

Progress in Mathematics
Volume 149

Series Editors
Hyman Bass
Joseph Oesterlé
Alan Weinstein

J.-M. Souriau

Structure of Dynamical Systems

A Symplectic View of Physics

Translated by C. H. Cushman-de Vries
Translation editors R. H. Cushman
G. M. Tuynman

Springer Science+Business Media, LLC

Jean-Marie Souriau
UFR de Mathématiques
Université de Provence
Marseille
France

C. H. Cushman-de Vries
Dordrecht
The Netherlands

Translation edited by R. H. Cushman and G. M. Tuynman

Richard H. Cushman
Department of Mathematics
University of Utrecht
Utrecht
The Netherlands

Gijs M. Tuynman
UFR de Mathématiques
Université des Sciences
et Technologies de Lille
Villeneuve d'Ascq
France

Library of Congress Cataloging-in-Publication Data
Souriau, J.-M. (Jean-Marie),
[Structure des systèmes dynamiques. English]
Structure of dynamical systems : a symplectic view of physics / J.
-M. Souriau ; translated by C. H. Cushman-de Vries ; R. H. Cushman,
G. M. Tuynman, translation editors.
p. cm. -- (Progress in mathematics ; v. 149)
ISBN 978-1-4612-6692-1 ISBN 978-1-4612-0281-3 (eBook)
DOI 10.1007/978-1-4612-0281-3
1. Mechanics. 2. Statistical mechanics. 3. Quantum theory.
4. Symplectic manifolds. 5. Mathematical physics. I. Title.
II. Series: Progress in mathematics (Boston, Mass.) : vol. 149.
QC125.2.S6813 1997
530.15 ' 6362--dc21 97-300
CIP

Printed on acid-free paper
© 1997 Springer Science+Business Media New York

Originally published by Birkhäuser Boston in 1997
Softcover reprint of the hardcover 1st edition 1997

Copyright is not claimed for works of U.S. Government employees.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the copyright owner.
Permission to photocopy for internal or personal use of specific clients is granted by Birkhäuser Boston for libraries and other users registered with the Copyright Clearance Center (CCC), provided that the base fee of $6.00 per copy, plus $0.20 per page is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923, U.S.A. Special requests should be addressed directly to Springer Science+Business Media, LLC.

ISBN 978-1-4612-6692-1
Typset in LaTeX by R. H. Cushman.
Drawings rendered in postscript by C. Sacré.

9 8 7 6 5 4 3 2 1

Editors' preface

It is with great pleasure that we are able to provide the reader with a translation of Souriau's classical text "Structure des Systèmes Dynamiques" on mechanics. We have added the subtitle "a symplectic view of physics", which is close to the title first proposed by the author.

Compared to the French version, we have made several changes in notation for the English edition, all of which have the author's approval. The most important change is that the symbol ∇ for the exterior derivative of differential forms has been replaced by the symbol d. Consequently in order to avoid confusion, the symbols d, d',... for vector fields have been replaced by δ, δ', ... We warn the reader that differential forms are always given in terms of their action on tangent vectors and not in terms of a dual basis.



We use this opportunity to thank C. Sacré of the University of Lille I for doing all the computer graphics.

R.H. Cushman

G.M. Tuynman

Author's preface to the French edition

I sincerely thank all those who have helped me with the design and the writing of this book by their remarks, their discussions, and their advice; in particular H. Bacry, P. Casal, R. Haag, F. Halbwachs, D. Kastler, A.A. Kirillov, L. Michel, M. Zerner, and also J. Breuneval, J. Elhadad, and H.H. Fliche who, in the course of a seminar at the "Faculté des Sciences de Marseille", have checked a large part of the computations and arguments. I also thank P. Lelong, who has graciously accepted this book in the series "Dunod Université", and I hope that all those whom I did not mention will forgive me.

Jean-Marie Souriau

Notre-Dame de Vaulx,
september 1968.

Table of Contents

Introduction xvii

I. Differential Geometry

§1. Manifolds

The definition of a manifold 3
Open sets 6
Differentiable maps 6
The tangent space 7
Submanifolds 10
Manifolds defined by an equation 12
Covering spaces 13
Quotient manifolds 14
Connectedness 14
Homotopy 16

§2. Derivations

Variables 18
Vector fields and derivations 19
Derivations of linear operators 22
The image of a vector field 23
Lie brackets 25

§3. Differential equations

The exponential of a vector field 27
The image of a differential equation 28
The derivative of the exponential map 29

§4. Differential forms

Covariant fields 31
The inverse image of a covariant field 31
The Lie derivative 33
Covariant tensor fields 34
p-Forms 35
The exterior derivative 36

§5. Foliated manifolds

Foliations 38
The quotient of a manifold by a foliation 42
Integral invariants 43
The characteristic foliation of a form 44

§6. Lie groups

Actions of a Lie group on a manifold 46
The Lie algebra of a Lie group 48
Orbits 49
The adjoint representation 49
Lie subalgebras and Lie subgroups 51
The stabilizer 53
Classical examples of Lie groups 53
Euclidean spaces 55
Matrix realizations 59

§7. The calculus of variations

Classical variational problems 62
Canonical variables 63
The Hamiltonian formalism 65
A geometrical interpretation of the canonical equations 66
Transformations of a variational problem 68
Noether's theorem 68

II. Symplectic Geometry

§8. 2-Forms

Orthogonality 73
Canonical bases 74
The symplectic group 77

§9. Symplectic manifolds

Symplectic and presymplectic manifolds 81
Symplectic structures arising from a 1-form 84
Poisson brackets 85
Induced symplectic structures 87

§10. Canonical transformations

Canonical charts 90
Canonical transformations 93
Canonical similitudes 94
Covering spaces of symplectic manifolds 96
Infinitesimal canonical transformations 97

§11. Dynamical Groups

The definition of a dynamical group 100
The cohomology of a dynamical group 104
The cohomology of a Lie group 107
The cohomology of a Lie algebra 108
Symplectic manifolds defined by a Lie group 111

III. Mechanics

§12. The geometric structure of classical mechanics

Material points 121
Systems of material points 122
Constraints 122
Describing forces 125

The evolution space 126
Phase spaces and the space of motions 127
The Lagrange 2-form 129
The Lagrange form for constrained systems 131
Changing the reference frame 133
The principle of Galilean relativity 135
Maxwell's principle 137
Potentials and the variational formalism 139
Geometric consequences of Maxwell's principle 141
An application: variation of constants 142
Galilean moments 144
Remarks 148
Examples of dynamical groups 149

§13. The principles of symplectic mechanics

Nonrelativistic symplectic mechanics 154
Moments, mass, and the center of mass 155
The center of mass decomposition 157
Minkowski space and the Poincaré group 163
Relativistic mechanics 166

§14. A mechanistic description of elementary particles

Elementary systems 173
A particle with spin 174
Remarks 179
A particle without spin 180
A massless particle 182
Remarks 184
Nonrelativistic particles 185
Mass and barycenter of a relativistic system 188
Inversions of space and time 189
A particle with nonzero mass 191
A massless particle 192

§15. Particle dynamics

A material point in an electromagnetic field 194
A particle with spin in an electromagnetic field 197
Systems of particles without interactions 200
Interactions 202

Scattering theory 205
Bounded scattering sources 210
Geometrical optics 212
Planar mirrors 214
Collisions of free particles 218

IV. Statistical Mechanics

§16. Measures on a manifold

Composite manifolds 223
Compact sets 224
Riesz spaces 227
Measures 229
The tensor product of measures 232
Examples of measures 234
Completely continuous measures 235
Examples of completely continuous measures 237
The support of a measure 241
Bounded measures 241
Integrable functions 246
The image of a measure 249
Examples 253
Random variables 256
Average values 258
Entropy and Gibbs measures 260
The Gibbs canonical ensemble of a dynamical group 265

§17. The principles of statistical mechanics

Statistical states 271
Hypotheses of the kinetic theory of gases 271
Equilibria of a conservative system 273
Ideal gases 279
A monatomic ideal gas 280
An arbitrary ideal gas 280
An ideal gas thermometer 281
Heat and work 283
Specific heat 284
Covariant statistical mechanics 287

Examples 291
The statistical equilibrium of an isolated system 293
Relativistic statistical mechanics 295
A relativistic ideal gas 296
Statistical equilibria of photons 298

V. A Method of Quantization

§18. Geometric quantization

Prequantum manifolds 307
Prequantization of a symplectic manifold 309
Prequantization of a symplectic manifold admitting a potential 312
Prequantization of a sphere S_2 313
Prequantization by "fusion" 315
Prequantization of a direct product 316
Prequantization of a relativistic particle with spin $\frac{1}{2}$ 318
Prequantization of a massless particle 321
Massless particle with spin $\frac{1}{2}$ 322
Massless particle with spin 1 324
Planck manifolds 325
Quantomorphisms 326
Homotopy and prequantization 328
Systems of elementary particles 331
Infinitesimal quantomorphisms 331
Quantization of dynamical groups 334
The Hilbert space of a prequantum manifold 339

§19. Quantization of dynamical systems

The correspondence principle 346
State vectors and observables 347
The formulation of Planck's condition 350
Stationary states 351
The formation of wave equations 352
The nonrelativistic material point 354
The relativistic material point 358
The nonrelativistic particle with spin $\frac{1}{2}$ 360
The relativistic particle with spin $\frac{1}{2}$ 362

The massless particle with spin $\frac{1}{2}$ 364
The massless particle with spin 1 366
Assemblées of particles 370
Creation and annihilation operators 373
Quantum states 378

Bibliography 387

Index 391

List of notation 405

Introduction

Lagrange's book *Mécanique Analytique*, published at the end of the 18^{th} century, is an incomplete work of which certain chapters are only sketched. This is the reason why it is known principally through the work of Poisson, Hamilton, and Jacobi, whose contributions have determined the *form* of analytical mechanics as it is taught today.

Formal perfection, in mathematics, is never negligible; it is even a necessary ingredient for progress. However, it seems that the classical formalization of analytical mechanics has overlooked an important part of Lagrange's ideas. Without doubt this happened because the mathematics of the first part of the 19^{th} century did not have the necessary scope; also because the successes of the theory (first in celestial mechanics and later in statistical mechanics) hid the necessity of questioning categories [1] so *classical* that they seemed *natural*. One had to wait until the 20^{th} century to learn that the words *time*, *space*, and *matter* do not have any direct physical meaning, but are only formal symbols of revisable physical theories, and nowadays are outdated.

Analytical mechanics is not an outdated theory, but it appears that the *categories* [1] which one classically attributes to it such as *configuration space*, *phase space*, *Lagrangian formalism*, *Hamiltonian formalism*, are, simply because they do not have the required covariance; in other words, because these categories are in contradiction with *Galilean relativity*. A fortiori, they are inadequate for the formulation of relativistic mechanics in the sense of Einstein.

We are now going to examine this idea by looking at the example of a material point subjected to a force $\mathbf{F}$. We will denote its mass by m and its position at time t by $\mathbf{r}$.

Here the *configuration space* is ordinary (3-dimensional) Euclidean space in which $\mathbf{r}$ evolves. The *phase space* is described by the pair $\begin{pmatrix} \mathbf{r} \\ \mathbf{v} \end{pmatrix}$ where $\mathbf{v}$ is the velocity of the particle; thus its dimension is 6. Newton's equation

$$\mathbf{F} = m \frac{d^2\mathbf{r}}{dt^2} \tag{I}$$

is a system of second order differential equations on the configuration space. One can rewrite it as a system of first order differential equations

[1] We use the word *category* in its epistomological sense (Aristotelian or Kantian, for example), and not in its algebraic sense.

on the phase space:

(II) $$\frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m} \qquad \frac{d\mathbf{r}}{dt} = \mathbf{v}\,.$$

In the case where the force $\mathbf{F}$ is derived from a potential w, one considers the quantity

(III) $$L = \tfrac{1}{2}\, m\, \mathbf{v}^2 - w\,,$$

called the *Lagrangian*. The integral of the Lagrangian between two arbitrary times t_0 and t_1

(IV) $$\mathcal{A} = \int_{t_0}^{t_1} L\, dt$$

is called the Hamiltonian *action*. The equations of motion (I) or (II) can be obtained by requiring that this integral is minimized (Hamilton's principle or the *principle of least action*). We can also give the equations of motion their *canonical* form

(V) $$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j} \qquad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}$$

by writing q_j $(j = 1, 2, 3)$ for the components of $\mathbf{r}$, p_j for the components of $m\mathbf{v}$, and by writing the *Hamiltonian*

(VI) $$H = \tfrac{1}{2}\, m\, \mathbf{v}^2 + w$$

as a function of the *canonical coordinates* p_j, q_j, and t.

The extension of equations (V) to more general cases constitutes the *Hamiltonian formalism*, which is used as the foundation of modern treatments of analytical mechanics.

However, the very definition of phase space (as well as of configuration space) assumes that a *choice of reference frame has been made* (for example, the earth). But the phase space of the same material point is *essentially* different for two observers in motion with respect to one another (even in the simplest case of uniform rectilinear motion) because the formulæ for the change of reference frame *depend on time*. The Hamiltonian H also depends on the observer, as does the variable t in the case of relativistic mechanics ("the relativity of simultaneity").

Fortunately, this *subjectivity* of the fundamental concepts has a simple remedy: *it suffices to treat the variable t like the others*. In the above example one should use the 7-dimensional space V over which varies the triple

(VII) $$y = \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \end{pmatrix}^{2}.$$

By "clearing denominators" in (II), these equations can be written in

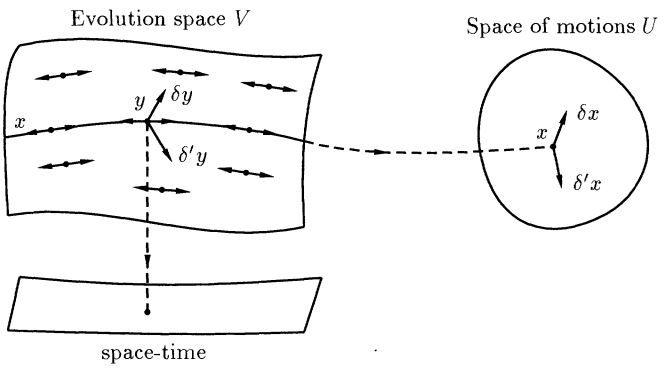
[2] y is sometimes called an *initial condition* of the motion.

the homogeneous form

(VIII) $$m\,\delta\mathbf{v} - \mathbf{F}\,\delta t = 0 \qquad\qquad \delta\mathbf{r} - \mathbf{v}\,\delta t = 0\,.$$

At each point y of V, these equations define the *tangent direction* to the curve x described by the point y during the evolution of the system (see Fig. 1). These curves thus are — according to a classical terminology — the *lines of force* of the field of directions defined by the equations (VIII).[3] Nowadays it is preferable to say *leaves* instead of *lines of force*, because the theory of foliated manifolds (§at p. 38) incorporates this problem in a more general and powerful formalism.

Fig 1



The space V, which we will call the *evolution space* (p. 127), and its field of directions (or "foliation") have an *objective* character. Of course, every observer can describe V by his favorite parameters.

It would seem natural to apply the same treatment to configuration space. In the case of a material point, this boils down to working on *space-time* and not on space and time separately. However, we encounter a serious conceptual difficulty as soon as we start treating a more complicated system. In the case of *two material points*, for example, should we replace the configuration space over which varies $\begin{pmatrix}\mathbf{r}_1\\ \mathbf{r}_2\end{pmatrix}$ by the 7-dimensional space described by $\begin{pmatrix}\mathbf{r}_1\\ \mathbf{r}_2\\ t\end{pmatrix}$, or by the 8-dimensional space described by $\begin{pmatrix}\mathbf{r}_1\\ t_1\\ \mathbf{r}_2\\ t_2\end{pmatrix}$, obtained by looking at the two points at different times t_1 and t_2? One of these choices seems natural in classical

[3] We assume that $\mathbf{F}$ is a known function of y, that is, a function of t, $\mathbf{r}$, and $\mathbf{v}$.

mechanics and the other in relativistic mechanics! Since it is more a matter of opinion than of physics, we simply propose to *give up alluding to configuration space*. We will see later that we do not lose anything by doing so.

Now that we represent a dynamical system by a foliation of the evolution space, we must ask ourselves what it is that is specific to mechanics in such a description. The answer to this question was given — implicitly — by Lagrange himself. The foliation is determined by an antisymmetric covariant second order tensor. We will denote it by σ and call it the *Lagrange form*.[4] The components of this tensor are expressions known as *Lagrange brackets*.

σ can be considered as a bilinear operator on tangent vectors of V. If we choose two such vectors (Fig. 1)

(IX) $$\delta y = \begin{pmatrix} \delta t \\ \delta \mathbf{r} \\ \delta \mathbf{v} \end{pmatrix} \qquad \delta' y = \begin{pmatrix} \delta' t \\ \delta' \mathbf{r} \\ \delta' \mathbf{v} \end{pmatrix},$$

σ associates to them a kind of (antisymmetric!) scalar product, which we will denote by $\sigma(\delta y)(\delta' y)$ and whose value is[5]

(X) $$\langle m\,\delta\mathbf{v} - \mathbf{F}\,\delta t,\ \delta'\mathbf{r} - \mathbf{v}\,\delta' t\rangle - \langle m\,\delta'\mathbf{v} - \mathbf{F}\,\delta' t,\ \delta\mathbf{r} - \mathbf{v}\,\delta t\rangle\,,$$

where the brackets $\langle\ ,\ \rangle$ denote the usual scalar product. Looking at the equations of motion (VIII), we see that the vector δy is *tangent* to the leaf x passing through y if it satisfies

(XI) $$\sigma(\delta y)(\delta' y) = 0 \qquad \forall\ \delta' y\,.$$[6]

This is what is meant by saying that the foliation is the *kernel* of the 2-form σ (p. 130).

One can find other tensors having the same kernel (for example, multiples of σ). Thus the description of a system by its Lagrange form, which we adopt henceforth, is *more complete* than its description by its foliation alone. The hypothesis that σ has a *direct physical significance* is essential, and can be tested by experiment. And thus, when carrying

[4] An antisymmetric covariant tensor of order p is called an *exterior differential form of degree p*, or simply a *p-form* (p. 31).

[5] In every new mechanical problem, we have to define the equivalent of this expression (see equations (12.45), (12.52), (14.4♣), (14.24♣), (14.29♣), (14.48), (14.53♣), (15.1), (15.26), etc.). We also treat the case of systems subjected to *ideal* holonomic or nonholonomic *constraints* (p. 124 and p. 132).

[6] The classical *principle of virtual work* is only an abbreviated form of formula (XI).

out a change of reference frame in classical mechanics, we also study the transformation formulæ of the 2-form σ and not just those of the equations of motion ((12.76), p. 136). This leads us to classify forces into two distinct types.[7]

Since the work of Henri Poincaré and Élie Cartan, the theory of differential forms (§4, p. 31) has become an essential instrument of modern (differential) geometry. This theory includes operations and specific formulæ that are very flexible and very powerful and allows us to express geometrically some properties whose classical denotation is quite cumbersome. For example, the fact that the force $\mathbf{F}$ in the example of a material point does not depend on velocities and derives from a potential can simply be written as

$$\mathrm{d}\sigma = 0\,, \tag{XII}$$

where the symbol d represents the operation of *exterior derivative* (p. 36).

In the diverse *dynamical systems* encountered in nature, condition (XII) is actually satisfied. We thus propose it as a *principle of mechanics*, under the name *Maxwell's principle*.[8]

Furthermore, Maxwell's principle is necessary and sufficient in order that, in the description of a system by means of a phase space, the equations of motion can be written in Hamilton's canonical form (V).[9] Thus Maxwell's principle contains the essence of the Hamiltonian formalism, but without the inconveniences of the usual formulation (p. 139).

Likewise, the *principle of least action* (Lagrangian formalism) disappears because we have given up the notion of configuration space in terms of which this principle is stated. However, from the days of Fermat and Maupertuis until the present, variational principles are among the most used guides in theoretical physics. Thus it is important to know that nothing is lost by replacing the Lagrangian formalism with Maxwell's principle.

The answer is supplied by experiment: there exist systems that satisfy Maxwell's principle and for which one can prove that they do not have an underlying Lagrangian formalism. One should think in particular of *polarized particles* or *particles with spin*.[10] One can see also that the

[7]Examples of forces of the first type are: *weight*, *electrostatic attraction*; of the second type: *Coriolis force*, *Lorentz force* (exerted by a magnetic field on a moving electrical charge).

[8]J.C. Maxwell did indeed formulate this principle in certain particular cases (equations (12.86), p. 138).

[9]With the proviso that the Hamiltonian H can be, in certain circumstances, a *multivalued* function on phase space.

[10](pages 174 and 186). It has been known for a long time how to write classical

Lagrangian, when it exists, does not always have the required covariance ((12.136), p. 148).

Thus "Maxwell's formalism" is better adapted to the description of nature.

We can define, without any ambiguity, a *motion* x of the system as being a *leaf* (or line of force) of V. *Causality* (or *determinism*) of the system is expressed by the fact that each *initial condition* y defines one and only one motion x.

We are now going to consider the *abstract set* of motions (p. 129). It is a space U of which *each point* x *represents a complete motion* (considered in its complete temporal development).

It can be shown that each motion x can be described by a system of n parameters, none superfluous, which are differentiable functions of the initial conditions, with the proviso that it may be necessary to use different systems of parameters (or "charts") to cover all of U (just as several charts are needed to represent the whole (planet) Earth).

Differential geometry provides us with a precise vocabulary (pages 3, 7, and 10) to state these properties:

(XIII) $\begin{cases} \text{The space of motions } U \text{ admits the } \textit{structure of a manifold} \text{ of dimension } n. \\ \text{The map } y \mapsto x \text{ (Fig. 1) is an } \textit{open differentiable} \text{ map from } V \text{ onto } U. \end{cases}$

Thus, in the example of a material point, the space of motions is a 6-dimensional manifold.[11]

It is important not to confuse *motions* with *trajectories* (which are the projections of motions on configuration space or phase space, and which are obtained by "forgetting" the temporal aspect of a motion). In general, the set of trajectories *does not admit* the structure of a manifold that can be derived from that of V or U, simply because there are not enough "single valued first integrals" to determine the trajectories.[12] The "entanglement" of the set of trajectories is the origin of what are called *ergodic problems*. This comparison between motions and trajectories

---

or relativistic equations for these particles (L.H. Thomas, 1927). Moreover, one uses them in practice (particularly for the measurement of magnetic moments). Several authors have searched in vain for a corresponding Lagrangian.

[11] In the case of the Kepler problem (p. 121), U is a 6-dimensional manifold (the "Kepler manifold"). The 6 parameters needed to describe a motion are called the *orbital elements*. It seems that the Kepler manifold is one of the first "abstract" manifolds considered by mathematicians.

[12] In many cases there exists only one such first integral: the *energy*. But there are exceptions: the Kepler trajectories form a 5-dimensional manifold.

shows how much the structures of differential geometry are important in physics.

The fact that the map $y \mapsto x$ is *differentiable* allows the following construction. After choosing a point y of V and tangent vectors δy and $\delta' y$ (formulæ IX), one can define (p. 7) corresponding vectors δx and $\delta' x$ *on the space* U. By analogy with the theory of surfaces, it is said that δx and $\delta' x$ are *tangent vectors to the manifold* U *at the point* x (Fig. 1).

Taking Maxwell's condition into account, it can be shown ((12.100), p. 141) that the tensor σ "pushes down to the space U",[13] which means that we can define a tensor on U, still denoted by σ, such that

(XIV) $$\sigma(\delta y)(\delta' y) \equiv \sigma(\delta x)(\delta' x)\,.$$

It thus suffices to know the 2-form σ on U to be able to compute it on V. Since δx is zero whenever δy is tangent to a leaf, it follows that the kernel (see XI above) of the 2-form σ on U is zero, that is, σ is *injective* on U.[14] One can prove (p. 141) that the 2-form σ on U *also* satisfies

(XV) $$\mathrm{d}\sigma = 0\,.$$

By a *symplectic manifold* (p. 81) we will mean any manifold on which an injective 2-form σ satisfying $\mathrm{d}\sigma = 0$ is defined. Thus the above results can be summarized by saying that

(XVI) *the space of motions is a symplectic manifold.*

Knowing the symplectic manifold U allows us, if we so desire, to go back to the manifold V.[15] The description of a dynamical system by means of its space of motions is thus *as complete* as the description provided by its evolution space. But it has an important advantage: the space of motions of a system *composed of* several independent parts is the *direct product* (p. 81) of the individual spaces of motions. The description of the evolution space of a composite system by means of the individual evolution spaces does not have the same simplicity (p. 200).

From now on (p. 154) we *give up the notion of evolution space* — which has served only as an epistomological intermediary. This will allow us to *enlarge classical mechanics*. We will be able to treat systems which have a symplectic space of motions, but which do not have a covariant

[13]The Lagrange form is said to be an "integral invariant" of the foliation of V (p. 43).

[14]From this it follows that the dimension n of U is *even* (p. 76).

[15]One can, if one wishes, consider V to be the direct product of U and the *time axis*. The Lagrange form on V can be computed using formula (XIV). Maxwell's principle is automatically satisfied because of (XV).

evolution space (p. 184). A typical example is the case of *photons*, which thus take their place in mechanics in accordance with Newton's wish.

Let us return to classical mechanics. As we have seen, the definition of U implies the *principle of causality*. Choosing U as the *initial category* of mechanics means allowing causality without restrictions. Causality thus loses its status as a *result* of the theory to become a *consistency condition* of it.

This approach may seem hazardous to some. However, it is indispensable for further progress. Nevertheless, let us recall that it concerns only one very particular aspect of causality. We will see in particular that the point of view adopted here is perfectly compatible with the kind of "indeterminism" encountered in statistical and quantum mechanics.

Let us also note that, by acting in this way, we do not prejudge the ease or difficulty of the problems posed by explicitly solving the equations of motion. Problems of asymptotic behavior, for example, reappear in due time, but from a new point of view (p. 274).

Directly considering U focuses attention on a situation one might call the *paradox of the physicist*. The physicist, who is very careful not to invent the world a priori, thinks he deals only with *real* facts. But the actual evolution of a material system, however complex it may be, is represented by a *single point* x of the manifold U.[16] Thus the other points have only a *virtual* existence. Nevertheless, only by giving a structure to the set of these virtualities can the physicist, to a certain degree, explain and predict reality.

Certain minds are reluctant to accept taking the virtual into account (it is there in particular that one finds the reason for the a priori rejection of the explanatory power of variational principles which appeal explicitly to "virtual motions"; one should note that the symplectic formulation contains *less of the virtual* than the variational formulation). It seems preferable to us to "resolve" this paradox by establishing a *dialectic between the real and the virtual*, which is made possible by the theory of groups.

It is clear that the laws of physics are discovered by *comparing* observations or experiments. This allows us to treat various partial aspects of reality as different virtualities. Now two distinct experiments necessarily occupy two different regions of space-time. We thus can compare them only by putting them into *correspondence*. The correspondences called upon in this way act on space-time as well as on its "contents".

[16] Henri Poincaré has observed that the laws of physics deal with all possible worlds, whereas "there exists only one copy" of the real world.

We will essentially presume that these correspondences preserve the structure of the virtual — otherwise a comparison of experiments would remain fruitless. This is why our scientific knowledge of nature derives necessarily from an explicit or implicit appeal to a set of isomorphisms of the virtual. These isomorphisms generally form a *group*.[17]

In the case of a dynamical system, we have assumed that the virtual makes up the space of motions U, and that its structure is symplectic. Its isomorphisms — or *symplectomorphisms* — are easily deduced. They are the transformations a such that

(XVII)
$$\begin{cases} (1) & a \text{ is a bijection from } U \text{ onto itself;} \\ (2) & a \text{ and its inverse are differentiable;} \\ (3) & a \text{ preserves the 2-form } \sigma: \\ & \qquad \sigma(\delta x)(\delta' x) \equiv \sigma(\delta[a(x)])(\delta'[a(x)])\,.^{18} \end{cases}$$

The symplectomorphisms of U obviously form a *group* ((10.15), p. 94), which we denote by $\mathrm{Can}(U)$.[19] It turns out — and this is an essential property of symplectic geometry — that this group is *extremely rich* [20] (much more so, for example, than the group of isomorphisms of the Euclidean structure, which is a finite dimensional group). In particular, every *quantity* capable of being *measured* (a "dynamical variable") can be associated, by an infinitesimal method (p. 98), to a 1-dimensional group of symplectomorphisms. Conversely, every *finite dimensional* group [21] of symplectomorphisms (which we will call a *dynamical group* [22]) most often can be associated with a system of *dynamical variables*, which are, by construction, "constants of motion". In this way ((11.12), p. 103) one

[17]This group differs according to the physical theory used. In the simplest cases (experiments repeated in the same laboratory) one uses the group of *time translations* (they are written as $\mathbf{r} \mapsto \mathbf{r}$, $t \mapsto t +$ constant in the reference frame of the laboratory). In general relativity the notion of a group is insufficient and one has to use the notion of a *pseudo group*, see Souriau [32]. We encounter groups corresponding to Newtonian and relativistic mechanics.

[18]These properties say that a is an isomorphism of successively the *set theoretic* structure of U (property (1)), of the *manifold* structure (properties (1) and (2); such isomorphisms are called *diffeomorphisms*), and finally of the *symplectic* structure.

[19]Because symplectomorphisms of the classical phase space are called "canonical transformations"; sometimes they are also called "contact transformations".

[20]When U is Hausdorff and connected, one can show that $\mathrm{Can}(U)$ acts transitively, that is, for every pair of points x, x' of U there exists a symplectomorphism a such that $a(x) = x'$, and even these exist in large numbers.

[21]Finite dimensional groups are called *Lie groups* (§6, p. 45).

[22]Having knowledge of a dynamical group is an important element in the description of a system; for example, the *harmonic oscillator* ((12.150) and (12.151), p. 151), or the *hydrogen atom* ((12.149), p. 150).

generalizes a theorem which Emmy Noether stated in the Lagrangian formalism (p. 68) and whose physical importance is considerable.

Given a dynamical group, the set of dynamical variables constructed in this fashion can be organized into a geometrical object μ which we will call a *moment*.[23]

The fact that such a moment is a priori only defined up to an additive constant brings into play an algebraic mechanism known as *cohomology* (p. 104). This mechanism will play an important role in the study and classification of dynamical systems.

The most immediate application of the notion of *dynamical group* is the study of *covariance* of mechanics. By stating that the structure of the set of motions of an *isolated* dynamical system is the same relative to every inertial reference frame (which is the natural formulation of the principle of Galilean relativity), one can show that a certain 10-dimensional group, called the *Galilei group*, is a dynamical group of such a system (p. 144). *This single fact* will have a great number of important physical consequences.

First of all, the ten components of the associated moment μ allow us to attach to every motion x physical quantities whose interpretation is easy: *energy*, *linear momentum*, *angular momentum*, and the *center of mass* ((12.140), p. 148). The *principle of equality of action and reaction* becomes a theorem ((12.139), p. 148). One can establish — without referring to a decomposition of the system into material points — some of the "general theorems" of mechanics: conservation of linear and angular momentum, conservation of energy. One can show that every motion of the system can be described by the uniform rectilinear motion of the center of mass (the principle of Galilean inertia) and the motion of the system "around" its center of mass (Koenig's theorem). In other words, the space of motions is *decomposable* (p. 161).[24]

The cohomology of the Galilei group allows us to associate to each dynamical system a *number* m, which is independent of the motion. m can be interpreted as the mass of the system ((12.136), p. 148).

One can also carry out a *dimensional analysis* of mechanics (p. 156).

---

[23]Mathematically, the moment μ is a *torsor* of the group, that is, an element of the dual of the Lie algebra. μ depends on the motion x. Editors' note: The map $x \mapsto \mu$ is called the *moment* or *momentum map*.

[24]Geometrically, one can see that U is the *product* of two symplectic manifolds. This fact arises from the existence of a 14-dimensional *dynamical group* which extends the Galilei group. In particular, this larger group contains subgroup SO(3), called the *group of spin*. We will see that this subgroup does not appear in relativistic mechanics.

We are led to choose as fundamental dimensions *action* A,[25] length L, and time T. The dimensional equations of mechanical quantities are better expressed with these dimensions than with the traditional dimensions M, L, and T. It is cohomology that gives mass the status of a derived dimension $M = ATL^{-2}$.

Without changing anything in the symplectic formulation of the theory, one can go from *classical mechanics* to *relativistic mechanics* simply by replacing the classical formulæ for the change of reference frame by the formulæ of Lorentz. This amounts to replacing the Galilei group as a dynamical group by the *Poincaré group*.[26] An analysis parallel to the one given for classical mechanics gives us — as theorems — the *principles of relativistic mechanics* ((13.70), p. 167). The *differences in structure* between the Galilei group and the Poincaré group give us *fundamental differences* between these two kinds of mechanics. Since the cohomology of the Poincaré group is zero, there is no longer a number which is characteristic for the system, such as the classical *mass*; the role of mass is played by the relativistic *energy*, which *no longer contains an arbitrary additive constant*, but which *depends on the motion* (pages 169 and 188). Furthermore, there is no *barycentric decomposition* because this decomposition was a consequence of the existence of a privileged subgroup of the Galilei group (p. 159), which does not have an equivalent in the Poincaré group.

A priori, the simplest dynamical systems are those on which the Poincaré group acts *transitively* (that is, every virtual motion can be derived from the real motion by a change of reference frame). Symplectic geometry allows us to completely *classify* these "elementary" systems.

The results of this classification (given in detail in §14 for the physically interesting cases) show that we obtain in this way the *elementary particles* that are actually observed. The classification is described by means of two numbers m and s for which we have a full physical interpretation: the (rest) *mass* and *spin* (or *intrinsic angular momentum*). In the massless case, a third number intervenes: the *helicity*, which can take only the values ± 1. The mechanical model thus constructed can be used for the *photon*, which is (circularly) polarized *to the right or left* according to the sign of the helicity. There is even place to describe hypothetical particles (tachyons) whose velocity is always larger than that

[25] $A = ML^2T^{-1}$ is the dimension of the Hamiltonian action (IV), whence its name. The role it plays here comes from the fact that it is the dimension of the *symplectic scalar product* (X). The fact that A is also the dimension of *Planck's constant* will play an essential role in quantization (p. 346).

[26] Also called the "inhomogeneous Lorentz group". It is 10-dimensional as is the Galilei group (p. 164).

of light ((14.37), p. 184).

Let us also point out that this theory allows us to classify the possible behaviors of these particles relative to *inversions of space and time* (p. 189), and to give a *nonrelativistic* description of them (p. 185). In particular, the nonrelativistic model of the photon depends (in addition to spin and helicity) on a new parameter called "color" (which can be interpreted as *spatial frequency* after quantization). One can see that *classical geometric optics* can be deduced from this model by neglecting spin (p. 212).

In these cases we see how much the real-virtual dialectic fertilizes the apparently sterile description of the motion of a system by a single point on a symplectic manifold.

Besides these theoretical aspects, "symplectic mechanics" seems to have numerous practical applications, of which we mention a few.

— First of all let us cite Lagrange's *perturbation theory* (p. 142), which plays a fundamental role in celestial mechanics.[27]

— The theory of a *particle in an electromagnetic field* contains a certain number of a priori hypotheses. By using the axioms of symplectic geometry, one can limit the arbitrariness in these hypotheses. The equations which we propose, in the classical as well as in the relativistic case (p. 194), seem to agree with experiment. Besides the well-known effects related to the (electric) charge (the electrostatic force and the Lorentz force), they describe the *magnetic effects related to spin* (the inertial effects of spin being known a priori and not by analogy with a rotating rigid body). In this way one can describe experiments concerning *spin precession* and *magnetic resonance*. One also recovers the equations of Thomas, Bargmann, Michel, and Telegdi, which are used for precise measurements of magnetic moments ((15.35) and (15.36), p. 199).

— *Scattering* theory plays a very important role in quantum mechanics. The *non-quantum scattering theory* which we propose (p. 205) is very general. It can be applied, for example, to the study of a *photon traversing an optical instrument* (the parameters on which the photon depends, *color* and *helicity*, allow us to anticipate *dispersion* and *rotational birefringence*, and we can derive important concepts of geometrical optics

[27]If one tries to find out, using the classical methods of differential calculus, how the coordinates of the orbit of a planet vary under a perturbing force, an extremely dense calculation leads to equations which were discovered by Lagrange and whose simplicity is striking . The search for understanding this simplicity certainly has been an essential factor for progress and discovery in theoretical mechanics.

such as the *angular eikonal* (p. 213)). Scattering theory can also be applied to collisions of particles (p. 218) (we obtain a priori the conservation laws which govern the phenomenon). In the theory of an *ideal gas*, scattering theory can take into account the collisions of the particles between themselves and with the walls of the box containing them (p. 272). We will study in detail the example of *reflection of a photon by a planar mirror*. The relativistic treatment of this phenomenon, taking the polarization of the photon into account, leads to some interesting conclusions (for example, the change of the helicity, the Doppler effect, etc. (p. 217)).

On a more speculative level, scattering theory can be used (in classical as well as in quantum mechanics) to study indirectly some problems which at present we do not know how to treat directly, the most "shocking" one being that of *relativistic interaction of particles*, a problem that probably can only be solved by giving up some of our preconceived ideas (p. 204).

Statistical mechanics can be formulated very simply by means of the space of motions U. It suffices to describe the behavior of the system not by a *point* of U, but by a *probability measure* on U, also called a "statistical state" (p. 271).[28]

The *entropy* of such a probability measure can be defined directly (p. 260). Imposing the condition that entropy is *maximal* (modulo certain conditions associated with a *dynamical group*), we define an *ensemble of statistical states* generalizing the *Gibbs canonical ensemble* and having some interesting mathematical properties (p. 265).

Classical statistical mechanics corresponds to the dynamical group of *time translations*. Assuming a priori that statistical equilibria belong to the associated Gibbs ensemble, we easily recover the concepts and principles of classical thermodynamics (temperature, energy, heat, work, entropy, thermodynamic potentials) and of the kinetic theory of gases (pressure, specific heats, Maxwell's velocity distribution, etc., see pages 276–287).

If we also allow Gibbs ensembles to be associated to various subgroups of the Galilei or Poincaré group, we describe *new phenomena* in a manner which seems to be consistent with experiment, such as *relativistic ideal gases* (p. 296), *rotation of celestial bodies* ((17.20), p. 295), and classical (p. 292) or relativistic *centrifuges* (p. 298). For example,

[28]The corresponding description in phase space is a complicated formalism, necessitating a partial differential equation in the sense of distributions.

one can predict the statistical orientation of the spin of particles inside a centrifuge ((17.113), p. 293).

In this *covariant* formulation of statistical mechanics, the role of classical *temperature* is played by a 4-vector (p. 297), which gives, in addition to the *proper temperature*, the *mean velocity* of the environment under study. The temporal orientation of the temperature vector explains why particles with negative mass (as well as tachyons) are excluded from the equilibria we know. One may conjecture that this vector represents the "arrow of time", and so find a means to escape the paradox that statistical irreversibility is based on reversible elementary laws. But it would be pleasant to have a formulation of thermodynamics *which is valid away from equilibrium states*, and which does not boil down to an upgraded version of "thermostatics".

Among the examples we discuss is that of the *statistical equilibrium of photons* enclosed in a box. Starting from the previously established model of a photon, taking into account their indistinguishability,[29] and treating the *number* of photons as a random variable,[30] we obtain a coherent description of *black body radiation*, which provides in particular the *Stefan-Boltzmann law* and the law of *radiation pressure* (p. 301). But the spectral distribution of black body radiation so obtained[31] differs quantitatively from the observed distribution (Planck's law). Thus we have reached the limits of the validity of statistical mechanics, which must be replaced by quantum mechanics.

One of the fundamental ideas of quantum mechanics is that of "replacing" the measurable quantities of classical mechanics (the dynamical variables) by *operators* (called *observables*) satisfying a certain number of conditions. In the formulation of Dirac, one should associate in a linear way to dynamical variables $u, v, \ldots$ operators $\widehat{u}, \widehat{v}, \ldots$ so that

(XVIII)
$$\begin{cases} \widehat{u} \circ \widehat{v} - \widehat{v} \circ \widehat{u} \equiv -i\hbar\, \widehat{[u,v]}_P \\ \widehat{1} = \text{the identity operator}\,.^{32} \end{cases}$$

[29] This indistinguishability has a precise symplectic formulation ((15.47), p. 202).

[30] Possible because of the notion of "Fock manifold" (p. 300), which we will meet again later on.

[31] Using a mechanistic model of the photon allows us to avoid divergences which render the Rayleigh-Jeans law incorrect (the "ultraviolet catastrophe").

[32] $\hbar$ is the quotient of Planck's constant by 2π. The expression $[u, v]_P$ denotes the *Poisson bracket* of u and v. It is a new dynamical variable whose definition is purely symplectic ((9.22), p. 86).

We give an *explicit solution* to "Dirac's problem" (XVIII) (p. 349). An ingredient of this solution is the construction of a manifold Y, associated in prescribed fashion to the symplectic manifold U of classical motions. We will say that Y is a *prequantum manifold* [33] and that the construction of Y is a *prequantization* of U (p. 307–310).

One can state necessary and sufficient conditions in order that the prequantization of the symplectic manifold U be possible (pages 310–312). It turns out that these conditions seem to be satisfied for dynamical systems encountered in nature.[34] Thus we have to investigate whether *prequantum manifolds* so constructed can play a role in the formulation of quantum mechanics.

Let us first note that the prequantization of the space of motions, which is unique in the case of an elementary particle ((18.135), p. 330), can have, in certain cases, *several essentially different solutions*. We provide a rule which allows us to *inventory the different prequantizations* ((18.133), p. 330). This rule is based on *homotopy theory* (p. 16). In this way we can show that a system of indistinguishable particles has *exactly two prequantizations*. Both prequantizations seem to be realized in nature ("Bose-Einstein statistics" and "Fermi-Dirac statistics"), contrary to other quantizations which have been proposed ("parastatistics") but which do not fit into this geometrical framework.

We will show that the usual *wave functions* of quantum mechanics can be described by means of functions defined *on the prequantum manifold itself* (one gets the usual spatio-temporal description by a Fourier transform). We will obtain the *Schrödinger equation* ((19.73), p. 357), the *Klein-Gordon equation* ((19.91), p. 359), the *Dirac equation* ((19.113), p. 363), the quantized *Maxwell equations* ((19.158), p. 369), and others ((19.100), p. 361 and (19.132), 365). In all of these examples we have a systematic way of defining the classical observables: *linear momentum*, *energy*, *angular momentum*, and *position*.

The case of an "assemblée" of an unknown number of identical particles can be treated by the same method starting from a very simple hypothesis: a classical motion of the assemblée is a *finite subset* of the space U of classical motions of a single particle.

The space U_Φ of motions of an assemblée as defined above also admits the structure of a symplectic manifold, which we will call the "Fock

[33] The *prequantum structure* of a manifold is uniquely determined by a 1-form ϖ (p. 308) satisfying certain axioms.

[34] In the case of elementary particles (classical or relativistic) it is necessary and sufficient that the spin (the intrinsic angular momentum) be an integer multiple of $\hbar/2$ ((18.57), p. 312 and (18.81), p. 321). This is in agreement with experiments.

manifold".

If we apply the Bose-Einstein or Fermi-Dirac prequantization to the Fock manifold, we see emerging the notions and algebraic relations constituting the formalism of "second quantization", namely, the *quantum vacuum* (p. 373) and the *creation and annihilation operators* (pages 373–378).

The notion of *variance* appears here in a new light, which we now outline.

The *isomorphisms of the prequantum structure* of Y (or "*quantomorphisms*") form a group. Each quantomorphism "descends" to the space U of classical motions and defines a *symplectomorphism* there.

Conversely one can try to "lift" the symplectomorphisms of U to quantomorphisms of Y. The problem of lifting a *dynamical group* to a group of quantomorphisms leads to some nontrivial situations.

For example, in the case of a linear system, the group of translations lifts to a nonabelian group: the *Weyl group* (p. 336). In the case of a nonrelativistic particle, the elements of the Galilei group are individually liftable, but the whole group *does not act* on the prequantum manifold Y and thus not on the solutions of the Schrödinger equation. V. Bargmann has shown the *cohomological* origin of this fact — an origin which is obvious here (p. 358). In the case of a relativistic particle with integer spin[35] one can make the *Poincaré group* act on the prequantum manifold Y. On the other hand, in the case of spin $\hbar/2$, it is not the Poincaré group that can be lifted, but a *covering space* (p. 13) of this group. This explains the paradoxical variance of the geometrical objects, called *spinors*, that are used to formulate the spatio-temporal form of the Dirac equation: spinors "only make a half turn when space makes a full turn".[36]

The description of a "state" is a fundamental problem in quantum mechanics. Initially, a state was a *unit vector* of the Hilbert space on which the observables act. Using the construction of the Hilbert space proposed here (pages 339–345), it is easy to associate to each "quantum state" a "statistical state", that is, a probability measure on the space of classical motions. In this way the *probabilistic interpretation* of quantum

[35]In the case of a photon one recovers in this way the *vectorial* nature of the Maxwellian potential A, which corresponds to a *scalar* function Ψ on the prequantum manifold Y (pages 366–369).

[36]One can look for other ways to explain this paradox. For example postulating a direct interpretation of the covering space of the Poincaré group (which boils down to modifying relativity in a manner which seems to be incompatible with general relativity) or describing the wave function by ordinary geometrical objects, chosen so that they describe the spinor *up to a sign* (but such a description destroys the linear character of the wave equation). Such artifices become unnecessary here.

mechanics (p. 348) can be formalized.

However, it was soon discovered that this idea of a state is not sufficient to cover all cases found in experiments. For example, the states of quantum chemistry are described by the formalism of a "density operator" (p. 382). More fundamentally, there exist states that have a vectorial description but that do not belong to the initially postulated Hilbert space.[37]

These difficulties disappear in the axiomatic formulation of quantum mechanics based on C^*-algebras. We propose here a solution that is roughly analogous. It consists of considering the *quantum states* as functions defined *on the group of quantomorphisms* of Y and satisfying certain axioms (p. 382). In particular, these states are "functions of positive type" on this group (p. 380).

This definition allows us to calculate *quantum average values* (p. 382). In certain cases one can associate to these states new Hilbert spaces and vectors in these spaces (due to the Gelfand-Naimark-Segal construction (p. 384)). This solves certain difficulties but poses new problems.

In any case, the method of quantization proposed here remains conjectural. It has the advantage of being intrinsic, of giving a precise sense to the correspondence principle (p. 346), and of uniting into one coherent view various aspects of quantum mechanics whose logical relations would not otherwise be evident. On the other hand, it leaves open many mathematical problems and questions of interpretation. We present this method of quantization above all as a subject of reflection for the researcher.

This book contains developments of pure mathematics which can be read independently from the rest. Reading them only requires knowledge at the undergraduate level.

Chapter I is an outline of *essential results from differential geometry* (see the table of contents for details). We only give proofs which bear directly on the understanding of the questions treated. The reader can find the omitted proofs in specialized treatises.

Chapter II is devoted to *symplectic geometry*. It contains not only the classical results of this theory (see the table of contents) but also a detailed study of theorems of *cohomological* type called for by the abstract notion of a dynamical group. It also treats the inverse procedure of constructing a symplectic manifold starting from a Lie group.[38]

[37] As is the case here for the *wave functions* of elementary particles (p. 357).

[38] This algorithm plays an essential role in recent research on the theory of unitary

§16 of chapter IV is devoted to the notion of a *measure* on a manifold. Those proofs that are not given, can be found in Bourbaki's treatise on integration. The text is illustrated with various examples: convolution on a Lie group, random variables, moments, Gaussian distribution, Gibbs ensemble of a dynamical group, and classical photometry.

§18 is, as others, purely mathematical. It seems that the "geometric quantization" treated there, may be useful in treating problems of unitary representation of Lie groups.[39] We use the opportunity to recall the essential definitions of this theory, in particular Hilbert spaces and C^*-algebras.

The physicist reader could start reading this book at chapter III, which starts with the most classical aspect of mechanics, followed by §17 (statistical mechanics), and then §19 (quantization of systems). He will see how certain *mathematical structures* appear naturally in the study of dynamical systems. The goal of this book would be attained if this would give him the desire to delve into the study of these structures and to use them as a tool in his research.

representations of Lie groups, following the works of, in particular, Dixmier, Kirillov, Bernat, Gelfand, Kostant, and Auslander.

[39]See Streater [34].

Chapter I

Differential Geometry

§1. Manifolds

The definition of manifolds

Let V be a set and let n be a positive integer. A *coordinate system* on V is a prescription which to n real numbers ${}^1x, {}^2x, \ldots, {}^nx$ assigns a point X of V; the ${}^jx$ are called the *coordinates* of X in the system F.[40]

It will be convenient to write

$$x = \begin{pmatrix} {}^1x \\ {}^2x \\ \vdots \\ {}^nx \end{pmatrix}, \qquad X \equiv F(x). \tag{1.1}$$

(1.2) Thus a coordinate system F will be considered as a map from $\mathbf{R}^n$ to V.[41] We will make the essential assumption that F is *injective* [42] and that $\mathrm{dom}(F)$ is *open*.[43]

EXAMPLE: If E is a real vector space of dimension n and if the vectors $S_1, S_2, \ldots, S_n$ constitute a basis of E, then one can define a coordinate system S on E by

$$S(x) = S_1\ {}^1x + S_2\ {}^2x + \cdots + S_n\ {}^nx\,, \tag{1.3}$$

or, using matrix notation,

$$S = [S_1\ S_2\ \ldots\ S_n]\,. \tag{1.4}$$

---

[40] Following tensor notation, we write the indices of the coordinates as superscripts; they are placed to the left for reasons which will become clear later.

[41] If E and E' are sets, then by a map from E to E' we will mean any map A defined on a subset of E (denoted by $\mathrm{dom}(E)$) onto a subset of E' (denoted by $\mathrm{im}(A)$).

[42] We will say that a map A is *injective* (one to one) if $[A(x) = A(x')] \Rightarrow [x = x']$. In such a case A possesses an inverse, written as A^{-1}, which is characterized by $\mathrm{dom}(A^{-1}) = \mathrm{im}(A)$, $\mathrm{im}(A^{-1}) = \mathrm{dom}(A)$, and

$$[A^{-1}(x) = y] \Longleftrightarrow [x = A(y)]\,.$$

It follows that A^{-1} is injective and that $[A^{-1}]^{-1} = A$.

[43] The *Euclidean norm* of $x \in \mathbf{R}^n$ is defined as the number

$$\|x\| = \sqrt{[{}^1x]^2 + \cdots + [{}^nx]^2}\,.$$

A *ball* (respectively *sphere*) of *center* y and *radius* r is the set of all x in $\mathbf{R}^n$ such that $\|x - y\| < r$ (respectively $\|x - y\| = r$); an *open* set of $\mathbf{R}^n$ is any union of balls.

We will simply say that the map S, defined by (1.3) or (1.4), is a *basis* of E. □

(1.5) If F and G are two coordinate systems (Fig. 1.I), it is clear that $F^{-1} \circ G$ is a map from $\mathbf{R}^n$ to $\mathbf{R}^n$;[44] we will say that F and G are *compatible* if $F^{-1} \circ G$ and its inverse $G^{-1} \circ F$ are differentiable.[45]

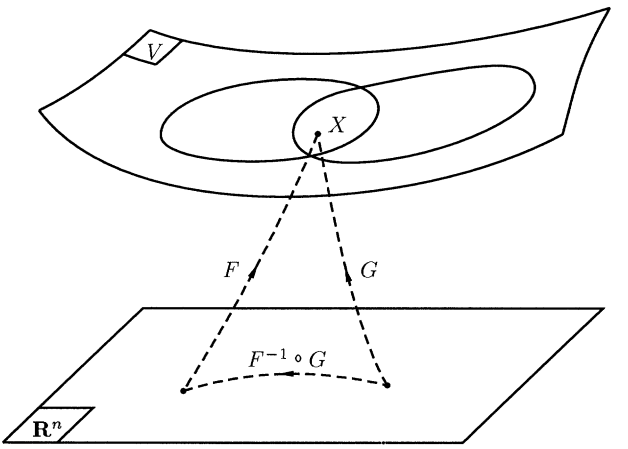


Fig 1.I

(1.6) An *atlas* of V is a set $\mathcal{A}$ of coordinate systems such that

a) The elements of $\mathcal{A}$ are pairwise compatible.

b) The sets $\mathrm{im}(F)$, $F \in \mathcal{A}$ cover V.

---

[44]We denote by $A \circ B$ or $A \cdot B$ the composition of the maps A and B; it is defined by

$$\{x \in \mathrm{dom}(A \circ B)\} \Longleftrightarrow \{x \in \mathrm{dom}(B), B(x) \in \mathrm{dom}(A)\} \Longrightarrow \{[A \circ B](x) = A(B(x))\}.$$

Of course, if $\mathrm{im}(B)$ and $\mathrm{dom}(A)$ do not intersect, $\mathrm{dom}(A \circ B)$ is empty. In such a case we will say that the map $A \circ B$ is *empty*; it does not constitute an exception, in particular not to the associative law

$$[A \circ B] \circ C = A \circ [B \circ C],$$

nor to the behavior of composition under inversion

$$[A \circ B]^{-1} = B^{-1} \circ A^{-1},$$

which is valid if A and B are injective.

[45]A map A from $\mathbf{R}^n$ to $\mathbf{R}^{n'}$ will be called *differentiable* if $\mathrm{dom}(A)$ is *open* and if for every $x \in \mathrm{dom}(A)$ the coordinates ${}^k y$ of $y \equiv A(x)$ have *continuous partial derivatives of all orders* with respect to the coordinates ${}^j x$ of x (we will have no occasion to use functions which are p *times differentiable* and which may not have partial derivatives of order higher than p). By convention we will say that the empty map is differentiable.

(1.7) Once an atlas $\mathcal{A}$ has been chosen, an *admissible coordinate system*, or for short a *chart of* V, is any coordinate system compatible with those of $\mathcal{A}$; it can be shown that *two charts are compatible*. It follows that the *set of all charts* is again an atlas and that this atlas is the largest atlas containing $\mathcal{A}$.

(1.8) Once we have defined in this way the set of all charts of V, we will say that V is a *manifold*, or, more precisely, that we have defined a *manifold structure*[46] on V; the number n is then called the *dimension* of V.

(1.9) EXAMPLE: If E is a vector space of dimension n, one can verify that the set of all bases of E (1.4) is an *atlas*; thus it gives E the structure of a manifold (of dimension n, as it should be).

(1.10) In particular $\mathbf{R}^n$ itself is a manifold of dimension n; one sees immediately that its charts are maps from $\mathbf{R}^n$ to $\mathbf{R}^n$ *which are differentiable as well as their inverses.* □

(1.11) EXAMPLE: Let S_n denote the sphere with center 0 and radius 1 in $\mathbf{R}^{n+1}$. One can form an atlas of S_n consisting of two charts F_+ and F_- by defining, for $x \in \mathbf{R}^n$:

$$\begin{aligned} {}^{j}[F_{\pm}(x)] &= \frac{2\,{}^{j}x}{1+\|x\|^2} \qquad [j = 1, 2, \ldots, n] \\ {}^{n+1}[F_{\pm}(x)] &= \pm\frac{1-\|x\|^2}{1+\|x\|^2}. \end{aligned}$$

This atlas gives S_n the structure of a manifold of dimension n. □

(1.12) Let V and V' be two manifolds of dimension n and n' respectively, possessing charts F and F' respectively. If we define

$$\Phi\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} F(x) \\ F'(x') \end{pmatrix},$$

we obtain a coordinate system Φ for the set theoretic product $V \times V'$; varying F and F' we obtain an atlas of $V \times V'$. In this way $V \times V'$ has a manifold structure of dimension $n+n'$; we will say that this is the *direct product manifold.*

This can be extended immediately to the product of p manifolds V_j; this yields a manifold whose dimension is the sum of those of the V_j.[47]

[46] One can give examples of sets V that possess two incompatible atlases; in such a case, these atlases define two different manifold structures on V (see (5.7)).

[47] Note that the manifold $\mathbf{R}^n$ (see (1.10)) is the direct product of n manifolds, all equal to $\mathbf{R}$.

Open sets

(1.13) DEFINITION: A subset E of a manifold V is *open* if E is the union of sets consisting of images of charts.[48]

(1.14) One can verify that: any union of open sets is open, the intersection of two open sets is open, the empty subset of V is open, and V is open in V.[49] □

(1.15) DEFINITION: Let F be a map from a manifold V to a manifold V'. F is *continuous* if for each open set $E' \subseteq V'$ there is an open set $E \subseteq V$ such that[50] $F^{-1}(E') = E \cap \mathrm{dom}(F)$. □

In the special case where $V = \mathbf{R}^n$ and $V' = \mathbf{R}^{n'}$, one recovers the usual definition of continuous functions of real variables.

(1.14) shows that:

$$\left[\begin{array}{l} F \text{ continuous} \\ \mathrm{dom}(F) \text{ open} \end{array}\right] \iff [E' \text{ open} \Longrightarrow F^{-1}(E') \text{ open}]. \tag{1.16}$$

Differentiable maps

(1.17) Let V and V' be two manifolds (of dimension n and n' respectively), and let A be a map from V to V'. We will say that A is *differentiable*[51] if for every chart F of V and every chart F' of V', the map $F'^{-1} \circ A \circ F$ (which is a map from $\mathbf{R}^n$ to $\mathbf{R}^{n'}$) is differentiable in the sense of (1.5[6]).

(1.18) EXAMPLE: The charts of a manifold and their inverses are differentiable. □

(1.19) EXAMPLE: If E is an open subset of a manifold, then the map 1_E is differentiable.[52] □

EXAMPLE: Let $V_1, V_2, \ldots, V_p$ be manifolds and let us denote by V their

[48] This definition is compatible with (1.2[4]) in case $E = \mathbf{R}^n$.

[49] This shows that *every manifold is a topological space*; see for example Bourbaki [10].

[50] We denote by $F^{-1}(E')$ the *inverse image* of E' under F, that is, the set of $x \in \mathrm{dom}(F)$ such that $F(x) \in E'$.

[51] This definition is compatible with (1.5[6]) in case $V = \mathbf{R}^n$ and $V' = \mathbf{R}^{n'}$.

[52] We denote by 1_E the identity map from E to E: $\mathrm{dom}(1_E) = E$ and $[x \in E] \Rightarrow [1_E(x) = x]$.

direct product (1.12). The j^{th} projection, denoted by ${}^{j}|$ and defined by

(1.20)
$$ {}^{j}\left|\begin{pmatrix} {}^{1}x \\ {}^{2}x \\ \vdots \\ {}^{n}x \end{pmatrix}\right. = {}^{j}x\,, $$

is a *differentiable map* from V onto V_j. □

(1.21) One can show that every differentiable map is *continuous* and that its domain of definition is *open*.

(1.22) Similarly if A and B are differentiable maps, then $A \circ B$ is differentiable.[53]

The tangent space

(1.23) THEOREM:

1) If V is a manifold of dimension n and x a point of V, one can define a *vector space of dimension* n, called the *tangent space to* V *at* x. We will denote it by D_x; its elements are called *tangent vectors* to V at x.
2) If A is a differentiable map from V to V' and if $x \in \mathrm{dom}(A)$, one can define a *linear map*, called the *tangent map* to A at x and denoted by $D(A)(x)$. $D(A)(x)$ maps the vector space D_x to the vector space $D_{A(x)}$.
3) If E is an open subset of V with $x \in E$ and $y \in D_x$, then
$$ D(1_E)(x)(y) = y\,. $$
4) Let ${}^{1}V$, ${}^{2}V, \ldots {}^{p}V$ be manifolds and let us choose a point ${}^{j}x$ in each ${}^{j}V$ and a tangent vector ${}^{j}y$ in each $D_{{}^{j}x}$. If we write
$$ x = \begin{pmatrix} {}^{1}x \\ {}^{2}x \\ \vdots \\ {}^{p}x \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} {}^{1}y \\ {}^{2}y \\ \vdots \\ {}^{p}y \end{pmatrix}\,, $$
then
$$ D({}^{j}|)(x)(y) = {}^{j}y\,.^{54} $$

[53] We assume of course that B takes its values in the manifold on which A is defined.
[54] See definition (1.20) of the projection ${}^{j}|$.

5) Let E and E' be finite dimensional vector spaces and let A be a differentiable map from E to E'. For $x \in \mathrm{dom}(A)$ and $y \in E$ we have

$$D(A)(x)(y) = \lim_{s \to 0} \frac{1}{s} \left[A(x + sy) - A(x)\right].$$

6) If A and B are differentiable maps[55] and if $x \in \mathrm{dom}(A \circ B)$, then

$$D(A \circ B)(x) = \left[D(A)(B(x))\right] \cdot \left[D(B)(x)\right]. \qquad \square$$

From this theorem the reader can easily deduce the following results.

(1.24) If a map F is *constant*, that is, $\forall x, y \in V : F(x) = F(y)$, it is differentiable and $D(F)(x) = 0$. $\square$

(1.25) If E is a finite dimensional vector space, the tangent space D_x *coincides with* E *for every* $x \in E$. This is true in particular for $E = \mathbf{R}^n$.

(1.26) A *diffeomorphism* from V to V' is an injective map A from V to V' such that A and A^{-1} are differentiable.[56] $\square$

Hence for $x \in \mathrm{dom}(A)$, $D(A)(x)$ is an *injective* linear map from D_x *onto* $D_{A(x)}$ and

(1.27) $$[D(A)(x)]^{-1} = D(A^{-1})(A(x)),$$

which shows that V and V' have *the same dimension.*

Let X be a point of a manifold V of dimension n. Then there exists a chart F and a point $x \in \mathbf{R}^n$ such that $F(x) = X$.

The operator

(1.28) $$S = D(F)(x)$$

is an injective linear map from $\mathbf{R}^n$ onto D_X and thus is a *basis* of the tangent space D_X in the sense of (1.4).

Writing

(1.29) $$|_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; |_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, |_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

one immediately obtains the following expression for the *basis vectors* of D_X

(1.30) $$S_j = S(|_j) = D(F)(x)(|_j).$$

[55] We assume that B takes its values in the manifold on which A is defined.

[56] Example: the charts of a manifold V of dimension n are diffeomorphisms from $\mathbf{R}^n$ to V.

(1.31) DEFINITION: V being a manifold, we denote by V^D the set of pairs

$$\begin{pmatrix} y \\ x \end{pmatrix} \qquad (x \in V,\ y \in D_x);$$

V^D is called the *tangent bundle* of V. The function which maps the point $\begin{pmatrix} y \\ x \end{pmatrix}$ to the point x is called the *projection* of V^D onto V. □

Let A be a differentiable map from a manifold V to a manifold V' (Fig. 1.II). We will denote by A^D the map defined by

$$(1.32) \qquad A^D \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} D(A)(x)(y) \\ A(x) \end{pmatrix} \qquad \forall \begin{pmatrix} y \\ x \end{pmatrix} \in V^D.$$

It is clear that A^D is a map from V^D to V'^D. It is sometimes called the *lift* of A on account of figure 1.II.

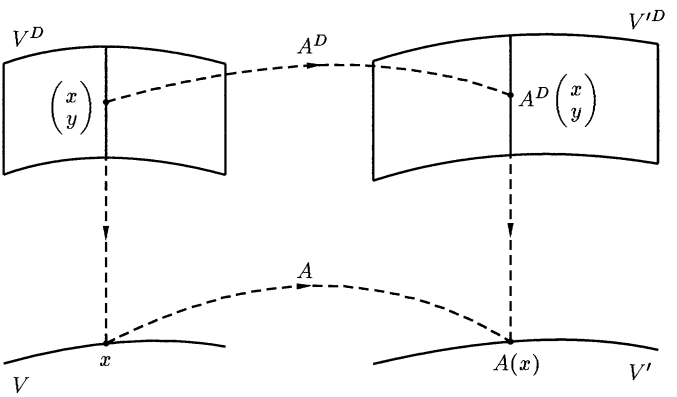


Fig 1.II

The formulæ (1.23) allow us to show immediately that

$$(1.33) \qquad \begin{aligned} [A \circ B]^D &= A^D \circ B^D \\ \left[A^{-1}\right]^D &= \left[A^D\right]^{-1} \quad \text{if } A \text{ is a diffeomorphism} \\ [1_E]^D &= 1_{E'} \quad \text{if } E \text{ is open in } V \text{ and } E' \text{ its inverse image under the projection of } V^D \text{ onto } V. \end{aligned}$$

In particular, let F be a *chart* of a manifold V. In the formula

$$F^D \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} D(F)(x)(y) \\ F(x) \end{pmatrix}$$

x and y are elements of $\mathbf{R}^n$. Thus the pair $\begin{pmatrix} y \\ x \end{pmatrix}$ is an element of $\mathbf{R}^{2n}$. Therefore F^D is a *coordinate system* for V^D. It can easily be verified that the set of all F^D is an atlas (definition (1.6)). Therefore:

(1.34) If V is a manifold of dimension n, V^D is a *manifold of dimension* $2n$. The set of all F^D (F being a chart of V) is an *atlas* of V^D. With this structure, the projection and the maps A^D (1.32) are differentiable. □

Thus V^D can henceforth be called the *manifold of tangent vectors* to V.

Submanifolds

(1.35) THEOREM: Let V and V' be two manifolds of dimension n and n' respectively (Fig. 1.III), let A be a differentiable map from V to V', let x be a point of $\mathrm{dom}(A)$, and let r be the rank of $D(A)(x)$.

1) If $r = n$, there exists an open subset E of V containing x such that A is *injective on* E and $r = n$ at every point of E.
2) If $r = n'$, there exists an open subset E' of V' containing $A(x)$ such that $E' \subseteq \mathrm{im}(A)$.[57]
3) If $r = n = n'$, there exists an open set E containing x such that the restriction A' of A to E is a *diffeomorphism.* □

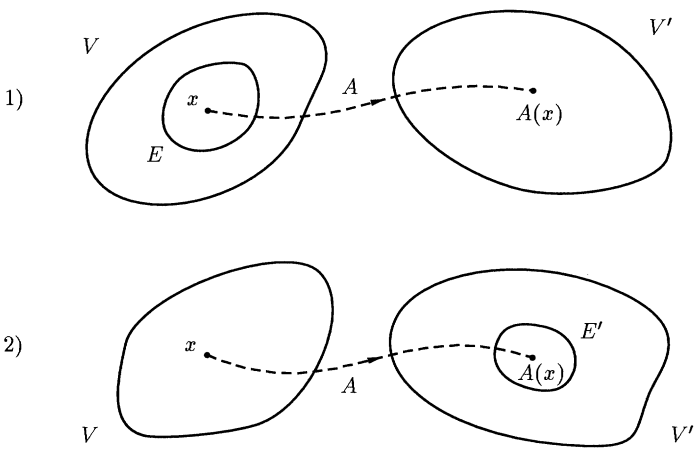


Fig 1.III

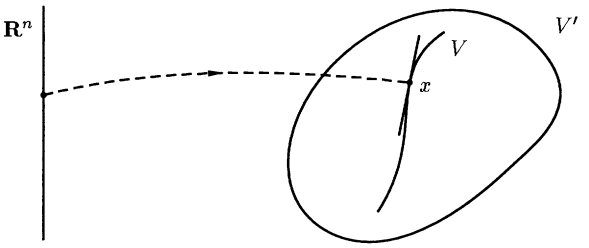
[57] $\mathrm{im}(A)$ is called a *neighborhood* of $A(x)$; as the image under A of an open subset of V is open in V', A is called an *open* map.

(1.36) DEFINITION: Let V and V' be two manifolds. An *injective immersion*[†] of V to V' is a map A from V to V' such that A is differentiable, injective, and for all $x \in \mathrm{dom}(A)$, $D(A)(x)$ is injective. □

The following results can be established immediately.

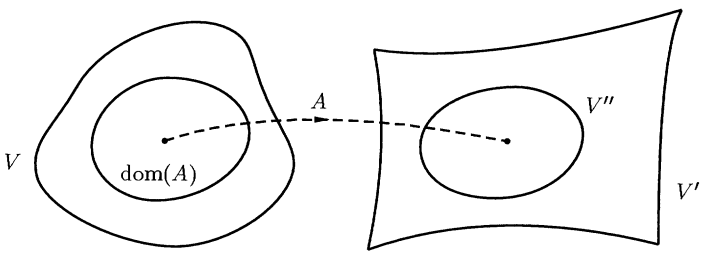
(1.37) A manifold V (of dimension n) is said to be a *submanifold* of a manifold V' (of dimension n') if the charts of V are injective immersions of $\mathbf{R}^n$ into V' (Fig. 1.IV); in particular V is a *subset* of V'.[†] At any point x of V the tangent space to V is a *vector subspace* of dimension n of the tangent space to V', and thus $n \leq n'$. □

Fig 1.IV



(1.38) Let V and V' be two manifolds of dimensions n and n' respectively. If A is an injective immersion of V into V', the set $V'' = \mathrm{im}(A)$ admits the structure of a *submanifold* of V', characterized by the fact that A is a *diffeomorphism* from V to V''. The charts of V'' are the $A \circ F$, where F is a chart of V (Fig. 1.V). □

Fig 1.V



(1.39) EXAMPLE: Let V' be a manifold. It follows from (1.35.2) that any submanifold V of V' with the *same dimension* as V' is *open* in V'.

[†]Editors' note: A map A is called an "immersion" without the adjective "injective" if A is differentiable and for all $x \in \mathrm{dom}(A)$, $D(A)(x)$ is injective.

[†]Editors' note: This notion of submanifold, though not uncommon, is *not* the standard definition of a submanifold. Here the canonical injection $i : V \to V'$ is an injective immersion. The standard definition requires in addition that the topology of V coincides with the topology induced by V', that is, $i : V \to V'$ is an embedding.

Conversely, if V is an open subset of V', 1_V is obviously an injective immersion of V into V', and thus its image V is a manifold of dimension equal to that of V'. One can verify immediately that the charts of V are those charts of V' whose images are contained in V. □

Thus from now on we can consider any *nonempty open subset* of a manifold as a *manifold of the same dimension.*

Manifolds defined by an equation

(1.40) THEOREM: Let V and V' be manifolds of dimension n and n' respectively with $n' < n$. Let A be a differentiable map from V to V', let y_0 be a point in V', and let V_0 be the set of points $x \in V$ satisfying the equation

$$A(x) = y_0 \qquad \text{(Fig. 1.VI).}$$

If

$$\operatorname{rank}(D(A)(x)) = n' \qquad \forall x \in V_0\,,$$

then V_0 admits a unique structure of a *submanifold of* V [58] *of dimension* $n - n'$. □

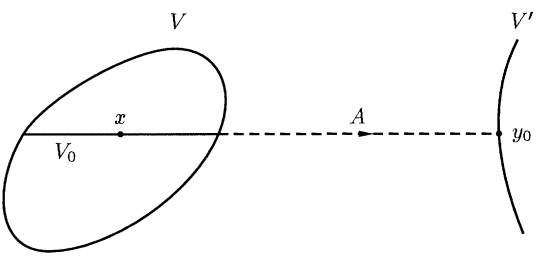


Fig 1.VI

It is clear that the charts of V_0 are the injective immersions F of $\mathbf{R}^{n-n'}$ into V satisfying

$$A(F(\xi)) = y_0 \qquad \forall \xi \in \operatorname{dom}(F)\,.$$

Theorem (1.40) shows that these charts form an atlas. Taking the derivative of the above equation, the reader can ascertain that

(1.40) THEOREM (continued): The tangent space to V_0 at x is the *kernel* of $D(A)(x)$. □

[58] See also (5.12).

(1.41) EXAMPLE: Let us take $V = \mathbf{R}^n$, $V' = \mathbf{R}$, $A(x) \equiv \|x\|^2$, and $y_0 = 1$. Then V_0 is the *sphere* S_{n-1}. We have already shown that S_{n-1} is an $n-1$ dimensional manifold and we have given an atlas consisting of two charts (1.11). It is easy to verify that these charts are injective immersions, and to construct the tangent space to S_{n-1} at a point x. □

Covering spaces

(1.42) DEFINITION: Let V be a manifold; a triple (W, G, P) is called a *covering space* of V provided that

a) W is a manifold of the same dimension as V;
b) G is a group of permutations of W such that

$$[g \in G, g(x) = x] \quad \Longrightarrow \quad [g = 1_W]\,;$$

c) P is a differentiable map from W onto V such that $D(P)(x)$ is injective for any $x \in W$ and

$$[P(x) = P(y)] \quad \Longleftrightarrow \quad [\exists g \in G, y = g(x)]\,. \qquad \square$$

Under these conditions the reader can easily verify that

(1.43) a) Any point x of W belongs to an open set E whose images under the elements of G are pairwise disjoint (Fig. 1.VII and 1.VIII);
b) The elements of G are diffeomorphisms. □

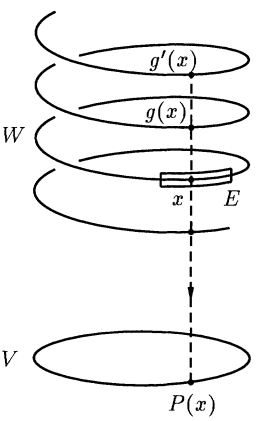


Fig 1.VII

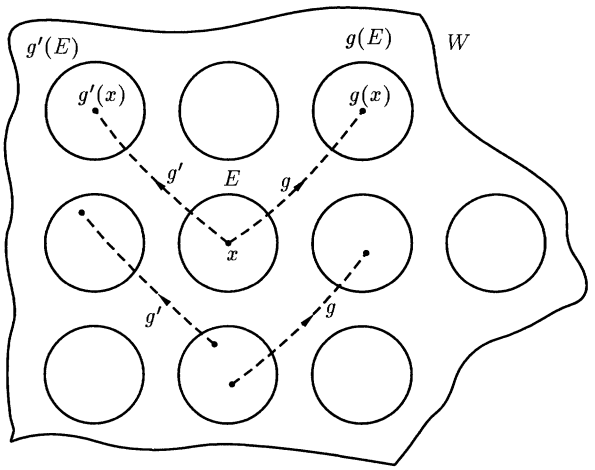


Fig 1.VIII

Quotient manifolds

(1.44) Conversely, let W be a set and G a group of permutations of W. Let us recall that the *orbit* of $x \in W$ under G is the subset $P(x)$ of W such that

$$[y \in P(x)] \iff [\exists g \in G : y = g(x)],$$

and thus $[y \in P(x)] \iff [P(y) = P(x)]$. The set $\mathrm{im}(P)$, denoted by W/G, is called the *quotient* of W by the group G.

(1.45) If W is a manifold of dimension n and if G satisfies the conditions (1.43) above,[59] it can be shown that the quotient W/G can be given the structure of a manifold of dimension n in such a way that (W, G, P) is a covering space of W/G. The injective maps of the form $P \circ F$, with F a chart of W, form an *atlas of the quotient manifold* W/G.

Connectedness

(1.46) DEFINITION: Let V be a manifold of dimension n; a subset E of V will be called a *cell* if there is a chart F such that $\mathrm{dom}(F) = \mathbf{R}^n$ and $\mathrm{im}(F) = E$.

It follows that each cell is open.

[59] One says that G is a *discrete group of diffeomorphisms*.

(1.47) EXAMPLE: One can easily verify that a nonempty ball in $\mathbf{R}^n$ is a cell. It follows that every point of a manifold V is contained in a cell in V.

(1.48) DEFINITION: Let E be an open subset of a manifold V; one says that E is *connected* if one cannot decompose E as the disjoint union of two nonempty open sets.[60]

EXAMPLE: It can be shown that any *cell* is connected.

(1.49) One can easily deduce from this that every manifold V decomposes as a union of pairwise disjoint connected open sets called *components*. This decomposition is *unique*: the component containing a point x is the *union of all connected open sets containing* x. It follows that a manifold is connected if it has *a single component*. □

(1.50) One can show that every direct product of connected manifolds is connected.

(1.51) If A is a continuous map from a connected manifold V to a manifold V', then $\mathrm{im}(A)$ is contained in a single component of V'.

(1.52) One can also show that two points x and y belong to the same component if there exists a path in V which joins x to y, that is, if there exists a continuous map F from the interval $[0,1]$ to V such that $F(0) = x$ and $F(1) = y$. The path $\mathrm{im}(F)$ is then completely contained in this component.



Fig 1.IX

[60] One says that E is the *disjoint* union of two sets E_1 and E_2 if $E = E_1 \cup E_2$ and $E_1 \cap E_2 = \emptyset$.

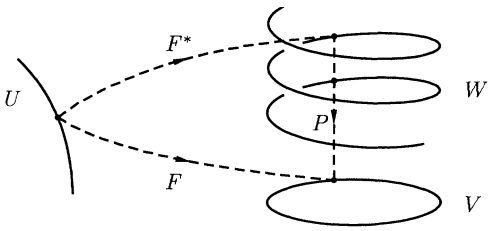
Homotopy

(1.53) A connected manifold U is called *simply connected* if for any continuous map F from U to a manifold V and for any covering space (W, G, P) of V, F can be factored as

$$F = P \circ F^*,$$

for some *continuous* map F^* from U to W (Fig. 1.X). □

Fig 1.X



(1.54) EXAMPLES: It can be shown that $\mathbf{R}$ is simply connected, that a direct product of simply connected manifolds is simply connected, and that the image $\mathrm{im}(A)$ of a simply connected manifold $\mathrm{dom}(A)$ by a diffeomorphism A is simply connected. It follows that each cell is simply connected. The sphere S_n too is simply connected for $n \geq 2$, although it is not a cell. □

(1.55) A covering space (W, G, P) of a manifold V is said to be *universal* if W is simply connected. One can show that *every connected manifold possesses a universal covering space.* □

(1.56) Let (W, G, P) be a universal covering space of a manifold V and let (W', G', P') be another covering space of V such that W' is connected (Fig. 1.XI).

One thus can factor P as $P' \circ P^*$. One can easily verify that G possesses a subgroup G^* such that (W, G^*, P^*) is a *covering space* of W', that G^* is a normal subgroup of G, and that the group G' is isomorphic to the quotient G/G^*. We thus see that a universal covering space can be considered as a *maximal* connected covering space.

(1.57) If (W', G', P') is also a universal covering space, it follows that P^* is injective and that the groups G and G' are isomorphic. The group G is thus defined, up to isomorphism, by V. It is called the *first homotopy group*[61]

[61] Alternatively the *fundamental group* or the *Poincaré group* of the connected manifold V.

(1.58) EXAMPLE: A simply connected manifold is a manifold whose homotopy group contains only the identity element. □

(1.59) EXAMPLE: Let G be the group (isomorphic to the additive group $\mathbf{Z}$ of all integers) consisting of the transformations g_k of $\mathbf{R}$ defined by $g_k(\varphi) = \varphi + 2k\pi \ \forall \varphi \in \mathbf{R}, \ \forall k \in \mathbf{Z}$. If we define $P(\varphi) = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$, then $(\mathbf{R}, G, P)$ is a universal covering space of the circle S_1 (notation (1.11)); *the first homotopy group of* S_1 *thus equals* $\mathbf{Z}$. □

(1.60) If (W_j, G_j, P_j) is a covering space of the manifold V_j $(j = 1, 2, \ldots, p)$, one immediately obtains a *product covering space* $(W_1 \times \cdots \times W_p, G_1 \times \cdots \times G_p, P)$ of the direct product $V_1 \times \cdots \times V_p$. If the given covering spaces are universal, then so is the direct product covering space. *Therefore the homotopy group of a direct product is the direct product of the homotopy groups.*

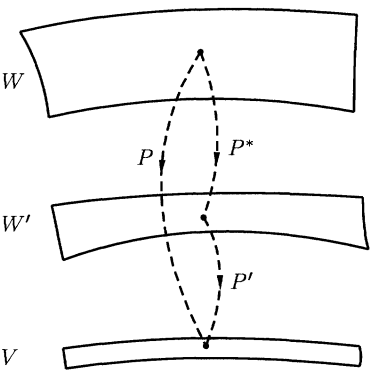


Fig 1.XI

§2. Derivations

Variables

Let us start by elucidating the more or less classical notion of a variable which we will use throughout the rest of this book.

(2.1) To define a system of variables, one starts by choosing a set E. Next, one calls *variable* any graphic symbol — for example, the letter y — to which one associates a map; we will refer to this map by writing the word "value" to the left of this symbol (for example, value y). This map is supposed to be defined on a subset of E, say E'. For $a \in E'$, we thus have defined the object $[\text{value}\, y](a)$, called *the value of the variable y at the point a.*

(2.2) In particular, an *independent variable* is a variable, say x, to which one associates the identity map of E:

$$[\text{value}\, x](a) = a \qquad \forall a \in E\,.^{62}$$

We will use the notation

(2.3) $$y \equiv z$$

to say that the maps $[\text{value}\, y]$ and $[\text{value}\, z]$ are equal, that is,

$$[\text{value}\, y](a) = [\text{value}\, z](a)$$

for every a for which one of the members of this equality exists.

We will admit that an algebraic expression F containing variables $y, z, \ldots$ defines a new variable by the rule

(2.4) $$[\text{value}\, F(y, z, \ldots)](a) = F([\text{value}\, y](a), [\text{value}\, z](a), \ldots) \qquad \forall a \in E\,.$$

As a particular instance of this rule, let us note the case of an algebraic expression *not containing any variables.* We then have

(2.5) $$[\text{value}\, F](a) = F \qquad \forall a\,,$$

and we will say that the variable F thus defined is *constant.*

[62] The object $[\text{value}\, y](a)$ can therefore be called "the value of y when the value of x is a".

(2.6) Another important special case is the one where the algebraic expression F can assume one of the two values "true" or "false". If one replaces certain objects in the algebraic expression F by variables, then, according to the rule (2.4), one obtains a new variable called a *boolean* variable. Boolean variables are also called *propositions*.

EXAMPLE: If x and y are two boolean variables, the expression $x = y$ is a proposition. This should not be confused with the expression $x \equiv y$.

(2.7) A variable z is called a *function* of a variable y if there exists a map F such that

$$z \equiv F(y) \qquad \text{(in the sense of (2.4))},$$

which can also be written as

$$[\text{value}\, z](a) = F([\text{value}\, y](a)) \qquad \forall a \in E\,,$$

or as

$$[\text{value}\, z] = F \circ [\text{value}\, y]\,.$$

It is clear that the map F is unique if one requires that its domain of definition be included in the set of values of y. One often denotes F as

(2.8) $$y \mapsto z\,,$$

so that the map $[\text{value}\, y]$ can also be denoted as $x \mapsto y$, x being the independent variable.

