm{GL}_n\left(\mathbb{R}\right)$. A $\mathbf{G}$-*structure* is a restriction $R_{\mathbf{G}}\left(B\right)$ of the principal bundle $R\left(B\right)$ to the group $\mathbf{G}$ (section 3.5.6(**II**)). From any given connection $\mathbf{P}$ on $R_{\mathbf{G}}\left(B\right)$, we can deduce a connection 1-form $\omega : P \to \mathfrak{g}$ in the same way as Corollary 7.33, and thus a gauge potential $\underline{\omega}$ on an arbitrary non-empty subset U of B by [7.64], as well as a linear connection $\mathbf{C}$ on B according to [7.45]. This connection $\mathbf{C}$ is said to be a $\mathbf{G}$-*connection* associated with the $\mathbf{G}$-structure of $R_{\mathbf{G}}\left(B\right)$ (the notion of $\mathbf{G}$-structure was introduced by Chern in 1953). A *moving* $\mathbf{G}$-*frame* is a section of class C^{∞} of $R_{\mathbf{G}}\left(B\right)$.

Conversely, every linear connection $\mathbf{C}$ on B uniquely determines a principal connection $\mathbf{P}$ on $R(B)$, again by Corollary 7.33. This connection $\mathbf{P}$ is fully determined if the space H_r of horizontal vectors is known at each point r of $R\left(B\right)$, giving the following result:

THEOREM 7.42.– $\mathbf{C}$ *is a* $\mathbf{G}$*-connection associated with the* $\mathbf{G}$*-structure of* $R_{\mathbf{G}}\left(B\right)$ *if and only if* H_r *is contained in the tangent space* $T_r\left(R_{\mathbf{G}}\left(B\right)\right)$ *for every frame* $r \in R_{\mathbf{G}}\left(B\right)$.

PROOF.– This condition is satisfied if and only if condition (iii) from Definition 7.32 is satisfied. ■

It follows that every differential manifold can be equipped with an $\mathrm{O}_n\left(\mathbb{R}\right)$-structure, or in other words, a Riemannian structure ([DIE 93], Volume 3, (20.7.13)). A differential manifold B can be equipped with an $\mathrm{SL}_n\left(\mathbb{R}\right)$-structure or an $\mathrm{SO}_n(\mathbb{R})$-structure if and only if it is orientable ([DIE 93], Volume 3, (20.7.5)). Furthermore, the next result follows from section 3.5.5 (**exercise**: see [COQ 02], section 3.4.3):

THEOREM 7.43.– *Let $\pi : P \to B$ be a principal bundle with structural group $\mathbf{G}$. The choice (when it exists) of a restriction of this bundle to a subbundle $\pi_Q : Q \to B$ with structural group $\mathbf{H}$, where $\mathbf{H}$ is a closed subgroup of $\mathbf{G}$, is not unique in general (section 3.5.6(**II**)). This choice is characterized by the choice of a global section on the fibration $P \times^{\mathbf{G}} \mathbf{G}/\mathbf{H}$ whose fibers are vector spaces isomorphic to the homogeneous space $\mathbf{G}/\mathbf{H}$ (section 3.5.5, Corollary 3.59).*

For example, we can pass from the structure of an arbitrary orientable n-dimensional differential manifold B to an oriented Riemannian manifold structure by reducing the bundle of frames $R(B)$ to the bundle of orthonormal frames $R_{\mathrm{SO}_n(\mathbb{R})}(B)$. This is equivalent to selecting from the set of all frames only the orthonormal frames, which transform via a change-of-basis matrix $\boldsymbol{g} \in \mathrm{SO}_n(\mathbb{R})$, thus preserving orientations. The dimension of the homogeneous space $\mathrm{GL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R})$ is

$$\dim(\mathrm{GL}_n(\mathbb{R})) - \dim(\mathrm{O}_n(\mathbb{R})) = n^2 - n(n-1)/2 = n(n+1)/2.$$

There are multiple possible reductions, corresponding to the multiple Riemannian metrics that can be chosen on B.

7.3.9. *Curvature and torsion forms of a principal connection*

(I)

DEFINITION 7.44.– *The* curvature form *of the principal connection $\mathbf{P}$ is the differential 2-form taking values in $\mathfrak{g}$ defined as the covariant exterior differential of the connection 1-form ω (Definition 7.32(2))*

$$\boxed{\mathbf{\Omega} = \mathbf{D}\omega.} \quad [7.72]$$

The connection $\mathbf{P}$ is said to be flat *if $\mathbf{\Omega} = 0$.*

REMARK 7.45.– *The forms ω and $\mathbf{\Omega}$ are invariant for the right action $(\mathbf{u}, \boldsymbol{g}) \mapsto \mathbf{u}.Ad(\boldsymbol{g})$ of $\mathbf{G}$ on $\mathfrak{g}$ (Theorem 7.34(1) and Remark 7.41).*

(II) IMPORTANT IDENTITIES As in section 7.3.4(**I**), we have the *first structure equation* (see [7.60]):

$$\boxed{\mathbf{\Omega} = d\omega + [\omega, \omega].} \quad [7.73]$$

The *second Bianchi identity* [7.63] can be stated as $\mathbf{D\Omega} = \mathbf{\Omega} \wedge \omega$. Since ω is vertical and $\mathbf{\Omega}$ is horizontal, $\mathbf{\Omega} \wedge \omega = 0$, so:

$$\boxed{\mathbf{D\Omega} = 0.}$$

(III) CASE OF THE BUNDLE OF $\mathbf{G}$-FRAMES If P is the bundle of $\mathbf{G}$-frames $R_{\mathbf{G}}$, U is a non-empty open subset of B, $\underline{\sigma}$ is a soldering form on U and $s : U \to R_{\mathbf{G}}$ is a local section, we can once again define the *soldering* 1*-form* σ on $R_{\mathbf{G}}$ such that $s^* (\sigma) = \underline{\sigma}$ (where $s^* (\sigma)$ is defined as the 1-form on U with components $s^* \left(\sigma^i\right)$), as well as the *torsion* 2*-form*, which satisfy the *second structure equation* (see [7.61]):

$$\mathbf{\Theta} := \mathbf{D}\sigma = d\sigma + \omega \wedge \sigma.$$

The first Bianchi identity [7.62] may now be written as

$$\boxed{\mathbf{D\Theta} = \mathbf{\Omega} \wedge \sigma.}$$

REMARK 7.46.– *i) We saw earlier that the 1-form ω is vertical (section 7.3.7(**I**)).*

*ii) The 1-form σ is horizontal (**exercise**), and so are the 2-forms $\mathbf{\Omega}$ and $\mathbf{\Theta}$ by Definition 7.40.*

*iii) The forms σ and $\mathbf{\Theta}$ are invariant under the canonical right action $(\boldsymbol{g}, \mathbf{u}) \mapsto \boldsymbol{g}^{-1}.\mathbf{u}$ of $\mathbf{G}$ on $\mathbb{R}^n$. Furthermore, σ is invariant under the right action of $\mathbf{G}$ on $R_{\mathbf{G}}(B)$ (**exercise**).*

*iv) Choosing the exterior product $\overline{\wedge}$ instead of $\wedge$ (section 4.2.4, Remark 4.11) modifies $\mathbf{D}$ and $\triangledown$ in the same way that it modifies d (section 5.5.1, Remark 5.23(2)). Thus, the two structure equations become $\mathbf{\Omega} = d\omega + \frac{1}{2}[\omega, \omega]$ and $\mathbf{\Theta} = d\sigma + \frac{1}{2}\omega \wedge \sigma$ ([KOB 69], Volume 1, p. 120) (**exercise**). Similar changes need to be made to the relation [7.74] stated below.*

(IV) CASE OF A LIE GROUP Let $\mathbf{G}$ be a Lie group, and view $\mathbf{G}$ as the principal bundle $\pi : \mathbf{G} \to \{b\}$, where $\{b\}$ is the manifold reduced to the single point b. There exists a *unique* principal connection on this principal bundle. Its 1-form ω taking values in $\mathfrak{g}$ is determined by the conditions:

1) $\omega_e = 1_{\mathfrak{g}}$;

2) the relation [7.71] is satisfied.

DEFINITION 7.47.– *This connection is known as the* Maurer–Cartan connection *of* $\mathbf{G}$*, and ω is said to be its* Maurer–Cartan form.

Maurer–Cartan's formula [5.26] reduces to:

$$\boxed{d\omega + [\omega, \omega] = 0 \Leftrightarrow \mathbf{D}\omega = 0,} \qquad [7.74]$$

so the Maurer–Cartan connection is flat.

(V) "Basic" vector fields Let ω be a connection 1-form on the bundle of $\mathbf{G}$-frames $\pi : R_{\mathbf{G}} \twoheadrightarrow B$, where the base B is an n-dimensional manifold. If $\xi \in \mathbb{R}^n$, we define the *"basic vector field"* $\mathfrak{b}(\xi)$ *corresponding to* ξ by the following condition: $\mathfrak{b}(\xi)_q$ is the unique horizontal vector at the point $q \in R_{\mathbf{G}}$ such that $\pi_*\left(\mathfrak{b}(\xi)_q\right) = q(\xi)$ (Remark 7.38(4)). Using the notation from section 7.3.7**(II)** we obtain the following:

Proposition 7.48.– *1)* $\sigma(\mathfrak{b}(\xi)) = \xi$.