Vector fields and derivations

(2.9) Let V be a manifold. A *vector field* on V is a map f defined on a subset E of V such that, for each $a \in E$, $f(a)$ is a tangent vector to V at a:

$$[a \in \text{dom}(f)] \implies [f(a) \in D_a]\,. \qquad \square$$

To define a *derivation* δ, one chooses a manifold V, a vector field f on V, and a variable x which ranges over V (meaning that V is the image of x [63]). To each variable y such that the map $x \mapsto y$ exists and is differentiable, one associates the new variable

(2.10) $$\delta y \equiv D(x \mapsto y)(x)(f(x))\,.$$

[63] $V = \text{im}([\text{value}\, x])$.

(2.11) EXAMPLE: If y is constant, we have $\delta y \equiv 0$. Substituting $y \equiv x$ into (2.10), we find $\delta x = f(x)$. □

(2.12) It follows that we can write *any vector field* in the form $x \mapsto \delta x$ for some derivation δ. □

Applying formula (1.23.6), one can see that

$$\delta[F(y)] \equiv D(F)(y)(\delta y)\,, \tag{2.13}$$

provided F is differentiable on the manifold where y takes its values. This formula applies to any derivation δ; it presupposes that F is a given differentiable function and not a variable.

If $z \equiv F(y)$ we will write

$$\frac{\partial z}{\partial y} \equiv D(F)(y)\,,^{64} \tag{2.14}$$

and thus we can write (2.13) in the form

$$\delta z \equiv \frac{\partial z}{\partial y}(\delta y)\,, \tag{2.15}$$

provided $y \mapsto z$ is differentiable. Notice that formulæ (1.23.6) and (1.27) show that

$$\frac{\partial z}{\partial x} \equiv \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} \tag{2.16}$$

if $x \mapsto y$ and $y \mapsto z$ are differentiable, and that

$$\frac{\partial y}{\partial z} \equiv \left[\frac{\partial z}{\partial y}\right]^{-1} \tag{2.17}$$

if $y \mapsto z$ is a diffeomorphism.

Let x be a variable whose values lie in a *direct product* of manifolds. Writing the projections ${}^{j}|$ explicitly (see (1.20)):

$${}^{j}x \equiv {}^{j}|(x)\,, \tag{2.18}$$

and using (1.23.4), one can see that

$$\delta \begin{pmatrix} {}^{1}x \\ \vdots \\ {}^{p}x \end{pmatrix} \equiv \begin{pmatrix} \delta\,{}^{1}x \\ \vdots \\ \delta\,{}^{p}x \end{pmatrix} \tag{2.19}$$

[64] Recall that $D(F)(y)$ is a linear map from the vector space D_y to the vector space D_z.

for any derivation δ. The tangent space to a direct product manifold is the direct product of the tangent spaces.

If $y \mapsto x$ is a differentiable map *with values in a direct product*, formula (2.19) can be written in matrix form as

$$\frac{\partial}{\partial y}\begin{pmatrix} {}^1x \\ \vdots \\ {}^px \end{pmatrix} \equiv \begin{pmatrix} \frac{\partial\, {}^1x}{\partial y} \\ \vdots \\ \frac{\partial\, {}^px}{\partial y} \end{pmatrix}. \tag{2.20}$$

In case of a differentiable map $x \mapsto y$ *defined on a direct product*, δy (which depends linearly on δx) can be written as

$$\delta y \equiv \frac{\partial y}{\partial\, {}^1x}(\delta\, {}^1x) + \cdots + \frac{\partial y}{\partial\, {}^px}(\delta\, {}^px)\,, \tag{2.21}$$

a form often called a "total differential". The $\frac{\partial y}{\partial\, {}^jx}$ defined by this formula are the *partial derivatives*. In matrix form this formula can be written as

$$\frac{\partial y}{\partial x} = \left[\frac{\partial y}{\partial [{}^1x]} \quad \cdots \quad \frac{\partial y}{\partial [{}^px]}\right]. \tag{2.22}$$

The above can be applied to $x \in \mathbf{R}^n$, from which one recovers the classical notion of partial derivative. In this particular case, it is convenient to define the *derivations* ∂_j by

$$\partial_j\, x \equiv |_j \qquad \text{(notation (1.29))}. \tag{2.23}$$



One thus can write $\partial_j y$ *instead of* $\frac{\partial y}{\partial [{}^jx]}$. With this notation one can make use of the properties of derivations. In the case where $x \mapsto X$ is a chart of a manifold, the associated basis S (1.28) is given by $S = \frac{\partial X}{\partial x}$. Hence the basis vectors are

$$S_j \equiv S(|_j) \equiv \frac{\partial X}{\partial x}(\partial_j\, x) \equiv \partial_j\, X\,. \tag{2.24}$$

If $x \mapsto y$ is a differentiable map *from a direct product to a direct product*, the preceding results allow us to write

$$\frac{\partial y}{\partial x} \equiv \begin{bmatrix} \frac{\partial\, {}^1y}{\partial\, {}^1x} & \cdots & \frac{\partial\, {}^1y}{\partial\, {}^px} \\ \cdots & \cdots & \cdots \\ \frac{\partial\, {}^{p'}y}{\partial\, {}^1x} & \cdots & \frac{\partial\, {}^{p'}y}{\partial\, {}^px} \end{bmatrix}. \tag{2.25}$$

For $x \in \mathbf{R}^p$ and $y \in \mathbf{R}^{p'}$ the entries of this matrix (2.25) are real numbers, which can be written as

$${}^j\left[\frac{\partial y}{\partial x}\right]_k \equiv {}^j|\cdot \frac{\partial y}{\partial x} \cdot |_k \equiv \frac{\partial\, {}^jy}{\partial\, {}^kx} \equiv \partial_k\, {}^jy\,.$$

Derivations of linear operators

Let E and E' be two finite dimensional vector spaces. Let x be a variable that ranges over E and A a variable that ranges over the vector space of all linear maps from E to E'. One can verify immediately that the map

$$\begin{pmatrix} A \\ x \end{pmatrix} \mapsto A(x)$$

is differentiable and that

$$\delta[A(x)] \equiv [\delta A](x) + A(\delta x)\,, \tag{2.26}$$

which reduces to $\delta[A(x)] \equiv A(\delta x)$ if A is constant.

Let us now assume that A is bilinear (and a variable). One finds

$$\begin{aligned} \delta[A(x)(y)] &\equiv \delta[A(x)](y) + A(x)(\delta y) \\ &\equiv [[\delta A](x) + A(\delta x)](y) + A(x)(\delta y) \\ &\equiv [\delta A](x)(y) + A(\delta x)(y) + A(x)(\delta y)\,, \end{aligned}$$

and more generally for a p-linear operator

$$\begin{aligned} \delta[A(x_1)(x_2)\dots(x_p)] &\equiv [\delta A](x_1)\dots(x_p) + \\ &+A(\delta x_1)(x_2)\dots(x_p) + \cdots + A(x_1)\dots(x_{p-1})(\delta x_p)\,. \end{aligned} \tag{2.27}$$

This formula has many applications, such as

$$\delta[A \cdot B] \equiv \delta A \cdot B + A \cdot \delta B\,, \tag{2.28}$$

where A and B are variable linear operators. If A is a variable injective linear map from a vector space to a vector space of the same dimension, one can deduce from (2.27) the formulæ

$$\frac{\delta[\det(A)]}{\det(A)} \equiv \mathrm{Tr}(A^{-1} \cdot \delta A) \tag{2.29}$$

and

$$\delta[A^{-1}] \equiv -A^{-1} \cdot \delta A \cdot A^{-1}\,. \tag{2.30}$$

The image of a vector field

(2.31) Let f be a vector field on a manifold V. The *graph* Γ of the field is by definition the set of pairs $\begin{pmatrix} f(x) \\ x \end{pmatrix}$, and thus Γ is a subset of the tangent bundle of V, denoted by V^D (see (1.31) and (1.34)). As this graph has at most one point in V^D over each point x of V, one says that Γ is a *section* of V^D.

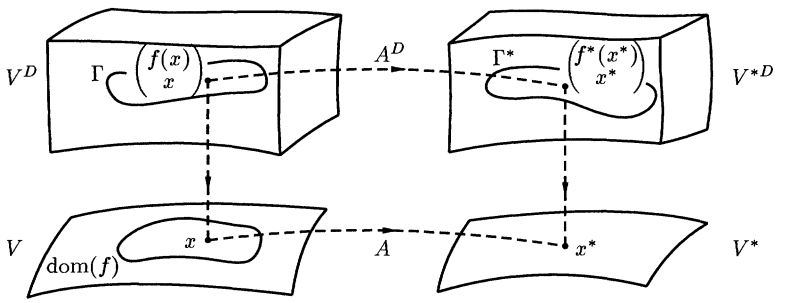


Fig 2.I

Let us now assume that A is a differentiable map from V to a manifold V^*, and let us denote by Γ^* the image of the graph Γ by the map A^D. If A is injective, one can see immediately that Γ^* is also a section, and thus is the graph of a field f^* on V^*; we will say that f^* is the *image*[†] *under A of the field f* and we shall write

(2.32) $$f^* = A_+(f)\,.^{\dagger}$$

Taking the image of f under another differentiable injective map B, one can verify immediately that

(2.33) $$[A \circ B]_+(f) = A_+(B_+(f))\,.$$

Replacing A by the inverse of a diffeomorphism B (from V^* to V),
(2.34) we define the *inverse image $B_-(f)$ of a field f by a diffeomorphism B* as being $[B^{-1}]_+(f)$. We have the following formulæ

(2.35) $$[A \circ B]_-(f) = B_-(A_-(f))$$

and

(2.36) $$B_+(B_-(f)) = f \qquad \text{if } \mathrm{dom}(f) \subseteq \mathrm{dom}(B^{-1}).$$

(2.37) To compute $A_+(f)$ in practice, it is important to note that the image of a vector field $x \mapsto \delta x$ under an injective map $x \mapsto x^*$ is the vector field $x^* \mapsto \delta x^*$. □

[†]Editors' note: Also called the *push forward* of the vector field f by the map A.
[†]Editors' note: Also denoted as $A_*(f)$.

This is simpler than the formulation directly in terms of vector fields, which is written as

$$A_+(f)(x^*) \equiv D(A)(A^{-1}(x^*))f(A^{-1}(x^*))\,. \tag{2.38}$$

If V is a vector space, every vector field f is a map from V to V, and thus we know how to determine whether f is continuous or differentiable. Unfortunately, this is no longer possible in the general case of a manifold because the tangent vector space D_x to which $f(x)$ belongs *varies with* x. One gets out of trouble by adopting the convention that

a vector field f is continuous (differentiable) if the map

$$x \mapsto \begin{pmatrix} f(x) \\ x \end{pmatrix} \tag{2.39}$$

is continuous (differentiable). □

Indeed, the above map takes V into V^D, which are manifolds (see (1.34)), and thus one can apply definitions (1.15) and (1.17).

(2.40) One can verify immediately that a *differentiable vector field* is defined on an open set and that its image (or its inverse image) under a diffeomorphism is differentiable.[65]

(2.41) To verify that a vector field is differentiable (continuous), it is sufficient to verify that its inverse images under the charts of an atlas are differentiable (continuous).

(2.42) Let f and f^* be two vector fields defined on manifolds of the same dimension V and V^*, respectively. One can ask whether these vector fields are *equivalent*, that is, whether there exists a diffeomorphism A such that $f^* = A_+(f)$. A similar problem is to know whether f and f^* are *locally equivalent* at a pair of points (x_0, x_0^*), that is, whether there exists a diffeomorphism A such that $A(x_0) = x_0^*$ and such that f^* coincides with $A_+(f)$ in an open set containing x_0^*.

Partial answers to these problems are given by the following theorems.

(2.43) THEOREM: Let f be a nonzero *constant* vector field on $\mathbf{R}^n$. In order that f be locally equivalent to a differentiable vector field f^* on a manifold of dimension n at a pair of points (x_0, x_0^*), it is necessary and sufficient that $f^*(x_0^*) \neq 0$.

[65] Its image under a differentiable map that is not a diffeomorphism need not be differentiable nor need it be defined on an open set.

(2.44) THEOREM: Let E be a finite dimensional vector space and let f be a differentiable vector field on E. In order that f be locally equivalent to the field $x \mapsto x$ on E at the pair $(0,0)$, it is necessary and sufficient that

$$\begin{cases} f(0) = 0\,, \\ D(f)(0) = 1_E\,. \end{cases}$$ □

Lie brackets

(2.45) THEOREM: Let $x \mapsto \delta x$ and $x \mapsto \delta' x$ be two differentiable vector fields on an open set E of a manifold V. Then there exists a derivation called the *Lie bracket* of δ and δ' denoted by $[\delta, \delta']$;[66] $[\delta, \delta']y$ is defined whenever the map $x \mapsto y$ is differentiable. If y takes values in a *fixed vector space*,[67] one has

$$\diamondsuit \qquad [\delta, \delta']y \equiv \delta[\delta' y] - \delta'[\delta y]\,.$$ □

If $F\,[\xi \mapsto x]$ is a chart of V, one deduces immediately from $\diamondsuit$ the formula

$$(2.46) \qquad [\delta, \delta']x = \frac{\partial x}{\partial \xi}(\delta\delta'\xi - \delta'\delta\xi) \qquad \text{if } x \in \mathrm{im}(F).$$

This allows us to *compute* the Lie bracket; it also shows that $x \mapsto [\delta, \delta']x$ is again a *differentiable vector field* on E. We will denote it by $[f, g]$ if f and g denote the vector fields $x \mapsto \delta x$ and $x \mapsto \delta' x$ respectively.

The differentiable vector fields on E obviously form a vector space. One can ascertain that the Lie bracket is bilinear, antisymmetric, and satisfies the *Jacobi identity*

$$(2.47) \qquad [f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0\,.$$

Let A be a diffeomorphism from E onto an open subset E^* of a manifold V^* and let us write $x^* \equiv A(x)$. The vector fields $x^* \mapsto \delta x^*$, $x^* \mapsto \delta' x^*$, and $x^* \mapsto [\delta, \delta']x^*$ are the images under A of the vector fields f, g, and $[f, g]$ respectively, whence

$$(2.48) \qquad [A_+(f), A_+(g)] = A_+([f, g])\,.$$

[66] We will sometimes write $[\delta, \delta']_L$.

[67] If y varies over a manifold then $\delta\delta' y$ and $\delta'\delta y$ *need not exist*; however, the left hand side of $\diamondsuit$ exists nonetheless.

(2.49) One says that two derivations δ and δ' *commute* if their Lie bracket is zero; one then has $\delta[\delta' y] \equiv \delta'[\delta y]$ for every variable y which takes its values in a fixed vector space. This is the case in particular for the derivations ∂_j defined in (2.23), taken pairwise, because

$$[\partial_j, \partial_k]\, x \equiv \partial_j\big|_k - \partial_k\big|_j \equiv 0\,.$$

§3. Differential equations

The exponential of a vector field

(3.1) THEOREM: Let V be a *Hausdorff* manifold.[68] Then there exists a map, called exp, which associates to every vector field f on V a map[69] $\exp(f)$ from V to itself such that, $\forall f$:

a) $\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \exp(tf)(x)$ is a differentiable map from $\mathbf{R} \times V$ to V;[70]

b) $[\delta x = 0] \Rightarrow \delta[\exp(tf)(x)] = f(\exp(tf)(x))\,\delta t$;[71]

c) $\exp(0f) = 1_{\mathrm{dom}(f)}$;

d) $\exp(-f) = [\exp(f)]^{-1}$;

e) $\exp(tf) \circ \exp(t'f) = \exp([t + t']f)$ if $tt' \geq 0$. □

From this it is easy to deduce the following result.

(3.2) THEOREM: Let f be a differentiable vector field on a Hausdorff manifold. Let x_0 be a point of $\mathrm{dom}(f)$ and t_0 a real number. If we define

$$\diamondsuit \qquad \Phi(t) \equiv \exp([t - t_0]f)(x_0)\,,$$

the map Φ has the following properties:

$$\heartsuit \qquad \begin{cases} \Phi \text{ is defined on an open interval;} \\ \Phi(t_0) = x_0; \\ x \equiv \Phi(t) \text{ satisfies the } \textit{differential equation } \dfrac{dx}{dt} \equiv f(x). \end{cases}$$

Conversely, each map Φ satisfying $\heartsuit$ satisfies $\diamondsuit$ for all $t \in \mathrm{dom}(\Phi)$. □

One can see that this theorem allows us to compute the map exp. For $\exp(f)(x)$ to exist, it is necessary and sufficient that there exists a solution Φ of the differential equation $\frac{dx}{dt} \equiv f(x)$, defined on an interval

---

[68] One says that a manifold is *Hausdorff* if for every pair (x, y) of distinct points of V there exists an open set containing x and an open set containing y whose intersection is empty. Examples: $\mathbf{R}^n$, a sphere. Later on we will encounter examples of non-Hausdorff manifolds (5.18).

[69] Possibly empty.

[70] One denotes by tf the product of the number t and the vector field f, defined by $[tf](x) = t[f(x)]$. In particular, $0f$ is the vector field which is identically zero on the domain of definition of f.

[71] One can write as well $\frac{\partial}{\partial t}[\exp(tf)(x)] \equiv f(\exp(tf)(x))$, where the left hand side is a *partial* derivative.

containing $[0,1]$, and such that $\Phi(0) = x$. This solution is *unique* and one has

$$\exp(f)(x) = \Phi(1)\,. \tag{3.3}$$

As a result, theorem (3.1) can be taken as a *definition* of exp.

(3.4) EXAMPLE: If E is a finite dimensional vector space, the identity map 1_E can be considered as a differentiable vector field on E; upon integrating the differential equation $\frac{dx}{dt} \equiv x$ one finds that

$$\exp(t\,1_E) = e^t\,1_E\,,$$

which explains the notation exp. □

(3.5) EXAMPLE: More generally, if A is a linear map from E to itself, one can verify that $\exp(A)$ is again a linear map from E to itself (see (6.17)). □

The image of a differential equation

Let A be a diffeomorphism from V to a manifold V^*. If x satisfies the differential equation $\frac{dx}{dt} \equiv f(x)$, it follows that $x^* \equiv A(x)$ satisfies $\frac{dx^*}{dt} \equiv f^*(x^*)$ with $f^* = A_+(f)$ (notation (2.32)). However, the map $t \mapsto x^*$ need not necessarily be defined on an interval unless $\mathrm{dom}(f) \subseteq \mathrm{dom}(A)$ (fig. 3.I). From this one deduces the formulæ

$$\begin{cases} A \circ \exp(f) \circ A^{-1} & \text{is an extension of } \exp(A_+(f)) \\ A \circ \exp(f) \circ A^{-1} = \exp(A_+(f)) & \text{if } \mathrm{dom}(f) \subseteq \mathrm{dom}(A). \end{cases} \tag{3.6}$$

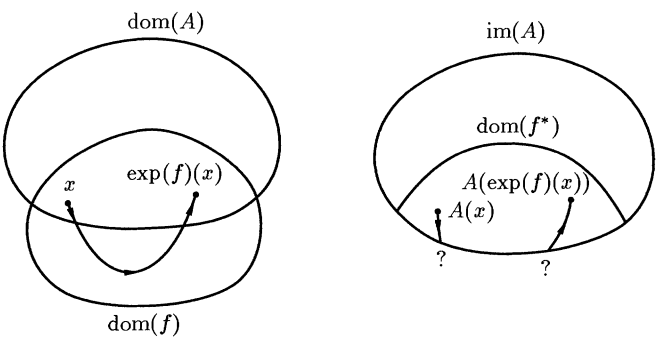


Fig 3.I

The derivative of the exponential map

Let f and g be two differentiable vector fields and let us assume that $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$. Let $\begin{pmatrix} t \\ y \end{pmatrix}$ be a variable ranging over the whole of $\mathbf{R} \times V$. Finally let us define $x \equiv \exp(-tf)(y)$, $\delta' \begin{pmatrix} t \\ y \end{pmatrix} \equiv \begin{pmatrix} 0 \\ g(y) \end{pmatrix}$, and $\delta \begin{pmatrix} t \\ y \end{pmatrix} \equiv \begin{pmatrix} 1 \\ f(y) \end{pmatrix}$.

The definition of the image of a vector field by a diffeomorphism shows that $\delta' x \equiv [\exp(-tf)]_+(g)(x)$. Applying the derivation δ to $y \equiv \exp(tf)(x)$ we find that $\delta x \equiv 0$.

On the other hand, we find that $[\delta, \delta'] \begin{pmatrix} t \\ y \end{pmatrix} \equiv \begin{pmatrix} 0 \\ [f,g](y) \end{pmatrix}$, whence

$$[\delta, \delta'] x \equiv [\exp(-tf)]_+([f,g])(x)\,.$$

In the case where V is a vector space, we have $[\delta, \delta'] x \equiv \delta[\delta' x] - \delta'[\delta x] \equiv \delta\, \delta' x$; evaluation yields the formula [72]

$$\frac{\partial}{\partial t} \left\{ [\exp(-tf)]_+(g)(x) \right\} \equiv [\exp(-tf)]_+([f,g])(x)\,. \tag{3.7}$$

In the case of an arbitrary manifold V, note that $[\exp(-tf)]_+(g)(x)$ varies over the *constant* vector space D_x, provided we vary only t and keep x fixed. It follows that one can compute the derivative written on the left hand side of (3.7); one can show that the formula *remains valid.*[73]

Let us now assume that the Lie bracket $[f,g]$ is *identically zero.* One can check that $\exp(-tf)_+(g)(x)$ is defined if $\exp(tf)(x)$ exists and thus in particular when t varies over an interval containing 0 (see (3.2)). Since its derivative with respect to t is zero, it equals its value at $t = 0$, whence the formula

$$[\exp(-tf)]_+(g)(x) = g(x)\,, \tag{3.8}$$

provided that the left hand side exists, $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$, and $[f,g] = 0$.

Let us note that this formula can also be written as

$$g(\exp(tf))(x) = D(\exp(tf))(x)(g(x)) \qquad ([f,g] = 0)\,, \tag{3.9}$$

which allows us to compute $D(\exp(tf))(x)$.

---

[72] $\frac{\partial}{\partial t}$ denotes here the derivative with respect to t, x being held constant.

[73] After choosing a chart for V in a neighborhood of x and applying the formulæ (2.48) and (3.6).

(3.10) THEOREM:[74] Let $f_1, \ldots, f_p$ be differentiable vector fields on an open subset E of a manifold V *for which the Lie brackets taken pairwise are zero*, and let x be a point of E. Then there exist positive numbers $a_1, \ldots, a_p$ such that

$$\exp({}^1t\, f_1 + \cdots + {}^pt\, f_p)(x)$$

exists if

$$\diamondsuit \qquad |{}^1t| < a_1, \ldots, |{}^pt| < a_p .$$

Under these conditions one has

$$\heartsuit \qquad \exp({}^1t\, f_1 + \cdots + {}^pt\, f_p)(x) = [\exp({}^1t\, f_1) \circ \cdots \circ \exp({}^pt\, f_p)](x) . \qquad \square$$

It is immediate that one can arbitrarily permute the factors $\exp({}^1t\, f_1)$, $\ldots$, $\exp({}^pt\, f_p)$ on the right hand side without changing the result.[75]

[74] This result can be established by using (3.8) and (3.6).

[75] Nevertheless there are cases where these maps *do not commute*; one has to pay attention to the conditions of the theorem.

§4. Differential forms

Covariant fields

(4.1) Let E be a vector space; a *covariant operator of degree p on E* is a map φ such that[76]

$$\varphi(y_1)(y_2)\dots(y_p) \in \mathbf{R} \qquad [\forall y_1, y_2, \dots y_p \in E]\,.$$

In the case $p = 1$, φ is simply a map from E to $\mathbf{R}$; by convention covariant operators *of degree 0 are real numbers*. □

(4.2) If φ is a covariant operator of degree p ($p \geq 1$) and $y \in E$, it is clear (due to the convention on the implicit brackets) that $\varphi(y)$ is a covariant operator of degree $p-1$; it is called the *contraction* of φ and y. □

(4.3) Let x be a variable which ranges over a manifold V; we will say that a map $x \mapsto \varphi$ defined on a subset of V is a *covariant field* (of degree p) if φ is a *covariant operator* of degree p on the *tangent space* D_x.[77] □

(4.4) It is clear that a covariant field of degree 0, also called a *scalar field*, is a map from V to $\mathbf{R}$.

The inverse image of a covariant field

Let f^* be a covariant field of degree p on a manifold V^* and let A be a differentiable map from a manifold V to V^*. Let us define

$$A_-(f^*)(x)(y_1)\dots(y_p) = f^*(A(x))\,(D(A)(x)(y_1))\dots(D(A)(x)(y_p))\,, \qquad \forall x \in V\,, \qquad \forall y_1, \dots y_p \in D_x\,. \tag{4.5}$$

It is clear that $A_-(f^*)$ is a covariant field of degree p on V; it is called the *inverse image*† of the field f^* by A.

---

[76] In this notation there are implicit brackets $[\dots[[\varphi(y_1)](y_2)]\dots](y_p)$ starting at the left.

[77] One also says that φ is a *differential form*; we shall reserve this name for the particular case (4.26).

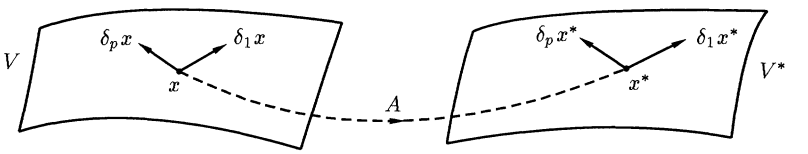
† Editors' note: Or *pull back*.

If we introduce variables x^*, φ^*, φ, and derivations δ_j such that $x^* \equiv A(x)$, $\varphi^* \equiv f^*(x^*)$, $\varphi \equiv A_-(f^*)(x)$, and $\delta_j x \equiv y_j$, we see that the inverse image of the covariant field $x^* \mapsto \varphi^*$ by the map $x \mapsto x^*$ is the covariant field $x \mapsto \varphi$ defined by

$$\varphi(\delta_1 x)\dots(\delta_p x) \equiv \varphi^*(\delta_1 x^*)\dots(\delta_p x^*)\,. \tag{4.6}$$

Formula (4.6), which is a simple transcription of (4.5), is easier to remember.

Fig 4.I



For $p = 0$ these formulæ obviously reduce to

$$A_-(f) = f \circ A\,, \qquad \varphi^* \equiv \varphi\,. \tag{4.7}$$

Composing two differentiable maps A and B, one immediately finds the formula (the analog of (2.35))

$$[A \circ B]_-(f) = B_-(A_-(f))\,, \tag{4.8}$$

valid this time for a covariant field f.

Definition (4.5) or (4.6) of the inverse image of f by A does *not* assume that A is a diffeomorphism or even that the manifolds V and V^* have the same dimension.

If A is a *diffeomorphism* we can define

$$A_+(f) = [A^{-1}]_-(f)\,; \tag{4.9}$$

$A_+(f)$ is called *the direct image of a covariant field*† of the covariant field f by A. Evidently we have

$$[A \circ B]_+(f) = A_+(B_+(f))\,, \tag{4.10}$$

provided, of course, that A and B are diffeomorphisms; we also have

$$B_+(B_-(f)) = f\,, \tag{4.11}$$

†Editors' note: or push forward

provided that $\mathrm{dom}(f) \subseteq \mathrm{dom}(B^{-1})$ (see (2.33) to (2.36)).

(4.12) If $x \mapsto \delta x$ is a vector field, its direct image by a diffeomorphism $x \mapsto x^*$ is the field $x^* \mapsto \delta x^*$ (2.37). If $x^* \mapsto \varphi^*$ is a covariant field of degree p ($p \geq 1$), formula (4.6) shows that *the inverse image of the contracted field* $x^* \mapsto \varphi^*(\delta x^*)$ *is the contracted field* $x \mapsto \varphi(\delta x)$.

(4.13) A field g of covariant operators of degree p defined on a vector space E will be called *differentiable* if the map (from E^{p+1} to $\mathbf{R}$)

$$\begin{pmatrix} x \\ y_1 \\ \vdots \\ y_p \end{pmatrix} \mapsto g(x)(y_1)\dots(y_p)$$

is differentiable. If g is defined on a manifold V, g will be called differentiable if its inverse images by the charts of V are differentiable. In such a case g should be defined on an open subset of V.

(4.14) One can verify that the inverse image of a differentiable field by a differentiable map is again a differentiable field.

The Lie derivative

(4.15) THEOREM: Let f and g be differentiable fields of vectors and covariant operators of degree p, respectively, defined on an open set E. Then there exists a field of covariant operators of degree p defined on E called the *Lie derivative* of the field g by (in the direction of) the field f and denoted by $[f, g]$, such that

$$\frac{\partial}{\partial t}[\exp(tf)_(g)(x)(\delta_1 x)\dots(\delta_p x)] \equiv \exp(tf)_([f,g])(x)(\delta_1 x)\dots(\delta_p x)\,.$$

□

As in (3.7), the notation $\frac{\partial}{\partial t}$ on the left hand side means that the derivative is taken with respect to t while the variables x, $\delta_1 x$, ..., $\delta_p x$ are held constant. The notation $[f, g]$, reminiscent of the Lie bracket, is justified by the analogy between (3.7) and (4.15). We will use concurrently another notation:

(4.16) if $f(x) \equiv \delta x$ and $g(x) \equiv \varphi$, we will write

$$\delta_L \varphi \equiv [f, g](x)\,.$$

For a single field $x \mapsto \varphi$ we are thus able to define several Lie derivatives $\delta_L\varphi$, $\delta'_L\varphi$, etc., where $x \mapsto \delta x$, $x \mapsto \delta' x$ etc., are differentiable vector fields.

If we apply the notation of (4.16) to the vector field $\delta' x \equiv \varphi$, we are lead to write

$$\delta_L \delta' x \equiv [\delta, \delta']\, x\,, \tag{4.17}$$

and thus the Lie bracket (2.45) can be interpreted as a Lie derivative.

In case $p = 0$, that is, if $x \mapsto \varphi$ is a scalar field, one finds immediately that

$$\delta_L \varphi = \delta\varphi \tag{4.18}$$

for every derivation δ for which the left hand side exists.

Covariant tensor fields

(4.19) A covariant operator φ of degree p defined on a vector space E is called a *covariant tensor* if $\varphi(y_1)\dots(y_p)$ depends in a linear way on each of the vectors $y_1, \dots, y_p$ (one also says that φ is a *multilinear* operator).

If S is a basis of E, the numbers

$$\varphi_{jk\dots m} = \varphi(S_j)(S_k)\dots(S_m) \tag{4.20}$$

are called the *components of the tensor φ with respect to the basis S*. The components of a tensor define it completely; indeed, multilinearity implies that

$$\varphi(y)(z)\dots(u) = \varphi_{jk\dots m}\, {}^j y\, {}^k z \cdots {}^m u\,, \tag{4.21}$$

where the numbers ${}^j y, {}^k z, \dots, {}^m u$ are the components of the vectors $y, z, \dots, u$ with respect to the basis S ($y = \sum_j S_j\, {}^j y, \dots, u = \sum_m S_m\, {}^m u$). In (4.21) we have used the *Einstein summation convention*, according to which the summation $\Sigma_{jk\dots m}$ is understood on the right hand side because the indices $j, k, \dots, m$ are each written twice (once as a subscript and once as a superscript).

If $x \mapsto X$ is a chart of a manifold V and $X \mapsto \varphi$ is a *field of covariant tensors* (of degree p), one usually denotes by $\varphi_{jk\dots m}$ the components of the tensor φ with respect to the basis $S \equiv \frac{\partial X}{\partial x}$. As in (2.23) we write $\partial_j x \equiv \big|_j$, whence $\partial_j X \equiv S_j$. Thus we have

$$\varphi_{jk\dots m} \equiv \varphi(\partial_j X)(\partial_k X)\dots(\partial_m X)\,. \tag{4.22}$$

(4.23) Let $x \mapsto \varphi$ be a *tensor field* of degree p ($p \geq 1$) and $x \mapsto \delta x$ a vector field. It is clear that by contraction one obtains a *tensor field of degree* $p-1$ given by $x \mapsto \varphi(\delta x)$. If $x \mapsto \varphi$ and $x \mapsto \delta x$ are differentiable, then the contracted field is also differentiable. One can establish the important formula

$$\delta'_L[\varphi(\delta x)] \equiv [\delta'_L\varphi](\delta x) + \varphi(\delta'_L \delta x)\,, \tag{4.24}$$

where $x \mapsto \delta' x$ denotes another differentiable vector field. By iteration one can deduce the formula

$$\begin{aligned} \delta'[\varphi(\delta_1 x)\dots(\delta_p x)] \;\equiv\; & [\delta'_L\varphi](\delta_1 x)\dots(\delta_p x) \\ & + \varphi(\delta'_L\delta_1 x)\dots(\delta_p x) \\ & + \cdots \\ & + \varphi(\delta_1 x)\dots(\delta'_L\delta_p x)\,, \end{aligned} \tag{4.25}$$

which is easy to remember because of its analogy with (2.27). On the left hand side we were able to get rid of the subscript L because of (4.18). This formula can be used to compute the Lie derivative $\delta'_L\varphi$ by isolating it from the right hand side of the equation.

p-Forms

(4.26) Let us now assume that a covariant tensor φ of order p is *antisymmetric*, that is, $\varphi(\delta_1 x)(\delta_2 x)\dots(\delta_p x)$ changes sign each time one interchanges two of the vectors $\delta_1 x, \delta_2 x, \dots, \delta_p x$.[78]



In such a case we will say that φ is an *exterior form of degree* p — in short a *p-form*.

(4.27) One can recognize that a tensor φ is antisymmetric by the fact that its components $\varphi_{jk\dots m}$ change sign when one interchanges the values of two of its indices $j, k, \dots, m$; hence a component is zero whenever two of these indices are equal.

(4.28) A 1-form will be any covariant tensor of degree 1 (also called a *covector*) and a 0-form will be any *scalar* (see (4.4)). Indeed for $p = 0$ or 1, the antisymmetry condition loses its meaning.

(4.29) If φ is a p-form ($p \geq 1$) and δx a vector, it is obvious that the contracted tensor $\varphi(\delta x)$ is a $[p-1]$-form, also called the *interior product* of φ and δx.

[78] Because of multilinearity, it is the same as assuming that $\varphi(\delta_1 x)\dots(\delta_p x)$ is zero as soon as two of these vectors are equal.

(4.30) One can show that the p-forms on a vector space E of dimension n form a vector space whose dimension is the binomial coefficient $\binom{n}{p} = \dfrac{n!}{p!(n-p)!}$ and that there is no nonzero p-form for $p > n$. The forms of degree n, also called forms of maximum degree, form a vector space of dimension 1. Hence, by choosing a nonzero one among them, which we denote by vol, all other forms of degree n are proportional to it.

(4.31) One can show that vol is injective and yields all the $[n-1]$-forms by contraction: if θ is an $[n-1]$-form, there exists a vector δx such that $\theta = \mathrm{vol}(\delta x)$; one might write $\delta x = \mathrm{vol}^{-1}(\theta)$.

The exterior derivative

(4.32) THEOREM: Let $x \mapsto \varphi$ be a differentiable field of p-forms on an open set E. Then there exists a differentiable field of $[p+1]$-forms on E, denoted by $x \mapsto \mathrm{d}\varphi$, such that

$$\begin{aligned}[\mathrm{d}\varphi](\delta x)(\delta_1 x)\dots(\delta_p x) \;\equiv\; & \delta\,[\varphi(\delta_1 x)(\delta_2 x)\dots(\delta_p x)] \\ & - \delta_1\,[\varphi(\delta x)(\delta_2 x)\dots(\delta_p x)] \\ & - \delta_2\,[\varphi(\delta_1 x)(\delta x)\dots(\delta_p x)] \\ & - \cdots \\ & - \delta_p\,[\varphi(\delta_1 x)\dots(\delta_{p-1} x)(\delta x)]\,,\end{aligned}$$

provided the derivations $\delta, \delta_1, \dots, \delta_p$ commute pairwise (2.49). The $[p+1]$-form $\mathrm{d}\varphi$ is called the *exterior derivative* of φ. □

For $p = 0$ (that is, the form φ is a scalar) the above definition can simply be written as

(4.33) $$[\mathrm{d}\varphi](\delta x) \equiv \delta\varphi \qquad \text{or} \qquad \mathrm{d}\varphi \equiv \frac{\partial\varphi}{\partial x}\,.$$

Using a chart and the associated derivations ∂_j, which commute pairwise (see (4.22)), formula (4.32) gives

(4.34) $$[\mathrm{d}\varphi]_{jk\ell\dots m} = \partial_j\varphi_{k\ell\dots m} - \partial_k\varphi_{j\ell\dots m} - \partial_\ell\varphi_{kj\dots m} - \cdots - \partial_m\varphi_{k\ell\dots j}\,,$$

which yields the *components* of $\mathrm{d}\varphi$. From this it follows that theorem (4.32) is a *definition* of $\mathrm{d}\varphi$, since it permits us to compute the exterior derivative of φ (see (4.21)).

For $p = 0, 1, 2$ formula (4.34) reads

$$
\begin{aligned}
[\mathrm{d}\varphi]_j &= \partial_j \varphi \\
[\mathrm{d}\varphi]_{jk} &= \partial_j \varphi_k - \partial_k \varphi_j \\
[\mathrm{d}\varphi]_{jk\ell} &= \partial_j \varphi_{k\ell} + \partial_k \varphi_{\ell j} + \partial_\ell \varphi_{jk} \,.
\end{aligned} \tag{4.35}
$$

(4.36) THEOREM: Let $x \mapsto \varphi$ be the *inverse image*[†] of a differentiable field of p-forms $x^* \mapsto \varphi^*$ by a differentiable map $A\,[x \mapsto x^*]$. Then $x \mapsto \varphi$ is a differentiable field of p-forms and the field $x \mapsto \mathrm{d}\varphi$ is the inverse image under A of $x^* \mapsto \mathrm{d}\varphi^*$.[79] □

In short, one says that the operations of "exterior derivative" and "inverse image" *commute*.

(4.37) POINCARÉ'S LEMMA: If $x \mapsto \varphi$ is a differentiable field of p-forms, then $\mathrm{d}\,[\mathrm{d}\varphi] = 0$. Conversely, if $x \mapsto \theta$ is a differentiable field of $[p+1]$-forms defined on a cell[80] and if $\mathrm{d}\theta \equiv 0$, then there exists a differentiable field $x \mapsto \varphi$ of p-forms on this cell such that $\theta \equiv \mathrm{d}\varphi$. □

One can add to Poincaré's lemma the following easily verified result.

(4.38) If a scalar field $x \mapsto \varphi$ is defined on a *connected* open set and if $\mathrm{d}\varphi \equiv 0$, then there exists a constant a such that $\varphi \equiv a$. □

(4.39) CARTAN'S THEOREM: If $x \mapsto \varphi$ is a differentiable field of p-forms ($p \geq 1$) and if $x \mapsto \delta x$ is a differentiable vector field, then[81]

$$\delta_L \varphi \equiv [\mathrm{d}\varphi]\,(\delta x) + \mathrm{d}\,[\varphi(\delta x)] \,. \qquad \square$$

Combining Poincaré's lemma and Cartan's theorem, one can verify immediately that

$$\delta_L\,[\mathrm{d}\varphi] = \mathrm{d}\,[\delta_L \varphi] \qquad (\varphi \text{ a } p\text{-form}, \; p \geq 0),$$

which can also be deduced from theorem (4.36) and the definition of $\delta_L \varphi$.

[†]Editors' note: also called the *pull back*.

[79]Of course, $\mathrm{d}\varphi^*$ denotes the exterior derivative of φ^* with respect to x^*, defined by putting a star * on all the letters x and φ in formula (4.32).

[80]For $p = 0$ one can replace this condition by the condition that $\mathrm{dom}(\theta)$ is *simply connected*. It is even sufficient that the *homotopy group* of $\mathrm{dom}(\theta)$ be *equal to its derived group*, that is, the homotopy group has no abelian quotient.

[81]For $p = 0$, formulæ (4.18) and (4.33) show that it is necessary to suppress the last term, which does not make sense.

§5. Foliated manifolds

Foliations

(5.1) Let V be a manifold of dimension n and x a variable ranging over V. Let $x \mapsto E$ be a map such that E is a vector subspace of *constant dimension* m of the tangent space D_x (in short, we will say that $x \mapsto E$ is a *field of vector spaces of dimension* m).

(5.2) We will say that a *field of vector spaces* is *differentiable* if one can cover its domain of definition by open sets on which $E \equiv \mathrm{im}(T)$, where $x \mapsto T$ is a differentiable field of bases of E.[82]

One can then prove the following theorems.

(5.3) THEOREM: If $x \mapsto \varphi$ is a differentiable tensor field of degree p and if the kernel $\ker(\varphi)$ [83] has constant dimension, then the field $x \mapsto \ker(\varphi)$ is differentiable. □

(5.4) THEOREM: If $x \mapsto E_1$ and $x \mapsto E_2$ are differentiable fields of vector spaces and if the dimension of $E_1 \cap E_2$ is constant, then the fields $x \mapsto E_1 \cap E_2$ and $x \mapsto E_1 + E_2$ are differentiable.[84] □

(5.5) DEFINITION: Let $x \mapsto E$ be a differentiable field of vector spaces defined on a manifold V. An *integral manifold* of this field is a manifold Φ such that:

$\diamondsuit$ $\begin{cases} \text{a) } \Phi \text{ is submanifold of } V.^{85} \\ \text{b) At each point } x \text{ of } \Phi \text{ the tangent space to } \Phi \text{ is } E. \end{cases}$

The field $x \mapsto E$ is called a *foliation* [86] of V if each point of V belongs to an integral manifold.

---

[82] One can write $T \equiv [T_1 T_2 \ldots T_m]$; the field $x \mapsto T$ is said to be differentiable if the vector fields $x \mapsto T_1, \ldots, x \mapsto T_m$ are differentiable.

[83] The *kernel* of φ is by definition the vector space of those δx such that $\varphi(\delta x) = 0$ (notation (4.29)).

[84] The *sum* $E_1 + E_2$ is a vector space of dimension $\dim(E_1) + \dim(E_2) - \dim(E_1 \cap E_2)$, and thus also of constant dimension.

[85] Definition (1.37).

[86] One also says that it gives V the structure of a *foliated manifold*.

(5.5) THEOREM (FROBENIUS): In order that the field $x \mapsto E$ be a foliation, the following condition[87] is necessary and sufficient:

$\heartsuit$ { If two vector fields $x \mapsto \delta x$ and $x \mapsto \delta' x$ satisfy $[\delta x \in E,\ \delta' x \in E]$ on an open set, then one also has $[\delta, \delta']x \in E$ on this open set.

PROOF: Let us indicate the main ideas of a proof of this theorem, which at the same time give a practical method of integration.

Let x_0 be a point of V. One can find a differentiable map $x \mapsto y$ from V to $\mathbf{R}^m$ (m being the dimension of E) such that $\ker(\frac{\partial y}{\partial x})$ is a complement to E at x_0. Since this property remains true in an open set Ω containing x_0, it follows that on Ω we have

$$\delta x \in E \quad \Longleftrightarrow \quad \delta x = S(\delta y)\,,$$

for some (uniquely determined) basis $S = [S_1 \dots S_m]$ of E; the resulting field $x \mapsto S$ is differentiable on Ω. On an integral manifold the map $x \mapsto y$ is, locally, a diffeomorphism whose inverse is a chart $F[y \mapsto x]$ satisfying $D(F)(y) \equiv S$. If $x \mapsto E$ is a foliation, one deduces that the Lie brackets of the vector fields $f_j[x \mapsto S_j]$ are pairwise 0. It is easy to verify the equivalence of this condition and $\heartsuit$ (on Ω).

Conversely, if condition ($\heartsuit$) is satisfied, choose a point $x_1 \in \Omega$ and write

$$x \equiv F(t) \equiv \exp({}^1t\, f_1 + \dots + {}^mt\, f_m)(x_1) \qquad [t \in \mathbf{R}^m]\,. \tag{5.6}$$

Obviously one has $F(0) = x_1$, and, according to theorem (3.10), $D(F)(t) \equiv S$. Shrinking its domain of definition if necessary, one can see that F is a chart of an integral manifold passing through x_1; thus V is foliated by integral manifolds. □

The study of integral manifolds is thus reduced to the computation of the exponential map figuring in (5.6), and thus to solving ordinary differential equations (see (3.3)).

One can also look for a chart $\varphi[z \mapsto x]$ of the manifold defined by the equation $y =$ constant and passing through x_0. This is a manifold of dimension $n - m$ (theorem (1.40)). Replacing x_1 by $\varphi(z)$ in (5.6), one can define a map $\begin{pmatrix} t \\ z \end{pmatrix} \mapsto x$; suitably shrinking its domain of definition, one can show that this map becomes a chart of V. From this one can deduce the following theorem.

[87] Called an *integrability condition*.

(5.7) THEOREM: Let V be a manifold of dimension n and $x \mapsto E$ a foliation of V ($\dim(E) \equiv m$). Then V possesses an atlas of charts

$$\begin{pmatrix} t \\ z \end{pmatrix} \mapsto x \qquad t \in \mathbf{R}^m,\ z \in \mathbf{R}^{n-m}$$

such that

♡ $$\delta x \in E \iff \delta z = 0\,,$$

and such that, for constant z, the map $t \mapsto x$ is a chart of an integral manifold. □

From this theorem it follows easily that the injective immersions $y \mapsto x$ of $\mathbf{R}^m$ into V such that $\operatorname{im}\left(\frac{\partial x}{\partial y}\right) \equiv E$ are compatible (1.5), and that they form an atlas of V. Hence V possesses *another structure of a manifold*, now of dimension m, *injectively immersed* in the old structure.



In this new structure the connected open sets are integral manifolds of E; the connected components (1.49) are called the *leaves* of E. It is easy to deduce from this the following theorem.

(5.8) THEOREM: Let $x \mapsto E$ be a foliation of a manifold V. Then

a) V possesses a partition into connected integral manifolds called *leaves*.
b) Each connected integral manifold is an *open subset of a leaf*.
c) If Φ is a connected manifold injectively immersed in V and if at every point of Φ the tangent space to Φ is contained in E, then Φ is *injectively immersed in a leaf of the foliation*.

(5.9) EXAMPLE: Let $x \mapsto E$ be a differentiable field of vector spaces *of dimension* 1 defined on a manifold V. It is easy to verify that the integrability condition is always satisfied, and thus V is foliated. The integral manifolds, which are of dimension 1, are called *lines of force*. □

(5.10) EXAMPLE: In particular, let f be a differentiable vector field on a manifold V. On the manifold $V \times \mathbf{R}$ one can define a differentiable field of vector spaces of dimension 1 by

$$\delta \begin{pmatrix} x \\ t \end{pmatrix} \in E \iff \delta x - f(x)\,\delta t = 0 \qquad \text{(Fig. 5.I)}\,.$$

It is clear that the *graph* of a solution of the differential equation $\frac{dx}{dt} \equiv f(x)$ is a line of force, but not every line of force is necessarily a graph. One can consider a line of force as a generalized solution of the differential equation. Note that leaves, which are maximal generalized solutions of the equation, exist *even if V is not Hausdorff* (compare with (3.2)). □

(5.11) EXAMPLE: Let $x \mapsto \mathbf{V}$ be a nonzero vector field on $\mathbf{R}^3$. An *orthogonal surface* for the vector field is a surface whose tangent space at each point x is orthogonal to $\mathbf{V}$. In order that through each point of $\mathbf{R}^3$ there exists an orthogonal surface for V, it is necessary that the field $x \mapsto \mathrm{orth}(\mathbf{V})$ be a foliation. Let us apply condition (5.5♡), which in this case becomes

$$\diamondsuit \qquad \begin{cases} \langle \mathbf{V}, \delta x \rangle \equiv 0 \\ \langle \mathbf{V}, \delta' x \rangle \equiv 0 \end{cases} \implies \langle \mathbf{V}, [\delta, \delta']x \rangle \equiv 0 .$$

Using (2.45♢) we obtain the identity

$$\langle \mathbf{V}, [\delta, \delta']x \rangle \equiv \langle \mathrm{curl}\, \mathbf{V}, \delta x \times \delta' x \rangle + \delta \langle \mathbf{V}, \delta' x \rangle - \delta'(\langle \mathbf{V}, \delta x \rangle) .$$

From this the reader can easily deduce that condition ♢ for the existence of an orthogonal surface for V can be written as $\langle \mathbf{V}, \mathrm{curl}\, \mathbf{V} \rangle \equiv 0$. □

(5.12) EXAMPLE: Let A be a differentiable map from a manifold V to a manifold V'. Define $E \equiv \ker(D(A))(x)$. If $m \equiv \dim(E)$ is constant, the field $x \mapsto E$ is differentiable and the integrability condition (5.5♡) is satisfied. Thus V is a foliated manifold. One can see immediately that the sets $A(x) =$ constant have a *unique structure* of a submanifold of V of dimension m, and that the connected components of each of these submanifolds are leaves.

This result completes, with different hypotheses, theorem (1.40) on manifolds defined by an equation. □

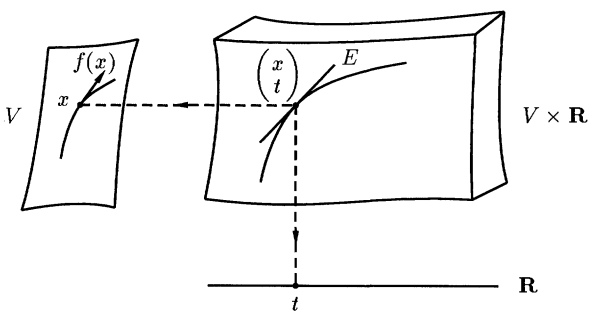


Fig 5.I

(5.13) THEOREM: Let V be a foliated manifold with f its foliation and let A be a diffeomorphism from V to V. We will say that A *preserves* the foliation if

$$[x \in \mathrm{dom}(A),\ \delta x \in f(x)] \implies [\delta[A(x)] \in f(A(x))] .$$

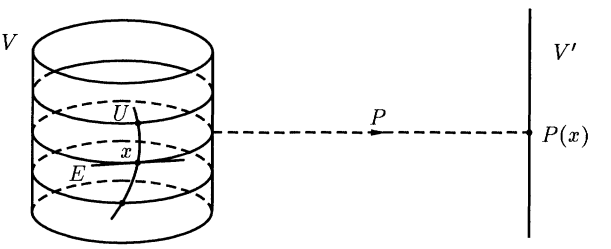
In order for A to be also a diffeomorphism of V *equipped with the structure of a manifold of dimension* m (notation of (5.7) and (5.8)), it is necessary and sufficient that A preserves the foliation.

If $\mathrm{dom}(A) = \mathrm{im}(A) = V$ and if A preserves the foliation, then the restriction of A to an arbitrary leaf Φ is a diffeomorphism from Φ onto a leaf Φ'. □

The quotient of a manifold by a foliation

(5.14) Let $x \mapsto E$ be a foliation of a manifold V of dimension n ($\dim(E) \equiv m$). We will say that a manifold U injectively immersed in V is *transverse* to the foliation if at each point of U its tangent space is complementary to E (Fig. 5.II); its dimension is necessarily $n - m$. We have seen (in (5.7)) that one can construct a manifold transverse to a foliation at each point of V.

Fig 5.II



(5.15) We will say that the foliation is *sectionable* if through each point of V there exists a transverse manifold U that is a *section* of the foliation, that is, U meets each leaf of the foliation in at most one point. In such a case we will say that U is a *transverse section*.

(5.16) Let us assume that the foliation is sectionable and let us denote by $P(x)$ the leaf that passes through the point x of V. One can show that the $P \circ F$, where F is a chart of a transverse section, form an atlas for the *set* V' of all leaves. Hence V' has the structure of a manifold of dimension $n - m$, called the *quotient manifold of V by the foliation*. It is clear that P is an open differentiable map from V onto V' and that $E \equiv \ker(D(P)(x))$; P is called the *projection* of V onto V'.

(5.17) Note that the sectionability condition for a foliation is a purely global condition: for each foliation of V one can cover V by open sets V_j on which the field $x \mapsto E$ induces a sectionable foliation (for example one can take the V_j to be the images of the charts (5.7)).

(5.18) Even if the manifold V is Hausdorff, it can happen that the quotient

manifold *is not*. For example, if $\begin{pmatrix} y \\ z \end{pmatrix}$ varies over $\mathbf{R}^2$ without the origin (Fig. 5.III) and if the foliation is defined by $\delta z = 0$, then the leaves are the lines $z = z_0$ ($z_0 \neq 0$) and the two half lines $D_+(z = 0,\ y > 0)$ and $D_-(z = 0,\ y < 0)$. It is easy to see that in the quotient manifold every open set containing D_+ has a nonempty intersection with every open set containing D_-.

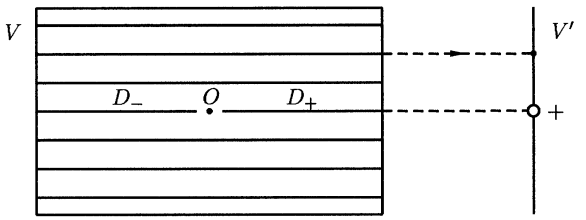


Fig 5.III

(5.19) Let V be a manifold with a sectionable foliation and let A be a diffeomorphism of V onto V which preserves the foliation. Theorem (5.13) shows that there is a permutation $\hat{A}$ of V' defined by

$$\hat{A}(P(x)) = P(A(x)) \qquad \forall x \in V\,;$$

one can show that $\hat{A}$ is a *diffeomorphism* of V'.

Integral invariants [88]

(5.20) Let $x \mapsto E$ be a sectionable foliation of a manifold V and let $x \mapsto \varphi$ be a differentiable field of p-forms on V. One says that φ is an *integral invariant* of the foliation if $x \mapsto \varphi$ is the inverse image under the projection onto V' of a differentiable field $x' \mapsto \varphi'$ of the quotient manifold V'. □

(5.21) If $x \mapsto \varphi$ is an integral invariant of degree p ($p \geq 1$), it is clear that $x \mapsto \mathrm{d}\varphi$ is also an integral invariant (because it is the inverse image of $x' \mapsto \mathrm{d}\varphi'$ (see (4.36))). It is also clear that $E \subseteq \ker(\varphi)$ (see definition (4.6) of the inverse image) and thus

$$\diamondsuit \qquad E \subseteq \ker(\varphi) \cap \ker(\mathrm{d}\varphi)\,.$$

Conversely one can show that condition ◇ is sufficient in order that φ be an integral invariant.

[88] Integral invariants *as defined by Élie Cartan.*

(5.22) An integral invariant φ *of degree* 0 is called a *first integral*; it is a scalar field which is *constant on the leaves*. Condition $\diamondsuit$, which does not make sense any more because $\ker(\varphi)$ is not defined, has to be replaced by

$$\heartsuit \qquad E \subseteq \ker(\mathrm{d}\varphi)\,.$$

(5.23) Note that formulæ $\diamondsuit$ or $\heartsuit$ remain meaningful if the foliation is not sectionable; in such a case it is said that $x \mapsto \varphi$ is an integral invariant for each open set V_j where the foliation is sectionable (see (5.17)).

The characteristic foliation of a form

Let $x \mapsto \varphi$ be a differentiable field of p-forms defined on a manifold V. Let us define

(5.24)
$$E' \equiv \ker(\varphi) \cap \ker(\mathrm{d}\varphi)\,,$$

and let us assume that *the dimension of* E' is *constant* and ≥ 1. Theorems (5.3) and (5.4) show that the field $x \mapsto E'$ is differentiable, so let us study the integrability condition (5.5$\heartsuit$).

If two differentiable fields of vectors $x \mapsto \delta x$ and $x \mapsto \delta' x$ are contained in E', that is,

$$\begin{cases} \varphi(\delta x) \equiv 0\,, & [\mathrm{d}\varphi](\delta x) \equiv 0\,, \\ \varphi(\delta' x) \equiv 0\,, & [\mathrm{d}\varphi](\delta' x) \equiv 0\,, \end{cases}$$

then one finds, using formula (4.24) and Cartan's theorem (4.39), that

$$\begin{aligned} \varphi([\delta', \delta]x) &\equiv \varphi(\delta'_L \delta x) \equiv \delta'_L[\varphi(\delta x)] - [\delta'_L \varphi](\delta x) \\ &\equiv -[\delta'_L \varphi](\delta x) \equiv -[\mathrm{d}\varphi(\delta' x) + \mathrm{d}[\varphi(\delta' x)]](\delta x) \equiv 0\,, \end{aligned}$$

and, due to Poincaré's lemma (4.37), that

$$\begin{aligned} [\mathrm{d}\varphi]([\delta', \delta]x) &\equiv \delta'_L[\mathrm{d}\varphi(\delta x)] - [\delta'_L \mathrm{d}\varphi](\delta x) \\ &\equiv -[\mathrm{d}[\mathrm{d}\varphi](\delta' x) + \mathrm{d}[\mathrm{d}\varphi(\delta' x)]](\delta x) \equiv 0\,, \end{aligned}$$

whence $[\delta, \delta']x \in E'$. It follows that the integrability condition is *automatically satisfied* and thus that $x \mapsto E'$ is a foliation, called the
(5.25) *characteristic foliation* of the form φ.

From (5.21) it follows that φ is an integral invariant of its characteristic foliation. If this foliation is sectionable then φ is the inverse image of a form $x' \mapsto \varphi'$ on V'. One can check immediately that $\ker(\varphi') \cap \ker(\mathrm{d}\varphi')$ is zero.

§6. Lie groups

(6.1) DEFINITION: Let G be a group, that is to say, a set on which is defined a *composition law*[†] $\times$, an *identity element* e and an *inverse* $a \mapsto a^{-1}$ such that $\forall a, b, c \in G$:

$$\begin{cases} a \times b \in G\,, \\ [a \times b] \times c = a \times [b \times c]\,, \\ a \times e = e \times a = a\,, \\ a^{-1} \in G\,, \\ a^{-1} \times a = a \times a^{-1} = e\,. \end{cases}$$

One says that G is a *Lie group* if G is a *manifold* and if the maps $\begin{pmatrix} a \\ b \end{pmatrix} \mapsto a \times b$ and $a \mapsto a^{-1}$ are *differentiable*.

(6.2) EXAMPLE: If E is a vector space of finite dimension n, it is a group with respect to the composition law $+$, identity element 0, and inverse $a \mapsto -a$. On E we have defined a manifold structure in (1.9). Since the maps $\begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + b$ and $a \mapsto -a$ are differentiable, E is a *Lie group*, called the *additive group* E. □

(6.3) EXAMPLE: Let us denote by $\mathrm{L}(E)$ the set of linear maps from E into E; $\mathrm{L}(E)$ is a vector space of dimension n^2 and thus is a manifold. $\mathrm{Gl}(E)$[89] denotes the set of elements a of $\mathrm{L}(E)$ such that $\det(a) \neq 0$; it is a group of permutations of E. Since it is an open subset of $\mathrm{L}(E)$[90] it is a manifold.[91] Since we already know that the maps $\begin{pmatrix} a \\ b \end{pmatrix} \mapsto a \cdot b$ and $a \mapsto a^{-1}$ are differentiable,[92] it follows that $\mathrm{Gl}(E)$ is a *Lie group of dimension* n^2. □

[†]Editors' note: Also called *multiplication*.

[89]The *General linear group* of E.

[90]It is the inverse image, under the continuous map det, of $\mathbf{R} \setminus \{0\}$, which is an open set (see (1.16)).

[91]See (1.39).

[92]See (2.28) and (2.30).

Actions of a Lie group on a manifold

(6.4) A Lie group is said to *act* on a manifold V if for each $a \in G$ there is a permutation $\underline{a}_V$ of V such that

$$\diamondsuit \qquad \underline{a \times b}_V = \underline{a}_V \circ \underline{b}_V \qquad \forall a, b \in G$$

and if the map from $G \times V$ to V given by $\begin{pmatrix} a \\ x \end{pmatrix} \mapsto \underline{a}_V(x)$ is differentiable.

(6.5) EXAMPLE: For any $a \in \mathrm{Gl}(E)$ we define $\underline{a}_E = a$. With this definition $\mathrm{Gl}(E)$ *acts* on E. □

EXAMPLE: Let G be an arbitrary Lie group. If for a and b in G we define

$$\underline{a}_G(b) = a \times b\,, \tag{6.6}$$

then G *acts on itself*. The $\underline{a}_G$ are called *translations*† of G. □

If G acts on a manifold V one can make G act on the tangent bundle V^D (1.31) of V by defining

$$\underline{a}_{[V^D]} = [\underline{a}_V]^D \qquad \forall a \in G \tag{6.7}$$

(see (1.32) to (1.34)).

If G acts on V one can check immediately that

$$\begin{cases} \underline{e}_V = 1_V & (e = \text{identity element of } G), \\ [\underline{a}^{-1}]_V = [\underline{a}_V]^{-1} & (\forall a \in G), \end{cases} \tag{6.8}$$

and thus that the $\underline{a}_V$ form a *group of diffeomorphisms* of V; we will denote this group by $\underline{G}_V$.

(6.9) If G acts on a vector space E and if the elements of $\underline{G}_E$ are linear, then the correspondence $a \mapsto \underline{a}_E$ is called a *linear representation* of G.[93]

If G acts on a manifold V we define

$$\hat{x}(a) = \underline{a}_V(x) \qquad [\forall x \in V\,, \quad \forall a \in G]\,. \tag{6.10}$$

It is immediate that $\hat{x}$ is a differentiable map from G to V and that $\hat{x}(e) = x$.

†Editors' note: $\underline{a}_G$ is also called the *left* translation of G by a.

[93]We consider here only finite dimensional representations.

Let $\mathcal{G}$ be the tangent space to G at e (Fig. 6.I) and let Z be an element of $\mathcal{G}$. *On every manifold V on which G acts* we can define a vector field Z_V[†] by

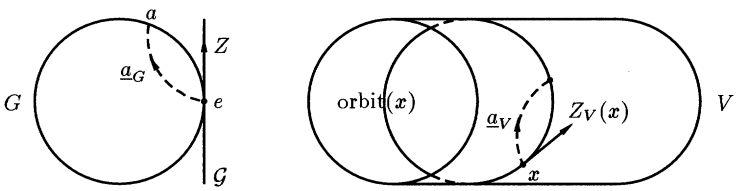
(6.11) $\diamondsuit$ $$Z_V(x) \equiv D(\hat{x})(e)(Z)$$

or, if one prefers, by

(6.11) $\heartsuit$ $$Z_V(x) = \delta[\underline{a}_V(x)] \qquad \text{for } a = e,\ \delta a = Z, \text{ and } \delta x = 0.$$

These formulæ can be applied in particular to the case $V = G$ (compare (6.6)), in which case Z_G is a vector field defined on the group G. Note that $Z_G(e) = Z$.

Fig 6.I



(6.12) THEOREM:

a) On every manifold V on which G acts, the vector field Z_V is *differentiable.*

b) If Z and Z' are elements of $\mathcal{G}$, then there is an element $[Z, Z']$ of $\mathcal{G}$ such that $[Z, Z']_V = [Z_V, Z'_V]$,[94,†] *independent of the manifold V on which the group G acts.*

c) If G acts on a *Hausdorff* manifold V,[95] then $\exp(Z_V)$ exists for every $Z \in \mathcal{G}$ and one has

$$\exp(Z_V) = \underline{\exp(Z_G)(e)}_V .^{96}$$

Conversely, if a vector field f satisfies $\exp(tf) \in \underline{G}_V$ for all real t, then there exists a $Z \in \mathcal{G}$ such that $f = Z_V$. □

[†]Editors' note: The vector field Z_V is called the *fundamental vector field* on the manifold V associated to $Z \in \mathcal{G}$.

[94]The right hand side is the *Lie bracket of the vector fields Z_V and Z'_V* (2.45); the left hand side is the vector field associated with $[Z, Z']$. Of course $[Z, Z']$ will also be called the *Lie bracket* of Z and Z'. Editors' note: The Lie bracket $[Z, Z']$ is also called the *commutator* of Z and Z'.

[†]Editors' note: Traditionally the Lie bracket on $\mathcal{G}$ is defined with the opposite sign: on Lie groups the vector fields Z_G as defined in the text are *right invariant*, whereas the Lie bracket is usually defined in terms of *left invariant* vector fields.

[95]See definition (3.1); one can show that every Lie group is a Hausdorff manifold.

[96]This shows in particular that $\exp(Z_G)$ is a translation of G ($\forall Z \in \mathcal{G}$).

The Lie algebra of a Lie group

(6.13) The set $\mathcal{G}$ (the tangent space to G at e) is called the *Lie algebra* of the Lie group G. One can verify that

a) $\mathcal{G}$ is a finite dimensional real vector space.

b) $\forall Z, Z' \in \mathcal{G}$ we have defined the commutator $[Z, Z']$, also denoted as $\mathrm{ad}(Z)(Z')$, which belongs to $\mathcal{G}$.

c) ad is an antisymmetric bilinear operator:

$$\begin{aligned} [Z, Z'] + [Z', Z] &\equiv 0 \\ [Z_1 + Z_2, Z'] &\equiv [Z_1, Z'] + [Z_2, Z'] \\ [sZ, Z'] &\equiv s\,[Z, Z'] \qquad [\forall s \in \mathbf{R}]\,. \end{aligned}$$

d) The Lie bracket satisfies the Jacobi identity [97]

$$[Z, [Z', Z'']] + [Z', [Z'', Z]] + [Z'', [Z, Z']] = 0\,,$$

which can also be written as

$$\mathrm{ad}([Z, Z']) = \mathrm{ad}(Z) \cdot \mathrm{ad}(Z') - \mathrm{ad}(Z') \cdot \mathrm{ad}(Z)\,. \qquad \square$$

(6.14) Conversely, a Lie algebra is any set $\mathcal{G}$ satisfying (6.13). It follows that the set $\mathcal{G}'$ of elements of $\mathcal{G}$ that are *sums of Lie brackets* is visibly a Lie algebra called the *derived algebra* of $\mathcal{G}$.

(6.15) EXAMPLE: In the case of the additive group E (notation (6.2)), the Lie algebra coincides with the vector space E, the Lie bracket being zero: $[Z, Z'] \equiv 0$. $\square$

EXAMPLE: In the case of the group $\mathrm{Gl}(E)$ (notation (6.3)), it is evident that its Lie algebra is $\mathrm{L}(E)$. For $Z \in \mathrm{L}(E)$ one has $Z_E = Z$ and [98]

(6.16) $$[Z, Z'] = Z' \cdot Z - Z \cdot Z'\,.$$

One has the series expansion

(6.17) $$\exp(Z) = 1 + Z + \frac{Z^2}{2!} + \cdots + \frac{Z^n}{n!} + \cdots\,,$$

[97] See (2.47).

[98] To obtain this formula with the traditional opposite sign, the Lie group and its Lie algebra have to act on the *other side*. Editors' note: Another way to obtain the opposite sign is to use the left-invariant vector fields to define the Lie bracket on $\mathcal{G}$, see the editors' note on page 47.

which converges for all $Z \in \mathrm{L}(E)$. Theorem (6.12) implies that $\exp(Z) \in \mathrm{Gl}(E)$, and thus that $\det(\exp(Z)) \neq 0$. Furthermore one can show[99] that

$$\text{(6.18)} \qquad \begin{cases} \det(\exp(Z)) = \mathrm{e}^{\mathrm{Tr}(Z)} \\ \exp(B \cdot Z \cdot B^{-1}) = B \cdot \exp(Z) \cdot B^{-1} \qquad \forall B \in \mathrm{Gl}(E)\,. \end{cases} \qquad \square$$

Orbits

(6.19) Let G be a Lie group acting on a manifold V (Fig. 6.I). The set $\underline{a}_V(x)$ as a varies over G is called the *orbit* of the point x under the group G; the orbit is thus the image of the map $\widehat{x}$ (notation (6.10)).

One can verify immediately that

$$[y \in \mathrm{orbit}(x)] \iff [\mathrm{orbit}(y) = \mathrm{orbit}(x)]\,,$$

and thus that the orbits form a *partition* of V.

(6.20) THEOREM: Let x be a point of V and let H be the vector space of all $Z_V(x)$ where Z varies over the Lie algebra $\mathcal{G}$ of G.

If the dimension p of H is nonzero, then the orbit Ω which passes through x is a *submanifold of V of dimension p*; its tangent space at x is H.[100]

For each $a \in G$ the map $\underline{a}_\Omega$, the restriction of $\underline{a}_V$ to the orbit Ω, *defines an action of the Lie group G on the manifold Ω.* $\square$

The adjoint representation

Let $\mathcal{G}$ be the Lie algebra of a Lie group G; let $Z \in \mathcal{G}$ and $a \in G$. We have defined the map $\widehat{a}$ (notation (6.10)) which in this case is a diffeomorphism of G given by

$$\text{(6.21)} \qquad \widehat{a}(b) = b \times a\,,$$

and is called *right translation* by a.

[99] Apply (2.29) and (3.6).

[100] Ω is an integral manifold of the field $x \mapsto H$; note that this field defines a foliation of V if the dimension p is constant on V.

(6.22) One can immediately verify that the vector field Z_G (notation (6.11)) is *right invariant*, that is, the *direct image* of Z_G under every right translation is equal to Z_G (definition (2.32) of direct image). Conversely, each right invariant vector field is of the form Z_G for some $Z \in \mathcal{G}$.

Let us now investigate the direct image of Z_G under the (left!) translation $\underline{a}_G$. One can easily check that it is a right invariant vector field, and thus of the form Z'_G for some $Z' \in \mathcal{G}$. Hence we can define a map $\underline{a}_\mathcal{G}$ from $\mathcal{G}$ to itself by (notation (2.32))

$$[\underline{a}_G]_+(Z_G) = [\underline{a}_\mathcal{G}(Z)]_G\,, \tag{6.23}$$

which may also be written as

$$\underline{a}_\mathcal{G}(Z) = \delta[a \times b \times a^{-1}]\,, \qquad \text{with } b = e\,,\ \delta b = Z\,,\ \text{and } \delta a = 0\,. \tag{6.24}$$

It is easy to verify that the map $a \mapsto \underline{a}_\mathcal{G}$ *defines an action of* G *on its Lie algebra, and that the* $\underline{a}_\mathcal{G}$ *are linear*. We conclude that $a \mapsto \underline{a}_\mathcal{G}$ is a linear representation of G on its Lie algebra $\mathcal{G}$, called the *adjoint representation*.

Note that formula (6.23) can be generalized to

$$[\underline{a}_V]_+(Z_V) = [\underline{a}_\mathcal{G}(Z)]_V\,, \tag{6.25}$$

which is valid for every manifold V on which G acts.

When $f = Z'_G$, $g = Z_G$, and $x = e$, formula (3.7) can be written, after some modifications, as

$$\frac{\partial}{\partial t}[\underline{\exp(-tZ'_G)(e)}_\mathcal{G}](Z) = \underline{\exp(-tZ'_G)(e)}_\mathcal{G}([Z', Z]) \qquad \forall Z, Z' \in \mathcal{G}\,.$$

By setting $t = 0$ and using the notation of (6.11) and (6.13b), the above formula becomes

$$Z_\mathcal{G} = -\mathrm{ad}(Z)\,. \tag{6.26}$$

Combining (2.48), (6.12), and (6.23), one can establish the formula

$$\underline{a}_\mathcal{G}([Z, Z']) \equiv [\underline{a}_\mathcal{G}(Z), \underline{a}_\mathcal{G}(Z')]. \tag{6.27}$$

EXAMPLE: In the case of the Lie group $\mathrm{Gl}(E)$, whose Lie algebra is $\mathrm{L}(E)$ (see (6.16)), the adjoint representation is given by

$$\underline{a}_{\mathrm{L}(E)}(Z) = a \cdot Z \cdot a^{-1} \qquad [a \in \mathrm{Gl}(E)\,,\ Z \in \mathrm{L}(E)]\,. \tag{6.28}$$

Lie subalgebras and Lie subgroups

(6.29) Let G be a Lie group and $\mathcal{G}$ its Lie algebra. A subset $\tilde{\mathcal{G}}$ of $\mathcal{G}$ is called a *Lie subalgebra* if

$$\begin{cases} \tilde{\mathcal{G}} \text{ is a vector subspace of } \mathcal{G}\,, \\ Z, Z' \in \tilde{\mathcal{G}} \implies [Z, Z'] \in \tilde{\mathcal{G}}\,. \end{cases}$$ □

(6.30) For an arbitrary element a of G we denote by E the set of all $Z_G(a)$ as Z varies over $\tilde{\mathcal{G}}$. E is a vector subspace of the tangent space to G at a; its dimension p, equal to that of $\tilde{\mathcal{G}}$, is constant. One can verify easily that the field $a \mapsto E$ is *differentiable* (5.2) and that it defines a *foliation* of G.[101] We will call it the *foliation associated to the subalgebra* $\tilde{\mathcal{G}}$.

(6.31) DEFINITION: Let G be a Lie group; we will say that a Lie group $\tilde{G}$ is a *Lie subgroup* of G if

a) as a group $\tilde{G}$ is a subgroup of G;[102]
b) as a manifold $\tilde{G}$ is a submanifold of G.[103] □

In particular, the Lie subgroups of G that have the same dimension as G are the *open subgroups* of G (see (1.39)).

One can see immediately that the Lie algebra $\tilde{\mathcal{G}}$ of a Lie subgroup $\tilde{G}$ of G is a *Lie subalgebra* of the Lie algebra of G. It is natural to ask whether conversely one can characterize $\tilde{G}$ by its Lie algebra $\tilde{\mathcal{G}}$. The following theorem deals with this question.

(6.32) THEOREM: Let G be a Lie group, e its identity element, and $\mathcal{G}$ its Lie algebra. Let $\tilde{\mathcal{G}}$ be a Lie subalgebra of $\mathcal{G}$ (6.29).

I) Let us denote by $\tilde{G}_I$ the leaf passing through e of the foliation associated to $\tilde{G}$ (6.30). Then
   a) $\tilde{G}_I$ is a *Lie subgroup* of G having $\tilde{\mathcal{G}}$ as its Lie algebra.
   b) There exists an open set Ω in $\tilde{\mathcal{G}}$ containing 0 such that the map $Z \mapsto \exp(Z_G)(e)$ is a *diffeomorphism* from Ω to $\tilde{G}_I$.[104]

---

[101] Because $\tilde{\mathcal{G}}$ is not only a vector subspace of $\mathcal{G}$ but also a Lie subalgebra.

[102] As a set $\tilde{G}$ is thus a subset of G; the multiplication on $\tilde{G}$ is inherited from G, and the identity elements of G and $\tilde{G}$ coincide.

[103] Definition (1.37).

[104] It follows that if S is a basis of $\tilde{\mathcal{G}}$, then the map $x \mapsto \exp(S(x)_G)(e)$, $[S(x) \in \Omega]$ is a *chart* of $\tilde{G}_I$; we will call it a *canonical chart*.

c) For $a \in \tilde{G}_I$ there is a finite number of elements ${}^1Z, {}^2Z, \ldots, {}^pZ$ of $\tilde{\mathcal{G}}$ such that

$$a = [\exp({}^1Z_G) \cdot \exp({}^2Z_G) \cdots \exp({}^pZ_G)](e) .$$

II) Let us denote by $\tilde{G}_{II}$ the set of elements a of G such that $[Z \in \tilde{\mathcal{G}}] \Rightarrow [\underline{a}_{\mathcal{G}}(Z) \in \tilde{\mathcal{G}}]$.[105] Then
   a) $\tilde{G}_{II}$ is a *subgroup* of G.
   b) $\tilde{G}_{II}$ is a union of leaves of the foliation associated to $\tilde{\mathcal{G}}$ and hence is a *submanifold* of G.
   c) With this group structure and this manifold structure $\tilde{G}_{II}$ is a *Lie subgroup* of G with Lie algebra $\tilde{\mathcal{G}}$.
   d) For $a \in \tilde{G}_{II}$, the map $Z \mapsto \exp(Z_G)(a)$ is a *diffeomorphism* from Ω to $\tilde{G}_{II}$.

III) Let $\tilde{G}$ be a *subgroup* of G. Then the following two properties are equivalent.
   a) $\tilde{G}$ is a *Lie subgroup* of G and as such has $\tilde{\mathcal{G}}$ as its Lie algebra.
   b) $\tilde{G}_I \subseteq \tilde{G} \subseteq \tilde{G}_{II}$.

   If one of these conditions is satisfied then $\tilde{G}$ is an *open* subgroup of $\tilde{G}_{II}$. $\tilde{G}_I$ is the *connected component*[106] of e in $\tilde{G}$ and constitutes a *normal subgroup*[107] of $\tilde{G}$. □

From these results one can easily deduce the following criterion.

(6.33) Let G be a Lie group, e its identity element, and $\mathcal{G}$ its Lie algebra. Let $\tilde{G}$ and $\tilde{\mathcal{G}}$ be subsets of G and $\mathcal{G}$, respectively. Then the following conditions are equivalent.

a) $\begin{cases} \tilde{G} \text{ is a subgroup of } G, \\ \tilde{\mathcal{G}} \text{ is a vector subspace of } \mathcal{G}, \\ \left[\begin{array}{l} Z \in \tilde{\mathcal{G}} \\ a \in \tilde{G} \end{array}\right] \implies \left[\begin{array}{l} \exp(Z_G)(e) \in \tilde{G} \\ \underline{a}_{\mathcal{G}}(Z) \in \tilde{\mathcal{G}} \end{array}\right]. \end{cases}$

b) $\begin{cases} \tilde{G} \text{ admits the structure of a Lie subgroup of } G \text{ with Lie} \\ \text{algebra } \tilde{\mathcal{G}}. \end{cases}$ □

(6.34) Evidently it follows that *any intersection of Lie subgroups* G_j of G is a Lie subgroup of G (and of the G_js) whose Lie algebra is the intersection of the Lie algebras of the G_j.

---

[105] $\tilde{G}_{II}$ is called the *stabilizer* of $\tilde{\mathcal{G}}$.

[106] See (1.49).

[107] That is, $[a \in \tilde{G}, b \in \tilde{G}_I] \Rightarrow [a \times b \times a^{-1} \in \tilde{G}_I]$.

(6.35) Taking $\tilde{\mathcal{G}} = \mathcal{G}$ one sees that *the connected component of the identity element of a Lie group is always a normal subgroup.*

The stabilizer

(6.36) THEOREM/DEFINITION: Let G be a Lie group with Lie algebra $\mathcal{G}$ acting on a manifold V. Let us define

$$\diamondsuit \quad \begin{cases} [a \in G_x] & \Longleftrightarrow & [\underline{a}_V(x) = x] \\ [Z \in \mathcal{G}_x] & \Longleftrightarrow & [Z_V(x) = 0] \end{cases} \qquad \forall a \in G\,,\ \forall x \in V\,,\ \forall Z \in \mathcal{G}\,.$$

Then G_x is a Lie subgroup of G, having $\mathcal{G}_x$ as its Lie algebra, which is called the *stabilizer* (or *isotropy group*) of x. We have

$$\heartsuit \quad \begin{cases} [a \in G_{\underline{b}(x)}] & \Longleftrightarrow & [b^{-1} \times a \times b \in G_x] \\ \left[Z \in \mathcal{G}_{\underline{b}(x)}\right] & \Longleftrightarrow & \left[\underline{(b^{-1})}_{\mathcal{G}}(Z) \in \mathcal{G}_x\right] \end{cases} \qquad \forall b \in G\,,$$

and thus stabilizers of different points on the same orbit are *conjugate* Lie groups. □

Classical examples of Lie groups

Let E be a real vector space of dimension n. We have seen that $\mathrm{Gl}(E)$ is a Lie group of dimension n^2 having $\mathrm{L}(E)$ as its Lie algebra.

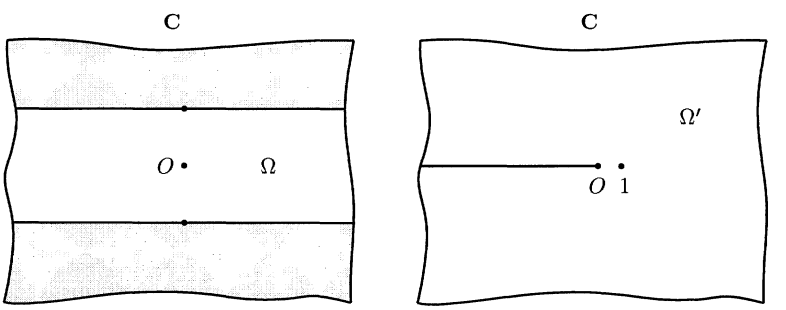


Fig 6.II

(6.37) If we define Ω to be the set of elements $Z \in \mathrm{L}(E)$ whose spectrum is contained in the strip $|\overline{z} - z| < 2\pi$ (Fig. 6.II), then the conditions of theorem (6.32.I.b) are satisfied. Hence the map

$$Z \mapsto \exp(Z) \qquad [Z \in \Omega]$$

is a diffeomorphism from the open set Ω onto an open subset Ω' of $\mathrm{Gl}(E)$. The elements of Ω' are those linear maps in $\mathrm{Gl}(E)$ that do not have negative eigenvalues (or zero eigenvalues, as they are injective).

The inverse diffeomorphism of (6.37) is called the *logarithm* and will be denoted by Log. One can show that

$$\mathrm{Log}(A) = \int_{-\infty}^{0} \{[s - A]^{-1} - [s - 1]^{-1}\}\, ds\,, \qquad A \in \Omega'\,. \tag{6.38}$$

From this formula one can deduce that

$$\mathrm{Log}(B \cdot A \cdot B^{-1}) \equiv B \cdot \mathrm{Log}(A) \cdot B^{-1} \qquad B \in \mathrm{Gl}(E) \tag{6.39}$$

$$\mathrm{Log}(A^{-1}) \equiv -\mathrm{Log}(A)\,. \tag{6.40}$$

(6.41) The connected component of e in $\mathrm{Gl}(E)$ (which is a normal subgroup according to (6.35)) is the set of elements of $\mathrm{L}(E)$ whose determinant is positive;[108] it evidently contains the open set Ω'.

One can check immediately that the elements a of $\mathrm{L}(E)$ that satisfy the *equation*

$$\det(a) = 1 \tag{6.42}$$

form a Lie subgroup of $\mathrm{Gl}(E)$, denoted by $\mathrm{Sl}(E)$.[109] $\mathrm{Sl}(E)$ is connected, its dimension is $n^2 - 1$, and its Lie algebra is made up of the elements of $\mathrm{L}(E)$ whose *trace is zero*.[110] □

Let E be a *complex* vector space of dimension n. We know that E can be considered as a *real vector space of dimension* $2n$; we will denote by the subscripts $\mathbf{C}$ and $\mathbf{R}$ concepts relating to the complex and the real structure of E, respectively. Thus $\mathrm{L}(E)_{\mathbf{C}}$ denotes the set of $\mathbf{C}$-linear maps from E to E. One can check that $\mathrm{L}(E)_{\mathbf{C}}$ is the set of elements of $\mathrm{L}(E)_{\mathbf{R}}$ that *commute* with the map $\underline{i}$:

$$\underline{i}(x) = ix \qquad \forall x \in E\,. \tag{6.43}$$

(6.44) One can easily see that $\mathrm{Gl}(E)_{\mathbf{C}}$ $[= \mathrm{Gl}(E)_{\mathbf{R}} \cap \mathrm{L}(E)_{\mathbf{C}}]$ is a *Lie subgroup* of $\mathrm{Gl}(E)_{\mathbf{R}}$ of dimension $2n^2$ with Lie algebra $\mathrm{L}(E)_{\mathbf{C}}$.

(6.45) For $a \in \mathrm{Gl}(E)_{\mathbf{C}}$, there exists a number φ such that $e^{i\varphi}a$ does not have a negative eigenvalue. One can easily see, using (6.38), that

$$b = \mathrm{Log}(e^{i\varphi}\, a) - i\varphi\, 1_E$$

[108] This fact can be deduced from Cartan's decomposition (6.70).

[109] The *Special linear group* of E.

[110] Apply formula (2.29).

is an element of $\mathrm{L}(E)_{\mathbf{C}}$ and that $\exp(b) = a$. It follows that the continuous map F given by $F(t) \equiv \exp(tb)$ satisfies $F(0) = 1_E$ and $F(1) = a$; as a result $\mathrm{Gl}(E)$ is connected.[111] Applying the formula

$$\mathrm{Tr}_{\mathbf{R}}(Z) = 2\,\mathrm{Re}(\mathrm{Tr}_{\mathbf{C}}Z)) \qquad [\forall Z \in \mathrm{L}(E)_{\mathbf{C}}] \tag{6.46}$$

and noting that formula (6.18) is valid for the complex structure as well as the real structure, one can see that

$$\begin{aligned} \det_{\mathbf{R}}(a) &= \exp(\mathrm{Tr}_{\mathbf{R}}(b)) = \exp(2\,\mathrm{Re}(\mathrm{Tr}_{\mathbf{C}}(b))) \\ &= |\exp(\mathrm{Tr}_{\mathbf{C}}(b))|^2 = |\det_{\mathbf{C}}(a)|^2 . \end{aligned}$$

The thus established formula

$$\det_{\mathbf{R}}(a) = |\det_{\mathbf{C}}(a)|^2 \tag{6.47}$$

is in fact true for every $a \in \mathrm{L}(E)_{\mathbf{C}}$ because both sides are zero for $a \notin \mathrm{Gl}(E)_{\mathbf{C}}$. □

(6.48) The set $\mathrm{Sl}(E)_{\mathbf{C}}$, defined as the set of all a with $a \in \mathrm{Gl}(E)_{\mathbf{C}}$ and $\det_{\mathbf{C}}(a) = 1$, is a *connected Lie group* of dimension $2[n^2 - 1]$ having as Lie algebra the set of all Z such that $Z \in \mathrm{L}(E)_{\mathbf{C}}$ and $\mathrm{Tr}_{\mathbf{C}}(Z) = 0$. $\mathrm{Sl}(E)_{\mathbf{C}}$ is a Lie subgroup of $\mathrm{Gl}(E)_{\mathbf{C}}$ and of $\mathrm{Sl}(E)_{\mathbf{R}}$ (see formula (6.47)). □

Euclidean spaces

(6.49) Let E be a finite dimensional *real* vector space. E is called a *Euclidean* space if on E there is a symmetric and injective tensor g of degree 2:

$$\begin{cases} g(X)(Y) = g(Y)(X) \in \mathbf{R} & [X, Y \in E] \\ [g(X) = 0]^{112} \iff [X = 0] . \end{cases}$$

A *complex* vector space of finite dimension n will be called *Euclidean* if

a) seen as a real vector space of dimension $2n$, E is Euclidean;

b) $g(iX)(iY) = g(X)(Y) \qquad \forall X, Y \in E$.

(6.50) EXAMPLE: $\mathbf{R}$ is a *real Euclidean space* if one defines

$$g(x)(y) = xy \qquad \forall x, y \in \mathbf{R} . \tag{□}$$

[111] See (1.52).

[112] $[g(X) = 0]$ can also be written as $[g(X)(Y) = 0\ \forall Y \in E]$.

(6.51) EXAMPLE: $\mathbf{C}$ is a *complex Euclidean space* if one defines

$$\begin{aligned} g(x+iy)(x'+iy') &= xx'+yy' && [\forall x,y,x',y' \in \mathbf{R}] \\ &= \mathrm{Re}\big(\overline{(x+iy)}\cdot(x'+iy')\big). \end{aligned}$$ □

EXAMPLE: A direct product of Euclidean spaces[113] is Euclidean if one defines

(6.52) $$g\begin{pmatrix} {}^1x \\ \vdots \\ {}^nx \end{pmatrix}\begin{pmatrix} {}^1y \\ \vdots \\ {}^ny \end{pmatrix} = g({}^1x)({}^1y)+\cdots+g({}^nx)({}^ny)\,.$$ □

(6.53) THEOREM: Let E and E' be two Euclidean spaces and let A be a linear map from E to E'. Then there exists a map $\overline{A}$ from E' to E, called the *transpose* of A, defined by

$$g(A(X))(Y) = g(X)(\overline{A}(Y)) \qquad \forall X \in E \text{ and } \forall Y \in E'\,.^{\dagger}$$

The transposition operation $\overline{\ }$ satisfies the following rules:

a) A linear $\Rightarrow$ $\overline{A}$ is linear;[114]

b) $g(X)(Y) = \mathrm{Re}(\overline{X}\cdot Y) \qquad [\forall X, Y \in E]$;[115]

c) $\overline{s}$ equals the complex conjugate of s for $s \in \mathbf{C}$;[116]

d) $\overline{1}_E = 1_E$ if E is Euclidean;

e) $\overline{A+B} \equiv \overline{A}+\overline{B}$;

f) $\overline{sA} \equiv \overline{s}\,\overline{A}$ for every scalar s;

g) $\overline{\overline{A}} \equiv A$;

h) $\overline{A\cdot B} \equiv \overline{B}\cdot\overline{A}$;[117]

i) $\mathrm{rank}(\overline{A}) = \mathrm{rank}(A)$;[118]

j) $\left[\overline{A}\right]^{-1} = \overline{A^{-1}}$ provided A maps a Euclidean space onto a Euclidean space of the same dimension;[119]

---

[113]All real or all complex (and then the direct product will be of the same kind); in particular, $\mathbf{R}^n$ and $\mathbf{C}^n$ are Euclidean spaces, real and complex respectively.

[†]Editors' note: On the left hand side g denote the tensor on E' and on the right hand side g denotes the tensor on E.

[114]$\mathbf{R}$ or $\mathbf{C}$ linear.

[115]This formula assumes that one has identified each element X of E with the linear map $X(s) = sX \quad \forall s \in \mathbf{R}$ (or $\forall s \in \mathbf{C}$, if E is complex).

[116]Thus $\overline{s} = s$ if s is real.

[117]This formula assumes of course that B takes its values in the Euclidean space $\mathrm{dom}(A)$.

[118]This formula is valid in the complex case as well as in the real case.

[119]In particular if A is a *basis* of E or if $A \in \mathrm{Gl}(E)$.

k) $\mathrm{Tr}(\overline{A}) = \overline{\mathrm{Tr}(A)}$, $\det(\overline{A}) = \overline{\det(A)}$, and $\exp(\overline{A}) = \overline{\exp(A)}$ for every $A \in \mathrm{L}(E)$;[118]

l) $\mathrm{Log}(\overline{A}) = \overline{\mathrm{Log}(A)}$ provided $\mathrm{Log}(A)$ exists;

m) ${}^{j}\left[\overline{M}\right]_{k} = \overline{{}^{k}M_{j}}$ if M is a matrix.[120] □

(6.54) If E is a Euclidean space, the number $\langle X, Y\rangle = \overline{X} \cdot Y$ is called the *scalar product* of X and Y in E. *If this number is real*, the scalar product coincides with $g(X)(Y)$ (6.53b). The usual properties of the scalar product, and in particular the formula

$$\langle Y, X\rangle \equiv \overline{\langle X, Y\rangle}\,, \tag{6.55}$$

follow immediately from the properties of the transposition operation.

(6.56) "Ordinary" vectors form a *Euclidean space* in the sense of (6.49), the scalar product $\langle X, Y\rangle$ being the "ordinary" scalar product; one identifies this Euclidean space with $\mathbf{R}^3$. □

It is immediate that the "scalar square" $\langle X, X\rangle$ of an element of E is always real. The Euclidean space E is called *positive* (respectively *negative*) if

$$[X \neq 0] \Rightarrow \left[\overline{X} \cdot X > 0\right] \qquad \text{(respectively } \overline{X} \cdot X < 0\text{)}. \tag{6.57}$$

"Ordinary" space $\mathbf{R}^3$ and the spaces $\mathbf{R}^n$ and $\mathbf{C}^n$ are positive. □

(6.58) Let E be a Euclidean space and suppose that $Z \in \mathrm{L}(E)$. Z is called *hermitian* (respectively *skew hermitian*) if $\overline{Z} = Z$ (respectively $\overline{Z} = -Z$).

If Z and Z' are skew hermitian one can note that

$$\overline{Z' \cdot Z - Z \cdot Z'} = \overline{Z} \cdot \overline{Z'} - \overline{Z'} \cdot \overline{Z} = -[Z' \cdot Z - Z \cdot Z']\,.$$

(6.59) Consequently, *the set of skew hermitian matrices form a Lie subalgebra* $\tilde{\mathcal{G}}$ of $\mathrm{L}(E)$ (see (6.16)).

Let us form the largest subgroup $\tilde{G}_{II}$ admitting $\tilde{\mathcal{G}}$ as a Lie group using the procedure of (6.32.II). One finds immediately that $\tilde{G}_{II}$ is the set of those $A \in \mathrm{Gl}(E)$ such that $\overline{A} \cdot A$ *commutes with every skew hermitian matrix*. One can show that this implies that there exists a scalar s (necessarily real) such that

$$\overline{A} \cdot A = s\, 1_E\,, \tag{6.60}$$

[120] In other words, one obtains the matrix $\overline{M}$ by taking the transpose of each of the elements of M and then interchanging rows and columns. As a matter of fact, this formula is an immediate consequence of $\overline{l_j} = {}^{j}l$ (notation (1.20) and (1.29)).

which can also be written as

$$\overline{A} = sA^{-1}\,. \tag{6.61}$$

Note that s is constant on each connected component of $\tilde{G}_{II}$.[121] It follows that the set of all A such that

$$\overline{A} = A^{-1} \tag{6.62}$$

forms an open subgroup of $\tilde{G}_{II}$, and thus a *Lie subgroup* of $\mathrm{Gl}(E)$ *having the skew hermitian matrices as its Lie algebra.*

(6.63) If E is *real*, this group is called the *orthogonal group* of E and is denoted by $\mathrm{O}(E)$; its dimension is $n(n-1)/2$. Formula (6.53k) shows that each element A of $\mathrm{O}(E)$ satisfies $\det(A) = \pm 1$. It follows that the set

$$\mathrm{SO}(E) = \mathrm{O}(E) \cap \mathrm{Sl}(E) \tag{6.64}$$

is an *open subgroup* of $\mathrm{O}(E)$. Hence $\mathrm{SO}(E)$, called the *special orthogonal group* of E, is another Lie group having the (real) skew hermitian matrices as its Lie algebra.

(6.65) If E is *positive*, one can show that every element of $\mathrm{SO}(E)$ can be written as $\exp(Z)$ for some skew hermitian matrix Z. Thus $\mathrm{SO}(3)$ is connected and $\mathrm{O}(E)$ has two *components*: on the one hand $\mathrm{SO}(E)$ and on the other the set of elements of determinant -1.[122] □

(6.66) If E is *complex*, the group defined by (6.62) is called the *unitary* group of E and is denoted by $\mathrm{U}(E)$; its dimension is n^2.

Every element A of $\mathrm{U}(E)$ satisfies $|\det(A)| = 1$; those elements whose determinant is equal to 1 form the Lie subgroup $\mathrm{SU}(E)$, called the *special unitary group*. The Lie algebra of $\mathrm{SU}(E)$ is the set of matrices Z such that

$$\overline{Z} = -Z \qquad \text{and} \qquad \mathrm{Tr}(Z) = 0\,, \tag{6.67}$$

and hence its dimension is $n^2 - 1$.

(6.68) If in addition E is *positive*, then each element of $\mathrm{U}(E)$ (respectively $\mathrm{SU}(E)$) can be written as $\exp(Z)$ with Z belonging to the Lie algebra of $\mathrm{U}(E)$ (respectively $\mathrm{SU}(E)$). It follows that $\mathrm{U}(E)$ (respectively $\mathrm{SU}(E)$) is *connected*.[123] □

[121] The *same subgroup* possesses *another Lie subgroup structure* having as its Lie algebra the set of matrices of the form $s\,1_E + Z$ with $\overline{Z} = -Z$.

[122] This is the case for $\mathrm{O}(\mathbf{R}^n)$ and $\mathrm{SO}(\mathbf{R}^n)$, usually denoted by $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ respectively.

[123] This is the case for $\mathrm{U}(\mathbf{C}^n)$ and $\mathrm{SU}(\mathbf{C}^n)$, usually denoted by $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ respectively.

(6.69) Let us now assume that $A \in \mathrm{O}(E)$ (respectively $\mathrm{U}(E)$) (with E an arbitary Euclidean space), and that $\mathrm{Log}(A)$ exists. Taking the logarithm of both sides of the equation $\overline{A} = A^{-1}$ and applying (6.40) and (6.53l) one can see that $\mathrm{Log}(A)$ is skew hermitian and thus belongs to the Lie algebra of $\mathrm{O}(E)$ (respectively $\mathrm{U}(E)$). As in (6.45) it follows that A belongs to the connected subgroup.

Returning to definition (6.37) of Ω', we see that *every element of* $\mathrm{O}(E)$ *that does not belong to the connected subgroup possesses a negative eigenvalue.*

Matrix realizations

(6.70) THEOREM (THE CARTAN DECOMPOSITION (for matrices)): Let M be an injective square matrix with real or complex entries for which we define

$$\heartsuit \qquad B = \tfrac{1}{2}\,\mathrm{Log}(\overline{M} \cdot M) \qquad \text{and} \qquad A = M \cdot \exp(-B)\,.$$

The matrices A and B always exist; they are real if M is real. They satisfy

$$\diamondsuit \qquad M = A \cdot \exp(B)\,, \qquad \overline{A} \cdot A = 1\,, \qquad \text{and} \qquad \overline{B} = B\,.$$

Moreover, the given matrices A and B are the only solutions of $\diamondsuit$, and the map $M \mapsto \begin{pmatrix} A \\ B \end{pmatrix}$ is a diffeomorphism.[124] □

This theorem shows that $\mathrm{Gl}(\mathbf{R}^n)$, also denoted by $\mathrm{Gl}(n)$, is diffeomorphic to $\mathrm{O}(n) \times \mathbf{R}^{n(n+1)/2}$, and that $\mathrm{Gl}(\mathbf{C}^n)_{\mathbf{C}}$, also denoted by $\mathrm{Gl}(n, \mathbf{C})$, is diffeomorphic to $\mathrm{U}(n) \times \mathbf{R}^{n^2}$.

(6.71) Let $S = [S_1 \ldots S_n]$ be a basis of a Euclidean space E. It is easy to see that $\overline{S} = \begin{bmatrix} \overline{S_1} \\ \vdots \\ \overline{S_n} \end{bmatrix}$, and thus that $\overline{S} \cdot S = G$ is an injective matrix called the *Gram matrix.*

One can easily construct an *orthonormal basis* S, that is, a basis such

---

[124] Of the general linear group onto the manifold of $\begin{pmatrix} A \\ B \end{pmatrix}$ such that $\overline{A} \cdot A \equiv 1$ and $\overline{B} \equiv B$.

that

$$G = \overline{S} \cdot S = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{bmatrix} \begin{matrix} \} p \\ \\ \\ \\ \} q \end{matrix} \quad \text{with } p+q=n\,. \tag{6.72}$$

The number p is the maximum dimension of a positive Euclidean subspace of E and thus does not depend on the choice of S; it is called the *index of inertia* of E.

We now assume that E is *hyperbolic*, that is, that $1 \leq p \leq n-1$.[125] Every element a of $\mathrm{Gl}(E)$ can be obviously be brought into the form

$$a = S \cdot M \cdot S^{-1}\,, \tag{6.73}$$

for some square matrix M. If $\overline{a} \cdot a = 1$, one can verify immediately that

$$M^{-1} = G \cdot \overline{M} \cdot G^{-1} \qquad [G \text{ as given in (6.72)}]\,. \tag{6.74}$$

The matrices M satisfying this equation form a Lie group, called $\mathrm{O}(p,q)$ in the real case and $\mathrm{U}(p,q)$ in the complex case, which is evidently isomorphic to $\mathrm{O}(E)$ (or $\mathrm{U}(E)$).[126]

Let M be an element of this group; by inverting and transposing (6.74) one can see that $\overline{M}$ belongs to the group, and thus $\overline{M} \cdot M$. The argument of (6.69) shows that $B = \frac{1}{2}\mathrm{Log}(\overline{M} \cdot M)$ belongs to the Lie algebra of this group, and hence the two factors of Cartan's decomposition (6.70$\diamondsuit$) of M belong to this group. One can verify immediately that B anti-commutes with the Gram matrix G, and that A commutes with G. Using these relations one arrives at the following result.

If M satisfies (6.74), then the same is true for the matrices in its Cartan decomposition, which can be written as

$$M = \left[\begin{array}{c|c} C & \\ \hline & D \end{array}\right] \cdot \exp\left(\left[\begin{array}{c|c} & \overline{E} \\ \hline E & \end{array}\right] \right)\,, \tag{6.75}\ \heartsuit$$

where C and D are square matrices of size p and q respectively satisfying $\overline{C} \cdot C = 1$ and $\overline{D} \cdot D = 1$, and where E is a $q \times p$ matrix.

[125] This is the same as saying that E is neither positive (for then one would have $\overline{S} \cdot S = 1$) nor negative (for the one would have $\overline{S} \cdot S = -1$). In these two cases the matrices A and M of (6.70) are simultaneously orthogonal or unitary.

[126] The correspondence $M \mapsto a$ is an isomorphism of groups and a diffeomorphism of manifolds; that is why it is called a Lie group isomorphism.

(6.76) Conversely, if C, D, and E are matrices satisfying the above conditions, one can see that M defined by $\heartsuit$ satisfies (6.74). Thus the diffeomorphism $M \mapsto \begin{pmatrix} C \\ D \\ E \end{pmatrix}$ shows that $\mathrm{O}(p,q)$ is diffeomorphic to $\mathrm{O}(p) \times \mathrm{O}(q) \times \mathbf{R}^{pq}$, and that $\mathrm{U}(p,q)$ is diffeomorphic to $\mathrm{U}(p) \times \mathrm{U}(q) \times \mathbf{R}^{2pq}$.[127] Taking (6.65) and (1.50) into account, one can see that $\mathrm{O}(p,q)$ has *four* connected components of which two form $\mathrm{SO}(p,q)$; likewise $\mathrm{U}(p,q)$ and $\mathrm{SU}(p,q)$ are connected. □

[127] One can corroborate these facts by a dimension count:

$$\tfrac{1}{2}\, n(n-1) = \tfrac{1}{2}\, p(p-1) + \tfrac{1}{2}\, q(q-1) + pq \quad \text{and} \quad n^2 = p^2 + q^2 + 2pq\,.$$

§7. The calculus of variations

Classical variational problems

Let us be given a *Lagrangian*, that is, a differentiable real valued function

$$\lambda \equiv \Lambda(t, x, x') \qquad [t \in \mathbf{R},\ x \in \mathbf{R}^n,\ \text{and } x' \in \mathbf{R}^n], \tag{7.1}$$

as well as two numbers t_0 and t_1. To a differentiable map $F(t \mapsto x)$ we associate the number

$$a = \int_{t_0}^{t_1} \Lambda(t, x, x')\, dt \qquad \left[x' \equiv \frac{dx}{dt}\right], \tag{7.2}$$

called the *action*.

(7.3) Let us now assume that F depends on a parameter u in such a way that the map $\begin{pmatrix} t \\ u \end{pmatrix} \mapsto x$ is differentiable and let us denote by δ the derivation with respect to u defined by $\delta u \equiv 1$ and $\delta t \equiv 0$. δ is called a *variation* of F, whence the term *calculus of variations*.

By means of an integration by parts, one can easily establish the formula

$$\delta a = \left[\frac{\partial \lambda}{\partial x'}(\delta x)\right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \left\{\frac{\partial \lambda}{\partial x} - \frac{d}{dt}\left[\frac{\partial \lambda}{\partial x'}\right]\right\}(\delta x)\, dt\,. \tag{7.4}$$

The first variational problems studied were of the following kind. One chooses two points x_0 and x_1 of $\mathbf{R}^n$ and one then searches among the collection of differentiable functions F satisfying

$$F(t_0) = x_0 \qquad \text{and} \qquad F(t_1) = x_1 \tag{7.5}$$

for one which *minimizes* the action a.

Using formula (7.4) one can show that if such a function $t \mapsto x$ exists, it has to satisfy

$$\frac{\partial \lambda}{\partial x} - \frac{d}{dt}\left[\frac{\partial \lambda}{\partial x'}\right] \equiv 0 \qquad \text{with } x' \equiv \frac{dx}{dt}\,. \tag{7.6}$$

□

In certain cases one can solve (7.5) and (7.6), and determine whether among the solutions of this system there exists one which minimizes

the action.[128] We will only study the equations (7.6), called *the Euler-Lagrange equations*,[129] without trying to find minima for the action. The Euler-Lagrange equations only depend on the Lagrangian (and not on the data t_0, x_0, t_1, x_1); their solutions are called *Lagrangian extrema.*

Canonical variables

It is useful to look for the Lagrangian extrema in *parametric* form. More precisely, we look for a *chart* for the *graph* of an extremum, that is, a differentiable map

$$s \mapsto \begin{pmatrix} x \\ t \end{pmatrix} \qquad s \in \mathbf{R} \tag{7.7}$$

such that $t \mapsto x$ is an extremum. If this map is defined on an interval and if s_0 and s_1 are points of this interval to which correspond $\begin{pmatrix} x_0 \\ t_0 \end{pmatrix}$ and $\begin{pmatrix} x_1 \\ t_1 \end{pmatrix}$ respectively, then — due to an ordinary change of variables — the corresponding action can be written as

$$a = \int_{s_0}^{s_1} \Lambda(t, x, \frac{\dot{x}}{\dot{t}})\, \dot{t}\, ds \qquad \left[\dot{x} \equiv \frac{dx}{ds} \text{ and } \dot{t} \equiv \frac{dt}{ds}\right]. \tag{7.8}$$

If we define

$$\begin{cases} q \equiv \begin{pmatrix} x \\ t \end{pmatrix} \text{ and } \dot{q} \equiv \begin{pmatrix} \dot{x} \\ \dot{t} \end{pmatrix} \\ \\ \ell \equiv L(q, \dot{q}) \equiv \Lambda(t, x, \frac{\dot{x}}{\dot{t}})\, \dot{t}\,, \end{cases} \tag{7.9}$$

(7.8) can also be written as

$$a = \int_{s_0}^{s_1} L(q, \dot{q})\, ds\,. \tag{7.10}$$

Applying formula (7.4) to this integral gives

$$\delta a = \left[\frac{\partial \ell}{\partial \dot{q}}(\delta q)\right]_{s_0}^{s_1} + \int_{s_0}^{s_1} \left\{\frac{\partial \ell}{\partial q} - \frac{d}{ds}\left[\frac{\partial \ell}{\partial \dot{q}}\right]\right\}(\delta q)\, ds\,. \tag{7.11}$$

We now will transform (7.11) into (7.16), taking into account that the new Lagrangian L defined by (7.9) does not contain the variable s and that it is visibly *homogeneous of degree* 1 in $\dot{q}$, that is,

$$L(q, \mu\dot{q}) \equiv \mu L(q, \dot{q}) \qquad \forall \mu \neq 0\,. \tag{7.12}$$

[128]It is also among the solutions of this system that one has to look for possible maxima of the action.

[129]They were written by Euler in 1744.

Differentiating (7.12) with respect to μ and then taking $\mu = 1$, one obtains *Euler's identity* for a homogeneous function, namely

$$\ell = p(\dot{q})^{\dagger} \tag{7.13}$$

with

$$p \equiv \frac{\partial \ell}{\partial \dot{q}}\,. \tag{7.14}$$

Differentiating (7.13) gives the identity

$$\frac{\partial \ell}{\partial q}(\delta q) \equiv [\delta p]\,(\dot{q}) \tag{7.15}$$

which, substituted into (7.11), gives

$$\delta a = [p(\delta q)]_{s_0}^{s_1} + \int_{s_0}^{s_1} \left\{ \delta p \left(\frac{dq}{ds} \right) - \frac{dp}{ds}(\delta q) \right\} ds \qquad \left[\dot{q} \equiv \frac{dq}{ds} \right]. \tag{7.16}$$

From definition (7.14) it is clear that p is an element of the dual space of $\mathbf{R}^{n+1}$, that is, a row vector

$$p \equiv [p_1\, p_2\, \ldots\, p_{n+1}]\,. \tag{7.17}$$

One can easily verify that p is *homogeneous of degree* 0 in $\dot{q}$, that is, p only depends on the direction of $\dot{q}$. In many cases this will allow us to eliminate $\dot{q}$ in equation (7.14), meaning that one can find a differentiable real valued function φ with nonzero derivative such that

$$[\varphi(p,q) = 0] \iff \left[\text{there exists } \dot{q} \text{ such that } p = \frac{\partial \ell}{\partial \dot{q}} \right]. \tag{7.18}$$

[130]

In such a case, the equation $\varphi(p,q) = 0$ is called *Jacobi's equation* of the variational problem.[131] One can show that the extrema are the lines of force of the system

$$\frac{\delta\, {}^1q}{\frac{\partial \varphi}{\partial p_1}} = \cdots = \frac{\delta\, {}^{n+1}q}{\frac{\partial \varphi}{\partial p_{n+1}}} = -\frac{\delta p_1}{\frac{\partial \varphi}{\partial\, {}^1q}} = \cdots = -\frac{\delta p_{n+1}}{\frac{\partial \varphi}{\partial\, {}^{n+1}q}} \tag{7.19}$$

[†]Editors' note: This should be read as the row vector p applied to the column vector $\dot{q}$ (see also (7.17)).

[130]This assumes that the problem is *regular*, that is, the kernel of $\frac{\partial p}{\partial \dot{q}}$ (which contains $\dot{q}$) has dimension 1.

[131]The *partial differential equation of Jacobi* is written as $\varphi(\frac{\partial \mathcal{S}}{\partial q}, q) \equiv 0$, where $\mathcal{S}$ is an unknown function of q.

for which $\varphi(p,q)$ (which is a first integral of this system) takes the value 0.

The Hamiltonian formalism

Looking again at the definitions in (7.9), one immediately finds the formulæ

$$\text{(7.20)}\qquad \begin{cases} {}^{j}q \equiv {}^{j}x\,, & p_j \equiv \dfrac{\partial\lambda}{\partial\,{}^{j}x'} & [j=1,2,\dots,n] \\ {}^{n+1}q \equiv t\,, & p_{n+1} \equiv \lambda - \dfrac{\partial\lambda}{\partial x'}(x') \equiv \lambda - \displaystyle\sum_{j=1}^{n} p_j\,{}^{j}x'\,. \end{cases}$$

The variational problem is called *Hamiltonian* if there exists a differentiable function H, called the *Hamiltonian function*, such that

$$\text{(7.21)}\qquad h \equiv \frac{\partial\lambda}{\partial x'}(x') - \lambda \equiv H(t,{}^{1}x,\dots,{}^{n}x,p_1,\dots,p_n)\,.$$

Under these conditions Jacobi's equation can be written as

$$\text{(7.22)}\qquad \varphi(p,q) \equiv p_{n+1} + H(t,{}^{1}q,\dots,{}^{n}q,p_1,\dots p_n) = 0\,.$$

Substituting this into (7.19) one immediately obtains

$$\text{(7.23)}\qquad \frac{\delta\,{}^{j}q}{\delta t} = \frac{\partial h}{\partial p_j}\,, \qquad \frac{\delta p_j}{\delta t} = -\frac{\partial h}{\partial\,{}^{j}q} \qquad [j=1,2,\dots,n]\,,$$

and

$$\text{(7.24)}\qquad \frac{\delta h}{\delta t} = \frac{\partial h}{\partial t}\,.$$

These *canonical equations* were written by Lagrange in 1808; they are called *Hamilton's equations.*

(7.25) Note that equation (7.24) can be omitted because it is a consequence of (7.21) and (7.23). Its main interest is to show that h is a first integral if the variable t does not appear in the Hamiltonian function H.

A geometrical interpretation of the canonical equations

Let us define

$$y \equiv \begin{pmatrix} \dot{q} \\ q \end{pmatrix} \qquad \text{and} \qquad \varpi(\delta y) \equiv p(\delta q)\,. \tag{7.26}$$

Since p is defined as a function of $\dot{q}$ and q (and hence of y) by (7.14), it is clear that (7.26) defines a field of 1-forms $y \mapsto \varpi$ on $\mathbf{R}^{2n+2}$ over which y varies; ϖ is called the *Cartan form.*[132] (7.27)

By introducing the *exterior derivative* $\mathrm{d}\varpi$ of the Cartan form[133] one can write (7.16) in the form

$$\delta a = [\varpi(\delta y)]_{s_0}^{s_1} + \int_{s_0}^{s_1} [\mathrm{d}\varpi]\,(\delta y)\left(\frac{dy}{ds}\right) ds\,. \tag{7.28}$$

This will allow us to extend the preceding results to variational problems *defined on a manifold* after we have proved the following theorem (Fig. 7.I).

(7.29) DEFINITION/THEOREM: Let q be a variable ranging over a manifold V. Let $\dot{q}$ be a tangent vector at q to V and define $y \equiv \binom{\dot{q}}{q}$.[134] We will call a function L a *Lagrangian* if $L(q, \dot{q}) \equiv \ell$ is differentiable[135] and homogeneous of degree 1 in $\dot{q}$, that is,

$$L(q, \mu\dot{q}) = \mu L(q, \dot{q}) \qquad \forall \mu \neq 0\,.$$

The 1-form ϖ on V^D given by

$$\varpi(\delta y) \equiv \frac{\partial \ell}{\partial \dot{q}}(\delta q)\ ^{136}$$

will be called the *Cartan form* of the Lagrangian L.

---

[132] With the notation of (7.20) and (7.21) the Cartan form is defined by $\varpi(\delta y) \equiv \left[\sum_{j=1}^{n} p_j\, \delta^j q\right] - h\,\delta t$.

[133] Definition (4.32) immediately gives $[\mathrm{d}\varpi]\,(\delta y)(\delta' y) = [\delta p]\,(\delta' q) - [\delta' p]\,(\delta q)$.

[134] The variable y ranges over the manifold V^D of tangent vectors to V (see (1.31) and (1.34)).

[135] On an open subset of V^D.

[136] The reader should notice that in general the notation $\dfrac{\partial \ell}{\partial q}$ makes *no sense whatsoever* on a manifold V.

In order that a map $s \mapsto q$ $(\dfrac{dq}{ds} \neq 0)$ be an *extremum*, it is necessary and sufficient that the map

$$s \mapsto y \equiv \begin{pmatrix} \dfrac{dq}{ds} \\ q \end{pmatrix}$$

satisfies identically

$$\frac{dy}{ds} \in \ker(\mathrm{d}\varpi) . \qquad \diamondsuit \qquad \square$$

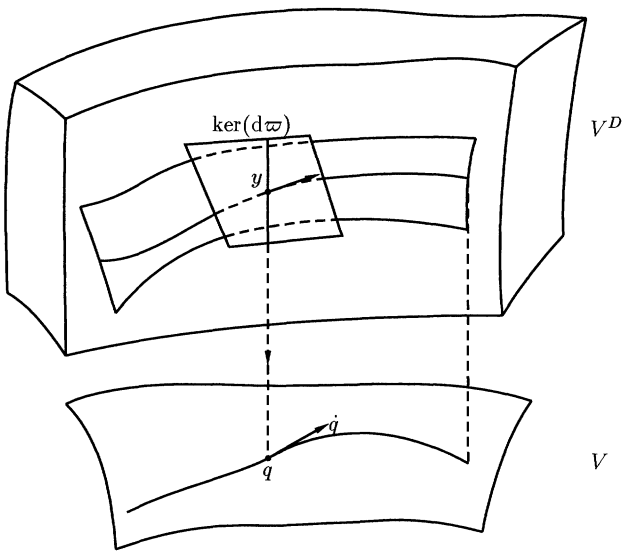


Fig 7.I

(7.30) If we assume that the vector space $\ker \mathrm{d}\varpi$ has constant dimension k, then the field $y \mapsto \ker \mathrm{d}\varpi$ is the *characteristic foliation* of the 2-form $\mathrm{d}\varpi$.[137] We will consider the leaves of this foliation to be (generalized) solutions of the variational problem defined by the Lagrangian.

(7.31) In particular, if the problem is *regular* (see (7.18)) k is equal to 2. In such a case each extremum on V is the projection of an integral manifold of the foliation.

[137]Because $\mathrm{d}[\mathrm{d}\varphi] = 0$ by Poincaré's lemma (see definition (5.24) and (5.25)).

Transformations of a variational problem

(7.32) Let $L(q)(\dot{q})$ be a Lagrangian on a manifold V (see (7.29)). It is clear that $L(q)$ is a covariant operator of degree 1, and thus that the Lagrangian L is a *covariant field* of degree 1 (see (4.3)).

It follows that we know how to define (Fig. 7.II) the inverse image L^* of L under a differentiable map A (4.5). By writing

$$(7.33) \qquad y^* \equiv \begin{pmatrix} \dot{q}^* \\ q^* \end{pmatrix}, \quad y \equiv \begin{pmatrix} \dot{q} \\ q \end{pmatrix} \equiv A^D(y^*) \equiv \begin{pmatrix} D(A)(q^*)(\dot{q}^*) \\ A(q^*) \end{pmatrix}$$

we find

$$(7.34) \qquad L^*(q^*)(\dot{q}^*) \equiv L(q)(\dot{q}),$$

which shows that L^* is *also a Lagrangian*. One can immediately check that the Cartan form ϖ^* of L^* is the inverse image under A^D of the Cartan form ϖ of L, namely

$$(7.35) \qquad \varpi^*(\delta y^*) \equiv \varpi(\delta y).$$

Theorem (4.36) then shows that the inverse image of $\mathrm{d}\varpi$ is equal to $\mathrm{d}\varpi^*$. Formulation (7.29$\diamondsuit$) of the Euler-Lagrange equations thus shows that the inverse image under A of an extremum of L^* is an extremum of L (even if the manifolds V^* and V do not have the same dimension).

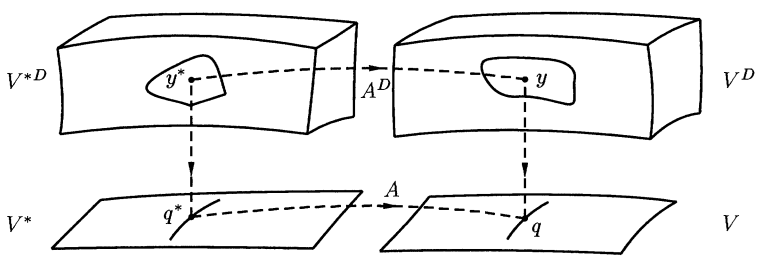


Fig 7.II

Noether's theorem

(7.36) THEOREM: Let L be a Lagrangian on a manifold V (notation (7.29)), and let G be a Lie group that acts on V (6.4). We will say that the group leaves the Lagrangian L *invariant* if the inverse image of L under the maps $\underline{a}_V$ is again L. Under these conditions, for every Z in the Lie

algebra $\mathcal{G}$ of G, the number $u \equiv \dfrac{\partial \ell}{\partial \dot{q}}(Z_V(q))$ is constant on each connected extremum.[138] □

In the next chapter we will study this theorem in detail within a more general framework (11.11). Let us only indicate how it can be applied to a classical Lagrangian. If we define $\delta q \equiv Z_V(q)$, then formulæ (7.20) and (7.21) give

$$u \equiv p_1 \, \delta \, {}^1q + \cdots + p_n \, \delta \, {}^nq - h \, \delta t \, . \tag{7.37}$$

It is this quantity u which is constant if the group leaves the Lagrangian invariant. The reader might verify that (7.25) (the constancy of the Hamiltonian h if the Lagrangian does not contain the variable t) is a special case of Noether's theorem.

[138] Recall that Z_V is the vector field on V associated to Z by formula (6.11). Incidently, one also has $u \equiv \varpi(Z_{V\,D}(y))$ (see (6.7)).

Chapter II

Symplectic Geometry

§8. 2-Forms

Orthogonality

(8.0) Let E be a real vector space of dimension n and let σ be a 2-form on E (see (4.26)).

(8.1) Two vectors X and Y of E are said to be *orthogonal* (with respect to σ) if the number $\sigma(X)(Y)$ is zero. This orthogonality relation is *symmetric, because the form σ is anti-symmetric* ($\sigma(Y)(X) \equiv -\sigma(X)(Y)$).

We will denote by $\mathrm{orth}(H)$ the set of vectors in E orthogonal to a vector subspace H of E:

$$(8.2) \qquad [X \in \mathrm{orth}(H)] \iff [\sigma(X)(Y) = 0 \quad \forall Y \in H]\,.$$

$\mathrm{orth}(H)$ is also a *vector subspace* of E. One can check immediately the following formulæ[139]

$$(8.3) \qquad \begin{aligned} &\text{a) } [H \subset \mathrm{orth}(H')] \iff [H' \subset \mathrm{orth}(H)] \\ &\text{b) } [H \subset H'] \implies [\mathrm{orth}(H') \subset \mathrm{orth}(H)] \\ &\text{c) } \mathrm{orth}(H + H') = \mathrm{orth}(H) \cap \mathrm{orth}(H') \\ &\text{d) } \mathrm{orth}(\mathrm{orth}(H) = H + \ker(\sigma) \\ &\text{e) } \ker(\sigma) = \mathrm{orth}(E) \\ &\text{f) } \dim(\mathrm{orth}(H)) = \dim(E) - \dim(H) + \dim(H \cap \ker(\sigma)). \end{aligned}$$ □

(8.4) We will say that a vector subspace H of E is *isotropic* (respectively, *coisotropic*) if $H \subset \mathrm{orth}(H)$ (respectively, $\mathrm{orth}(H) \subset H$). □

(8.5) For a vector subspace H to be isotropic it is necessary and sufficient that $\sigma(X)(Y) = 0\ \forall X, Y \in H$. It follows in particular that *all 1-dimensional vector subspaces are isotropic* (since $\sigma(X)(X) = 0\ \forall X \in E$).

Applying (8.3) one can verify that

$$(8.6) \qquad \begin{aligned} &\text{a)} \quad [H \text{ is isotropic}, H' \subset H] \implies [H' \text{ is isotropic}] \\ &\text{b)} \quad [H \text{ is coisotropic}, H \subset H'] \implies [H' \text{ is coisotropic}] \\ &\text{c)} \quad [H \text{ is isotropic}] \implies [\mathrm{orth(H) \text{ is coisotropic}}, \\ &\qquad\qquad H + \ker(\sigma) \text{ is isotropic}]\,. \end{aligned}$$ □

[139] where H and H' are vector subspaces of E and where the sign $+$ denotes the *sum* of the two vector subspaces: $[X \in H + H'] \iff [\exists Y \in H, \exists Z \in H',\ X = Y + Z]$.

(8.7) DEFINITION: We will say that a vector subspace H of E (notation (8.0)) is *self orthogonal* if $H = \mathrm{orth}(H)$.

(8.7) THEOREM: Every isotropic subspace of E is contained in a self-orthogonal subspace.

PROOF: Given an isotropic subspace of E, one could try to find out whether it is contained in a larger isotropic subspace, meaning in particular of larger dimension. Repeating this process (at most n times), one finds a *maximal isotropic subspace* H containing the given subspace.

Let V be an element of $\mathrm{orth}(H)$ and let us denote by $\{V\}$ the vector space spanned by V. We know that it is isotropic (8.5) and thus that $\{V\} \subset \mathrm{orth}(\{V\})$. Since $\{V\} \subset \mathrm{orth}(H)$, we thus have $H \subset \mathrm{orth}(\{V\})$ (8.3a). It follows that $H + \{V\} \subset \mathrm{orth}(H) \cap \mathrm{orth}(\{V\})$. But $\mathrm{orth}(H) \cap \mathrm{orth}(\{V\}) = \mathrm{orth}(H + \{V\})$ (see (8.3c)) and thus $H + \{V\}$ is isotropic. Since H is a maximal isotropic subspace, $H + \{V\} = H$ and hence $V \in H$. As V is an arbitrary element of $\mathrm{orth}(H)$, we have shown that $\mathrm{orth}(H) \subset H$. By hypothesis, $H \subset \mathrm{orth}(H)$ and thus H is self-orthogonal. □

Canonical bases

(8.8) THEOREM: Let H be a self-orthogonal subspace of E (notation (8.0) and (8.7)). Then $\ker(\sigma) \subset H$. If T is a basis[140] of $\ker(\sigma)$ and if one completes T to a basis $[T\,U]$ of H, then there exists a basis $S = [T\,U\,V]$ of E such that the matrix of components of σ with respect to S is given by

$$J = \left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \\ \hline 0 & \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & 0 \end{array}\right] \begin{matrix} \} q \\ \} p \\ \} p \end{matrix}$$

$$\underbrace{\qquad}_{q} \quad \underbrace{\qquad}_{p} \quad \underbrace{\qquad}_{p}$$

where $q = \dim(\ker(\sigma))$, $q + p = \dim(H)$, and $q + 2p = n$. We will say

[140] Notation (1.4).

that S is a *canonical basis*.

PROOF: Since $H \subset E$, we have $\mathrm{orth}(E) \subset \mathrm{orth}(H)$ (8.3a) or, in other words, $\ker(\sigma) \subset H$ (because $H = \mathrm{orth}(H)$). We thus can complete a basis T of $\ker(\sigma)$ to a basis $[T\,U]$ of H. Next we complete this basis to a basis $S' = [T\,U\,V']$ of E. We then have:

$$\begin{aligned} H &= \mathrm{im}(T) + \mathrm{im}(U)\,, & \mathrm{im}(T) \cap \mathrm{im}(U) &= \{0\}\,, \\ E &= H + \mathrm{im}(V')\,, & H \cap \mathrm{im}(V') &= \{0\}\,. \end{aligned}$$

As $\mathrm{im}(T) = \mathrm{orth}(E)$ and $H = \mathrm{orth}(H)$ by hypothesis, we see that

$$\begin{aligned} H = \mathrm{orth}(H) &= \mathrm{orth}(\mathrm{im}(T) + \mathrm{im}(U)) \\ &= \mathrm{orth}(\mathrm{im}(T)) \cap \mathrm{orth}(\mathrm{im}(U)) = \mathrm{orth}(\mathrm{im}(U))\,, \end{aligned}$$

whence

$$\diamondsuit \qquad \mathrm{orth}(\mathrm{im}(U)) \cap \mathrm{im}(V') = \{0\}\,.$$

We also have

$$\begin{aligned} \mathrm{im}(U) \cap \mathrm{orth}(\mathrm{im}(V')) &= \mathrm{im}(U) \cap H \cap \mathrm{orth}(\mathrm{im}(V')) \\ &= \mathrm{im}(U) \cap [\mathrm{orth}(H) \cap \mathrm{orth}(\mathrm{im}(V'))] \\ &= \mathrm{im}(U) \cap \mathrm{orth}(H + \mathrm{im}(V')) \\ &= \mathrm{im}(U) \cap \mathrm{orth}(E) = \mathrm{im}(U) \cap \mathrm{im}(T)\,, \end{aligned}$$

or in other words

$$\heartsuit \qquad \mathrm{im}(U) \cap \mathrm{orth}(\mathrm{im}(V')) = \{0\}\,.$$

Formula $\diamondsuit$ says that the matrix M with entries ${}^{j}M_{k} = \sigma(U_j)(V'_k)$ is injective.[141]

Likewise, formula $\heartsuit$ shows that the matrix M' with entries ${}^{k}M'_{j} = \sigma(V'_k)(U_j)$ is also injective. Moreover, ${}^{k}M'_{j} = -\,{}^{j}M_{k}$, that is, $M' = -\overline{M}$ (6.51m). Hence the matrix M, which is injective as well as its transpose $\overline{M}$, is necessarily square. And thus the spaces $\mathrm{im}(V')$ and $\mathrm{im}(U)$ have the same dimension, say p.

Inverting M we find a square matrix M^{-1}. If we define $V'' = V' \cdot M^{-1}$, then V'' is a new basis of $\mathrm{im}(V')$ and by construction we have

$$\sigma(U_j)(V''_k) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k\,. \end{cases}$$

[141] Because: $[M(x) = 0] \Longrightarrow [\sigma(U_j)(\sum_k V'_k\,{}^{k}x) = 0\ \forall j] \Longrightarrow [V'(x) \in \mathrm{orth}(\mathrm{im}(U))] \Longrightarrow [V'(x) \in \mathrm{orth}(\mathrm{im}(U)) \cap \mathrm{im}(V')] \Longrightarrow [V'(x) = 0] \Longrightarrow [x = 0]$.

If we define

$$\clubsuit \qquad V_k = V_k'' + \tfrac{1}{2} \sum_j U_j \, \sigma(V_j'')(V_k'') \qquad [k = 1, 2, \dots, p],$$

we find immediately

$$\sigma(T_m)(T_{m'}) = 0, \quad \sigma(T_m)(U_j) = 0, \quad \sigma(T_m)(V_k) = 0,$$

$$\sigma(U_j)(U_{j'}) = 0, \quad \sigma(U_j)(V_k) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases},$$

$$\sigma(V_k)(V_{k'}) = 0\,.$$

These results show that if we define

$$S = [T_1 \dots T_q \, U_1 \dots U_p \, V_1 \dots V_p]$$

then S is a *basis* of E.[142] They also show that the components of σ with respect to the basis S, that is, the numbers $\sigma(S_\ell)(S_{\ell'})$,[143] are the entries of the matrix J in the statement of the theorem. □

(8.9) COROLLARY: The rank of a 2-form is an even number. If σ is a 2-form of rank $2p$ on a vector space E of dimension n, then all self-orthogonal subspaces of E with respect to σ have *the same dimension*, namely, $n-p$.

PROOF: The rank of σ equals $\dim(E) - \dim(\ker(\sigma))$, that is, using the preceding notation, $n - q\,(= 2p)$. One sees that the dimension of H, which equals $q + p\,(= n - p)$, does not depend on the choice of H but only on the form σ. □

(8.10) DEFINITION: A real vector space E on which is defined an *injective* 2-form is called a *symplectic vector space*.

(8.10) THEOREM: If E is symplectic, its dimension is *even* $(= 2p)$. Moreover, E possesses a basis S, called a *canonical basis*, such that[144]

$$\diamondsuit \qquad \sigma(S(x))(S(y)) = \overline{x} \cdot J \cdot y \qquad \forall x, y \in \mathbf{R}^{2p}$$

with

$$\heartsuit \qquad J = \left[\begin{array}{c|c} \begin{matrix} & & 0 \\ & \iddots & \\ -1 & & \end{matrix} & \begin{matrix} 1 & & \\ & \iddots & \\ & & 1 \end{matrix} \\ \hline \begin{matrix} & & -1 \\ & \iddots & \\ 0 & & \end{matrix} & \end{array}\right] \begin{matrix} \} p \\ \} p \end{matrix}$$

[142] The formulæ ♣ show that S is the product of S' and an injective triangular matrix.

[143] See definition (4.20).

[144] Notation (6.52) applied to the Euclidean space $\mathbf{R}^{2p}$.

Every self-orthogonal† subspace of E has dimension p and every basis of a Lagrangian subspace can be completed to a canonical basis of E.

PROOF: This is a special case of (8.8). It suffices to note that σ is injective if and only if the dimension of its kernel is 0. □

In the case of a *symplectic* vector space E, one can verify immediately that formulæ (8.3,4,5,6) can be strengthened to

(8.11)
$$\begin{aligned} &\text{a)} \quad \operatorname{orth}(\operatorname{orth}(H)) = H, \\ &\text{b)} \quad \operatorname{orth}(H \cap H') = \operatorname{orth}(H) + \operatorname{orth}(H'), \\ &\text{c)} \quad [H \text{ isotropic}] \iff [\operatorname{orth}(H) \text{ is coisotropic}], \\ &\text{d)} \quad \dim(H) + \dim(\operatorname{orth}(H)) = \dim E. \end{aligned}$$ □

The symplectic group

(8.12) DEFINITION: Let E be a symplectic vector space of dimension $2p$. The group $\operatorname{Sp}(E)$ of elements A of $\operatorname{Gl}(E)$ such that

$$\diamondsuit \qquad \sigma(A(X))(A(Y)) = \sigma(X)(Y) \qquad \forall X, Y \in E$$

is called the *symplectic group*.[145]

(8.13) THEOREM: Let S be a canonical basis of a symplectic vector space E and let A be an element of $\operatorname{Gl}(E)$. Then

$$[A \in \operatorname{Sp}(E)] \iff [A \cdot S = \text{canonical basis}].$$

PROOF: It suffices to compare (8.10◇) with (8.12◇). □

(8.14) THEOREM: Let U and U' be bases of two *isotropic subspaces* of the same dimension of a symplectic vector space. Then there exists an element A of the symplectic group such that $U' = A \cdot U$.

PROOF: One can complete U' and U to bases of Lagrangian subspaces (theorem 8.7). Then one can complete these bases to canonical bases S' and S (8.10). It then suffices to take $A = S' \cdot S^{-1}$ (8.13). □

†Editors' note: A self-orthogonal subspace of half the dimension of a symplectic vector space is called a *Lagrangian* subspace.

[145]One sometimes defines other "symplectic groups" which are subgroups of this one. We will not encounter them.

(8.15) COROLLARY: If U and U' are *nonzero* vectors of a symplectic vector space E, then there exists an element A of $\mathrm{Sp}(E)$ such that $U' = A \cdot U$.

PROOF: U' and U can be considered as bases of subspaces of dimension 1, which are isotropic (8.5). □

In other words, $\mathrm{Sp}(E)$ acting on E has two orbits: $\{0\}$ on the one hand and $E \setminus \{0\}$ on the other hand.

Let us represent an element A of $\mathrm{Sp}(E)$ in a canonical basis S by

$$A = S \cdot M \cdot S^{-1} . \tag{8.16}$$

Then (8.12◇) shows that $\sigma(S \cdot M \cdot x)(S \cdot M \cdot y) = \sigma(S \cdot x)(S \cdot y)$ for all $x, y \in \mathbf{R}^{2p}$, which, using (8.10◇), is equivalent to $\overline{M \cdot x} \cdot J \cdot M \cdot y = \overline{x} \cdot J \cdot y$. From this we get

$$\overline{M} \cdot J \cdot M = J . \tag{8.17}$$

Since (8.10♡) shows that

$$\overline{J} = J^{-1} = -J , \tag{8.18}$$

equation (8.17) can be written as

$$M^{-1} = J \cdot \overline{M} \cdot J^{-1} .^{146} \tag{8.19}$$

Matrices M satisfying this condition (which, one should note, is independent of the choice of the canonical basis S) are called *symplectic matrices*. Obviously, they form a *group isomorphic to* $\mathrm{Sp}(E)$.

(8.20) THEOREM: The set of symplectic matrices is a *Lie group* having as its Lie algebra the set of matrices Z such that

$$\diamondsuit \qquad -Z = J \cdot \overline{Z} \cdot J^{-1} .$$

PROOF: Let $\tilde{G}$ be the set of symplectic matrices and $\tilde{\mathcal{G}}$ the set of matrices satisfying (8.20◇). One can verify immediately that they satisfy condition (6.33a). Hence $\tilde{G}$ is a Lie subgroup of $\mathrm{Gl}(\mathbf{R}^{2p})$. □

(8.21) Expanding the Lie algebra equation (8.20◇), one can see that it can be written as

$$Z = \begin{bmatrix} A & C \\ B & -\overline{A} \end{bmatrix} \qquad B = \overline{B} , \; C = \overline{C} ,$$

[146] Compare with (6.74).

A, B, C being real matrices of order p. Hence the dimension of the group $\tilde{G}$ is $p^2 + \frac{1}{2}p(p+1) + \frac{1}{2}p(p+1) = p(2p+1)$. It follows that $\mathrm{Sp}(E)$ is a Lie subgroup of $\mathrm{Gl}(E)$ of dimension $p(2p+1)$.

(8.22) THEOREM: For $A \in \mathrm{Sp}(E)$, if $\mathrm{Log}(A)$ exists, then $\mathrm{Log}(A)$ belongs to the Lie algebra of $\mathrm{Sp}(E)$.

PROOF: It obviously suffices to verify this for symplectic matrices. The result follows by taking the logarithm of (8.19) and applying formulæ (6.39) and (6.40). □

Due to (8.18) one can see that $\mathbf{R}^{2p}$ possesses the structure of a *positive complex Euclidean space* of complex dimension p, obtained by adjoining, to its Euclidean structure (6.52), the complex structure defined by

$$[a + bi]\, x = ax + bJ(x) \qquad \forall a, b \in \mathbf{R}\,,\ \forall x \in \mathbf{R}^{2p}\,. \tag{8.23}$$

The elements of the associated *unitary* group (6.66) are the real matrices A of order $2p$ which are orthogonal and commute with J (see (6.43)):

$$\overline{A} = A^{-1}\,, \qquad A \cdot J = J \cdot A\,, \tag{8.24}$$

which can also be written as

$$\overline{A} = A^{-1}\,, \qquad A^{-1} = J \cdot \overline{A} \cdot J^{-1}\,. \tag{8.25}$$

It follows that $\mathrm{U}(\mathbf{R}^{2p})$ consists of *symplectic matrices* and one can easily verify that $\mathrm{U}(\mathbf{R}^{2p})$ is a *Lie subgroup* of the group of symplectic matrices. Thus:

(8.26) The group $\mathrm{Sp}(E)$ (notation (8.12)) possesses a Lie subgroup isomorphic to $\mathrm{U}(p)$. □

Let us apply *Cartan's decomposition* (6.70) to a symplectic matrix M:

$$M = A \cdot \exp(B) \qquad \overline{A} = A^{-1}\,,\ \overline{B} = B\,. \tag{8.27}$$

We know that $B = \frac{1}{2}\mathrm{Log}(\overline{M} \cdot M)$. Using (8.19) it follows easily that $\overline{M}$ too is a symplectic matrix, hence so is $\overline{M} \cdot M$. We deduce that $B \in \tilde{\mathcal{G}}$ (theorem (8.22)) and thus that $\exp(B)$ is a symplectic matrix (8.20). Hence $A = M \cdot (\exp(B))^{-1}$ is a symplectic matrix. Since A is orthogonal, we conclude that $A \in \mathrm{U}(\mathbf{R}^{2p})$. Moreover, expanding $B \in \tilde{\mathcal{G}}$, $B = \overline{B}$ (see (8.21)), we find

$$B = \begin{bmatrix} M & N \\ N & -M \end{bmatrix} \qquad M = \overline{M}\,,\ N = \overline{N}\,. \tag{8.28}$$

Conversely for $A \in \mathrm{U}(\mathbf{R}^{2p})$, if B belongs to the real vector space of dimension $p(p+1)$ defined by (8.28), it is immediate that $A \cdot \exp(B)$ is symplectic. Since the correspondence $M \mapsto \begin{pmatrix} A \\ B \end{pmatrix}$ is a diffeomorphism (6.70), it follows that:

(8.29) The symplectic group $\mathrm{Sp}(E)$ of a symplectic vector space E of dimension $2p$ is diffeomorphic to $\mathrm{U}(p) \times \mathbf{R}^{p(p+1)}$, and thus is connected. □

(8.30) THEOREM: The determinant of a symplectic matrix equals 1.

PROOF: In Cartan's decomposition $M = A \cdot \exp(B)$ we have

$$\begin{aligned} \det(\exp(B)) &= \exp(\mathrm{Tr}(B)) && (6.18) \\ &= 1 && (8.28) \end{aligned}$$

and

$$\begin{aligned} \det(A) &= |\det_{\mathbf{C}}(A)|^2 && (6.47) \\ &= 1\,.^{147} && (6.66) \end{aligned}$$

□

[147] One can also take the determinant of (8.19), which gives $\det(M) = \pm 1$. As $\det(M)$ is continuous on the connected manifold of symplectic matrices, $\det(M)$ cannot take both values $+1$ and -1. Thus it keeps the same value as for $M = 1$. There also exists a purely algebraic proof of this fact using exterior multiplication of forms.

§9. Symplectic manifolds

Symplectic and presymplectic manifolds

(9.1) DEFINITION: A *symplectic manifold* U is a manifold on which is defined a differentiable field of 2-forms $x \mapsto \sigma$[148] satisfying the following conditions:

$$\begin{cases} \sigma \text{ is injective}^{149} \\ \mathrm{d}\sigma \equiv 0\,.^{150} \end{cases}$$ □

(9.2) At each point x of a symplectic manifold, the 2-form σ_x gives the tangent space D_x the structure of a *symplectic vector space* (8.10). *Thus the dimension of a symplectic manifold is even.*

For future reference we recall that the condition $\mathrm{d}\sigma \equiv 0$ can be written in a chart as

(9.3) $$\partial_j \sigma_{k\ell} + \partial_k \sigma_{\ell j} + \partial_\ell \sigma_{jk} \equiv 0$$

(formula (4.35)).

(9.4) EXAMPLE: Let E be a *symplectic vector space* (8.10). On E we can obviously define the *constant* map $x \mapsto \sigma$ and we can interpret it as a field of injective 2-forms (since at each point x of E the tangent space is equal to E). Taking a basis of E as a chart (1.9), the components σ_{jk} of σ are constant and thus formula (9.3) shows that $\mathrm{d}\sigma = 0$. In other words, E is a *symplectic manifold.* □

(9.5) EXAMPLE: Let $V = {}^1V \times {}^2V \times \cdots \times {}^nV$ be a *direct product of symplectic*

[148] We will denote this field by $x \mapsto \sigma_U$ whenever we need to specify on which manifold we are working. σ_U will be called the *Lagrange form* on U.

[149] Let us note the analogy with Riemannian manifolds on which is defined a field $x \mapsto g$, where g is also an injective tensor of degree 2. But g is *symmetric* whereas σ is *antisymmetric*.

[150] Later on we will see that this equation implies the existence of charts in which the components of σ are *constant* (10.8). The analog of this in Riemannian geometry is the condition that the Riemann-Christoffel tensor is zero.

manifolds. A point x in V is by definition an n-tuple

$$x \equiv \begin{pmatrix} {}^1x \\ {}^2x \\ \vdots \\ {}^nx \end{pmatrix}$$

such that ${}^jx \in {}^jV$ for $j = 1, 2, \ldots, n$. Recall that V is a manifold whose dimension is the *sum of the dimensions of the* ${}^jV$, that the *tangent space to* V at x is the *direct product* of the tangent spaces to the ${}^jV$ at ${}^jx$, and that for each derivation δ we have (2.19)

$$\delta x \equiv \begin{pmatrix} \delta\, {}^1x \\ \delta\, {}^2x \\ \vdots \\ \delta\, {}^nx \end{pmatrix}.$$

For any derivations δ and δ' we now define

$$\sigma_V(\delta x)(\delta' x) \equiv \sigma_{[{}^1V]}(\delta[{}^1x])(\delta'[{}^1x]) + \cdots + \sigma_{[{}^nV]}(\delta[{}^nx])(\delta'[{}^nx])\,. \tag{9.6}$$

One can verify directly that this form σ_V is injective and that its exterior derivative is zero (it suffices to apply definition (4.32)). We thus obtain the following result.

(9.7) With the convention (9.6) every direct product of symplectic manifolds is a symplectic manifold. □

(9.8) DEFINITION/THEOREM: A *presymplectic* manifold is a manifold V on which is defined a differentiable field of 2-forms $y \mapsto \sigma$ [151] such that

$$\begin{cases} \ker(\sigma) \text{ has } \textit{constant dimension} > 0 \text{ [152]} \\ \mathrm{d}\sigma \equiv 0\,. \end{cases}$$

In that case the field $y \mapsto \ker(\sigma)$ is called the *characteristic foliation* of the 2-form σ, the leaves of which will be called the *leaves of* V. We will say that V is sectionable if this foliation is sectionable.[153]

PROOF: It suffices to apply definition (5.25) of a characteristic foliation to the field $x \mapsto \ker(\sigma) \cap \ker(\mathrm{d}\sigma)$, which here becomes $x \mapsto \ker(\sigma)$, since $\mathrm{d}\sigma \equiv 0$. □

[151] As in (9.1), σ will be called the *Lagrange form* and will be denoted by σ_V if needed.

[152] If this dimension were zero, V would be symplectic.

[153] Recall that this condition is always satisfied *locally* (5.17).

(9.9) Let us now consider a presymplectic manifold V and let x be the leaf which passes through a point y of V (Fig. 9.I). If V is sectionable (5.15), we know that the set U of leaves admits the structure of a manifold (U is called the *quotient manifold* of V by the foliation), that the map $[y \mapsto x]$ (called the *projection* of V onto U) is differentiable, that $\ker\left(\frac{\partial x}{\partial y}\right) = \ker(\sigma)$, and that $\operatorname{im}\left(\frac{\partial x}{\partial y}\right)$ is the tangent space to U at x. It follows immediately that the dimension of U is the difference of the dimensions of V and of $\ker(\sigma)$, that is, the *rank* of σ (also called the *rank* of V).

Because it is the characteristic foliation of σ, we know that σ is an *integral invariant*, and thus σ_V is the inverse image, under the projection, of a 2-form σ_U on U. Using derivations δ and δ' this can be written as

$$\sigma_V(\delta y)(\delta' y) \equiv \sigma_U(\delta x)(\delta' x)\,.$$

Let us now look at the kernel of σ_U. We immediately find that

$$\begin{aligned} [\delta x \in \ker(\sigma_U)] &\iff [\sigma_V(\delta y)(\delta' y) = 0 \qquad\quad \forall \delta' y] \\ &\iff [\delta y \in \ker(\sigma_V)] \iff \left[\delta y \in \ker\left(\frac{\partial x}{\partial y}\right)\right] \\ &\iff [\delta x = 0]\,. \end{aligned}$$

Thus σ_U is *injective*.

Moreover, since the inverse image of $\mathrm{d}\sigma_U$ under the projection equals $\mathrm{d}\sigma_V$ (theorem (4.36)), which is *zero*, it follows immediately that $\mathrm{d}\sigma_U = 0$. We thus have proved theorem (9.10).

(9.10) THEOREM: Let V be a sectionable presymplectic manifold and let U be the set of leaves of V (Fig. 9.I). Then U is a *symplectic manifold* and the projection $y \mapsto x$, which sends a point y of V to the leaf x of V containing y, is a differentiable map from V onto U such that

$$\diamondsuit \qquad \ker\left(\frac{\partial x}{\partial y}\right) = \ker(\sigma_V)\,, \qquad\qquad \operatorname{im}\left(\frac{\partial x}{\partial y}\right) = D_x\,.$$

The inverse image of the Lagrange form on U under the projection is the Lagrange form on V:

$$\heartsuit \qquad \sigma_V(\delta y)(\delta' y) = \sigma_U(\delta x)(\delta' x) \qquad\quad \forall \delta y\,,\ \delta' y\,. \qquad \square$$

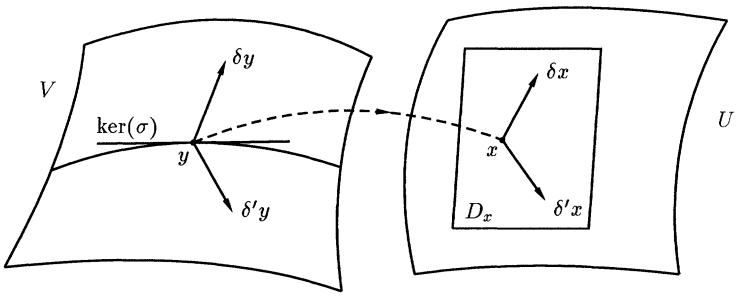


Fig 9.I

Symplectic structures arising from a 1-form

(9.11) Let $y \mapsto \varpi$ be a differentiable field of 1-forms on a manifold V, and let us define $\sigma = \mathrm{d}\varpi$. By Poincaré's lemma (4.37) we know that $\mathrm{d}\sigma \equiv 0$. If $\ker(\sigma)$ has constant dimension, V is symplectic or presymplectic. If V is presymplectic and sectionable, the manifold of leaves is symplectic.

(9.12) EXAMPLE: Let E be a vector space of dimension p and let $V = E^* \times E$. A variable point of V will be denoted by $y \equiv \begin{pmatrix} P \\ Q \end{pmatrix}$, $P \in E^*$, $Q \in E$. If we define $\varpi(\delta y) \equiv P(\delta Q)$, then the 2-form σ is given by

$$\sigma(\delta y)(\delta' y) = [\delta P](\delta' q) - [\delta' P](\delta Q)\,.$$

One can immediately verify that σ is injective and thus that V is a symplectic manifold of dimension $2p$. Occasionally one says that V is the *phase space* of E. Note that V is a *symplectic vector space* (9.4). □

(9.13) EXAMPLE: If ϖ is the *Cartan form of a variational problem* (7.27), the form $\sigma \equiv \mathrm{d}\varpi$ is never injective. If the kernel of σ has constant dimension, we know that the leaves are the solutions of the variational problem (7.30). If this foliation is sectionable, *the set U of solutions has the structure of a symplectic manifold.* □

(9.14) On the other hand, if V is a symplectic or presymplectic manifold, it might happen that there exists a 1-form ϖ such that $\sigma \equiv \mathrm{d}\varpi$.[154] In such a case we will say that V *admits a potential* and that ϖ is a *potential* of σ. One actually encounters manifolds which admit a potential (9.12) and others which do not (18.36).

[154]The converse of Poincaré's lemma shows that this is always the case if V is a cell (4.37).

Poisson brackets

(9.15) DEFINITION: Let V be a symplectic manifold. If a scalar field $x \mapsto u$ on V is differentiable, we will say that the real variable u is a *dynamical variable*.[155] □

At each point x, the expression $\frac{\partial u}{\partial x}$, also denoted by $\mathrm{d}u$,[156] is a 1-form, that is, an element of the dual of the tangent space D_x.

The Lagrange form σ of V can be considered as a linear map from D_x to its dual D_x^*. Because this map is injective and D_x and D_x^* have the same dimension, σ maps D_x *onto* D_x^*. Thus there exists a vector, which we will call the *symplectic gradient* of u and which we will denote by $\operatorname{grad} u$,[157] such that

$$-\mathrm{d}u \equiv \sigma(\operatorname{grad} u)\,. \tag{9.16}$$

This formula can also be written as

$$\operatorname{grad} u \equiv -\sigma^{-1}(\mathrm{d}u) \tag{9.17}$$

or as

$$\delta u \equiv \sigma(\delta x)(\operatorname{grad} u) \qquad \text{for any vector field } x \mapsto \delta x. \tag{9.18}$$

(9.19) It is easy to verify[158] that the vector field $x \mapsto \operatorname{grad} u$ is *differentiable* (on the open domain of definition of $x \mapsto u$).

We define a *derivation* $\underset{u}{\delta}$ *associated to the variable* u by

$$\underset{u}{\delta} x \equiv \operatorname{grad} u. \tag{9.20}$$

(9.21) THEOREM: For every dynamical variable u, the Lie derivative $[\underset{u}{\delta}]_L \sigma$ is zero.

PROOF: It is sufficient to apply Cartan's formula (4.39) and Poincaré's lemma (4.37):

$$[\underset{u}{\delta}]_L \sigma \equiv [\mathrm{d}\sigma](\underset{u}{\delta} x) + \mathrm{d}\,[\sigma(\underset{u}{\delta} x)] \equiv 0 + \mathrm{d}\,[-\mathrm{d}u] \equiv 0\,. \qquad \square$$

[155] One thus generalizes a notion used in analytical mechanics.

[156] See (4.33).

[157] By analogy with the Euclidean or Riemannian case. The minus sign is there for practical reasons.

[158] By taking a chart of V and looking at the components of $\operatorname{grad} u$ in this chart (see (2.41)).

(9.22) DEFINITION: Let u and v be two dynamical variables. The *Poisson bracket* of u and v is the dynamical variable

$$[u,v]_P \equiv \sigma(\operatorname{grad} u, \operatorname{grad} v)\,. \qquad \square$$

Using the rules already known to the reader, one can establish the following formulæ in which u, v, and w are arbitrary dynamical variables:

(9.23)
a) $[u,v]_P \equiv -[\mathrm{d}u](\operatorname{grad} v) \equiv [\mathrm{d}v](\operatorname{grad} u) \equiv \underset{u}{\delta} v \equiv -\underset{v}{\delta} u$
b) $[u,v]_P + [v,u]_P \equiv 0$
c) $[u, v+w]_P \equiv [u,v]_P + [u,w]_P$
d) $[u, vw]_P \equiv [u,v]_P \times w + v \times [u,w]_P$
e) $[u, f(v)]_P = f'(v)[u,v]_P$ (f any differentiable function)
f) $v = \text{constant} \implies [u,v]_P \equiv 0\,.$

By applying the Lie derivative $[\underset{u}{\delta}]_L$ to the identity $\mathrm{d}v \equiv -\sigma(\underset{v}{\delta} x)$, and taking into account (4.40), (4.18), (9.23a), (4.24), (9.21), and (4.17), one finds

(9.24) $$\operatorname{grad}[u,v]_P \equiv [\underset{u}{\delta}, \underset{v}{\delta}]_L\, x\,,$$

which can also be written as

(9.25) $$\underset{[u,v]_P}{\delta} = [\underset{u}{\delta}, \underset{v}{\delta}]_L\,.$$

We have used the subscripts P and L to clearly distinguish Poisson brackets from Lie brackets.

If we apply the derivation (9.25) to a third dynamical variable w, we immediately find (see (2.45))

$$\underset{[u,v]_P}{\delta} w = \underset{u}{\delta}[\underset{v}{\delta} w] - \underset{v}{\delta}[\underset{u}{\delta} w]\,,$$

which, due to (9.19), may be written as

$$[[u,v],w] = [u,[v,w]] - [v,[u,w]]$$

or, using antisymmetry (9.23b), as

(9.26) $$[u,[v,w]] + [v,[w,u]] + [w,[u,v]] \equiv 0\,.$$

In these last two formulæ the subscript P is implicit.

(9.27) The identity (9.26), discovered by Jacobi, and formulæ (9.23) show that the set $\mathcal{G}$ of dynamical variables defined on an open subset of V satisfy all the axioms of a *Lie algebra* (6.14) except one: the dimension of $\mathcal{G}$ is *infinite*.

Induced symplectic structures

(9.28) Let V be a symplectic or presymplectic manifold and let V' be a submanifold of V (Fig. 9.II). Furthermore, let y be a variable point of V' and let H be the tangent space to V' at y.

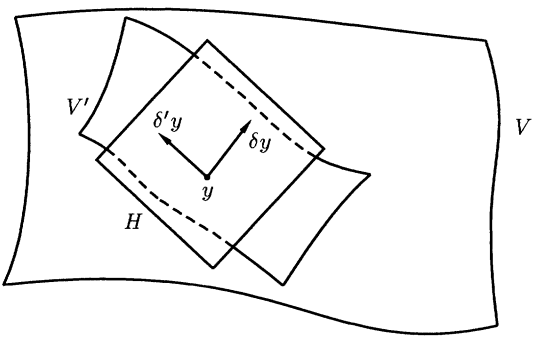


Fig 9.II

On V' we define the 2-*form* $\sigma_{V'}$ *induced* by the Lagrange form σ_V on V by

(9.29) $$\sigma_{V'}(\delta y)(\delta' y) = \sigma_V(\delta y)(\delta' y) \qquad \text{if } \delta y, \delta' y \in H\,.$$

One can verify immediately that $\mathrm{d}\sigma_{V'} \equiv 0$.[159] Thus V' is a symplectic or presymplectic manifold provided $\ker(\sigma_{V'})$ has constant dimension. Note in passing that

(9.30) $$\ker(\sigma_{V'}) \equiv H \cap \mathrm{orth}(H)\,.$$

(9.31) EXAMPLE: In particular, if V is (pre)symplectic and if V' is *open* in V (see (1.39)), then V' is a *(pre)symplectic manifold of the same dimension as* V. □

(9.32) EXAMPLE: If V is presymplectic and if V' is *transverse* to the foliation of V (5.14), then V' is symplectic. □

[159] One can prove this without calculation by observing that $\sigma_{V'}$ is the inverse image of σ_V under the map $1_{V'}$ (seen as a differentiable map from V' to V). It then suffices to apply theorem (4.36).

(9.33) DEFINITION: We will say that the submanifold V' is *isotropic*, respectively *coisotropic*, if H is isotropic, respectively coisotropic, at each point y of V'. □

(9.34) V' is isotropic if and only if $\sigma_{V'} \equiv 0$.

(9.35) If V' is *coisotropic*, then $\ker(\sigma_{V'}) \equiv H \cap \mathrm{orth}(H) \equiv \mathrm{orth}(H)$ (since by hypothesis $\mathrm{orth}(H) \subset H$). As $\ker(\sigma_V) \subset \mathrm{orth}(H)$, formula (8.3f) shows that

$$\diamondsuit \qquad \dim(\ker(\sigma_{V'})) = \dim(V) - \dim(V') + \dim(\ker(\sigma_V)) .$$

It follows that $\dim(\ker(\sigma_{V'}))$ is constant and thus V' is presymplectic or possibly even symplectic.

As $\ker(\sigma_{V'})$ is contained in H, its dimension is at most the dimension of V'. Hence the inequality

$$\dim(V') \geq \tfrac{1}{2} \left[\dim(V) + \dim(\ker(\sigma_V))\right] \tag{9.36}$$

is valid whenever V' is coisotropic. Note that the right hand side of (9.36) equals $\dim(V) - \frac{1}{2}\mathrm{rank}(\sigma_V)$, which is an integer (theorem (8.9)).

(9.37) THEOREM: Let V be a symplectic manifold of dimension $2p$ and let $x \mapsto u \equiv \begin{bmatrix} {}^1u \\ \vdots \\ {}^qu \end{bmatrix}$ be a differentiable map from V to $\mathbf{R}^q$. Suppose that the equation $u =$ constant defines a submanifold V' of dimension $2p - q$ (theorem (1.40)). Then V' is *coisotropic* if and only if the ${}^ju$s have Poisson brackets that are pairwise zero on V': $[{}^ju, {}^ku]_P \equiv 0 \ \forall j, k$. In such a case one says that the dynamical variables ${}^ju$ are *in involution* (on V').

PROOF: At a point x of V the tangent space to V' is $H = \ker\left(\frac{\partial u}{\partial x}\right)$ (theorem 1.40). A tangent vector δx at x belongs to $\mathrm{orth}(H)$ if $[\delta' x \in H] \Longrightarrow [\sigma(\delta x)(\delta' x) = 0]$, that is, if $\ker\left(\frac{\partial u}{\partial x}\right) \subset \ker(\sigma(\delta x))$. By a well known theorem of linear algebra this relation is equivalent to the existence of a "Lagrange multiplier" $\lambda = [\lambda_1 \ldots \lambda_q]$ such that $\sigma(\delta x) = \lambda \cdot \frac{\partial u}{\partial x}$. Using the notation of (9.16) this can be written as $x = -\{\lambda_1 \,\mathrm{grad}[{}^1u] + \cdots + \lambda_q \,\mathrm{grad}\,[{}^qu]\}$. This tells us that the q vectors $\mathrm{grad}\,{}^ju$ generate the vector space $\mathrm{orth}(H)$, whose dimension is $2p - (2p - q) = q$ (theorem 8.11d). Hence they constitute a *basis* of $\mathrm{orth}(H)$. We conclude that $\mathrm{orth}(H) \subset H$ holds if the $\mathrm{grad}\,{}^ju$ lie in $H = \ker\left(\frac{\partial u}{\partial x}\right)$, that is, if $\frac{\partial\, {}^ku}{\partial x}(\mathrm{grad}\,{}^ju) = 0$, which can also be written as $[{}^ju, {}^ku]_P = 0$ (9.23a). □

(9.38) If the dynamical variables are in involution, the inequality (9.36) shows

that $q \leq p$. If $q = p$, formula (9.35$\diamondsuit$) shows that V' is simultaneously isotropic and coisotropic, that is, *Lagrangian* (definition (8.7)).

(9.39) EXAMPLE: Let E be a real vector space of dimension p, let $Q \mapsto \alpha$ be a differentiable scalar field on E, and define $P \equiv \dfrac{\partial \alpha}{\partial Q}$, $y \equiv \begin{pmatrix} P \\ Q \end{pmatrix}$. It is clear that $Q \mapsto y$ is an *injective immersion* of E into $V = E^* \times E$, and thus y varies over a submanifold V' of dimension p of V (1.38). But V is a symplectic manifold of dimension $2p$ with symplectic form

$$\sigma_V(\delta y)(\delta' y) \equiv \delta P(\delta' Q) - \delta' P(\delta Q)$$

(see (9.12)). One immediately deduces that

$$\sigma_{V'}(\delta y)(\delta' y) \equiv \mathrm{d}[\mathrm{d}\alpha](\delta Q)(\delta' Q) \equiv 0\,,$$

which shows that V' is *isotropic*, that is, $H \subset \mathrm{orth}(H)$. Using (8.11d) it follows that H and $\mathrm{orth}(H)$ have the same dimension p and hence that $H = \mathrm{orth}(H)$, that is, H is Lagrangian.

We thus can associate a Lagrangian submanifold of $E \times E^*$ to any differentiable function α on E. □

APPLICATION: The above result shows that we have solved a *first order partial differential equation* of the form

(9.40) $$\varphi\left(\frac{\partial \alpha}{\partial Q}, Q\right) = 0$$ [160]

if we can find all *Lagrangian submanifolds* of the manifold V'' defined by the equation $\varphi(P, Q) = 0$; V'' is *coisotropic* (9.37). These Lagrangian submanifolds are called the *characteristics* of the equation. □

[160] This is the case for *Jacobi's equation* of a variational problem (7.18). Note that the equations of the extremals written in (7.19) are simply the equations of the leaves of the presymplectic manifold V'' (which is sometimes called *Jacobi's manifold*).

§10. Canonical transformations

Canonical charts

(10.1) DEFINITION: Let $F\ [\xi \mapsto x]$ be a chart of a symplectic manifold V. F is a called a *canonical chart* if for every ξ the basis $S = \frac{\partial x}{\partial \xi}$ [161] of the tangent space D_x is *canonical* (8.10). □

(10.2) An equivalent definition is that the components σ_{jk} of the tensor σ in the chart F [162] are constant and equal to the entries ${}^{j}[J]_{k}$ of the matrix (8.10♡):

$$J = \left[\begin{array}{cc|cc} & & 1 & \\ & 0 & & \ddots \\ & & & 1 \\ \hline -1 & & & \\ & \ddots & & 0 \\ & -1 & & \end{array}\right].$$

(10.3) EXAMPLE: Let E be a vector space of dimension n and let T be a basis of E. Let us write a vector Q in E and a vector P in E^* as

$$Q \equiv T \cdot \begin{bmatrix} {}^{1}q \\ \vdots \\ {}^{n}q \end{bmatrix} \qquad\qquad P \equiv [p_1 \ \dots \ p_n] \cdot T^{-1}.$$

It then is easy to check that the map

$$\begin{pmatrix} p_1 \\ \vdots \\ p_n \\ {}^{1}q \\ \vdots \\ {}^{n}q \end{pmatrix} \mapsto \begin{pmatrix} P \\ Q \end{pmatrix}$$

is a canonical chart of the symplectic manifold $E^* \times E$ (9.12). □

[161] See (2.24).
[162] See (4.20).

By analogy with this example, it is standard to write

$$\xi \equiv \begin{pmatrix} p_1 \\ \vdots \\ p_n \\ {}^1q \\ \vdots \\ {}^nq \end{pmatrix} \tag{10.4}$$

for an arbitrary canonical chart $\xi \mapsto x$ of a symplectic manifold V. The variables $p_j, {}^jq$ are then called *canonical coordinates* of V. When this is the case, one can easily verify the formulæ

$$\sigma(\delta x)(\delta' x) \equiv \sum_{j=1}^{n} \delta[p_j]\,\delta'[{}^jq] - \delta'[p_j]\,\delta[{}^jq]\,, \tag{10.5}$$

$$\underset{u}{\delta}[{}^jq] \equiv \frac{\partial u}{\partial[p_j]} \qquad\qquad \underset{u}{\delta}[p_j] \equiv -\frac{\partial u}{\partial[{}^jq]}\,, \tag{10.6}$$

$\underset{u}{\delta}$ being the derivation associated with the dynamical variable u (9.20), and

$$[u, v]_P = \sum_j \frac{\partial u}{\partial[p_j]}\frac{\partial v}{\partial[{}^jq]} - \frac{\partial u}{\partial[{}^jq]}\frac{\partial v}{\partial[p_j]}\,. \tag{10.7}$$

(10.8) DARBOUX'S THEOREM: Let V be a symplectic manifold. Then the *canonical charts* of V form an *atlas* of V.