2) $\rho_{\boldsymbol{g}*}(\mathfrak{b}(\xi)) = \mathfrak{b}\left(\boldsymbol{g}^{-1}.\xi\right)$, $\forall \boldsymbol{g} \in \mathbf{G}$ *(see section 7.3.5).*

3) If Y_1, Y_2 *are horizontal vector fields, then (see section 7.3.7**(II)**)*

$$\nu([Y_1, Y_2])(q) = -\boldsymbol{\Omega}(Y_{1q} \wedge Y_{2q}).q,$$

where $\Omega(Y_{1q} \wedge Y_{2q}).q$ *is the Killing field associated with* $\boldsymbol{\Omega}(Y_{1q} \wedge Y_{2q}) \in \mathfrak{g}$ *at the point q (section 6.4.1**(IV)**, Definition 6.69).*

4) If $\mathfrak{b}_1, \mathfrak{b}_2$ *are "basic vector fields", then*

$$\eta([\mathfrak{b}_1, \mathfrak{b}_2](q)) = \mathfrak{b}(-\Theta(\mathfrak{b}_{1q} \wedge \mathfrak{b}_{2q}))_q.$$

Proof.– 1), 2): **exercise**. 3): Since Y_1, Y_2 are horizontal and ω is vertical, the first structure equation implies that

$$\boldsymbol{\Omega}(Y_{1q} \wedge Y_{2q}) = d\omega(Y_1 \wedge Y_2)_q \underset{(5.26)}{=} -\omega([Y_1, Y_2])(q),$$

which implies (3) by Definition 7.32(2).

4) Since $\mathfrak{b}_1, \mathfrak{b}_2$ are horizontal, the second structure equation implies that

$$\begin{aligned}\boldsymbol{\Theta}(\mathfrak{b}_{1q} \wedge \mathfrak{b}_{2q}) &= d\sigma(\mathfrak{b}_1 \wedge \mathfrak{b}_2)(q) \\ &\underset{(5.26)}{=} \mathcal{L}_{\mathfrak{b}_1}.\sigma(\mathfrak{b}_2)(q) - \mathcal{L}_{\mathfrak{b}_2}.\sigma(\mathfrak{b}_1)(q) - \sigma([\mathfrak{b}_1, \mathfrak{b}_2])(q) \\ &\underset{\text{Remark 7.46(ii)}}{=} -\sigma(\eta([\mathfrak{b}_1, \mathfrak{b}_2]))(q).\end{aligned}$$

There exists a unique $\xi \in \mathbb{R}^n$ such that $\eta([\mathfrak{b}_1, \mathfrak{b}_2])(q) = \mathfrak{b}(\xi)(q)$, so $\mathfrak{b}(-\boldsymbol{\Theta}(\mathfrak{b}_{1q} \wedge \mathfrak{b}_{2q})) = \mathfrak{b}(\sigma(\eta([\mathfrak{b}_1, \mathfrak{b}_2])(q))) = \mathfrak{b}(\xi(q))$ by (1), which finally implies that $\mathfrak{b}(-\boldsymbol{\Theta}(\mathfrak{b}_{1q} \wedge \mathfrak{b}_{2q})) = \eta([\mathfrak{b}_1, \mathfrak{b}_2])(q)$. ■

7.3.10. *Cartan connections*

(I) $\mathbf{G}$-connections (section 7.3.8) are special cases of so-called *Cartan* connections [EHR 51, KOB 57]. Some Cartan connections are not $\mathbf{G}$-connections: most notably *projective* connections and *conformal* connections.

(II) For example, consider a projective connection. In the words of É. Cartan [CAR 24]:

"A projectively connected manifold (or space) is a numerical manifold that, in the immediate neighborhood of each point, has all the characteristics of a projective space, while also being equipped with a rule that connects the two small pieces surrounding two infinitely close points into a single projective space. [...] This rule defines the projective connection of the manifold. [...] Analytically, we arbitrarily choose a frame defining a system of projective coordinates in the projective space attached to each point $\mathbf{a}$ of the manifold. [...] The connection between the projective spaces attached to two infinitely close points $\mathbf{a}$ and $\mathbf{a}'$ is expressed analytically by a homographic transformation".

This citation is worth comparing with the remarks of section 7.1 on affine or linear connections.

We will once again use the notation from section 3.5.7**(III)**. It can be shown that the center $\mathbf{Z}'_{n+1}$ of $\mathrm{PSL}_{n+1}(\mathbb{R})$ is

$$\mathbf{Z}'_{n+1} = \begin{cases} I_{n+1} \text{ if } n+1 \text{ is odd} \\ \pm I_{n+1} \text{ if } n+1 \text{ is even} \end{cases}$$

The stabilizer ([P1, section 2.2.8**(I)**) of the origin $[(1,0,...,0)]$ of $\{1\} \times \mathbb{R}^n$ is

$$\mathbf{G}' = \left\{ \boldsymbol{g}' \in \mathrm{PSL}_{n+1}(\mathbb{R}) : \boldsymbol{g}' = \begin{pmatrix} z & p \\ 0 & a \end{pmatrix} \text{ mod.} \mathbf{Z}'_{n+1} \right\},$$

where the diagonal blocks z and a have dimensions 1×1 and $n \times n$, respectively. Indeed, we have:

$$[(1,0,...,0)] \, .\boldsymbol{g}' = [(z,0,...,0)] = [(1,0,...,0)] \, .$$

The mapping $\mathbf{G} \to \mathbf{P}_n : \boldsymbol{g} \mapsto [(1,0,...,0)] \, .\boldsymbol{g}$ therefore induces a diffeomorphism $\mathbf{G}/\mathbf{G}' \tilde{\to} \mathbf{P}_n$ by taking quotients, allowing us to identify $\mathbf{P}_n$ with the homogeneous space $\mathbf{G}/\mathbf{G}'$. According to Cartan, a projectively connected n-dimensional manifold B is a differential manifold whose tangent space at each point is isomorphic to the projective space $\mathbf{P}_n$. Any such manifold can be equipped with a "bundle of projective frames" (where each fiber is a base of $\mathbf{P}_n$) and is a principal bundle $\pi : P \twoheadrightarrow B$ with structural group $\mathbf{G} = \mathrm{PSL}_{n+1}(\mathbb{R})$.

(III) Now, let us generalize: let B be an n-dimensional differential manifold, $\mathbf{G}$ a Lie group and $\mathbf{G}'$ a closed subgroup of $\mathbf{G}$ (and hence a Lie subgroup by the Cartan–von Neumann theorem (Theorem 2.76)), and suppose that the homogeneous space $\mathbf{F} = \mathbf{G}/\mathbf{G}'$ is n-dimensional. Let $\pi : P \twoheadrightarrow B$ be a principal bundle with structural group $\mathbf{G}$ and consider the vector bundle $F = \left(P \times^{\mathbf{G}} \mathbf{F}, B, \pi_F\right)$ with fibers of type $\mathbf{F}$ and structural group $\mathbf{G}$ associated with P (Lemma-Definition 3.58(ii) and Corollary 3.59).

Next, reduce the structural group of P from $\mathbf{G}$ to $\mathbf{G}'$ (section 3.5.6(**II**)), write $\pi' : P' \twoheadrightarrow B$ for the reduced principal bundle thus obtained, and write $j : P' \hookrightarrow P$ for the canonical injection. Suppose that a principal connection $\mathbf{P}$ is given in P, and let ω be its 1-form taking values in $\mathfrak{g}$ (Definition 7.31(2)). The induced form $\omega' = j^*(\omega)$ (Definition 4.27, section 4.4.2(**I**)) is a differential 1-form on P' taking values in $\mathfrak{g}$.

DEFINITION 7.49.– *The differential 1-form ω' is called a* Cartan connection *on B with fiber $\mathbf{F}$ if, at every point $q' \in P'$, $\omega'_{q'}$ is a linear isomorphism $T_{q'}(P') \tilde{\rightarrow} \mathfrak{g}$.*

Equivalently, we can define ω' as follows: it is a 1-form on P' taking values in $\mathfrak{g} = \mathrm{Lie}(\mathbf{G})$ that satisfies the following conditions (see [7.71]):

1) $\omega'(Z_X) = X \quad (X \in \mathfrak{g}')$;

2) $\rho^*(\boldsymbol{g}').\omega' = \mathrm{Ad}\left(\boldsymbol{g}'^{-1}\right).\omega' \quad (\boldsymbol{g}' \in \mathbf{G}')$;

3) $\omega'_q : T_q(P') \rightarrow \mathfrak{g}$ is an isomorphism (Cartan's condition).

Unlike a principal connection in a principal bundle $\pi : P \twoheadrightarrow B$ with structural group $\mathbf{G}$, a Cartan connection on B taking values in $\mathfrak{g}$ is only equivariant under the action of $\mathbf{G}'$ (condition 2 above); condition 3 or Cartan's condition is expressed by saying that ω' is an *absolute parallelism* on P.

The curvature $\boldsymbol{\Omega}'$ of the Cartan connection (taking values in $\mathfrak{g}$) is given by:

$$\boldsymbol{\Omega}' = \mathbf{D}\omega' := d\omega' + [\omega', \omega']$$

(see [7.72] and [7.73]).

Consider the induced 1-form $\varpi' : T(P') \rightarrow \mathfrak{g}/\mathfrak{g}' : h' \mapsto \omega'(h') + \mathfrak{g}'$ for which the following diagram commutes:

$$\begin{array}{ccc} T(P') & \xrightarrow{\omega'} & \mathfrak{g} \\ & \overset{\varpi'}{\searrow} & \downarrow \mu \\ & & \mathfrak{g}/\mathfrak{g}' \end{array}$$

where μ is the canonical surjection. By condition 1, $\varpi'(Z_X) = 0$, so ϖ' is horizontal. Condition 2 implies that ϖ' is equivariant and defines a morphism of bundles $T(M) \rightarrow P \times^{\mathbf{G}'} \mathfrak{g}/\mathfrak{g}'$ (Definition 3.44) that is in fact an *isomorphism* by condition 3. The vector space $\mathfrak{g}/\mathfrak{g}'$ has dimension $n' = n - \dim(\mathbf{G}')$, so $\mathfrak{g}/\mathfrak{g}' \cong \mathbb{R}^{n'}$. The 1-form ϖ' satisfying all of these conditions is said to be a *soldering form*, which generalizes the definition given in section 7.3.9**(III)**.

The Cartan connection is said to be *reductive* if the Lie algebra $\mathfrak{g}$ is a *direct product* $\mathfrak{g} = \mathfrak{g}' \times \mathfrak{h}$ (section 5.4.1**(II)**), where $\mathfrak{h} \cong \mathfrak{g}/\mathfrak{g}'$. If so, we can uniquely write $\omega' = \omega_0 + \varpi'_0$, where ω_0 (respectively ϖ'_0) is a 1-form taking values in $\mathfrak{g}'$ (respectively $\mathfrak{h}$); we can identify ϖ'_0 with ϖ', and ω_0 is a principal connection on the bundle π'. The 1-form ϖ' is often interpreted as a field of frames (see **(IV)** below).

EXAMPLE 7.50.– *Let* $\mathbf{G}$ *be a Lie group and* $\mathbf{G}'$ *a closed subgroup of* $\mathbf{G}$*, and consider the principal bundle* $\pi = \mathbf{G} \times^{\mathbf{G}'} \mathbf{G}$*,* $\pi : \mathbf{G} \twoheadrightarrow \mathbf{G}/\mathbf{G}'$ *with structural group* $\mathbf{G}'$ *(Lemma-Definition 3.58). This principal bundle is trivializable (Lemma-Definition 3.53). The Maurer–Cartan form* ω *of* $\mathbf{G}$ *(Definition 7.47) can be uniquely extended to the connection 1-form of a Cartan connection (again written as* ω*) for* $\mathbf{G}$ *on* $\mathbf{G} \times^{\mathbf{G}'} \mathbf{G}$*. This connection, again said to be the* Maurer–Cartan connection*, is flat, and therefore satisfies the Maurer–Cartan formula [7.74].*

By definition, the *torsion* of the connection is $\bar{\mathbf{\Theta}}' = \mu(\mathbf{\Omega}')$ [AMB 53, TAK 55], ([SHA 97], Chapter 5, section 3). Thus, the Cartan connection is said to be *torsion-free* if (and only if) $\mathbf{\Omega}'$ takes values in $\mathfrak{g}'$.

In particular, we can define the torsion of a projective connection. In the words of É. Cartan [CAR 24]: "A projective manifold is said to be torsion-free if the infinitesimal displacement associated with an arbitrary, infinitely small, closed contour that starts from an arbitrary point $\mathbf{a}$ of this manifold and returns to this point leaves the point $\mathbf{a}$ (geometrically) invariant". This can be compared with the relation [7.59].

(IV) A Cartan connection can be conceptually illustrated as follows [WIS 10]: imagine a hamster in a ball that moves around on a torus B (Figure 7.7). The ball has symmetry group $\mathbf{G} = \mathrm{SO}_3(\mathbb{R})$. Let $\mathbf{G}' = \mathrm{SO}_2(\mathbb{R})$ be the stabilizer of the point of contact between the ball and the torus (which is a two-dimensional manifold). The hamster can orient itself within the ball according to the group $\mathbf{G}'$. A change of orientation of the hamster translates to a rotation (condition 1), but the direction taken by the hamster is not absolute, since it needs to be anchored to a frame (equivariance condition 2). The homogeneous space $\mathbf{G}/\mathbf{G}'$ (the fiber) is the ball (mathematically, the sphere $\mathbb{S}^2$). The Cartan connection takes values in $\mathfrak{g} = \mathfrak{so}(3)$ and the principal bundle $\pi' : P' \twoheadrightarrow B$ has structural group $\mathbf{G}'$. The configuration space of the hamster on the torus is the principal bundle π'. Any point of this space is a direction toward which the hamster can move, together with the position of the contact point. Mathematically, this is translated by the isomorphism $\mathfrak{so}_3(\mathbb{R}) \cong \mathfrak{so}_2(\mathbb{R}) \oplus \mathbb{R}^2$. If we assume that the ball rolls without slipping on the torus ("soldering"), the movement of the ball is fully determined by that of the hamster (condition 3). The connection taking values in $\mathfrak{so}_3(\mathbb{R})$ describes the "infinitesimal rotation" of the ball when the hamster moves by "infinitesimal displacements" belonging to $\mathfrak{g}' = \mathfrak{so}_2(\mathbb{R})$. This is a reductive Cartan connection (see **(III)**).

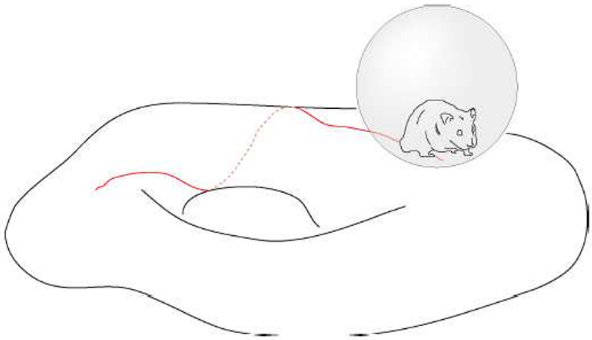


Figure 7.7. *The hamster in its ball. For a color version of this figure, see www.iste.co.uk/bourles/fundamentals3.zip*

7.4. Riemannian geometry

7.4.1. *Levi-Cività connection*

On a pure n-dimensional pseudo-Riemannian manifold equipped with a metric $\mathbf{g} = \sum_{i,j=1}^{n} g_{ij}.d\xi^i \otimes d\xi^j$ (section 4.5.1, [4.41]), we can define the Christoffel symbols $\Gamma_{ij}^k = \left\{ \begin{array}{c} k \\ ij \end{array} \right\}$ and $[ij, k]$ by the relations [7.15, 7.16, 7.17, 7.18]. Since $g_{ij} = g_{ji}$, these symbols are symmetric in (i, j). Hence, the connection thus obtained is torsion-free by [7.37]. The symbols $\left\{ \begin{array}{c} k \\ ij \end{array} \right\}$ determine a linear connection $\mathbf{C}$ on B (Definition 7.7) by the relation [7.22]

DEFINITION 7.51.– *The linear connection* $\mathbf{C}$ *above is called the Levi-Cività connection.*

THEOREM 7.52.– (Levi-Cività) *On any pseudo-Riemannian manifold* B*, there exists a unique torsion-free connection, the Levi-Cività connection.*

PROOF.– We have already established the existence of this connection constructively. Uniqueness can be shown as follows: let $\mathbf{C}'$ be a torsion-free linear connection on B, Γ_{ij}^k its coefficients in the local coordinates $\xi^1, ..., \xi^n$ and $[ij, k]$ the Riemann–Christoffel symbols of the first kind of the Levi-Cività connection.

The Lie derivative along $\frac{\partial}{\partial \xi^i}$ of the *scalar* field $\left\langle \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle$ coincides with the covariant derivative along $\frac{\partial}{\partial \xi^i}$, so

$$\frac{\partial g_{jk}}{\partial \xi^i} = \frac{\partial}{\partial \xi^i} \left\langle \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle = \nabla_{\frac{\partial}{\partial \xi^i}} \left\langle \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle$$
$$= \left\langle \nabla_{\frac{\partial}{\partial \xi^i}} \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle + \left\langle \frac{\partial}{\partial \xi^j} | \nabla_{\frac{\partial}{\partial \xi^i}} \frac{\partial}{\partial \xi^k} \right\rangle .$$

Since the connection $\mathbf{C}'$ is torsion-free, $\nabla_{\frac{\partial}{\partial \xi^i}} \frac{\partial}{\partial \xi^j} = \nabla_{\frac{\partial}{\partial \xi^j}} \frac{\partial}{\partial \xi^i}$. Applying a circular permutation to the indices, we therefore deduce from [7.17] that $\left\langle \nabla_{\frac{\partial}{\partial \xi^i}} \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle = [ij, k]$. Furthermore, by [7.23],

$$\left\langle \nabla_{\frac{\partial}{\partial \xi^i}} \frac{\partial}{\partial \xi^j} | \frac{\partial}{\partial \xi^k} \right\rangle = \sum_l \Gamma_{ij}^l . \underbrace{\left\langle \frac{\partial}{\partial \xi^k} | \frac{\partial}{\partial \xi^l} \right\rangle}_{g_{kl}},$$

which implies that $[ij, k] = \sum_l \Gamma_{ij}^l . g_{kl}$, so $\Gamma_{ij}^l = \sum_k g^{kl} . [ij, k] = \left\{ \begin{array}{c} l \\ ij \end{array} \right\}$. ■

LEMMA 7.53.– (Ricci) *Let B be a pseudo-Riemannian manifold,* $\mathbf{g}$ *its metric and* ∇ *the covariant exterior differential of a linear connection* $\mathbf{C}$ *on B (section 7.2.2). Then,* $\mathbf{C}$ *is a* $\mathbf{G}$*-connection if and only if* $\nabla \mathbf{g} = 0$*. In particular, this equality is satisfied when* $\mathbf{C}$ *is the Levi-Cività connection.*

PROOF.– We will show that $\nabla \mathbf{g} = 0$ if $\mathbf{C}$ is the Levi-Cività connection (for the other claims, see ([DIE 93], Volume 4, (20.9.5)). By [7.27],

$$\nabla_k g_{ij} = \frac{\partial g_{ij}}{\partial \xi^k} - \sum_p g_{pj} . \left\{ \begin{array}{c} p \\ ki \end{array} \right\} - \sum_q g_{iq} . \left\{ \begin{array}{c} q \\ kj \end{array} \right\}$$
$$= \frac{\partial g_{ij}}{\partial \xi^k} - [ik, j] - [jk, i]$$

as a result of [7.15]. Thus, $\nabla_k g_{ij} = 0$ by [7.17]. ■

7.4.2. *Geodesics*

(I) GEODESIC EQUATION If $c:[t_1,t_2]\to B$ is a curve of class C^1 in B (equipped with the Levi-Cività connection), then it has length s_2-s_1, where $s_i = s(t_i)$ $(i=1,2)$ and s is the *arclength* defined by:

$$\left(\frac{ds}{dt}\right)^2 = \sum_{i,j} g_{ij}.\frac{d\xi^i}{dt}.\frac{d\xi^j}{dt} \Leftrightarrow ds^2 \underset{\text{notation}}{=} \sum_{i,j} g_{ij}.d\xi^i.d\xi^j. \qquad [7.75]$$