PROOF: Let x_0 be a point of V. Poincaré's lemma states that there exists a field of 1-forms $x \mapsto \varpi$, differentiable in an open set containing x_0 (for example, a *cell* (1.47)), such that $\sigma \equiv \mathrm{d}\varpi$. Let us choose a chart $\Phi\,[\xi \mapsto x]$ of V such that $\Phi(0) = x_0$. Then we have (4.35)

$$\diamondsuit \qquad\qquad \sigma_{jk} \equiv \partial_j \varpi_k - \partial_k \varpi_j\,.$$

Of course, for each dynamical variable u, we can replace ϖ by $\varpi + \mathrm{d}u$ (that is, ϖ_j by $\varpi_j + \partial_j u$) without changing $\diamondsuit$. By taking u to be a quadratic polynomial in the coordinates ${}^j\xi$, one can arrange that the additional conditions

$$\heartsuit \qquad\qquad \varpi_j = 0 \qquad \text{and} \qquad \partial_j \varpi_k + \partial_k \varpi_j = 0$$

hold *at the origin*. If we define $\delta x \equiv 2\,\sigma^{-1}(\varpi)$, then $x \mapsto \delta x$ is a differentiable vector field (see the arguments leading to definition (9.16) of the symplectic gradient), and evidently we have

$$\clubsuit \qquad\qquad \mathrm{d}[\sigma(\delta x)] \equiv 2\,\sigma\,.$$

Expanding the conditions ♡ one can see that they may be written as[163]

$$\begin{cases} \delta\xi = 0 \\ \frac{\partial}{\partial\xi}[\delta\xi] = 1_{\mathbf{R}^{2p}} \end{cases} \qquad \text{at } \xi = 0\,.$$

One should recognize these conditions as the hypotheses of theorem (2.44). Thus there exists a diffeomorphism $\xi \mapsto \xi^*$ from $\mathbf{R}^{2p}$ to $\mathbf{R}^{2p}$ preserving 0 such that $\delta\xi^* \equiv \xi^*$.[164] Now choose a *canonical basis* S_0 of the tangent space to V at x_0 (8.10) and define M_0 to be the invertible matrix such that the basis $\frac{\partial x}{\partial \xi^*}$ (2.24) equals $S_0 \cdot M_0$ at $\xi = 0$. Finally define $\widetilde{\xi} \equiv M_0^{-1} \cdot \xi^*$. One can verify immediately that the map $\widetilde{\Phi}$ $[\widetilde{\xi} \mapsto x]$ is a chart, that $\widetilde{\Phi}(0) = x_0$, that $D(\widetilde{\Phi})(0) = S_0$, and that $\delta\widetilde{\xi} \equiv \widetilde{\xi}$.

By expanding ♣ in the chart $\widetilde{\Phi}$, we find

$$^j\widetilde{\xi}\,[\partial_k \sigma_{j\ell} + \partial_\ell \sigma_{kj}] \equiv 0\,,$$

which, taking (9.3) into account, may be written as $^j\widetilde{\xi}\,\partial^j[\sigma_{k\ell}] = 0$. This formula, valid for a ball with center at 0, shows that the components σ_{jk} are *constant* on the rays $^j\widetilde{\xi} = {}^jA\,\mathrm{e}^t$, which are solutions of the differential equation $\frac{d\widetilde{\xi}}{dt} \equiv \widetilde{\xi}$. Taking the limit $t \to -\infty$, it follows that, in the ball in question, the $\sigma_{k\ell}$ have the same value as at the origin. Hence they are equal to the entries of the matrix J (10.2) because $\frac{dx}{d\widetilde{\xi}}$ is the canonical basis S_0 at $\widetilde{\xi} = 0$. We have thus constructed a canonical chart whose image contains an arbitrary point x_0 of V. We conclude that the canonical charts form an atlas (definition (1.6)). □

(10.9) Let V be a *presymplectic* manifold (9.8). If y_0 is a point of V, we know how to construct a chart F $\left[\begin{pmatrix} t \\ z \end{pmatrix} \mapsto y\right]$ of V such that $y_0 \in \mathrm{im}(F)$ and such that $\delta z = 0 \Leftrightarrow \sigma(\delta y) = 0$ (theorem (5.8) applied to the characteristic foliation of σ).

We also know that $\mathrm{im}(F)$ is a sectionable foliated manifold (5.17), that $z \mapsto x$ is a chart of the quotient manifold U (x being the leaf which passes through y), and that U is symplectic (theorem 9.10). Darboux's theorem shows that there exists an atlas of canonical charts $z_j \mapsto x$ of U, one of which contains in its image the leaf passing through y_0. Let us denote this chart by $\zeta \mapsto x$. Thus Φ $\left[\begin{pmatrix} t \\ \zeta \end{pmatrix} \mapsto y\right]$ is a chart of V such that $y_0 \in \mathrm{im}(\Phi)$ and such that at each point $D(\Phi)\begin{pmatrix} t \\ \zeta \end{pmatrix}$ is a *canonical basis* in the sense of (8.8). The components of σ in the chart Φ are the

[163] We write $2p$ for the dimension of V.

[164] One can prove that $\xi^* \equiv \lim_{t\to\infty} \mathrm{e}^t[\exp(-tf)(\xi)]$ with $f(\xi) \equiv \delta\xi$ is a solution.

entries of the matrix J defined in (8.8). *In this context too we will say that Φ is a canonical chart.* Since the point y_0 is arbitrary, it follows that *presymplectic* manifolds also possess an atlas of canonical charts.

Canonical transformations

(10.10) DEFINITION: Let V and V' be two symplectic (or presymplectic) manifolds. A *canonical transformation* [165] from V to V' is any *diffeomorphism* A such that

$$\begin{cases} \mathrm{dom}(A) \subseteq V, \quad \mathrm{im}(A) \subseteq V' \\ \text{the inverse image under } A \text{ of the Lagrange form on } V' \\ \text{equals the Lagrange form on } V.^{166} \end{cases}$$

If there exists a canonical transformation A such that $\mathrm{dom}(A) = V$ and $\mathrm{im}(A) = V'$, we will say that V and V' are *symplectomorphic* and that A is a *symplectomorphism* from V to V'. □

EXAMPLE: Let E be a *symplectic vector space* and let A be a diffeomorphism from E to E. For A to be canonical we deduce from (4.5) the condition

$$\sigma(D(A)(x)(\delta x))(D(A)(x)(\delta' x)) \equiv \sigma(\delta x)(\delta' x)\,.$$

Using the notation of (8.12), the above equation can be written as

(10.11) $$D(A)(x) \in \mathrm{Sp}(E) \qquad \forall x \in \mathrm{dom}(A)\,.$$

In particular, if we define

(10.12) $$A(x) \equiv B(x) + C \qquad [B \in \mathrm{Sp}(E)\,,\ C \in E]\,,$$

then A is a symplectomorphism from E to E.[167] □

EXAMPLE: For each positive integer p, $\mathbf{R}^{2p}$ can be seen as a symplectic vector space by defining

(10.13) $$\sigma(\delta\xi)(\delta'\xi) \equiv \sum_{j,k} \sigma_{jk}\, \delta[^j\xi]\, \delta'[^k\xi] \qquad [\xi \in \mathbf{R}^{2p}]\,,$$

[165] Or *local symplectomorphism.*

[166] Abuse of language. In fact we mean the Lagrange form on the open set $\mathrm{dom}(A)$ induced from the Lagrange form on V in the sense of (9.28).

[167] We will see that there are many other symplectomorphisms (10.23, 10.34). This is an important difference with the case of a Euclidean space: the only diffeomorphisms from E to E which preserve the symmetric tensor g defining the Euclidean metric are the "displacements" $x \mapsto B(x) + C$ ($B \in O(E)$, $C \in E$).

the σ_{jk} being the entries of the matrix J (8.10♡).

If V is a symplectic manifold of dimension $2p$, we thus see that the *canonical charts of V are the same as the canonical transformations from* $\mathbf{R}^{2p}$ *to* V. □

Keeping (4.10) and (4.11) in mind, we can see immediately that:

(10.14) The inverse of a canonical transformation and the composition of two canonical transformations are canonical transformations. □

(10.15) Let V be a symplectic, respectively presymplectic, manifold. From (10.14) it follows that the set of all symplectomorphisms from V to V form a *group*, which we will call $\mathrm{Can}(V)$.

(10.16) THEOREM: Let V and V' be two symplectic (or presymplectic) manifolds of the same dimension n and of the same rank (9.9). Let x_0 and x'_0 be points of V and V', respectively. Then there exists a canonical transformation A from V to V' such that $A(x_0) = x'_0$.

PROOF: We know that there exists a canonical chart Φ of V such that $x_0 \in \mathrm{im}(\Phi)$ (10.8, 10.9), and thus $x_0 = \Phi(\xi_0)$, $\xi_0 \in \mathbf{R}^n$. Likewise we have $x'_0 \equiv \Phi'(\xi'_0)$, where Φ' is a canonical chart of V' and $\xi'_0 \in \mathbf{R}^n$.

Φ and Φ' are canonical transformations (10.13), as well as the translation $T\,[\xi \mapsto \xi+\xi'_0-\xi_0]$ (10.12). Hence the same is true for $A = \Phi' \circ T \circ \Phi^{-1}$ (10.14) for which we have $A(x_0) = x'_0$. □

(10.17) If V is a presymplectic manifold and A a canonical transformation from V to V, one can verify immediately that A *preserves* the foliation of V (definition (5.13)).

Let us suppose that $A \in \mathrm{Can}(V)$ and that the foliation is sectionable. Then theorem (5.19) shows that there exists a diffeomorphism $\widehat{A}$ of the quotient U of V by its foliation such that $\widehat{A}(P(x)) \equiv P(A(x))$, where $P(x)$ denotes the *leaf* passing through the arbitrary point x of V. We know that U is a symplectic manifold (9.10); it is easy to check that $\widehat{A} \in \mathrm{Can}(U)$.

Canonical similitudes

(10.18) DEFINITION: Let V and V' be two symplectic (or presymplectic) manifolds and let A be a diffeomorphism from V to V'. We will say that A is a *canonical similitude* if the inverse image under A of the Lagrange form on V' is the product of the Lagrange form on V and a constant a. If $a = -1$, then A is said to be *anti-canonical*. □

Let φ and φ' denote the fields defining the Lagrange forms on V and V' respectively. The condition that A is a canonical similitude can be written as (see (4.5)):

(10.19) $$\varphi'(A(x))(D(A)(x)(\delta x))(D(A)(x)(\delta' x)) \equiv a\,\varphi(x)(\delta x)(\delta' x)\,.^{168}$$

(10.20) It is clear that the composition and inverses of canonical similitudes are canonical similitudes, and that the composition of two anti-canonical transformations is a canonical transformation.

(10.21) As is the case for canonical transformations, canonical similitudes of a presymplectic manifold V preserve the foliation of V. If V is sectionable, then the "global" canonical similitudes of V induce on U, the quotient of V by its foliation, global canonical similitudes (see (10.17)).

(10.22) THEOREM: Let V be a symplectic manifold and let A be a diffeomorphism from V to itself. A is *anti-canonical* if and only if the *graph* of A is a *Lagrangian* submanifold of $V \times V$.†

PROOF: The graph of A, which is the image of the injective immersion $x \mapsto \begin{pmatrix} A(x) \\ x \end{pmatrix}$, is a submanifold of $V \times V$; $V \times V$ is itself symplectic (9.7) and has twice the dimension of V. For this graph to be Lagrangian it is sufficient (and necessary) that it is isotropic (see (9.38)). From this the result follows trivially. □

Let E be a finite dimensional vector space. We have seen (9.12) that the associated *phase space* $\mathcal{E} = E^* \times E$ is symplectic. Moreover, $\mathcal{E}^2$ is symplectomorphic to the phase space of E^2. Now recall that we know an algorithm which yields Lagrangian submanifolds for an arbitrary phase space (9.39). Applying this construction to E^2, we obtain anti-canonical

[168] One might ask whether this condition can be satisfied with *a variable*. This is obviously the case, *whatever A is*, if V and V' have dimension 2, since then all 2-forms are proportional.

Now assume that V and V' are symplectic manifolds of dimension > 2. Taking into account the conditions $d\sigma_V = 0$ and $d\sigma_{V'} = 0$, one can show that $\frac{\partial a}{\partial x} \equiv 0$ and thus that a is constant on each connected component of the open set $\mathrm{dom}(A)$. Hence A coincides locally with a canonical similitude.

One might compare these results with those for *conformal transformations* of Euclidean spaces, for which dimension 2 also plays an exceptional role.

† Editors' note: Denoting by V^- the symplectic manifold V whose symplectic form equals the negative of the symplectic form on V, this theorem can be reformulated as saying that A is canonical if and only if its graph is a Lagrangian submanifold of $V \times V^-$.

transformations of $\mathcal{E}$; composing these transformations with the standard anti-canonical transformation

$$\begin{pmatrix} P \\ Q \end{pmatrix} \mapsto \begin{pmatrix} -P \\ Q \end{pmatrix} \qquad P \in E^*, Q \in E\,,$$

we obtain canonical transformations. We are thus lead to the following classical statement, which is easily verified.

(10.23) Let E be a finite dimensional vector space and let $\begin{pmatrix} Q \\ Q' \end{pmatrix} \mapsto \alpha$ be a differentiable map from E^2 to $\mathbf{R}$. If we define

$$P \equiv \frac{\partial \alpha}{\partial Q}\,, \qquad P' \equiv -\frac{\partial \alpha}{\partial Q'}\,,$$

and if the maps $\begin{pmatrix} Q \\ Q' \end{pmatrix} \mapsto \begin{pmatrix} P \\ Q \end{pmatrix}$ and $\begin{pmatrix} Q \\ Q' \end{pmatrix} \mapsto \begin{pmatrix} P' \\ Q' \end{pmatrix}$ are diffeomorphisms, then the map $\begin{pmatrix} P \\ Q \end{pmatrix} \mapsto \begin{pmatrix} P' \\ Q' \end{pmatrix}$ is a canonical transformation of the phase space $E^* \times E$. □

(10.24) EXAMPLE: For $E = \mathbf{R}$ and $\alpha \equiv q\sinh(q')$, we obtain the following symplectomorphism of $\mathbf{R}^2$:

$$\begin{pmatrix} p' \\ q' \end{pmatrix} \equiv \begin{pmatrix} -q\sqrt{1+p^2} \\ \mathrm{Log}(p+\sqrt{1+p^2}) \end{pmatrix}$$

with inverse

$$\begin{pmatrix} p \\ q \end{pmatrix} \equiv \begin{pmatrix} \sinh(q') \\ -p'/\cosh(q') \end{pmatrix},$$

which is visibly not of the form (10.12). □

Covering spaces of symplectic manifolds

With the definition of covering spaces (1.42) in mind, the reader can immediately verify the following theorem.

(10.25) THEOREM: Let V be a symplectic manifold and let (W, G, P) be a covering space of V. If we define σ_W to be the inverse image of σ_V under the projection P, W becomes a *symplectic manifold* and the elements of the group G are symplectomorphisms of W. □

Conversely, using (1.45) one can show:

(10.26) Let W be a symplectic manifold and let G be a *discrete* subgroup of $\mathrm{Can}(W)$. Then the *quotient manifold* $V = W/G$ (which has the same

dimension as W) possesses a *unique symplectic structure* such that σ_W is the inverse image of σ_V under the projection P. □

(10.27) EXAMPLE: Let U be a symplectic manifold and let n be a positive integer. Furthermore, let W be the set of n-tuples $x \equiv \begin{pmatrix} {}^1x \\ \vdots \\ {}^nx \end{pmatrix}$ of points of U such that ${}^1x, {}^2x, \ldots, {}^nx$ are *all distinct*. It is easy to verify that W is an open subset of the direct product U^n, which is a symplectic manifold (9.7).

If Γ is a group of permutations of the integers $1, 2, \ldots, n$ and if for $\gamma \in \Gamma$ we define

$$\hat{\gamma} \begin{pmatrix} {}^1x \\ \vdots \\ {}^nx \end{pmatrix} \equiv \begin{pmatrix} {}^{\gamma^{-1}(1)}x \\ \vdots \\ {}^{\gamma^{-1}(n)}x \end{pmatrix},$$

then one can easily verify that the set $\hat{\Gamma}$ of maps $\hat{\gamma}$ is a discrete subgroup of $\mathrm{Can}(V)$. Thus the quotient manifold $V = W/\hat{\Gamma}$ is a *symplectic manifold* whose dimension is the same as the dimension of W, that is, n times the dimension of U. □

Infinitesimal canonical transformations

(10.28) DEFINITION: Let V be a Hausdorff symplectic (or presymplectic) manifold and let f be a differentiable vector field on V. f is called an *infinitesimal canonical transformation* if the diffeomorphism $\exp(tf)$ is a canonical transformation (for every t for which $\exp(tf)$ exists). □

If we denote by φ the field $x \mapsto \sigma$ defining the Lagrange form on V, then f is an infinitesimal canonical transformation if and only if (notation (4.5))

$\diamondsuit$ $$\exp(tf)_(\varphi)(x) = \varphi(x)$$

for every t and x such that the left hand side of $\diamondsuit$ exists, that is, such that $\exp(tf)(x)$ exists. For a fixed x the allowed values of t form an interval containing 0 (3.2). Thus equation $\diamondsuit$ is equivalent to

$$\frac{\partial}{\partial t}[\exp(tf)_(\varphi)(x)] \equiv 0\,,$$

or, taking into account the definition of the Lie derivative (4.15), to the condition $\exp(tf)_([f, \varphi]) \equiv 0$, which can also be written as

$$[f, \varphi] \equiv 0\,.$$

Hence, using the notation of (4.16) we have proved:

(10.29) THEOREM: A vector field $x \mapsto \delta x$ on a symplectic or presymplectic manifold V is an infinitesimal canonical transformation if and only if $\delta_L \sigma \equiv 0$. □

(10.30) If V is not Hausdorff, one can cover it by open sets (cells for example) that are. The formula $\delta_L \sigma \equiv 0$ then expresses the fact that $x \mapsto \delta x$ is an infinitesimal canonical transformation on each of these open sets. Even in this case we will say that $x \mapsto \delta x$ is an infinitesimal canonical transformation of V. □

Using Cartan's formula (4.39) it follows that $\delta_L \sigma \equiv 0$ can also be written as

(10.31) $$\mathrm{d}[\sigma(\delta x)] \equiv 0 .$$

Applying Poincaré's lemma (4.37), we immediately deduce:

(10.32) COROLLARY: Let $f[x \mapsto \delta x]$ be a differentiable vector field on a symplectic or presymplectic manifold V. The condition

$\diamondsuit$ there exists a differentiable scalar field $[x \mapsto u]$ such that $$\sigma(\delta x) \equiv -\mathrm{d}u$$

is *sufficient* for f to be an infinitesimal canonical transformation. It is *necessary* if $\mathrm{dom}(f)$ is *simply connected*. □

Using the *symplectic gradient* (9.16), we can state in particular:

(10.33) COROLLARY: Let V be a *symplectic* manifold and let $f[x \mapsto \delta x]$ be a differentiable vector field on V. The condition

$\diamondsuit$ there exists a dynamical variable u such that $$f(x) \equiv \mathrm{grad}(u)$$ [169]

is sufficient for f to be an infinitesimal canonical transformation. It is necessary if $\mathrm{dom}(f)$ is *simply connected*. □

These results show the *extreme richness* of the set of canonical transformations, as does the following theorem.

(10.34) THEOREM: For every differentiable function u on $\mathbf{R}^{2p}$ with bounded first order partial derivatives (these functions form an infinite dimensional vector space), the field $f[x \mapsto \mathrm{grad}\, u]$ (defined by the symplectic structure (10.13) of $\mathbf{R}^{2p}$) is an infinitesimal canonical transformation and

$$\exp(tf) \in \mathrm{Can}(\mathbf{R}^{2p}) \qquad \forall t \in \mathbf{R} .$$

[169] One can also write $\delta x \equiv \underset{u}{\delta} x$ (notation (9.20)).

Furthermore, one can restrict attention to functions u which vanish outside a given ball (these functions still form an infinite dimensional vector space). The corresponding symplectomorphisms permute points of the ball but leave other points fixed. □

Let $f\,[x \mapsto \delta x]$ be a differentiable vector field on a symplectic or presymplectic manifold V. Looking at the proof of (10.29) one can easily establish the *equivalence* of the following two assertions:

(10.35)
- ♢ There exists a constant α such that $\delta_L \sigma \equiv \alpha\,\sigma$.
- ♡ For each t, the map $A = \exp(tf)$ is a canonical similitude.[170]

In such a case we will call f an *infinitesimal canonical similitude.*

Now let us suppose that $\alpha \neq 0$ (if not, f would be an infinitesimal canonical transformation). Using Cartan's formula, condition ♢ may be written as

$$\sigma \equiv \mathrm{d}\varpi \tag{10.36}$$

with

$$\varpi \equiv \frac{1}{\alpha}\,\sigma(\delta x)\,. \tag{10.37}$$

Thus if $\mathrm{dom}(f) = V$, formula (10.37) gives us a *potential* ϖ for the Lagrange form σ_V (definition (9.14)).

[170]Note that the constant a (definition (10.18)) equals $\mathrm{e}^{\alpha t}$.

§11. Dynamical Groups

The definition of a dynamical group

(11.1) Let V be a symplectic (or presymplectic) manifold. We will call a Lie group G acting on V a *dynamical group* of V if and only if[171]

$$\underline{a}_V \in \mathrm{Can}(V) \qquad \forall a \in G\,. \qquad \square$$

(11.2) EXAMPLE: If E is a symplectic vector space, the Lie group $\mathrm{Sp}(E)$ (8.19) acts on E and its elements are canonical transformations (10.12). Thus $\mathrm{Sp}(E)$ is a dynamical group of E. $\square$

(11.3) EXAMPLE: If G is a dynamical group of a presymplectic manifold V and if $\widetilde{G}$ is a *Lie subgroup* of G (6.13), then $\widetilde{G}$ is evidently a dynamical group of V. For instance, (8.26) shows that the unitary group $\mathrm{U}(p)$ — or its Lie subgroup $\mathrm{SU}(p)$ — is a dynamical group for every symplectic vector space of dimension $2p$. $\square$

(11.4) EXAMPLE: Let A be a symplectomorphism from a symplectic manifold V to a symplectic manifold V' and let G be a dynamical group of V. If we define

$$\underline{a}_{V'} = A \circ \underline{a}_V \circ A^{-1} \qquad \forall a \in G\,,$$

it follows that G acts on V' and that the $\underline{a}_{V'}$ are symplectomorphisms of V'. And thus G becomes a dynamical group of V' as well.

If $V' = V$, then, by varying the symplectomorphism A, we see that a dynamical group can act in infinitely many ways on V. $\square$

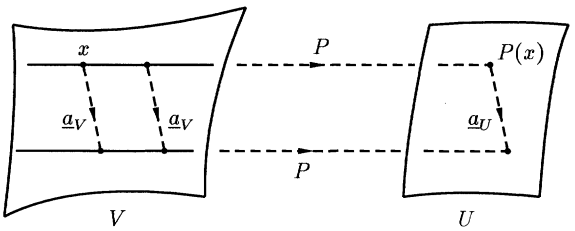
(11.5) EXAMPLE: Let V be a sectionable presymplectic manifold (9.8), let U be the quotient symplectic manifold of V by its foliation (9.10) (Fig. 11.I), and let G be a dynamical group of V. We have seen (10.17) that for each $a \in G$, the map $\underline{a}_V$ preserves the foliation and that there exists a symplectomorphism of U, which we will denote here by $\underline{a}_U$, such that

(11.6)
$$\underline{a}_U(P(x)) \equiv P(\underline{a}_V(x))\,,$$

where $P(x)$ is the leaf of the foliation which passes through the point x of V.

[171] Notation: $\underline{a}_V$ (6.4), $\mathrm{Can}(V)$ (10.15).

Fig 11.I



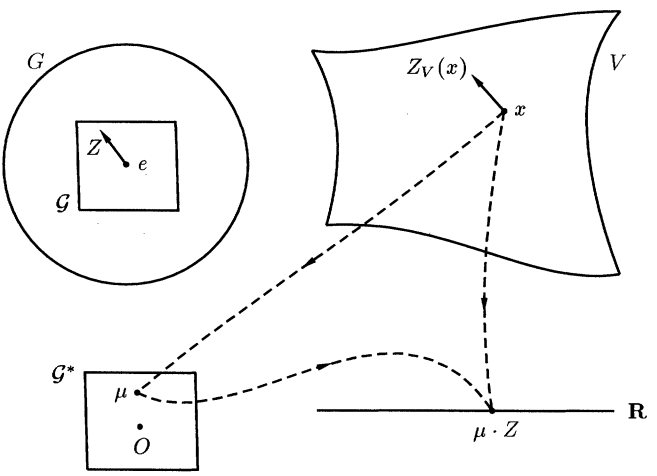
One can easily verify that this definition of $\underline{a}_U$ defines an *action* of G on U (6.4), and thus that G is a dynamical group of U. □

(11.7a) DEFINITION: Let G be a Lie group with Lie algebra $\mathcal{G}$. We will call an element of the dual space $\mathcal{G}^*$ of $\mathcal{G}$ a *torsor*.[172]

(11.7b) DEFINITION: If G is a dynamical group of a symplectic or presymplectic manifold V (Fig. 11.II), we will say that a variable torsor μ is a *moment* of the G-action,† if there exists a differentiable map $x \mapsto \mu$ from V to $\mathcal{G}^*$ such that

$$\diamondsuit \qquad \sigma(Z_V(x)) \equiv -\mathrm{d}[\mu \cdot Z] \qquad \text{for every constant } Z \in \mathcal{G}\,.^{173} \qquad \square$$

Fig 11.II



[172] To every torsor μ there corresponds a field $[x \mapsto \varpi]$ of 1-forms on G which is invariant under right translation and which takes the value μ when x is the identity element. These 1-forms are called *Maurer-Cartan forms*. The "systems of forces" in elementary mechanics are the torsors of the group of Euclidean displacements (see (12.21)).

† Editors' note: Also called a *moment map* or *momentum map* for the G-action.

[173] Z_V is the vector field on V defined in (6.11).

(11.8) THEOREM:

a) Let G be a dynamical group of a symplectic manifold V which possesses a moment μ. If one *adds a constant torsor* to the moment μ, one again obtains a moment. All moments of G are obtained this way if V is connected.

b) If V is *Hausdorff and simply connected*,[174] *then every dynamical group of V possesses a moment.*

c) If G is a dynamical group of a symplectic manifold V and if the Lie algebra of G *coincides with its derived algebra, then G possesses a moment.*

PROOF: (a) It is obvious that (11.7$\diamondsuit$) does not change if we add a constant to μ. Conversely, if μ and μ' satisfy $\diamondsuit$, then $\mathrm{d}[[\mu' - \mu] \cdot Z] \equiv 0$. Thus $[\mu' - \mu] \cdot Z$ is constant if V is connected (4.38). Hence $\mu' - \mu =$ constant because the constant Z is arbitrary.

(b) Now suppose that V is Hausdorff and simply connected and choose a basis $S = [S_1 \dots S_n]$ of the Lie algebra $\mathcal{G}$ of G. For $j = 1, 2, \dots, n$ and for t real we know that $\exp(t[S_j]_V)$ is a diffeomorphism of V given by the action of the group G (6.12) and thus is a canonical transformation. Hence $[S_j]_V$ is an infinitesimal canonical transformation (10.28). Consequently, there exists a differentiable scalar field $x \mapsto \mu_j$ on V such that $\sigma([S_j]_V(x)) \equiv -\mathrm{d}\mu_j$ (10.32). It now suffices to define $\mu \equiv [\mu_1 \dots \mu_n] \cdot S^{-1}$ to satisfy (11.7$\diamondsuit$).

(c) Let Z and Z' be two elements of the Lie algebra of a dynamical group G, and define $\delta x = Z_V(x)$ and $\delta' x = Z'_V(x)$. Then we know that the Lie derivatives $\delta_L \sigma$ and $\delta'_L \sigma$ are zero. Let us now compute the Lie derivative $\delta_L[\sigma(\delta' x)]$. On the one hand one finds, using (4.24) and (6.12),

$$\delta_L[\sigma(\delta' x)] \equiv [\delta_L \sigma](\delta' x) + \sigma(\delta_L \, \delta' x) \;\equiv\; \sigma([Z, Z']_V(x)) \,,$$

and on the other hand, using (4.39) and (10.31), one finds

$$\delta_L[\sigma(\delta' x)] \equiv \mathrm{d}[\sigma(\delta' x)](\delta x) + \mathrm{d}[\sigma(\delta' x)(\delta x)] \;\equiv\; \mathrm{d}[\sigma(\delta' x)(\delta x)] \,,$$

whence the identity, *valid for every dynamical group*,

$$\heartsuit \qquad \sigma([Z, Z']_V(x)) \equiv -\mathrm{d}[\sigma(Z_V(x))(Z_{V'}(x))] \,.$$

If the Lie algebra $\mathcal{G}$ of G equals its derived algebra (6.14), it follows that $\mathcal{G}$ possesses a basis S whose elements are Lie brackets: $S_j = [Z_j, Z'_j]$. Thus, if we define $\mu_j = \sigma(Z_j(x))(Z'_j(x))$, the identity $\heartsuit$ shows that $\mu \equiv [\mu_1 \dots \mu_n] \cdot S^{-1}$ is a moment. □

[174] These hypotheses are by no means necessary. For example, it suffices that the homotopy group of each component of V does not admit an abelian quotient.

If V is *symplectic*, definition (11.7$\diamondsuit$) of a moment can also be written, using the symplectic gradient (9.16), as

$$Z_V(x) \equiv \operatorname{grad}[\mu \cdot Z] \equiv \underset{\mu \cdot Z}{\delta} x \,. \tag{11.9}$$

(11.10) THEOREM: Let V be a symplectic or presymplectic manifold. Let $x \mapsto \varpi$ be a field of 1-forms such that $\sigma_V \equiv \mathrm{d}\varpi$ [175] and let G be a Lie group acting on V which leaves ϖ *invariant*.[176] Then G is a *dynamical group* of V and possesses a moment μ defined by

$$\diamondsuit \qquad \mu \cdot Z \equiv \varpi(Z_V(x)) \qquad \text{for every } Z \text{ in the Lie algebra of } G.$$

PROOF: Theorem (4.36) shows immediately that the inverse image of $[x \mapsto \sigma_V]$ under the operations $\underline{a}_V$ of the group remains $[x \mapsto \sigma_V]$, because $\sigma_V \equiv \mathrm{d}\varpi$. Thus G is indeed a dynamical group.

Now let Z be an element of the Lie algebra of G and define $\delta x \equiv Z_V(x)$. Since the inverse image of $x \mapsto \varpi$ under $\exp(tZ_V)$ is still $x \mapsto \varpi$ (because $\exp(tZ_V)$ is an action of the group), one can show as in (10.29) that the Lie derivative $\delta_L \varpi$ is zero. Applying Cartan's formula, we obtain $[\mathrm{d}\varpi](\delta x) + \mathrm{d}[\varpi(\delta x)] \equiv 0$, that is, $\sigma(Z_V(x)) + \mathrm{d}[\varpi(Z_V(x))] \equiv 0$. Taking (11.7$\diamondsuit$) into account, we obtain $\diamondsuit$. □

(11.11) EXAMPLE: Let us consider a variational problem satisfying (7.30), for example, a regular variational problem. If a group G leaves invariant the Lagrangian (7.36), it also leaves invariant *Cartan's form* ϖ (7.29) and thus we can apply the preceding theorem. Note that the quantity (11.7$\diamondsuit$) is the one which figures in *Noether's theorem* (7.36) and is often taken as the *definition of a moment*. This explains why the following generalization is still called

(11.12) NOETHER'S THEOREM: Let V be a *presymplectic* manifold and let μ be a moment of a dynamical group of V. Then μ is constant on each leaf of the characteristic foliation of V.

PROOF: It suffices to show that, for every Z in the Lie algebra, the function $\mu \cdot Z$ is a first integral, that is, according to (5.22), that $\ker(\sigma) \subset \ker(\mathrm{d}[\mu \cdot Z])$, or, according to (11.7$\diamondsuit$) that $\ker(\sigma) \subset \ker(\sigma(Z_V(x)))$, that is, that $[\sigma(y) = 0] \Rightarrow [\sigma(Z_V(x))(y) = 0]$, which follows from the antisymmetry of σ. □

[175] ϖ is a *potential* of σ_V (9.14).

[176] This means, of course, that the inverse image of the field $[x \mapsto \varpi]$ under the operations $\underline{a}_V$ of the group remains $[x \mapsto \varpi]$.

(11.13) Let $P(x)$ be the leaf of the characteristic foliation which passes through x (Fig. 11.III). Noether's theorem says that μ depends only on $P(x)$. Moreover, if V is sectionable, we know that $P(x)$ varies over the symplectic manifold U, which is the quotient of V by its characteristic foliation, and that G is also a dynamical group of U. One can verify easily that μ *is also a moment of the dynamical group* G *acting on* U.

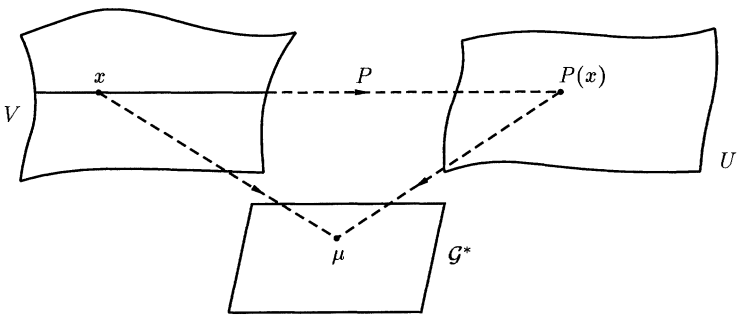


Fig 11.III

The cohomology of a dynamical group

Let $a \mapsto \underline{a}_E$ be a *linear representation* of a Lie group G (6.9). If μ lies in the vector space E^* dual to E and if we define

(11.14) $$\underline{a}_{E^*}(\mu) = \mu \cdot \underline{a^{-1}}_E \,,$$

one can check immediately that $a \mapsto \underline{a}_{E^*}$ is also a linear representation of G, called the *dual representation*[†] of the preceding one.

In particular, the *coadjoint representation* of G is the dual representation of the adjoint representation (6.24). Thus we obtain an *action of any Lie group* G *on its torsors* (11.7) given by

(11.15) $$\underline{a}_{\mathcal{G}^*}(\mu) = \mu \cdot \underline{a^{-1}}_{\mathcal{G}} \qquad \forall a \in G \,,\ \forall \mu \in \mathcal{G}^* \,.$$

If Z belongs to the Lie algebra of G, the associated vector field $Z_{\mathcal{G}^*}$ (definition (6.11)) is given by

(11.16) $$Z_{\mathcal{G}^*}(\mu) = \mu \cdot \mathrm{ad}(Z) \,.$$

This result follows immediately from (6.26).

[†]Editors' note: Also called the *contragredient representation*.

(11.17) THEOREM: (See Fig. 11.IV.) Let V be a *connected* symplectic (or pre-symplectic) manifold and let G be a dynamical group of V possessing a moment μ (11.7). Finally let ψ denote the map $x \mapsto \mu$ from V to the space $\mathcal{G}^*$ of torsors of G. Then

a) There exists a differentiable map θ from G to $\mathcal{G}^*$ defined by[177]

$\diamondsuit$ $$\theta(a) \equiv \psi(\underline{a}_V(x)) - \underline{a}_{\mathcal{G}^*}(\psi(x))\,.$$

b) The map θ satisfies the condition

$\heartsuit$ $$\theta(a \times b) \equiv \theta(a) + \underline{a}_{\mathcal{G}^*}(\theta(b))\,.$$

c) The derivative $f = D(\theta)(e)$, where e is the identity element of G, is a 2-form on the Lie algebra $\mathcal{G}$ of G which satisfies

$\clubsuit$ $$f(Z)([Z', Z'']) + f(Z')([Z'', Z]) + f(Z'')([Z, Z']) \equiv 0\,.$$

d) The following identities hold:

$\spadesuit$ $$D(\psi)(x)(Z_V(x)) \equiv \psi(x) \cdot \mathrm{ad}(Z) + f(Z)$$

$\sharp$ $$\sigma(Z_V(x))(Z'_V(x)) \equiv \mu \cdot [Z, Z'] + f(Z)(Z')\,. \qquad \square$$

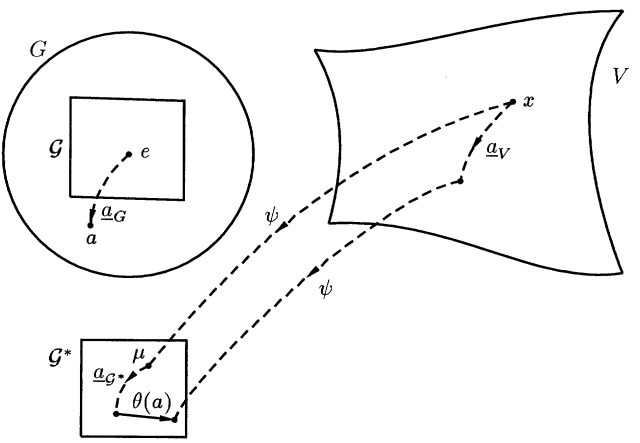


Fig 11.IV

PROOF: (a) Let φ be the field $x \mapsto \sigma$ on V. Definition (11.7$\diamondsuit$) of a moment can evidently be written as

♭ $\varphi(x)(Z_V(x))(f(x)) \equiv -D(\psi)(x)(f(x)) \cdot Z$, for every vector field f.

[177] With variables $a, b \in G$, $x \in V$, and $Z, Z', Z'' \in \mathcal{G}$.

By hypothesis the field φ is invariant under every operation $\underline{a}_V$ of the group G (11.1). Thus we have (4.5)

$$\natural \qquad \begin{aligned} &\varphi(x)(Z_V(x))(f(x)) \equiv \\ &\quad \equiv \varphi(\underline{a}_V(x))(D(\underline{a}_V)(x)(Z_V(x)))(D(\underline{a}_V)(x)(f(x)))\,. \end{aligned}$$

By definition of the image of a vector field under a map (2.38), we have

$$\bowtie \qquad D(\underline{a}_V)(x)(f(x)) \equiv [\underline{a}_V]_+(f)(\underline{a}_V(x))\,,$$

and likewise

$$\begin{aligned} D(\underline{a}_V)(x)(Z_V(x)) &\equiv [\underline{a}_V]_+(Z_V)(\underline{a}_V(x)) \\ &\equiv [\underline{a}_\mathcal{G} \cdot Z]_V(\underline{a}_V(x)) \qquad \text{(theorem (6.25))}. \end{aligned}$$

Substituting this into $\natural$ we get

$$\varphi(x)(Z_V(x))(f(x)) \equiv \varphi(\underline{a}_V(x))([\underline{a}_\mathcal{G} \cdot Z]_V(\underline{a}_V(x)))([\underline{a}_V]_+(f)(\underline{a}_V(x)))\,.$$

The right hand side of this expression is also the result of substituting $x \mapsto \underline{a}_V(x)$, $Z \mapsto \underline{a}_\mathcal{G} \cdot Z$, and $f \mapsto [\underline{a}]_+(f)$ into the left hand side of $\flat$. Hence this substitution does not change the right hand side of $\flat$ either, and thus

$$\begin{aligned} D(\psi)(x)(f(x)) \cdot Z &\equiv D(\varphi)(\underline{a}_V(x))([\underline{a}_V]_+(f)(\underline{a}_V(x))) \cdot \underline{a}_\mathcal{G} \cdot Z \\ &\equiv D(\psi)(\underline{a}_V(x))(D(\underline{a}_V(x)(f(x))) \cdot \underline{a}_\mathcal{G} \cdot Z \\ &\qquad \text{due to } \bowtie \\ &\equiv D(\psi \cdot \underline{a}_V)(x)(f(x)) \cdot \underline{a}_\mathcal{G} \cdot Z \quad (1.23,6). \end{aligned}$$

Since f is arbitrary, the above equation can be written as

$$\mathrm{d}[\psi(x) \cdot Z] \equiv \mathrm{d}[(\psi \cdot \underline{a}_V)(x) \cdot \underline{a}_\mathcal{G} \cdot Z]\,,$$

or, since V is connected (4.31), as

$$[\psi(x) - [\psi \cdot \underline{a}_V](x) \cdot \underline{a}_\mathcal{G}] \cdot Z = \text{constant}\,,$$

or, since Z is arbitrary, as

$$\psi(x) - \psi(\underline{a}_V(x)) \cdot \underline{a}_\mathcal{G} = \text{constant}\,.$$

Multiplying the above equation on the right by $\underline{a^{-1}}_\mathcal{G} = [\underline{a}_\mathcal{G}]^{-1}$ (6.8) and using definition (11.15) of the coadjoint representation, it follows that $\psi(\underline{a}_V(x)) - \underline{a}_{\mathcal{G}^*}(\psi(x))$ does not depend on x. Thus we may call it $\theta(a)$. To verify that θ is differentiable it suffices to choose a fixed point x in V, whence (a).

(b) Formula ♡ is obtained by successively replacing, in ◇, a by $a \times b$, x by $\underline{b}_V(x)$, and a by b, and then comparing the obtained results.

(d) Using notation (6.10), equation ◇ can be written as

$$\theta(a) \equiv \psi(\hat{x}(a)) - \widehat{\psi(x)}(a)\,.$$

Thus we can compute $\delta[\theta(a)]$ for $a = e$ and $\delta a = Z'$ by applying (6.11), giving

$$f(Z') = D(\psi)(x)(Z'_V(x)) - Z'_{\mathcal{G}^*}(\psi(x))\,,$$

which is ♠ due to (11.16). Next, ♭ gives

$$f(Z')(Z) = -\varphi(x)(Z_V(x))(Z'_V(x)) - \psi(x)\cdot \mathrm{ad}(Z')(Z)\,,$$

whence (d).

(c) Obviously f is a 2-form. As far as the identity ♣ is concerned, it will be proved in (11.23). □

(11.18) We know that one can pass from one moment to another by adding a constant torsor (11.8). If we change μ into $\mu^* \equiv \mu - \mu_0$ with $\mu_0 \in \mathcal{G}^*$, we observe that θ and f are changed into θ^* and f^*:

$$\theta^*(a) \equiv \theta(a) + \underline{a}_{\mathcal{G}^*}(\mu_0) - \mu_0$$

and

$$f^*(Z) \equiv f(Z) + Z_{\mathcal{G}^*}(\mu_0)\,.$$

Of course the results of theorem (11.17) remain applicable to θ^* and f^*. This might suggest the ideas which follow.

The cohomology of a Lie group

(11.19) DEFINITION/THEOREM:† Let G be a Lie group and let $a \mapsto \underline{a}_E$ be a linear representation of G. An *E-cocycle* of G is any differentiable map

†Editors' note: This definition gives group cohomology in dimension 1. More generally an E-cochain in dimension n is a map $\theta : G^n \to E$ whose differentiability depends on circumstances. Given such a θ one defines an E-cochain $\Delta\theta$ in dimension $n+1$ by the formula

$$(\Delta\theta)(a, b_1, \ldots, b_n) = \underline{a}_E(\theta(b_1, \ldots, b_n)) + (-1)^{n+1}\theta(a, b_1, \ldots, b_{n-1}) - \theta(a \times b_1, b_2, \ldots, b_n) + \sum_{i=1}^{n-1}(-1)^{i+1}\theta(a, b_1, \ldots, b_i \times b_{i+1}, \ldots, b_n)\,.$$

θ is said to be an E-cocycle if $\Delta\theta = 0$; it is an E-coboundary if it is of the form $\theta = \Delta\mu$ for an E-cochain μ in dimension $n-1$. And "of course" any coboundary is a cocycle.

θ from G to E such that

$$\heartsuit \qquad \theta(a \times b) \equiv \theta(a) + \underline{a}_E(\theta(b)) \qquad \forall a, b \in G\,.$$

If for $\mu \in E$ we define

$$\diamondsuit \qquad \Delta(\mu)(a) = \underline{a}_E(\mu) - \mu \qquad \forall a \in G\,,$$

then $\Delta(\mu)$ is an E-cocycle of G called the *coboundary* of μ.

Two cocycles are said to be *cohomologous* if their difference is a coboundary.[178] Being cohomologous is an equivalence relation in the vector space of E-cocyles.[179]

PROOF: Immediate. □

(11.20) Using this language, the map θ (11.17$\diamondsuit$) is a *$\mathcal{G}^*$-cocycle of G* whose cohomology class *does not depend* on the choice of the moment. Thus the cohomology class of θ is a *characteristic property of the dynamical group G.*[180]

(11.21) EXAMPLE: Under the hypotheses of (11.11) one can check that the moment μ defined by (11.10$\diamondsuit$) leads to a θ which is zero. Thus the associated cohomology class is *zero*. Hence, if the cohomology class of a *dynamical group* G is not zero, the hypotheses of (11.11) cannot be satisfied, that is, there cannot exist a G-invariant potential.[181]

The cohomology of a Lie algebra

(11.22) THEOREM: Let G be a Lie group, e its identity element, and $\mathcal{G}$ its Lie algebra. Furthermore, let $a \mapsto \underline{a}_E$ be a linear representation of G. Then a) $Z \mapsto Z_E$ (notation (6.11)) is a linear map from $\mathcal{G}$ to $\mathrm{L}(E)$ satisfying

$$\diamondsuit \qquad [Z, Z']_E = Z'_E \cdot Z_E - Z_E \cdot Z'_E \qquad \forall Z, Z' \in \mathcal{G}\,.$$

[178] Thus a cocycle is cohomologous to 0 if it is a coboundary.

[179] The quotient (the set of equivalence classes) possesses an obvious vector space structure. It is called the *cohomology space* of the representation. This space reduces to $\{0\}$ if every E-cocycle is a coboundary.

[180] Abuse of language. One should say a characteristic property of the group G and the map $a \mapsto \underline{a}_V$ which defines the action of G on V.

[181] Therefore it is *impossible* that the symplectic structure is obtained from a *G-invariant Lagrangian* according to the procedure of (11.11) and (7.29).

b) If θ is an E-cocycle of G and if we define $f = D(\theta)(e)$, then f is a linear map from $\mathcal{G}$ to E satisfying

$$\clubsuit \qquad D(\theta)(a)(Z_G(a)) = Z_E(\theta(a)) + f(Z) \qquad \forall a \in G\,,\ \forall Z \in \mathcal{G}$$

$$\spadesuit \qquad \theta(\exp(Z_G)(e)) = \left[\int_0^1 \exp(tZ_E)\,dt\right](f(Z)) \qquad \forall Z \in \mathcal{G}$$

$$\heartsuit \qquad f([Z, Z']) = Z'_E(f(Z)) - Z_E(f(Z')) \qquad \forall Z, Z' \in \mathcal{G}\,.$$

c) If G is connected and if f is zero, then θ is zero.

PROOF: (b) Let us write the definition of a cocycle (11.19♡) in the form

$$\theta(\hat{b}(a)) - \theta(a) \equiv \widehat{\theta(b)}(a) \qquad \text{(notation (6.10))}.$$

Applying the derivation δ with $a = e$, $\delta a = Z$, and $\delta b = 0$, we find, using (6.11):

$$D(\theta)(b)(Z_G(b)) - D(\theta)(e)(Z) = Z_E(\theta(b))\,,$$

that is, after replacing b by a, the identity ♣.

From ♣ we deduce immediately

$$\frac{d}{dt}\left[\exp(-tZ_E)(\theta(\exp(tZ_G)(e)))\right] \equiv \exp(-tZ_E)(f(Z))\,.$$

Since $\theta(e) = 0$ (just take $a = b = e$ in (11.19♡)), this gives ♠.

Now let us define $\delta a \equiv Z_G(a)$, $\delta' a \equiv Z'_G(a)$ and $\delta'' z \equiv [Z, Z']_G(a)$; note that the derivation δ'' is the bracket of δ and δ' (6.12). As the image of $\theta(a)$ lies in the fixed vector space E, we have $\delta''[\theta(a)] \equiv \delta\delta'[\theta(a)] - \delta'\delta[\theta(a)]$ (2.45). This formula can be expanded, using ♣ several times, and will give us ♡ when we take $a = e$.

(a) In the special case when θ is the coboundary of μ we have $f(Z) \equiv Z_E(\mu)$, and ♡ reduces to ♢.

(c) If G is connected, every element of G can be written as (6.32.I.c)

$$a = \exp({}^1Z_G)\cdots\exp({}^pZ_G)(e)\,.$$

If $f = 0$, ♠ shows that $\theta((Z_G)(e)) = 0$; it follows that $\theta(a) = 0$ for $p = 1$. The general case follows by induction on p and applying (11.19♡). □

(11.23) In (11.17) f is antisymmetric. Taking (11.16) into account, formula ♡ above takes the more symmetric form (11.17♣).

These results motivate the following.

(11.24) DEFINITION/THEOREM:† Let $\mathcal{G}$ be a *Lie algebra*. A vector space E is called a *$\mathcal{G}$-module* if there has been given a linear map $Z \mapsto Z_E$ from $\mathcal{G}$ to $\mathrm{L}(E)$ such that [182]

$$\clubsuit \qquad [Z, Z']_E \equiv Z'_E \cdot Z_E - Z_E \cdot Z'_E \,.$$

If E is a $\mathcal{G}$-module, an *E-cocycle of $\mathcal{G}$* will be any *linear* map of $\mathcal{G}$ to E satisfying

$$\heartsuit \qquad f([Z, Z']) \equiv Z'_E(f(Z)) - Z_E(f(Z')) \,.$$

If for $\mu \in E$ we define

$$\diamondsuit \qquad \delta(\mu)(Z) \equiv Z_E(\mu) \,,$$

then $\delta(\mu)$ is an E-cocycle of $\mathcal{G}$ called the *coboundary* of μ.[183]

Two E-cocycles of $\mathcal{G}$ are said to be *cohomologous* if their difference is a coboundary. Being cohomologous is an equivalence relation in the vector space of E-cocycles of $\mathcal{G}$.[184]

PROOF: Trivial. □

(11.25) DEFINITION: A Lie algebra is called *semisimple* if the tensor k defined by

$$k(Z)(Z') = \mathrm{Tr}(\mathrm{ad}(Z) \cdot \mathrm{ad}(Z')) \qquad \forall Z, Z' \in \mathcal{G} \quad \text{(see (6.13))}$$

is injective (k is called the *Killing form*). A Lie group G is called *semisimple* if it is *connected* and *its Lie algebra is semisimple*. □

These definitions permit us to formulate the next lemma and its corollary.

---

†Editors' note: This definition gives Lie algebra cohomology in dimension 1. More generally an *E-cochain* in dimension n is an n-linear antisymmetric map $f : \mathcal{G}^n \to E$. For f one defines an E-cochain δf in dimension $n+1$ by

$$(\delta f)(Z_0, \ldots, Z_n) = \sum_{i=0}^{n} (-1)^i Z_i \Big(f(Z_0, \ldots, \widehat{Z_i}, \ldots Z_n \Big) + \\ + \sum_{i<j} (-1)^{i+j} f([Z_i, Z_j], Z_0, \ldots, \widehat{Z_i}, \ldots, \widehat{Z_j}, \ldots, Z_n) \,,$$

where the hat means that the corresponding element should not appear in the list. f is said to be an E-cocycle if $\delta f = 0$; it is an E-coboundary if it is of the form $f = \delta\mu$ for an E-cochain μ in dimension $n-1$. And again every E-coboundary is an E-cocycle. Nota Bene: This formula uses the standard sign of the Lie bracket, which is not the one used in this book.

[182]The map $Z \mapsto Z_E$ is called a *representation* of the Lie algebra $\mathcal{G}$.

[183]This should not be confused with the coboundary defined in (11.19◇).

[184]The same remarks as in the footnote at (11.19).

(11.26) WHITEHEAD'S LEMMA: Let $\mathcal{G}$ be a semisimple Lie algebra. If E is a finite dimensional $\mathcal{G}$-module, every E-cocycle of $\mathcal{G}$ is a coboundary.

PROOF: See Jacobson.[185]

(11.27) COROLLARY: Let G be a semisimple Lie group and let $a \mapsto \underline{a}_E$ be a linear representation of G. Then every E-cocycle is a coboundary.

PROOF: Let θ be an E-cocycle and define $f = D(\theta)(e)$. According to Whitehead's lemma there exists a $\mu \in E$ such that $f = \delta(\mu)$. Writing $\theta^* = \theta - \Delta(\mu)$ it follows that θ^* is an E-cocycle satisfying $D(\theta^*)(e) = 0$, and thus, using (11.22c), that θ^* is zero, whence $\theta = \Delta(\mu)$. □

Symplectic manifolds defined by a Lie group

Let θ be an E-cocycle of a Lie group G (hypotheses (11.19)). If we define

$$\underline{a}_{E_\theta}(\mu) = \underline{a}_E(\mu) + \theta(a) \qquad \forall a \in G\,,\ \forall \mu \in E\,, \tag{11.28}$$

it follows immediately that the map $a \mapsto \underline{a}_{E_\theta}$ defines an *action* of G on E. We use the notation E_θ to denote the *same manifold* as E, but with a *different G-action* (if θ is not zero).

Let $\mathcal{G}$ be the Lie algebra of G. For $Z \in \mathcal{G}$ the associated vector field Z_{E_θ} (definition (6.11)) is given by

$$Z_{E_\theta}(\mu) = Z_E(\mu) + f(Z) \qquad \forall \mu \in E\,, \tag{11.29}$$

where $f = D(\theta)(e)$ is the cocycle of the Lie algebra associated with θ (11.22).

(11.30) DEFINITION: Let G be a Lie group, $\mathcal{G}$ its Lie algebra, and $\mathcal{G}^*$ the dual of $\mathcal{G}$. A differentiable map θ from G to $\mathcal{G}^*$ will be called a *symplectic cocycle* of G if

$\diamondsuit$ $\qquad \theta(a \times b) \equiv \theta(a) + \underline{a}_{\mathcal{G}^*}(\theta(b)) \qquad$ [θ is a $\mathcal{G}^*$-cocycle]

$\heartsuit$ $\qquad f = D(\theta)(e)$ is antisymmetric. □

(11.31) If θ is a symplectic cocycle, every cohomologous cocycle is also symplectic (because the coboundary of a torsor is symplectic). Thus the cohomology classes of symplectic cocycles form a vector subspace (which we will call

[185] Jacobson [18].

the *symplectic cohomology space* of G) of the cohomology space of the coadjoint representation (11.19) (see Fig. 11.V).

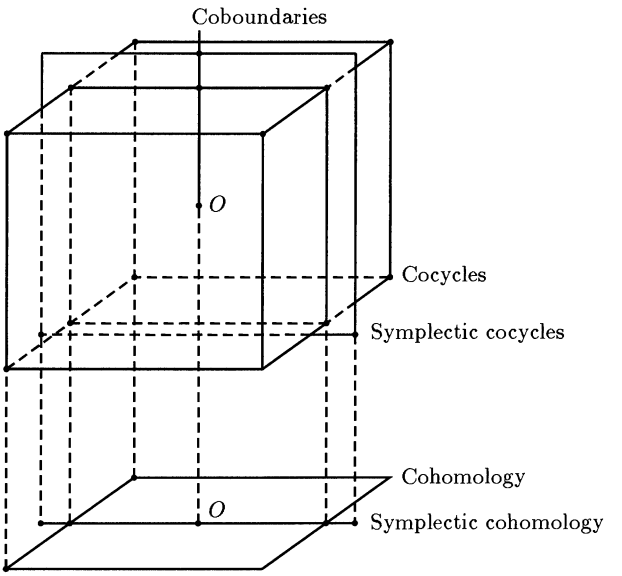


Fig 11.V

(11.32) If θ is a symplectic cocycle of G, then $f = D(\theta)(e)$ satisfies the identity (11.22♡). Taking into account (11.16) and the antisymmetry of f, this identity can be written in the more symmetric form

(11.33) $$f(Z)([Z', Z'']) + f(Z')([Z'', Z]) + f(Z'')([Z, Z']) = 0\,.$$[186]

(11.34) THEOREM: Let G be a Lie group, $\mathcal{G}$ its Lie algebra, and θ a symplectic cocycle of G. Furthermore, let U be an orbit of the action $a \mapsto \underline{a}_{\mathcal{G}^*_\theta}$ (notation (11.28)) and let μ be a variable point in U. Then U is a *submanifold* of $\mathcal{G}^*$, the space of torsors of G. A vector $\delta\mu$ is tangent to U at μ if there exists a $Z \in \mathcal{G}$ such that

♡ $$\delta\mu = \mu \cdot \mathrm{ad}(Z) + f(Z) \qquad (f = D(\theta)(e))\,.$$

Moreover, the dimension of U (assumed to be nonzero) is *even* and U admits the structure of a *symplectic manifold* whose Lagrange form σ_U is given by

♢ $$\sigma_U(\delta'\mu)(\delta\mu) = \delta'\mu(Z) \qquad \text{if } \delta\mu = \mu \cdot \mathrm{ad}(Z) + f(Z)\,.$$

[186] If θ is a coboundary $\theta = \Delta(\mu)$, we have $f(Z)(Z') \equiv \mu[Z, Z']$, and we recover the Jacobi identity.

Finally, G, acting on U, is a dynamical group and each point $\mu \in U$ is its own moment.†

PROOF: We know that U is a submanifold of $\mathcal{G}^*$ on which G acts (6.20), and that a vector $\delta\mu$ tangent to U at μ is of the form $Z_{\mathcal{G}^*_\theta}$. Taking (11.29) and (11.16) into account, we find ♡.

If $\delta'\mu$ is tangent to U at μ, there exists a Z' such that $\delta'\mu = \mu \cdot \mathrm{ad}(Z') + f(Z')$, and thus

$$\delta'\mu(Z) = \mu[Z', Z] + f(Z)(Z') \;=\; -\mu[Z, Z'] - f(Z')(Z) \;=\; -\delta\mu(Z')\,.$$

It follows that $\delta'\mu(Z)$ depends on Z only through the intermediary $\delta\mu$, hence we may write $\delta'\mu(Z) = \sigma_U(\delta'\mu)(\delta\mu)$. It is clear that σ_U is a 2-form on the tangent space to U at μ; σ_U is injective because

$$[\sigma_U(\delta'\mu) = 0] \quad\Longleftrightarrow\quad [\delta'\mu(Z) = 0 \quad \forall Z \in \mathcal{G}] \quad\Longleftrightarrow\quad [\delta'\mu = 0]\,.$$

We conclude that the dimension of U is even (9.2).

Let φ denote the field $\mu \mapsto \sigma_U$. Since $\mu \cdot \mathrm{ad}(Z) + f(Z)$ equals $Z_U(\mu)$,[187] definition ♢ of σ_U may be written as

$$\clubsuit \qquad \varphi(\mu)(Z'_U(\mu))(Z_U(\mu)) = Z'_U(\mu) \cdot Z\,.$$

If we calculate the inverse image of φ by $\underline{a}_U$, a being an element of G, we obtain, using definition (4.5)

$$\spadesuit \qquad \begin{aligned}&[\underline{a}_U]_-(\varphi)(\mu)(Z_U(\mu))(Z'_U(\mu)) = \\ &\quad = \varphi(\underline{a}_U(\mu))(D(\underline{a}_U)(\mu)(Z'_U(\mu)))(D(\underline{a}_U)(\mu)(Z_U(\mu)))\,.\end{aligned}$$

Noting that $\underline{a}_U$ is induced from $\underline{a}_{\mathcal{G}^*_\theta}$ and applying definition (11.28) of $\underline{a}_{\mathcal{G}^*_\theta}$ and definition (11.15) of the coadjoint representation, we obtain

$$\sharp \qquad D(\underline{a}_U)(\mu)(Z'_U(\mu)) = Z'_U(\mu) \cdot \underline{a^{-1}}_{\mathcal{G}}\,.$$

On the other hand, using (2.38) and (6.25), we obtain

$$\flat \qquad D(\underline{a}_U)(\mu)(Z_U(\mu)) = [\underline{a}_U]_+(Z_U)(\underline{a}_U(\mu)) \;=\; [\underline{a}_{\mathcal{G}}(Z)]_U(\underline{a}_U(\mu))\,.$$

Using ♯ and ♭, equation ♠ becomes

$$\varphi(\underline{a}_U(\mu))(Z'_U(\mu) \cdot \underline{a^{-1}}_{\mathcal{G}})([\underline{a}_{\mathcal{G}} \cdot Z]_U(\underline{a}_U(\mu)))\,,$$

†Editors' note: That is, the restriction to U of the identity map is a moment map for the G-action.

[187]Because $\underline{a}_U$ is induced on U by $\underline{a}_{\mathcal{G}^*_\theta}$ and Z_U by $Z_{\mathcal{G}^*_\theta}$, which in turn is given by (11.29).

which, due to ♣, is equivalent to

$$Z'_U(\mu)\cdot \underline{a^{-1}_{\mathcal{G}}}\cdot \underline{a}_{\mathcal{G}}\cdot Z = \varphi(\mu)(Z'_U(\mu))(Z_U(\mu))\,.$$

Thus $[\underline{a}_U]_-(\varphi)=\varphi$, that is, the field $\mu\mapsto\sigma$ is *invariant under the action of the group*. In particular, if we take $a=\exp(tZ_U)$, we find from formula (4.15)

$$0=\frac{\partial}{\partial t}\{[\exp(tZ_U)]_-(\varphi)(\mu)\}=[\exp(tZ_U)]_-([Z_U,\varphi])(\mu))\,,$$

which shows that the Lie derivative $[Z_U,\varphi]$ is *zero*. Taking $\delta\mu\equiv Z_U(\mu)$, we thus find $\delta_L\sigma_U=0$.

On the other hand, note that definition ♢ can also be written as

$$\sigma_U(\delta\mu)(\delta'\mu)\equiv -[\delta'\mu](Z)\ \equiv\ -\delta'[\mu(Z)]\ \equiv\ -\mathrm{d}[\mu(Z)](\delta'\mu)\,,$$

that is,

♮ $$\sigma_U(Z_U(\mu))\equiv -\mathrm{d}[\mu(Z)]\,.$$

Cartan's formula then gives

$$0\equiv\delta_L\sigma_U\ \equiv\ [\mathrm{d}\sigma_U](\delta\mu)+\mathrm{d}[\sigma_U(\delta\mu)]\ \equiv\ [\mathrm{d}\sigma_U](Z_U(\mu))-\mathrm{d}[\mathrm{d}[\mu(Z)]]\,,$$

or, according to Poincaré's lemma, $[\mathrm{d}\sigma_U](Z_U(\mu))\equiv 0$. As $Z_U(\mu)$ is an arbitrary tangent vector to U, we find $\mathrm{d}\sigma_U\equiv 0$. Thus the injective form σ_U defines a symplectic structure on U. Since the action of the group G leaves σ_U invariant, G is a dynamical group. Formula ♮ shows that μ is its own moment (see (11.7)). □

(11.35) We now have at our disposal an algorithm which associates to each Lie group G and each symplectic cocycle θ of G (*especially* to $\theta\equiv 0$) a family of symplectic manifolds on each of which G acts canonically and transitively.

(11.36) If we add a coboundary to θ, the new orbits are obtained from the old ones by a translation, which induces at the same time a symplectomorphism between them. *Thus the symplectic structure of the orbits depends only on the cohomology class of* θ.

(11.37) If the symplectic cohomology of G is zero, it suffices to study the orbits corresponding to the case $\theta=0$ (that is, *the orbits of the coadjoint representation*) to have a symplectic model of all possible cases. This is the case if G is *semisimple*, since then its cohomology is zero by Whitehead's lemma (11.27) and thus a fortiori its symplectic cohomology (11.31).

From a *local* symplectic point of view, the construction (11.34) is *universal*. This is what is expressed in the following theorem.

(11.38) THEOREM: Let V be a symplectic manifold on which a connected dynamical group G acts transitively having μ as moment. Denote by ψ the map $x \mapsto \mu$ and let θ be the associated cocycle (11.17♢).

1) ψ maps V onto a manifold U, *of the same dimension*, constructed from the group G by algorithm (11.34). Moreover

♡ $$\psi(\underline{a}_V(x)) \equiv \underline{a}_U(\psi(x))\,.$$

2) If ψ is injective, it is a *symplectomorphism* from V to U. If not, one can cover V by open sets on which ψ is a *canonical transformation* from V to U.
3) For ψ to be a *symplectomorphism*, it suffices that the *stabilizer* G_μ of an element μ of U (on which G acts according to (11.28)) is *connected*.

PROOF: Let x_0 be a point of V. Since G acts transitively, each $x \in V$ is of the form $\underline{a}_V(x_0)$, or equivalently, $\widehat{x}_0(a)$. It follows that V is connected because $\widehat{x}_0$ is continuous (theorem (1.51)).

Thus (11.17♢) defines a map θ from V to $\mathcal{G}^*$, which is a symplectic cocycle. This allows us to apply (11.34). Formula (11.17♢) can then be written as

♢ $$\psi(\underline{a}_V(x)) = \underline{a}_{\mathcal{G}^*_\theta}(x)\,.$$

Taking $x = x_0$ it follows that ψ maps V to the orbit U passing through $\mu_0 = \psi(x_0)$ and that we can write ♢ in the form (11.38♡). Likewise formula (11.17♠) can be written as

♣ $$D(\psi)(x)(Z_V(x)) \equiv Z_U(\mu)$$

and (11.17♯) as

♠ $$\sigma_V(Z_V(x))(Z'_V(x)) \equiv \sigma_U(Z_U(\mu))(Z'_U(\mu))\,,$$

which shows that σ_V is the inverse image of σ_U under ψ.

Since V and U are symplectic, ♠ shows that $[Z_V(x) = 0] \Leftrightarrow [Z_U(\mu) = 0]$. Thus the linear maps $Z \mapsto Z_V(x)$ and $Z \mapsto Z_U(\mu)$ have the same kernel. Hence they have the same rank and thus V and U have the same dimension.

As $D(\psi)(x)$ is injective everywhere, theorem (1.35) shows that we can cover V by open sets on which ψ is a diffeomorphism, and thus a canonical transformation (10.10).

Furthermore, let us suppose that the stabilizer of an element of U is connected and let x be a point of V, $\mu = \psi(x)$. Then $\heartsuit$ shows that the stabilizer G_x is a subgroup of G_μ and thus is a Lie subgroup of G (6.36), (6.34). Since it has the same dimension as G_μ (6.36), it is an open subgroup. But G_μ is connected because according to (6.36$\heartsuit$) all stabilizers are conjugate. Using (6.32.III) we see that $G_x = G_\mu$. Let us now suppose that $\psi(x) = \psi(y)$ and let b be an element of G such that $y = \underline{b}_V(x)$. Then we have $\psi(x) = \psi(y) = \psi(\underline{b}_V(x)) = \underline{b}_U(\psi(x))$. Hence $b \in G_\mu$ and thus $b \in G_x$. Therefore $y = \underline{b}_V(x) = x$, that is, ψ is injective. □

(11.39) EXAMPLE: Let E be a Euclidean space of dimension n and let G be the connected component of e in $O(E)$. This G is the connected Lie group which we have studied in §6; its Lie algebra $\mathcal{G}$ is the set of $Z \in \mathrm{L}(E)$ such that $\overline{Z} = -Z$. We define

$$\diamondsuit \qquad g(Z)(Z') = -\tfrac{1}{2}\operatorname{Tr}(Z \cdot Z')\,,$$

which gives $\mathcal{G}$ the structure of a Euclidean space of dimension $\frac{1}{2}n(n-1)$.

An elementary calculation shows that the Killing form (11.25) equals $[4-2n]g$. Thus, if $n \geq 3$, G is semisimple. Hence it suffices to study the orbits of the coadjoint representation (11.37). □

EXAMPLE: Let us treat the case $E = \mathbf{R}^3$ and $G = \mathrm{SO}(3)$. The sphere S_2 defined by $\langle x, x\rangle \equiv 1$, $x \in \mathbf{R}^3$ possesses a symplectic structure σ defined by

(11.40) $$\sigma(\delta x)(\delta' x) \equiv \lambda\langle x, \delta x \times \delta' x\rangle\,,$$

λ being a nonzero constant. Indeed, this σ is injective and its exterior derivative is certainly zero as it is a 3-form on a 2-dimensional manifold. One can easily show that for $a \in \mathrm{SO}(3)$ the map $\underline{a}_E$ leaves σ invariant. Thus SO(3) is a dynamical group. Moreover, it acts transitively on the 2-sphere.

An arbitrary element Z of the Lie algebra of SO(3) can be written in the form $j(z)$ defined as

(11.41) $$j(z)(y) = z \times y \qquad \forall y \in \mathbf{R}^3\,.$$

From this one can easily deduce the moment

(11.42) $$\mu \equiv g(-\lambda x)j^{-1}\,.$$[188]

Since $x \mapsto \mu$ is injective, the sphere S_2 equipped with the symplectic structure (11.40) is *symplectomorphic to an orbit of the coadjoint representation* (11.38). Letting λ vary over all positive real numbers yields all possible coadjoint orbits. □

[188] Notation (6.49).

(11.43) EXAMPLE: Let E be a vector space of dimension n seen as an additive abelian group. A symplectic cocycle θ of E is simply a 2-form on E. The orbits of the action associated with θ are the image of θ (which is a vector subspace of even dimension of the dual E^* of E) and its translates. □

Chapter III

Mechanics

§12. The geometric structure of classical mechanics

Material points

(12.1) Classical mechanics considers the world as made up of *material points*. A material point, characterized by its (nonzero) *mass* m, occupies a *geometric point* $\mathbf{r}$ at each *instant of time* t. The *velocity* of a point is by definition the vector $\mathbf{v} = \frac{d\mathbf{r}}{dt}$. The *force* $\mathbf{F}$ to which a material point is subjected is assumed to be known, and one writes *Newton's equations*

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F}\,. \tag{12.2}$$

(12.3) EXAMPLE: In the case of a Coulomb field, the assumption is

$$\mathbf{F} \equiv -k\,m\,\frac{\mathbf{r}}{r^3}\,,$$

k being a given constant ($r = \|\mathbf{r}\|$ is the distance from the material point to the origin O chosen in space). By differentiating, one can immediately verify that the vectors $\mathbf{l} = \mathbf{r} \times \mathbf{v}$ and $\mathbf{e} = \frac{1}{k}\,\mathbf{l} \times \mathbf{r} + \frac{\mathbf{r}}{r}$ are constant and that they are orthogonal.

Eliminating $\mathbf{v}$ one finds that

$$\langle \mathbf{l}, \mathbf{r} \rangle \equiv 0 \qquad \text{and} \qquad \langle \mathbf{e}, \mathbf{r} \rangle + \frac{l^2}{k} \equiv r\,. \tag{12.4}$$

These relations are the equations of the trajectory of a particle in a Coulomb field. For nonzero $\mathbf{l}$ one can check that this trajectory is a conic section with *focus* at O, *eccentricity* $e = \|\mathbf{e}\|$, *latus rectum* $\frac{l^2}{ke}$, and lies in a plane perpendicular to $\mathbf{l}$. For $e = 0$ this conic is a circle.

For $0 < e < 1$ the trajectory is an ellipse, and from (12.4) one can deduce that there exists a number u, called the *eccentric anomaly*, such that

$$\mathbf{r} = \frac{\mathbf{e}}{ek}\,\frac{l^2}{1-e^2}\,(e - \cos u) + \frac{\mathbf{e} \times \mathbf{l}}{ek}\,\frac{l}{\sqrt{1-e^2}}\,\sin u\,. \tag{12.5}$$

Substituting this expression for $\mathbf{r}$ into the definition of $\mathbf{l}$, one finds a differential equation for u which can be immediately integrated to give

Kepler's equation

$$u - e \sin u \equiv k^2 \frac{[1-e^2]^{3/2}}{l^3}[t - t_0], \tag{12.6}$$

t_0 being a constant of integration.

Knowing the parameters $\mathbf{l}, \mathbf{e}, t_0$, equations (12.5) and (12.6) enable us to calculate $\mathbf{r}$ as a function of t. Thus the motion of a particle in a Coulomb field is completely known. This motion is in accordance with the empirical laws discovered by Kepler (between 1605 and 1618) for the motion of the planets. □

Systems of material points

To study the collective motion of a system of material points, each marked by a number j ($j = 1, 2, \ldots, N$), one writes the system of equations

$$\frac{d\mathbf{r}_j}{dt} = \mathbf{v}_j \qquad \text{and} \qquad m_j \frac{d\mathbf{v}_j}{dt} = \mathbf{F}_j\,. \tag{12.7}$$

Celestial mechanics, founded by Newton (*Philosophiæ Naturalis Principia Mathematica*, 1687) consists, to a first and very good approximation, of the N-body problem. One considers the celestial bodies which constitute the solar system as material points, and one applies (12.7) with

$$\mathbf{F}_j = C \sum_{\substack{k \\ [k \neq j]}} m_j m_k \frac{\mathbf{r}_k - \mathbf{r}_j}{\|\mathbf{r}_k - \mathbf{r}_j\|^3}\,, \tag{12.8}$$

C being a universal constant (Newton's constant).

It is not necessary to summarize the successes of Newton's theory. The extreme precision of astronomical observations has not been able, until recently, to falsify this theory.

Constraints

Experience has shown that it is legitimate to apply the principles (12.7) of classical mechanics to very diverse systems. However, it is quite often difficult to know the forces $\mathbf{F}_j$, and thus one is obliged to make approximations.

One of the most used approximations consists of considering a system as being *constrained*, that is, of writing constraint equations

$$u_\lambda(t, \mathbf{r}_1, \ldots, \mathbf{r}_N) = 0 \qquad [\lambda = 1, 2, \ldots, p], \tag{12.9}$$

which are assumed to hold for *all* motions of the system. For example, a *rigid body* is a system of material points whose mutual distances are assumed to be constant, in spite of the fact that a real "rigid body" is elastic and susceptible to vibrations. We will write the constraint equations in the form

$$u \equiv \begin{pmatrix} u_1(t, \mathbf{r}_1, \dots, \mathbf{r}_N) \\ \vdots \\ u_p(t, \mathbf{r}_1, \dots, \mathbf{r}_N) \end{pmatrix} = 0 \tag{12.10}$$

by introducing a variable u with values in $\mathbf{R}^p$. Differentiating (12.10) we obtain

$$u_t\, \delta t + \sum_j u_{\mathbf{r}_j}(\delta \mathbf{r}_j) = 0 \tag{12.11}$$

with

$$u_t = \frac{\partial u}{\partial t} \qquad \text{and} \qquad u_{\mathbf{r}_j} = \frac{\partial u}{\partial \mathbf{r}_j} \qquad \text{(notation (2.21))}. \tag{12.12}$$

To treat a dynamical problem involving constraints, one of course always assumes the validity of the equations

$$\frac{d\mathbf{r}_j}{dt} - \mathbf{v}_j = 0\,, \tag{12.13}$$

which, in order to be compatible with (12.10), gives

$$u' \equiv u_t + \sum_j u_{\mathbf{r}_j}(\mathbf{v}_j) = 0\,.^{189} \tag{12.14}$$

On the other hand, in the presence of constraints one rewrites Newton's law in the form

$$m_j \frac{d\mathbf{v}_j}{dt} = \mathbf{F}_j + \mathbf{F}'_j\,, \tag{12.15}$$

$\mathbf{F}_j$ being the "applied" force, which is supposed to be known, and $\mathbf{F}'_j$ being an additional term called the "constraint force", whose existence ensures that the equations are compatible.

In general this compatibility condition is not sufficient to determine the constraint forces. One needs to make diverse empirical hypotheses (for example, laws of friction), whose successes are quite random.

Nevertheless, one of these hypotheses leads to consistent results compatible with a certain idealization of experiments, namely, the hypothesis

[189] When this condition is satisfied, one says that the *velocities are compatible with the constraints*. If one imposes a constraint which is incompatible with the velocities, one produces a shock.

of *ideal constraints*, which is stated in terms of the *principle of virtual work*.

(12.16) If the vectors $\delta \mathbf{r}_j$ satisfy

$$\sum_j u_{\mathbf{r}_j}(\delta \mathbf{r}_j) = 0\,,^{190}$$

then one must have

$$\sum_j \langle \mathbf{F}'_j, \delta \mathbf{r}_j \rangle = 0\,.^{191} \qquad \square$$

(12.17) EXAMPLE: Let us study the principle of virtual work in the case of a *rigid body*. Toward this goal, it is useful to note that permutations of the ordinary Euclidean space $\mathbf{R}^3$ that preserve distances (such permutations are called "displacements") can be considered as belonging to the Lie group of matrices

$$\begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix} \qquad [A \in \mathrm{O}(3),\ \mathbf{b} \in \mathbf{R}^3]$$

which acts on $\mathbf{R}^3$ according to

$$\underline{\begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix}}_{\mathbf{R}^3} (\mathbf{r}) = A(\mathbf{r}) + \mathbf{b}\,.$$

One can show that a system of virtual displacements $\delta \mathbf{r}_j$ *is compatible with the constraints defining a rigid body* if there exists a Z in the Lie algebra $\mathcal{G}$ of G such that

(12.18) $$\delta \mathbf{r}_j = Z_{\mathbf{R}^3}(\mathbf{r}_j) \qquad \forall j \quad \text{(notation (6.11))}.$$

The vector field $Z_{\mathbf{R}^3}$ is called an *infinitesimal displacement*.

Thus the principle of virtual work (12.16) can be written as

(12.19) $$\Theta_I + \Theta_F = 0\,,$$

with

(12.20) $$\begin{cases} \Theta_F(Z) = \sum_j \langle \mathbf{F}_j, Z_{\mathbf{R}^3}(\mathbf{r}_j) \rangle \\ \Theta_I(Z) = -\sum_j \langle m_j \dfrac{d\mathbf{v}_j}{dt}, Z_{\mathbf{R}^3}(\mathbf{r}_j) \rangle \end{cases} \qquad \forall Z \in \mathcal{G}\,.$$

[190] One says that the vectors $\delta \mathbf{r}_j$ are *virtual displacements compatible with the constraints such as they are defined at each instant.*

[191] This expression is called the *virtual work of the constraint forces* performed during the virtual displacement of the system defined by the $\delta \mathbf{r}_j$.

Θ_I and Θ_F are visibly linear maps from $\mathcal{G}$ to $\mathbf{R}$, that is, they are *torsors of the group of displacements*, in the sense of (11.7). Θ_F is called the *torsor of forces applied to the rigid body* and Θ_I the *torsor of inertial forces.*

It is easy to write equation (12.19) in the form

(12.21)
$$\begin{cases} \dfrac{d}{dt}\sum_j m_j \mathbf{v}_j \equiv \sum_j \mathbf{F}_j \\ \dfrac{d}{dt}\sum_j m_j \mathbf{r}_j \times \mathbf{v}_j \equiv \sum_j \mathbf{r}_j \times \mathbf{F}_j \,, \end{cases}^{\dagger}$$

in which appear the classical quantities *linear momentum*, *angular momentum*, *resultant of forces*, and *resultant of torques*. The equations (12.21) are generally considered as *the principles of the mechanics of a rigid body.* □

(12.22) EXAMPLE: In the case of a point moving on a curve or a surface, the principle of virtual work expresses the condition that the constraining forces are always perpendicular to the curve (or surface). Thus the constraint of slipping is *ideal* when it takes place *without friction.* □

Describing forces

When we say that the forces that act on the material points of a dynamical system are "known", we mean that there exist functions Φ_j such that

(12.23)
$$\mathbf{F}_j \equiv \Phi_j(t, \mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N)\,.$$

In other words, each force may depend on *time*, *position*, and *velocity* of each of the material points.

It can happen of course that some of these variables do not appear explicitly in the description of these forces.

For example, in the N-body problem (12.8) the forces depend neither on time nor on velocity.

(12.24) A priori other expressions for the forces are possible. For example, in the case of *hereditary* systems one assumes that each force depends (by

---

†Editors' note: For the first equation use $Z = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$ with arbitrary $b \in \mathbf{R}^3$, for the second use $Z = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$ with arbitrary A in the Lie algebra of O(3).

means of an integral expression) on the whole past motion of the system. However, such hypotheses are rather artificial, destined to give a phenomenological description of an imperfectly known system.[192] Moreover, their consistency is not automatically assured.

In fact, the principles of mechanics *implicitly assume* the existence of expressions for the forces of the form (12.23). We will make this hypothesis from now on.

The evolution space

Under the above hypothesis, the differential equations of mechanics, also called the *equations of motion*, can be written as

$$\frac{d\mathbf{r}}{dt} - \mathbf{v}_j \equiv 0\,, \qquad m_j \frac{d\mathbf{v}_j}{dt} - \Phi_j(t, \mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{v}_1, \dots, \mathbf{v}_N) \equiv 0\,. \tag{12.25}$$

It is easy to bring these equations into the standard form of §5. They tell us that the $\delta\mathbf{r}_j$ and $\delta\mathbf{v}_j$ are linear functions of δt with coefficients depending on t, the $\mathbf{r}_k$ and the $\mathbf{v}_k$. It thus suffices to consider (Fig. 12.I) the space of $(2N+1)$-tuples [193]

$$y \equiv \begin{bmatrix} t \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} \tag{12.26}$$

and the field $[y \mapsto \mathcal{E}]$ of 1-dimensional vector subspaces defined by

$$\{\delta y \in \mathcal{E}\} \Leftrightarrow \left\{ \begin{array}{ll} \delta\mathbf{r}_j - \mathbf{v}_j\,\delta t = 0 & \\ m_j\,\delta\mathbf{v}_j - \Phi_j(t, \mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{v}_1, \dots, \mathbf{v}_N)\,\delta t = 0 & \forall j \end{array} \right\} \tag{12.27}$$

in order that the solutions of the equations of motion be the *lines of force* of the field $[y \mapsto \mathcal{E}]$ (see (5.10)).

(12.28) *Henceforth we assume* that the expressions Φ_j describing the forces are *differentiable* (in the sense of §1). Implicit in this assumption (which

[192] As for example a box containing a gyroscope whose existence we are ignorant of, or a controlled system if we do not know the nature of the controlling mechanism.

[193] y is called an *initial condition* of the system.

does not seem to be contradicted by experiments) is that the set V on which the field $y \mapsto \mathcal{E}$ is defined, is an *open* subset of the vector space

$$\mathbf{R} \times \underbrace{\mathbf{R}^3 \times \cdots \times \mathbf{R}^3}_{N} \times \underbrace{\mathbf{R}^3 \times \cdots \times \mathbf{R}^3}_{N},$$

to which the variable y (12.26) belongs a priori.