THEOREM 7.54.– *In a pseudo-Riemannian manifold B, consider a curve $c:[t_1,t_2]\to B$ of class C^2 with endpoints s_1,s_2. The following conditions are equivalent:*

i) c is a geodesic.

ii) $\delta E(c)=0$, where E is the "energy"

$$E(\xi)=\int_{t_1}^{t_2}\sum_{i,j} g_{ij}.\dot{\xi}^i(t).\dot{\xi}^j(t).dt=\int_{t_1}^{t_2}\left(\frac{ds}{dt}\right)^2.dt.$$

iii) The curve c is traveled at constant speed and $\delta L(c)=0$, where L is the length:

$$L(\xi)=\int_{t_1}^{t_2}\sqrt{\sum_{i,j} g_{ij}.\dot{\xi}^i(t).\dot{\xi}^j(t)}.dt=\int_{s_1}^{s_2} ds.$$

iv) $$\delta\int_{s_1}^{s_2} ds = 0.$$

PROOF.– Let $\varphi=\sum_{i,j} g_{ij}.\dot{\xi}^i(t).\dot{\xi}^j(t)$. If c is traveled at constant speed (i.e. $ds/dt = const.$), then its parametrization is affine, i.e. $t=as+b$, $a\neq 0$. Therefore, $\varphi=a$ for $x=c$. The Euler–Lagrange equation for the stationarity of L is (section 1.3.2**(II)**)

$$\frac{1}{2\sqrt{\varphi}}\frac{\partial\varphi}{\partial\xi^i}=\frac{d}{dt}\left(\frac{1}{2\sqrt{\varphi}}.\frac{\partial\varphi}{\partial\dot{\xi}_i}\right),$$

so (ii) is equivalent to (iii) and to (iv). Now, write

$$\underbrace{\sum_{p,q} \frac{\partial g_{pq}}{\partial \xi^i} .\dot{\xi}^p.\dot{\xi}^q}_{\partial\varphi/\partial\xi^i} = \underbrace{2\sum_{p,q} \frac{\partial g_{ip}}{\partial \xi^q} .\dot{\xi}^p.\dot{\xi}^q + 2\sum_k g_{ik}.\ddot{\xi}^k}_{d/dt(\partial\varphi/\partial\dot{x}^i)}$$

$$= \sum_{p,q} \left(\frac{\partial g_{ip}}{\partial \xi^q} .\dot{\xi}^p.\dot{\xi}^q + \frac{\partial g_{iq}}{\partial \xi^p} .\dot{\xi}^p.\dot{\xi}^q \right) + 2\sum_k g_{ik}.\ddot{\xi}^k.$$

In light of [7.17], this is equivalent to $\sum_k g_{ik}.\ddot{\xi}^k + \sum_{p,q} \Gamma_{pq,i}.\dot{\xi}^p.\dot{\xi}^q = 0$, and by [7.15], setting $v = \dot{x}$, it is also equivalent to $\ddot{\xi}^k + \sum_{p,q} \Gamma^k_{pq}.\dot{\xi}^p.\dot{\xi}^q = 0$, or to $\nabla_v v = 0$, which shows that (ii) is equivalent to (i) by section 7.2.4**(III)**. ■

(II) MINIMAL GEODESICS ON A RIEMANNIAN MANIFOLD

DEFINITION 7.55.– *We say that a geodesic is* minimal *(respectively* locally minimal*) if it minimizes (respectively locally minimizes) L.*

A *minimal* geodesic, and similarly a *locally minimal* geodesic, can only exist in a *Riemannian* manifold, since this is the only type of manifold where the weak Legendre condition (Theorem 1.46) is satisfied (**exercise**).

In Euclidean space, every geodesic is minimal: the geodesics are the straight lines. On a sphere, the geodesics are arcs of great circles. An arc along the equator is only a minimal geodesic if it extends over at most 180° of longitude.

Consider the cylinder with parameters (θ, z), where $x = R.\cos\theta$, $y = R.\sin\theta$. On this cylinder, consider the curve $\theta = \theta(t)$, $z = z(t)$. Then, $ds^2 = dx^2 + dy^2 + dz^2 = \left(\dot{\theta}^2 + \dot{z}^2\right) dt^2$, so $\left(\frac{ds}{dt}\right)^2 = \dot{\theta}^2 + \dot{z}^2$. The Euler–Lagrange equation implies that $\ddot{\theta} = \ddot{z} = 0$, so $\theta = a.t + b$, $z = c.t + d$. The geodesics are therefore the circular helices, the circles obtained by cutting the cylinder with a horizontal plane and the generating lines of the cylinder.

7.4.3. *Flat pseudo-Riemannian manifolds*

THEOREM 7.56.– (Riemann) *Consider a pseudo-Riemannian manifold equipped with its Levi-Cività connection. This manifold is locally isometric to $\mathbb{R}^n$ equipped with its pseudo-Euclidean metric if and only if its Riemann–Christoffel curvature tensor is zero.*

PROOF.– (A) Necessary condition: the coefficients of the connection are given by [7.5].

1) Let us evaluate the quantity $\partial_j \Gamma^l_{ki} - \partial_k \Gamma^l_{ij}$; this can be written as

$$\underbrace{\sum_{\alpha} \frac{\partial^3 b^\alpha}{\partial y^j \partial y^k \partial y^i} \frac{\partial y^l}{\partial b^\alpha} + \sum_{\alpha} \frac{\partial^2 b^\alpha}{\partial y^k \partial y^i} \frac{\partial^2 y^l}{\partial b^\alpha \partial y^j}}_{L_1} - L_2,$$

where L_2 is obtained from L_1 by permuting the indices j and k. The first term of L_1 therefore cancels with the first term of L_2. The second term of L_1 is equal to

$$\sum_{\alpha} \frac{\partial^2 b^\alpha}{\partial y^k \partial y^i} \frac{\partial y^l}{\partial b^\alpha} \delta_{l,j}$$

and therefore cancels with the second term of L_2.

2) The quantity $\Gamma^h_{ki}.\Gamma^l_{jh}$ can be written as

$$\sum_{\alpha,\beta} \frac{\partial^2 b^\alpha}{\partial y^i \partial y^k} \frac{\partial y^h}{\partial b^\alpha} \frac{\partial^2 b^\beta}{\partial y^h \partial y^j} \frac{\partial y^l}{\partial b^\alpha},$$

an expression that is symmetric in the indices j, k. Hence, the same is true for $\sum_h \left(\Gamma^h_{ki}.\Gamma^l_{jh} - \Gamma^h_{ji}.\Gamma^l_{kh}\right)$, a quantity that must therefore be zero.

B) Sufficient condition:

1') If $\mathbf{\Omega} = 0$, Proposition 7.48(3) shows that the bracket of two horizontal vector fields is horizontal. Hence, the contact distribution H is involutive (section 5.7.5, Definition 5.81(3)), and therefore integrable by the Frobenius theorem (Theorem 5.87). At a point $b \in B$, choose an orthonormal frame $q \in \pi^{-1}(b)$ and let Φ be the leaf of H passing through q (section 5.7.5**(VII)**). This leaf Φ is locally the image of a section $s : U \to R_{\mathbf{G}}(B)$ ($\mathbf{G} = \mathrm{SO}_n(\mathbb{R})$) such that $s(b) = q$.

2') Consider the "basic vector fields" $\mathfrak{b}(e_i)$ $(i = 1, ..., n)$, where $(e_i)_{1 \leq i \leq n}$ is the canonical basis of $\mathbb{R}^n$. The $\mathfrak{b}(e_i)$ are horizontal, so each bracket $[\mathfrak{b}(e_i), \mathfrak{b}(e_j)]$ is horizontal (see (1')). But the Levi-Cività connection has zero torsion (Theorem 7.52). Hence, by Proposition 7.48(4), $\eta([\mathfrak{b}(e_i), \mathfrak{b}(e_j)]) = 0$, so $[\mathfrak{b}(e_i), \mathfrak{b}(e_j)] = 0$. Let $X_i(b) = \pi_*(\mathfrak{b}(e_i))(s(b))$; then $[X_i, X_j] = 0$ by the functor nature of the bracket (section 5.4.4**(I)**, Remark 5.18). Furthermore,

$$\pi_*(\mathfrak{b}(e_i)(s(b))) = s(b)(e_i) = q_i,$$

so $s = (X_1, ..., X_n)$, where $[X_i, X_j] = 0$. By Theorem 5.77, there exists a coordinate system $\xi^1, ..., \xi^n$ such that $X_i = \frac{\partial}{\partial \xi^i}$.

3’) To show that this coordinate system is as required, we still need to show that $s(b)$ is orthonormal for every b, or alternatively that $s(b_2)$ can be obtained by parallel transport from $q = s(b_1)$ along a curve $c : [0,1] \to U$, where U is an open subset of $\mathbb{R}^n$, $b_1 = c(0)$, $b_2 = c(1)$. However, we can perform a parallel transport along c from q by choosing a lifting $\widehat{c}$ of c such that $\widehat{c}(0) = q$, in which case the point transported from q is now $\widehat{c}(1)$. It is clear that $\widehat{c} = s \circ c$ is an appropriate lifting to do this. ■

7.4.4. *Ricci tensor and Einstein tensor*

These two tensors (or more precisely tensor fields) play a fundamental role in general relativity. Consider an n-dimensional pseudo-Riemannian manifold B equipped with its Levi-Cività connection.

(I) RICCI TENSOR Consider the Riemann–Christoffel tensor $\mathbf{r}$ with components $R^l_{i,jk}$. Contracting (section 4.2.1**(II)**) the indices l, j yields the twice covariant antisymmetric tensor with components

$$R_{ik} = \sum_j R^j_{i,jk}.$$

This tensor is symmetric by [7.36] and [7.19]. It is called the *Ricci tensor*, denoted **Ric**. We can write this tensor in mixed form (applying the rule **(R)** from section 5.6.5**(II)**) as follows:

$$R^j_i = \sum_k g^{jk}.R_{ik}.$$

The invariant

$$R = \sum_{i,k} R^i_i$$

is called the *Ricci curvature*. In two dimensions, it can be shown that $R = 2K$, where K is the *Gaussian curvature*, which is equal to $\frac{1}{r^2}$ for a sphere of radius $r > 0$[14].

Furthermore, by lowering the contravariant index l of the Riemann–Christoffel $\mathbf{r}$ (again applying the rule (**R**)), we obtain the tensor of type $(0,4)$ whose components $R_{li,jk}$ can be expressed in terms of the Riemann–Christoffel symbol of the first kind according to the following relation, deduced from [7.36]:

$$R_{li,jk} = \partial_j \Gamma_{lk,i} - \partial_k \Gamma_{lj,i} + \sum_h \left(\Gamma^h_{ki}.\Gamma_{lj,h} - \Gamma^h_{ji}.\Gamma_{lk,h} \right).$$

14 See the Wikipedia articles on *Scalar curvature* and *Gaussian curvature*.

It can easily be checked (**exercise**) that $R_{li,jk}$ satisfies the following relations (and no others):

$$R_{li,jk} = R_{jk;li} = -R_{il,jk} = -R_{li,kj},$$

$$R_{li,jk} + R_{lj,ki} + R_{lk,ij} = 0.$$

Therefore, $R_{li,jk}$ has $\frac{n^2(n^2-1)}{12}$ independent components, and R_{ik} has $\frac{n(n+1)}{2}$. The Riemann–Christoffel tensor $\mathbf{r}$ is therefore fully determined by the Ricci tensor $\mathbf{Ric}$ if and only if $\frac{n^2(n^2-1)}{12} \leq \frac{n(n+1)}{2}$, i.e. $n \leq 3$.

DEFINITION 7.57.– *The pseudo-Riemannian manifold is said to be* flat *if* $\mathbf{r} = 0$ *(see Definition 7.44) and* Ricci-flat *if* $\mathbf{Ric} = 0$.

The above shows that every flat pseudo-Riemannian manifold is Ricci-flat, but the converse only holds if the dimension is ≤ 3.

(II) EINSTEIN TENSOR When the torsion is zero, the second Bianchi identity (section 7.3.4(**II**)) is given by:

$$\nabla_l R^h_{i,jk} + \nabla_j R^h_{i,kl} + \nabla_k R^h_{i,lj} = 0.$$

Suppose that $k = h$. The antisymmetry in (i, k) of $R^h_{i,jk}$ implies that:

$$-\nabla_l R_{ij} + \nabla_j R_{il} + \sum_h \nabla_h R^h_{i,jl} = 0.$$

By multiplying by g^j_i and contracting the indices i, j, we deduce that:

$$-\nabla_l R + \sum_j R^j_l + \sum_h R^h_l = 0,$$

and by Ricci's lemma (Lemma 7.53), it follows that

$$\sum_i \nabla_i S^i_k = 0, \qquad [7.76]$$

where

$$\boxed{S_{ik} := R_{ik} - \frac{1}{2} g_{ik}.R.}$$

The tensor field $\mathbf{S}$ with components S_{ik} is called the *Einstein tensor*. By section 5.5.2**(II)**, [7.76] can be rewritten as a conservation relation (Remark 5.29)

$$\boxed{\mathrm{div}\,(\mathbf{S}) = 0,}$$

where div $(\mathbf{S})$ is the tensor of type $(0,1)$ whose k-th component is div $(\mathbf{S}_k)$.

É. Cartan showed the following result in 1922 ([CAR 22b], section 40):

THEOREM 7.58.– *On an Einstein manifold (section 4.5.1), consider a twice covariant symmetric tensor field* $\mathbf{Q}$ *whose components* Q_{ik} *only depend on the* g_{jl} *and their partial derivatives of first and second orders, with a linear dependence on the second-order partial derivatives. The following conditions are equivalent:*

i) div $(\mathbf{Q}) = 0$*;*

ii) the Q_{ik} *are of the form*

$$Q_{ik} = \mu\,(S_{ik} + \Lambda.g_{ik})\,,$$

where μ *and* Λ *are constants.*

Einstein's equation can be stated as follows[15]:

$$S_{ik} + \Lambda.g_{ik} = \chi.T_{ik}, \qquad [7.77]$$

where Λ is the cosmological constant, $\chi > 0$ is a constant and the T_{ik} are the components of the momentum-energy tensor $\mathbf{T}$; the latter is conservative, so div $(\mathbf{T}) = 0$; it is a symmetric tensor of type $(0,2)$ that characterizes the energy distribution[16]. The cosmological constant Λ (introduced at a very late stage by Einstein who then abandoned the idea) is necessarily zero if the Einstein manifold is required to be Ricci-flat when $\mathbf{T} = 0$. However, as we saw above, this does not necessarily imply that this manifold must be flat (we could imagine an empty – and hence Ricci-flat – Einstein manifold traversed by a gravitational wave that imprints a curvature upon it).

(III) NEWTONIAN APPROXIMATION To understand the Einstein equation [7.77], it is useful to consider its Newtonian approximation. When $\Lambda = 0$, by the rule **(R)** from section 5.6.3**(II)**), [7.77] is equivalent to:

$$R_i^j - \frac{1}{2}\delta_i^j.R = \chi.T_i^j. \qquad [7.78]$$

15 Einstein announced this equation on November 25, 1915 at the Prussian Academy of Sciences in Berlin and published it in an article dated 1916 [EIN 52c]. Hilbert (after lengthy discussions with Einstein) had presented an equivalent formulation based on a variational principle to the Royal Academy of Sciences Göttingen on November 20, 1915 ([SAU 09], Chapter 1, p. 28–46), without however claiming priority ([PAI 82], section 14d).

16 See the Wikipedia articles on the *Cosmological constant* and the *Energy-momentum tensor*.

(α) **CASE OF A REGION Ω WITHOUT MATTER** In any such region, $T_{ik} = 0$, so [7.77] becomes $S_{ik} = 0$. Suppose that the gravitation field is sufficiently weak in Ω to ensure that $g_{ij} = \eta_{ij} + h_{ij}$, where the η_{ij} are the normal components of the metric tensor of Minkowski space-time[17], i.e. $(\eta_{ij}) = \text{diag}\,(1,-1,-1,-1)$, and where $|h_{ij}| \ll 1$. In Ω, consider a particle with coordinates x^i $(0 \leq i \leq 4)$, where $x^0 = ct$, and assume that its speed is small relative to light, so that

$$\left|\frac{dx^i}{ds}\right| \ll 1 \ (1 \leq i \leq 3)\,.$$

With these assumptions, $\left(ds/dx^0\right)^2 = \sum_i g_{ij}\left(dx^i/dx_0\right)\left(dx^j/dx_0\right) \simeq g_{00}$ by [7.75], and, by [7.18],

$$\Gamma^p_{ij} \simeq \begin{cases} -\frac{1}{2}\sum_k g^{pk}\partial_k g_{00} \underset{\textbf{(R)}}{=} -\frac{1}{2}\partial_p g_{00} \text{ if } (i,j) = (0,0) \\ \qquad\qquad = 0 \text{ otherwise} \end{cases} \qquad [7.79]$$

where $\partial_i := \partial/\partial x_i$. Therefore,

$$\Gamma^p_{00}\left(\frac{\partial x^0}{\partial s}\right)^2 \simeq -\frac{1}{2g_{00}}\partial_p g_{00}.$$

Furthermore, $\frac{d^2x^p}{ds^2} \simeq \frac{d^2x^p}{(dx_0)^2} = \frac{1}{c^2}\frac{d^2x^p}{(dt)^2}$, so, for $1 \leq p \leq 3$, the geodesic equation [7.30] becomes:

$$\frac{d^2x^p}{dt^2} = -\frac{c^2}{2}\partial_p g_{00}.$$

This is Newton's equation $\frac{d^2x}{dt^2} = \overrightarrow{E}$, $\overrightarrow{E} = -\overrightarrow{\nabla}U$ with $U = \frac{c^2}{2}h_{00} + const.$, or, up to an additive constant:

$$g_{00} = 1 + \frac{2U}{c^2}.$$

REMARK 7.59.– *Questions of priority regarding the theory of relativity have been under discussion for many years. Lorentz ([LOR 21], p. 298) said in 1914 that Poincaré "has obtained a perfect invariance of the equations of electrodynamics [i.e. the Lorentz equations] and formulated the 'special relativity postulate', an expression that he was the first to use." This postulate was stated in 1904 as follows [POI 04]: "The laws of physical phenomena must be the same for a stationary*

17 This could be called *Poincaré–Minkowski space-time*, since [POI 06] preceded [MIN 52]. See Remark 7.59 below.

observer as for one carried along in a uniform motion of translation." Poincaré, in his Note de Comptes Rendus at the French Academy of Sciences of June 5, 1905 [POI 06], corrected the Lorentz transformations and showed that they form a group under which the quadratic form with coefficients η_{ij} is invariant. Einstein formulated the principle of relativity as follows in his famous article [EIN 52a], submitted on June 30, 1905: "The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems of coordinates in uniform translatory motion." In order to derive the Lorentz equations as rectified by Poincaré (of which Einstein later claimed to be unaware when he wrote his first article in 1905; his rigorous proof eventually appeared in [EIN 07]), Einstein added the principle of independence of the speed c of light with respect to the motion of the source, and specified in [EIN 52b] that this second principle is a consequence of Maxwell's equations. It was shown in 1968 that the principle of relativity and the hypotheses of the homogeneity, isotropy and causality of space-time are necessary and sufficient (without the postulate of invariance of c) to establish the kinematic equations of special relativity (i.e. the Lorentz–Poincaré equations) if they are assumed to form a group (see [LÉV 76]).