(12.29) We conclude that V possesses the structure of a manifold of dimension $6N + 1$. We will say that V is the *evolution space* of the system. □

Furthermore, the hypothesis of differentiability of the Φ_j shows that the field $y \mapsto \mathcal{E}$ itself is *differentiable* in the sense of (5.2). We thus can apply the results of §5 and especially (5.5), (5.10), and (5.9).

(12.30) The field $y \mapsto \mathcal{E}$ (12.27) gives the evolution space V the structure of a *foliated manifold*. Through each point y of V passes a *line of force*,[194] and the leaf through y is the *maximal connected line of force* passing through y. Henceforth we shall consider the leaf through y as the *motion* of the system defined by the initial condition y. □

Phase spaces and the space of motions

(12.31) Since the variable t appears in the expression (12.26) for y, there exists a map $[y \mapsto t]$ from the evolution space V to the real line of time (Fig. 12.I).

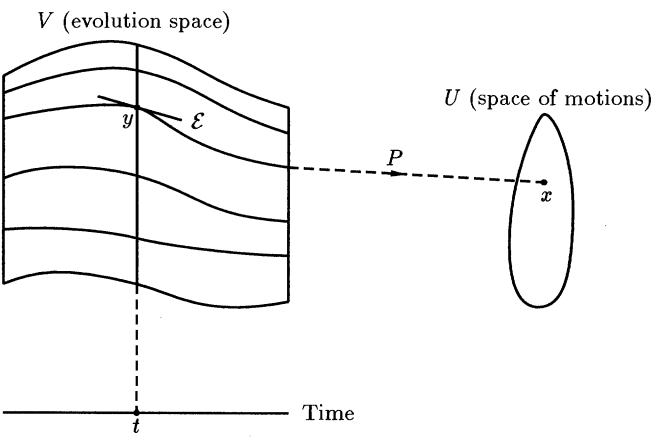


Fig 12.I

The subset of the evolution space V defined by the equation $t = t_0$ is called the *phase space at time* t_0. It can be immediately verified that

[194] That is, an *integral manifold* of dimension 1.

the map $[y \mapsto t]$ is *differentiable* (1.17) and that the phase spaces are *submanifolds* of V of dimension $6N$. In particular, it is clear that the map

$$\begin{bmatrix} m_1 x_1' \\ m_1 y_1' \\ m_1 z_1' \\ \vdots \\ m_N x_N' \\ m_N y_N' \\ m_N z_N' \\ x_1 \\ y_1 \\ z_1 \\ \vdots \\ x_N \\ y_N \\ z_N \end{bmatrix} \longmapsto \begin{bmatrix} t_0 \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} \tag{12.32}$$

is a *chart* for the phase space at time t_0.[195]

(12.33) Although it is a standard object of study, we will *not* use the "nontemporal" phase space constructed from the previously defined evolution space by neglecting to specify the time t_0.[196]

(12.34) If we take $\delta t = 0$ in the equation $\delta y \in \mathcal{E}$ (see (12.27)), we find immediately that $\delta \mathbf{r}_j = 0$ and $\delta \mathbf{v}_j = 0$ (because the masses are assumed to be nonzero) and thus $\delta y = 0$. Consequently, $\mathcal{E} \cap \ker\left(\frac{\partial t}{\partial y}\right) = \{0\}$, that is, each phase space at time $t = t_0$ is *transverse* to the foliation of V (definition (5.14)). On the other hand, we know that this foliation equips V with the structure of a manifold of dimension 1, whose connected components are just the leaves [(5.7) and (5.8)]. The transversality condition expresses the fact that $\frac{\partial t}{\partial y}$ is not zero on $\mathcal{E}$, the tangent space of V in this new structure. In other words, $\frac{\partial t}{\partial y}$ *does not vanish on this* 1*-dimensional manifold*. From this one can deduce that the map $y \mapsto t$ is a *diffeomorphism* on each connected component and thus that each leaf possesses a

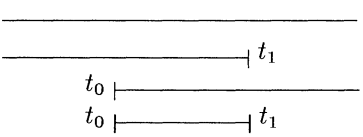
[195]The variables x_j, y_j, z_j, respectively x_j', y_j', z_j', are the Cartesian coordinates of $\mathbf{r}_j$, respectively $\mathbf{v}_j$. The introduction of the masses m_j, for the moment purely whimsical, will be useful in (12.103).

[196]A more precise definition is that one *identifies* two initial conditions y_1 and y_2 when they differ "only by their date", just as one usually identifies points of space at different times. However, this identification is *incompatible with Galilean invariance*. One can make this identification only by relating it to an arbitrarily chosen reference frame — the earth, for example.

chart $t \mapsto y$. One can also show that the domain of the map $t \mapsto y$ is an open *interval* bounded or not (from either above or below).

This result can also be established by a direct application of the general theorem (3.2) on differential equations. It allows us to classify the motions of a mechanical system into four types:

(12.35)



eternal
finishing
beginning
bounded in time.

(12.36) All four cases are possible, as can be seen in the example of the (attracting) Coulomb field: the linear motions, for which the vector **l** (12.3) is zero, give rise to collisions of the moving point with the center of attraction.

(12.37) Furthermore, we see that each phase space at time $t = t_0$ intersects a leaf in at most one point and thus it constitutes a *transverse section* of the foliation of V (5.15). In other words, the foliation of V is *sectionable* (5.15), since through each point y of V passes a phase space at some time t_0. Applying (5.16) one arrives at the following result (Fig. 12.I).

(12.38) The space U of motions of a dynamical system admits the *structure of a manifold; it is the quotient of the evolution space V by its foliation.* If $P(y)$ denotes the motion defined by the initial condition y, then P is a *differentiable map* from V onto U. The restriction of P to a phase space at time $t = t_0$ is a *diffeomorphism* from this phase space *to* U.[197]

(12.39) U is a manifold of *dimension* $6N$. If we choose a time t_0 and if we denote by F_{t_0} the chart (12.32) of the phase space at time t_0, then $P \circ F_{t_0}$ is a *chart of the space of motions* (5.16). By varying t_0 we evidently obtain an *atlas* of U.

The Lagrange 2-form

Let us return to the evolution space V and let us define a priori

(12.40)
$$\sigma(\delta y)(\delta' y) = \sum_j \big(\langle m_j\, \delta \mathbf{v}_j - \mathbf{F}_j\, \delta t, \delta' \mathbf{r}_j - \mathbf{v}_j\, \delta' t \rangle - \langle m_j\, \delta' \mathbf{v}_j - \mathbf{F}_j\, \delta' t, \delta \mathbf{r}_j - \mathbf{v}_j \delta t \rangle \big) .$$

[197] But not necessarily *onto* U. One misses motions which have *already finished* or *have not yet begun* at time t_0.

It is clear that σ is a 2-*form* on V (definition (4.26)), and that the differential equation of motion $\delta y \in \mathcal{E}$ (notation (12.27)) implies

$$\sigma(\delta y)(\delta' y) = 0 \qquad \forall\, \delta' y\,, \tag{12.41}$$

which, using (4.2), can also be written as

$$\sigma(\delta y) = 0\,, \tag{12.42}$$

or as

$$\delta y \in \ker(\sigma)\,. \tag{12.43}$$

The 2-form σ was introduced by Lagrange in his study of celestial mechanics in 1808,[198] in a different language of course. Its explicit expression (12.40) can be found in Gallissot.[199]

For reasons which will be given later on, we generalize formula (12.40) by *choosing* vectors $\mathbf{B}_j$, which we assume to be differentiable functions of y on V, and by defining

$$\mathbf{E}_j \equiv \mathbf{F}_j + \mathbf{B}_j \times \mathbf{v}_j \qquad \forall j \tag{12.44}$$



and

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;\equiv\; & \textstyle\sum_j \langle m_j\, \delta \mathbf{v}_j - \mathbf{E}_j\, \delta t,\, \delta' \mathbf{r}_j - \mathbf{v}_j\, \delta' t\rangle \\ & - \textstyle\sum_j \langle m_j\, \delta' \mathbf{v}_j - \mathbf{E}_j\, \delta' t,\, \delta \mathbf{r}_j - \mathbf{v}_j\, \delta t\rangle + \sum_j \langle \mathbf{B}_j,\, \delta \mathbf{r}_j \times \delta' \mathbf{r}_j\rangle\,.\end{aligned} \tag{12.45}$$

This σ is still a 2-form, which we call the *Lagrange form*, and the field $y \mapsto \sigma$ is differentiable on the evolution space V.

Using some elementary transformations one finds that (12.45) is equivalent to

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;\equiv\; & \textstyle\sum_j \langle m_j\, \delta \mathbf{v}_j - [\mathbf{E}_j - \mathbf{B}_j \times \mathbf{v}_j]\, \delta t,\, \delta' \mathbf{r}_j - \mathbf{v}_j\, \delta' t\rangle \\ & - \textstyle\sum_j \langle m_j\, \delta' \mathbf{v}_j - \mathbf{E}_j\, \delta' t + \mathbf{B}_j \times \delta' \mathbf{r}_j,\, \delta \mathbf{r}_j - \mathbf{v}_j\, \delta t\rangle\,,\end{aligned} \tag{12.46}$$

[200]

which shows that the equation $\sigma(\delta y)(\delta y') = 0$ $[\forall\, \delta' y]$ can be written as

$$\begin{cases} m_j\, \delta \mathbf{v}_j - \mathbf{F}_j\, \delta t = 0 \\ \delta \mathbf{r}_j - \mathbf{v}_j\, \delta t = 0 \end{cases} \qquad \forall j\,. \tag{12.47}$$

(12.48) It follows that *the equations of motion can be written as*

[198] See Lagrange [22] and [23].

[199] See Gallissot [15].

[200] Note that the antisymmetry of σ is not obvious anymore in the expression (12.46).



$$\sigma(\delta y) = 0$$

and that *the vector space* $\mathcal{E}$ *of* (12.27) *equals* $\ker(\sigma)$. □

(12.49) This result is independent of the choice of the $\mathbf{B}_j$. It applies in particular to the case $\mathbf{B}_j \equiv 0$ (and hence $\mathbf{E}_j \equiv \mathbf{F}_j$), where the Lagrange form reduces to the expression given in (12.40).

The Lagrange form for constrained systems

Let us consider a system subjected to constraints, for which we use the notation introduced in (12.9) to (12.15). We leave it to the reader to establish the following algebraic results in which the Lagrange form σ is involved.

(12.50) THEOREM: Assume that

◇ the rank of the linear map $[u_{\mathbf{r}_1}\ u_{\mathbf{r}_2}\ \dots\ u_{\mathbf{r}_N}]$ equals p.[201]

Then

a) The equation $u_t\,\delta t + \sum_j u_{\mathbf{r}_j}(\delta \mathbf{r}_j) = 0$ (12.11) can be written as $\delta y \in H$, where H is a vector space of dimension $6N + 1 - p$.

b) The vector space $\mathrm{orth}(H)$ is isotropic.[202]

c) The compatibility condition $u' = 0$ (12.14) can be written as $\ker(\sigma) \subset H$ or (using b)) as H is coisotropic.[203]

d) $H' = \ker\left(\frac{\partial u'}{\partial y}\right)$ is a vector space of dimension $6N + 1 - p$ and the dimension of $K = H \cap H'$ equals $6N + 1 - 2p$.

e) Taking into account the compatibility condition $u' \equiv 0$, the *principal of virtual work* (12.16) and the equations *defining the velocities* $\frac{d\mathbf{r}_j}{dt} - \mathbf{v}_j = 0$ (12.13) can be written as $\delta y \in \mathrm{orth}(H)$.

f) $H' \cap \mathrm{orth}(H) = K \cap \mathrm{orth}(K)$, and the dimension of this vector space, which we will call $\mathcal{E}'$, equals 1. Its intersection with $\ker\left(\frac{\partial t}{\partial y}\right)$ is $\{0\}$.[204]

□

[201] There are cases in which one can realize this condition by suppressing certain constraint equations that are a consequence of others (for example N points forming a rigid body when $N \geq 5$). In the opposite case, when this condition is violated, one has a system which is susceptible to jamming (for example, a rod of length ℓ whose extremities move on a circle of diameter ℓ).

[202] Definitions (8.2) and (8.4) applied to the Lagrange form σ (12.45).

[203] That is, $\mathrm{orth}(H) \subset H$ (definition (8.4)).

[204] The results of (f) make use of the fact that the masses m_j all have *the same sign*.

(12.51) The above results show in what way the consequences of the principle of virtual work are coherent under condition $\diamondsuit$, namely, the constraint equations $u = 0$ and $u' = 0$ determine a submanifold V' of dimension $6N + 1 - 2p$ of the evolution space V, and V' has K as its tangent space (theorem (1.40)).

Let us denote by σ' the 2-form induced on V' by σ:

(12.52) $$\sigma'(\delta y)(\delta' y) = \sigma(\delta y)(\delta' y) \qquad \text{if } y \in V',\ \delta y \in K, \text{ and } \delta' y \in K\ .$$

The formula $\mathcal{E}' = K \cap \operatorname{orth}(K)$ shows that $\mathcal{E}' = \ker(\sigma')$ (see (9.30)). Since the equations of motion with ideal constraints can be written as $\delta y \in \mathcal{E}'$ and since $\dim \mathcal{E}' = 1$ and $\dim(\mathcal{E}' \cap \ker\left(\frac{\partial t}{\partial y}\right)) = 0$, we can apply to V', σ', and $\mathcal{E}'$ the results established above for V, σ, and $\mathcal{E}$, respectively. This means in particular that the motions, which are the lines of force of the field $y \mapsto \mathcal{E}'$, belong to one of the types (12.35), that the *phase space* at time $t = t_0$ is a transverse section, and that the space of (constrained) motions admits the structure of a manifold (of dimension $6N - 2p$).

(12.53) This is why we will say that V' is *the evolution space of the constrained system* and that σ' is its *Lagrange form*.

(12.54) Applying the above discussion to the motion of a rigid body, one finds that the dimension of the evolution space is 13 and that the dimension of the space of motions is 12.[205]

(12.55) Besides the constraints we have studied up until now — which are called *holonomic* — one encounters constraints (especially in problems such as rolling without friction) which are expressed by the condition

$$u_t\, \delta t + \sum_j u_{\mathbf{r}_j}\, \delta \mathbf{r}_j = 0\ .$$

This condition, called a *nonholonomic constraint*, is analogous to (12.14) but without the hypothesis that u_t and the $u_{\mathbf{r}_j}$ are partial derivatives of a variable u; one only assumes them to be differentiable functions of t and the $\mathbf{r}_k$. On this kind of constraint we also impose the rules (12.13) to (12.16). It turns out that *theorem* (12.50) *is still valid*. Thus the motion of a system subjected to *ideal nonholonomic constraints* is still defined by the equation $\delta y \in \mathcal{E}'$. However, in general $\mathcal{E}'$ is no longer the kernel of a 2-form σ' induced by σ.

[205] These numbers become 11 and 10 if the points of the rigid body lie on a line.

Changing the reference frame

(12.56) An *event* X or "a point of space-time" is characterized by a pair $\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}$ consisting of the *position* $\mathbf{r}$ of the event, which belongs to the Euclidean space $\mathbf{R}^3$, and its *date* t, which is a number. In fact, the same event might be denoted by another pair $\begin{pmatrix}\mathbf{r}^*\\ t^*\end{pmatrix}$ if for example we change the choice of the origin. We will call a map $\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix} \mapsto X$ a *reference frame*. If $\mathcal{R}$ and $\mathcal{R}^*$ are two *reference frames*, we can write

$$X \equiv \mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix} \equiv \mathcal{R}^*\begin{pmatrix}\mathbf{r}^*\\ t^*\end{pmatrix}. \tag{12.57}$$

Thus there exists a formula for the *change of reference frame*, which can be written as

$$\begin{pmatrix}\mathbf{r}^*\\ t^*\end{pmatrix} \equiv (\mathcal{R}^*)^{-1} \circ \mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}. \tag{12.58}$$

In classical mechanics it is assumed to be possible to choose, once and for all, units of length and time as well as to orient space and time. These assumptions are expressed mathematically by the fact that the formula for the change from one reference frame to another is given by

$$\begin{aligned}\mathbf{r}^* &\equiv A(\mathbf{r}+\mathbf{R})\\ t^* &\equiv t + t_0\,,\end{aligned} \tag{12.59}$$

where t_0 is a number, $\mathbf{R}$ an element of $\mathbf{R}^3$, and A a *rotation*, that is, an element of the group SO(3) (notation (6.64), (6.65)).

We assume that t_0 is *fixed*, but that A and $\mathbf{R}$ *may depend on* t; more precisely, we assume that the maps $t \mapsto A$ and $t \mapsto \mathbf{R}$ are differentiable.

As $A \in \mathrm{SO}(3)$, we have the identities

$$\overline{A} \cdot A \equiv 1 \qquad \text{(notation (6.53))} \tag{12.60}$$

and

$$A(\mathbf{x}) \times A(\mathbf{y}) \equiv A(\mathbf{x} \times \mathbf{y}) \qquad (\mathbf{x}, \mathbf{y} \in \mathbf{R}^3)\,, \tag{12.61}$$

which can also be written as

$$j(A(\mathbf{x})) \equiv A \cdot j(\mathbf{x}) \cdot A^{-1}\,. \tag{12.62}$$

For this, recall that we denoted by $j(\mathbf{x})$ the operator "vector product by $\mathbf{x}$" defined by

$$j(\mathbf{x})(\mathbf{y}) \equiv \mathbf{x} \times \mathbf{y}\,, \qquad \text{or} \qquad j\begin{pmatrix}p\\ q\\ r\end{pmatrix} \equiv \begin{pmatrix}0 & -r & q\\ r & 0 & -p\\ -q & p & 0\end{pmatrix}. \tag{12.63}$$

By differentiating (12.60) one can show that there exists a vector $\mathbf{\Omega}$ such that

$$\frac{dA}{dt} \equiv -A \cdot j(\mathbf{\Omega}) . \tag{12.64}$$

The vector $\mathbf{\Omega}$, which might also depend on t, is called the *instantaneous axis of rotation* of the new reference frame with respect to the old one.

We now apply the formulæ (12.59) to the points of the system to obtain

$$\begin{aligned} t^* &\equiv t + t_0 \\ \mathbf{r}_j^* &\equiv A(\mathbf{r}_j + \mathbf{R}) . \end{aligned} \tag{12.65}$$

If we introduce the *drag velocity* $\mathbf{U}_j$ and the *drag acceleration* $\mathbf{\Gamma}_j$ of the j^{th} point, which are defined by

$$\begin{aligned} \mathbf{U}_j &\equiv \mathbf{\Omega} \times [\mathbf{r}_j + \mathbf{R}] - \frac{d\mathbf{R}}{dt} \\ \mathbf{\Gamma}_j &\equiv \left[j(\frac{d\mathbf{\Omega}}{dt}) + j(\mathbf{\Omega})^2 \right] (\mathbf{r}_j + \mathbf{R}) - \frac{d^2\mathbf{R}}{dt^2} , \end{aligned} \tag{12.66}$$

and if we define

$$\begin{aligned} \mathbf{v}_j^* &\equiv A(\mathbf{v}_j - \mathbf{U}_j) \\ \mathbf{B}_j^* &\equiv A(\mathbf{B}_j + 2m_j \mathbf{\Omega}) \\ \mathbf{E}_j^* &\equiv A(\mathbf{E}_j - \mathbf{B}_j \times \mathbf{U}_j - m_j \mathbf{\Gamma}_j) , \end{aligned} \tag{12.67}$$

then one can verify that the Lagrange form σ in the new variables is simply obtained by making the substitutions

$$t \mapsto t^* , \quad \mathbf{r}_j \mapsto \mathbf{r}_j^* , \quad \mathbf{v}_j \mapsto \mathbf{v}_j^* , \quad \mathbf{E}_j \mapsto \mathbf{E}_j^* , \quad \mathbf{B}_j \mapsto \mathbf{B}_j^* \tag{12.68}$$

in expression (12.45) of σ (the masses m_j being preserved).

(12.69) It is to be noted that the values of $\mathbf{v}_j^*$, $\mathbf{B}_j^*$, and $\mathbf{E}_j^*$ given in (12.67) are *the only ones* which allows us to complete (12.65) in such a way that we have this property.

The principle of Galilean relativity

(12.70) Now there do exist physically privileged reference frames called *inertial reference frames*. When one passes from one inertial reference frame to another, the instantaneous axis of rotation and the drag acceleration are 0. From this one can easily deduce that the change of inertial reference frame (12.59) can be written as

$$\text{(12.71)} \qquad \begin{pmatrix} \mathbf{r}^* \\ t^* \\ 1 \end{pmatrix} \equiv \begin{bmatrix} A & \mathbf{b} & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \qquad \begin{array}{l} A \in \mathrm{SO}(3)\,, \quad \mathbf{b} \in \mathbf{R}^3\,, \\ \mathbf{c} \in \mathbf{R}^3\,, \quad e \in \mathbf{R}\,, \end{array}$$

A, $\mathbf{b}$, $\mathbf{c}$, e being *constant*. Conversely, if one passes from an inertial reference frame to another reference frame using the above formula, the second frame is also inertial.

In classical mechanics, one generally admits, under the name of the *principle of Galilean relativity*, the following hypothesis:

(12.72) If V is the evolution space of an *isolated* dynamical system,[206] then the field $y \mapsto \sigma$, which defines the Lagrange form on V, is *the same in all inertial reference frames.* □

Let us denote by G the set of matrices a considered in (12.71), namely

$$\text{(12.73)} \qquad a \equiv \begin{bmatrix} A & \mathbf{b} & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l} A \in \mathrm{SO}(3)\,, \quad \mathbf{b} \in \mathbf{R}^3\,, \\ \mathbf{c} \in \mathbf{R}^3\,, \quad e \in \mathbf{R}\,. \end{array}$$



It is easy to verify that these matrices form a Lie group which is homeomorphic to $\mathrm{SO}(3) \times \mathbf{R}^7$ (and thus is connected and of dimension 10). This group is called the *Galilei group*. Its Lie algebra $\mathcal{G}$ is the set of matrices

$$\text{(12.74)} \qquad Z \equiv \begin{bmatrix} j(\boldsymbol{\omega}) & \boldsymbol{\beta} & \boldsymbol{\gamma} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{array}{ll} \boldsymbol{\omega} \in \mathbf{R}^3\,, & \boldsymbol{\beta} \in \mathbf{R}^3 \\ \boldsymbol{\gamma} \in \mathbf{R}^3\,, & \varepsilon \in \mathbf{R}\,. \end{array}$$

(12.75) The Galilei group is a *Lie subgroup* (6.31) of the group of matrices $\mathrm{Gl}(\mathbf{R}^3 \times \mathbf{R}^2)$ (criterion (6.33) can be applied).

Applying formulæ (12.66) and (12.67) for the change of reference frame, one can see that the principle of Galilean relativity can be formulated in the following way.

[206] This means that the action of the rest of the universe on this system is *negligible* (the solar system is an example of an isolated system). This concept suggests some delicate questions (think of *Mach's principle*).

(12.76) Let a be an element of the Galilei group (notation (12.73)) and let V be the evolution space of an isolated dynamical system whose coordinates are given with respect to an inertial reference frame. If we define

$$\diamondsuit \qquad y \equiv \begin{bmatrix} t \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} \qquad \text{and} \qquad \underline{a}_V(y) \equiv y^* \equiv \begin{bmatrix} t^* \\ \mathbf{r}_1^* \\ \vdots \\ \mathbf{r}_N^* \\ \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_N^* \end{bmatrix}$$

with

$$t^* \equiv t + e\,, \qquad \mathbf{r}_j^* \equiv A(\mathbf{r}_j) + \mathbf{b}t + \mathbf{c}\,, \qquad \mathbf{v}_j^* \equiv A(\mathbf{v}_j) + \mathbf{b} \qquad \forall j\,,$$

then $\underline{a}_V$ is a *diffeomorphism* from V to V. The field $\Phi\,[y \mapsto \sigma]$ defining the Lagrange 2-form satisfies

$$\clubsuit \qquad \Phi(y)(\delta y)(\delta' y) \equiv \Phi(y^*)(\delta y^*)(\delta' y^*)\,,$$

which can also be written as [207]

$$\spadesuit \qquad \begin{aligned} &[y \mapsto \mathbf{E}_j](y^*) \equiv A(\mathbf{E}_j) + A(\mathbf{B}_j) \times \mathbf{b} \\ &[y \mapsto \mathbf{B}_j](y^*) \equiv A(\mathbf{B}_j)\,. \end{aligned}$$

□

(12.77) One can easily verify that the map $a \mapsto \underline{a}_V$ (12.76$\diamondsuit$) defines an *action* of the Galilei group on the evolution space V, in the sense of (6.4).

(12.78) For a constrained isolated system we complement the principle of Galilean relativity by assuming that, in addition to the Lagrange form, the constraint equations are also the same in all inertial reference frames. If the constraints are holonomic, it is easy to see that (12.76) and (12.77) still apply to the evolution space V' of the constrained isolated system as well as to its Lagrange form σ' (12.53).

[207] Here $[y \mapsto \mathbf{E}_j](y^*)$ obviously denotes the variable $\Phi_j(y^*)$, where Φ_j is the map $[y \mapsto \mathbf{E}_j]$ which associates $\mathbf{E}_j$ to y (see (2.8)).

Maxwell's principle

(12.79) MAXWELL'S PRINCIPLE:[208] The Lagrange form σ of a dynamical system has *zero exterior derivative* on the evolution space: $\mathrm{d}\sigma \equiv 0$.

If we substitute definition (12.45) of the form σ into definition (4.32) of the exterior derivative and expand it, then after some computations, we obtain

$$\begin{array}{lll} \dfrac{\partial \mathbf{E}_j}{\partial \mathbf{v}_k} \equiv 0 & \dfrac{\partial \mathbf{B}_j}{\partial \mathbf{v}_k} \equiv 0 & \forall\, j,k \\ \dfrac{\overline{\partial \mathbf{E}_k}}{\partial \mathbf{r}_j} - \dfrac{\partial \mathbf{E}_j}{\partial \mathbf{r}_k} \equiv 0 & \dfrac{\partial \mathbf{B}_j}{\partial \mathbf{r}_k} \equiv 0 & \forall\, j \neq k \\ \operatorname{curl} \mathbf{E}_k + \dfrac{\partial \mathbf{B}_k}{\partial t} \equiv 0 & \operatorname{div} \mathbf{B}_k \equiv 0 & \forall\, k\,.^{209} \end{array} \tag{12.80}$$

(12.81) The system (12.80) is *equivalent* to Maxwell's principle $\mathrm{d}\sigma \equiv 0$. If (12.80) is satisfied in a single reference frame, it will be satisfied in *any other reference frame* (even if it is not an inertial reference frame).

(12.82) In the definition of the Lagrange form σ, we have arbitrarily chosen vectors $\mathbf{B}_j$ from which we defined vectors $\mathbf{E}_j$ by

$$\mathbf{E}_j \equiv \mathbf{F}_j + \mathbf{B}_j \times \mathbf{v}_j\,,$$

assuming that the forces $\mathbf{F}_j$ were known.

Thus we can, if we want to, formulate Maxwell's principle as follows:

"The vectors $\mathbf{B}_j$ can be chosen so that $\mathrm{d}\sigma \equiv 0$".

It is easy to show that in this case the choice of the $\mathbf{B}_j$ is unique: Maxwell's principle, if it applies, removes the arbitrariness in the definition of the Lagrange form which existed up until now.

One can easily verify that Maxwell's principle holds in the following examples.

---

[208] We give it this name on account of (12.86), but also because one of the equations of (12.80), namely, $\dfrac{\overline{\partial \mathbf{E}_k}}{\partial \mathbf{r}_j} - \dfrac{\partial \mathbf{E}_j}{\partial \mathbf{r}_k} = 0$, is sometimes called *Maxwell's reciprocity law*.

[209] The *curl* of the vector $\mathbf{E}_k$ can be defined by $\dfrac{\partial \mathbf{E}_k}{\partial \mathbf{r}_k} - \dfrac{\overline{\partial \mathbf{E}_k}}{\partial \mathbf{r}_k} \equiv j(\operatorname{curl} \mathbf{E}_k)$, and the *divergence* of $\mathbf{B}_k$ can be defined by $\operatorname{div} \mathbf{B}_k \equiv \operatorname{Tr}\left(\dfrac{\partial \mathbf{B}_k}{\partial \mathbf{r}_k}\right)$.

EXAMPLE: The N-body problem (12.8), given in an inertial frame by

$$\mathbf{B}_j \equiv 0\,, \qquad \mathbf{E}_j \equiv C \sum_{\substack{k \\ [k\neq j]}} m_j\, m_k \frac{\mathbf{r}_k - \mathbf{r}_j}{\|\mathbf{r}_k - \mathbf{r}_j\|^3}\,. \tag{12.83}$$

□

EXAMPLE: A *mass point* in a vacuum under the influence of *gravity*. In a reference frame fixed to the earth this is described by

$$\mathbf{E} \equiv m\,\mathbf{g}\,, \qquad \mathbf{B} \equiv 2m\,\boldsymbol{\Omega}\,, \tag{12.84}$$

where $\mathbf{g}$ is the *acceleration due to gravity* and $\boldsymbol{\Omega}$ is the *rotation vector of the earth*. □

EXAMPLE: *Charged particles in an exterior electromagnetic field* for which we take

$$\begin{aligned} \mathbf{E}_j &\equiv q_j\,\mathbf{E}(t,\mathbf{r}_j) + \sum_{\substack{k \\ [k\neq j]}} q_j\, q_k \frac{\mathbf{r}_j - \mathbf{r}_k}{\|\mathbf{r}_j - \mathbf{r}_k\|^3} \\ \mathbf{B}_j &\equiv q_j\,\mathbf{B}(t,\mathbf{r}_j)\,. \end{aligned} \tag{12.85}$$

Here q_j is the electric charge of the j^{th} particle and $\mathbf{E}(t,\mathbf{r})$ and $\mathbf{B}(t,\mathbf{r})$ are the *electric field* and the *magnetic induction* at time t and position $\mathbf{r}$, expressed in electrostatic units. Applying the formulæ (12.80) gives *Maxwell's equations*

$$\operatorname{curl}\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} \equiv 0 \qquad \text{and} \qquad \operatorname{div}\mathbf{B} \equiv 0\,, \tag{12.86}$$

which are indeed satisfied by the fields. They express the *law of induction* and *the absence of magnetic charges* (magnetic "monopoles").

(12.87) In (12.85) the electrostatic forces between the particles are taken into account. However, Maxwell's principle obliges us to neglect small terms such as magnetic interactions. But note that this kind of term cannot be written without introducing the constant c, the speed of light, into the equations, that is, without mixing classical mechanics and special relativity. In classical mechanics charged particles are *subjected* to a magnetic field, *they do not participate in its creation.*

The formulæ (12.85) are valid in an inertial reference frame. If we change the reference frame, the formulæ (12.67) show that the vectors $\mathbf{B}$ and $\mathbf{E}$ have to be replaced by

$$\begin{aligned} \mathbf{B}^* &\equiv A(\mathbf{B}) \\ \mathbf{E}^* &\equiv A(\mathbf{E} - \mathbf{B}\times\mathbf{U})\,, \end{aligned} \tag{12.88}$$

where $\mathbf{U}$ is the drag velocity. Consequently, a purely magnetic field ($\mathbf{B} \neq 0$, $\mathbf{E} \equiv 0$) in a given inertial reference frame appears to be *accompanied by an electric field* in another inertial frame which is moving with respect to the first. This fact is dictated by rational mechanics and is *not* a relativistic effect. □

(12.89) *On the contrary*, Maxwell's principle *does not apply* to problems such as the motion of a projectile when we take into account air resistance, which is a function of velocity. *But* we know that this is only an empirical approximation, meant to replace a detailed study of the mechanics of the atmosphere itself. □

(12.90) By induction from these few examples, we accept henceforth Maxwell's principle as a new *law of mechanics*.

Potentials and the variational formalism

One can verify that the equations (12.80) are *locally* equivalent to the existence of a scalar field $v(t, \mathbf{r}_1, \ldots, \mathbf{r}_N)$ and vectors $\mathbf{A}_j(t, \mathbf{r}_j)$, called *potential* and *vector potential*, respectively, such that

$$\begin{aligned} \mathbf{E}_j &\equiv -\overline{\frac{\partial v}{\partial \mathbf{r}_j}} - \frac{\partial \mathbf{A}_j}{\partial t} \\ \mathbf{B}_j &\equiv \operatorname{curl} \mathbf{A}_j \,. \end{aligned} \tag{12.91}$$

We now consider the *Lagrangian* (notation of §7)

$$\lambda \equiv \sum_j \left\{ \tfrac{1}{2} m_j \|\mathbf{r}'_j\|^2 + \langle \mathbf{A}_j, \mathbf{r}'_j \rangle \right\} - v + \frac{\partial \varphi}{\partial t} + \sum_j \frac{\partial \varphi}{\partial \mathbf{r}_j} \cdot \mathbf{r}'_j \,, \tag{12.92}$$

where $\mathbf{v}_j \equiv \mathbf{r}'_j \equiv \frac{d\mathbf{r}_j}{dt}$ and where φ is an arbitrary differentiable function of $t, \mathbf{r}_1, \ldots, \mathbf{r}_N$ called a *gauge*. The *Cartan form* ϖ of the corresponding variational problem (see (7.29)) is given by

$$\varpi(\delta y) \equiv \sum_j \langle \mathbf{p}_j, \delta \mathbf{r}_j \rangle - h \, \delta t \,, \tag{12.93}$$

with

$$\mathbf{p}_j \equiv m_j \mathbf{v}_j + \mathbf{A}_j + \overline{\frac{\partial \varphi}{\partial \mathbf{r}_j}}$$

and

$$h \equiv \tfrac{1}{2} \sum_j m_j \|\mathbf{v}_j\|^2 + v - \frac{\partial \varphi}{\partial t} \,.$$

One can check immediately that

(12.94) $$\sigma \equiv \mathrm{d}\varpi\,,$$

that is, ϖ is a *potential* of σ (9.14). Conversely, the *most general* Lagrangian

(12.95) $$\lambda = \Lambda(t, \mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{r}'_1, \ldots, \mathbf{r}'_N)$$

whose Cartan form is a potential of σ, is the one given by (12.92).

Thus the *equations of motion* $\delta y \in \ker(\sigma)$ (see (12.48)) coincide with the Euler-Lagrange equations for the Lagrangian λ (theorem (7.29)). Consequently, the motions of the system are *extremals* of the integral

(12.96) $$\int_{t_0}^{t_1} \lambda\, dt\,,$$

called the *Hamiltonian action*.

(12.97) We shall not dwell on the practical applications of this fact (called "Hamilton's principle"), as they have been developed at length in treatises on mechanics (one should think of the use of the Euler-Lagrange equations (7.6), or of the equations deduced from Jacobi's equation (7.19) or from Hamilton's canonical equations (7.23)).

(12.98) It is tempting to take Hamilton's principle as the foundation of mechanics,[211] especially because it dispenses with having to *formulate* Maxwell's principle.

However, there are two fundamental drawbacks to such an approach.

a) The existence of a potential and a vector potential in the Lagrangian is guaranteed locally but not globally.[212]
b) These potentials are not completely determined because the gauge φ is arbitrary.[213]

Furthermore, we will see that the symmetry of the *Lagrangian* λ is often less than the symmetry of the *Lagrange form* σ (see (12.136)).



This is why we will not use any variational principle from now on.

[210] h is called the *Hamiltonian* of the system (see also (7.21)).

[211] See for example Landau and Lifschitz [24].

[212] In particular, we will show that in the case of a *particle with spin*, there is no globally defined potential, whereas Maxwell's principle remains valid (see (18.56)).

[213] One can postulate that the *linear momentum* $\mathbf{p}_j$ of a particle does not depend on the position of other particles. This implies that φ is of the form $\sum_j \varphi_j(t, \mathbf{r}_j)$ (but each φ_j remains arbitrary). In the case where all $\mathbf{B}_j$ are zero, one can reduce $\mathbf{p}_j$ to the form $m_j \mathbf{v}_j$. But there still remains an arbitrary function of time which can be added to the potential v, the Lagrangian λ, and the Hamiltonian h.

Geometric consequences of Maxwell's principle

(12.99) We have seen that the evolution space V of a dynamical system is a *foliated manifold*, the foliation being given by $y \mapsto \mathcal{E} \equiv \ker(\sigma)$. When Maxwell's principle $\mathrm{d}\sigma \equiv 0$ holds, we obtain $\mathcal{E} \equiv \ker(\sigma) \cap \ker(\mathrm{d}\sigma)$, that is, $y \mapsto \mathcal{E}$ is the *characteristic foliation* of σ (5.25). This shows that the Lagrange form σ is an *integral invariant* (5.21). Since the characteristic foliation is sectionable (12.38), σ is the inverse image of a 2-form on the quotient, that is, the space of motions.

This argument also applies to a system with *ideal holonomic constraints* (see (12.52)), because the Lagrange form σ' on the constrained evolution space V' (12.53) is *induced* from the Lagrange form of the unconstrained system (see (9.29)). Hence we have proved the following theorem, which uses the language of §9 (see especially (9.10)).

(12.100) THEOREM: Let V be the evolution space of a dynamical system satisfying Maxwell's principle, eventually supplemented with ideal holonomic constraints. Then

a) The Lagrange form σ gives V the structure of a *presymplectic manifold*.

b) Let x denote the motion of the system defined by an initial condition y (Fig. 12.II). Then the map $y \mapsto x$ is *differentiable*. On the space of motions U there exists a 2-form, which we shall also call the *Lagrange form* and denote by σ, defined by

$$\sigma(\delta y)(\delta' y) \equiv \sigma(\delta x)(\delta' x)\,. \qquad \diamondsuit$$



This 2-form gives the space of motions the structure of a *symplectic manifold*. □

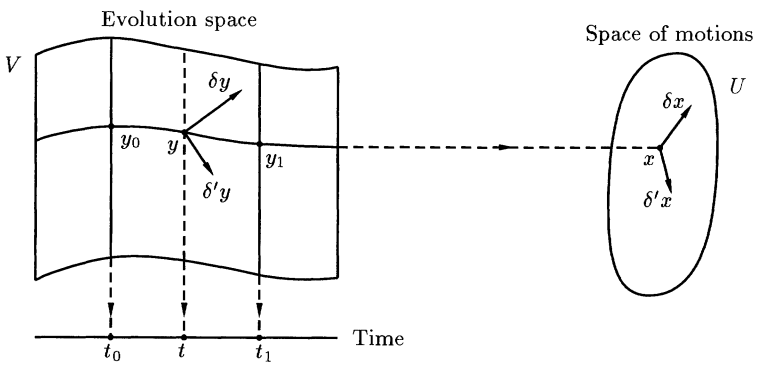


Fig 12.II

(12.101) Let t_0 be a instant of time. We have seen that the associated *phase space*

at time $t = t_0$, defined in V by the equation $t = t_0$, is a transverse section (12.37). It follows that σ induces a *symplectic structure* (9.32) on each phase space. The corresponding Lagrange form σ is obtained by setting $\delta t = \delta' t = 0$ in expression (12.45), which gives

(12.102) $$\sigma(\delta y)(\delta' y) = \sum_j \langle m_j\, \delta \mathbf{v}_j, \delta' \mathbf{r}_j \rangle - \langle m_j\, \delta' \mathbf{v}_j, \delta \mathbf{r}_j \rangle + \langle \mathbf{B}_j, \delta \mathbf{r}_j \times \delta' \mathbf{r}_j \rangle\,.$$

(12.103) *If all* $\mathbf{B}_j$ *are zero* and if there are no constraints, we see that a chart of the phase space defined in (12.32) is canonical (see (10.1), (10.5)).

(12.104) The transformation $y_0 \mapsto x$, where y_0 varies over the phase space at time $t = t_0$ and where x denotes the motion passing through y_0, is a *canonical transformation* (10.10). It follows that the same is true for the transformation $y_0 \mapsto y_1$, where y_0 and y_1 denote the intersections of a motion x with the phase spaces at times $t = t_0$ and $t = t_1$, respectively. This is *Poisson's theorem.*

(12.105) If all motions of a system have the same duration, especially if they are all *eternal* (12.35), then the transformations in (12.104) are *symplectomorphisms* (10.10). In such a case, the space of motions and all phase spaces are *symplectomorphic.*

An application: variation of constants

(12.106) Suppose that we have to solve a problem of dynamics and let us assume that we know how to solve the auxiliary problem obtained by *subtracting a variable* Ω, a function of y, from the *Hamiltonian* h (notation (12.93)). Denoting by ϖ' and σ' the Cartan and Lagrange forms, respectively, of the auxiliary problem, we can immediately deduce from (12.93) that

(12.107) $$\varpi(\delta y) \equiv \varpi'(\delta y) + \Omega\, \delta t\,,$$

whence, by exterior differentiation

$$\sigma(\delta y)(\delta' y) \equiv \sigma'(\delta y)(\delta' y) + \delta\Omega\ \delta' t - \delta'\Omega\ \delta t\,,$$

or, by introducing the motion x *of the auxiliary problem* corresponding to the initial condition y,

(12.108) $$\sigma(\delta y)(\delta' y) \equiv \sigma'(\delta x)(\delta' x) + \delta\Omega\ \delta' t - \delta'\Omega\ \delta t\,.$$

Here σ' is the Lagrange form on the symplectic manifold U, the space of motions of the auxiliary problem (see (12.100)).

The equations of motion $\sigma(\delta y) = 0$ follow from (12.108) and take the form

$$\sigma'(\delta x)(\delta' x) \equiv \delta'\Omega\ \delta t \qquad \text{if } \delta' t \equiv 0\,, \tag{12.109}$$

or, using the notation of (9.16),

$$\frac{dx}{dt} \equiv \operatorname{grad}\Omega\,. \tag{12.110}$$

If we knew a canonical chart for U, (10.6) shows that the above equation can be written as

$$\frac{d}{dt}\,{}^{j}q \equiv \frac{\partial\Omega}{\partial p_j} \qquad \frac{d}{dt}\,p_j \equiv -\frac{\partial\Omega}{\partial\,{}^{j}q}\,. \tag{12.111}$$

(12.112) EXAMPLE: Suppose we have to solve a problem in which all vector potentials $\mathbf{A}_j$ are zero (notation (12.93)). We choose $\Omega \equiv h$, and we note that the auxiliary problem is simply solved by $\mathbf{r}_j = \text{constant}$ and $\mathbf{v}_j = \text{constant}$.[214] We thus can identify the canonical coordinates of the phase space at an arbitrary time (notably the coordinates (12.32)) with the canonical coordinates of the space of motions U. We then recognize that (12.111) are simply Hamilton's equations (7.23). □

(12.113) EXAMPLE: Suppose we have to study the motion of a small planet or a comet which is perturbed by one or several big planets. Here we take Ω equal to the *perturbation potential* so that the auxiliary problem becomes the *Keplerian motion* of the celestial object under study around the sun.

The equations (12.111) define the evolution of 6 parameters characterizing the Keplerian motion, called "orbital elements". Since these elements vary slowly, we can solve the equations (12.111) quite accurately using approximation techniques. Knowing x and t, Kepler's formulæ enable us to calculate y and thus to obtain the perturbed motion.

This is the method of *variation of constants*, due to Lagrange, which is also called *perturbation theory*. It is essential for celestial mechanics. □

[214] Since the auxiliary problem does not satisfy the dynamical equations, no paradox arises from fact that $\frac{d\mathbf{r}_j}{dt} \neq \mathbf{v}_j$

Galilean moments

(12.114) Let V be the evolution space of a dynamical system and U its space of motions (Fig. 12.III). We assume that the system is *isolated* and that it satisfies Maxwell's principle. We know that V is a *presymplectic manifold* (12.100), and that the Galilei group G *acts* on V preserving the Lagrange form (12.76). Thus G is a *dynamical group* for V (definition (11.1)) and we can apply the results of §11.

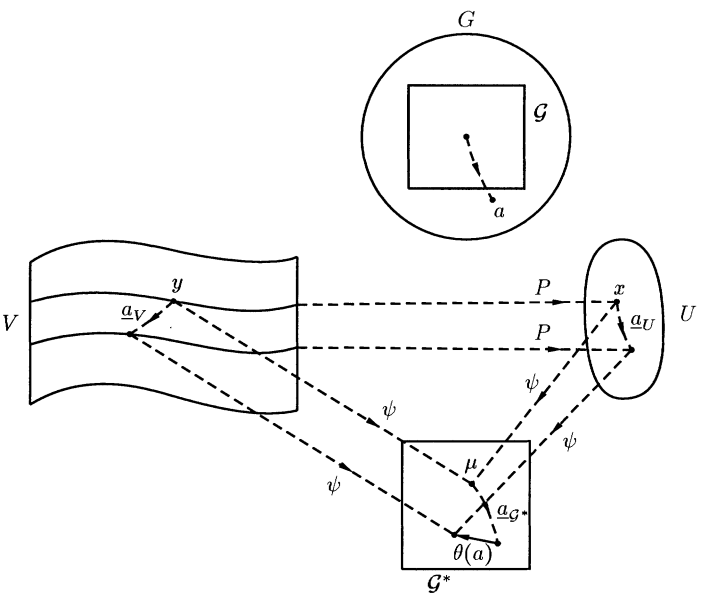


Fig 12.III

Let

$$a = \begin{bmatrix} A & \mathbf{b} & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{matrix} A \in \mathrm{SO}(3)\,, \quad \mathbf{b} \in \mathbf{R}^3\,, \\ \mathbf{c} \in \mathbf{R}^3\,, \quad e \in \mathbf{R} \end{matrix} \tag{12.115}$$

be an element of the Galilei group; it acts on V according to the formulæ (12.76◇), namely,

$$\begin{aligned} t &\mapsto t + e \\ \mathbf{r}_j &\mapsto A(\mathbf{r}_j) + \mathbf{b}\,t + \mathbf{c} \\ \mathbf{v}_j &\mapsto A(\mathbf{v}_j) + \mathbf{b} \end{aligned} \tag{12.116}$$

and thus defines a canonical transformation $\underline{a}_V$ of V.

Let

$$y \equiv \begin{bmatrix} t \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix}$$

be a variable point in V and denote by $x \equiv P(y)$ the motion defined by the initial condition y. We know that the *Galilei group also acts on the space of motions* U according to

$$\underline{a}_U(P(y)) \equiv P(\underline{a}_V(y))\,, \tag{12.117}$$

and thus constitutes a dynamical group of U (11.6).

Let

$$Z = \begin{bmatrix} j(\boldsymbol{\omega}) & \boldsymbol{\beta} & \boldsymbol{\gamma} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} \boldsymbol{\omega} \in \mathbf{R}^3\,, & \boldsymbol{\beta} \in \mathbf{R}^3\,, \\ \boldsymbol{\gamma} \in \mathbf{R}^3\,, & \varepsilon \in \mathbf{R} \end{matrix} \tag{12.118}$$

be an element of the Lie algebra $\mathcal{G}$ of G (notation (12.74)). We know that Z defines a vector field Z_V on V and also a vector field Z_U on U. Definition (6.11) shows that we can calculate Z_V using the formula $Z_V(y) \equiv \delta[\underline{a}_V(y)]$ with $a = 1$ and $\delta a = Z$. This immediately leads to

$$\begin{aligned} \delta t &\equiv \varepsilon \\ \delta \mathbf{r}_j &\equiv \boldsymbol{\omega} \times \mathbf{r}_j + \boldsymbol{\beta}\, t + \boldsymbol{\gamma} \\ \delta \mathbf{v}_j &\equiv \boldsymbol{\omega} \times \mathbf{v}_j + \boldsymbol{\beta}\,. \end{aligned} \tag{12.119}$$

We can now apply *Noether's theorem* in its general symplectic form: there exists (at least locally) a map $\psi\,[y \mapsto \mu]$ from V to the dual $\mathcal{G}^*$ of the Lie algebra $\mathcal{G}$ defined by

$$\sigma(Z_V(y)) = -\,\mathrm{d}[\mu \cdot Z]\,. \tag{12.120}$$

(12.121) The *torsor* μ is the *moment* of the Galilei group acting on V and is a *first integral of the equations of motion*, that is, $\psi(y)$ depends only on the motion $x = P(y)$ (11.12). In other words, there exists a map $x \mapsto \mu$ (which as a matter of fact defines a moment for the action of G on U (11.13)) [Fig. 12.III].

Since μ acts in a linear way on $\mathcal{G}$, we can write

$$\begin{aligned} \mu(Z) \equiv \langle \mathbf{l}, \boldsymbol{\omega} \rangle - \langle \mathbf{g}, \boldsymbol{\beta} \rangle + \langle \mathbf{p}, \boldsymbol{\gamma} \rangle + E\,\varepsilon\,, \\ \mathbf{l} \in \mathbf{R}^3\,,\ \mathbf{g} \in \mathbf{R}^3\,,\ \mathbf{p} \in \mathbf{R}^3\,,\ E \in \mathbf{R}\,; \end{aligned} \tag{12.122}$$

we will denote the torsor μ defined this way by

$$\mu \equiv \{\mathbf{l},\, \mathbf{g},\, \mathbf{p},\, E\}\,. \tag{12.123}$$

Definition (12.120) of μ can be rewritten as

$$\sigma(\delta' y)(\delta y) \equiv \delta' \mu(Z) \qquad \forall\, \delta' y\,, \tag{12.124}$$

with δy being given by (12.119). Looking at formula (12.46) for σ, we can find the value of $\delta'\mu$. Moreover, we know *a priori* that it is a "total derivative", which can be integrated. This determines μ up to an additive constant.

EXAMPLE: Let us consider a material point of unit mass not subjected to any forces. A calculation gives immediately the following solution of (12.124)

$$\mu \equiv \{\mathbf{r} \times \mathbf{v},\, \mathbf{r} - \mathbf{v}\,t,\, \mathbf{v},\, \tfrac{1}{2}\,\|\mathbf{v}\|^2\}\,. \tag{12.125}$$

It is clear that μ is constant on each motion, which is uniform and rectilinear.

Let us apply theorem (11.17). If we denote by ψ the map $y \mapsto \mu$, we know that there should exist a map θ_0 from G to $\mathcal{G}^*$ such that

$$\theta_0(a) \equiv \psi(\underline{a}_V(y)) - \underline{a}_{\mathcal{G}^*}(\psi(y)) \qquad a \in G\,,\ y \in V\,. \tag{12.126}$$

A calculation gives

$$\theta_0(a) \equiv \{\mathbf{c} \times \mathbf{b},\, \mathbf{c} - \mathbf{b}\,e,\, \mathbf{b},\, \tfrac{1}{2}\,\|\mathbf{b}\|^2\}\,. \tag{12.127}$$

We know that θ_0 satisfies (11.17♡):

$$\theta_0(a \cdot a') \equiv \theta_0(a) + \underline{a}_{\mathcal{G}^*}(\theta_0(a'))\,, \tag{12.128}$$

that is, θ_0 is a $\mathcal{G}^*$-cocycle of the Galilei group G (11.19), and its derivative at the origin is a 2-form f_0 satisfying

$$f_0(Z)([Z', Z'']) + f_0(Z')([Z'', Z]) + f_0(Z'')([Z, Z']) \equiv 0\,. \tag{12.129}$$

This identity can be verified immediately using the explicit expression for f_0 deduced from (12.127):

$$f_0(Z)(Z') = \langle \boldsymbol{\beta}, \boldsymbol{\gamma}' \rangle - \langle \boldsymbol{\beta}', \boldsymbol{\gamma} \rangle\,. \tag{12.130}$$

Applying definition (11.19) one can easily verify that θ_0 *is not a coboundary*. Thus θ_0 defines a *nonzero symplectic cohomology class* of the Galilei group (11.30, 11.31). □

(12.131) Conversely, one can look directly for the symplectic cocycles of the Galilei group. Employing proposition (11.22c) and identity (11.33), a lengthy but straightforward calculation[215] shows that *the dimension of the symplectic cohomology space of the Galilei group is* 1. In other words, every symplectic cocycle θ is obtained from the cocycle θ_0 (12.127) by the formula



$$\theta(a) \equiv \underline{a}_{\mathcal{G}^*}(\mu_0) - \mu_0 + m\,\theta_0(a)\,, \tag{12.132}$$

where μ_0 is a torsor and m is a number *representing the cohomology class* of θ. Its derivative f at $a = 1$ is

$$f(Z)(Z') \equiv \mu_0[Z, Z'] + m f_0(Z)(Z')\,. \tag{12.133}$$

Let us now return to the general case of an isolated dynamical system satisfying Maxwell's principle and let us apply (11.17#), replacing f by its value (12.133). If the system possesses a moment μ, one can choose this map (subtracting μ_0 (which is permitted) in view of (12.133)) so that

$$\sigma(Z_V(y))(Z'_V(y)) \equiv \mu[Z, Z'] + m f_0(Z)(Z')\,. \tag{12.134}$$

If in the system (12.124) and (12.134) we replace σ, μ, and f_0 by their expressions (12.46), (12.123), and (12.130), respectively, we obtain the following set of equations:

(12.135)

$$\spadesuit \qquad m = \sum_j m_j$$

$$\heartsuit \qquad \mathbf{B}_j = 0 \qquad \forall\, j$$

$$\diamondsuit \qquad \sum_j \mathbf{E}_j = 0 \qquad \sum_j \mathbf{r}_j \times \mathbf{E}_j = 0$$

$$\clubsuit \qquad \mathbf{l} = \sum_j \mathbf{r}_j \times m_j\,\mathbf{v}_j \qquad \mathbf{g} = \sum_j m_j[\mathbf{r}_j - \mathbf{v}_j\,t] \qquad \mathbf{p} = \sum_j m_j \mathbf{v}_j$$

$$\bowtie \qquad \delta E = \delta\Big\{\tfrac{1}{2}\sum_j m_j \|\mathbf{v}_j\|^2\Big\} - \sum_j \langle \mathbf{E}_j, \delta \mathbf{r}_j\rangle \qquad \forall\, \delta y\,.$$

[215] See Bargmann [4]. The point of view is different, but the algebraic calculations are the same.

Remarks

(12.136) Equation ♠ shows that the *total mass* $\sum_j m_j$ *of the system* can be interpreted as characterizing the *cohomology class* of the Galilei group. Since this class is *never zero*, it follows that *the Lagrangian in Hamilton's principle is not invariant under the Galilei group*, whatever the choice of potentials and gauge (see (11.21)). □

(12.137) Equation ♡ shows that the Lagrange form of an isolated system reduces to the "classical" expression (12.40) with the $\mathbf{E}_j$ being equal to the forces $\mathbf{F}_j$. □

(12.138) We thus see that classical mechanics seems to be incapable of describing, for instance, a system of magnets moving freely in space, because the magnetic field $\mathbf{B}$ created by the system has to be zero.

In fact, this paradox can be resolved by assuming that the magnets are not made up of material points but of *particles with spin endowed with a magnetic moment*. We will see later on (15.53) that we can describe in this way the *magnetostatic* forces between magnets without introducing special relativity. Furthermore, it is known that *ferromagnetism* is effectively due to the intrinsic magnetic moment of the electrons in a magnet and not to their orbital motion. □

(12.139) The equations ♢ express the fact that the *total force* $\sum_j \mathbf{F}_j$ and the *total torque* $\sum_j \mathbf{r}_j \times \mathbf{F}_j$ are zero, that is, the Euclidean *torsor* of these forces is *zero* (see (12.20), (12.21)). One can show that this is equivalent to the existence of a decomposition

$$\mathbf{F}_j = \sum_{\substack{k \\ [k \neq j]}} \mathbf{F}_{jk}$$

such that $\mathbf{F}_{jk} + \mathbf{F}_{kj} = 0$ and such that $\mathbf{F}_{jk}$ is parallel to $\mathbf{r}_j - \mathbf{r}_k$. One should recognize that this is the *principle of equality of action and reaction*, as formulated by Newton. It appears here as a *consequence* of Galilean relativity and Maxwell's principle. □

(12.140) The equalities ♣ enable us to recognize $\mathbf{l}$ as the *angular momentum* and $\mathbf{p}$ as the *linear momentum* of the system. As for $\mathbf{g}$, one can write it in the form

(12.141) $$\mathbf{g} \equiv m\,\mathbf{R} - \mathbf{p}\,t$$

by introducing the *center of mass* $\mathbf{R}$.

That $\mathbf{p}, \mathbf{l}$, and $\mathbf{g}$ are constant is a classical theorem. In particular the constancy of $\mathbf{g}$ tells us that the motion of the center of mass is rectilinear and uniform. This is the *principle of inertia* of Galilei. □

Formula $\bowtie$ can be written as

(12.142) $$E \equiv \tfrac{1}{2} \sum_j m_j \|\mathbf{v}_j\|^2 + v(\mathbf{r}_1, \ldots, \mathbf{r}_N) + \text{constant}\,,$$

where v is a function such that $\overline{\frac{\partial v}{\partial \mathbf{r}_j}} = -\mathbf{E}_j$, that is, v is a *potential* (definition (12.91)). Note that *time does not appear in the expression for* v. In E one should recognize the quantity called *energy*, which is the sum of the *potential energy* v and the *kinetic energy* $\frac{1}{2} \sum_j m_j \|\mathbf{v}_j\|^2$. The fact that E is a first integral is called the *principle of conservation of energy*.

(12.143) We see that the structure of the Galilei group does not permit us to specify the additive constant in formula (12.142) for the energy,[216] nor does it assures us that E is *globally* defined: it could happen that the energy is a multivalued function.[217] The fact that such a multivalued function is not found in nature can be credited to relativistic mechanics. □

(12.144) Without going into details, let us mention that the above results are also valid for an *isolated system subjected to ideal holonomic constraints*, provided of course that the constraints also satisfy the principle of Galilean relativity in the form (12.78).

Examples of dynamical groups

(12.145) For particular dynamical systems there may exist several Lie groups acting on the space of motions leaving the Lagrange form invariant. These are *dynamical groups* in the sense of (11.1) but also according to terminology used by physicists. Let us give some examples without going into the details of the calculations.

(12.146) Consider two dynamical systems which evolve independently of each other. Clearly the two systems taken together *again* form a system. It is immediate that *the space of motions of the composite system* is *the direct product of the spaces of motions of the component systems, a direct product in the sense of the direct product of two symplectic manifolds* (9.7).

[216] This follows from the fact that (12.134) only gives the value of μ on the *derived algebra* of the Lie algebra $\mathcal{G}$. The derived algebra of $\mathcal{G}$ is defined by $\varepsilon = 0$ (notation (12.118)).

[217] Abuse of language meaning that E is *defined on the universal covering space* of V, which we suppose to be connected. This covering space is a simply connected symplectic manifold (see (10.25)).

On the other hand, the evolution space of the composite system *is not* the direct product of the evolution spaces of the component systems. This follows in particular from the fact that the dimension of an evolution space is odd.

(12.147) If G is a dynamical group for each component system, then $G \times G$, the direct product of G with itself, is a dynamical group for the composite system: it suffices to let an element a_1 of G act on a motion of the first system and an element a_2 of G on one of the second system.

In the case where the composite system is made up of two isolated systems, one can see that the Galilean moment of the composite system is the *sum* of the Galilean moments of the component systems. This is true separately for angular momentum, linear momentum, energy, and the quantity $\mathbf{g}$. □

(12.148) Consider a material point subject to a *central force* $\mathbf{F} \equiv f(r)\,\mathbf{r}$, where f is a differentiable function. One can check immediately that the evolution space admits $\mathrm{SO}(3) \times \mathbf{R}$ as a *dynamical group*: it suffices to set $\mathbf{b} = \mathbf{c} = 0$ in equations (12.115) and (12.116). This dynamical group is a Lie subgroup of the Galilei group. And of course $\mathrm{SO}(3) \times \mathbf{R}$ is also a dynamical group for the space of motions U. □

(12.149) The previous example can be applied in particular to the case of Keplerian motions (where $f(r) \equiv -k\,m\,r^{-3}$) which was studied in (12.3). Notice that in this case the space of motions *is not Hausdorff*, because to each rectilinear motion x ending in a collision with the center of attraction at time t_0, there corresponds *another* motion x' starting at time t_0 such that every open subset of U containing x intersects each open set containing x'. If we agree to *identify* these motions, we obtain a Hausdorff space of motions U_0.

Now let us consider the set U_1 of bounded Keplerian motions (having circular or elliptical trajectories, or even a segment of a straight line). One can show that U_1 is an open subset of U_0 diffeomorphic to $S_2 \times S_2 \times S_1 \times \mathbf{R}$, and that it admits $\mathrm{SO}(4) \times \mathbf{R}$ as a dynamical group.[218] This group has a moment μ, which can be denoted as $\{\mathbf{l}, \mathbf{a}, E\}$ and which satisfies $\langle \mathbf{l}, \mathbf{a} \rangle \equiv 0$ and $E \equiv -\frac{1}{2[\|\mathbf{a}\|^2 + \|\mathbf{l}\|^2]}$. Thus the quantity E, which represents the standard choice for the energy, is negative for the motions in U_1.

The quantity $\mathbf{l}$ is the angular momentum already encountered in (12.3).[219] The vector $\mathbf{a}$ is related to the "Lenz vector" $\mathbf{e}$ (12.3) by $\mathbf{a} = \mathbf{e}/\sqrt{-2E}$. The existence of the "accidental" first integral $\mathbf{e}$ can

[218] See Bacry et al. [3]

[219] Units are chosen such that $k = 1$ and $m = 1$.

thus be linked to the existence of a particular dynamical group. □

(12.150) When one wants to study vibrations of a dynamical system in the neighborhood of an equilibrium point, one often uses the well known method of *linearization*.[220] This method consists of replacing the given system by a "slightly different" system, called a *linear oscillator* or *harmonic oscillator*, which is easier to study. Since Maxwell's principle is assumed to hold for the given system, it follows that it also holds for the associated harmonic oscillator. Moreover, its space of motions U is a *symplectic vector space* in the sense of (8.10). Hence we know a dynamical group G of U: the affine symplectic group, that is, the set of matrices $a = \begin{bmatrix} A & B \\ 0 & 1 \end{bmatrix}$ ($A \in \mathrm{Sp}(U)$, $B \in U$) acting on U according to $\underline{a}_U(x) = A(x) + B$ (see (10.12)). If the dimension of U is $2p$ then the dimension of G is $p(2p+3)$ (see (8.21)).

Consequently, every Lie subgroup of G is also a dynamical group of the harmonic oscillator (11.3). For example, the group of translations of U, that is, $A \equiv 1$; or the symplectic group $\mathrm{Sp}(U)$, that is, $B \equiv 0$; or subgroups of $\mathrm{Sp}(U)$ isomorphic to $\mathrm{U}(p)$ (see (8.26)) or $\mathrm{SU}(p)$, whose dimensions are p^2 and $p^2 - 1$, respectively ((6.66), (6.67)). □

(12.151) As an example of a harmonic oscillator, let us consider a point attracted by a central force $\mathbf{F}$ proportional to the distance, that is,

$$\diamondsuit \qquad \mathbf{F} = -m\,\omega^2\,\mathbf{r}\,.$$

Introducing two constant vectors $\mathbf{A}$ and $\mathbf{B}$ we can integrate the equations of motion to obtain

$$\heartsuit \qquad \mathbf{r} \equiv \mathbf{A}\,\cos\omega t + \mathbf{B}\,\sin\omega t\,.$$

We thus can identify a motion x with the pair $\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}$ and we find immediately that

$$\clubsuit \qquad \sigma(\delta x)(\delta' x) \equiv m\,\omega[\langle \delta\mathbf{B}, \delta'\mathbf{A}\rangle - \langle \delta'\mathbf{B}, \delta\mathbf{A}\rangle]\,,$$

which exhibits the *symplectic vector space* structure of the space of motions. In this case the affine symplectic group has dimension 27; a possible subgroup isomorphic to U(3) is given by

$$\spadesuit \qquad \mathbf{A} + i\mathbf{B} \mapsto a(\mathbf{A} + i\mathbf{B}) \qquad [a \in \mathrm{U}(3)]\,. \qquad \square$$

[220] See Landau and Lifschitz [24].

(12.152) Let G be a dynamical group of a system which possesses a moment μ. *If the symplectic cohomology class of μ is zero*, we can choose the moment so that

$$\diamondsuit \qquad [\mu(Z), \mu(Z')]_{\text{Poisson}} \equiv \mu\left([Z, Z']_{\text{Lie}}\right) \qquad \forall\, Z, Z' \text{ constant.}$$

This is an immediate consequence of (11.17#) with $f = 0$, of definition (9.22) of the Poisson bracket, and of definition (11.9) of the moment of a dynamical group.[221]

Let us now assume additionally that G is connected and abelian, and that its orbits have the same dimension n as G. Then the orbits are isotropic (9.33) and $\ker\left(\frac{\partial \mu}{\partial x}\right)$ is the symplectic orthogonal of the tangent space to the orbit through x (11.17♠). Thus the equations $\mu =$ constant define coisotropic submanifolds of V of dimension $2p - n$ (see (8.6c), (8.3f), and (1.40)), which are thus presymplectic (9.35). Actually, each orbit is a *submanifold* of a (sub)manifold $\mu =$ constant. More precisely, each orbit is a *leaf* of the characteristic foliation (definition (9.8)) of the submanifold $\mu =$ constant. □

(12.153) The above discussion applies in particular to the case of *conservative* systems, that is, systems such that the group of time translations ($t \mapsto t +$ constant, $\mathbf{r}_j \mapsto \mathbf{r}_j$, $\mathbf{v}_j \mapsto \mathbf{v}_j$) preserves the Lagrange form. The moment of this action is, up to a sign, the *energy* E of the system. If the space of motions does not contain equilibrium points, then the equations $E =$ constant define presymplectic submanifolds of dimension $2p - 1$, for which the orbits of G are leaves of the characteristic foliation. Thus each leaf defines a *nondated* motion (sometimes called a *stationary state* of the system). If a submanifold $E =$ constant is *sectionable*, then the orbits corresponding to this value of the energy form a symplectic manifold of dimension $2p - 2$ (theorem (9.10)). □

(12.154) The above discussion applies in particular to Kepler's problem. The set of stationary states corresponding to a negative value of the energy is symplectomorphic to the symplectic direct product of two spheres of dimension two, each having the symplectic structure defined in (11.40) with the same value for λ. These two spheres are described by the vectors $\mathbf{e} + \mathbf{l}\sqrt{-2E}$ and $\mathbf{e} - \mathbf{l}\sqrt{-2E}$, respectively. It follows that $\mathrm{SO}(3) \times \mathrm{SO}(3)$ is a dynamical group for this manifold (see (11.40) and (12.147)). □

[221] By abuse of language one sometimes says that G is a "dynamical group" if there exists a differentiable map $x \mapsto \mu$ defined on an open subset of the space of motions U to the dual $\mathcal{G}^*$ of the Lie algebra of G satisfying $\diamondsuit$. But this notion only makes use of the property that the *dimension* of U is $2p$ (since two symplectic manifolds of the same dimension are locally symplectomorphic); its relevance for physics seems to be rather limited.

(12.155) Likewise, in the case of the harmonic oscillator considered in (12.151), the energy is given by $E \equiv \frac{1}{2} m \omega^2 [\, \|\mathbf{A}\|^2 + \|\mathbf{B}\|^2]$. The set of stationary states corresponding to a positive value of E forms a symplectic manifold of dimension 4 which is diffeomorphic to the set of complex 1-dimensional vector subspaces of $\mathbf{C}^3$, that is, the complex projective plane. The group U(3) is a dynamical group for this manifold. In fact, only the quotient $\mathrm{U}(3)/\mathrm{U}(1)$, or equivalently $\mathrm{SU}(3)/\mathbf{Z}_3$, acts effectively. □

(12.156) In certain dynamical problems one also encounters *groups of canonical similitudes*.

For example, the substitutions $\mathbf{r}_j \mapsto a^{-2}\, \mathbf{r}_j$, $\mathbf{v}_j \mapsto a\, \mathbf{v}_j$, $t \mapsto a^{-3}\, t$ are canonical similitudes of the evolution space of the N-body problem, and thus also of its space of motions (10.21). By differentiation one obtains the infinitesimal canonical similitude

$$\diamondsuit \qquad \delta \mathbf{r}_j \equiv -2\mathbf{r}_j\,, \qquad \delta \mathbf{v}_j \equiv \mathbf{v}_j\,, \qquad \delta t \equiv -3t\,,$$

which provides us with a *potential* ϖ for the Lagrange form (10.37). The 1-form ϖ is an *integral invariant of the equations of motion*.

§13. The principles of symplectic mechanics

Nonrelativistic symplectic mechanics

(13.0) Consider a classical dynamical system. An elementary description of such a system consists of decomposing it into material points and formulating the laws describing the forces to which these points are subjected. If these laws satisfy Maxwell's principle (12.90), the procedures which we have studied in the preceding section (namely, the construction of the evolution space and its Lagrange form, the integration of the equations of motion, and the passage to the quotient) lead to the construction of a symplectic manifold: the space of motions.

We will now *give up* the notions of material point and force and will take as *axioms of mechanics* some essential properties of the space of motions:[222]

(13.1)

I. The space of motions of a dynamical system is a *connected symplectic manifold.*

II. If several dynamical systems evolve independently, the manifold of motions of the composite system is the *symplectic direct product* of the spaces of motions of the component systems.

III. If a dynamical system is isolated, its manifold of motions admits the Galilei group as a dynamical group.

As one can see, we not only give up the notion of a material point but also the notion of an evolution space (and thus a fortiori the notions of phase space, "configuration space", etc.).[223] Euclidean space and time itself are only indirectly invoked through the Galilei group.

Since the principles (13.1) are valid in classical mechanics (see (12.100), (12.146), and (12.117), respectively), they do *not break* with classical mechanics, but simply *enlarge* it. They permit us to treat new dynamical systems, as we shall see in examples later on.

Of course it is the analogy with the classical cases which will guide us in our physical interpretation of these new systems. □

[222] Definitions: symplectic manifold (9.1), direct product of symplectic manifolds (9.7), Galilei group (12.73), and dynamical group (11.1).

[223] Giving up these notions is familiar in quantum mechanics.

(13.2) Consider an isolated dynamical system. We shall suppose that its *structure* is completely characterized by its symplectic properties and by the manner in which the Galilei group G acts on it. In more precise terms, two isolated dynamical systems having spaces of motions U and U' will be called *isomorphic* if there exists a *symplectomorphism* Φ from U to U' such that

$$\Phi(\underline{a}_U(x)) = \underline{a}_{U'}(\Phi(x)) \qquad \forall\, x \in U\,,\ \forall\, a \in G\,.$$

When this is the case, we will say that the dynamical systems have *the same model* and that Φ is an *isomorphism* from the first system to the second.

Moments, mass, and the center of mass

The study of Galilean moments of an isolated system made in the preceding section remains valid as long as the explicit expression of the Lagrange form (12.45) does not play any role in it. Thus the *Galilei group* G of matrices

(13.3)
$$a \equiv \begin{bmatrix} A & \mathbf{b} & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l} A \in \mathrm{SO}(3)\,, \quad \mathbf{b} \in \mathbf{R}^3\,, \\ \mathbf{c} \in \mathbf{R}^3\,, \quad e \in \mathbf{R} \end{array}$$

acts on the space of motions U of an isolated dynamical system while preserving the Lagrange form (13.1.III). To each element of the Lie algebra $\mathcal{G}$ of G

(13.4)
$$Z \equiv \begin{bmatrix} j(\boldsymbol{\omega}) & \boldsymbol{\beta} & \boldsymbol{\omega} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{array}{l} \boldsymbol{\omega}\,, \boldsymbol{\beta}\,, \boldsymbol{\gamma} \in \mathbf{R}^3\,, \\ \varepsilon \in \mathbf{R} \end{array}$$

there corresponds a vector field Z_U on U which is an infinitesimal canonical transformation. The associated *moment* (definition (11.7)) has been denoted by

(13.5)
$$\mu \equiv \{\mathbf{l},\, \mathbf{g},\, \mathbf{p},\, E\}\,,$$

and its action on the Lie algebra $\mathcal{G}$ is given by

(13.6)
$$\mu(Z) \equiv \langle \mathbf{l},\, \boldsymbol{\omega} \rangle - \langle \mathbf{g},\, \boldsymbol{\beta} \rangle + \langle \mathbf{p},\, \boldsymbol{\gamma} \rangle - E\,\varepsilon\,.$$

(13.7) *We now make the fundamental assumption that this moment exists.*[224] We thus know that the dynamical group G possesses a *symplectic cohomology class* (§11), which can be labeled by a number m (12.132). On

[224] One might say that this is a new principle (see (12.143)). Newtonian mechanics can borrow this principle from relativistic mechanics without being alienated from it (see (13.81)).

account of (12.135♠), we will call m the *mass* of the system. Thus a priori we can consider systems with *zero* or *negative* mass.

As in (12.134) we may choose an additive constant for μ so that

(13.8) $$\sigma(Z_U(x))(Z'_U(x)) \equiv \mu([Z, Z']) + m\{\langle \boldsymbol{\beta}, \boldsymbol{\gamma}' \rangle - \langle \boldsymbol{\beta}', \boldsymbol{\gamma} \rangle\} .$$

(13.9) The remarks on the formulæ (12.135) furnish us with a physical interpretation of the moment μ. The vector $\mathbf{l}$ is the *angular momentum* of the system, $\mathbf{p}$ its *linear momentum*, and E its *energy*. As for the quantity $\mathbf{g}$, it defines the center of mass (or barycenter) of the system: it is the point $\mathbf{R}$ having a uniform and rectilinear motion defined by the equation

(13.10) $$\mathbf{R} \equiv \frac{\mathbf{p}\, t + \mathbf{g}}{m} ,$$

provided of course that the mass m is nonzero. One can verify that this motion does not depend on the chosen inertial reference frame.

(13.11) The quantities $\mathbf{l}$, $\mathbf{g}$, and $\mathbf{p}$ are completely defined by (13.8). As for the energy E, it is defined up to an arbitrary additive constant.

(13.12) For a system composed of various isolated systems, one can verify — as a consequence of (13.1.II) — that the quantities m, $\mathbf{l}$, $\mathbf{g}$, $\mathbf{p}$, and E are *additive*. It follows in particular that the center of mass of the composite system is given by the usual formula

$$\mathbf{R} = \frac{\sum_j m_j \mathbf{R}_j}{\sum_j m_j}$$

at every time t. □

If we change the units of measurement of the quantities $\mathbf{r}$, t, and σ by

$$\mathbf{r} \mapsto \frac{\mathbf{r}}{L} , \qquad t \mapsto \frac{t}{T} , \qquad \sigma \mapsto \frac{\sigma}{A} ,$$

the preceding formulæ tell us how the other variables transform. These transformations can be summarized by the following "dimensional equations":

(13.13)

r	L
t	T
σ	A

$\boldsymbol{\omega}$	1
$\boldsymbol{\beta}$	LT^{-1}
$\boldsymbol{\gamma}$	L
ε	T

$\mathbf{l}$	A
$\mathbf{g}$	ATL^{-1}
$\mathbf{p}$	AL^{-1}
E	AT^{-1}

m	ATL^{-2}

If, as is customary, we denote by M the dimension of mass, we see that $A = ML^2T^{-1}$. A quantity having this dimension is called an *action*,

because it agrees with the dimension of the Hamiltonian action (12.96). Replacing A by ML^2T^{-1} in (13.13), we recover the usual dimensions of the given quantities of classical mechanics. We conclude that they remain valid in symplectic mechanics.

The center of mass decomposition

(13.14) THEOREM: Let G be a dynamical group acting on a connected symplectic manifold V and possessing a moment μ. Let θ be the corresponding cocycle (notation (11.17)).

1) Let $\tilde{G}$ be a Lie subgroup of G with Lie algebra $\tilde{\mathcal{G}}$. Then $\tilde{G}$ is a *dynamical group* of V possessing a moment $\tilde{\mu}$ induced from μ:

$$\diamondsuit \qquad \tilde{\mu}(Z) \equiv \mu(Z) \qquad \forall\, Z \in \tilde{\mathcal{G}}\,.$$

2) Now suppose additionally that $\tilde{G}$ is a *normal* subgroup of G. Denote by $\tilde{\psi}$ the map $x \mapsto \tilde{\mu}$, by $\underline{a}_{\tilde{\mathcal{G}}^*}$ the representation of G dual to the representation induced on $\tilde{\mathcal{G}}$ by the adjoint representation of G,[225] and by $\tilde{\theta}$ the map

$$\tilde{\theta}(a)(Z) \equiv \theta(a)(Z) \qquad \forall Z \in \tilde{\mathcal{G}}\,.$$

Then

$$\heartsuit \qquad \tilde{\psi}(\underline{a}_U(x)) - \underline{a}_{\tilde{\mathcal{G}}^*}(\tilde{\psi}(x)) \equiv \tilde{\theta}(a) \qquad \forall\, a \in G\,.^{226}$$

PROOF: Immediate. □

(13.15) THEOREM: Let G be a dynamical group of a connected symplectic manifold V possessing a moment μ (notation (11.17)). Assume that G is *connected* and *abelian* and that the 2-form f (notation (11.17c)) is *injective*. Then

a) θ is an *isomorphism* between the Lie group G and the *additive group* $\mathcal{G}^*$.[227]

b) If we define $\sigma(\mu)(\mu') \equiv \mu f^{-1}(\mu')$, $\mathcal{G}^*$ becomes a symplectic vector space.

[225] As $\tilde{G}$ is a normal subgroup, $[Z \in \tilde{\mathcal{G}},\, a \in G] \Rightarrow [\underline{a}_{\mathcal{G}}(Z) \in \tilde{\mathcal{G}}]$ (see (6.23)). If we define $\underline{a}_{\tilde{\mathcal{G}}}(Z) \equiv \underline{a}_{\mathcal{G}}(Z)\ \forall Z \in \tilde{\mathcal{G}}$, then $\underline{a}_{\tilde{\mathcal{G}}}$ is indeed a representation of G on $\tilde{\mathcal{G}}$.

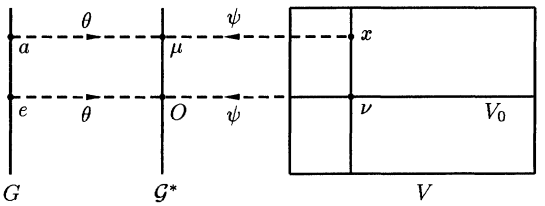
[226] This shows that $\tilde{\theta}$ is a $\tilde{\mathcal{G}}^*$-cocycle of G.

[227] It is simultaneously an isomorphism of groups and a diffeomorphism of manifolds.

c) The map ψ $[x \mapsto \mu]$ induces a *symplectomorphism* between each G-*orbit in* V and $\mathcal{G}^*$.

d) V is *symplectomorphic* either to $\mathcal{G}^*$ or to the *direct product* $\mathcal{G}^* \times V_0$, where V_0 is the submanifold of V defined by $\mu = 0$ with the induced symplectic structure.

Fig 13.I



PROOF: Since G is abelian, its adjoint representation is trivial ($\underline{a}_{\mathcal{G}}(Z) \equiv Z$) and thus also its coadjoint representation. Hence the Lie bracket on $\mathcal{G}$ is 0. Formulæ (11.17♡, ♢, ♠, ♯) then become

$$\heartsuit \qquad \theta(a \times b) \equiv \theta(a) + \theta(b)$$

$$\diamondsuit \qquad \psi(\underline{a}_V(x)) \equiv \psi(x) + \theta(a)$$

$$\spadesuit \qquad D(\psi)(x)(Z_V(x)) \equiv f(Z)$$

$$\sharp \qquad \sigma(Z_V(x))(Z'_V(x)) \equiv f(Z)(Z') .$$

Since f, the derivative of θ at the identity element e, is injective, there exists an open subset E of G containing e on which the restriction of θ is a diffeomorphism (theorem (1.35)). Thus $\mathrm{im}(\theta)$ contains a ball with center 0 in $\mathcal{G}^*$. As $\theta(a^2) \equiv 2\theta(a)$ according to ♡, we see that $\mathrm{im}(\theta)$ is in fact all of $\mathcal{G}^*$.

Let Γ be the group of translations $\underline{a}_G$ such that $\theta(a) = 0$. The elements of Γ are diffeomorphisms of G. One can verify immediately that the images of E under the elements of Γ are pairwise disjoint. Since the conditions (1.43) are satisfied, theorem (1.45) shows that (G, Γ, θ) is a *covering space* of $\mathcal{G}^*$ (theorems (1.56), (1.57)). As $\mathcal{G}^*$ is simply connected (1.54), Γ can only contain the identity element. Thus θ is a diffeomorphism from G onto $\mathcal{G}^*$. Formula ♡ shows that θ is also an isomorphism of groups, whence (a). Formula ♢ then shows that ψ is a bijection between G-orbits in V and $\mathcal{G}^*$. Let $\delta x \equiv Z_V(x)$ and $\delta' x \equiv Z'_V(x)$ be two tangent vectors to the G-orbit through x. Formula ♠ shows that $\delta\mu = f(Z)$ and $\delta'\mu = f(Z')$ and formula ♯ shows that

$$\sigma(\delta x)(\delta' x) = f(Z)(Z') = \delta\mu(f^{-1}(\delta'\mu)) = \sigma(\delta\mu)(\delta'\mu) .$$

Hence (b) and (c) follow.

If the dimension of V equals that of G, then the G-orbits are submanifolds of V of the same dimension, hence are open and pairwise disjoint. Since V is connected, there is but a single orbit. Thus ψ is a symplectomorphism between V and $\mathcal{G}^*$.

Now suppose that the dimension $2n$ of V is different from the dimension $2p$ of G. It is clear that the rank of $D(\psi)(x)$ equals $2p$ and that each equation $\psi(x) = \mu_0$ defines a submanifold V_{μ_0} of dimension $2(n-p)$ of V (theorem (1.40)). Formula $\diamondsuit$ shows that each G-orbit and each (sub)manifold V_{μ_0} intersect at a single point. In particular, the intersection of the G-orbit through x with V_0 is

$$\clubsuit \qquad \nu = \underline{\theta^{-1}(-\psi(x))}_V(x)\,.$$

Thus the map $x \mapsto \nu$ is differentiable.