(β) **MOMENTUM-ENERGY TENSOR DUE TO A MACROSCOPIC SYSTEM OF PARTICLES IN MOTION** If ρ is the specific mass of the system (the sum of the mass per unit volume of each particle at rest), this symmetric tensor of type $(2, 0)$ has the contravariant components $T^{ij} = \rho c^2 u^i u^j$, where $u^i := dx^i/ds$ $(0 \leq i \leq 4)$. If $\left|u^i\right| \ll 1$ for $1 \leq i \leq 3$, then $T^{ij} \simeq 0$ if $(i, j) \neq (0, 0)$ and $T^{00} = T^0_0 = \rho c^2$; this is the famous Poincaré–Einstein equation[18] $E = mc^2$ considered per unit volume. The approximations of (α) remain valid and show that, for $p \geq 1$, $\Gamma^p_{00} \simeq -\frac{1}{c^2}\partial_p U$, $R^0_0 \simeq -\sum_{p\geq 1} \partial_p \Gamma^p_{00} = \frac{1}{c^2}\overrightarrow{\nabla}^2 U$ and $R^j_i \simeq 0$ if $(i, j) \neq (0, 0)$. By index contraction, Einstein's equation [7.78] implies that

$$\frac{1}{2} R^0_0 = \chi T^0_0 \Rightarrow \overrightarrow{\nabla}^2 U = 2c^4 \chi \rho,$$

and we recover Poisson's formula ([5.48], section 5.6.4) by setting $\chi = 8\pi G/c^4$, where G is the universal constant of gravitation, as done in ([EIN 52c], equation 69).

18 See ([POI 00b], p. 260), ([EIN 52b], p. 641) and [IVE 52] for a comparison of these two articles.

This page intentionally left blank

References

[ABR 63] ABRAHAM R., *Lectures of Smale on Differential Topology*, Notes at Columbia University, New York, 1963.

[ABR 83] ABRAHAM R., MARSDEN J.E., RATIU T., *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley, 1983.

[ADA 03] ADAMS R.A., FOURNIER J.F., *Sobolev Spaces*, Elsevier, 2003.

[ALE 87] ALEXEEV V., TIKHOMIROV V., FOMINE S., *Optimal Control*, Plenum Publishing Corporation, 1987.

[AMB 53] AMBROSE W., SINGER I.M., "A theorem on holonomy", *Transactions of the American Mathematical Society*, vol. 75, no. 3, pp. 428–443, 1953.

[BIN 13] BINCER A.M., *Lie Groups and Algebras – A Physicist's Perspective*, Oxford University Press, 2013.

[BIR 36] BIRKHOFF G., "Lie groups simply isomorphic with no linear group", *Bulletin of the American Mathematical Society*, vol. 42, pp. 883–888, 1936.

[BIR 38] BIRKHOFF G., "Analytical groups", *Transactions of the American Mathematical Society*, vol. 43, pp. 61–101, 1938.

[BIS 68] BISHOP R.S., GOLDBERG S.I., *Tensor Analysis on Manifolds*, MacMillan, 1968.

[BOC 59] BOCHNER S., *Lectures on Fourier Integrals*, Princeton University Press, 1959.

[BOR 01] BOREL A., *Essays in the History of Lie Groups and Algebraic Groups*, American Mathematical Society, 2001.

[BOU 67] BOURBAKI N., *Théories Spectrales*, Hermann, 1967.

[BOU 69] BOURBAKI N., *Integration*, Springer, 1969.

[BOU 74] BOURBAKI N., *General Topology*, Springer, 1974.

[BOU 76] BOURBAKI N., *Functions of a Real Variable*, Springer, 1976.

[BOU 81] BOURBAKI N., *Topological Vector Spaces*, Springer, Masson, 1981.

[BOU 82a] BOURBAKI N., *Variétés différentielles et analytiques – Fascicule de résultats*, Hermann, 1982.

[BOU 82b] BOURBAKI N., *Lie Groups and Lie Algebras*, Springer, 1982.

[BOU 10] BOURLÈS H., *Linear Systems*, ISTE Ltd, London and John Wiley & Sons, New York, 2010.

[BOU 12] BOURBAKI N., *Algebra*, Springer, 2012.

[BOU 16] BOURBAKI N., *Topologie algébrique, Chapters 1–4*, Springer, 2016.

[BRU 56] BRUHAT F., "Sur les représentations induites des groupes de Lie", *Bulletin de la Société Mathématique de France*, vol. 84, pp. 97–205, 1956.

[BRU 61] BRUHAT F., "Distributions sur un groupe localement compact et applications à l'étude des représentations des groupes $\mathfrak{p}$-adiques", *Bulletin de la Société Mathématique de France*, vol. 89, pp. 43–75, 1961.

[CAR 22a] CARTAN É., *Leçons sur les invariants intégraux*, Hermann, 1922.

[CAR 22b] CARTAN É., *Les équations de la gravitation d'Einstein*, Gauthier-Villars, 1922.

[CAR 23] CARTAN É., "Les variétés à connection conforme", *Annales de la Société Polonaise de Mathématiques*, vol. 2, pp. 171–221, 1923.

[CAR 24] CARTAN É., "Sur les variétés à connexion projective", *Bulletin de la Société Mathématique de France*, vol. 52, pp. 205–241, 1924.

[CAR 25] CARTAN É., "Sur les variétés à connexion affine et la théorie de la relativité généralisée" (I) Première partie : *Annales scientifiques de l'É.N.S.*, 3rd series, vol. 40, pp. 325–412, 1923; (II) Première partie (suite) : *Annales scientifiques de l'É.N.S.*, 3rd series, vol. 41, pp. 1–25, 1924; (III) Deuxième partie : *Annales scientifiques de l'É.N.S.*, 3rd series, vol. 42, pp. 17–88, 1925.

[CAR 26] CARTAN É., "Les groupes d'holonomie des espaces généralisés", *Acta Mathematica*, vol. 48, nos 1–2, pp. 1–42, 1926.

[CAR 27] CARATHÉODORY C., *Vorlesungen über reelle Funktionen*, 2nd edition, Springer, 1927.

[CAR 35] CARTAN É., *La méthode du repère mobile, la théorie des groupes continus et les espaces généralisés*, Hermann, 1935.

[CAR 45] CARTAN É., *Les systèmes différentiels extérieurs et leurs applications géométriques*, Hermann, 1945.

[CAR 51] CARTAN É., *Leçons sur la géométrie des espaces de Riemann*, 2nd edition, Gauthiers-Villars, 1951.

[CAR 52] CARTAN É., *La théorie des groupes finis et continus et l'analysis situs*, Gauthier-Villars, new edition, 1952.

[CAR 66] CARLESON L., "On convergence and growth of partial sums of Fourier series", *Acta Mathematica*, vol. 116, pp. 135–157, 1966.

[CAR 70] CARTAN H., *Differential Forms*, Hermann, 1970.

[CAR 99] CARTAN É., "Sur certaines expressions différentielles et le problème de Pfaff", *Annales de l'É.N.S.*, vol. 16, pp. 239–332, 1899.

[CHE 46] CHEVALLEY C., *Theory of Lie Groups I*, Princeton University Press, 1946.

[CHI 01] CHIRIKJIAN G.S., KYATKIN A.B., *Engineering Applications of Noncommutative Harmonic Analysis*, CRC Press, 2001.

[CHR 69] CHRISTOFFEL E.B., "Über die Transformation der homogenen Differentialausdrücke zweiten Grades", *Journal für die Reine und Angewandte Mathematik* (*Journal de Crelles)*, vol. 70, no. 1, pp. 46–70, 1869.

[CLA 90] CLARKE F.H., *Optimization and Nonsmooth Analysis*, SIAM, 1990.

[COH 65] COHN P.M., *Lie Groups*, Cambridge University Press, 1965.

[CON 95] CONNES A., *Noncommutative Geometry*, Academic Press, 1995.

[COQ 02] COQUEREAUX R., Espaces fibrés et connexions – Une introduction aux géométries classiques et quantiques de la physique théorique, Centre de Physique Théorique de Luminy-Marseille, 2002.

[COU 58] COURANT R., HILBERT D., *Methods of Mathematical Physics*, vols 1–2, Interscience Publishers, 1958.

[DAR 13] DARBOUX G., *Leçons sur la théorie générale des surfaces*, vols 1–4, Gauthier-Villars, 2nd edition, 1913.

[DEI 77] DEIMLING K., *Ordinary Differential Equations in Banach Spaces*, Springer-Verlag, 1977.

[DER 84] DE RHAM G., *Differential Manifolds*, Springer, 1984.

[DES 19] DE SAINT GERVAIS H.P., "Analysis situs". Available at: http://analysis-situs.math.cnrs.fr/, 2019.

[DIE 63] DIEUDONNÉ J., *La géométrie des groupes classiques*, Springer-Verlag, 1963.

[DIE 73] DIEUDONNÉ J., *Sur les groupes classiques*, 3rd edition, Hermann, 1973.

[DIE 89] DIEUDONNÉ J., *A History of Algebraic and Differential Topology, 1900–1960*, Birkhäuser, 1989.

[DIE 93] DIEUDONNÉ J., *Treatise on Analysis*, Academic Press, 1993.

[DIX 81] DIXMIER J., *Von Neumann Algebras*, North-Holland Publishing Company, 1981.

[EEL 58] EELLS J., "On the geometry of function spaces", *Symposium International Topological Algebraic*, Mexico, pp. 303–308, 1958.

[EEL 66] EELLS J., "A setting for global analysis", *Bulletin of the American Mathematical Society*, vol. 72, no. 5, pp. 751–807, 1966.

[EEL 70] EELLS J., ELWORTHY D., "On the differential topology of hilbertian manifolds", in CHERN S.S., SMALE S. (eds), *Global Analysis*, Part 2, American Mathematical Society, pp. 41–44, 1970.

[EHR 51] EHRESMANN C., "Les connexions dans un espace fibré différentiable", *Colloque de Topologie*, pp. 29–55, George Thone (Liège), 1950; Masson (Paris), 1951.

[EIN 52a] EINSTEIN A, "Zur Elektrodynamik bewegter Körper", *Annalen der Physik*, vol. 17, no. 10, pp. 891–921, 1905. English translation: A. Einstein, H.A. Lorentz, H. Weyl, H. Minkowski, *The Principle of Relativity*, Chapter 3, pp. 35–65, Dover, 1952.

[EIN 52b] EINSTEIN A., "Ist die Trägheiteines Körpers son seinem Energiegehalt abhängig?", *Annalen der Physik*, vol. 18, no. 13, pp. 639–641, 1905. English translation: A. Einstein, H.A. Lorentz, H. Weyl, H. Minkowski, *The Principle of Relativity*, Chapter 4, pp. 67–71, Dover, 1952.

[EIN 52c] EINSTEIN A., "Die Grundlagen der allgemeinen Relativitätstheorie", *Annalen der Physik*, vol. 49, no. 7, pp. 769–822, 1916. English translation: A. Einstein, H.A. Lorentz, H. Weyl, H. Minkowski, *The Principle of Relativity*, Chapter 7, pp. 109–164, Dover, 1952.

[EIN 54] EINSTEIN A., *The Meaning of Relativity*, 6th edition, Princeton University Press, 1954.

[EIN 07] EINSTEIN A., "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen", *Jahrbuch der Radioaktivität und Elektronik*, vol. 4, pp. 411–461, 1907.

[FLA 98] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Elsevier, 1998.

[FLI 97] FLIESS M., LÉVINE J., MARTIN P., ROUCHON P., Deux applications de la géométrie locale des diffiétés, *Annales de l'Institut Poincaré*, section A, vol. 66, no. 3, pp. 275–292, 1997.

[GEL 63] GELFAND I.M., FOMIN S.V., *Calculus of Variations*, Prentice-Hall, 1963.

[GEO 82] GEORGI H., *Lie Algebras in Particle Physics*, Addison-Wesley, 1982.

[GIL 98] GILBARG D., TRUDINGER N.S., *Elliptic Partial Differential Equations of the Second Order*, Springer, 1998.

[GLÖ 06] GLÖCKNER H., "Implicit functions from topological vector spaces to Banach spaces", *Israël Journal of Mathematics*, vol. 155, pp. 205–252, 2006.

[GOD 15] GODEMENT R., *Analysis IV – Integration and Spectral Theory, Harmonic Analysis, the Garden of Modular Delights*, Springer, 2015.

[GOD 17] GODEMENT R., *Introduction to the Theory of Lie Groups*, Springer, 2017.

[HAL 77] HALE J., *Theory of Functional Differential Equations*, Springer, 1977.

[HAL 80] HALL A.R., *Philosophers at War*, Cambridge University Press, 1980.

[HAR 51] HARISH-CHANDRA, "On some applications of the universal enveloping algebra of a semisimple lie algebra", *Transactions of the American Mathematical Society*, vol. 70, pp. 28–99, 1951.

[HAR 66] HARISH-CHANDRA, "Discrete series for semisimple Lie groups", *I: Acta Mathematica*, vol. 113, pp. 241–318, 1965 and II vol. 116, pp. 1–111, 1966.

[HAT 02] HATCHER A., *Algebraic Topology*, Cambridge University Press, 2002.

[HAU 68] HAUSNER M., SCHWARTZ J.T., *Lie Groups, Lie Algebras*, Gordon and Breach, 1968.

[HAW 00] HAWKINS T., *Emergence of the Theory of Lie Groups – An Essay in the History of Mathematics 1869–1926*, Springer, 2000.

[HIL 57] HILLE E., PHILLIPS R.S., *Functional Analysis and Semi-groups*, American Mathematical Society, 1957.

[HIS 49] HISAAKI Y., "Unitary representations of locally compact groups. Reproduction of Gelfand–Raikov's theorem", *Osaka Mathematical Journal*, vol. 1, no. 1, pp. 81–89, 1949.

[HOC 65] HOCHSCHILD G., *The Structure of Lie Groups*, Holden-Day, Inc., 1965.

[HOG 71] HOGDE-NLEND H., *Théorie des bornologies et applications*, Springer, 1971.

[HUN 67] HUNT R., "On the convergence of Fourier series, orthogonal expansions and their continuous analogues", *Proceeding Conference, Edwardsville, Ill*, Southern Illinois University Press, pp. 235–255, 1967.

[IRE 90] IRELAND K., ROSEN M., *A Classical Introduction to Modern Number Theory*, 2nd edition, Springer, 1990.

[IVE 52] IVES H.E., "Derivation of the mass-energy relation", *Journal of the Optical Society of America*, vol. 42, no. 8, pp. 540–543, 1952.

[JAC 38] JACOBSON N., "Simple Lie algebras over a field of characteristic zero", *Duke Mathematical Journal*, vol. 4, no. 3, pp. 534–551, 1938.

[JAC 53] JACOBSON N., *Lectures in Abstract Algebra. II. Linear Algebra*, Springer, 1953.

[JAC 62] JACOBSON N., *Lie Algebras*, Interscience Publishers, 1962.

[KAH 66] KAHANE J.-P., ATZNELSON Y.K., "Sur les ensembles de divergence des séries trigonométriques", *Studia Mathetamica*, vol. 26, pp. 305–306, 1966.

[KAT 79] KATZ V.T., "The history of Stokes' theorem", *Mathematics Magazine*, vol. 52, no. 3, pp. 146–156, 1979.

[KHO 72] KHOAN V.-K., *Distributions, analyse de Fourier, opérateurs aux dérivées partielles*, vols 1–2, Vuibert, 1972.

[KIR 76] KIRILLOV A., *Elements of the Theory of Representations*, Springer-Verlag, 1976.

[KNO 51] KNOPP K., *Theory and Application of Infinite Series*, Blackie & Son Ltd, 1951.

[KOB 57] KOBAYASHI S., "Theory of connections", *Annali di Matematica Pura ed Applicada*, vol. 43, no. 1, pp. 119–194, 1957.

[KOB 69] KOBAYASHI S., NOMIZU K., *Foundations of Differential Geometry*, Interscience Publishers, vol. 1, 1963, vol. 2, 1969.

[KOL 77] KOLMOGOROV A., FOMINE S., *Eléments de la théorie des fonctions et de l'analyse fonctionnelle*, Mir, 1977.

[KÖT 79] KÖTHE G., *Topological Vector Spaces I & II*, Springer-Verlag, 1979.

[KRE 76] KREE P., "Introduction aux théories des distributions en dimension infinie", *Mémoires de la S.M.F.*, vol. 46, pp. 143–162, 1976.

[KRI 97] KRIEGEL A., MICHOR P., *The Convenient Setting of Global Analysis*, American Mathematical Society, 1997.

[KUI 65] KUIPER N.H., "The homotopy type of the unitary group of Hilbert space", *Topology*, vol. 3, pp. 19–30, 1965.

[LAN 62] LANG S., *Introduction to Differentiable Manifolds*, John Wiley & Sons, 1962.

[LAN 85] LANG S., $SL_2(\mathbb{R})$, Springer, 1985.

[LAN 99a] LANG S., *Algebra*, 3rd edition, Addison-Wesley, 1999.

[LAN 99b] LANG S., *Fundamentals of Differential Geometry*, Springer, 1999.

[LEB 82] LEBORGNE D., *Calcul différentiel et géométrie*, Presses universitaires de France, 1982.

[LEE 02] LEE J.M., *Introduction to Smooth Manifolds*, Springer, 2002.

[LEV 17] LEVI-CIVITÀ T., "Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana", *Rendiconti del Circolo Matematico di Palermo*, vol. 42, pp. 73–205, 1917.

[LEV 27] LEVI-CIVITÀ T., *The Absolute Differential Calculus (Calculus of Tensors)*, Blackie & Son Limited, 1927.

[LÉV 76] LÉVY-LEBLOND J.-M., "One more derivation of the Lorentz transformation", *American Journal of Physics*, vol. 44, no. 3, pp. 271–277, 1976.

[LIC 46] LICHNEROWICZ A., *Éléments de calcul tensoriel*, Armand Colin, 1946.

[LIC 55] LICHNEROWICZ A., *Théories relativistes de la gravitation et de l'électromagnétisme*, Masson, 1955.

[LOR 21] LORENTZ H.A., "Deux mémoires de Henri Poincaré sur la physique mathématique", *Acta Mathematica*, vol. 38, pp. 293–308, 1921.

[MAD 97] MADSEN I., TORNEHAVE J., *From Calculus to Cohomology*, Cambridge University Press, 1997.

[MAI 62] MAISSEN B., "Lie-gruppen mit Banachräumen als Parameterräumen", *Acta Mathematica*, vol. 108, pp. 229–269, 1962.

[MAR 74] MARSDEN J., *Applications of Global Analysis in Mathematical Physics*, Publish or Perish, Inc., 1974.

[MAU 50] MAUTNER F.I., "Unitary representations of locally compact groups II", *Annals of Mathematics*, vol. 52, no. 3, pp. 528–556, 1950.

[MAU 55] MAUTNER F.I., "Note on the Fourier inversion formula on groups", *Transactions of the American Mathematical Society*, vol. 78, no. 2, pp. 371–384, 1955.

[MCC 01] MCCONNELL J.C., ROBSON J.C., *Noncommutative Noetherian Rings*, corrected edition, American Mathematical Society, 2001.

[MEY 64] MEYERS N.G., SERRIN J., "H=W", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 51, pp. 1055–1056, 1964.

[MIN 52] MINKOWSKI H., "Raum und Zeit", *Jahresberichte der Deutschen Mathematiker-Vereinigung*, vol. 18, Leipzig, 1909. English translation: A. Einstein, H.A. Lorentz, H. Weyl, H. Minkowski, *The Principle of Relativity*, Chapter 5, pp. 73–91, Dover, 1952.