And thus the map $x \mapsto \begin{pmatrix}\mu\\ \nu\end{pmatrix}$ is a diffeomorphism between V and $\mathcal{G}^* \times V_0$, because the inverse map $\begin{pmatrix}\mu\\ \nu\end{pmatrix} \mapsto \underline{\theta^{-1}(\mu)}_V(\nu)$ is also differentiable. Hence we can write

$$\sigma(\delta x)(\delta' x) \equiv \sigma\left(\frac{\partial x}{\partial \mu}\,\delta\mu + \frac{\partial x}{\partial \nu}\,\delta\nu\right)\left(\frac{\partial x}{\partial \mu}\,\delta'\mu + \frac{\partial x}{\partial \nu}\,\delta'\nu\right).$$

Since the vector $\frac{\partial x}{\partial \mu}\,\delta\mu$ is tangent to a G-orbit, it can be written as $Z_V(x)$. Formula $\heartsuit$ shows that the vector $\frac{\partial x}{\partial \nu}\,\delta\nu$ is tangent to the manifold V_{μ_0} passing through x and thus is symplectically perpendicular to the previous vector (due to definition (11.7) of a moment). Employing (c) and the fact that the maps $\underline{a}_V$ are symplectomorphisms, the above equation becomes

$$\sigma(\delta x)(\delta x') \equiv \sigma(\delta\mu, \delta'\mu) + \sigma(\delta\nu)(\delta'\nu)\,.$$

This shows that V_0 is a symplectic manifold and that the map $x \mapsto \begin{pmatrix}\mu\\ \nu\end{pmatrix}$ is a symplectomorphism between V and $\mathcal{G}^* \times V_0$. □



It happens that the Galilei group G has a *normal abelian subgroup* $\tilde{G}$ consisting of the elements of the form

$$a = \begin{bmatrix} 1 & \mathbf{b} & \mathbf{c} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{b}\,,\,\mathbf{c} \in \mathbf{R}^3\,, \tag{13.16}$$

whose Lie algebra $\tilde{\mathcal{G}}$ is the set of elements

$$Z = \begin{bmatrix} 0 & \boldsymbol{\beta} & \boldsymbol{\gamma} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \boldsymbol{\beta}\,,\,\boldsymbol{\gamma} \in \mathbf{R}^3\,. \tag{13.17}$$

Now consider an isolated dynamical system (in the general sense of (13.1)). Theorem (13.14) gives the following results.

THEOREM: $\tilde{G}$ is a dynamical group of the system and has a moment $\tilde{\mu}$ induced from the Galilean moment $\mu \equiv \{\mathbf{l}, \mathbf{g}, \mathbf{p}, E\}$ by

$$\tilde{\mu}(Z) \equiv \langle \mathbf{p}, \boldsymbol{\gamma} \rangle - \langle \mathbf{g}, \boldsymbol{\beta} \rangle \qquad \forall\, Z \in \tilde{\mathcal{G}}\,. \tag{13.18}$$

We shall denote $\tilde{\mu}$ as

$$\tilde{\mu} \equiv \{\mathbf{g}, \mathbf{p}\}\,. \tag{13.19}$$ □

The map $\tilde{\theta}$ defined by

$$\tilde{\theta}(a) \equiv m\,\{\mathbf{c}, \mathbf{b}\} \tag{13.20}$$

satisfies (13.14♡). The corresponding symplectic cocycle $\tilde{f}$ of the Lie algebra $\tilde{\mathcal{G}}$ is given by

$$\tilde{f}(Z)(Z') \equiv m\,[\langle \boldsymbol{\beta}, \boldsymbol{\gamma}' \rangle - \langle \boldsymbol{\beta}', \boldsymbol{\gamma} \rangle] \qquad \forall\, Z, Z' \in \tilde{\mathcal{G}}\,. \tag{13.21}$$

If m, which is the generalized mass of the system in the sense of (13.7), *is not zero*, then $\tilde{f}$ is injective and we can apply theorem (13.15): the Lie group $\tilde{G}$ is isomorphic to the additive group $\mathbf{R}^6$ (which can be verified immediately), $\tilde{\mathcal{G}}^*$ becomes a symplectic vector space by defining

$$\sigma(\{\delta\mathbf{g}, \delta\mathbf{p}\})(\{\delta'\mathbf{g}, \delta'\mathbf{p}\}) \equiv \frac{1}{m}[\langle \delta\mathbf{p}, \delta'\mathbf{g} \rangle - \langle \delta'\mathbf{p}, \delta\mathbf{g} \rangle]\,, \tag{13.22}$$

and the map $\tilde{\psi}\ [x \mapsto \tilde{\mu}]$ is a symplectomorphism between each $\tilde{G}$-orbit and $\tilde{\mathcal{G}}^*$. A priori there are two possible cases:

CASE 1. The dimension of the space of motions U is 6. Then U must be symplectomorphic to $\tilde{\mathcal{G}}^*$. Using as variables the time t, the center of mass $\mathbf{R}$ (13.10), and its velocity $\mathbf{V} = \mathbf{p}/m$, we find that

$$\begin{aligned} \sigma(\delta x)(\delta' x) &= \sigma(\delta\tilde{\mu})(\delta'\tilde{\mu}) \\ &= \langle m\,\delta\mathbf{V},\ \delta'\mathbf{R} - \mathbf{V}\,\delta' t \rangle - \langle m\,\delta'\mathbf{V},\ \delta\mathbf{R} - \mathbf{V}\,\delta t \rangle\,. \end{aligned} \tag{13.23}$$

We conclude that the system is *isomorphic*, in the sense of (13.2), *to a free material point.* □

CASE 2. The dimension of U is ≥ 8, in which case there exists a symplectomorphism χ defined by

$$\chi(x) \equiv \begin{pmatrix} \tilde{\mu} \\ \nu \end{pmatrix}, \tag{13.24}$$

between U and the direct product $\widetilde{\mathcal{G}}^* \times U_0$.

(13.25) Due to (13.23), $\widetilde{\mathcal{G}}^*$ can be interpreted as the manifold of *motions of the center of mass*, assumed to have (total) mass m.

(13.26) U_0 is the (sub)manifold of the space of motions for which $\tilde{\mu}$ is zero, that is, for which the center of mass remains fixed at the origin of the chosen inertial reference frame. In other words, it is the set of motions *around* the center of mass. The variable ν is called the *proper motion*. □

One can verify that the Galilei group *preserves this decomposition*, that is, one can define an action of G on $\widetilde{\mathcal{G}}^*$ and on U_0 so that

$$\chi(\underline{a}_U(x)) \equiv \begin{pmatrix} \underline{a}_{\widetilde{\mathcal{G}}^*}(\tilde{\mu}) \\ \underline{a}_{U_0}(\nu) \end{pmatrix} \tag{13.27}$$

with

$$\begin{cases} \underline{a}_{\widetilde{\mathcal{G}}^*}(\tilde{\mu}) \equiv \tilde{\mu} \cdot \underline{a^{-1}}_{\mathcal{G}} + \tilde{\theta}(a)\,^{228} \\ \underline{a}_{U_0}(\nu) \equiv \underline{F(a)}_U(\nu)\,, \end{cases} \tag{13.28}$$

where F is the homomorphism from G to itself defined by

$$[\tilde{\theta}(F(a)) = 0 \qquad \text{and} \qquad a \cdot F(a)^{-1} \in \widetilde{G}] \qquad\qquad \forall\, a \in G \tag{13.29}$$

or, equivalently, by

$$F\left(\begin{bmatrix} A & \mathbf{b} & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix}\right) \equiv \begin{bmatrix} A & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix}. \tag{13.30}$$

(13.31) It follows that G acts on U_0 only through $G_0 = \mathrm{im}(F)$, which is the Lie subgroup of G of elements preserving U_0. G_0 is isomorphic to the quotient group $G/\widetilde{G}$ (because of (13.29)) and to $\mathrm{SO}(3) \times \mathbf{R}$ (because of (13.30)).

Furthermore, G acts by symplectomorphisms on each of the factors of $\widetilde{\mathcal{G}}^* \times U_0$. Consequently, the direct product $G \times G$ is *also a dynamical group* of U: it suffices to define $\underline{\begin{pmatrix} a \\ b \end{pmatrix}}_U(x)$ by

$$\chi\left(\underline{\begin{pmatrix} a \\ b \end{pmatrix}}_U(x)\right) = \begin{pmatrix} \underline{a}_{\widetilde{\mathcal{G}}^*}(\tilde{\mu}) \\ \underline{b}_{U_0}(\nu) \end{pmatrix} \qquad\qquad \forall\, a, b \in G\,. \tag{13.32}$$

[228] It is the adjoint representation of G which appears on the right hand side.

The initial action (13.1.III) is the *diagonal* action of (13.32), that is, $a \equiv b$. More precisely, taking (13.31) into account, it follows that $G \times G_0$ is a dynamical group of U once we define

$$\chi\Big(\underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{U}(x)\Big) = \begin{pmatrix} \underline{a}_{\widetilde{\mathcal{G}}^*}(\widetilde{\mu}) \\ \underline{b}_{U_0}(\nu) \end{pmatrix} \qquad \forall\, a \in G\,,\ \forall\, b \in G_0\,. \tag{13.33}$$

We then conclude:

(13.34) If the mass of an isolated system is not zero, it admits the product of the Galilei group G and the group $\mathrm{SO}(3) \times \mathbf{R}$ as a dynamical group. □

It follows that the Galilean moment is the sum of an *orbital momentum* and a *proper angular momentum*. A calculation gives

$$\begin{cases} \mu_{\text{orbital}} = \{\frac{1}{m}\mathbf{g} \times \mathbf{p},\, \mathbf{g},\, \mathbf{p},\, \frac{1}{2m}\|\mathbf{p}\|^2\} \\ \mu_{\text{proper}} = \{\mathbf{l}_{\text{proper}},\, 0,\, 0,\, E_{\text{proper}}\}\,, \end{cases} \tag{13.35}$$

where $\mathbf{l}_{\text{proper}}$ and E_{proper} depend only on the proper motion ν. Using the variables (13.23), the orbital momentum may be written as [229]

$$m\,\{\mathbf{R} \times \mathbf{V},\, \mathbf{R} - \mathbf{V}\,t,\, \mathbf{V},\, \tfrac{1}{2}\,\|\mathbf{V}\|^2\}\,. \tag{13.36}$$

□

(13.37) EXAMPLE: The above results apply first of all to classical mechanics, where they are well known. In the case of the two body problem (for example, a double star or the hydrogen atom), one can characterize a proper motion by studying only one of the two bodies which appears to be attracted to the fixed center of mass according to Newton's law. The 6-dimensional manifold U_0 has been considered already in (12.149). □

(13.38) EXAMPLE: The above results also apply to the classical mechanics of *constrained systems* (see (12.144)). In the case of a *free rigid body* (the problem of Euler-Poinsot), U_0 is diffeomorphic to $\mathrm{SO}(3) \times \mathbf{R}^3$ and hence is 6-dimensional. There is an exception when the rigid body is *linear* (for example, a diatomic molecule). Then U_0 is the space S_2^D of tangent vectors to a 2-dimensional sphere (notation (1.31)), and thus is 4-dimensional. □

(13.39) EXAMPLE: In the nonclassical case of a *particle with spin* (see (14.48)) U_0 is a *sphere* S_2 equipped with the symplectic structure defined in (11.40).

[229] Compare with (12.125).

Minkowski space and the Poincaré group

Let us go back to the language of (12.57). An event X, that is, a point of space-time E_4, can be described in a reference frame $\mathcal{R}$ by its position $\mathbf{r}$ and its date t. This is written as

(13.40) $$X \equiv \mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \qquad \mathbf{r} \in \mathbf{R}^3\,,\ t \in \mathbf{R}\,.$$

(13.41) A *Lorentz (reference) frame* is an inertial reference frame in which the unit of time is chosen so that the velocity of light in vacuum equals 1.[230] □

(13.42) Let $\mathcal{R}$ be a Lorentz frame. Space-time E_4 can obviously be given a *vector space* structure by assuming that the map $\mathcal{R}$ is linear.[231] But then it can be given the structure of a *Euclidean space* (in the sense of (6.49)) by

(13.43) $$g(\delta X)(\delta X) \equiv \delta t^2 - \|\delta \mathbf{r}\|^2\,.$$

The tensor g is completely characterized by this formula, knowing that it is bilinear and symmetric (6.49). It follows immediately that

(13.44) $$g(\mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix})(\mathcal{R}\begin{pmatrix} \mathbf{r}' \\ t' \end{pmatrix}) \equiv t\,t' - \langle \mathbf{r},\, \mathbf{r}' \rangle\,.$$

(13.45) Space-time E_4 with the structure defined above is called *Minkowski space*. □

Employing the *transpose* (definition (6.53)), formula (13.44) can be written as

(13.46) $$\overline{\mathcal{R}} \cdot \mathcal{R} = \begin{pmatrix} -1_{\mathbf{R}^3} & 0 \\ 0 & 1 \end{pmatrix},$$

where $1_{\mathbf{R}^3}$ denotes the identity map from $\mathbf{R}^3$ to itself, that is, the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(13.47) This shows that E_4 is a hyperbolic space of dimension 4 and index of inertia 1 (theorem (6.72)). □

---

[230] For example, one might take the unit of time to be $3.335635\,10^{-11}$ seconds when the unit of length is the centimeter.

[231] In which case the equalities $\mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} + \mathcal{R}\begin{pmatrix} \mathbf{r}' \\ t' \end{pmatrix} \equiv \mathcal{R}\begin{pmatrix} \mathbf{r}+\mathbf{r}' \\ t+t' \end{pmatrix}$ and $\lambda\mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \equiv \mathcal{R}\begin{pmatrix} \lambda\mathbf{r} \\ \lambda t \end{pmatrix}$ are taken as the *definition* of the operations $X + X'$ and λX in E_4.

(13.48) The *Lorentz group* is the orthogonal group $O(E_4)$ of Minkowski space (definition (6.63)). Theorem (6.76) shows that $O(E_4)$ is a 6-dimensional Lie group having four connected components. The *restricted Lorentz group* is the connected component of the identity element of the Lorentz group. It is an open, connected, normal Lie subgroup of $O(E_4)$, which is still 6-dimensional (6.35).

Using theorem (6.75) one can easily show that every element of the restricted Lorentz group can be written as

$$L \equiv \mathcal{R} \cdot \exp \begin{pmatrix} 0 & \mathbf{b} \\ \overline{\mathbf{b}} & 0 \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \cdot \mathcal{R}^{-1} \qquad A \in \mathrm{SO}(3)\,,\ \mathbf{b} \in \mathbf{R}^3\,. \tag{13.49}$$

More precisely, the map $\begin{pmatrix} A \\ \mathbf{b} \end{pmatrix} \mapsto L$ defined by (13.49) is a *diffeomorphism* from $\mathrm{SO}(3) \times \mathbf{R}^3$ *onto* the restricted Lorentz group.

Similarly, an element of the *Lie algebra* of the Lorentz group (see (6.62)) can be written as

$$\Lambda \equiv \mathcal{R} \cdot \begin{pmatrix} j(\boldsymbol{\omega}) & \boldsymbol{\beta} \\ \overline{\boldsymbol{\beta}} & 0 \end{pmatrix} \cdot \mathcal{R}^{-1} \qquad \boldsymbol{\omega} \in \mathbf{R}^3\,,\ \boldsymbol{\beta} \in \mathbf{R}^3\,. \tag{13.50}$$ □

The *Poincaré group*[232] is the group of matrices

$$a \equiv \begin{pmatrix} L & C \\ 0 & 1 \end{pmatrix} \qquad L \in \mathrm{O}(E_4)\,,\ C \in E_4\,, \tag{13.51}$$

acting on Minkowski space according to

$$\underline{a}(X) \equiv L \cdot X + C\,. \tag{13.52}$$

(13.53) It is clear that the Poincaré group is a 10-dimensional Lie group diffeomorphic to $O(E_4) \times E_4$, and thus it also has four connected components (1.50). The connected component of the identity, called the *restricted Poincaré group*, is obtained by choosing L to lie in the restricted Lorentz group. Thus the restricted Poincaré group is diffeomorphic to $\mathrm{SO}(3) \times \mathbf{R}^7$ and hence is diffeomorphic to the Galilei group (see (12.73)).

An element of the *Lie algebra of the Poincaré group* will be denoted by

$$Z \equiv \begin{pmatrix} \Lambda & \Gamma \\ 0 & 0 \end{pmatrix} \qquad \overline{\Lambda} = -\Lambda\,,\ \Gamma \in E_4\,. \tag{13.54}$$

[232] Also called the "inhomogeneous Lorentz group".

Using (13.50) and defining[233]

$$\mathcal{R}' = \begin{pmatrix} \mathcal{R} & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \Gamma \equiv \mathcal{R} \begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix}, \tag{13.55}$$

we can also write

$$Z \equiv \mathcal{R}' \cdot \begin{pmatrix} j(\omega) & \beta & \gamma \\ \overline{\beta} & 0 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix} \cdot (\mathcal{R}')^{-1} \qquad \begin{matrix} \omega\,,\, \beta\,,\, \gamma \in \mathbf{R}^3\,, \\ \varepsilon \in \mathbf{R}\,. \end{matrix} \tag{13.56}$$

A *torsor of the Poincaré group* will be denoted by

$$\mu \equiv \{M,\, P\}\,, \qquad \text{with} \quad \overline{M} = -M\,,\ P \in E_4 \tag{13.57}$$

and will be defined by the identity[234]

$$\mu(Z) \equiv -\tfrac{1}{2}\,\mathrm{Tr}(M \cdot \Lambda) - \overline{P} \cdot \Gamma\,. \tag{13.58}$$

Representing M and P in the reference frame $\mathcal{R}$ by

$$M \equiv \mathcal{R} \begin{bmatrix} j(\mathbf{l}) & \mathbf{g} \\ \overline{\mathbf{g}} & 0 \end{bmatrix} \mathcal{R}^{-1} \quad \text{and} \quad P \equiv \mathcal{R} \begin{bmatrix} \mathbf{p} \\ E \end{bmatrix} \qquad \begin{matrix} \mathbf{l}\,,\, \mathbf{g}\,,\, \mathbf{p} \in \mathbf{R}^3\,, \\ E \in \mathbf{R}\,, \end{matrix} \tag{13.59}$$

formula (13.58) becomes[235]

$$\mu(Z) \equiv \langle \mathbf{l},\, \omega \rangle - \langle \mathbf{g},\, \beta \rangle + \langle \mathbf{p},\, \gamma \rangle - E\,\varepsilon\,. \tag{13.60}$$

□

A direct verification allows us to establish the following propositions.

(13.61) PROPOSITION: The Lie algebra $\mathcal{G}$ of the Poincaré group is equal to its derived algebra.

(13.62) PROPOSITION: The symplectic cohomology space of the Poincaré group is zero.[236]

---

[233] $\mathcal{R}'$ is a linear map from $\mathbf{R}^5$ onto $E_4 \times \mathbf{R}$.

[234] For every real Euclidean space E one can define a Euclidean structure on the space E' of *anti-hermitian* linear maps from E to itself by defining $g(M)(\Lambda) = -\frac{1}{2}\,\mathrm{Tr}(M \cdot \Lambda)$ $\forall M, \Lambda \in E'$.

[235] Compare with (12.122).

[236] More generally, the $\mathcal{G}^*$-cohomology of the Poincaré group is zero. Note that the $\mathcal{G}$-cohomology is not zero: there exists a nonzero $\mathcal{G}$-cocycle θ defined by $\theta\begin{pmatrix} L & C \\ 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}$ which is not a $\mathcal{G}$-coboundary.

Relativistic mechanics



We postulate that in special relativity the change from one Lorentz frame $\mathcal{R}$ to another $\mathcal{R}^*$ is given by

$$\mathcal{R} = \underline{a} \cdot \mathcal{R}^* , \tag{13.63}$$

where a is an element of the *restricted Poincaré group.*[237]

From (13.63) we immediately deduce that formula (13.43) (13.64)

$$g(\delta X)(\delta X) = \delta t^2 - \|\delta \mathbf{r}\|^2$$

is valid in *all Lorentz frames.* This explains the fact that, in an arbitrary inertial reference frame, the velocity of light is the same in all directions (the Michelson-Morley experiment).

But it is then necessary to abandon the classical formulæ for the change of reference frame. If we write

$$X \equiv \mathcal{R} \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \equiv \mathcal{R}^* \begin{pmatrix} \mathbf{r}^* \\ t^* \end{pmatrix}, \tag{13.65}$$

then the preceding formulæ show that we must have

$$\begin{pmatrix} \mathbf{r}^* \\ t^* \\ 1 \end{pmatrix} \equiv \alpha \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \tag{13.66}$$

with

$$\alpha = (\mathcal{R}')^{-1} \cdot a \cdot \mathcal{R}' = \left[\begin{array}{cc|c} & & \mathbf{c} \\ & \mathcal{L} & \\ & & e \\ \hline 0 & 0 & 1 \end{array}\right] \qquad \mathbf{c} \in \mathbf{R}^3 , \, e \in \mathbf{R} \tag{13.67}$$

$$\mathcal{L} = \exp \begin{pmatrix} 0 & \mathbf{b} \\ \overline{\mathbf{b}} & 0 \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{b} \in \mathbf{R}^3 , \, A \in \mathrm{SO}(3) .$$

In particular, if $\mathbf{b} \neq 0$ it follows that t^* depends not only on t but also on $\mathbf{r}$: in special relativity *the notion of simultaneity depends on the choice of reference frame.*

[237] The case of the full Poincaré group will be considered later on (14.46).

EXAMPLE: Taking $A = 1$, $\mathbf{c} = 0$, $e = 0$, and $\mathbf{b} = \begin{pmatrix} \varphi \\ 0 \\ 0 \end{pmatrix}$ we find

$$\begin{pmatrix} \mathbf{r}^* \\ t^* \end{pmatrix} \equiv \begin{bmatrix} \cosh\varphi & 0 & 0 & \sinh\varphi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\varphi & 0 & 0 & \cosh\varphi \end{bmatrix} \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix}. \tag{13.68}$$

This formula can be interpreted as describing a reference frame moving in a straight line and with constant velocity with respect to another one ("Einstein's train"), the velocity being $v = \tanh\varphi$ (and thus less than the speed of light). Replacing $\cosh\varphi$ and $\sinh\varphi$ by their values $1/\sqrt{1-v^2}$ and $v/\sqrt{1-v^2}$, respectively, we obtain Lorentz's formulæ, whose interpretation leads to the well known notions of "slowing down of clocks" and of "Lorentz contraction". □

(13.69) A fortiori, the formulæ (12.59) giving the transition to an *accelerating* or *rotating* reference frame are no longer valid. It took a long time before it was realized that these formulæ *do not have any relativistic counterpart*.[238]

Special relativity has an immediate consequence for mechanics: in order to apply the principle of relativity (12.72), we have to replace the classical formulæ (12.71), which involve the Galilei group, by (13.63). Consequently, we will modify the axioms (13.1) of symplectic mechanics (changing only axiom III).

(13.70)

I. The space of motions of a dynamical system is a connected symplectic manifold.

II. If several dynamical systems evolve independently, the manifold of motions of the composite system is the symplectic direct product of the spaces of motions of the component systems.

III. If a dynamical system is isolated, its manifold of motions admits the *restricted Poincaré group* as a dynamical group.

We will treat the above axioms as the principles of *relativistic mechanics*.

[238] When one wants to describe in classical mechanics an elastic body that is reasonably rigid, it is customary to consider it, in first approximation, to be a *rigid body*. However, the *definition* of a rigid body is not compatible with special relativity, because the definition of the distance between two moving points involves the notion of simultaneity, which depends (as we have seen) on the choice of reference frame. There is no contradiction here: the theory of elasticity can be described consistently in special or general relativity (see J.-M. Souriau [32]). One simply has to give up an approximation that is inconsistent with the structure of the theory.

(13.71) Of course we also adapt definition (13.2) of *isomorphism* between two isolated dynamical systems: an isomorphism will be a symplectomorphism Φ satisfying

$$\Phi(\underline{a}_V(x)) = \underline{a}_{V'}(\Phi(x)) \qquad \forall a \in G\,,$$

where G is now the Poincaré group.

(13.72) Let U be the symplectic manifold of motions of an isolated relativistic system. Since the restricted Poincaré group is a dynamical group of U and since the derived algebra of the Lie algebra $\mathcal{G}$ of G is $\mathcal{G}$ itself (13.61), it follows from theorem (11.8) that G has a moment μ, which is a torsor of G, and for which we will use the notation (13.57) to (13.60). As the symplectic cohomology of G is zero (13.62), we may *choose* the additive constant in μ so that the associated symplectic cocycle is zero; in other words (see (11.17◇))

(13.73) $$\psi(\underline{a}_U(x)) \equiv \underline{a}_{\mathcal{G}^*}(\psi(x)) \qquad \forall a \in G\,,$$

where ψ denotes the map $[x \mapsto \mu]$.

Formula (11.17♯) then shows that

(13.74) $$\sigma(Z_V(x))(Z'_V(x)) \equiv \mu[Z, Z']\,.$$

One might consider the above formula as a *definition* of μ.

(13.75) Note that theorem (13.24) (which gave a decomposition of U into a direct product and permitted us to define a dynamical group of dimension 14 for each *isolated classical* dynamical system with nonzero mass), *does not hold in the relativistic case.* This follows from the fact that the Poincaré group G does not have an abelian subgroup satisfying the hypotheses of theorem (13.15), and thus certainly not a normal such subgroup.

(13.76) The Galilei group and the Poincaré group are two Lie subgroups of the group of affine transformations of space-time. Thus their intersection $\tilde{G}$ is a Lie subgroup of each of them having as its Lie algebra the intersection $\tilde{\mathcal{G}}$ of their Lie algebras (theorem (6.34)). We find immediately [239]

(13.77)
$$a \in \tilde{G} \iff a = \begin{bmatrix} A & 0 & \mathbf{c} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix}$$

$$Z \in \tilde{\mathcal{G}} \iff Z = \begin{bmatrix} j(\boldsymbol{\omega}) & 0 & \boldsymbol{\gamma} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix},$$

[239] In view of the standard abuse of identifying space-time transformations with their matrix representatives, we recommend calling $\tilde{G}$ the Aristotle group.

and thus $\widetilde{G}$ is 7-dimensional.

Theorem (13.14) shows that $\widetilde{G}$ is a dynamical group not only for a classical system but also for a relativistic system. It also tells us how to compute the moment $\tilde{\mu}$, for which we find

$$\tilde{\mu}(Z) \equiv \langle \mathbf{l}, \boldsymbol{\omega} \rangle + \langle \mathbf{p}, \boldsymbol{\gamma} \rangle - E\,\varepsilon \qquad \forall Z \in \widetilde{\mathcal{G}} \tag{13.78}$$

in terms of the classical notation (13.6) as well as in terms of the relativistic notation (13.59).

(13.79) Consequently, the *relativistic* quantities $\mathbf{l}$, $\mathbf{p}$, and E can also be called *angular momentum*, *linear momentum*, and *energy*, respectively, as in (13.9). This is why in Minkowski space the vector (see (13.59))

$$P \equiv \mathcal{R}\begin{pmatrix} \mathbf{p} \\ E \end{pmatrix} \tag{13.80}$$

is called the *energy-momentum* vector (or 4-*momentum*, or *Lorentz momentum*, etc).

(13.81) Note that the energy E of a relativistic system is defined *globally and without an additive constant*, once we have adopted the convention (13.73). □

If G is the restricted Poincaré group and X is a point of space-time, then the *isotropy group* G_X of X (see (6.36)) is given by

$$[a \in G_X] \iff a = \begin{bmatrix} L & [1-L]X \\ 0 & 1 \end{bmatrix}, \tag{13.82}$$

where L belongs to the restricted Lorentz group. The Lie algebra of G_X is given by

$$[Z \in \mathcal{G}_X] \iff Z = \begin{bmatrix} \Lambda & -\Lambda X \\ 0 & 0 \end{bmatrix} \qquad (\overline{\Lambda} = -\Lambda)\,. \tag{13.83}$$

Returning to our relativistic dynamical system (13.72), G_X is a dynamical group whose moment μ_X is given by

$$\mu_X(Z) = \mu(Z) \qquad \forall Z \in \mathcal{G}_X\,, \tag{13.84}$$

(theorem (13.84)). Taking (13.82) and (13.58) into account, we find

$$\mu_X(Z) \equiv -\tfrac{1}{2}\,\mathrm{Tr}\,(M_X \cdot \Lambda)\,, \tag{13.85}$$

with

$$M_X = M + P \cdot \overline{X} - X \cdot \overline{P}\,. \tag{13.86}$$

The anti-hermitian operator M_X, which characterizes the moment μ_X, is called the *Lorentz moment with respect to the point* X. Formula (13.86) expresses the Lorentz moment with respect to X in terms of X itself, the 4-momentum P, and in terms of M, which is the Lorentz moment with respect to the origin. Using the notation of (13.59) we get more precisely

$$\begin{aligned} \mathbf{l}_X &= \mathbf{l} + \mathbf{p} \times \mathbf{r} \\ \mathbf{g}_X &= \mathbf{g} + \mathbf{p}\, t - \mathbf{r}\, E\,. \end{aligned} \tag{13.87}$$

□

We now give some useful formulæ concerning the geometry of oriented Minkowski space.[240]

First of all there exists a nonzero 4-form on E_4, denoted by vol (see (4.30), (4.31)) which can be defined by

$$\mathrm{vol}(X_1)(X_2)(X_3)(X_4) = \det \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{r}_4 \\ t_1 & t_2 & t_3 & t_4 \end{pmatrix} \tag{13.88}$$

$$[X_j \equiv \mathcal{R} \begin{pmatrix} \mathbf{r}_j \\ t_j \end{pmatrix} \qquad j = 1,2,3,4]\,,$$

whatever our choice for the Lorentz frame $\mathcal{R}$.

We will call the vector

$$j(X_1)(X_2)(X_3) = \overline{\mathrm{vol}(X_1)(X_2)(X_3)} \tag{13.89}$$

the *vector product of the three vectors* X_1, X_2, and X_3. It is orthogonal to each X_i and is zero if two of them coincide.

For two vectors X_1 and X_2, the map $j(X_1)(X_2)$ is anti-hermitian and satisfies [241]

$$j(X_1)(X_2) \equiv \mathcal{R} \begin{pmatrix} j(\mathbf{r}_2 t_1 - \mathbf{r}_1 t_2) & \mathbf{r}_1 \times \mathbf{r}_2 \\ \overline{\mathbf{r}_1 \times \mathbf{r}_2} & 0 \end{pmatrix} \mathcal{R}^{-1}\,, \tag{13.90}$$

$$j(LX_1)(LX_2) = \det(L)\, L\, j(X_1)(X_2) \overline{L} \qquad \text{for } L \in \text{Lorentz group}, \tag{13.91}$$

$$\begin{aligned} j(A)(B) \cdot j(X)(Y) \;\equiv\; & [X \cdot \overline{Y} - Y \cdot \overline{X}] \cdot [A \cdot \overline{B} - B \cdot \overline{A}] \\ & + [\overline{A} \cdot X \cdot \overline{B} \cdot Y - \overline{A} \cdot Y \cdot \overline{B} \cdot X]\, 1_{E_4}\,. \end{aligned} \tag{13.92}$$

There also exists a *linear map* $*$ transforming an anti-hermitian operator F into an anti-hermitian operator $*(F)$ such that:

$$\overline{X}_1 \cdot *(F) \cdot X_2 \equiv -\tfrac{1}{2} \mathrm{Tr}(j(X_1)(X_2) \cdot F) \tag{13.93}$$

[240] See Souriau [31]. Here we choose the opposite orientation.

[241] Inside the matrix, j denotes the operator of the *three dimensional* vector product.

$$*(X_1 \cdot \overline{X}_2 - X_2 \cdot \overline{X}_1) \equiv j(X_1)(X_2) \tag{13.94}$$

$$*(*(F)) \equiv -F \tag{13.95}$$

$$* \left(j(X_1)(X_2)\right) \equiv X_2 \cdot \overline{X}_1 - X_1 \cdot \overline{X}_2 \tag{13.96}$$

$$[*(F)]^2 \equiv F^2 - \tfrac{1}{2} \operatorname{Tr}(F^2)\, 1_{E_4} \tag{13.97}$$

$$*(L \cdot F \cdot \overline{L}) \equiv L \cdot *(F) \cdot \overline{L} \det(L) \quad \text{for } L \in \text{Lorentz group} \tag{13.98}$$

$$* \left(\mathcal{R} \cdot \begin{bmatrix} j(\omega) & \beta \\ \overline{\beta} & 0 \end{bmatrix} \cdot \mathcal{R}^{-1} \right) \equiv \mathcal{R} \cdot \begin{bmatrix} j(-\beta) & \omega \\ \overline{\omega} & 0 \end{bmatrix} \cdot \mathcal{R}^{-1} . \tag{13.99}$$

Finally there exists a number $\operatorname{pf}(F)$, called the *Pfaffian* of F, defined by

$$* (F) \cdot F = F \cdot [*(F)] = \operatorname{pf}(F) \cdot 1_{E_4} \qquad (\overline{F} \equiv -F) , \tag{13.100}$$

which has the following properties

$$\begin{gathered} \operatorname{pf}(X_1\overline{X_2} - X_2\overline{X_1} + X_3\overline{X_4} - X_4\overline{X_3}) \equiv \operatorname{vol}(X_1)(X_2)(X_3)(X_4) \\ \operatorname{pf}(A \cdot F \cdot \overline{A}) \equiv \det(A) \operatorname{pf}(F) \qquad [\forall A \in L(E_4)] \\ \det(F) = -[\operatorname{pf}(F)]^2 \\ \operatorname{pf} \left(\mathcal{R} \cdot \begin{bmatrix} j(\omega) & \beta \\ \overline{\beta} & 0 \end{bmatrix} \cdot \mathcal{R}^{-1} \right) = \langle \omega, \beta \rangle . \end{gathered} \tag{13.101}$$

□

Combining (13.86) with (13.94) we find that the moment M_X with respect to the point X satisfies

$$* (M_X) = *(M) + j(P)(X) . \tag{13.102}$$

Hence if we define

$$W = *(M) \cdot P , \tag{13.103}$$

then

$$* (M_X) \cdot P = W \qquad \forall X . \tag{13.104}$$

The vector W is called the *polarization* of the dynamical system. Its definition (13.103) shows that W is *orthogonal* to P, because $*(M)$ is anti-hermitian. Note that we have

$$W = \mathcal{R} \begin{pmatrix} \mathbf{p} \times \mathbf{g} + \mathbf{l}\, E \\ \langle \mathbf{l}, \mathbf{p} \rangle \end{pmatrix} \tag{13.105}$$

and

$$\operatorname{pf}(M_X) \equiv \operatorname{pf}(M) + \overline{W} \cdot X . \tag{13.106}$$

□

By hypothesis (13.73), the Poincaré group acts on its torsors by the coadjoint representation, which can be computed using (11.15) and (6.28).

With the notation $a = \begin{bmatrix} L & C \\ 0 & 1 \end{bmatrix}$ and $\mu = \{M, P\}$ already used in (13.51) and (13.57) we find

$$\underline{a}(\mu) = \{M', P'\}$$

with

$$\begin{aligned} M' &\equiv L\,M\,\overline{L} + C\overline{P}\,\overline{L} - L\,P\,\overline{C} \\ P' &\equiv L\,P\,. \end{aligned} \tag{13.107}$$

Let us also note the formulæ

$$M'_{X'} \equiv L \cdot M_X \cdot \overline{L} \tag{13.108}$$

$$W' \equiv L \cdot W\,, \tag{13.109}$$

where X' is the image of X under $\underline{a}$, namely, $X' = L \cdot X + C$, and where W' is the new polarization $W' = *(M') \cdot P'$.

It follows from (13.107) and (13.109) that the two numbers

$$\begin{aligned} C_2 &= \overline{P} \cdot P \\ C_4 &= \overline{W} \cdot W \end{aligned} \tag{13.110}$$

are *invariant* under the Poincaré group; they are called *Casimirs*. Due to (13.97) we have

$$C_4 = \overline{P} \cdot [\tfrac{1}{2}\,\mathrm{Tr}(M^2) - M^2] \cdot P\,, \tag{13.111}$$

and thus C_2 and C_4 are polynomials in μ of degree 2 and 4, respectively. □

In relativity the dimensions L and T (notation (13.13)) *coincide* because of (13.41). Here are the dimensional equations of the other variables:

(13.112)

X	L
σ	A

Λ	1
Γ	L

M	A
P	AL^{-1}

W	A^2L^{-1}
C_2	A^2L^{-2}
C_4	A^4L^{-2}

.

§14. A mechanistic description of elementary particles

Elementary systems

(14.1) Let U be the symplectic manifold of motions of an isolated relativistic dynamical system. This means that the restricted Poincaré group G acts on U. We will say that the system is "*elementary*" if G acts transitively.[242] □

We are now going to study these elementary dynamical systems. Later on we will see that the *elementary particles* which one observes in nature seem, when they are free, to be elementary systems — at least in approximation.

(14.2) Let us denote by x a motion of the system, that is, a point of U. Let ψ be the map which sends x to its moment μ, and which is normalized by condition (13.73)

(14.3) $$\psi(\underline{a}_U(x)) \equiv \underline{a}_{G^*}(\psi(x))\,.$$

When x varies over U, μ varies over *an orbit of the coadjoint representation* ($a \mapsto \underline{a}_{G^*}$) of the Poincaré group. Theorem (11.38) shows that ψ is locally a symplectomorphism between U and the coadjoint orbit, equipped with the symplectic structure defined by algorithm (11.34) with $\theta = 0$. Thus U and the orbit have *the same dimension.*

We are going to classify the elementary systems according to the type[243]

[242] That is, $[x, x' \in U] \Rightarrow [\exists\, a \in G,\ \underline{a}_U(x) = x']$. If x and x' are two motions of the system, then there are two Lorentz frames $\mathcal{R}$ and $\mathcal{R}'$ such that the motion x observed in the reference frame $\mathcal{R}$ is identical to the motion x' observed in the reference frame $\mathcal{R}'$. One might say that a system is elementary *if it does not have any other structure than its space-time structure.*

[243] A vector $X = \mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}$ is

timelike	if $\overline{X} \cdot X > 0$	$(\lvert t\rvert > \lvert\mathbf{r}\rvert)$
spacelike	if $\overline{X} \cdot X < 0$	$(\lvert t\rvert < \lvert\mathbf{r}\rvert)$
lightlike or *null*	if $\overline{X} \cdot X = 0,\ X \neq 0$	$(\lvert t\rvert = \lvert\mathbf{r}\rvert \neq 0)$.

The timelike and lightlike vectors fall into two classes: *future* (with $t > 0$) and *past* (with $t < 0$). An element L of the Lorentz group transforms a timelike vector into a timelike vector and a lightlike vector into a lightlike vector. L is *orthochronous* if it preserves the future and past; it is *antichronous* if it exchanges the future with the

of their energy-momentum vector P and of their polarization vector W. □

Case I. A particle with spin

(14.4) DEFINITION: We will call an elementary dynamical system *a (relativistic) particle with spin* if its energy-momentum P and polarization W satisfy

$$\overline{P} \cdot P > 0 \qquad \text{and} \qquad W \neq 0 .$$

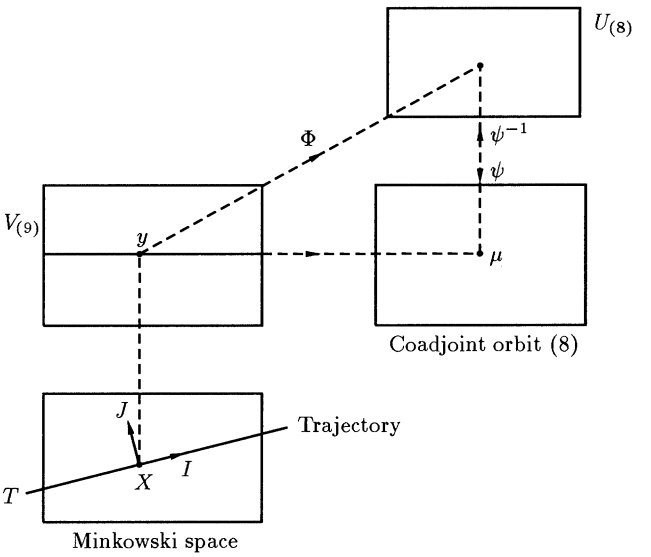


Fig 14.I

(14.4) THEOREM: For relativistic particles with spin we have the following collection of results.

a) $\overline{W} \cdot W$ is negative and the numbers

$$\diamondsuit \qquad m = \operatorname{sign}(E)\sqrt{\overline{P} \cdot P}\ ^{244} \qquad \text{and} \qquad s = \sqrt{\frac{-\overline{W} \cdot W}{\overline{P} \cdot P}}$$

do not depend on the motion. They are called the *mass* [245] and *spin*

past. The set of orthochronous elements of the Lorentz group form an open subgroup of $O(E_4)$ and is the union of two connected components of $O(E_4)$. The restricted Lorentz group is the intersection of the orthochronous group and $SO(E_4)$.

[244] E is the *energy* measured in an arbitrary Lorentz frame (notation (13.59)).

of the particle, respectively.[246]

b) Given m and s ($m \neq 0$, $s > 0$), there exists *a unique model*[247] of a particle with spin of mass m and spin s. The dimension of the space U of motions is 8. If μ is the moment (for the restricted Poincaré group G) of a motion x, then the map $x \mapsto \mu$ is a symplectomorphism from U to a coadjoint orbit of G.

c) Let V be the 9-dimensional manifold described by the 3-tuple

$$y = \begin{pmatrix} X \\ I \\ J \end{pmatrix},$$

where X is a point of Minkowski space, and where I and J are 4-vectors satisfying the relations

$$\begin{aligned} &\overline{I} \cdot I \equiv 1\,, && I \text{ future pointing} \\ &\overline{I} \cdot J \equiv 0 \\ &\overline{J} \cdot J \equiv -1\,. \end{aligned} \qquad \heartsuit$$

A motion x can be characterized by a point y of V such that the map $\Phi\,[y \mapsto x]$ is differentiable and such that the moment $\mu \equiv \{M, P\}$ is given by

$$M \equiv s\,\Omega + m[X \cdot \overline{I} - I \cdot \overline{X}] \qquad \text{and} \qquad P \equiv m\,I\,, \qquad \sharp$$

with

$$\Omega \equiv j(I)(J) \qquad \text{(notation (13.89))}. \qquad \spadesuit$$

The polarization then is

$$W \equiv s\,m\,J\,. \qquad \bowtie$$

d) The Lagrange form of U is given by

$$\sigma(\delta x)(\delta' x) \equiv -s\,\mathrm{Tr}(\delta\Omega \cdot \Omega \cdot \delta'\Omega) + m[\overline{\delta X}\,\delta' I - \overline{\delta' X}\,\delta I]\,,^{248} \qquad \clubsuit$$

and the group G acts on U according to

$$\underline{a}_U \left(\Phi \begin{pmatrix} X \\ I \\ J \end{pmatrix} \right) \equiv \Phi \begin{pmatrix} LX + C \\ LI \\ LJ \end{pmatrix} \qquad \text{(notation (13.51))}. \qquad \flat$$

---

[245]The use of the term "mass", which up until now has been defined only in *non-relativistic* mechanics (13.7), will be justified later (14.49).

[246]The dimensions of m and s are AL^{-1} and A, respectively (see (13.112)).

[247]In the sense of the notion "model" as given in (13.2) and (13.71).

[248]Note the following identity, which is a consequence of (13.92):

$$\mathrm{Tr}(\delta\Omega \cdot \Omega \cdot \delta'\Omega) \equiv \mathrm{vol}(I)(J)(\delta I)(\delta' I) - \mathrm{vol}(I)(J)(\delta J)(\delta' J)\,.$$

This allows us to verify by a direct calculation that Maxwell's principle $\mathrm{d}\sigma = 0$ holds (which is here a hypothesis).

e) One has

$$\Phi\begin{pmatrix} X \\ I \\ J \end{pmatrix} = \Phi\begin{pmatrix} X' \\ I' \\ J' \end{pmatrix} \iff \begin{bmatrix} \exists\lambda \in \mathbf{R}, & X' = X + \lambda I \\ & I' = I \\ & J' = J \end{bmatrix}.$$

Thus the set of all X corresponding to a given motion is a *straight line* T in Minkowski space parallel to I; it is called the *trajectory* of the particle.[249]

PROOF: Every nonzero vector which is orthogonal to a timelike vector is spacelike. Consequently, the numbers m and s defined by $\diamondsuit$ are not zero; in fact $s > 0$. These numbers are independent of the motion of the system because of (13.107) and (13.109).[250] Thus the vectors I and J defined by $P \equiv m\,I$ and $W \equiv sm\,J$ satisfy $\heartsuit$.

Let T be the set

$$[X \in T] \iff [M_X \cdot P = 0]\,. \tag{14.5}$$

Using expression (13.86) of the moment M_X with respect to X:

$$M_X \equiv M + P \cdot \overline{X} - X \cdot \overline{P}\,, \tag{14.6}$$

one can easily verify that

$$[X \in T] \iff \left[\exists\lambda \in \mathbf{R}\,, \quad X = \frac{M \cdot P}{\overline{P} \cdot P} + \lambda P\right], \tag{14.7}$$

that is, T is a straight line parallel to P and thus parallel to I.

It follows from (14.6) and (14.7) that M_X does not depend on X when X varies over T; let us denote it by $s\,\Omega$. We have

$$\overline{\Omega} \equiv -\Omega\,, \qquad \Omega \cdot I \equiv 0\,, \qquad \text{and} \qquad *(\Omega) \cdot I \equiv J\,, \tag{14.8}$$

from which one can easily deduce that

$$\Omega \equiv j(I)(J)\,. \tag{14.9}$$

Hence, replacing M_X by its value $s\,\Omega$ in (14.6) we obtain the equations $\sharp$ and $\spadesuit$.

Thus we have established that the data X, I, and J (or T and J) define the moment $\mu \equiv \{M, P\}$. These data will determine the motion x

[249] Thus the motion is characterized by the *trajectory* T *and the vector* J, which is orthogonal to T and whose length-squared is -1.

[250] The sign of the energy E does not depend on the motion because the restricted Lorentz group is contained in the orthochronous group.

if the map $x \mapsto \mu$ is injective. This will follow from theorem (11.38) once we have shown that the stabilizer of a moment μ_0 under the coadjoint representation is *connected*.

Toward this goal we first verify that

$$\underline{a}_V \begin{pmatrix} X \\ I \\ J \end{pmatrix} \equiv \begin{pmatrix} LX + C \\ LI \\ LJ \end{pmatrix} \qquad \text{(notation (13.51))} \tag{14.10}$$

defines an action of the restricted Poincaré group G on the manifold V (definition (c)), that G acts *transitively* on V,[251] and that

$$\theta \cdot \underline{a}_V \equiv \underline{a}_{G^*} \cdot \theta\,, \tag{14.11}$$

where θ is the map $y \mapsto \mu$.

We thus obtain an element μ_0 of the coadjoint orbit by choosing X_0, I_0, and J_0 arbitrarily in V; for example

$$X_0 = 0\,, \qquad I_0 = \mathcal{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad J_0 = \mathcal{R} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

This choice leads to $\mu_0 = \{M_0, P_0\}$ with

$$P_0 = \mathcal{R} \begin{pmatrix} 0 \\ m \end{pmatrix} \qquad \text{and} \qquad M_0 = \mathcal{R} \cdot \begin{pmatrix} 0 & -s & 0 & 0 \\ s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \mathcal{R}^{-1}\,. \tag{14.12}$$

Using (13.107) and (13.67) it is easy to deduce that the isotropy group of μ_0 is the set of

$$a = \mathcal{R}' \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & e \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot (\mathcal{R}')^{-1} \qquad \varphi, e \in \mathbf{R}\,, \tag{14.13}$$

[251] If I_1, J_1, I_2, J_2 are two pairs of vectors satisfying $\heartsuit$, one can construct two orthonormal bases $S_1 = [I_1\, J_1\, J_1'\, J_1'']$ and $S_2 = [I_2\, J_2\, J_2'\, J_2'']$ such that

$$\mathrm{vol}(I_1)(J_1)(J_1')(J_1'') = \mathrm{vol}(I_2)(J_2)(J_2')(J_2'') = 1\,.$$

Then $L = S_2 \cdot S_1^{-1}$ is an element of the restricted Lorentz group which satisfies $LI_1 = I_2$ and $LJ_1 = J_2$.

which is indeed a connected Lie group. It follows that $x \mapsto \mu$ is a diffeomorphism. This shows that x, just as μ, depends only on T and J, that the map $y \mapsto x$ is differentiable (because $y \mapsto \mu$ and $\mu \mapsto x$ are), and that U has dimension 8 (because the dimension of the isotropy group is 2 and the dimension of G is 10). Thus there exists a unique model of a relativistic particle with spin of mass m and spin s. Finally, note that (14.13) and (14.3) give equation $\flat$.

Now note that the formulæ (14.12) insure the *existence* of a particle with spin of mass m and spin s ($m \neq 0$, $s > 0$). To see this, it suffices to take the coadjoint orbit of μ_0 as the *space of motions*. Since this orbit is a symplectic manifold and admits G as a dynamical group (theorem (11.34)), the axioms (13.70) of relativistic mechanics hold. Looking at $\diamondsuit$ we see that the numbers m and s appearing in (14.12) are indeed the mass and the spin of the particle, respectively. It only remains to compute the Lagrange form σ. Since $x \mapsto \mu$ is a symplectomorphism, we have

$$\sigma(\delta x)(\delta' x) \equiv \sigma(\delta\mu)(\delta'\mu)\,,$$

or, due to (11.34$\diamondsuit$),

$$\sigma(\delta x)(\delta' x) = \delta\mu(Z) \qquad \text{if} \quad \delta'\mu = \mu \cdot \mathrm{ad}(Z)\,. \tag{14.14}$$

Taking $Z = \begin{bmatrix} \Lambda & \Gamma \\ 0 & 0 \end{bmatrix}$, we find

$$\delta' M = \Lambda M - M\Lambda + \Gamma \overline{P} - P\overline{\Gamma}\,, \qquad \delta' P = \Lambda P\,, \qquad \delta'\Omega = \Lambda\cdot\Omega - \Omega\cdot\Lambda\,, \tag{14.15}$$

and

$$\sigma(\delta x)(\delta' x) = \delta\mu(Z) \;=\; -\tfrac{1}{2}\,\mathrm{Tr}(\delta M\cdot\Lambda) - \overline{\delta P}\,\Gamma \qquad \text{by (13.58)}$$

or, using the expressions $\sharp$ for M and P:

$$\sigma(\delta x)(\delta' x) = -\frac{s}{2}\,\mathrm{Tr}(\delta\Omega\cdot\Lambda) + m\,[\overline{\delta X}\,\delta' I - \overline{\delta' X}\,\delta I]\,. \tag{14.16}$$

The map Ω can be interpreted as a projection onto the plane orthogonal to I and J, followed by a rotation through $\pi/2$ in this plane. With the aid of (13.92) one can verify that Ω satisfies

$$\overline{\Omega} \equiv -\Omega \equiv \Omega^3 \tag{14.17}$$

and

$$\Omega\cdot\delta\Omega\cdot\Omega \equiv \Omega^2\cdot\delta\Omega + \delta\Omega\cdot\Omega^2 + \delta\Omega \equiv 0\,. \tag{14.18}$$

Hence[252]

$$\begin{aligned}\mathrm{Tr}(\delta\Omega\cdot\Lambda) &= -\mathrm{Tr}(\Omega^2\cdot\delta\Omega\cdot\Lambda)-\mathrm{Tr}(\delta\Omega\cdot\Omega^2\cdot\Lambda)\\ &= -2\,\mathrm{Tr}(\Omega^2\,\delta\Omega\,\Lambda) = -2\,\mathrm{Tr}(\Lambda\cdot[\Omega^2\,\delta\Omega-\Omega\;\delta\Omega\;\Omega])\\ &= -2\,\mathrm{Tr}([\Lambda\,\Omega-\Omega\,\Lambda]\,\Omega\;\delta\Omega) = -2\,\mathrm{Tr}(\delta'\Omega\;\Omega\;\delta\Omega)\,.\end{aligned}$$

Substituting this into (14.16) gives ♣. □

Remarks

(14.19) Let us use the same symbol σ for the inverse image of the Lagrange form σ under the map $y\mapsto x$ from V to U:

$$\sigma(\delta y)(\delta' y)\equiv\sigma(\delta x)(\delta' x)\,.$$

With this definition, V becomes a presymplectic manifold admitting the restricted Poincaré group as a dynamical group (due to (14.10)). Each leaf of the characteristic foliation of σ on V is the inverse image of a motion under the map $y\mapsto x$. Since there exists a differentiable map $y\mapsto t$ from V to $\mathbf{R}$ (through the intermediary of the map $y\mapsto X$ and the choice of a Lorentz frame), we see that V is an *evolution space* of the particle (see Fig. 12.I and (12.77)). And we obtain this evolution space without having to require its existence in the axioms (13.70) of relativistic mechanics.

Let X be a point of Minkowski space which varies over the trajectory T of a motion x. Writing

(14.20)
$$X\equiv\mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix},$$

we see that the point $\mathbf{r}$ has a *uniform and rectilinear motion* with velocity

(14.21)
$$\mathbf{v}=\frac{\mathbf{p}}{E}\,,$$

(because T is parallel to P). Definition $\diamondsuit$ of m shows that the velocity of the particle is less than the velocity of light and gives *Einstein's formula*

(14.22)
$$E=\frac{m}{\sqrt{1-\|\mathbf{v}\|^2}}\,.$$

Note that the polarization remains constant throughout the motion. □

[252]Recall the identities $\mathrm{Tr}(A\cdot B)\equiv\mathrm{Tr}(B\cdot A)$ and $\mathrm{Tr}(\overline{A})\equiv\mathrm{Tr}(A)$.

(14.23) If the particle is *at rest*, that is, $\mathbf{r} \equiv$ constant and $\mathbf{v} = 0$, its energy E is equal to m and its linear momentum $\mathbf{p}$ is zero. It follows that its angular momentum $\mathbf{l}_X$ with respect to the point X of the trajectory (notation (13.87)) is a *constant vector* of *length* s; it is called the *intrinsic angular momentum* or (vector) *spin* of the particle. Recall that the length s of this vector is also called spin (scalar spin if one wants to be more precise).

Case II. A particle without spin

(14.24) DEFINITION: A relativistic *particle without spin* (or a *relativistic material point*) is an elementary dynamical system such that

$$\overline{P} \cdot P > 0 \qquad \text{and} \qquad W \equiv 0 .$$

(14.24) THEOREM: For a particle without spin we have the following results.

a) The number

$$\diamondsuit \qquad m = \operatorname{sign}(E)\sqrt{\overline{P} \cdot P}$$

does not depend on the motion. It is called the *mass* of the particle.

b) Given m ($m \neq 0$), there exists a *unique model* of a particle without spin of mass m. The space of motions U has dimension 6. The map $x \mapsto \mu$, where x is a motion and μ its moment, is a diffeomorphism from U onto a coadjoint orbit of G.

c) Let V be the 7-dimensional manifold described by

$$y \equiv \begin{pmatrix} X \\ I \end{pmatrix},$$

where X is a point of Minkowski space and I is a 4-vector such that

$$\heartsuit \qquad \overline{I} \cdot I \equiv 1 \qquad I \text{ future pointing.}$$

A motion x can be characterized by a point y from V such that the map $\Phi[y \mapsto x]$ is differentiable and such that

$$\# \qquad M \equiv m[X \cdot \overline{I} - I \cdot \overline{X}] \qquad \text{and} \qquad P \equiv m\, I .$$

d) The Lagrange form is given by

$$\clubsuit \qquad \sigma(\delta x)(\delta' x) = m\,[\overline{\delta X} \cdot \delta' I - \overline{\delta' X} \cdot \delta I],$$

and the group G acts on U according to

$$\flat \qquad \underline{a}_U\Big(\Phi \begin{pmatrix} X \\ I \end{pmatrix}\Big) \equiv \Phi \begin{pmatrix} LX + C \\ LI \end{pmatrix}.$$

e) We have

$$\Phi\begin{pmatrix} X \\ I \end{pmatrix} = \Phi\begin{pmatrix} X' \\ I' \end{pmatrix} \iff \begin{bmatrix} \exists\lambda \in \mathbf{R}\,, & X' = X + \lambda I \\ & I' = I \end{bmatrix}.$$

Thus the set of all X corresponding to a given motion is a straight line T in Minkowski space parallel to I; it is called the trajectory of the particle.[253]

PROOF: The argument follows along the lines of the proof of (14.4). We leave the details to the reader. Let us only note that we can choose

(14.25) $$P_0 = \begin{pmatrix} 0 \\ m \end{pmatrix} \qquad \text{and} \qquad M_0 = 0$$

as a moment of a particular motion (instead of (14.12)) and that the stabilizer subgroup of this moment is the set of

(14.26) $$a = \mathcal{R}' \cdot \begin{pmatrix} A & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix} \cdot (\mathcal{R}')^{-1} \qquad A \in \mathrm{SO}(3)\,,\ e \in \mathbf{R}\,.$$

Thus this stabilizer is also connected; it is 4-dimensional. □

(14.27) As in the case of a particle with spin, we see that V plays the role of an evolution space and that the trajectory of the particle can be interpreted, in a Lorentz frame, as a uniform rectilinear motion with velocity $\mathbf{v} = \frac{\mathbf{p}}{E}$ which is less than the velocity of light. The energy of the particle is given by Einstein's formula (14.22)

(14.28) $$E = \frac{m}{\sqrt{1 - \|\mathbf{v}\|^2}}\,,$$

but its intrinsic angular momentum $\mathbf{l}_X$ is zero.

[253]Thus the motion is entirely determined by its trajectory.

Case III. A massless particle

(14.29) DEFINITION: A *massless particle* [254] is an elementary dynamical system such that

$$\overline{P} \cdot P = \overline{W} \cdot W = 0$$

with both P and W nonzero.

(14.29) THEOREM: For a massless particle we have the following results.

a) There exist three numbers η, χ, and s, with $\eta = \pm 1$, $\chi = \pm 1$, and $s > 0$, such that for every motion of the particle:

$\diamondsuit$ $$\operatorname{sign}(E) \equiv \eta \qquad \text{and} \qquad W = \chi s\, P\,;$$

η is the *sign of the energy*, χ is the *helicity* [255] and s the *spin* [256] of the particle.

b) There exists a *unique model* of a massless particle for which the quantities η, χ, and s assume given values. The dimension of the manifold of motions U is 6. If μ is the moment of a motion x, then the map $[x \mapsto \mu]$ is a symplectomorphism of U onto a coadjoint orbit of G.

c) Let V be the 9-dimensional manifold described by the 3-tuple

$$y \equiv \begin{pmatrix} X \\ I \\ J \end{pmatrix},$$

where X is a point of Minkowski space, and I, J are two 4-vectors satisfying

$\heartsuit$ $$\begin{aligned} &\overline{I} \cdot I = 0\,, && I \text{ future pointing} \\ &\overline{I} \cdot J = -1 \\ &\overline{J} \cdot J = 0\,. \end{aligned}$$

A motion x can be characterized by a point y of V such that the map $\Phi\,[y \mapsto x]$ is differentiable and such that the moment $\mu \equiv \{M, P\}$ is given by

$\sharp$ $$M = \chi s\, \Omega + \eta\,[X\overline{I} - I\overline{X}] \qquad \text{and} \qquad P \equiv \eta\, I$$

[254] If we pretend that the mass is given by (14.4$\diamondsuit$), we see that a "massless particle" would have to be characterized by the single condition $\overline{P} \cdot P = 0$, without W necessarily being nonzero and lightlike. However, the known massless particles behave according to the model given here. Thus it is customary to assume *implicitly* that the conditions $\overline{W} \cdot W = 0$ and $W \neq 0$ hold.

[255] The particle is said to be left polarized if $\chi = -1$ and right polarized if $\chi = 1$.

[256] The dimension of spin is again A, as in case I (see (13.112)).

with

♠ $$\Omega \equiv j(I)(J)\,.$$

The polarization then is

⋈ $$W \equiv \eta\chi s\, I\,.$$

d) The Lagrange form of U is given by

♣ $$\sigma(\delta x)(\delta' x) \equiv -\chi s\,\mathrm{Tr}(\delta\Omega\cdot\Omega\cdot\delta'\Omega) + \eta\,[\overline{\delta X}\cdot\delta' I - \overline{\delta' X}\cdot\delta I]\,,$$

and the group G acts on U according to

♭ $$\underline{a}_U\Big(\Phi\begin{pmatrix} X \\ I \\ J \end{pmatrix}\Big) \equiv \Phi\begin{pmatrix} LX+C \\ LI \\ LJ \end{pmatrix} \qquad \text{(notation (13.51))}.$$

e) We have:

$$\Phi\begin{pmatrix} X' \\ I' \\ J' \end{pmatrix} = \Phi\begin{pmatrix} X \\ I \\ J \end{pmatrix} \Longleftrightarrow \left[\begin{array}{lrl} & \overline{I}\cdot Z &= 0 \\ & X' &= X+\eta\chi s\, Z \\ \exists\, Z & I' &= I \\ & J' &= J + j(I)(J)(Z) + \dfrac{\overline{Z}\cdot Z}{2} I \end{array}\right].$$

Hence the set of all X corresponding to a single motion x is a 3-plane orthogonal to I; we shall call it the *wave front* of the particle.[257]

PROOF: Let us give a sketch of the proof. In a Euclidean space with inertia index 1, two vectors that are lightlike and orthogonal are parallel to each other. Then (a) follows from (13.107) and (13.110).

The *wave front* of the particle can be defined directly by one of the equations

(14.30) $$M_X\cdot P = 0 \qquad \text{or} \qquad \mathrm{pf}(M_X) = 0\,.$$

If X belongs to this wave front, then M_X can be written in the form $j(P)(Y)$ for some vector Y. The identity $*(M_X)P = \chi s\, P$ shows that Y can be chosen to be lightlike (by adding a multiple of P to it). Defining $I = \eta\, P$ and $J = \frac{\chi\eta}{s} Y$, we obtain the equations ♡, ♠, and ⋈, and finally (e).

Defining

(14.31) $$\underline{a}_V\begin{pmatrix} X \\ I \\ J \end{pmatrix} \equiv \begin{pmatrix} LX+C \\ LI \\ LJ \end{pmatrix}$$

[257] In a Lorentz frame the wave front can be interpreted as a 2-plane in ordinary space which moves parallel to itself with normal velocity equal to the velocity of light.

we obtain a transitive action of G on V. Since $\Phi(\underline{a}_V(y)) \equiv \underline{a}_U(\Phi(y))$, we can choose the particular moment μ_0:

$$M_0 = \mathcal{R} \cdot \begin{pmatrix} 0 & -\chi s & 0 & 0 \\ \chi x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \mathcal{R}^{-1} \qquad \text{and} \qquad P_0 = \mathcal{R} \begin{pmatrix} 0 \\ 0 \\ \eta \\ \eta \end{pmatrix}. \tag{14.32}$$

This allows us to show that the isotropy group of μ_0 is diffeomorphic to $\mathrm{SO}(2) \times \mathbf{R}^3$ and thus is connected. Theorem (11.38) then gives (b) and (c).

Finally, formula ♣ is established like (14.4♣), knowing that here too Ω satisfies (14.17) and (14.18). □

Remarks

(14.33) Let us choose a Lorentz frame $\mathcal{R}$. By a *trajectory* of a massless particle we mean the set T of points X of Minkowski space such that $\mathbf{g}_X = 0$ (notation (13.87)). Such a trajectory is a straight line parallel to I and contained in the wave front. In the reference frame $\mathcal{R}$ the trajectory appears as a uniform and rectilinear motion with velocity $\mathbf{v}$ such that $\|\mathbf{v}\| = 1$.[258] Note the equations

$$\mathbf{p} = E\,\mathbf{v} \qquad \text{(Einstein)} \tag{14.34}$$

and

$$\mathbf{l}_X = \chi s\,\mathbf{v} \qquad \text{if} \quad X \in T\,. \tag{14.35}$$

Again $\mathbf{l}_X$ is called the *intrinsic angular momentum* or *spin*; its length is s. We see that the helicity χ determines whether the spin is in the same direction as the velocity or opposite.

(14.36) ☡ But beware, *the trajectory T depends upon the choice of the reference frame $\mathcal{R}$.*[259] By varying the latter, T sweeps out the entire *wave front* of the motion. □

(14.37) We will not study other kinds of elementary systems, which do not correspond to known particles. Let us only note that $\overline{P} \cdot P < 0$ (particles with imaginary mass) would correspond to particles that move *faster*

[258] It is generally accepted that *photons* are massless particles. Thus the "velocity of a photon" is equal to the "velocity of light".

[259] The equation $X \in T$ can be written as $M_X \cdot \mathbf{u} = 0$, where $\mathbf{u}$ is the unit vector $\mathcal{R}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

than light and that could have, in a particular Lorentz frame, *infinite* velocity ("tachyons"). □

Nonrelativistic particles

Let us start by considering a *relativistic particle without spin* (14.24) and let us label a point y in the evolution space V by the variables $\mathbf{r}$, t, and $\mathbf{v}$ where

$$\text{(14.38)} \qquad \begin{cases} \mathbf{v} = \text{velocity of the particle (14.27)} \\ X \equiv \mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix}. \end{cases}$$

One can verify immediately that

$$\text{(14.39)} \qquad \sigma(\delta y)(\delta' y) = \langle \delta \frac{m\,\mathbf{v}}{\sqrt{1-v^2}},\, \delta'\mathbf{r} - \mathbf{v}\,\delta t' \rangle - \langle \delta' \frac{m\,\mathbf{v}}{\sqrt{1-v^2}},\, \delta\mathbf{r} - \mathbf{v}\,\delta t \rangle ,$$

and that an element a of the Poincaré group acts on V (using notation (13.67)) according to

$$\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \mapsto \mathcal{L}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} + \begin{pmatrix} \mathbf{c} \\ e \end{pmatrix}$$

$$\text{(14.40)} \qquad \begin{pmatrix} \dfrac{\mathbf{v}}{\sqrt{1-v^2}} \\ \dfrac{1}{\sqrt{1-v^2}} \end{pmatrix} \mapsto \mathcal{L}\begin{pmatrix} \dfrac{\mathbf{v}}{\sqrt{1-v^2}} \\ \dfrac{1}{\sqrt{1-v^2}} \end{pmatrix}$$

with

$$\mathcal{L} = \exp\begin{pmatrix} 0 & \mathbf{b} \\ \overline{\mathbf{b}} & 0 \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}.$$

Let us *change the units of length and time without changing the unit of action*. Such a transformation can be written as

$$\text{(14.41)} \qquad \begin{aligned} m &\equiv \frac{m^*}{M}, \\ \mathbf{r} \equiv \frac{\mathbf{r}^*}{L}, \quad t &\equiv \frac{t^*}{T}, \quad \mathbf{v} \equiv \frac{T}{L}\mathbf{v}^*, \\ \mathbf{c} = \frac{\mathbf{c}^*}{L}, \quad e &= \frac{e^*}{T}, \quad \mathbf{b} = \frac{T}{L}\mathbf{b}^*, \end{aligned}$$

where M, L, and T are numbers satisfying

$$\text{(14.42)} \qquad M\,L^2\,T^{-1} = 1$$

(see (13.13)).

If we substitute these values into the formulæ (14.39) and (14.40) and then let L/T *tend toward infinity*, these formulæ become

(14.43) $$\sigma(\delta y)(\delta' y) = \langle m^* \, \delta \mathbf{v}^*, \, \delta' \mathbf{r}^* - \mathbf{v}^* \, \delta' t^* \rangle - \langle m^* \, \delta' \mathbf{v}^*, \, \delta \mathbf{r}^* - \mathbf{v}^* \, \delta t^* \rangle$$

and

(14.44) $$\begin{pmatrix} \mathbf{r}^* \\ t^* \\ \mathbf{v}^* \end{pmatrix} \mapsto \begin{pmatrix} A\mathbf{r}^* + \mathbf{b}^* \, t^* + \mathbf{c}^* \\ t^* + e^* \\ A\mathbf{v}^* + \mathbf{b}^* \end{pmatrix}.$$

In other words, we recover the equations of §12 defining the nonrelativistic *model* of a free material point of mass m.

(14.45) We conclude that the method outlined above allows us to consider a Newtonian material point as an approximation of a relativistic particle without spin when the number $c = L/T$, which is the velocity of light in the new units, is very large. □

(14.46) Let us recall that condition (14.42) expresses the fact that the values of the 2-form σ do not change (see (13.13)).

(14.47) In the case of a particle with spin (14.4), let us introduce as an auxiliary variable the unit vector $\mathbf{u}$ which has the same direction as $\mathbf{l}_X$ (14.23). Arguing in a similar manner as above, we obtain the following limiting equations:

$$\sigma(\delta y)(\delta' y) = -s \, \langle \mathbf{u}, \, \delta \mathbf{u} \times \delta' \mathbf{u} \rangle + \langle m^* \delta \mathbf{v}^*, \, \delta' \mathbf{r}^* - \mathbf{v}^* \, \delta' t^* \rangle - \langle m^* \, \delta' \mathbf{v}^*, \, \delta \mathbf{r}^* - \mathbf{v}^* \, \delta t^* \rangle$$

(14.48) $$\begin{pmatrix} \mathbf{r}^* \\ t^* \\ \mathbf{v}^* \\ \mathbf{u} \end{pmatrix} \mapsto \begin{pmatrix} A\mathbf{r}^* + \mathbf{b}^* \, t^* + \mathbf{c}^* \\ t^* + e^* \\ A\mathbf{v}^* + \mathbf{b}^* \\ A\mathbf{u} \end{pmatrix}.$$

These equations define a model of an isolated classical dynamical system. We will take them as the *definition of a nonrelativistic particle with spin*.

(14.49) The system (14.48) is easily interpreted in terms of §13. The *mass* of the particle is m^*, its *barycenter*, situated at the point $\mathbf{r}^*$ at time t^*, moves uniformly and rectilinearly with velocity $\mathbf{v}^*$. Its *proper motion* is characterized by the unit vector $\mathbf{u}$, which is *constant* during the motion. Its *proper angular momentum* (see (13.35)) is

(14.50) $$\mathbf{l}_{\text{proper}} = s \, \mathbf{u} \, .$$

Thus the length of $\mathbf{l}_{\text{proper}}$ equals the spin s. Its *proper energy* is zero.

If one looks a priori for classical dynamical systems (in the sense of (13.1)) on which the Galilei group acts transitively, one arrives at the following results.

(14.51) CASE a): If the mass of the system is not zero, only the models already found above exist, namely, the material point and the nonrelativistic particle with spin.

CASE b): So let us study the case $m = 0$. Using the notation of (12.115) and (12.123), the coadjoint representation of the Galilei group is given by the formula

(14.52)
$$\begin{aligned} &\underline{a}_{G^*}\{\mathbf{l}, \mathbf{g}, \mathbf{p}, E\} = \\ &\quad = \{A\mathbf{l} + A\mathbf{g} \times \mathbf{b} + \mathbf{c} \times A\mathbf{p},\ A[\mathbf{g} - \mathbf{p}\,e],\ A\mathbf{p},\ E + \langle \mathbf{b}, A\mathbf{p}\rangle\}\,. \end{aligned}$$

From (14.52) we deduce that the numbers $\|\mathbf{p}\|^2$ and $\|\mathbf{p} \times \mathbf{g}\|^2$ are characteristic for the system.

Let us consider the special case $\mathbf{p} \neq 0$ and $\mathbf{p} \times \mathbf{g} = 0$. We see that then the number $\langle \mathbf{l}, \mathbf{p}\rangle$ is also a characteristic of the system. If we assume that it is nonzero we obtain the following unique model.

(14.53) Let χ, s, and k be three real numbers with $\chi = \pm 1$, $s > 0$, and $k > 0$. We will call χ, s, and k the *helicity*, *spin* and *color* [260] of the particle. A motion of the particle can be described by a *unit vector* $\mathbf{u}$, a *time* t, a *straight line* D parallel to $\mathbf{u}$, and the *energy* E of the particle. The Galilei group acts on these quantities as follows:

$$\begin{aligned} t &\mapsto t + e\,, & \mathbf{u} &\mapsto A\mathbf{u}\,, \\ E &\mapsto E + k\,\langle \mathbf{b}, A\mathbf{u}\rangle\,, & \mathbf{r} &\mapsto A\mathbf{r} + \mathbf{b}\,t + \mathbf{c} \qquad (\mathbf{r} \in D)\,. \end{aligned}$$

The corresponding Galilean moments are

$$\heartsuit \qquad \mathbf{p} = k\,\mathbf{u}\,, \quad \mathbf{g} = -\mathbf{p}\,t\,, \quad \mathbf{l} = \chi s\,\mathbf{u} + \mathbf{r} \times \mathbf{p} \qquad (\mathbf{r} \in D)\,,$$

and the Lagrange form is given by

$$\begin{aligned} \clubsuit \qquad \sigma(\delta x)(\delta' x) \;&=\; k\,[\langle \delta\mathbf{u}, \delta'\mathbf{r}\rangle - \langle \delta'\mathbf{u}, \delta\mathbf{r}\rangle] + \\ &\quad + \delta t\ \delta' E - \delta' t\ \delta E - \chi s\,\langle \mathbf{u}, \delta\mathbf{u} \times \delta'\mathbf{u}\rangle\,. \end{aligned}$$

□

(14.54) The above model can be considered as a nonrelativistic limit of the relativistic massless particle described in (14.29), but a different kind of limit than the one encountered in (14.43) and (14.48): a single relativistic particle gives rise to a *family* of *distinct* nonrelativistic particles, each

[260] In the case of the photon we will see that k is indeed the *color* in the usual sense of the word: k is inversely proportional to its "wave length" (19.43).

one labeled by a color k. The straight line D and the time t can be considered as characterizing the instantaneous motion of the particle. Thus in Galilean mechanics the velocity of light is *infinite*.[261]

(14.55) As for tachyons (14.37), it does not seem that one can obtain a nonrelativistic limit for them.

Mass and barycenter of a relativistic system

(14.56) Let U be the space of motions of an isolated relativistic dynamical system. Assume that for a motion x of the system the energy-momentum vector P is timelike. We then know that the moment $\mu = \{M, P\}$ of the Poincaré group can be seen as the moment of a particle described by (14.4) or (14.24). By *mass* and *spin* of this system we will mean the mass and spin of this particle, namely

$$m = \text{sign}(E)\sqrt{\overline{P} \cdot P} \qquad \text{and} \qquad s = \sqrt{\frac{-\overline{W} \cdot W}{\overline{P} \cdot P}}\,, \tag{14.57}$$

where s might be zero (if the polarization W vanishes).

In the same spirit, by the *relativistic barycenter* or *center of mass* of the system we will mean the point which varies over the trajectory of this fictitious particle, a trajectory which is given by the equation (see (14.4))

$$M_X \cdot P = 0\,. \tag{14.58}$$

We thus use a language close to that of classical mechanics. But there is an important difference: the *mass* depends, in general, on the chosen motion x.[262]

EXAMPLE: Consider a system consisting of particles of positive mass m_j without interaction. An immediate consequence of the axioms (13.70) is that the moment of the system is the sum of the moments of the individual particles:

$$M = \sum_j M_j \qquad \text{and} \qquad P = \sum_j P_j\,. \tag{14.59}$$

One can show that any sum of future pointing vectors is future pointing. Hence this composite system too has positive mass. If we choose a

[261] That is why the Doppler effect does not take place, which enables the color to be an invariant.

[262] More precisely, on the *orbit* of x

Lorentz frame in which the center of mass has zero velocity, that is,

$$\sum_j \frac{m_j \mathbf{v}_j}{\sqrt{1 - \|\mathbf{v}_j\|^2}} = 0\,, \tag{14.60}$$

where $\mathbf{v}_j$ is the velocity of the j^{th} particle in this reference frame, we find immediately that

$$m = \sum_j \frac{m_j}{\sqrt{1 - \|\mathbf{v}_j\|^2}}\,. \tag{14.61}$$

Thus we see that the total mass of the system is greater than the sum of the constituent masses, and in fact depends on the motion of the system. □

Inversions of space and time

Let U be the space of motions of an isolated relativistic dynamical system. Up until now we have assumed that the *restricted* Poincaré group acts on U while preserving the Lagrange form (13.70).

We also know that the *full* Poincaré group has 4 connected components (13.53). It is generated by the restricted Poincaré group G and by the following two operations, namely, *space inversion*[†]

$$I_s\left(\mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}\right) = \mathcal{R}\begin{pmatrix}-\mathbf{r}\\ t\end{pmatrix}, \tag{14.62}$$

and *time reversal*

$$I_t\left(\mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}\right) = \mathcal{R}\begin{pmatrix}\mathbf{r}\\ -t\end{pmatrix}. \tag{14.63}$$

(14.64) One might ask whether the *full* Poincaré group too is a dynamical group of U. It should be noted that general relativity argues in favor of this. Thus let us consider a symplectic manifold U (not necessarily connected) which admits the full Poincaré group G' as a dynamical group.

It is clear that the restricted Poincaré group G is also dynamical group of U and even that it is a dynamical group of *each connected component* of U (theorem (1.51)). We know that the action of G has a moment defined up to an additive constant *on each component* of U, and that this constant can be chosen such that

$$\psi(\underline{a}_U(x)) = \underline{a}_{\mathcal{G}^*}(\psi(x)) \qquad \forall\, a \in G\,,\ \forall\, x \in U\,, \tag{14.65}$$

[†]Editors' note: also called *parity reversal*.

where ψ is the map that assigns to each point x of U its moment μ. It is clear that μ is also a moment of the group G' because the definition of moment (11.7) only involves the Lie algebra of the group.

TECHNICAL DISCUSSION: Let I be an arbitrary element of G' and define

$$F(x) \equiv \psi(\underline{I}_U(x)) - \underline{I}_{\mathcal{G}^*}(\psi(x)) .$$

The proof of (11.7) shows that $F(x)$ is constant on each component of U.

Now let a be an arbitrary element of G. Then formula (14.65) shows that

$$\underline{a}_{\mathcal{G}^*}(F(x)) = \psi(\underline{a \times I}_U(x)) - \underline{a \times I}_{\mathcal{G}^*}(\psi(x)) ,$$

or, writing $b = I^{-1} \times a \times I$,

$$\underline{a}_{\mathcal{G}^*}(F(x)) = \psi(\underline{I \times b}_U(x)) - \underline{I \times b}_{\mathcal{G}^*}(\psi(x)) .$$

Theorem (6.35) shows that $b \in G$. Applying (14.65) again we obtain

$$\underline{a}_{\mathcal{G}^*}(F(x)) = F(\underline{b}_U(x)) .$$

Theorem (1.51) shows that $\underline{b}_U(x)$ belongs to the same component of U as x. Thus $F(\underline{b}_U(x)) = F(x)$, whence

$$\underline{a}_{\mathcal{G}^*}(F(x)) = F(x) \qquad \forall\, a \in G\,, \forall\, x \in U\,.$$

This expresses the fact that the *coboundary* of the torsor $F(x)$ of the group G is *zero* (see definition (11.19)). From this it is elementary to deduce (using, for example, expression (13.107) of the coadjoint representation) that $F(x)$ is zero. □

We thus have proved the identity

$$\psi(\underline{I}_U(x)) = \underline{I}_{\mathcal{G}^*}(\psi(x)) \qquad \forall\, x \in U\,,\ \forall\, I \in G'\,, \tag{14.66}$$

which extends the validity of (14.65) to the full Poincaré group.

This formula applies when $I = I_s$ (space inversion) or when $I = I_t$ (time reversal). Calculating $\underline{I_s}_{\mathcal{G}^*}$ and $\underline{I_t}_{\mathcal{G}^*}$ gives (notation (13.59)):

$$\begin{cases} \underline{I_s} : \mathbf{l} \to \mathbf{l}\,, & \mathbf{g} \to -\mathbf{g}\,, & \mathbf{p} \to -\mathbf{p}, & E \to E\,, \\ \underline{I_t} : \mathbf{l} \to \mathbf{l}\,, & \mathbf{g} \to -\mathbf{g}\,, & \mathbf{p} \to \mathbf{p}\,, & E \to -E\,. \end{cases} \tag{14.67}$$

(14.68) Let us now extend definition (14.1) of an elementary system by postulating that it is the full Poincaré group G' which acts transitively and canonically on U. □

(14.69) From (14.66) it follows immediately that the moment $\mu \equiv \psi(x)$ varies over a coadjoint orbit *of* G' when x varies over U. If this orbit is not connected, then *neither is* U since ψ is continuous (theorem (6.51)).

On the other hand, we know that every orbit in U of the restricted Poincaré group G is contained in a connected component U_0 of U. Since G and G' have the same Lie algebra, these orbits have the same dimension as U (theorem (6.20)). They thus are pairwise disjoint open sets filling out U_0. But this is impossible if there is more than one orbit because U_0 is connected. Hence each connected component of U is an orbit of the restricted Poincaré group and thus defines *an elementary system in the restricted sense* (14.1).

(14.70) It follows that the space of motions of an elementary system for the full Poincaré group is obtained by taking the *sum* [263] of the spaces of motions of several elementary systems in the restricted sense (14.1). □

Let us treat some examples.

A particle with nonzero mass m

(14.71) Let us begin by observing that the two Casimirs $\overline{P} \cdot P$ and $\overline{W} \cdot W$ are constant on every coadjoint orbit of G' and hence on U (14.69). If we assume that $\overline{P} \cdot P > 0$, then each connected component of U corresponds to a *particle* of mass $m = \pm\sqrt{\overline{P} \cdot P}$ and spin $s = \sqrt{\frac{-\overline{W} \cdot W}{\overline{P} \cdot P}}$. Equation (14.67) shows that time reversal I_t changes the sign of the energy and thus the sign of the mass (14.4◇). Consequently *it transforms every motion of a particle of mass* m *into a motion of a particle of mass* $-m$.

On the other hand, space inversion I_s preserves the mass m and the spin s. Hence the associated orbit of G' has *two components*.

(14.72) If ψ is bijective, then U has two components as well. But it is possible that U has four components. A model of this possibility can be constructed as follows:

$$U = \text{the set of pairs } x \equiv (\mu, \varepsilon) \quad [\mu \text{ lies in a coadjoint orbit and } \varepsilon = \pm 1]$$

(14.73)
$$\sigma(\delta x)(\delta' x) \equiv \sigma(\delta\mu)(\delta'\mu)$$
$$\underline{a}_U(\mu, \varepsilon) = \big(\underline{a}_U(\mu), \chi_s(a)\varepsilon\big), \text{ where } \chi_s(a) \text{ is the } \textit{spatial character} \text{ of } a.^{264}$$ □

[263] By the *sum* of several sets one means their *union*, provided that their pairwise intersections are empty.

There is a way to safeguard the connectedness of U and thus to avoid particles of negative mass. First one postulates that ψ is injective and then one lets the *full* Poincaré group G' act on the *unique* component where $m > 0$ by defining

(14.74) $$\underline{a}(\mu) = \chi_t(a)\, \underline{a}_{G^*}(\mu)\,,$$

where $\chi_t(a)$ is the *temporal character* of a.[265]

(14.75) However, we know that these transformations cannot be canonical (14.71). More precisely, if $\chi_t(a) = -1$ (one says that a is *antichronous*), the transformation $\underline{a}$ defined by (14.74) is anticanonical (definition (10.18)).

We thus have to modify the axioms of symplectic mechanics by allowing anticanonical transformations to belong to a dynamical group.

(14.76) There is another way to arrive at the same result. It consists of *excluding* the antichronous transformations by looking only at the subgroup G'' defined by $\chi_t(a) = +1$. G'', which is called the *orthochronous* group, is indeed a dynamical group (without anticanonical elements) of a connected manifold.

It should be obvious that these various conventions have to be judged according to their ability to explain experiments.[266] □

A massless particle

(14.77) Let us assume that $\overline{P} \cdot P = 0$ and $\overline{W} \cdot W = 0$ while P and W are nonzero. Then each component of U corresponds to a massless particle in the sense of (14.29). These particles have the *same spin* s, but the formulæ (14.67) show that space inversion changes the *helicity* and time reversal the *sign of the energy*. It follows that the coadjoint orbit has four components and consequently U as well. Hence ψ is necessarily a bijection.

---

[264] By definition $\chi_s(a)$ equals $+1$ on the components of the group G' containing 1 and I_t and equals -1 on the components containing I_s and $I_t \cdot I_s$. It is called a *character* because it satisfies $\chi_s(ab) \equiv \chi_s(a)\chi_s(b)$.

[265] $\chi_t(a)$ equals 1 on the components of G' containing 1 and I_s and equals -1 on the components containing I_t and $I_t \cdot I_s$. It also satisfies $\chi_t(ab) \equiv \chi_t(a)\chi_t(b)$.

[266] As it happens, recent experiments seem to show that the artifice (14.74) does not apply universally in physics. If — as general relativity suggests — it is the *full* Poincaré group, that is the dynamical group of real systems, it is not possible to call particles with negative mass into question. Thus one might hope to find them in nature, although statistical mechanics shows us they must be rare (see (17.149)).

(14.78) *Experiments show that the known massless particles* (that is, photons and neutrinos) *appear with two helicities* (that is, are polarized to the *left* or to the *right* (see(15.105)). This is an inductive argument[†] to include space inversion in the invariance group of mechanics; otherwise the existence of two massless particles, differing only by their helicity, would be a mere coincidence.

(14.79) We thus have to give up in general the idea of describing a massless particle by a connected manifold. However, one can eliminate the components with negative energy by using one of the above procedures: the introduction of anticanonical transformations by equation (14.74), or the exclusion of antichronous transformations by (14.76). Still more radically, one can restrict oneself to the connected group G (the restricted Poincaré group) and return to the previous description (14.29).

[†]Editors' note: As opposed to a deductive argument.

§15. Particle dynamics

A material point in an electromagnetic field

In classical mechanics (12.85) we have already met the *nonisolated* system consisting of a material point subjected to an electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$. This system is conveniently described by the Lagrange form σ:

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;\equiv\; & \langle m\,\delta\mathbf{v} - q\,\mathbf{E}\,\delta t,\ \delta'\mathbf{r} - \mathbf{v}\,\delta' t\rangle \\ & - \langle m\,\delta'\mathbf{v} - q\,\mathbf{E}\,\delta' t,\ \delta\mathbf{r} - \mathbf{v}\,\delta t\rangle + q\,\langle \mathbf{B}, \delta\mathbf{r}\times\delta'\mathbf{r}\rangle\,,\end{aligned} \tag{15.1}$$

where q is a number called the *electric charge*.

Let X be a point of space-time with coordinates $\mathbf{r}$ and t:

$$X \equiv \mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix} \qquad \text{(notation (12.56))}, \tag{15.2}$$

and define[267]

$$\mathcal{F}(\delta X)(\delta' X) = \langle \mathbf{E},\ \delta\mathbf{r}\ \delta' t - \delta'\mathbf{r}\ \delta t\rangle + \langle \mathbf{B},\ \delta\mathbf{r}\times\delta'\mathbf{r}\rangle\,. \tag{15.3}$$

Let x_0 be the motion of a material point with initial condition $y = \begin{pmatrix} t\\ \mathbf{r}\\ \mathbf{v}\end{pmatrix}$ if it were not subjected to the electromagnetic field. Then x_0 varies over the symplectic manifold of motions U_0 of the free particle. We know that

$$\sigma(\delta x_0)(\delta' x_0) = \langle m\,\delta\mathbf{v},\ \delta'\mathbf{r} - \mathbf{v}\,\delta' t\rangle - \langle m\,\delta'\mathbf{v},\ \delta\mathbf{r} - \mathbf{v}\,\delta t\rangle\,, \tag{15.4}$$

and thus we get

$$\sigma(\delta y)(\delta' y) \equiv \sigma(\delta x_0)(\delta' x_0) + q\,\mathcal{F}(\delta X)(\delta' X)\,, \tag{15.5}$$

which can be interpreted as follows (Fig. 15.I): the Lagrange form of the evolution space V is the sum of

1) the inverse image of the Lagrange form of U_0 under the map $[y \mapsto x_0]$, and

[267] It is clear that $\mathcal{F}$ is a 2-form on space-time. Conversely, every 2-form $\mathcal{F}$ on space-time is defined by a formula of the form (15.3) where $\mathbf{E}$ and $\mathbf{B}$ take their values in $\mathbf{R}^3$.

2) the inverse image of the 2-form $q\mathcal{F}$ under the map $y \mapsto X$.

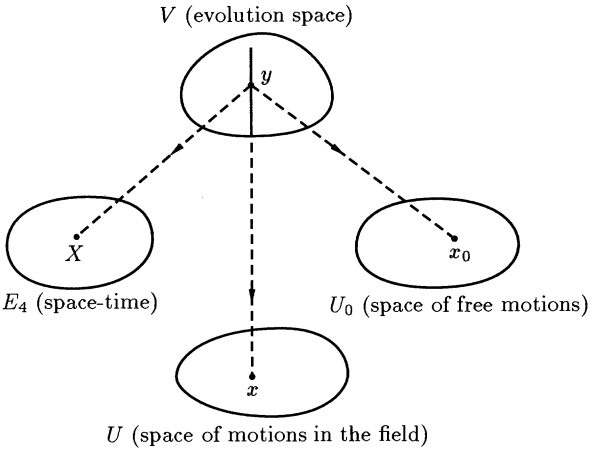


Fig 15.I

We know that the electric and magnetic fields satisfy Maxwell's equations (see (12.86))

$$\operatorname{curl}\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \qquad \text{and} \qquad \operatorname{div}\mathbf{B} = 0\,, \tag{15.6}$$

which can be written as

$$\mathrm{d}\mathcal{F} \equiv 0\,. \tag{15.7}$$

This shows directly why the 2-form σ defined by (15.5) has an exterior derivative equal to zero. We also know that the kernel of σ is 1-dimensional and that the space of motions U is a symplectic manifold (of dimension 6) such that

$$\sigma(\delta x)(\delta' x) \equiv \sigma(\delta y)(\delta' y)\,, \tag{15.8}$$

where x is the *motion of the particle under the influence of a given electromagnetic field* with initial condition y. Finally let us recall that the *equations of motion* $\delta y \in \ker(\sigma)$ can be written as

$$\frac{d\mathbf{r}}{dt} \equiv \mathbf{v} \qquad \text{and} \qquad \frac{d}{dt}[m\mathbf{v}] \equiv q\,[\mathbf{E} - \mathbf{B} \times \mathbf{v}]\,. \tag{15.9}$$

□

Now if we want to take relativity into account, we use the extremely simple idea of preserving *the same scheme*, replacing only the classical space of motions by the relativistic space of motions defined in (14.24). This implies that, instead of (15.5), we define

$$\sigma(\delta y)(\delta' y) = m\,[\overline{\delta X} \cdot \delta' I - \overline{\delta' X} \cdot \delta I] + q\mathcal{F}(\delta X)(\delta' X)\,, \tag{15.10}$$

where I is the unit future pointing vector tangent to the trajectory of the free motion. Taking into account the constraint (14.24♡), that is,

$$\overline{I} \cdot I = 1 \,, \tag{15.11}$$

we can immediately write the equations of motion $\delta y \in \ker(\sigma)$ in the form [268]

$$\delta X \text{ parallel to } I \qquad \text{and} \qquad m\,\delta I + q\,F \cdot \delta X \equiv 0 \,, \tag{15.12}$$

where F is the linear operator defined by

$$F(\delta X) \equiv -\overline{\mathcal{F}(\delta X)} \,, \tag{15.13}$$

or by

$$\mathcal{F}(\delta X)(\delta' X) \equiv \overline{\delta X} \cdot F \cdot \delta' X \,. \tag{15.14}$$

Note that F is anti-Hermitian:

$$\overline{F} \equiv -F \,. \tag{15.15}$$

Depending on the context, either the 2-form $\mathcal{F}$ or the operator F is called the *electromagnetic field*.

We deduce from (15.12) that I is also tangent to the trajectory of the particle in the electromagnetic field — which is not evident a priori. It is often convenient to parameterize the trajectory by "proper time" θ so that

$$\frac{dX}{d\theta} \equiv 1 \,. \tag{15.16}$$

With this the second equation in (15.12) becomes

$$\frac{dI}{d\theta} = -\frac{q}{m} F \cdot I \,. \tag{15.17}$$

The vector $\frac{dI}{d\theta}$ can be interpreted as the (four dimensional) *curvature* of the trajectory. A priori we know that it is a vector orthogonal to the trajectory and thus spacelike. The condition of orthogonality is automatically satisfied by the right hand side of equation (15.17), because F is anti-Hermitian.[269]

[268] We have to write $\delta'[\overline{I} \cdot I] = 0 \Rightarrow \sigma(\delta y)(\delta' y) = 0$. Using a Lagrange multiplier λ, it suffices to write $\sigma(\delta y)(\delta' y) + \lambda\,\delta'[\overline{I} \cdot I] = 0 \;\forall\, \delta' I$.

[269] The number $\gamma = \sqrt{-\frac{\overline{dI}}{d\theta} \cdot \frac{dI}{d\theta}}$ can be interpreted as the *proper acceleration* of the trajectory, that is, as its acceleration in a Lorentz frame in which it has zero velocity at the instant under consideration.

Using notation with 3-dimensional vectors, we have

(15.18) $$I = \mathcal{R}\begin{pmatrix} \mathbf{v} \\ 1 \end{pmatrix} \frac{1}{\sqrt{1 - \|\mathbf{v}\|^2}}$$

and

(15.19) $$F = \mathcal{R} \cdot \begin{pmatrix} j(\mathbf{B}) & -\mathbf{E} \\ \overline{-\mathbf{E}} & 0 \end{pmatrix} \cdot \mathcal{R}^{-1} .$$

In this notation the equations of motion [270] look like

(15.20) $$\frac{d\mathbf{r}}{dt} = \mathbf{v} \qquad \text{and} \qquad \frac{d}{dt}\left[\frac{m\,\mathbf{v}}{\sqrt{1 - \|\mathbf{v}\|^2}}\right] = q\,[\mathbf{E} - \mathbf{B} \times \mathbf{v}] .$$

Mnemonically, one might note that (15.20) is obtained from (15.9) by replacing the *classical momentum* $m\mathbf{v}$ by the *relativistic momentum* $m\mathbf{v}/\sqrt{1 - \|\mathbf{v}\|^2}$.

(15.21) As in the case of classical mechanics (see (12.37)) the fact that δt does not vanish on the integral curves of motion shows that the characteristic foliation of the evolution space is sectionable and thus that the space of motions U is a manifold. Thus U is symplectic and formula (15.8) remains valid.

(15.22)  It turns out that the equations (15.20) are correct. They are the ones used in particle accelerator calculations, where relativistic effects are important. This justifies a posteriori the expression (15.10) of the Lagrange form. One says that this formula is *phenomenological*, since it accounts for the phenomena that are studied, without being entirely a consequence of theoretical considerations.

A particle with spin in an electromagnetic field

A priori one could extend the above formalism to the case of a particle with spin by simply replacing the evolution space of a material point by the one defined in (14.4) (or (14.48) in the nonrelativistic case).

But experiments show that it is advisable to add a new term to the Lagrange form (15.5), namely

(15.23) $$\mu\, \mathrm{d}\varpi ,$$

[270] Expanding (15.12) one also obtains the equation $\frac{d}{dt}\left[\frac{m}{\sqrt{1-\mathbf{v}^2}}\right] = q\,\langle \mathbf{E}, \mathbf{v} \rangle$. But this equation is a consequence of (15.20).

where μ is a constant which we will interpret later on, and where ϖ is the 1-form

$$\varpi(\delta y) \equiv -\tfrac{1}{2}\,\mathrm{Tr}\,(\Omega \cdot F)\,\overline{I}\,\delta X\,. \tag{15.24}$$

Poincaré's theorem assures that the condition $d\sigma \equiv 0$ still holds.

In the nonrelativistic case, we replace (15.24) by its value at $\mathbf{v} = 0$, that is,

$$\varpi(\delta y) = \langle \mathbf{u},\, \mathbf{B}\rangle\, \delta t\,, \tag{15.25}$$

where $\mathbf{u}$ is the unit vector in the direction of the intrinsic angular momentum (notation (14.47)). Thus we have (see (14.48))

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;=\; & \langle m\,\delta\mathbf{v} - q\,\mathbf{E}\,\delta t,\, \delta'\mathbf{r} - \mathbf{v}\,\delta' t\rangle \\ & - \langle m\,\delta'\mathbf{v} - q\,\mathbf{E}\,\delta' t, \delta\mathbf{r} - \mathbf{v}\,\delta t\rangle + q\,\langle \mathbf{B},\, \delta\mathbf{r}\times\delta'\mathbf{r}\rangle \\ & - s\,\langle \mathbf{u},\, \delta\mathbf{u}\times\delta'\mathbf{u}\rangle + \mu[\delta\langle \mathbf{B},\,\mathbf{u}\rangle\;\delta' t - \delta'\langle \mathbf{B},\,\mathbf{u}\rangle\;\delta t]\,,\end{aligned} \tag{15.26}$$

[271]

from which we obtain the equations of motion

$$\left[\begin{array}{ll} \clubsuit & \dfrac{d\mathbf{r}}{dt} = \mathbf{v} \\[2ex] \heartsuit & \dfrac{d}{dt}[m\,\mathbf{v}] = q\,[\mathbf{E} - \mathbf{B}\times\mathbf{v}] + \dfrac{\overline{\partial \mathbf{B}}}{\partial \mathbf{r}}\cdot\mu\,\mathbf{u} \\[2ex] \diamondsuit & \dfrac{d}{dt}[s\,\mathbf{u}] = \mu\,\mathbf{u}\times\mathbf{B}\,. \end{array}\right. \tag{15.27}$$

The above equations express the fact that the particle is *magnetized* and has a *magnetic moment* $\mu\,\mathbf{u}$ parallel to its (intrinsic) angular momentum $s\,\mathbf{u}$. These equations have been verified experimentally many times. The term $\frac{\overline{\partial \mathbf{B}}}{\partial \mathbf{r}}\cdot\mu\,\mathbf{u}$ of equation $\heartsuit$ shows that the deviation of a particle moving through a spatially rapidly changing magnetic field *depends on the direction of the spin* (*the Stern-Gerlach experiment*). Equation $\diamondsuit$ shows that the spin of a particle subject to a constant magnetic field $\mathbf{B}$ rotates about this field with angular velocity $\mu\,\|\mathbf{B}\|/s$ (the phenomenon of *precession of spin*). The same equation also shows that a particle subjected to a field $\mathbf{B}_1 + \mathbf{B}_2\cos(\omega t)$, where $\mathbf{B}_1$ and $\mathbf{B}_2$ are orthogonal, undergoes a resonance effect if $\omega = \mu\|\mathbf{B}_1\|/s$. This phenomenon, called *magnetic resonance*, allows magnetic moments to be measured. □ (15.28)

In the relativistic case we have (see (14.4♣), (15.14), and (15.24)):

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;\equiv\; & -s\,\mathrm{Tr}(\delta\Omega\,\Omega\,\delta'\Omega) + m\,[\overline{\delta X}\cdot\delta' I - \overline{\delta' X}\cdot\delta I] \\ & + q\,\overline{\delta X}\cdot F\cdot\delta' X + \mu\,[\delta(\alpha\,\overline{I})\,\delta' X - \delta'(\alpha\,\overline{I})\,\delta X]\,,\end{aligned} \tag{15.29}$$

[271] Here the new terms come from the *magnetic Hamiltonian* $h = -\langle \mu\,\mathbf{u},\, \mathbf{B}\rangle$ (see (12.93)).

with

$$\alpha \equiv -\tfrac{1}{2}\,\mathrm{Tr}(\Omega \cdot F)\,. \tag{15.30}$$

To form the equations of motion, we have to take into account the constraints (14.4♡)

$$\Omega \equiv j(I)(J)\,, \qquad \overline{I} \cdot I \equiv 1\,, \qquad \text{and} \qquad \overline{J} \cdot J \equiv -1\,, \tag{15.31}$$

which can be done using Lagrange multipliers. The formulæ so obtained are rather complicated; linearized near $F = 0$ [272] they take the form

$$\begin{aligned} \frac{dX}{d\theta} &= I + \left[\frac{\mu}{m} - \frac{s\,q}{m^2}\right] \Omega \cdot F \cdot I - \frac{\mu\, s}{m^2}\,\Omega \cdot \overline{\frac{\partial \alpha}{\partial X}} \\ \frac{dI}{d\theta} &= -\frac{q}{m}\,F \cdot I - \frac{\mu}{m}\,[1 - I \cdot \overline{I}] \cdot \overline{\frac{\partial \alpha}{\partial X}} \\ \frac{dJ}{d\theta} &= -\frac{\mu}{s}\,F \cdot J + \left[\frac{q}{m} - \frac{\mu}{s}\right] I\,[\overline{J} \cdot F \cdot I] + \frac{\mu}{m}\,I\,\overline{J} \cdot \overline{\frac{\partial \alpha}{\partial X}}\,. \end{aligned} \tag{15.32}$$

Note that the above equations preserve the constraints (15.31) identically, but that I is no longer tangent to the trajectories (as is the case for a particle without spin). Consequently the parameter θ (defined by $\delta\theta = \overline{I}\,\delta X$) is not strictly equal to the proper time; however, it differs from it only by terms of second order in F. (15.33)

In the case of a charged particle one usually introduces the quantity

$$g = 2\,\frac{\mu\, m}{q\, s} \tag{15.34}$$

called the *gyromagnetic ratio* or *Landé g-factor*.

Using this notation and assuming that the electromagnetic field is *constant*, which implies that $\frac{\partial \alpha}{\partial X} \equiv 0$, [273] the equations (15.32) become

$$\begin{aligned} \frac{dX}{d\theta} &= I + \frac{q\, s}{2\, m^2}[g - 2]\,\Omega\, F\, I \\ \frac{dI}{d\theta} &= -\frac{q}{m}\,F\, I \\ \diamondsuit \qquad \frac{dJ}{d\theta} &= -\frac{q}{2\, m}\,\left[g\, F\, J + [g - 2]\, I\,[\overline{J}\, F\, I]\right]. \end{aligned} \tag{15.35}$$

[272] This approximation amounts to neglecting the *squares* of the dimensionless quantities $\mu\,\alpha/m$ and $q\, s\,\alpha/m^2$. Under usual experimental conditions these quantities are indeed very small. For example, for a magnetic field of 10^4 Gauss, they are of the order of magnitude of 10^{-10} for an electron or 10^{-13} for a nucleon.

[273] The notation $\frac{\partial \alpha}{\partial X}$ denotes a partial derivative with respect to X (the variables I, J, and hence Ω being held constant): $\frac{\partial \alpha}{\partial X}(\delta X) = -\frac{1}{2}\,\mathrm{Tr}\,(\Omega \cdot \delta F)$.

(15.36) The magnetic moment is called *normal* if $\mu = \frac{q\,s}{m}$, that is, if $g = 2$. This is the case to a very high degree for the electron and the muon. Equation $\diamondsuit$, established by Bargmann, Michel, and Telegdi,[274] is used to measure $g - 2$, that is, the *anomalous magnetic moment*. The consistency of experiments measuring the anomalous magnetic moment gives an a posteriori justification for the equations (15.32) and thus for the phenomenological expression (15.26) of σ.

Systems of particles without interactions

The space of motions U of a system of particles without interactions is the *direct product* of the spaces $U_1, \ldots, U_n$ of motions of the individual particles (see (13.1) and (13.12)). Thus a motion x of the system can be written as

$$x \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad x_1 \in U_1, \ldots, x_n \in U_n\,, \tag{15.37}$$

and the Lagrange form of U is given by

$$\sigma(\delta x)(\delta' x) = \sigma(\delta x_1)(\delta' x_1) + \cdots + \sigma(\delta x_n)(\delta' x_n)\,. \tag{15.38}$$

These formulæ apply as well to (a system of) free particles, as to particles subjected to a field.

By the *synchronous evolution space* of the system of particles we will mean the set V of

$$y \equiv \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \tag{15.39}$$

where $y_1, \ldots, y_n$ are the initial conditions of the various particles all having *the same time* t.[275] V admits the structure of a manifold such that $\dim(V) = \dim(U) + 1$; it is presymplectic if we define

$$\sigma(\delta y)(\delta' y) = \sigma(\delta y_1)(\delta' y_1) + \cdots + \sigma(\delta y_n)(\delta' y_n) \tag{15.40}$$

and then possesses U as a quotient manifold. In the case of nonrelativistic particles without spin it is easy to identify V with the classical evolution space (12.29).

[274] See Bargmann, Michel, and Telegdi [5] and also Thomas [36].

[275] Recall that an initial condition of a particle consists of an event, that is, a position $\mathbf{r}$ and a time t.

(15.41) Likewise we define an *asynchronous evolution space* V' by giving up the requirement that the initial conditions y_j are synchronized. Hence V' is the *direct product* of the evolution spaces V_j of the various particles and is still a presymplectic manifold with quotient U. However, the dimension of V' is the dimension of U augmented by n (one chooses an independent time t_j for each of the particles).

(15.42) We now consider the case of *identical* particles (with the same mass, the same spin, and subjected to the same field); their spaces of motions $U_1, \ldots, U_n$ are thus the same. Hence we should have $U = [U_1]^n$.

But experiments have shown[276] *that one cannot distinguish two motions that differ only by exchanging two particles.* Thus the actual space of motions of the system is not $[U_1]^n$ but the *quotient manifold* constructed in (10.27), which we denote by $U = [U_1]^n/S_n$. Here S_n denotes the group of permutations of the integers $1, 2, \ldots, n$, called the *symmetric group of degree n*.

Let us recall the construction of (10.27). The formula

$$(15.43)\qquad \hat{\gamma}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{\gamma^{-1}(1)} \\ \vdots \\ x_{\gamma^{-1}(n)} \end{pmatrix} \qquad \forall\, \gamma \in S_n \text{ and } \forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in [U_1]^n\,,$$

defines an action of the symmetric group S_n on $[U_1]^n$ in the sense that $\widehat{\gamma \circ \gamma'} = \hat{\gamma} \circ \hat{\gamma}'$. This action defines an equivalence relation on $[U_1]^n$ for which the equivalence class $\hat{x}$ of x is defined by

$$(15.44)\qquad [\hat{x} = \hat{x'}] \iff [\,\exists\, \gamma \in S_n \;:\; x' = \hat{\gamma}(x)\,]\,.$$

We then obtain:

(15.45) The manifold U of motions of a system of n identical particles is the set of all equivalence classes $\hat{x}$, where x is an n-tuple whose components $x_1, x_2, \ldots, x_n$ are *all different*. The Lagrange form σ of U is given by

$$\sigma(\delta\hat{x}, \delta'\hat{x}) \equiv \sigma(\delta x)(\delta' x) \equiv \sigma(\delta x_1)(\delta' x_1) + \cdots + \sigma(\delta x_n)(\delta' x_n)\,. \qquad \square$$

(15.46) Let $\widetilde{[U_1]^n}$ be the set of x such that $x_1, \ldots, x_n$ are all different, which is an open subset of U_1^n. Furthermore let S_n be the group of $\hat{\gamma}$ acting on U_1^n and let P be the map $x \mapsto \hat{x}$. Then the triple

$$\{\widetilde{[U_1]^n},\, S_n,\, P\}$$

[276] See (17.169).

is a *covering space* of U in the sense of (1.42) (see (10.26) and (10.27)). The dimension of U is thus *equal* to that of U_1^n, that is, n times the dimension of U_1.

(15.47) The fact that the space of motions is $\widetilde{[U_1]}^n/S_n$ and not $\widetilde{[U_1]^n}$ is a mathematical way of expressing the fact that the particles are *indiscernible*. Contrary to common opinion, this formulation *does not involve quantum mechanics*.



Interactions

One of the major problems of physics consists of finding a correct description of mutual interactions of a system of particles. A "classical" method is to consider the *synchronous evolution space* V of the system (see (15.39) above) and to modify the Lagrange form by the substitution

$$\sigma(\delta y)(\delta' y) \to \sigma(\delta y)(\delta' y) + \delta t\ \delta' v - \delta' t\ \delta v\,, \tag{15.48}$$

where v is a differentiable function of y called the *interaction potential*. Since the added term is an exterior derivative, the condition $\mathrm{d}\sigma \equiv 0$ remains valid. With a little bit of luck, the rank of σ will not change (it will remain equal to the maximum compatible with the fact that V is odd dimensional). V is then presymplectic and the quotient U will be the space of motions of the interacting particles.

This classical method works successfully in the case of celestial mechanics, where it leads, among others, to Lagrange's formulæ for the *variation of constants* (see (12.106)). □

Let us apply the above reasoning to a system of nonrelativistic particles with spin. We will use the notation of (14.48) and each particle will be indexed by a subscript j. We obtain

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;=\; & \sum_j m_j \langle \delta \mathbf{v}_j,\, \delta' \mathbf{r}_j - \mathbf{v}_j\, \delta' t\rangle - m_j \langle \delta' \mathbf{v}_j,\, \delta \mathbf{r}_j - \mathbf{v}_j\, \delta t\rangle \\ & - s_j \langle \mathbf{u}_j, \delta \mathbf{u}_j \times \delta' \mathbf{u}_j\rangle + \delta t\ \delta' v - \delta' t\ \delta v\,.\end{aligned} \tag{15.49}$$

We now assume that v *does not depend* on the velocities $\mathbf{v}_j$. With this assumption the equations of motion $\delta y \in \ker(\sigma)$ become

$$\begin{aligned}\frac{d\mathbf{r}_j}{dt} &= \mathbf{v}_j \\ m_j \frac{d\mathbf{v}_j}{dt} &= -\overline{\frac{\partial v}{\partial \mathbf{r}_j}} \\ s_j \frac{d\mathbf{u}_j}{dt} &= \mathbf{u}_j \times \overline{\frac{\partial v}{\partial \mathbf{u}_j}}\,.\end{aligned} \tag{15.50}$$

If the system is isolated, the *principle of Galilean relativity* gives some supplementary conditions. Using equations (12.124) and (12.134), it follows that v must satisfy[277]

$$\frac{\partial v}{\partial t} \equiv 0\,, \qquad \sum_j \frac{\partial v}{\partial \mathbf{r}_j} \equiv 0\,, \qquad \sum_j \mathbf{r}_j \times \frac{\overline{\partial v}}{\partial \mathbf{r}_j} + \mathbf{u}_j \times \frac{\overline{\partial v}}{\partial \mathbf{u}_j} \equiv 0\,, \tag{15.51}$$

and that the Galilean moments (notation (12.122)) are given by

$$\begin{aligned} \mathbf{l} &\equiv \sum_j \mathbf{r}_j \times m_j\,\mathbf{v}_j + s_j\,\mathbf{u}_j\,, & \mathbf{g} &\equiv \sum_j m_j\,[\mathbf{r}_j - t\,\mathbf{v}_j]\,, \\ \mathbf{p} &\equiv \sum_j m_j\,\mathbf{v}_j\,, & E &\equiv v + \tfrac{1}{2}\sum_j m_j\,\|\mathbf{v}_j\|^2\,. \end{aligned} \tag{15.52}$$

According to the generalized version of Noether's theorem (11.12) these are constants of motion.

The above formulæ (15.49–52) are capable of describing *electrostatic* and *magnetostatic* forces between particles by choosing for v the classical expression for the energy of a system of electrically charged magnetic dipoles, namely

$$\begin{aligned} v \;=\; & \sum_{\substack{j,k \\ [j\neq k]}} \frac{q_j\,q_k}{\|\mathbf{r}_j - \mathbf{r}_k\|} + \\ & + \sum_{\substack{j,k \\ [j\neq k]}} \mu_j\,\mu_k \left[\frac{\langle \mathbf{u}_j,\, \mathbf{u}_k \rangle}{\|\mathbf{r}_j - \mathbf{r}_k\|^3} - 3\,\frac{\langle \mathbf{u}_j, \mathbf{r}_j - \mathbf{r}_k \rangle \langle \mathbf{u}_k, \mathbf{r}_j - \mathbf{r}_k \rangle}{\|\mathbf{r}_j - \mathbf{r}_k\|^5} \right], \end{aligned} \tag{15.53}$$

where q_j is the electric charge and μ_j is the magnetic moment of the j^{th} particle.

(15.54) Notice that the magnetic interaction terms (those that contain the factor $\mu_j\,\mu_k$) are negligible in a first approximation,[278] an approximation in which the $\mathbf{u}_j$ are constant. Thus the space of motions is the direct product of *the space of motions of the interacting particles considered without spin* with as many *symplectic spheres* as there are particles with spin in the system. □

(15.55) This method applies without difficulty to the case of a system of identical particles, provided that we work on the covering space (see (15.46)). In

[277] After a possible addition of a function of time only, which does not modify the value (15.49) of σ.

[278] In the case of the hydrogen atom, these terms are responsible for the ***hyperfine*** structure of the spectrum as well as for the 21 cm wavelength emission line used in radio astronomy.

such a case v should be a *symmetric* function of the particles (as is clearly the case for expression (15.53)).

The above formulæ do not account for small, but measurable, effects such as the Laplace force, to which a charged particle moving in the magnetic field created by other particles is subjected. We have already stated that these effects are relativistic (see (12.87)).

(15.56) We are thus confronted with the problem of finding a *relativistic model for a system of particles in interaction.* It is certain that such a model cannot be constructed using the preceding method. The reason is that the *synchronous evolution space* depends on the chosen Lorentz frame, because the notion of simultaneity depends on it.

It is thus natural to try to use the *asynchronous* evolution space. Let us discuss the simplest case: that of a system of *nonrelativistic* particles without spin subjected to forces $\mathbf{F}_j$.

The Lagrange form of the asynchronous evolution space is given by

$$\begin{aligned}\sigma(\delta y)(\delta' y) \;=\; \sum_j \Big(&\langle m_j\,\delta \mathbf{v}_j - \mathbf{F}_j\,\delta t_j, \delta' \mathbf{r}_j - \mathbf{v}_j\,\delta' t_j\rangle \\ &- \langle m_j\,\delta' \mathbf{v}_j - \mathbf{F}_j\,\delta' t_j,\ \delta \mathbf{r}_j - \mathbf{v}_j\,\delta t_j\rangle\Big)\,.\end{aligned} \tag{15.57}$$

A priori the forces $\mathbf{F}_j$ can be taken to be functions of the t_k, $\mathbf{r}_k$ and $\mathbf{v}_k$. But expanding Maxwell's condition $\mathrm{d}\sigma \equiv 0$ gives the relations

$$\begin{aligned} &\frac{\partial \mathbf{F}_j}{\partial \mathbf{v}_k} \equiv 0 && && \forall\, j,k \\ &\frac{\partial \mathbf{F}_j}{\partial t_k} \equiv 0 && \frac{\partial \mathbf{F}_j}{\partial \mathbf{r}_k} \equiv 0 && \text{if } j \neq k \\ &\frac{\partial \mathbf{F}_j}{\partial \mathbf{r}_j} - \overline{\frac{\partial \mathbf{F}_j}{\partial \mathbf{r}_j}} \equiv 0 && && \forall\, j\,. \end{aligned} \tag{15.58}$$

(15.59)  Comparing the above equations with (12.80), we see that *replacing the synchronous evolution space by the asynchronous one prohibits the existence of interactions*, since Maxwell's reciprocity law $\overline{\frac{\partial \mathbf{F}_k}{\partial \mathbf{r}_j}} - \frac{\partial \mathbf{F}_j}{\partial \mathbf{r}_k} \equiv 0$ is replaced by the condition $\frac{\partial \mathbf{F}_k}{\partial \mathbf{r}_j} \equiv 0\ \forall\, j \neq k$. On the other hand, relativistic invariance prohibits the use of the synchronous evolution space, as we have already seen.

(15.60) The above considerations explain why *one does not know a relativistic model of interacting particles*, not even in such a simple cases as the hydrogen atom.[279]

[279] One can find an approximate model by assuming that the mass of the electron is

Actually, these arguments *do not prohibit* the existence of such a model. They simply suggest that one should give up the overly Newtonian idea — and above all the overly Aristotelian idea — that the real behavior of a system ("energeia") can be described by means of permanent particles ("hulè") whose motions are simply affected by an interaction ("morphè"). Thus the dualism between "form" and "matter" can only be considered as a *nonrelativistic approximation*, whence the speculative interest in methods that allow us to avoid it — such as scattering theory.

Scattering theory

(15.61) Let U be the space of motions of a dynamical system (Fig. 15.II), let V be the evolution space of the system, and let P be the projection of V onto U. We now suppose (just as we have done to take into account the action of a field on a particle, or the interaction of a system of particles) that a new dynamical system is defined by *modifying* the Lagrange form σ of V (for example, by adding a supplementary term), but that σ_V *is not modified in a certain open subset* Ω of V. In such a case one says that the modified system is a *scattered system* with respect to the initial system. We will call the complement of Ω the *scattering source*.

One might assume, if one wants to, that Ω is the *largest open subset* of V where the Lagrange forms of the initial system and of the scattered system coincide;[280] its complement is then called the *support* of the scattering source.

We will denote by $\tilde{U}$ the space of *scattered motions* and by $\tilde{P}$ the projection from V onto $\tilde{U}$.

We now make the fundamental assumption that the evolution space V is sectionable for the scattered system as well as for the initial system, and thus that U and $\tilde{U}$ are symplectic manifolds.

Let $y \in \Omega$. The initial condition y defines an unscattered motion $x = P(y)$ and a scattered motion $\tilde{x} = \tilde{P}(y)$. The motions x and $\tilde{x}$ can be

---

negligible in comparison with that of the proton, which thus is assumed to be fixed. But *there is no interaction* in this case: only the action of the proton on the electron.

Another method consists of studying the interaction of a proton and an electron with the electromagnetic field separately. But the model so constructed is not realistic because it leads to loss of energy by radiation, and this grants the atom only a very short lifetime. One often says that the observed stability of the atom is a quantum effect. But the relativistic quantum treatment of the problem is also incomplete.

[280] In other words, the union of open sets having this property; it is indeed an open set (theorem 1.14)).

considered as *leaves* of V for its two foliations. It is easy to verify that the intersection of one of these two leaves with Ω is an *integral manifold* of Ω (seen as a submanifold of V with a foliation), and thus a *union of leaves* of Ω. In particular, the leaf of Ω containing y (indicated by the heavy curve in Fig. 15.II) is contained in x and $\tilde{x}$ at the same time.

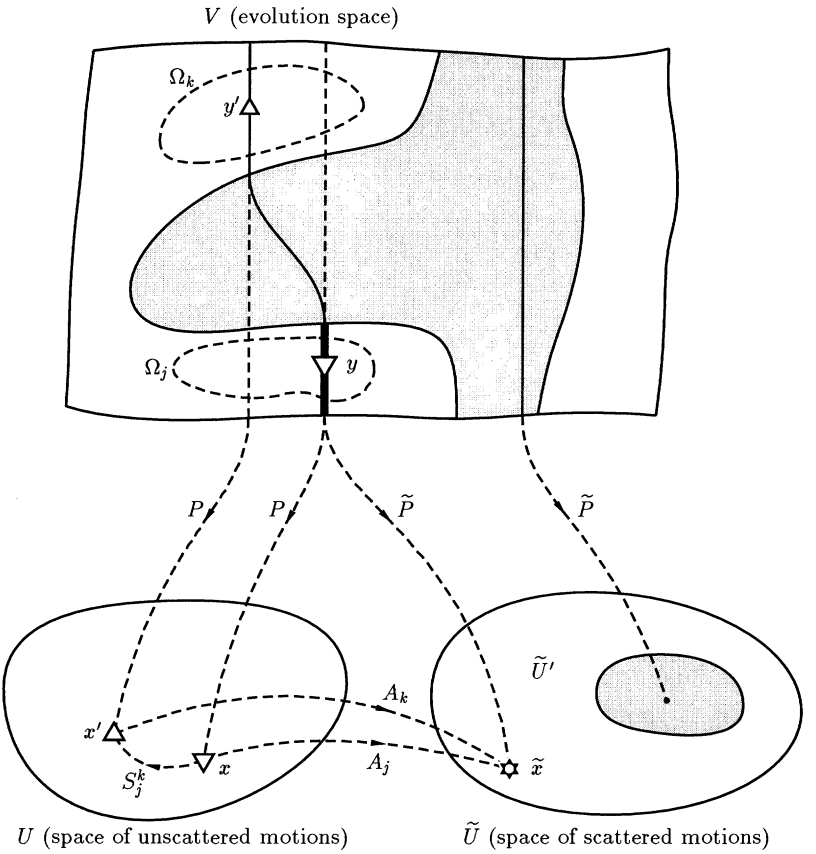


Fig 15.II

Using these remarks one can prove the following theorem.

(15.62) THEOREM: There exists a family of open sets Ω_j, whose union equals Ω, such that

$$[y, y' \in \Omega_j] \Rightarrow [\{P(y) = P(y')\} \implies \{\tilde{P}(y) = \tilde{P}(y')\}]\,. \qquad \square$$

An immediate corollary is that there exists an injective map A_j from U to $\tilde{U}$ defined by

(15.63) $$[A_j(x) = \tilde{x}] \implies [\exists\, y \in \Omega_j :\ P(y) = x \text{ and } \tilde{P}(y) = \tilde{x}]\,.$$

It can be proven (by choosing a transversal section of the foliation passing through y) that A_j is a local diffeomorphism. The formula

$$\sigma(\delta x)(\delta' x) = \sigma(\delta y)(\delta' y) = \sigma(\delta\tilde{x})(\delta'\tilde{x}) \tag{15.64}$$

then shows that A_j is a local *symplectomorphism*.

We will call a scattered motion which is entirely contained in the scattering source a *constrained motion*. As the sets Ω_j cover Ω, the union of the images $\mathrm{im}(A_j)$ is the set $\tilde{U}'$ of *unconstrained motions*. Thus $\tilde{U}'$ is an open subset of $\tilde{U}$ (which could coincide with $\tilde{U}$).

If we define

$$S_j^k = A_k^{-1} \circ A_j \qquad \forall j,k \in I\,,^{281} \tag{15.65}$$

or, what amounts to the same thing (Fig. 15.II),

$$[S_j^k(x) = x'] \Longleftrightarrow \left[\begin{array}{cc} & \exists y \in \Omega_j \ \exists y' \in \Omega_k \ : \\ \tilde{P}(y) = \tilde{P}(y')\,, & P(y) = x\,, \quad P(y') = x' \end{array}\right], \tag{15.66}$$

then we obtain immediately:

$$\begin{array}{ll} S_j^k \text{ is a local symplectomorphism from } U \text{ to } U & \forall j,k \in I \\ S_j^j = 1_{E_j}, \quad \text{where } E_j \text{ is an open subset of } U & \forall j \in I \\ [S_k^j]^{-1} = S_j^k & \forall j,k \in I \\ S_j^k \circ S_\ell^j < S_\ell^k & \forall j,k,\ell \in I\,.^{282} \end{array} \tag{15.67}$$



Conversely, if we know the maps S_j^k, we know the space $\tilde{U}'$ of unconstrained motions. This is a consequence of the following theorem.

(15.68) THEOREM: Let U be a symplectic manifold and I an indexing set. Suppose that there exists a set of maps S_j^k *satisfying* (15.67). Then:

1) There exists a symplectic manifold U' and local symplectomorphisms A_j from U to U' such that

$$\begin{array}{ll} A_k^{-1} \circ A_j = S_j^k & \forall j,k \in I \\ \bigcup_j \mathrm{im}(A_j) = U'\,. \end{array}$$

◇

[281] I denotes the indexing set of the family Ω_j. The map S_j^k could be empty.

[282] We use the notation $A < B$, where A and B are maps, to mean that B is an *extension* of A, that is, $[x \in \mathrm{dom}(A)] \Rightarrow [B(x) = A(x)]$.

2) U' is *uniquely determined up to a global symplectomorphism*: if a manifold U'^* and local symplectomorphisms A_j^* from U to U'^* also satisfy $\diamondsuit$, then there exists a *global symplectomorphism* B from U' to U'^* such that

$$\heartsuit \qquad A_j^* = B \circ A_j \qquad \forall\, j \in I\,.$$

PROOF: Consider the set of pairs $\begin{pmatrix} j \\ x \end{pmatrix}$ with $j \in I$ and $x \in E_j$.† On this set we define a relation $\sim$ by

$$\left[\begin{pmatrix} j \\ x \end{pmatrix} \sim \begin{pmatrix} j' \\ x' \end{pmatrix}\right] \iff \left[x' = S_j^{j'}(x)\right].$$

From (15.67) it follows trivially that $\sim$ is an equivalence relation.

If we denote by $A_j(x)$ the equivalence class of $\begin{pmatrix} j \\ x \end{pmatrix}$ and by U' the union of the $\mathrm{im}(A_j)$, then the A_j are injective maps from U to U' satisfying $\diamondsuit$. It follows that the set of all $A_j \circ F$, with $j \in I$ and F a chart of U, form an atlas of U' turning U' into a manifold such that the A_j are local diffeomorphisms. Since the S_j^k are local symplectomorphisms, the fields of 2-forms σ_j defined on the open sets $\mathrm{im}(A_j)$ by

$$\sigma_j(\delta[A_j(x)])(\delta'[A_j(x)]) \equiv \sigma(\delta x)(\delta' x)$$

are compatible and are thus the restrictions of a field of 2-forms σ on U'. This σ defines a symplectic structure on U' such that the A_j are local symplectomorphisms (this is an immediate verification), whence (1).

Let us now assume the hypotheses of (2). Then the $B_j = A_j^* \circ A_j^{-1}$ are local symplectomorphisms from U' to U'^*. For $x \in \mathrm{dom}(B_j) \cap \mathrm{dom}(B_k)$, and writing x in the form $A_\ell(y)$, we deduce from $\diamondsuit$ that $B_j(x) = B_k(x)$. Hence the B_j have a common extension B defined on U'. In the same way the B_j^{-1} have a common extension B^* defined on U'^*. One then notices that $B^* = B^{-1}$ and that B is a symplectomorphism from U' to U'^* satisfying $\heartsuit$, whence (2). □

(15.69) We will say that a Lie group G is a *symmetry of a scattering source* if

$$\begin{cases} G \text{ acts on } V \\ G \text{ is a } \textit{dynamical group for the scattered system} \\ \quad \textit{as well as for the unscattered system.} \end{cases}$$ □

From (11.6) we know that G is also a dynamical group of U and $\tilde{U}$. It is easy to show that one can enlarge Ω so that it is invariant under G. Thus G acts also on Ω and on the space of unconstrained motions $\tilde{U}'$.

†Editors' note: Recall that E_j is the domain of definition of S_j^j.

To conveniently express the properties of G we introduce the following definition.

(15.70) DEFINITION: A map F from one manifold to another is called a *step function* if

$$[y \in \mathrm{im}(F)] \implies [F^{-1}(y) \text{ is an open set}]\,.^{283,\dagger} \qquad \square$$

One can check that the following two characterizations of a step function are equivalent to the given one.

(15.71) $\mathrm{dom}(F)$ is open and F is constant on each connected component of $\mathrm{dom}(F)$. $\square$

(15.72) F is differentiable and $\Big[[x \in \mathrm{dom}(F)] \Rightarrow [D(F)(x) = 0]\Big]$. $\square$

(15.73) THEOREM: (Hypotheses and notation of (15.61) to (15.72).) Let G be a symmetry group of a scattering source. Then

1) The maps

$$\begin{aligned} a &\mapsto [\underline{a}_U \circ A_j \circ \underline{a^{-1}}_U](x) \\ a &\mapsto [\underline{a}_U \circ S_j^k \circ \underline{a^{-1}}_U](x) \end{aligned} \qquad a \in G$$

are step functions ($\forall\, j, k \in I,\ \forall\, x \in U$).

2) If $\mathcal{G}$ is the Lie algebra of G and if $Z \in \mathcal{G}$, then

$$\begin{aligned} D(A_j)(x)(Z_U(x)) &= Z_{\widetilde{U}}(A_j(x)) && \text{if } x \in \mathrm{dom}(A_j)\,, \\ D(S_j^k)(x)(Z_U(x)) &= Z_U(S_j^k(x)) && \text{if } x \in \mathrm{dom}(S_j^k)\,. \end{aligned}$$

3) Let us assume that G admits moments $\mu \equiv \psi(x)$ and $\widetilde{\mu} \equiv \widetilde{\psi}(\widetilde{x})$ for the unscattered and scattered systems, respectively. Then the maps

$$\begin{aligned} x &\mapsto \widetilde{\psi}(A_j(x)) - \psi(x) \\ x &\mapsto \psi(S_j^k(x)) - \psi(x) \end{aligned} \qquad x \in U$$

are step functions ($\forall\, j \in I$).

4) If Z belongs to the derived algebra of $\mathcal{G}$, then

$$\psi(S_j^k(x))Z = \psi(x)Z \qquad \forall\, x \in \mathrm{dom}(S_j^k)\,.$$

[283] The inverse image $F^{-1}(y)$ is the set of x such that $F(x) = y$.

[†] Editors' note: In measure theory the definition of a step function is slightly weaker; one only asks that $F^{-1}(y)$ is measurable.

PROOF: Left to the reader.[284] □

The following examples will show the physical significance of this theorem.

Bounded scattering sources

(15.74) Let us consider a particle subjected to a field which is zero or negligible outside the ball $\|\mathbf{r}\| < R$ (in a given inertial reference frame) and let us take as the space of motions of the unscattered system the set U of rectilinear uniform motions of the free particle satisfying the following two conditions:

(15.75)
1) The velocity of the particle is not zero.
2) The shortest distance of the particle to the origin is less than R.

Then U is an open set in the manifold of free motions (relativistic or not), and thus U is a symplectic manifold.

Let Ω be the set of initial conditions y of the particle (defined, as is known, by a time t, a position $\mathbf{r}$, a velocity $\mathbf{v}$ and possibly a spin direction $\mathbf{u}$) that are compatible with (15.75) and for which $\|\mathbf{r}\| > R$. Ω is an open set where the Lagrange form for the unscattered particle coincides with that of the scattered particle.

It follows immediately from the definition of U (15.75) that

(15.76) $$[y \in \Omega] \implies [\langle \mathbf{r}, \mathbf{v} \rangle \neq 0] .$$

We are thus able to partition Ω into two disjoint open sets Ω_{out} and Ω_{in} defined by

(15.77) $$\begin{aligned} y \in \Omega_{\text{out}} &\iff \langle \mathbf{r}, \mathbf{v} \rangle > 0 \\ y \in \Omega_{\text{in}} &\iff \langle \mathbf{r}, \mathbf{v} \rangle < 0 . \end{aligned}$$

(15.78) The conditions (15.62) are satisfied with I the set of two elements $\{\text{in}, \text{out}\}$. It is standard to denote by S the symplectomorphism $S^{\text{out}}_{\text{in}}$. We thus have

(15.79) $$\begin{aligned} S^{\text{out}}_{\text{out}} &= 1_{\Omega_{\text{out}}} & S^{\text{out}}_{\text{in}} &= S \\ S^{\text{in}}_{\text{out}} &= S^{-1} & S^{\text{in}}_{\text{in}} &= 1_{\Omega_{\text{in}}} . \end{aligned}$$

Using theorem (15.68) we see that if we know S, we know all about the space of scattered motions. Note that $\text{dom}(S)$ is not necessarily equal

[284] To prove (4) one could use (11.17♯).

to U nor need it be connected.[285] □

(15.80) Now let us suppose that the scattering source is *static*, that is, the field is invariant under time translations. Then the group G of time translations[286] is a *symmetry group of the scattering source* in the sense of (15.69). According to theorem (15.73), the map $a \mapsto \underline{a}_U \circ S \circ \underline{a}^{-1}{}_U(x)$ is a step function ($a \in \mathbf{R}$). For $x \in \mathrm{dom}(S)$ one can see that the above map is defined on all of $\mathbf{R}$. Since $\mathbf{R}$ is connected, this map must be constant (lemma (15.71)), showing that S *commutes with time translations.*

(15.81) Let $\psi(x)$ be the moment of the group G corresponding to the motion x. Then, up to a sign, $\psi(x)$ is *the energy* (see (12.122)). Theorem (15.73.3) shows that the map $x \mapsto \psi(S(x)) - \psi(x)$ is a step function: crossing the scattering source adds an amount to the energy which depends only on the connected component of $\mathrm{dom}(S)$ to which x belongs.[287]

(15.82) It happens fairly often that the scattering source is *conservative*, that is, $\psi(S(x)) \equiv \psi(x)$. Let us mention some cases where this necessarily holds:

a) If $\mathrm{dom}(S)$ is connected.
b) If one can extend G to a group of symmetries whose Lie algebra equals its derived algebra (see (15.73.4)).
c) Finally, in certain cases where the previous explanations are not applicable,[288] it can happen that the scattering source is conserva-

[285] In the case of a photon scattered by an astronomical telescope (contained in the sphere $\|\mathbf{r}\| < R$), $\mathrm{dom}(S)$ has three components, described as the free motions of the photon whose phase "in" is such that

a) the photon enters by the lens and exits through the eyepiece.
b) the photon enters by the eyepiece and exits through the lens.
c) the photon avoids the telescope.

Furthermore, each of these three components splits in two if one allows polarized photons to enter the telescope. Two photons having the same initial trajectories but differing in their helicity, could have different trajectories after scattering. This is confirmed by experiments: an ordinary light source observed through a quartz crystal produces two images, both of which are circularly polarized but in opposite senses (*rotary birefringence* or *double refraction*).

[286] G can be considered as the additive group $\mathbf{R}$ acting on V according to

$$\underline{a}_V \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \\ \mathbf{u} \end{pmatrix} \equiv \begin{pmatrix} t+a \\ \mathbf{r} \\ \mathbf{v} \\ \mathbf{u} \end{pmatrix}.$$

[287] Thus an electron crossing an electron microscope sees its energy increased by a constant which is characteristic of the difference in potential between the entrance and the exit.

[288] For example, when the scattering source is described by an additive potential

tive because it is possible to *extend* S *complex analytically* so that it becomes defined on a connected set. □

We now give some examples of scattering sources having other symmetries.

(15.83) EXAMPLE: An centered optical system has a rotational symmetry about its optical axis. □

(15.84) EXAMPLE: SO(3) is a symmetry group for the scattering of a particle by a central force field. Theorem (15.73.4) shows that S preserves the corresponding moment (which is the angular momentum $\mathbf{l}$ using the notation of (12.122)). □

(15.85) EXAMPLE: The photon, *in its nonrelativistic approximation*, has *infinite velocity* (14.54). The time t of its motion is the moment of the group of translations of energy E (notation (14.53)). The fact that the entry and exit times of a photon in an optical instrument are the same can be related to the fact that the scattering is independent of E. □

Geometrical optics

(15.86) This last example (15.85) gives us the opportunity to apply procedure (12.152) to the dynamical group ($E \to E+$constant). This shows us that at *each time* t there exists a symplectic manifold whose points are the orbits of the above dynamical group acting on the space of motions of the nonrelativistic photon. Such an orbit ξ can be described by a unit vector $\mathbf{u}$ (notation (14.53)), defining the direction and sense of propagation of the photon, and a straight line D parallel to $\mathbf{u}$ (on which a point $\mathbf{r}$ should be chosen). Thus ξ is what is called a *light ray*.

(15.87) The set of light rays forms a manifold V of dimension 4.[289] The Lagrange form of V can be deduced from (14.53♣); up to a constant

---

v (as in (15.48)). A priori it could happen that v is multivalued and yet defines a symplectic structure. In that case the scattering source will not be conservative.

[289]The manifold V consists of all oriented lines in space. One can show that V is *simply connected*. (15.88) tells us that the map ($\mathbf{r} \to \mathbf{r}$, $\mathbf{u} \to -\mathbf{u}$), which reverses the orientation of a given line, is *anticanonical*. This map generates a *discrete group* of two elements. The quotient manifold (see (1.45)) consists of all *unoriented lines* in space and thus has V as its *universal covering space* (3.55).

factor k it is given by [290]

$$\sigma(\delta\xi)(\delta'\xi) \equiv \langle \delta\mathbf{u},\, \delta'\mathbf{r}\rangle - \langle \delta'\mathbf{u},\, \delta\mathbf{r}\rangle - \frac{\chi\, s}{k}\,\langle \mathbf{u},\, \delta\mathbf{u}\times\delta'\mathbf{u}\rangle\,. \tag{15.88}$$

Now let us consider a light ray ξ which enters an optical instrument. Since S is a local symplectomorphism (notation (15.78)), the emerging light ray $\xi' \equiv S(\xi)$ will be related to ξ by

$$\begin{aligned}\langle \delta\mathbf{u},\, \delta'\mathbf{r}\rangle - \langle \delta'\mathbf{u},\, \delta\mathbf{r}\rangle - \frac{\chi\, s}{k}\,\langle \mathbf{u},\, \delta\mathbf{u}\times\delta'\mathbf{u}\rangle \\ \equiv \langle \delta\mathbf{u}',\, \delta'\mathbf{r}'\rangle - \langle \delta'\mathbf{u}',\, \delta\mathbf{r}'\rangle - \frac{\chi\, s}{k}\langle \mathbf{u}',\, \delta\mathbf{u}'\times\delta'\mathbf{u}'\rangle\,.\end{aligned} \tag{15.89}$$

This formula shows that *it is impossible to obtain an exact point photograph* ($\mathbf{r}'$ = constant) *of a star at infinity* ($\mathbf{u}$ = constant), as long as the quantity $\frac{2\pi s}{k}$ (which later we will interpret as the *wavelength* of the photon (19.43)) cannot be neglected. If we do neglect it, thus replacing (15.88) by

$$\sigma(\delta\xi)(\delta'\xi) \equiv \langle \delta\mathbf{u},\, \delta'\mathbf{r}\rangle - \langle \delta'\mathbf{u},\, \delta\mathbf{r}\rangle\,, \tag{15.90}$$

we recover the *standard approximation of geometrical optics*. The conservation of (15.90) as the photon passes through an instrument is a property known as *Lagrange's theorem* (which is usually deduced from *Fermat's principle*).

On the manifold of rays V there exists a potential ϖ for σ defined by [291]

$$\varpi(\delta\xi) \equiv -\langle \mathbf{r},\, \delta\mathbf{u}\rangle\,. \tag{15.91}$$

Hence, there exists, at least locally, a *dynamical variable* W such that

$$\langle \mathbf{r}',\, \delta\mathbf{u}'\rangle - \langle \mathbf{r},\, \delta\mathbf{u}\rangle \equiv \delta W \tag{15.92}$$

for the passage of the photon through the instrument. W is called the *angular eikonal* [292] or *angular characteristic*. It is used in modern methods of computation concerning optical instruments.

---

[290] Recall that $\chi(=\pm 1)$ is the *helicity*, s the *spin*, and k the *color* of the photon (14.53).

[291] The potential ϖ' on the manifold V' of pairs $\eta \equiv (\mathbf{u}, \mathbf{r})$ defined by $\varpi'(\delta\eta) \equiv \langle \mathbf{u},\, \delta\mathbf{r}\rangle$ is more classical (it coincides with Cartan's form of Fermat's variational principle). However, ϖ' is not an integral invariant as is ϖ. In other words, ϖ' *cannot be defined on the manifold of rays* V.

[292] The classical *eikonal* function, which is deduced from the 1-form ϖ' (previous footnote), is not defined on the space of rays.

Planar mirrors

Let N be a vector in space-time satisfying

$$\overline{N} \cdot N = -1 \,. \tag{15.93}$$

It is easy to verify that the Poincaré group G has a Lie subgroup G' consisting of matrices (notation (13.51))

$$\begin{bmatrix} L & C \\ 0 & 1 \end{bmatrix} \qquad (\overline{L} \cdot L = 1) \,, \tag{15.94}$$

that satisfy the conditions

$$\overline{N} \cdot L = \overline{N} \qquad \text{and} \qquad \overline{N} \cdot C = 0 \,, \tag{15.95}$$

and whose Lie algebra thus consists of matrices

$$\begin{bmatrix} \Lambda & \Gamma \\ 0 & 0 \end{bmatrix} \qquad (\overline{\Lambda} + \Lambda = 0) \tag{15.96}$$

such that

$$\overline{N} \cdot \Lambda = 0 \qquad \text{and} \qquad \overline{N} \cdot \Gamma = 0 \,. \tag{15.97}$$

G' is the set of elements of G that preserve the half-space-time defined by

$$\overline{N} \cdot X > 0 \,, \tag{15.98}$$

which can be interpreted in 3-space as *one side of a stationary plane or of a plane undergoing uniform translation.*

Suppose that a scattering source has G' as a symmetry group. One can easily verify that the Lie algebra of G' is equal to its derived algebra. Theorem (15.73.4) then shows that the symplectomorphisms S_j^k of the scattering process preserve the moments of G'. These moments are deduced from those of G by (13.14◇). From (15.97) it follows that the only allowed transformation of the moment $\{M, P\}$ of the Poincaré group (notation (13.57)) is of the form

$$M \to M + V \cdot \overline{N} - N \cdot \overline{V} \,, \qquad P \to P + \alpha N \,, \tag{15.99}$$

where V is a vector and α a number.

If the type of the particle in the scattering process is known, we know some of the relations satisfied by M and P (see the classification of particles (14.4, 14.24, and 14.29)). Since these relations must be satisfied before and after scattering, they provide relations for V and α. We also have differential relations expressing the fact that the maps S_j^k are

symplectomorphisms. In two interesting cases — a particle without spin and a massless particle — we deduce that the transformation (15.99) takes one of the two forms

(15.100)
$$\text{a)}\quad P \to P\,, \qquad M \to M + [N \cdot \overline{P} - P \cdot \overline{N}]\, f(\overline{N} \cdot P)$$
$$\text{b)}\quad P \to AP\,, \qquad M \to A \cdot [M + [N \cdot \overline{P} - P \cdot \overline{N}]\, f(\overline{N} \cdot P)] \cdot \overline{A}\,,$$

where f is a real valued function and A is the linear map defined by

(15.101)
$$A = 1 + 2N \cdot \overline{N}\,.$$

Case (a) corresponds to transmission and case (b) to reflection. □

Let us consider the case of a *planar mirror* occupying the region $\overline{N} \cdot X \geq 0$, on which photons (massless particles) impinge. Choosing a Lorentz frame in which the mirror does not move, we can write

(15.102)
$$N = \mathcal{R}\begin{pmatrix} \mathbf{n} \\ 0 \end{pmatrix},$$

where $\mathbf{n}$ (Fig. 15.III) is the outward directed unit vector normal to the mirror.

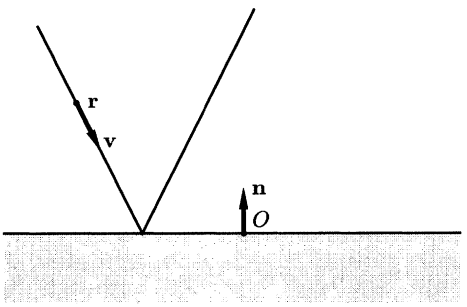


Fig 15.III

(15.103) A motion of the free photon can be described by a *time* t, a *position* $\mathbf{r}$, its *velocity* $\mathbf{v}$ (of length 1), its *energy* E, and its *helicity* $\chi = \pm 1$. The point $\mathcal{R}\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix}$ belongs to the *trajectory* of the free photon relative to the reference frame $\mathcal{R}$ (definition (14.33)).

A computation gives

(15.104)
$$P = \mathcal{R}\begin{bmatrix} \mathbf{v}\,E \\ E \end{bmatrix}$$
$$M = \mathcal{R} \cdot \begin{bmatrix} j(\chi\, s\, \mathbf{v} + \mathbf{r} \times \mathbf{v}\,E) & [\mathbf{r} - \mathbf{v}\,t]\,E \\ E\,[\mathbf{r} - \mathbf{v}\,t] & 0 \end{bmatrix} \cdot \mathcal{R}^{-1}\,.$$

The experimental laws of reflection show that reflection is described by the transformation

$$t \to t\,, \quad \mathbf{r} \to \Phi(\mathbf{r})\,, \quad \mathbf{v} \to \Phi(\mathbf{v})\,, \quad E \to E\,, \quad \chi \to -\chi\,, \tag{15.105}$$

[293]

where Φ is the *symmetry with respect to the mirror plane*, that is,

$$\Phi = 1 - 2\,\mathbf{n} \cdot \overline{\mathbf{n}}\,. \tag{15.106}$$

Observing that the linear map A (15.101) can be written as

$$A = \mathcal{R} \cdot \begin{bmatrix} \Phi & \\ & 1 \end{bmatrix} \cdot \mathcal{R}^{-1} \tag{15.107}$$

and using (15.104), it follows that the formulæ of reflection can be written as

$$P \to A \cdot P \qquad M \to A \cdot \overline{M} \cdot A\,, \tag{15.108}$$

which is a special case of (15.100): case b) with $f \equiv 0$. We thus can treat a mirror by scattering theory. The set Ω of initial conditions of a photon situated *in front* of the mirror is partitioned into open sets Ω_θ which are defined *with the aid of a time* θ by the conditions

$$\langle \mathbf{n}, \mathbf{r} \rangle > 0 \qquad \langle \mathbf{n}, \mathbf{r} + \mathbf{v}[\theta - t] \rangle > 0\,, \tag{15.109}$$

which express the fact that the photon does not hit the mirror in the time interval between t and θ. The relations (15.62) are satisfied (the indexing set here being time) and (15.66) gives us

$$\begin{aligned} S_\theta^{\theta'}(x) &= x \qquad \text{if} \quad \langle \mathbf{n}, \mathbf{r} + \mathbf{v}[\theta - t] \rangle > 0 \quad \text{and} \\ &\qquad\qquad\qquad\quad \langle \mathbf{n}, \mathbf{r} + \mathbf{v}[\theta' - t] \rangle > 0 \\ S_\theta^{\theta'}(x) &= S(x) \qquad \text{if} \quad \langle \mathbf{n}, \mathbf{r} + \mathbf{v}[\theta - t] \rangle > 0 \quad \text{and} \\ &\qquad\qquad\qquad\quad \langle \mathbf{n}, \mathbf{r} + \mathbf{v}[\theta' - t] \rangle < 0, \end{aligned} \tag{15.110}$$

where S denotes the involutive map defined on the space of free motions by (15.108).

From theorem (15.68) we know that these relations *completely define the (nonconnected) space of motions of a photon in the presence of a mirror*. Using the formulæ (15.110) one can directly verify the results (15.67) and (15.73). □

[293] A photon polarized to the left before reflection is polarized to the right after reflection. This is the reason why we had to adopt the point of view (14.78), according to which the photon is described by a *nonconnected* space of motions.

The above formulæ are *relativistic*. They allow us to study, for example, reflection in a mirror which is undergoing uniform translation with normal velocity $\mathbf{n}\,\beta$. It suffices to replace (15.102) by

(15.111) $$N = \mathcal{R}\begin{pmatrix} \mathbf{n}\cosh\varphi \\ \sinh\varphi \end{pmatrix} \qquad \tanh\varphi = \beta\,.$$

In this way we find that the angle of incidence α and the angle of reflection α' are not equal anymore, but are related by

(15.112) $$\cos\alpha' = \frac{\cos\alpha + \dfrac{2\beta}{1+\beta^2}}{1 + \dfrac{2\beta}{1+\beta^2}\cos\alpha}\,.$$

Moreover, the energy E of the incident photon and the energy E' of the reflected photon are related by

(15.113) $$\frac{E'}{E} = 1 + \frac{2\beta}{1-\beta^2}\,[\beta + \cos\alpha]$$

(*the Doppler effect due to reflection*). Finally the point $\mathbf{r}'$ of emission is *shifted* with respect to the point $\mathbf{r}$ of impact by the amount

(15.114) $$\mathbf{r}' - \mathbf{r} = \frac{2\chi\, s\,\beta}{1+\beta^2+2\beta\cos\alpha}\,\frac{\mathbf{v}\times\mathbf{n}}{E}\,,$$

whose orientation depends on the helicity χ. Its length is less than $\beta\,\lambda/\pi$, where λ is the "wavelength" $2\pi s/E$. This effect is due to the non-intrinsic nature of the trajectory (see (14.36)).

(15.115) Let us note that translating the mirror parallel to itself does not change the reflection (because it does not change the 4-vector N), whereas it does change the substance of the mirror. *This is an example where the scattering process is more symmetrical than the scattering source.*

In other cases where the formulæ (15.100) do not apply (for example, the phenomenon of a *mirage*, which does not change the sign of the helicity, or a charged particle passing through a planar capacitor, which changes its energy), *the symmetry of the scattering process is certainly less than the symmetry of the scattering source.*

Collisions of free particles

(15.116) Consider a system of free particles without interactions and whose masses are *nonzero*. Let U be the space of those motions for which the velocities of the particles are distinct and let V be the corresponding subset of the synchronous evolution space (definition (15.39)). Recall that if $y_0 \in V$, then it defines a *time* t_0 and a motion of the system $x_0 = P(y_0)$ in U.

Finally let us choose a positive number ρ and let us define the sets Ω_{in} and Ω_{out} by

(15.117)
$$\left[y_0 \in \Omega \underset{\text{out}}{\text{in}} \right] \iff \left[\begin{array}{ll} \text{for the motion } P(y_0) : \text{if} & t \lesseqgtr t_0, \\ \text{then} & \|\mathbf{r}_j - \mathbf{r}_k\| > \rho \qquad \forall j,k \end{array} \right].$$

One can check that Ω_{in} and Ω_{out} are *open subsets* of V.

Let us now assume that the particles interact and that the interaction can be described by means of a potential v (see (15.48)) which is zero if the mutual distances between any two particles is greater than ρ. One can then apply scattering theory to the open set $\Omega = \Omega_{\text{in}} \cup \Omega_{\text{out}}$. The sets Ω_{in} and Ω_{out} satisfy the conditions (15.62). It follows that (notation (15.65))

(15.118)
$$\begin{array}{llll} S^{\text{out}}_{\text{out}} = 1_{\Omega_{\text{out}}} & & S^{\text{out}}_{\text{in}} = S \\ S^{\text{in}}_{\text{out}} = S^{-1} & & S^{\text{in}}_{\text{in}} = 1_{\Omega_{\text{in}}} . \end{array}$$

where S is a local symplectomorphism from U to itself which characterizes the scattering of the particles, or, using a different terminology, their *collisions*. The map S assigns to each motion x *before* (a possible) *collision* the motion x' *after the collision* according to the rule (see (15.66)):

(15.119)
$$[S(x) = x'] \iff \left[\begin{array}{cc} \exists y \in \Omega_{\text{in}}, & \exists y' \in \Omega_{\text{out}} : \\ P(y) = x, & P(y') = x', \quad \widetilde{P}(y) = \widetilde{P}(y') \end{array} \right],$$

where $\widetilde{P}$ denotes the correspondence between an initial condition and the corresponding motion of the *interacting* particles.

(15.120) Using this formula one can verify that the map S does not change if one *increases* ρ. Thus S is characteristic for the interactions.

(15.121) Even if one does not know the *relativistic* formulation of interactions (see (15.60)), one might venture the following hypothesis: there exists a local symplectomorphism S, satisfying (15.119), which acts on the space $\widetilde{U}$ of motions of relativistic free particles without parallel trajectories in space-time.[294] One then should verify that S *does not depend* on the

[294]The particles can be identical (15.42) or not.

chosen Lorentz frame (although this frame is used implicitly in the definition of Ω_{in} and Ω_{out}). If this is the case, the maps (15.118) define the symplectic manifold of unconstrained motions (due to theorem (15.68)). Since the system is assumed to be isolated, the principle (13.70) of relativity must be applicable. It then follows that the map S *commutes* with the elements of the restricted Poincaré group G (which acts on $\tilde{U}$). Since the Lie algebra of G is equal to its derived algebra, theorem (15.73) shows that S preserves the moment $\mu = \{M, P\}$ of the system, which, according to (14.59), is equal to the sum of the corresponding moments of each of the particles. In particular, the total energy is conserved; one says that the collisions are *elastic*. □

Even if we are unable to find S by means of a computation, we know a priori that

(15.122)
$$\begin{cases} S \text{ is a local symplectomorphism of } \tilde{U}, \\ S \text{ commutes with the action of the restricted Poincaré group on } \tilde{U}, \\ S \text{ preserves the quantities } P = \sum_j P_j \text{ and } M = \sum_j M_j, \end{cases}$$

which limits its arbitrariness. There are numerous experimental verifications of (15.122). They allow us to test the principles which we have proposed for relativistic mechanics and are all the more valuable because the study of the map S is at present the *only access* we have to the study of relativistic non-elementary systems. □

(15.123) It is not impossible that knowing the symplectomorphism S will allow us also to study the constrained motions by the technique of *analytic continuation*. For this to be possible it is necessary that the manifold of motions and the map S are analytic. There are certain indications which suggest that this is indeed true for particular types of interactions.

Chapter IV

Statistical Mechanics

§16. Measures on a manifold

Composite manifolds

In this chapter we are going to extend the notion of a *manifold* by using the following conventions.

(16.1) By $\mathbf{R}^0$ we will denote the set consisting only of the number 0. The two subsets of $\mathbf{R}^0$ ($\mathbf{R}^0$ and the empty set) will be called *open*. A map A from $\mathbf{R}^n$ to $\mathbf{R}^0$ will be called *differentiable* if it is defined on an open set. For $x \in \mathrm{dom}(A)$, $D(A)(x)$ will be the (unique) map from $\mathbf{R}^n$ to $\mathbf{R}^0$. Every map A from $\mathbf{R}^0$ to $\mathbf{R}^n$ will be called differentiable. For $x \in \mathrm{dom}(A)$, we will denote by $D(A)(x)$ the map identically zero from $\mathbf{R}^0$ to $\mathbf{R}^n$. □

(16.2) Let V be a set and F a map. We will say that F is a *system of coordinates of* V if there exists an integer n ($n = 0, 1, 2, \ldots$) such that:

F is a map from $\mathbf{R}^n$ to V.
F is injective.
$\mathrm{dom}(F)$ is open in $\mathbf{R}^n$.

It is clear that the integer n is completely determined by F; it is called the *dimension* of the coordinate system F. □

(16.3) With these conventions the following definitions from chapter I still make sense and extend the notions involved:

- *compatible systems of coordinates* (1.5),
- *atlas* (1.6),
- *charts* (1.7).

By *manifold* we still mean any set V equipped with an atlas. The manifold structure on V is defined by the set of all charts (which is again an atlas). We can also extend the following concepts:

- an *open subset* of a manifold (1.13),
- a *continuous* (respectively *differentiable*) map from a manifold to a manifold (1.15) (respectively (1.18)),
- *connected* open sets (1.48) and *connected components* (1.49).

However, the concept of *dimension* needs to be modified.

(16.4) If V is a manifold and x is a point of V, then all charts F such that $x \in \mathrm{im}(F)$ have the same dimension n (16.2). The number n is called the *dimension of* V *at the point* x. □

(16.5) We will say that a manifold V is *pure* if the dimension, say n, is the same at all points. This n is then called the "dimension of V".

(16.6) The manifolds which have been studied up until now are thus *pure* manifolds of dimension ≥ 1.

(16.7) If V is a *pure* manifold of dimension 0, every subset of V is open.[295]

(16.8) A manifold which is not pure will be called a *composite manifold*, a name which is explained by the following theorem.

(16.9) THEOREM: Let V be a manifold. For each integer $n \geq 0$, let V_n be the set (possibly empty) of points x of V where the dimension equals n. Then V_n is an open subset of V and V_n possesses (in the sense of (1.39)) the structure of a pure manifold of dimension n.

PROOF: One can verify that the set of charts of V whose dimension is n is an atlas of V_n. Thus this gives V_n a pure manifold structure of dimension n. □

(16.10) We see that the sets V_n *partition* V into *disjoint open sets*. It follows immediately that each connected component of V (1.49) is contained in exactly one of the V_n. Thus a composite manifold has at least two components.

Conversely, if the V_j are pairwise disjoint pure manifolds of dimension n_j, one can give their union V a manifold structure by calling any chart of a V_j a chart of V. V will be a composite manifold if the n_j are not all equal.

(16.11) These results reduce the study of manifolds in the generalized sense (16.3) to the study of pure manifolds ((16.6) and (16.7)). Thus the new definition, which we adopt from now on, does not produce any new mathematical results, but instead is a convenient language for the investigations which will follow.

Compact sets

(16.12) DEFINITION: Let V be a manifold and Φ a subset of V. Φ is said to be *closed* if the *complement* of Φ in V is *open*.

[295] This fact is expressed by saying that V is a *discrete* topological space.

(16.12) THEOREM: Let V be a manifold, then

V and the empty subset are closed.
The union of two closed subsets of V is closed.
Every intersection of closed subsets of V is closed.

PROOF: The verification is immediate (see (1.14)). □

(16.13) DEFINITION: The *closure* of a subset Φ_0 of a manifold V is the intersection of all closed subsets of V that contain Φ_0. □

(16.14) It follows immediately from (16.12) that the closure of Φ_0 is *the smallest closed set containing* Φ_0.

If $V = \mathbf{R}^n$, the closure Φ of a set Φ_0 can be characterized as follows:

(16.15) $$[x \in \Phi] \iff [\text{every ball with center } x \text{ intersects } \Phi_0]\,.$$

(16.16) DEFINITION: Let V be a *Hausdorff* manifold and let K be a subset of V. A family of open subsets Ω_j $(j \in I)$ is called an (open) *cover*† of K if their union contains K:[296]

$$\left[\forall j \in I:\ \Omega_j \subseteq V,\ \Omega_j \text{ open}\right] \qquad \text{and} \qquad \left[K \subset \bigcup_{j \in I} \Omega_j\right].$$

K is called *compact* if every open cover of K contains a finite open cover. □

(16.17) It is clear that every *finite* subset of V is compact and that the union of two compact subsets of V is compact.

(16.18) A manifold V, considered as a subset of itself, can be *compact* (for example the sphere S_n) *or not* (for example a vector space).

(16.19) THEOREM: If V is a Hausdorff manifold and K is a compact subset of V, then K is closed.

(16.20) THEOREM (BOREL-LEBESGUE): Let K be a subset of $\mathbf{R}^n$. Then

$$[K \text{ is compact}] \iff [K \text{ is closed and bounded}]\,.^{297}$$

†Editors' note: Not to be confused with the notion of covering space (1.42).

[296]Recall that the *indexing set* I is arbitrary and thus not necessarily finite or countable. It is clear that every subset of V possesses an open cover, namely, the set of all open subsets whose indexing set is itself.

[297]A subset of $\mathbf{R}^n$ is *bounded* if it is contained in a ball.

(16.21) THEOREM: Let $V_1, \ldots, V_n$ be Hausdorff manifolds. If K_j is a compact subset of V_j for $j \in [1, 2, \ldots, n]$, then the *direct product* $K = K_1 \times K_2 \times \cdots \times K_n$ is a *compact subset of the product manifold* $V = V_1 \times \cdots \times V_n$.

(16.22) THEOREM: Let V and V' be two Hausdorff manifolds and let K be a compact subset of V. If A is a continuous map from K to V' (definition (1.15)), the image of A is a *compact* subset of V'. If in addition A is injective, then A^{-1} is *continuous*. □

Applying the above theorem with $V' = \mathbf{R}$, one can easily see that (due to (16.20)) every continuous real valued function on a compact set K is *bounded from above and below*, that is,

$$\exists\, a, b \in \mathbf{R} \;:\quad [a \leq f(x) \leq b \quad \forall\, x \in K] \tag{16.23}$$

and that f *attains its bounds*, that is,

$$\exists\, x_0,\, x_1 \in K \;:\quad [\, f(x_0) \leq f(x) \leq f(x_1) \qquad \forall\, x \in K\,]\,. \tag{16.24}$$

(16.25) A subset K of a Hausdorff manifold V is said to be *relatively compact* if the closure of K is compact.[298] Every subset of a relatively compact set is relatively compact. Every Hausdorff manifold V admits an *open cover* Ω_j such that the Ω_j are *relatively compact*.[299] □

(16.26) DEFINITION: Let V and V' be two Hausdorff manifolds. We will say that a map A from V to V' is *proper* if the inverse image under A of every relatively compact open subset of V' is a relatively compact open subset of V.†

(16.26) THEOREM: If A is a proper map from V to V', $\mathrm{dom}(A)$ is an open subset of V and A is continuous. If in addition B is a proper map from V' to V'' then $B \circ A$ is a proper map from V to V''.

---

[298] Example: a subset of $\mathbf{R}^n$ is relatively compact if it is bounded (see(16.20)).

[299] One can rephrase this theorem by saying that V is *locally compact*. Various concepts which we will investigate on Hausdorff manifolds can be extended to all *locally compact topological spaces* (see Bourbaki [9] and Naimark [27]).

† Editors' note: Since V and V' are locally compact and Hausdorff, this condition is equivalent to the standard definition that A be continuous and the inverse image of a compact set is compact.

Riesz spaces

(16.27) DEFINITION: E is called a *Riesz space* if the following four conditions hold.

1) E is a real vector space;
2) On E there is defined a *partial order* $\geq$:[300]

$$\diamondsuit \qquad \begin{cases} [X \geq Y,\ Y \geq Z] & \Longrightarrow \quad [X \geq Z] \\ [X \geq Y,\ Y \geq X] & \Longleftrightarrow \quad [X = Y]\,; \end{cases}$$

3) The partial order $\geq$ satisfies:

$$\begin{cases} [X \geq Y] & \Longleftrightarrow \quad [X - Y \geq 0] \\ [X \geq 0,\ s \geq 0 \quad (s \in \mathbf{R})] & \Longrightarrow \quad [sX \geq 0]\,; \end{cases}^{301}$$

4) Every pair X, Y of vectors of E possesses an *upper bound* $\sup(X,Y)$, which is an element of E such that

$$\clubsuit \qquad \begin{bmatrix} Z \geq X \\ Z \geq Y \end{bmatrix} \quad \Longleftrightarrow \quad [Z \geq \sup(X,Y)]\,. \qquad \square$$

(16.28) One can easily verify that the upper bound $\sup(X,Y)$ is *completely determined* by $\clubsuit$.

For any Riesz space one can immediately verify the relations

$$(16.29) \qquad \begin{cases} [X \geq Y] & \Longleftrightarrow \quad [-Y \geq -X] \\ [X \geq 0,\ Y \geq 0] & \Longrightarrow \quad [X + Y \geq 0]\,. \end{cases}$$

It follows that every pair of vectors X and Y possesses a *lower bound* $\inf(X,Y)$, characterized by[302]

$$(16.30) \qquad \begin{bmatrix} Z \leq X \\ Z \leq Y \end{bmatrix} \quad \Longleftrightarrow \quad [Z \leq \inf(X,Y)]\,,$$

and that it can be obtained by

$$(16.31) \qquad \inf(X,Y) = -\sup(-X,-Y)\,.$$

[300] A relation $\geq$ is called a *partial order* if it satisfies $\diamondsuit$. Note that it is not required that every pair of vectors X, Y satisfies one of the relations $X \geq Y$ or $Y \geq X$.

[301] X is called *positive* if $X \geq 0$.

[302] Of course we write $Z \leq X$ for $X \geq Z$.

If X is an element of a Riesz space, we define

(16.32) $$X^+ = \sup(X,0)\,, \qquad X^- = \sup(-X,0)\,, \qquad |X| = \sup(X,-X)\,.^{303}$$

One can easily establish the formulæ:

(16.33) $$X \equiv X^+ - X^- \qquad |X| \equiv X^+ + X^-$$

(16.34) $$|s\,X| = |s|\,|X| \qquad \forall\, s \in \mathbf{R}$$

(16.35) $$|X+Y| \leq |X| + |Y|$$

(16.36) $$\Big|\,|X| - |Y|\,\Big| \leq |X-Y|$$

(16.37) $$\sup(X,Y) = \tfrac{1}{2}\,[X+Y+|X-Y|]\,.$$

(16.38) THEOREM/DEFINITION: Let E be a Riesz space and E' a vector subspace of E.

1) The following conditions are equivalent.
   a) $[X,Y \in E'] \implies [\sup(X,Y) \in E']\,.$
   b) $[X \in E'] \implies [X^+ \in E']\,.$
   c) $[X \in E'] \implies [|X| \in E']\,.$
2) When one of these conditions is satisfied we will say that E' is a *Riesz subspace* of E. If E' is a Riesz subspace of E, then with the induced operations E' is Riesz space.