[MIS 73] MISNER W., THORNE K.S., WHEELER J.A., *Gravitation*, W.H. Freeman and Company, 1973.

[NAR 73] NARASIMHAN R., *Analysis on Real and Complex Manifolds*, Masson, 1973.

[NOM 56] NOMIZU K., *Lie Groups and Differential Geometry*, The Mathematical Society of Japan, 1956.

[OMO 78] OMORI H., "On Banach Lie groups acting on finite dimensional manifolds", *Tohoku Mathematical Journal*, vol. 30, pp. 223–250, 1978.

[PAI 82] PAIS A., *"Subtle is the Lord" The Science and the Life of Albert Einstein*, Oxford University Press, 1982.

[PAL 63] PALAIS R., "Morse theory on Hilbert manifolds", *Topology*, vol. 2, pp. 299–340, 1963.

[PAL 64] PALAIS R., SMALE S., "A generalized Morse theory", *Bulletin of the American Mathematical Society*, vol. 70, no. 1, pp. 175–172, 1964.

[PAL 66] PALAIS R., "Homotopy theory of infinite differential manifolds", *Topology*, vol. 5, pp. 1–16, 1966.

[PAL 68] PALAIS R., *Foundations of Global Analysis*, W.A. Benjamin, Inc., 1968.

[PAU 58] PAULI W., *Theory of Relativity*, Pergamon Press, 1958.

[PER 95] PERRIN D., *Géométrie algébrique – Une introduction*, InterÉditions/CNRS Éditions, 1995.

[PER 96] PERRIN D., *Cours d'algèbre*, Ellipses, 1996.

[PET 91] PETROVSKY I.G., *Lectures on Partial Differential Equations*, Dover, 1991.

[POI 00a] POINCARÉ H., "Analysis situs", *Journal de l'École Polytechnique*, vol. 1, pp. 1–121, 1895. "Compléments à l'Analysis Situs", *Rendiconti del Circolo Matematico di Palermo*, vol. 13, pp. 285–343, 1899. "Second complément à l'Analysis Situs", *Proceedings of the London Societ*y, vol. 32, pp. 277–308, 1900.

[POI 00b] POINCARÉ H., "La théorie de Lorentz et le principe de réaction", *Archives néerlandaises*, series 2, vol. 5, pp. 252–278, 1900.

[POI 04] POINCARÉ H., "L'état actuel de la Science et l'avenir de la physique mathématique", *Bulletin des Sciences Mathématiques*, Paris, December 1904.

[POI 06] POINCARÉ H., "Sur la dynamique de l'électron", *Comptes rendus de l'Académie des Sciences*, vol. 140, p. 1504, June 5, 1905; *Rendiconti del Cercolo Matematico di Palermo*, vol. 21, pp. 129–176, 1906.

[PON 62] PONTRYAGIN L.S., BOLTYANSKII V.G., GAMKRELIZE R.V. *et al.*, *The Mathematical Theory of Optimal Processes*, John Wiley & Sons, 1962.

[POP 12] POPESCU-PAMPU P., "La dualité de Poincaré", *Images des mathématiques (CNRS)*, November 28, 2012.

[POP 16] POPESCU-PAMPU P., *What is the Genus?*, Springer, 2016.

[RIC 00] RICCI G., LEVI-CIVITÀ T., "Méthodes de calcul différentiel absolu et leurs applications", *Mathematische Annalen*, vol. 54, pp. 125–201, 1900.

[SAM 01] SAMELSON H., "Differential forms, the early days; or the stories of Deahna's theorem and Volterra's theorem", *The American Mathematical Monthly*, vol. 108, no. 6, pp. 521–530, 2001.

[SAT 86] SATTINGER D.H., WEAVER O.L., *Lie Groups and Algebras with Applications to Physics*, Springer-Verlag, 1986.

[SAU 09] SAUER T., MAJER U. (eds), *David Hilbert's Lectures on the Foundations of Physics 1915—1927*, Springer, 2009.

[SCH 58] SCHWARTZ L., "Théorie des distributions à valeurs vectorielles", *Annales de l'Institut Fourier*, Part 2, vol. 8, pp. 1–209, 1958.

[SCH 66] SCHWARTZ L., *Théorie des distributions*, 3rd edition, Hermann, 1966.

[SCH 89] SCHECHTER E., "A survey of local existence theories for abstract nonlinear initial value problems", in GILL T.L., ZACHARY W.W. (eds), *Nonlinear Semigroups, Partial Differential Equations and Attractors*, vol. 1394, pp. 136–184, Springer, 1989.

[SCH 93] SCHWARTZ L., *Analyse*, Hermann, 1993.

[SCH 99] SCHAEFER H.H., WOLFF M.P., *Topological Vector Spaces*, 2nd edition, Springer, 1999.

[SCH 11] SCHNEIDER P., *p-Adic Lie Groups*, Springer, 2011.

[SEG 50] SEGAL I.E., "An extension of Plancherel's formula for separable unimodular locally compact groups", *Annals of Mathematics*, vol. 52, pp. 272–292, 1950.

[SER 72] SERGERAERT F., "Un théorème de fonctions implicites sur certains espaces de Fréchet et quelques applications", *Annales Scientifiques de l'É.N.S.*, 4th edition, vol. 5, no. 4, pp. 599–660, 1972.

[SHA 97] SHARPE R.W., *Differential Geometry – Cartan's Generalization of Klein's Erlangen Programme*, Springer, 1997.

[SPI 99] SPIVAK M., *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, 1999.

[TAK 55] TAKIZAWA S., "On Cartan connexions and their torsions", *Memoirs of the College of Science, University of Kyoto*, series A, vol. 29, no. 3, pp. 199–217, 1955.

[TIT 62] TITS J., "Groupes simples et géométries associées", *Proceedings of the International Congress of Mathematicians*, Stockholm, pp. 197–221, 1962.

[TON 59] TONNELAT M.-A., *Les principes de la théorie électromagnétique et de la relativité*, Masson, 1959.

[TON 65] TONNELAT M.-A., *Les théories unitaires de l'électromagnétisme et de la gravitation*, Gauthier-Villars, 1965.

[TOR 73] TORUŃCZYK H., "Smooth partition of unity on some non-separable Banach spaces", *Studia Mathematica*, vol. 46, pp. 41–51, 1973.

[TRÈ 67] TRÈVES F., *Topological Vector Spaces, Distributions and Kernels*, Academic Press, 1967.

[VAR 77] VARADARAJAN V.S., *Harmonic Analysis on Real Reductive Groups*, Springer, 1977.

[VAR 84] VARADARAJAN V.S., *Lie Groups, Lie Algebras, and Their Representations*, Springer-Verlag, 1984.

[VAR 89] VARADARAJAN V.S., *An Introduction to Harmonic Analysis on Semisimple Lie Groups*, Cambridge University Press, 1989.

[VEB 33] VEBLEN O., *Projective Relativity Theory*, Springer, 1933.

[VLA 79] VLADIMIROV V.S., *Generalized Functions in Mathematical Physics*, Mir Publishers, 1979.

[WEI 38] WEIL A., *L'Intégration dans les Groupes Topologiques et ses Applications*, Hermann, 1938.

[WEI 74] WEIL A., *Basic Number Theory*, 3rd edition, Springer-Verlag, 1974.

[WEY 38] WEYL H., "Cartan on groups and differential geometry", *Bulletin of the American Mathematical Society*, vol. 44, no. 9, Part 1, pp. 598–601, 1938.

[WEY 52] WEYL H., *Space–Time–Matter*, Dover, 1952.

[WEY 53] WEYL H., *The Classical Groups – Their Invariants and Representations*, 2nd edition, Princeton University Press, 1953.

[WHI 44] WHITNEY H., "The selfintersections of a smooth n-manifold in $2n$-space", *Annals of Mathematics*, vol. 45, no. 2, pp. 220–246, 1944.

[WHI 57] WHITNEY H., *Geometric Integration Theory*, Princeton University Press, 1957.

[WHI 65] WHITTLESEY E., "Analytic functions in Banach spaces", *Proceedings of the American Mathematical Society*, vol. 16, no. 3, pp. 1077–1083, 1965.

[WIS 10] WISE D.K., "MacDowell-mansouri gravity and Cartan geometry", *Classical and Quantum Gravity*, vol. 27, no. 15, 2010.

[YAM 50] YAMABÉ H., "On an arcwise connected subgroup of a Lie group", *Osaka Journal of Mathematics*, vol. 2, no. 1, pp. 13–14, 1950.

[YOS 80] YOSIDA K., *Functional Analysis*, 6th edition, Springer-Verlag, 1980.

Cited Authors

Index

C, D

E, F

G, H

I, J, K

L, M

N, O, P

Q, R, S

T, V

W, Y

This page intentionally left blank

Mathematics and Statistics

Published in three volumes, this series, *Fundamentals of Advanced Mathematics*, presents mathematical elements that make up the foundations of a number of contemporary scientific methods: modern systems theory, physics and engineering.

This third volume is dedicated to differential and integral calculus, in their local and global components. "Local" differential calculus is discussed in the framework of Banach and Fréchet spaces. We examine the "global" approach in how one can replace these spaces with differential or analytic manifolds modeled on them.

The generalized Stokes theorem, other than its usual applications, offers us a view on the dualities of Hodge, Poincaré and de Rham. Lie's group and algebra theories allow us to develop harmonic analysis in a very general context; this analysis encompasses both the Fourier transforms and the Fourier series.

The study of connections and tensor calculus leads to spaces equipped with curvature and torsion, generalizing the Riemann and Lorentz spaces of general relativity.

Henri Bourlès has taught automatics at engineering schools and at universities for over 30 years. He is Full Professor and Chair at the *Conservatoire National des Arts et Métiers* in Paris, France, and his research focuses on systems theory.

ELSEVIER

ISTE PRESS

www.iste.co.uk

9 781785 482502