PROOF: Follows immediately from (16.32) and (16.37). □

(16.39) DEFINITION: Let E be a Riesz space. An element Φ of the dual space E^* of E[304] is said to be *positive* if

$\diamondsuit$ $$[X \geq 0] \implies [\Phi(X) \geq 0]\,.$$

A *measure* on E will be an element of E^* which is the *difference of two positive elements.*

(16.39) THEOREM: The set E' of measures on E is a vector subspace of E^*. Moreover E' is itself a Riesz space such that

$\heartsuit$ $$\left[\begin{array}{l}\Phi \in E^* \\ \Phi \text{ is positive}\end{array}\right] \iff \left[\begin{array}{l}\Phi \in E' \\ \Phi \geq 0\end{array}\right]\,.$$ □

[303] These formulæ apply in particular to the Riesz space $E = \mathbf{R}$. But one should remember that in general $|X|$ (the "absolute value of X") is an element of E and *not an element of* $\mathbf{R}$.

[304] By definition the space E^* is the set of linear maps from E to $\mathbf{R}$. E^* is also called the space of *linear forms* on E.

The reader will find the proof in Bourbaki[305] as well as a proof of the formula

$$X \geq 0 \implies |\Phi|(X) = \sup_{|Y| \leq X} \Phi(Y) \qquad \forall \Phi \in E' . \tag{16.40}$$

(16.41) THEOREM: $[|\Phi| \leq \Phi',\ |X| \leq X'] \implies |\Phi(X)| \leq \Phi'(X')$.

PROOF: On the one hand

$$\Phi'(X') - |\Phi|(|X|) = [\Phi' - |\Phi|](X') + |\Phi|(X' - |X|) \geq 0 ,$$

while on the other hand

$$\begin{aligned} |\Phi(X)| &= |[\Phi^+ - \Phi^-](X^+ - X^-)| \\ &= |[\Phi^+(X^+) + \Phi^-(X^-)] - [\Phi^+(X^-) + \Phi^-(X^+)]| \\ &\leq [\Phi^+ + \Phi^-](X^+ + X^-) = |\Phi|(|X|) . \end{aligned}$$

□

Measures

(16.42) DEFINITIONS: Let V be a Hausdorff manifold. We will denote by $\mathcal{C}(V)$ the set of *continuous* maps from V to $\mathbf{R}$ (= the set of "continuous functions on V"). For $f \in \mathcal{C}(V)$ the *support* of f is the *closure* of the set of points x in V such that $f(x) \neq 0$. $\mathcal{K}(V)$ will be the set of functions in $\mathcal{C}(V)$ whose support is *compact*.

(16.43) THEOREM: The space $\mathcal{C}(V)$ is a vector space. If we define $\forall\, f, g \in \mathcal{C}(V)$

$$[f \geq g] \iff [f(x) \geq g(x) \quad \forall\, x \in V] ,$$

then $\mathcal{C}(V)$ becomes a *Riesz space*. □

The verification is immediate and leads to the following formulæ

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) \leq 0 \end{cases} \qquad f^-(x) = \begin{cases} -f(x) & \text{if } f(x) \leq 0 \\ 0 & \text{if } f(x) \geq 0 \end{cases} \tag{16.44}$$

$$|f|(x) = |f(x)| \qquad \forall\, x \tag{16.45}$$

$$\begin{aligned} \sup(f,g)(x) &= \sup(f(x), g(x)) \\ &= \tfrac{1}{2}[\, f(x) + g(x) + |f(x) - g(x)|\,] \qquad \forall\, x . \end{aligned} \tag{16.46}$$

[305] See Bourbaki [9].

Since the support of $|f|$ is equal to the support of f, one can verify immediately, using (16.38), that

(16.47) $\mathcal{K}(V)$ is a Riesz subspace of $\mathcal{C}(V)$. □

One can prove the following lemma.

(16.48) LEMMA (URYSOHN): Let K be a compact subset of a Hausdorff manifold V. Then there exists an $f \in \mathcal{K}(V)$ such that

$$[0 \leq f(x) \leq 1 \qquad \forall x \in V] \qquad \text{and} \qquad [f(x) = 1 \qquad \forall x \in K]\,. \quad \square$$

One can also prove the following more detailed result.[306]

(16.49) THEOREM: Let Ω_j be a finite open cover of a compact set K. Then there exist *differentiable* functions f_j on V such that

$$\begin{cases} f_j \geq 0 & \forall j\,, \\ \text{support}(f_j) \subseteq \Omega_j & \forall j\,, \\ \sum_j f_j(x) \leq 1 \qquad \text{and} \qquad \left[\sum_j f_j(x) = 1 \qquad \forall x \in K\right]. \end{cases}$$

(16.50) DEFINITION: By a *measure* on the manifold V (presumed to be Hausdorff) we mean any measure on the Riesz space $\mathcal{K}(V)$ (definition (16.39)).



In other words, *a measure μ on V is a linear form on the vector space of continuous functions with compact support, which is either positive or the difference of two positive measures.*

One can prove the following theorem,[307] which provides a new definition of a measure on V.

(16.51) THEOREM: Let V be a Hausdorff manifold and let μ be a linear form on $\mathcal{K}(V)$. Then the following statements are equivalent.

1) μ is a measure.
2) For each compact subset K of V there exists a number M such that

$\diamondsuit$ $\qquad |\mu(f)| \leq M\|f\|$[308] $\qquad$ if $f \in \mathcal{K}(V)$ and $\text{support}(f) \subseteq K$. □

[306] See Bourbaki [9] or Schwartz [28].
[307] See Bourbaki [9].
[308] The notation $\|f\| = \sup_{x \in V} |f(x)|$ will be defined in (16.110$\diamondsuit$).

(16.52) The set of measures on V form a Riesz space (theorem (16.39)). To each measure μ we thus can associate the positive measures μ^+, μ^-, and $|\mu|$. Applying the formulæ (16.40) and (16.41) we obtain

$$[f \geq 0] \implies [|\mu|(f) = \sup_{|g| \leq f} \mu(g)] \qquad f, g \in \mathcal{K}(V) \tag{16.53}$$

$$[|\mu| \leq \mu',\ |f| \leq f'] \implies |\mu(f)| \leq \mu'(f')\,. \tag{16.54}$$

For $f, g \in \mathcal{C}(V)$ we will denote by $f \times g$ the usual product

$$[f \times g](x) = f(x)\, g(x) \qquad \forall\, x \in V\,. \tag{16.55}$$

It is clear that $f \times g \in \mathcal{C}(V)$. Thus the operation $\times$ gives $\mathcal{C}(V)$ the structure of a *commutative algebra*.[309]

For a measure μ on V and a continuous function f on V we define

$$[\mu \times f](g) = \mu(f \times g) \qquad \forall\, g \in \mathcal{K}(V)\,. \tag{16.56}$$

(16.57) One can easily verify that $\mu \times f$ is again a measure called the *product of the measure μ and the continuous function f*. $\mu \times f$ is positive if μ and f are both positive. Since the map $(\mu, f) \to \mu \times f$ is bilinear and associative (on the right), one says that *the algebra $\mathcal{C}(V)$ acts on the space of measures on V*.

(16.58) THEOREM: (Notation (16.56).) $|\mu \times f| = |\mu| \times |f|$.

PROOF: Using formula (16.53) we have to show that for every $g \geq 0$ in $\mathcal{K}(V)$ the numbers

$$X = \sup_{k \in I} \mu(k) \qquad \text{and} \qquad X' = \sup_{k' \in I'} \mu(k')$$

are equal, where the sets I and I' are defined by

$$[k \in I] \iff [\exists h \in \mathcal{K}(V)\ :\ |h| \leq g\,,\ k = f \times h]$$

and

$$[k' \in I'] \iff [k' \in \mathcal{K}(V)\,,\ |k'| \leq |f|\, g]\,.$$

It is clear that $I \subseteq I'$ and thus $X \leq X'$.

To prove $\geq$, let K be the support of g and let M be the number associated with μ and K by (16.51$\diamondsuit$), that is,

$$\heartsuit\ \ [\ell \in \mathcal{K}(V)\,,\ \text{support}(\ell) \subseteq K\,,\ |\ell(x)| \leq \lambda\ \forall\, x] \implies [|\mu(\ell)| \leq M\, \lambda]\,.$$

[309] That is, the operation $\times$ is *bilinear*, *associative*, and *commutative*.

If ε is a positive number, let us denote by φ the continuous function whose graph (see figure (16.I)) is a line segment joining two pieces of a hyperbola. For each $k' \in I'$ let us define

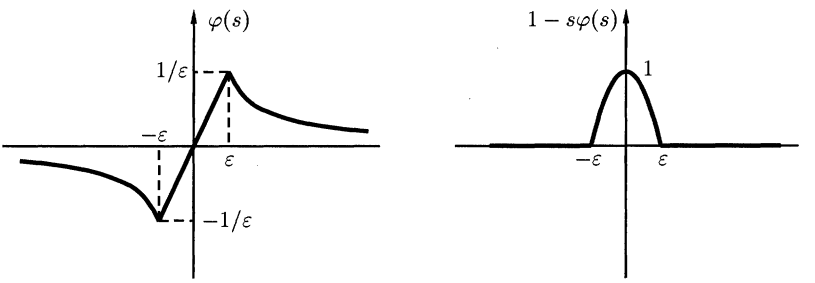
$$h(x) = k'(x)\,\varphi(f(x))\,, \qquad k = f \times h\,, \qquad \text{and} \qquad \ell = k' - k\,.$$

One can easily verify that $h \in \mathcal{K}(V)$ and that $|h| \leq g$. Thus $k \in I$ and $|\mu(k)| \leq X$. On the other hand we have

$$[\ell(x) \neq 0] \implies [1 - f(x)\,\varphi(f(x)) \neq 0] \implies [|f(x)| \leq \varepsilon]$$
$$\implies |\ell(x)| \leq |k'(x)| \leq |f(x)|\,g(x) < \varepsilon\,\|g\| \quad \text{(notation (16.110}\diamondsuit\text{))}.$$

Since $\ell \in \mathcal{K}(V)$ and $\text{support}(\ell) \subseteq K$, it follows from $\heartsuit$ that $|\mu(\ell)| \leq \varepsilon\,M\,\|g\|$. Hence $\mu(k') = \mu(k) + \mu(\ell) \leq X + \varepsilon\,M\,\|g\|$ and $X' \leq X + \varepsilon\,M\,\|g\|$. Since ε is arbitrary, we conclude that $X' = X$. □

Fig 16.I



The tensor product of measures

(16.59) THEOREM: Let V and V' be two Hausdorff manifolds having measures μ and μ', respectively. Then there exists a measure $\mu \otimes \mu'$ on $V \times V'$, called the *tensor product* of μ and μ', which is defined by

$$\diamondsuit \qquad [\mu \otimes \mu'](f) = \mu\left(x \mapsto \mu'\left(x' \mapsto f\begin{pmatrix} x \\ x' \end{pmatrix}\right)\right) \qquad \forall\, f \in \mathcal{K}(V \times V')\,.$$

This theorem implicitly contains the following results. The product $V \times V'$ is a Hausdorff manifold. For $f \in \mathcal{K}(V \times V')$ and $x \in V$, the map $x' \mapsto f\begin{pmatrix} x \\ x' \end{pmatrix}$ belongs to $\mathcal{K}(V')$. Similarly the map $x \mapsto \mu'\left(x' \mapsto f\begin{pmatrix} x \\ x' \end{pmatrix}\right)$ belongs to $\mathcal{K}(V)$. The reader will find the complete proofs in Bourbaki. □

(16.60) THEOREM: $[\mu \geq 0,\ \mu' \geq 0] \implies [\mu \otimes \mu' \geq 0]$. □

A simple calculation gives the following theorem:

(16.61) THEOREM: The tensor product $\otimes$ of measures is bilinear and associative.

Verification of associativity leads to the formula

$$[\mu \otimes \mu' \otimes \mu''](f) = \tag{16.62}$$
$$= \mu\left(x \mapsto \mu'\left(x' \mapsto \mu''\left(x'' \mapsto f\begin{pmatrix} x \\ x' \\ x'' \end{pmatrix}\right)\right)\right).$$

For $g \in \mathcal{C}(V)$ and $g' \in \mathcal{C}(V')$ we define

$$[g \otimes g']\begin{pmatrix} x \\ x' \end{pmatrix} \equiv g(x)\, g'(x'). \tag{16.63}$$

It is immediate that for every measure μ on V and every measure μ' on V' we have

$$[\mu \otimes \mu'](g \otimes g') = \mu(g)\, \mu'(g') \qquad \forall g \in \mathcal{K}(V)\ \forall g' \in \mathcal{K}(V'). \tag{16.64}$$

Conversely, one can prove the following theorem, see Bourbaki.

(16.65) THEOREM: The *only* measure ν on $V \times V'$ satisfying the condition

$$\nu(g \otimes g') = \mu(g)\, \mu'(g') \qquad \forall g \in \mathcal{K}(V)\ \forall g' \in \mathcal{K}(V')$$

is the tensor product $\mu \otimes \mu'$. □

From this theorem one can easily deduce that the formula

$$[\mu \otimes \mu']\left(\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto y\right) = [\mu' \otimes \mu]\left(\begin{pmatrix} x' \\ x \end{pmatrix} \mapsto y\right) \tag{16.66}$$

is valid if μ and μ' are measures and if the map $\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto y$ belongs to $\mathcal{K}(V \times V')$.

One should also note the following formulæ:

$$[\mu \times f] \otimes [\mu' \times f'] = [\mu \otimes \mu'] \times [f \otimes f'] \qquad \forall f \in \mathcal{C}(V)\ \forall f' \in \mathcal{C}(V') \tag{16.67}$$

$$|\mu \otimes \mu'| = |\mu| \otimes |\mu'| \tag{16.68}$$

$$\text{support}(\mu \otimes \mu') = \text{support}(\mu) \times \text{support}(\mu') \qquad \text{(see (16.102))}. \tag{16.69}$$

Examples of measures

(16.70) Let V be a manifold of dimension 0. Since each subset of V is open (16.7), one can easily verify that V is Hausdorff, that every map from V to $\mathbf{R}$ is continuous, that every compact subset of V is finite, and thus that $\mathcal{K}(V)$ is the set of maps from V to $\mathbf{R}$ which are zero except on a finite subset. Therefore for each $g \in \mathcal{K}(V)$ one can define a number $I(g)$ by

$$\diamondsuit \qquad I(g) = \sum_{x \in V} g(x) \,.$$

One can verify immediately that I is a *positive measure* on V and that *every* measure on V is of the form

$$\heartsuit \qquad I \times f \,, \qquad f \in \mathcal{C}(V) \,. \qquad \square$$

(16.71) Let λ be a *nondecreasing* function on $\mathbf{R}$. For $f \in \mathcal{K}(\mathbf{R})$ one can show that the sum

$$\sum_j f(x_j)\,[\lambda(x_{j+1}) - \lambda(x_j)] \,,$$

where x_j is an increasing sequence which encloses the support of f, has a limit S as the largest of the numbers $x_{j+1} - x_j$ tends to zero. The limit S is called the *Stieltjes integral* and is written as

$$(16.72) \qquad \int_{-\infty}^{\infty} f(x)\, d\lambda(x) \,.$$

One can easily verify that the map $f \mapsto S$ is a positive measure on $\mathbf{R}$. Conversely, one can show that every positive measure on $\mathbf{R}$ is obtained by this procedure with the aid of a nondecreasing function λ. Thus one can obtain every measure on $\mathbf{R}$ as a Stieltjes integral involving a function λ which is the difference of two strictly increasing functions ("a function of *bounded variation*"). $\square$

In particular, the *Lebesgue measure* I on $\mathbf{R}$ is defined by the formula

$$(16.73) \qquad I(f) = \int_{-\infty}^{\infty} f(x)\, dx \qquad \forall\, f \in \mathcal{K}(\mathbf{R}) \,;$$

it clearly is a positive measure. $\square$

If I is the Lebesgue measure on $\mathbf{R}$, the tensor product I_n of n measures equal to I is a measure on $\mathbf{R}^n$, still called the Lebesgue measure; it is defined by

$$(16.74) \qquad I_n(f) = \int_{-\infty}^{\infty} d\,{}^1x \,\ldots \int_{-\infty}^{\infty} d\,{}^nx \; f \begin{pmatrix} {}^1x \\ \vdots \\ {}^nx \end{pmatrix} \qquad \forall\, f \in \mathcal{K}(\mathbf{R}^n) \,.$$

According to theorem (16.66) one can arbitrarily change the order of integration in (16.74). One often uses the following notation for the multiple integral $I_n(f)$:

(16.75) $$\int f(x)\, d^{\,1}x\, d^{\,2}x \ldots d^{\,n}x\,. \qquad \square$$

Let V be a Hausdorff manifold and let a be a point of V. The *Dirac measure at a* is the measure δ_a defined by

(16.76) $$\delta_a(f) = f(a) \qquad \forall\, f \in \mathcal{K}(V)\,. \qquad \square$$

Let A be a homeomorphism [310] from a Hausdorff manifold V to a manifold V^*. One can verify immediately the following theorem.

(16.77) THEOREM: If μ is a measure on V, there exists a measure $A^+(\mu)$ on V^* called the *push forward of* μ *by* A, which is defined by

$$A^+(\mu)(f^*) = \mu(f^* \circ A) \qquad \forall\, f^* \in \mathcal{K}(V^*)\,.$$

We will generalize this concept below in (16.149). $\square$

EXAMPLE: The push forward of the measure $\mu \otimes \mu'$ by the map $\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto \begin{pmatrix} x' \\ x \end{pmatrix}$ is the measure $\mu' \otimes \mu$ (notation (16.64)). This is just another way of formulating theorem (16.66).

Completely continuous measures

(16.78) DEFINITION: Let E be a vector space of dimension n. A *density* on E is a function φ defined on the set of *bases* of E satisfying

$$\varphi(S \cdot M) = \varphi(S)\,|\det(M)|$$

for every basis S and every matrix $M \in \mathrm{Gl}(\mathbf{R}^n)$. $\square$

One can easily verify that densities on E form a vector space of dimension 1; it is even a Riesz space when we define $\varphi \geq 0$ if $\varphi(S) \geq 0$ $\forall\, S$.

(16.79) Let V be a pure Hausdorff manifold of dimension $n\,(n \geq 1)$ and let $x \mapsto \varphi$ be a *field of densities* on V, that is, for $x \in V$, φ is a density

[310] That is, a continuous bijection whose inverse is also continuous. For example, a diffeomorphism.

on the tangent space D_x of V at x. Let us assume that this field is *continuous*, that is, the correspondence $\xi \mapsto \varphi(D(F)(\xi))$ is continuous for every chart $F[\xi \mapsto x]$ of an atlas of V. This definition of continuity does not depend on the choice of the atlas. We are going to associate to this field of densities a measure μ on V.

First suppose that f is a real valued function on V whose support is compact and is contained in the image of a chart $F[\xi \mapsto x]$. We then define[311]

$$\mu(f) = \int f(x)\,\varphi\left(\frac{\partial x}{\partial \xi}\right) d^{\,1}\xi\, d^{\,2}\xi \ldots d^{\,n}\xi\,. \tag{16.80}$$

The "change of variables formula in multiple integrals" shows that the above integral *does not depend on the choice of the chart* F, which justifies the notation $\mu(f)$.

Now let A be an atlas of V and let f be a continuous function on V with compact support. One can show[312] that f can be written as a finite sum

$$f = \sum_{j=1}^{N} f_j\,, \tag{16.81}$$

where each term f_j is a continuous function whose support is compact and contained in the image of one of the charts of the atlas. One can verify that the sum $\sum_j \mu(f_j)$[313] does not depend on either the choice of the atlas or the decomposition (16.81). Thus we can define

$$\mu(f) = \sum_j \mu(f_j)\,. \tag{16.82}$$

This defines a function μ on $\mathcal{K}(V)$. It is immediate that μ is a measure on V which is positive if the density φ is positive.[314] Conversely, if μ is positive, one can show that $\varphi \geq 0 \quad \forall x$.[315] To summarize:

(16.83) THEOREM: Let $x \mapsto \varphi$ be a *continuous field of densities* on a pure Hausdorff manifold V. Corresponding to this field of densities there is a *measure* μ on V, *characterized* by formula (16.80), which is valid

[311] One calculates this integral over all of $\mathbf{R}^n$ after extending f by 0 outside the domain of the chart (see (16.100)).

[312] Apply (16.16) and (16.49).

[313] Of which each term is defined by (16.80).

[314] See (16.49) and (16.81).

[315] If there exists an x_0 where $\varphi < 0$, one can find an open set Ω such that $x_0 \in \Omega$ and $[x \in \Omega] \Rightarrow [\varphi < 0]$, because of the continuity of $x \mapsto \varphi$. One can find an $f \in \mathcal{K}(V)$, having its support contained in Ω and in the image of a chart, such that $f \geq 0$ and $f \neq 0$. Using (16.80) one can verify that $\mu(f) < 0$. This then is a contradiction.

whenever f has compact support contained in the image of a chart $\xi \mapsto x$. We write

$\diamondsuit$ $$\mu(f) = \int_V f(x)\,\varphi(dx)\,.$$

Moreover

$\heartsuit$ $$[\mu \geq 0] \iff [\varphi \geq 0 \qquad \forall x]\,.$$ □

(16.84) A measure μ will be called *completely continuous measure*[316] if it is defined by a continuous field of densities $x \mapsto \mu$. It is clear that the correspondence $[x \mapsto \varphi] \mapsto \mu$ is *linear*. From (16.83$\heartsuit$) it follows that it is also *injective*.

Examples of completely continuous measures

(16.85) The Lebesgue measure on $\mathbf{R}^n$ (16.75) is completely continuous because it can be defined by means of the "Lebesgue density" φ

$$\varphi(S) \equiv |\det(S)|\,.$$

Note that a basis S of $\mathbf{R}^n$ is simply an invertible matrix. □

(16.86) Let φ be a continuous field of densities on a manifold V and let μ be the associated measure:

$$\mu(f) = \int_V f(x)\,\varphi(dx)\,.$$

If $A\,[x \mapsto x^*]$ is a diffeomorphism from V onto a manifold V^*, then one can verify without difficulty that the push forward $A^+(\mu)$ (16.77) is given by the formula

$\diamondsuit$ $$A^+(\mu)(f^*) = \int_{V^*} f^*(x^*)\,\varphi^*(dx^*) \qquad \forall f^* \in \mathcal{K}(V^*)\,,$$

where $[x^* \mapsto \varphi^*]$ is the field of densities on V^* defined by

$\heartsuit$ $$\varphi^*(S^*) \equiv \varphi\left(\frac{\partial x}{\partial x^*} S^*\right),$$

which is called *the push forward by A of the density φ*. The field $[x^* \mapsto \varphi^*]$ is also continuous, and hence $A^+(\mu)$ is a completely continuous measure. □

[316] Sometimes one extends this name to all elements of $L^1(V)$ (see (16.129) below).

(16.87) Let $x = \begin{pmatrix} {}^1x \\ \vdots \\ {}^nx \end{pmatrix}$ be a variable point of the direct product $V = V_1 \times \cdots \times V_n$ of Hausdorff manifolds V_j. Assume that on each V_j we have a continuous field ${}^jx \mapsto \varphi_j$ of densities to which are associated the measures μ_j:

$$\diamondsuit \qquad \mu_j(f_j) = \int_{V_j} f_j({}^jx)\,\varphi_j(d\,{}^jx) \qquad \text{(do not sum over } j\text{!)}$$

If $S^{(j)}$ is a basis of the tangent space of V_j at ${}^jx$, one can easily verify that

$$\heartsuit \qquad S = \begin{bmatrix} S^{(1)} & 0 & \cdots & 0 \\ 0 & S^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S^{(n)} \end{bmatrix}$$

is a basis of the tangent space of V at x (see (1.23)) and that there exists a *continuous field of densities* $[x \mapsto \varphi]$ on the direct product V characterized by

$$\clubsuit \qquad \varphi(S) \equiv \varphi_1(S^{(1)}) \times \varphi_2(S^{(2)}) \times \cdots \times \varphi_n(S^{(n)})$$

for arbitrary $S^{(j)}$. We will write

$$\varphi = \varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_n\,, \tag{16.88}$$

or, in the notation of (16.83$\diamondsuit$)

$$\varphi(dx) \equiv \varphi_1(d\,{}^1x) \cdots \varphi_n(d\,{}^nx)\,. \tag{16.89}$$

(16.90) Using (16.80) and (16.64) one can verify that the measure associated with the field $x \mapsto \varphi_1 \otimes \cdots \otimes \varphi_n$ is the *tensor product of measures* associated with the densities $\varphi_1, \ldots, \varphi_n$. Thus *the tensor product of completely continuous measures is a completely continuous measure*, hence

$$\begin{aligned} &\int_{V_1 \times \cdots \times V_n} f(x)[\varphi_1 \otimes \cdots \otimes \varphi_n]\,(dx) \;= \\ &\qquad = \int_{V_1} \varphi_1(d\,{}^1x) \int_{V_2} \varphi_1(d\,{}^2x) \int_{V_n} \varphi_n(d\,{}^nx)\,f(x)\,. \end{aligned} \tag{16.91}$$

Theorem (16.66) allows one to change the order of integration.

(16.92) Let G be a Lie group with Lie algebra $\mathcal{G}$ and let φ_0 be a density on the vector space $\mathcal{G}$. One can easily verify that on G there exists a unique field of densities $a \mapsto \varphi_L$ which is *invariant* under the left translations

$\underline{a}_G$ of G [317] (definition (6.6)) and whose value at $a = e$ is φ_0. This field is given by

$$\diamondsuit \qquad \varphi_L(S) \equiv \varphi_0([D(\underline{a}_G)(e)]^{-1} \cdot S)\,.$$

Looking at $\diamondsuit$ we see that this field is continuous. Thus it defines a measure μ_L on G,

$$\heartsuit \qquad \mu_L(f) = \int_G f(x)\,\varphi_L(dx)\,,$$

which is *invariant under left translations*.[318] The measure μ_L is called a (left invariant) *Haar measure* on G. It is defined up to a constant factor by the choice of φ_0 (see (16.78)). One can show that these Haar measures are the only measures that are invariant under left translations.

In an analogous way one can define a density φ_R which is invariant under right translations (6.21) and which coincides with φ_0 at the point e. In this way we obtain a *right invariant* Haar measure μ_R. One can easily verify that

$$\clubsuit \qquad \mu_R = \mu_L \times \lambda\,,$$

where λ is the function on G defined by

$$\spadesuit \qquad \lambda(a) \equiv |\det(\underline{a}_{\mathcal{G}})|\,,$$

with $a \mapsto \underline{a}_{\mathcal{G}}$ the *adjoint representation* of G (6.24).

The function λ is called the *modular function* of G. When $\lambda(a) \equiv 1$, G is called *unimodular group*. Clearly every abelian Lie group is unimodular. Similarly one can show that every *compact* or *semisimple* Lie group is unimodular. For unimodular groups the same Haar measure is invariant under left and right translations. □

(16.93) The Lebesgue measure is invariant under left and right translations of the additive group $\mathbf{R}^n$. For the Lie group of real matrices $a \equiv \begin{bmatrix} x & y \\ 0 & 1/x \end{bmatrix}$ the modular function is x^2. □

(16.94) LEMMA: Let E be a *Euclidean* (respectively *symplectic*) vector space. There exists a positive density φ on E called the *Euclidean* (respective *symplectic*) *density* characterized by

$$\diamondsuit \qquad \varphi(S) = 1 \qquad \text{for every } \textit{orthonormal} \text{ (respectively, } \textit{canonical}\text{) basis of } E.$$

[317] That is, φ_L equals its push forward (16.86) under these translations.

[318] That is, the push forward of μ_L under left translations, in the sense of (16.77), equals μ_L.

PROOF: It suffices to choose an orthonormal (respectively canonical) basis S_0. Then $\diamondsuit$ and (16.78) give for each basis S

$$\varphi(S) = |\det(S_0^{-1} \cdot S)| . \tag{16.95}$$

Conversely, taking (16.95) as a definition of φ, we see that φ is a positive density. If S is also an orthonormal (respectively symplectic) basis, we know that $|\det(S_0^{-1} \cdot S)| = 1$ ((6.63) respectively (8.30)), whence $\diamondsuit$. □

(16.96) Let V be a *Hausdorff symplectic manifold.* At each point x of V the preceding lemma allows us to define a symplectic density φ. By taking a canonical atlas one can show that the field $x \mapsto \varphi$ is continuous. We thus obtain a measure μ defined by

$$\mu(f) \equiv \int_V f(x)\varphi(dx) .$$



The measure μ is called the *Liouville measure* on V. It is a *positive* measure.

If the support of f is contained in the image of a canonical chart $[\xi \mapsto x]$, it follows immediately from (16.80) and (16.94$\diamondsuit$) that

$$\mu(f) = \int \cdots \int f(x)\, dp_1 \ldots dp_n\, d[^1q] \ldots d[^nq] , \tag{16.97}$$

using the usual notation (10.4) for canonical coordinates.

With the aid of the above formula and (16.87) one can prove the following theorem.

(16.98) THEOREM: If μ_j is the Liouville measure of the symplectic manifold V_j, then the Liouville measure of the symplectic direct product manifold (9.7) is the tensor product (16.45) of the measures μ_j. □

Let V be a Hausdorff symplectic manifold and A an element of $\mathrm{Can}(V)$ (notation (10.15)). Comparing (16.94) with (16.86$\heartsuit$) one can verify that the push forward by A of the symplectic density φ on V equals φ. Hence by applying (16.86) we obtain *Liouville's theorem.*

(16.99) THEOREM (LIOUVILLE): The Liouville measure on a Hausdorff symplectic manifold V is *invariant* under every symplectomorphism of V.

The support of a measure

(16.100) Let Ω be an open subset of a Hausdorff manifold V. We know that Ω is a Hausdorff manifold of the same dimension as V. For $f \in \mathcal{K}(\Omega)$ the function $\widehat{f}$ on V defined by

$$\diamondsuit \qquad \widehat{f}(x) = \begin{cases} f(x) & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega \end{cases}$$

is an extension of f. One can verify that $\widehat{f} \in \mathcal{K}(V)$.

Thus, if μ is a measure on V, we can define a map μ_Ω by $\mu_\Omega(f) \equiv \mu(\widehat{f})$. One can verify that μ_Ω is a measure on Ω, which is positive if μ is positive. The measure μ_Ω is called the measure *induced* by μ on the open set Ω.

(16.101) THEOREM: Let Ω be an open set. The map $[\mu \mapsto \mu_\Omega]$ is *linear* and is increasing as a map between the Riesz spaces of measures on V and Ω. Moreover, $|\mu_\Omega| \equiv |\mu|_\Omega$. □

(16.102) The union of the open sets Ω_j satisfying $[\mu_{\Omega_j} = 0]$ is an open set Ω (possibly empty). One can verify that $\mu_\Omega = 0$. Thus Ω is the *largest open subset* such that $\mu_\Omega = 0$. The complement of Ω, which is closed, is called the *support* of the measure μ.

(16.103) EXAMPLES: The support of the Dirac measure δ_a is the point a. The support of the measure corresponding to a continuous field of densities $x \mapsto \varphi$ is the closure of the set of x such that $\varphi(x) \neq 0$ (see (16.42)). □

(16.104) THEOREM:[319] Every function $f \in \mathcal{K}(V)$ which vanishes on the support of a measure μ satisfies $\mu(f) = 0$.

Bounded measures

(16.105) DEFINITION: Let E be a (real or complex) vector space. E is called a *normed vector space* if to each $X \in E$ is associated a real number $\|X\|$, called the *norm* of X, such that

$$\diamondsuit \qquad \begin{cases} \|X\| > 0 & \text{if } X \neq 0 \\ \|sX\| = |s| \times \|X\| & \text{if } X \in E \text{ and } s \text{ a scalar}^{320} \\ \|X + Y\| \leq \|X\| + \|Y\| & \forall\, X, Y \in E\,. \end{cases}$$

[319] See Bourbaki [9].

[320] $|s|$ denotes the absolute value of the real or complex number s. It follows that $[\|X\| = 0] \Leftrightarrow [X = 0]$.

EXAMPLE: $\mathbf{R}$ (or $\mathbf{C}$) is evidently a normed vector space, as is $\mathbf{R}^n$ (1.2).

(16.106) DEFINITION: A *Banach space*, also called a *complete* normed space, is a normed vector space where every *absolutely convergent* series is *convergent*:[321]

$$\left[X_j \in E \quad \forall j = 1, 2, \ldots : \ \exists a \in \mathbf{R} \ \forall n : \ \sum_{j=1}^{n} \|X_j\| \leq a\right] \implies$$

$$\implies \quad \left[\exists X \in E \ \forall \varepsilon > 0 \ \exists N \ \forall n \geq N : \ \|X - \sum_{j=1}^{n} X_j\| \leq \varepsilon\right].$$

EXAMPLE: One can show that every *finite dimensional* normed vector space is a Banach space.

(16.107) DEFINITION: Let E and E' be two normed vector spaces, and let $L(E, E')$ be the set of linear maps from E to E'. A map $A \in L(E, E')$ is called *bounded* if the number

$$\diamondsuit \qquad \|A\| = \sup_{\|X\| \leq 1} \|A(x)\|$$

is finite. If that is the case, we have

$$\heartsuit \qquad \|A(x)\| \leq \|A\| \times \|X\| \qquad \forall X \in E\,.$$

(16.107) THEOREM: The set of bounded elements of $L(E, E')$ is a *normed vector space*. It is a Banach space if E' is a Banach space.

PROOF: Elementary. □

(16.108) EXAMPLE: Let E be a normed vector space. The *topological dual space* of E is the set E' of bounded elements of the dual space E^* ($= L(E, \mathbf{R})$ or $L(E, \mathbf{C})$). Thus E' is a Banach space. One can prove the Hahn-Banach theorem:

(16.109) $$\|X\| = \sup_{C \in E',\ \|C\| \leq 1} |C(X)| \qquad \forall X \in E\,.$$ □

(16.110) THEOREM: Let V be a Hausdorff manifold and let $\mathcal{K}(V)$ be the vector space of continuous functions on V with compact support (16.42). If we define

$$\diamondsuit \qquad \|f\| = \sup_{x \in V} |f(x)| \qquad \forall f \in \mathcal{K}(V)\,,$$

[321] One abbreviates this as $\sum_{j=1}^{\infty} \|X_j\| < \infty \Rightarrow \sum_{j=1}^{\infty} X_j = X$. Note that $\|\sum_{j=1}^{\infty} X_j\| \leq \sum_{j=1}^{\infty} \|X_j\|$.

then $\mathcal{K}(V)$ becomes a *normed* vector space.

PROOF: The fact that equation $\diamondsuit$ defines a finite number $\|f\|$ follows from theorem (16.24). The verification of the axioms (16.105$\diamondsuit$) then is trivial. □

(16.111) It follows immediately from (16.51) that the *topological dual space* of the *normed vector space* $\mathcal{K}(V)$ is the set of *bounded measures*. Thus $[\mathcal{K}(V)]'$ is a Banach space (16.107). □

(16.112) THEOREM: Every measure μ with compact support is bounded.

PROOF: Let K be the support of μ and let f be one of those functions whose existence is assured by theorem (16.48). It follows that $\mu = \mu \times f$. Because of theorem (16.133) below, μ is bounded. □

(16.113) In particular, *on a compact manifold all measures are bounded*. On the contrary, the Lebesgue measure of $\mathbf{R}^n$ is not bounded.

(16.114) THEOREM:

a) If μ and μ' are bounded measures on V and V', respectively, then $\mu \otimes \mu'$ is a bounded measure on $V \times V'$ and

$$\|\mu \otimes \mu'\| = \|\mu\| \times \|\mu'\| .$$

b) $$[\mu \otimes \mu' \neq 0,\ \mu \otimes \mu' \text{ bounded}] \implies [\mu \text{ and } \mu' \text{ bounded}] .$$

PROOF: The result follows directly from (16.59) and (16.64). □

(16.115) THEOREM: For a measure μ,

$$[\ \mu \text{ is bounded}] \iff [|\mu| \text{ is bounded}] \implies [\|\mu\| = \| \, |\mu| \, \|] .$$

PROOF: Applying (16.54) gives $|\mu(f)| \leq |\mu|(|f|)$, hence $\|\mu\| \leq \| \, |\mu| \, \|$. Applying (16.53) gives $|\mu|(f) \leq |\mu|(|f|) = \sup_{|g| \leq |f|} \mu(g)$, hence $\| \, |\mu| \, \| \leq \|\mu\|$. □

(16.116) THEOREM: If μ and ν are bounded measures, then

$$[\mu \geq 0,\ \nu \geq 0] \implies \|\mu + \nu\| = \|\mu\| + \|\nu\| .$$

PROOF: We know that $\|\mu + \nu\| \leq \|\mu\| + \|\nu\|$ (16.105$\diamondsuit$). On the other hand, for any $\varepsilon > 0$:

$$\exists f, g \in \mathcal{K}(V) :\ \|f\| \leq 1 ,\ \|g\| \leq 1 ,\ \mu(f) \geq \|\mu\| - \varepsilon , \nu(g) \geq \|\nu\| - \varepsilon .$$

If $h = \sup(|f|, |g|)$, we have $\|h\| \leq 1$, $[\mu + \nu](h) \geq \|\mu\| + \|\nu\| - 2\varepsilon$, and hence $\|\mu + \nu\| \geq \|\mu\| + \|\nu\| - 2\varepsilon$. □

Using (16.33) one can deduce from (16.116) that for every bounded measure μ:

$$\|\mu\| = \|\mu^+\| + \|\mu^-\| . \tag{16.117}$$

(16.118) Gathering the results (16.38), (16.107) and (16.115) together we see that in the Riesz space of measures on a manifold V, the set of bounded measures form a *Riesz subspace*. □

If μ and ν are bounded positive measures, it follows immediately from (16.116) that the number $\|\mu\| - \|\nu\|$ only depends on the difference $\mu - \nu$. Hence, the *mass* of a bounded measure μ, defined by

$$\text{mass}(\mu) = \|\mu^+\| - \|\mu^-\| \tag{16.119}$$

satisfies

$$\mu \geq 0,\ \nu \geq 0,\ \mu \text{ and } \nu \text{ bounded} \implies \|\mu\| - \|\nu\| = \text{mass}(\mu - \nu) . \tag{16.120}$$

From this formula and (16.116) one can easily deduce that the map "mass" is linear:

$$\begin{cases} \text{mass}(\mu + \nu) \equiv \text{mass}(\mu) + \text{mass}(\nu) \\ \text{mass}(s\,\mu) \equiv s\ \text{mass}(\mu) \qquad\qquad \forall\, s \in \mathbf{R} . \end{cases} \tag{16.121}$$

Note also the formulæ

$$\begin{gathered} |\ \text{mass}(\mu)\ | \leq \|\mu\| \\ \text{mass}(\mu) = \|\mu\| \iff \mu \geq 0 . \end{gathered} \tag{16.122}$$

(16.123) THEOREM: If μ and μ' are bounded measures on manifolds V and V', respectively, then

$$\text{mass}(\mu \otimes \mu') = \text{mass}(\mu) \times \text{mass}(\mu') .$$

PROOF: It suffices to note that $\mu \otimes \mu'$ is the difference of the measures

$$\mu^+ \otimes \mu'^+ + \mu^- \otimes \mu'^- \qquad \text{and} \qquad \mu^+ \otimes \mu'^- + \mu^- \otimes \mu'^+ ,$$

which are positive (16.60) and bounded (16.114), and then to apply (16.114) and (16.116). □

(16.124) DEFINITION: Let Ω be an open subset of V and let μ be a measure on V. Ω is called *μ-measurable* if the measure μ_Ω (definition (16.100)) is

bounded. The number $\text{mass}(\mu_\Omega)$ is called the *μ-measure* of Ω and is in general denoted by $\mu(\Omega)$. □

One can easily verify that Ω is μ-measurable whenever μ is bounded or whenever Ω is relatively compact. Moreover,

(16.125) $$\Omega \cap \Omega' = \emptyset \implies \mu(\Omega \cup \Omega') = \mu(\Omega) + \mu(\Omega') .$$

If $\mu \geq 0$ we also have

(16.126) $$\begin{cases} \mu(\Omega) \geq 0 \\ \Omega \subseteq \Omega' \end{cases} \implies \mu(\Omega) \leq \mu(\Omega') .$$

(16.127) EXAMPLE: In $\mathbf{R}^n$, a *ball* B of radius r ($r > 0$): $x \in B \Leftrightarrow \|x\| < r$ is a relatively compact open set. If μ is the Lebesgue measure on $\mathbf{R}^n$, then the number $\mu(B)$ is called the *volume* of B. One can show that it is given by [322]

$$\mu(B) = \frac{\pi^{n/2}}{[n/2]!} r^n .$$ □

EXAMPLE: A *sphere* S of radius r: $x \in S \Leftrightarrow \|x\| = r$ is a compact submanifold of $\mathbf{R}^n$. The Euclidean density (16.94) defines a measure μ on S. The number $\mu(S) = \|\mu\|$ is called the *area* of the sphere. A calculation shows that

$$\|\mu\| = \frac{n\pi^{n/2}}{[n/2]!} r^{n-1} .$$ □

(16.128) The idea of a measure can be extended to a class of sets much larger than that of open sets [323] but we will not use this extension. □

(16.129) Let V be a pure Hausdorff manifold. We define $L^1(V)$ to be the set of bounded measures μ on V satisfying: for every $\varepsilon > 0$ there exists a completely continuous measure μ_0 on V such that $\|\mu - \mu_0\| < \varepsilon$. □

We leave the verification of the following theorems as an exercise for the reader.

(16.130) THEOREM: $L^1(V)$ is a Banach space.

[322] We define $x! = \Gamma(1 + x) = \int_0^\infty e^{-t} t^x \, dt$. For odd n, the quantity $[n/2]!$ is calculated by induction starting from $[-\frac{1}{2}]! = \sqrt{\pi}$ using the formula $x! = x([x-1]!)$.

[323] See Bourbaki [9].

(16.131) THEOREM: $L^1(V)$ is a Riesz subspace of the space of bounded measures on V.[324]

(16.132) THEOREM: $L^1(V) \otimes L^1(V') \subseteq L^1(V \otimes V')$.[325]

Integrable functions

(16.133) LEMMA: If μ is a measure on V and if $f \in \mathcal{K}(V)$, then $\mu \times f$ (notation (16.56)) is a *bounded* measure and

$$\diamondsuit \qquad \text{mass}(\mu \times f) = \mu(f)\,.$$

PROOF: If μ and f are positive, it is easy to verify that $\mu \times f$ is positive and that[326]

(16.134) $$\|\mu \times f\| = \mu(f)\,.$$

One obtains $\diamondsuit$ by replacing μ and f in (16.134) by differences of positive ones, and then applying the properties of linearity. □

(16.135) DEFINITION: Let μ be a measure on V and let f be a continuous function on V. f is called *μ-integrable* if the measure $\mu \times f$ is bounded.

(16.135) COROLLARY: The linear map μ defined on $\mathcal{K}(V)$ can be *extended linearly* to the vector space of continuous μ-integrable functions by *defining*

$$\mu(f) = \text{mass}(\mu \times f)\,. \qquad \square$$

Of course, in the case of a measure defined by a density φ (16.83), we will still use the notation

(16.136) $$\int_V f(x)\,\varphi(dx)$$

to denote the number $\mu(f)$, even if the support of f is not compact. In that case the integral is said to be *convergent*.

(16.137) EXAMPLE: Let V be a manifold of dimension 0 and let I be the measure defined in (16.70). We know that every function f defined on V is continuous; f is said to be *summable* if f is I-integrable. In that case the

[324] One should first establish the identity $\| \, |\mu| - |\mu_0| \, \| \leq \|\mu - \mu_0\|$, which is a consequence of (16.36) and (16.54).

[325] Abuse of notation. It signifies: $\mu \in L^1(V)$, $\mu' \in L^1(V') \Rightarrow \mu \otimes \mu' \in L^1(V \times V')$. Use (16.90) and (16.114).

[326] Use (16.48).

number $I(f)$ is written as $\sum_{x\in V} f(x)$, but this is in general an "infinite sum".

In the particular case where V is the set of positive integers, one can show that f is summable if the *series*

$$f(1) + f(2) + \cdots + f(n) + \cdots$$

is *absolutely convergent*; and then its sum equals $I(f)$. □

(16.138) EXAMPLE: Let f be a continuous function on a bounded open interval $a < x < b$ of $\mathbf{R}$. If f is integrable with respect to the Lebesgue density, its integral is denoted by $\int_a^b f(x)\,dx$. One can show that this integral is the limit of $F(\beta) - F(\alpha)$ (where F is a primitive of f) when β tends to b and α tends to a. In this way

$$\int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} = \pi\,.$$

If f is not positive, the limit of $F(\beta) - F(\alpha)$ *might exist without* f *being integrable*. □

(16.139) THEOREM: If the measures μ, μ' and the continuous functions f, f' satisfy the conditions

$$|\mu| \le \mu'\,, \qquad |f| \le f'\,, \qquad \text{and} \qquad f' \text{ is } \mu'\text{-integrable}\,,$$

then f is μ-integrable and $|\mu(f)| \le \mu'(f')$.

PROOF: For $g \in \mathcal{K}(V)$ with $|g| \le 1$ we have $|f \times g| \le f'$. Hence, using (16.54),

$$\big|[\mu \times f](g)\big| = \big|\mu(f \times g)\big| \le \mu'(f')\,.$$

Hence $\mu \times f$ is bounded and $\|\mu \times f\| \le \mu'(f')$. It now suffices to apply the inequality (16.122) to $\mu \times f$. □

(16.140) THEOREM: If f is μ-integrable, then $|f|$ is $|\mu|$-integrable.

PROOF: We have [f is μ-integrable] $\Leftrightarrow$ [$\mu \times f$ is bounded] $\Rightarrow$ [$|\mu \times f|$ is bounded] (16.115) $\Leftrightarrow$ [$|\mu| \times |f|$ is bounded] (16.58). □

(16.141) COROLLARY: [$\mu \ge 0$, $f \in \mathcal{C}(V)$, $f \ge 0$, f is μ integrable] $\Rightarrow$[$\mu(f) \ge 0$]. □

(16.142) By combining theorems (16.139) and (16.140) one can see that the four hypotheses: [f is μ-integrable], [$|f|$ is μ-integrable], [f is $|\mu|$-integrable], and [$|f|$ is $|\mu|$-integrable] are equivalent. They imply that

$\diamondsuit$ $$\mu(f) = \mu(f^+) - \mu(f^-)\,.$$ □

(16.143) THEOREM: Let μ and μ' be measures on manifolds V and V', respectively, and let $f \in \mathcal{C}(V)$ and $f' \in \mathcal{C}(V')$. Then

$$[f \text{ is } \mu\text{-integrable and } f' \text{ is } \mu'\text{-integrable}] \implies$$
$$\implies \left[\begin{array}{l} f \otimes f' \text{ is } \mu \otimes \mu'\text{-integrable} \\ [\mu \otimes \mu'](f \otimes f') = \mu(f) \times \mu'(f') \,. \end{array} \right]$$

PROOF: A simple corollary of (16.67) and (16.114). □

(16.144) Let F be a *continuous map from V to a finite dimensional vector space E*. F is *μ-integrable* [327] if the map $x \mapsto C(F(x))$ is μ-integrable for every element C of the dual space of E. In that case there exists an element of E, denoted by $\mu(F)$, or by $\int_V F(x)\,\varphi(dx)$ in the case (16.83), and defined by

(16.145) $$C(\mu(F)) = \mu(C \circ F) \qquad \forall\, C \in E^* \,.$$

To establish this result, it suffices to choose a basis S of E. If we define

(16.146) $$F(x) \equiv S \begin{pmatrix} {}^1f(x) \\ \vdots \\ {}^nf(x) \end{pmatrix},$$

one can verify immediately that

$$[F \text{ is } \mu\text{-integrable}] \iff [{}^jf \text{ is } \mu\text{-integrable } \forall\, j = 1, \ldots, n]$$

(16.147) $$\implies \left[\mu(F) = S \begin{pmatrix} \mu({}^1f) \\ \vdots \\ \mu({}^nf) \end{pmatrix} \right]. \qquad \square$$

(16.148) THEOREM: Let V be a Hausdorff manifold and let μ be a measure on V. If f and g are continuous functions [328] on V such that

$$\begin{cases} f \text{ is } \mu\text{-integrable}\,, \\ f(x) = g(x) \qquad\qquad \forall\, x \in \text{support}(\mu)\,, \end{cases}$$

then g is μ-integrable and $\mu(g) = \mu(f)$.

[327] The reader can find in Bourbaki an extension of this notion to the case where E is a Banach space.

[328] The theorem is proved under the assumption that f and g are real valued. It can be extended immediately to the case where they have values in a finite dimensional vector space.

PROOF: The continuous function $h = g - f$ is zero on the support of μ. Thus $\forall k \in \mathcal{K}(V)$ the function $h \times k$, which has compact support, is also zero on the support of μ. Hence, by (16.104) $\mu(h \times k) = 0$. Since k is arbitrary, we have $\mu \times h = 0$. Therefore $\mu \times g = \mu \times f$. Hence $\|\mu \times g\| = \|\mu \times f\| < \infty$, and thus $\mu(g) = \text{mass}(\mu \times g) = \text{mass}(\mu \times f) = \mu(f)$. □

The image of a measure

(16.149) DEFINITION: Let V and V^* be two Hausdorff manifolds and let A be a continuous map from an open subset of V to V^*.

a) We will say that the measure μ is *A-transportable* [329] if

$$\heartsuit \qquad [f^* \in \mathcal{K}(V^*)] \implies [f^* \circ A \text{ is } \mu_{\text{dom}(A)}\text{-integrable}] .$$

(16.149) THEOREM:

b) In the Riesz space of measures on V, those which are A-transportable form a Riesz subspace.

c) If μ is A-transportable, there exists a measure $A^+(\mu)$ on V, called the *image of* μ under A, which is defined by

$$\diamondsuit \qquad A^+(\mu)(f^*) = \mu_{\text{dom}(A)}(f^* \circ A) \qquad [\forall f^* \in \mathcal{K}(V^*)] .$$

d) The map A^+ is *linear* and *increasing*.[330]

PROOF: The verification of (b) and of the linearity of A^+ is elementary, taking into account (16.101) and (16.57). If $\mu \geq 0$, it follows that $A^+(\mu)$ is a positive linear form on $\mathcal{K}(V^*)$ and thus is a measure. For every A-transportable measure μ we have $A^+(\mu) = A^+(\mu^+) - A^+(\mu^-)$, and thus $A^+(\mu)$ is a measure. The fact that A^+ is increasing is then immediate. □

(16.150) THEOREM: Under the hypotheses of (16.149) let us assume that μ is A-transportable and that $\mu \geq 0$. Then formula (16.149$\diamondsuit$)

$$\diamondsuit \qquad A^+(\mu)(f^*) = \mu_{\text{dom}(A)}(f^* \circ A)$$

remains valid for every $f^* \in \mathcal{C}(V^*)$ for which either the left or right hand side of $\diamondsuit$ exists.[331]

---

[329] Synonym: *A is μ-proper.*

[330] From the Riesz space of A-transportable measures on V to the Riesz space of measures on V^*.

[331] That is, if f^* is $A^+(\mu)$-integrable.

PROOF: (a) Let us first suppose that $f^* \geq 0$. Taking definitions (16.135) and (16.122) into account, the theorem to be proved can be written as $X = Y$, with

$$X = \sup_h \mu_{\mathrm{dom}(A)}([f^* \circ A] \times h)\,,$$
$$Y = \sup_{g^*,h} \mu_{\mathrm{dom}(A)}([[f^* \times g^*] \circ A] \times h)\,,$$

where the functions g^* and h range over the sets defined by

$$[g^* \in \mathcal{K}(V),\ 0 \leq g^* \leq 1] \qquad \text{and} \qquad [h \in \mathcal{K}(\mathrm{dom}(A)),\ 0 \leq h \leq 1]\,.$$

It is obvious that $Y \leq X$ because $[[f^* \times g^*] \circ A] \times h \leq [f^* \circ A] \times h$. On the other hand, for every h let us denote by K the support of h (which is a compact subset of $\mathrm{dom}(A)$) and by K^* the image of K under A, which is a compact subset of V^* because A is continuous (theorem (16.22)). According to Urysohn's lemma, there exists a function g^* satisfying $g^* \in \mathcal{K}(V^*)$ and $0 \leq g^* \leq 1$ such that $[x^* \in K^*] \Rightarrow [g^*(x^*) = 1]$. For $x \in \mathrm{dom}(A)$ we have

$$[h(x) \neq 0] \implies [x \in K] \implies [A(x) \in K^*] \implies [g^*(A(x)) = 1]\,,$$

and thus $h = [g^* \circ A] \times h$. Therefore $[f^* \circ A] \times h = [[f^* \times g^*] \circ A] \times h$, and thus $X \leq Y$.

(b) Let us now remove the hypothesis $f^* \geq 0$. We note that

$$[f^*]^+ \circ A = [f^* \circ A]^+ \qquad \text{and} \qquad [f^*]^- \circ A = [f^* \circ A]^- \qquad \text{(see (16.44))}.$$

It follows that the left and right hand sides of $\diamondsuit$, if they exist, are equal to (theorem (16.142$\diamondsuit$))

$$A^+(\mu)([f^*]^+) - A^+(\mu)([f^*]^-)$$

and

$$\mu_{\mathrm{dom}(A)}([f^*]^+ \circ A) - \mu_{\mathrm{dom}(A)}([f^*]^- \circ A)\,,$$

respectively. Thus they are equal since the result established in a) is valid for the positive functions $[f^*]^+$ and $[f^*]^-$. □

(16.151) THEOREM: If $A^+(\mu)$ and $B^+(A^+(\mu))$ exist, then

$$[B \circ A]^+(\mu) = B^+(A^+(\mu))\,.$$

PROOF: In case $\mu \geq 0$, it suffices to expand the expressions $[B \circ A]^+(\mu)(f)$ and $B^+(A^+(\mu))(f)$ according to formula (16.150$\diamondsuit$). One gets the general

case by considering μ as the difference of two positive measures μ^+ and μ^-, to which the formula is applicable, and applying (16.149). □

(16.152) THEOREM: Under the hypotheses of (16.149), if μ is A-transportable, the support of $A^+(\mu)$ is contained in the closure of the image under A of the support of μ.

PROOF: It suffices to verify that, if Ω is an open subset of V^* which does not intersect the image of the support of μ under A, then $[A^+(\mu)]_\Omega = 0$. Now, using theorem (16.151), we find $[A^+(\mu)]_\Omega = [1_\Omega]^+(A^+(\mu)) = C^+(\mu)$, with $C = 1_\Omega \circ A$. The domain of definition of C is an open subset of V which does not intersect the support of μ, and thus $\mu_{\mathrm{dom}(C)} = 0$ (16.102). Hence, by theorem (16.149c), we have $C^+(\mu) = 0$. □

(16.153) THEOREM: Under the hypotheses of (16.149), every bounded measure μ is A-transportable and $\|A^+(\mu)\| \leq \|\mu\|$.

PROOF: For $f^* \in \mathcal{K}(V^*)$ and $g \in \mathcal{K}(\mathrm{dom}(A))$ one can verify that

$$\|\hat{k}\| \leq \|f^*\| \times \|g\|\,,$$

with $k = [f^* \circ A] \times g$ and where the hat denotes the extension by zero in the complement of $\mathrm{dom}(A)$ (16.100◇). Hence $\| \mu_{\mathrm{dom}(A)} \times [f^* \circ A] \| \leq \|\mu\| \times \|f^*\|$, which immediately leads to the desired result. □

(16.154) We see that A^+ defines an *increasing bounded linear* map from the Banach-Riesz space of bounded measures on V to the corresponding space on V^*.

(16.155) THEOREM: If A is a homeomorphism from V to V^* and if μ is a bounded measure on V, then $\|A^+(\mu)\| = \|\mu\|$.

PROOF: Let μ^* be $A^+(\mu)$. Theorem (16.151) shows that $\mu = [A^{-1}]^+(\mu^*)$. Thus we have $\|\mu^*\| \leq \|\mu\|$ and $\|\mu\| \leq \|\mu^*\|$. □

(15.156) THEOREM: Under the hypotheses of (16.149), if $\mathrm{dom}(A) = V$ and if μ is bounded, then $\mathrm{mass}(A^+(\mu)) = \mathrm{mass}(\mu)$.

PROOF: (a) Let us first assume that $\mu \geq 0$. If we define $f^*(x^*) = 1$ $[\forall x^* \in V^*]$, then $f = f^* \circ A$ satisfies $f(x) = 1$ $[\forall x \in V]$. Since the mass of a bounded measure is equal to the value of this measure on the constant function 1 (definition (16.135)), we have $\mathrm{mass}(\mu) = \mu(f)$ and $\mathrm{mass}(A^+(\mu)) = A^+(\mu)(f^*)$. These two numbers are equal according to (16.150).

(b) In the general case, it suffices to apply the result of (a) to the

positive measures μ^+ and μ^- and to use the linearity of the maps A^+ and "mass". □

(16.157) THEOREM: Let V and V^* be two Hausdorff manifolds and let A be a proper map[332] from V to V^*. Then every measure μ on V is A-transportable.

PROOF: (a) Let us assume that $\mu \geq 0$ and let us define $\Omega = \mathrm{dom}(A)$. For $f^* \in \mathcal{K}(V^*)$ the set of $x^* \in V^*$ such that $f^*(x^*) \neq 0$ is an open set (being the inverse image under f^* of an open subset of $\mathbf{R}$) and is relatively compact (because it is contained in the compact support of f^*). Thus its inverse image under A is an open relatively compact set Ω' (since A is proper). Note that $x \in \Omega' \Leftrightarrow f(x) \neq 0$, with $f = f^* \circ A$.

Let $\overline{\Omega'}$ be the (compact) closure of Ω'. Due to Urysohn's lemma (16.48)

$$\exists\, h \in \mathcal{K}(V):\ h \geq 0\,, \qquad \text{and} \qquad [x \in \overline{\Omega'}] \Rightarrow [h(x) = 1]\,.$$

For $x \in \Omega$ we have either $f(x) = 0$ or $x \in \Omega'$, whence $f(x) = f(x)\,h(x)$.

If we take $g \in \mathcal{K}(\Omega)$ with $\|g\| \leq 1$, then $\forall\, x \in \Omega$

$$\Big|[f \times g](x)\Big| = \Big|f(x)\,h(x)\,g(x)\Big| = \Big|f^*(A(x))\,h(x)\,g(x)\Big| \leq \|f^*\|\,h(x)\,.$$

This inequality is also valid for the extension $\widehat{f \times g}$, which is 0 outside Ω (16.100), just because $h \geq 0$. Thus $|\widehat{f \times g}| \leq \|f^*\|\,h$, whence

$$\Big|\,[\mu_\Omega \times f]\,(g)\Big| = \Big|\mu_\Omega(f \times g)\Big| = \Big|\mu(\widehat{f \times g})\Big| \leq \mu\big(\big|\widehat{f \times g}\big|\big) \leq \|f^*\|\,\mu(h)\,,$$

which shows that $\mu_\Omega \times f$ is *bounded*, that is, that f is μ_Ω-integrable. Thus we have $\mu_{\mathrm{dom}(A)}(f^* \circ A) = \mu_\Omega(f) = \mathrm{mass}(\mu_\Omega \times f)$ (definition (16.135)). Hence formula (16.149◇) defines an operator $A^+(\mu)$ on $\mathcal{K}(V^*)$ which is obviously linear. Moreover, if $f^* \geq 0$, $A^+(\mu)(f^*) \geq 0$, and thus $A^+(\mu)$ is a positive measure on V.

(b) The general case is obtained by writing $\mu = \mu^+ - \mu^-$, applying (a) to the positive measures μ^+ and μ^-, and using (16.149). □

(16.158) We thus see that if A is proper from V to V^*, A^+ defines an increasing linear map from the Riesz space of measures on V to those on V^*. This map extends the one defined for *bounded* measures, which we examined in (16.154).

(16.159) In particular, if A is a homeomorphism from V to V^*, then A is proper and $A^+(\mu)$ exists for every measure μ on V. In this way, we recover the definition given in (16.77).

[332] Definition (16.26). Recall that $\mathrm{dom}(A)$ is open and that A is continuous.

Examples

CONVOLUTION ON A LIE GROUP

(16.160) Let G be a Lie group and let μ and μ' be bounded measures on G. The *convolution of* μ and μ' is the image of the measure $\mu \otimes \mu'$ under the map $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x \times y$; it is a bounded measure, denoted[333] by $\mu \times \mu'$. One can easily verify that the *convolution operation* $\times$ is an associative and bilinear operation[334, 335] on the space B of bounded measures on G. And thus B is an *algebra.*†

Moreover, we know that B is a Banach space (16.111) and that convolution satisfies[336]

(16.161) $$\|\mu \times \nu\| \leq \|\mu\| \times \|\nu\| .$$

Hence B is a *Banach algebra.*†

For $\mu \in B$ let μ^* be the measure which is the image of μ under the map $[x \mapsto x^{-1}]$. Obviously μ^* belongs to B, but we also have (see (16.154)):

(16.162) $$[\mu + \mu']^* \equiv \mu^* + \mu'^*$$

(16.163) $$[s\,\mu]^* \equiv s\,[\mu^*] \qquad \forall\, s \in \mathbf{R} .$$

Since $x \mapsto x^{-1}$ is a diffeomorphism from G to G, we have[337]

(16.164) $$\|\mu^*\| = \|\mu\| .$$

Moreover the identities $[x^{-1}]^{-1} \equiv x$ and $[x \times y]^{-1} \equiv y^{-1} \times x^{-1}$ in G lead to the formulæ

(16.165) $$[\mu^*]^* \equiv \mu$$

(16.166) $$[\mu \times \nu]^* \equiv \nu^* \times \mu^* .$$

The formulæ (16.162) to (16.166) can be summarized by saying that the Banach algebra B has *an involution.*† One can show that $L^1(G)$ is a

[333] In the case of an additive group, one usually uses the notation $\mu * \mu'$.

[334] The verification shows that $\mu \times \mu' \times \mu''$ is the image of $\mu \otimes \mu' \otimes \mu''$ under the map $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x \times y \times z$.

[335] If G is abelian, convolution is commutative on G.

†Editors' note: An algebra is a vector space with a multiplication that is associative and bilinear.

[336] Apply (16.114) and (16.153).

†Editors' note: A Banach space with an associative and bilinear multiplication is a Banach algebra if the multiplication satisfies (16.161).

[337] Theorem (16.155).

†Editors' note: An algebra having an involution is also called a $*$-*algebra*.

Banach subalgebra of B which obviously has an involution. $L^1(G)$ is usually called the *group algebra* of G.

Let us finally note that the formula

(16.167) $$\text{mass}(\mu \times \nu) = \text{mass}(\mu) \cdot \text{mass}(\nu)\,,$$

is valid in B[338] and also on its subalgebra $L^1(G)$. □

PHOTOMETRY

(16.168) The goal of classical photometry is to determine the *flux of light* crossing a given surface (in either direction), when this surface is illuminated by various light sources. The flux is a number which is measured, for example, in *lumens*.

Obviously the theory has to indicate how the flux depends on the surface. It turns out that one can formulate this dependence in the following fashion.

(16.169) A *photometric state* is a *positive measure* μ on the manifold V of light rays.[339]

If Ω is the set of light rays which cross a given surface in a given direction,[340] the corresponding light flux is equal to $\mu(\Omega)$.[341] □

Without completely developing the scientific basis of lighting, let us outline some of the uses of the axioms (16.169).

(16.170) Let S be a bounded light source. The set Ω of light rays emitted by this source that cross a sphere surrounding S does not depend on the choice of the sphere. The number $\mu(\Omega)$ is the *flux* of the source. It is thus measured in lumens.

(16.171) If we assume that the measure μ is completely continuous, then the corresponding density is proportional to the *symplectic density on* V.[342] The factor of proportionality will be denoted by $\lambda(\xi)/\pi$ (where ξ is a variable light ray). The number $\lambda(\xi)$ is called the *brilliance* (or *luminance*). It is measured in lumens $\otimes$ meter^{-2}, called *blondels* in this context.

(16.172) A photometric state is said to be *uniform* if the luminance is constant. One often considers states *induced on an open set by a uniform state*

[338] Use (16.123) and (16.156).

[339] Definition (15.87). The dimension of V is 4.

[340] Recall that a light ray is an oriented line (15.87).

[341] Definition (16.124). $\mu(\Omega)$ equals $\|\mu_\Omega\|$ because μ is positive (16.122), and thus $\mu(\Omega) \geq 0$.

[342] Definition (16.94). We choose for V the symplectic structure defined by (15.90).

(16.100). For example, the sky observed through a skylight or the radiation of heat in the interior of an isothermic furnace (see (17.151) below).

(16.173) Let us now consider a screen, whose surface is a submanifold of dimension 2 in $\mathbf{R}^3$. If we denote by $\mathbf{r}$ the point where a light ray ξ pierces this screen (properly oriented), one can show that the map $\xi \mapsto \mathbf{r}$ is *proper*. The image of a photometric state under this map is exactly the *image* projected on the screen (in the photographic sense of the term). If this image is completely continuous, the corresponding density can be written as $E(\mathbf{r})\,\varphi$, where φ denotes the *Euclidean density* of the tangent plane.[343] The number $E(\mathbf{r})$ is called the *illumination* of the screen at the point $\mathbf{r}$; it is measured in lumens $\otimes$ meter^{-2} (this time called *lux*). If the photometric state is completely continuous, the illumination is given by

(16.174) $$E(\mathbf{r}) = \frac{1}{\pi}\int_H \lambda(\xi)\,\cos I\,\varphi'(d\mathbf{u})\,,$$

where H is the hemisphere which describes the unit vectors $\mathbf{u}$ (notation (15.86)) corresponding to the light rays ξ that arrive on the screen at $\mathbf{r}$, where φ' is the Euclidean density on the tangent space to the sphere,[344] and where I is the *angle of incidence* ($\cos I = |\langle \mathbf{n}, \mathbf{u}\rangle|$, $\mathbf{n}$ being the unit normal vector to the screen).

(16.175) If the photometric state is uniform, it follows that the illumination is constant and *equal to the luminance* (due to the factor $1/\pi$ in definition (16.171)).

(16.176) An illuminated screen usually behaves as a light source. If it is *perfectly diffusing*, the brilliance of a ray emitted from a point $\mathbf{r}$ of the screen equals the illumination $E(\mathbf{r})$. Thus a perfectly diffusing sphere placed in uniform lighting is invisible.

(16.177) If one observes light sources with photometric state μ through an optical instrument and if one can ignore absorption and partial reflections inside the instrument, one generally assumes that the energy of a bundle of light rays (an open subset of V) which passes through the instrument is conserved. It follows that the photometric state of the rays that did pass through the instrument is equal to $S^+(\mu)$[345], where S denotes the local symplectomorphism of scattering (see (15.87) to (15.90)). Since S conserves the symplectic density (see (16.99)), it follows that luminance

[343] Definition (16.94). $\varphi(d\mathbf{r})$ is called the "element of area".

[344] $\varphi'(d\mathbf{u})$ is called the "element of solid angle".

[345] Notation (16.149). In the usual examples the map S is proper.

is conserved for rays that pass through the instrument.[346] □

Random variables

(16.178) DEFINITION: Let V be a Hausdorff manifold. A positive bounded *measure* on V of mass 1 is called a *probability measure* on V. □

Let μ be a probability measure on V and let Ω be an open subset of V. The number

(16.179) $$p = \mu(\Omega)$$ [347]

is called the *probability* of the set Ω. One can easily verify that

(16.180) $$0 \leq p \leq 1 .$$

The formulæ (16.125) and (16.126) give some of the usual properties of the concept of probability. However, establishing a complete correspondence between the axioms for the theory of probability (due to Kolmogorov) and definition (16.179) necessitates an extension of the concept of measure to non-open subsets, an extension we did not describe (see (16.128)). Therefore we will be content to work directly with definition (16.178).



(16.181) EXAMPLE: Let a be a point of a manifold V. The Dirac measure δ_a is obviously a probability measure on V. □

(16.182) EXAMPLE: Let μ be a probability measure defined on a manifold V *of dimension* 0. For $j \in V$, the set whose sole element is j is open (16.7); let us denote its probability by p_j. One can verify immediately that (notation (16.137))

♢ $$\sum_{j \in V} p_j = 1$$

and that for every $g \in \mathcal{K}(V)$

♡ $$\mu(g) = \sum_{j \in V} p_j \, g(j) .$$

Hence the probability measure is entirely defined by the "elementary probabilities" p_j. □

(16.183) DEFINITION: Let x be a variable taking values in a manifold V. The variable x is called a *random variable* if a probability measure on V is associated to x (called the "probability measure of x").

[346] But not the illumination of the retina, if the open set Ω of rays passing through the instrument does not contain the open set Ω' of rays which pass through the pupil.

[347] Definition (16.124). Every open set is μ-measurable, because μ is bounded.

(16.184) THEOREM: Let V and V^* be two Hausdorff manifolds, let A be a continuous map from V to V^*, and let μ be a probability measure on V. Then the image $A^+(\mu)$ is a probability measure on V^*.

PROOF: We know that $A^+(\mu)$ is a measure (16.153), which is positive (16.149) and of mass 1 (16.156). □

(16.185) CONVENTION: If x is a random variable having μ as its probability measure, we will consider $A(x)$ also as a random variable, having $A^+(\mu)$ as its probability measure. Theorem (16.151) insures the consistency of this convention.[348]

(16.186) EXAMPLE: Let $X \equiv \begin{pmatrix} x \\ x' \end{pmatrix}$ be a random variable taking values in the direct product $V \times V'$ of two manifolds. Since the maps $X \mapsto x$ and $X \mapsto x'$ are continuous, it follows that x and x' *are random variables.* □

(16.187) THEOREM: If μ and μ' are probability measures on the manifolds V and V' respectively, then the tensor product $\mu \otimes \mu'$ is a probability measure on $V \times V'$.

PROOF: Use (16.60) and (16.123). □

(16.188) In the notation of (16.186) x and x' are called *independent* random variables if the probability measure of $\begin{pmatrix} x \\ x' \end{pmatrix}$ is the tensor product of two probability measures μ and μ'. In that case one can verify[349] that μ and μ' are probability measures of x and x' respectively.

(16.189) Consider now in particular a Lie group G. If the variable $\begin{pmatrix} x \\ x' \end{pmatrix}$ ranges over G^2, the random variable $x \times x'$ ranges over G. If x and x' are independent, it is immediate that the probability measure of $x \times x'$ is the convolution (16.160) of the probability measures of x and x'. This applies in particular to the case when G is a vector space, seen as an additive group ("addition of random variables").

[348] Let x be a random variable. If there exist continuous maps $x \mapsto y$ and $y \mapsto z$, then defining z as a random variable by means of the map $x \mapsto z$ or through the intermediary of y is the same.

[349] An excellent exercise!

Average values

(16.190) DEFINITION: Let x be a random variable taking values in a finite dimensional vector space E and let μ be its probability measure. The *mean value* of x denoted by average(x) is the element of E such that

$$\text{average}(x) = \mu(1_E)\,,$$

which is defined if 1_E is μ-integrable.[350]

(16.191) THEOREM: Let x be a random variable taking values in a manifold V, let μ be its probability measure, and let F be a continuous map from V to a finite dimensional vector space E. Then average$(F(x)) = \mu(F)$, provided F is μ-integrable.

PROOF: In fact $F^+(\mu)(1_E) = \mu(1_E \circ F) = \mu(F)$. □

(16.192) EXAMPLE: Let x be a random variable taking values in a vector space E. The *p^{th}-order moment* of x is the tensor (if it exists)

$$\text{average}\,(\underbrace{x \otimes x \cdots \otimes x}_{p})\,.^{351}$$

(16.193) DEFINITION: Let E be a finite dimensional vector space and let Q be a subset of E. Q is called *convex* if

$$[x \in Q\,,\ y \in Q\,,\ 0 \leq s \leq 1] \quad \Longrightarrow \quad [sx + [1-s]y \in Q]\,.$$

(16.193) THEOREM: For a finite dimensional vector space E we have the following results.

◇ Every intersection of convex subsets of E is convex.

Let H be a subset of E and let Q be the *intersection of all closed half spaces* containing H:

♡ $$[x \in Q] \quad \Longleftrightarrow \quad [\forall\, \Gamma \in E^* :\ \Gamma(x) \leq \sup_{y \in H} \Gamma(y)]\,.$$

Then Q is *the smallest closed convex set* containing H.

[350] Definition (16.144). 1_E is the map $[x \mapsto x]$, not $[x \mapsto 1]$.

[351] It is convenient to consider the number 1 as the moment of order 0 and the average of x as the moment of order 1. For the definition of the tensor product $x \otimes x \cdots \otimes x$, see for example Souriau [31], §13.

PROOF: The verification of $\diamondsuit$ is trivial as is the fact that the set Q, defined by $\heartsuit$, contains H and is closed and convex. To show that Q is the smallest, the reader is referred to Bourbaki[352] who treats the case of a Banach space. □

(16.194) COROLLARY: (Notation (16.190)) If the support of x[353] is contained in a closed convex set Q, then

$$\text{average}(x) \in Q \qquad \text{(provided average}(x)\text{ exists).}$$

PROOF: Let H be the support of μ. For $\Gamma \in E^*$ define $a = \sup_{y \in H} \Gamma(y)$. Defining $f(x) = \min(\Gamma(x), a)$, we obviously have $f \leq a$ and $[x \in H] \Rightarrow [f(x) = \Gamma(x)]$. Using (16.145) and (16.148) one can see that

$$\begin{aligned}\Gamma(\text{average}(x)) = \Gamma(\mu(1_E)) &= \mu(\Gamma \circ 1_E) = \\ &= \mu(\Gamma) = \mu(f) \leq \mu(a) = a\,.\end{aligned}$$

Thus the average of x belongs to the set defined by (16.193$\heartsuit$), which is the smallest closed convex set containing H. □

(16.195) COROLLARY: Using the hypotheses of (16.191), if Q is a closed convex set, if $\text{im}(F) \subseteq Q$, and if average$(F(x))$ exists, then average$(F(x)) \in Q$.

PROOF: It suffices to use (16.152). □

(16.196) EXAMPLE: Let x be a random variable and define $F(x) \equiv \begin{pmatrix} x \\ x \otimes x \end{pmatrix}$; F takes its values in the vector space of pairs $\begin{pmatrix} {}^1M \\ {}^2M \end{pmatrix}$ with ${}^1M$ a contravariant tensor of degree 1 on E and ${}^2M$ a symmetric contravariant tensor of degree 2 on E. One can verify that the subset Q of pairs $\begin{pmatrix} {}^1M \\ {}^2M \end{pmatrix}$ such that[354]

$$1 + 2\ {}^1M(C) + {}^2M(C)(C) \geq 0 \qquad \forall C \in E^*\,,$$

is a closed convex set containing $\text{im}(F)$. We thus have average$(F(x)) \in Q$, which can be written as

$$1 + 2\ {}^1M(C) + {}^2M(C)(C) \geq 0 \qquad \forall C \in E^*\,,$$

where ${}^1M$ and ${}^2M$ now denote the first and second order *moments of the random variable* x. This property can also be written in the form

$$\Phi(C)(C) \geq 0 \qquad \forall C \in E^*\,,$$

[352] See Bourbaki [8].
[353] Abuse of language meaning the support of the probability measure of x.
[354] For the notation see Souriau [31].

where Φ denotes the tensor ${}^2M - {}^1M \otimes {}^1M$. Φ is called the *variance tensor* of x.

Entropy and Gibbs measures

(16.197) DEFINITION:

a) Let V be a Hausdorff manifold, let $x \mapsto \varphi$ be a continuous field of densities such that $\varphi > 0 \ \forall x$,[355] and let λ be the corresponding measure (16.83).

b) If μ is a *completely continuous probability measure* on V, there exists a function $f \in \mathcal{C}(V)$ such that $f \geq 0$ and $\mu = \lambda \times f$. The λ-*entropy* of the probability measure μ will be the number

$$s = \begin{cases} \lambda(H \circ f) & \text{if } H \circ f \text{ is } \lambda\text{-integrable} \\ -\infty & \text{if } H \circ f \text{ is not } \lambda\text{-integrable}, \end{cases}$$

where H is the function

$$H(\rho) = \begin{cases} 0 & \text{if } \rho = 0 \\ -\rho \log \rho & \text{if } \rho > 0. \end{cases}$$ □

With some abuse of notation we thus can write

(16.198) $$s = \int_V -\rho \log \rho \, \varphi(dx) \qquad [\rho \equiv f(x)].$$

It is useful to know the graph of H, which is given in figure (16.II).

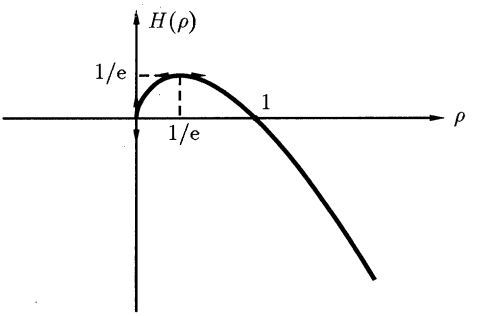


Fig 16.II

(16.199) DEFINITION: Let V be a Hausdorff manifold. Assume that the following are given.

[355] We will say that $x \mapsto \varphi$ is *strictly positive*.

a) A continuous and strictly positive field of densities $x \mapsto \varphi$, defining a measure λ on V.
b) A continuous map Ψ from V to a finite dimensional vector space E.

By a (generalized) *Gibbs measure* we will mean a probability measure ζ such that

$$\diamondsuit \quad \begin{cases} \exists\, z \in \mathbf{R}\ \exists\, Z \in E^* :\ \zeta = \lambda \times f \quad \text{with} \quad f(x) \equiv e^{-[z+Z(\Psi(x))]} \\ \Psi \text{ is } \zeta\text{-integrable.} \end{cases}$$

(16.200) THEOREM: The λ-entropy of a Gibbs measure exists and is equal to

$$\heartsuit \qquad s = z + Z(M)\,,$$

where we introduce

$$\clubsuit \qquad M = \zeta(\Psi)\,.$$

All *other* completely continuous probability measures ζ' satisfying $\zeta'(\Psi) = M$ *have entropy that is strictly less than that of a Gibbs measure.*

PROOF: Since the graph of the function H (Fig. 16.II) is concave (the second derivative is $H''(\rho) = -1/\rho$), it lies below its tangent line at a point ρ_0, that is,

$$[\rho \geq 0\,,\ \rho_0 > 0\,,\ \rho \neq \rho_0] \quad \Longrightarrow \quad [H(\rho) < H(\rho_0) + [\rho - \rho_0]H'(\rho_0)]\,.$$

Replacing H' by its value, the above inequality becomes

$$H(\rho) < \rho_0 - \rho - \rho \log(\rho_0)\,.$$

Hence, if two completely continuous probability measures have densities $\rho\varphi$ and $\rho_0\varphi$, respectively, then the entropy s of the first probability measure satisfies[356]

$$s \leq \int -\rho \log(\rho_0)\, \varphi(dx)\,.$$

If the second probability measure is a Gibbs measure, that is, $\rho_0 \equiv e^{-[z+Z(\Psi(x))]}$, then

$$s \leq \int_V \rho[z + Z(\Psi(x))]\, \varphi(dx) = z + Z(M)\,.$$

Here equality holds if and only if $\rho \equiv \rho_0$ (theorem (16.84)). □

[356] Use that the two probability measures have the same mass 1 and thus that $\int_V [\rho_0 - \rho]\, \varphi(dx) = 0$.

Now let ζ be a Gibbs measure defined by means of z and Z (notation (16.199)). If we recall that $\text{mass}(\zeta) = 1$ and that Ψ is ζ-integrable, we see that the two integrals

$$I_0 = \int_V e^{-Z(\Psi(x))}\,\varphi(dx) \tag{16.201}$$

$$I_1 = \int_V \Psi(x)\, e^{-Z(\Psi(x))}\,\varphi(dx) \tag{16.202}$$

are *convergent* and that

$$z = \log(I_0) \tag{16.203}$$

$$M = \frac{I_1}{I_0}\,. \tag{16.204}$$

(16.205) Conversely, if for a given Z in E^* the integrals I_0 and I_1 are convergent, then I_0 is positive and thus we can define a number z by (16.203). And then the density $e^{-[z+Z(\Psi(x))]}\,\varphi$ is the density of a Gibbs measure.

(16.206) EXAMPLE: (*Gaussian normal measure*) Let us consider the case $V = \mathbf{R}^n$, λ the Lebesgue measure, and $\Psi(x) \equiv \begin{pmatrix} x \\ x \otimes x \end{pmatrix}$. Any element Z of the dual of E [357] can be characterized uniquely by an $a \in \mathbf{R}^n$ and a symmetric matrix H such that

$$Z(\Psi(x)) \equiv \overline{a} \cdot x + \tfrac{1}{2}\,\overline{x} \cdot H \cdot x\,.$$

One can verify that the integral I_0 is convergent if and only if the matrix H is positive.[358] In that case the associated Gibbs measure is called the *Gaussian normal measure*. One can easily compute I_0 by using the change of variable $x^* = H^{1/2}x + H^{-1/2}a$,[359] which gives

$$z = \tfrac{1}{2}\,[\,\overline{a} \cdot H^{-1} \cdot a - \log(\det(H)) + n\log(2\pi)\,]\,. \tag{16.207}$$

If H is positive, I_1 also converges, and thus one can compute M, which is defined by the first and second order moments of the probability measure (16.199). A computation shows that the first order moment is equal to $-H^{-1} \cdot a$ and that the components of the *variance tensor* (16.196) are equal to the entries of the matrix H^{-1}. The second order moment follows immediately.

---

[357] Here E denotes the vector space of pairs $\begin{pmatrix} {}^1m \\ {}^2m \end{pmatrix}$ with ${}^im$ a (symmetric) contravariant tensor of degree i on $\mathbf{R}^n$ (see (16.196)).

[358] See Souriau [31].

[359] That is, finding the *image* of the probability measure under the map $x \mapsto x^*$.

Formula (16.200♡) gives the *entropy* as

$$s = \frac{n}{2}\log(2\pi e) - \tfrac{1}{2}\log(\det(H))\,. \tag{16.208}$$

We thus can associate to each probability measure on $\mathbf{R}^n$ having an *injective* fluctuation,† a *Gaussian normal measure having the same first and second order moments.* The given probability measure has entropy *strictly less* than the value given in (16.208), unless it already is a Gaussian normal measure. □

(16.209) THEOREM: The map $Z \mapsto M$ defined by (16.204) is *injective*, unless the image of Ψ is an affine subspace of dimension less than that of E.

PROOF: Let us assume that two different values Z and Z' lead to the same M. Theorem (16.200) shows that each of these two probability measures gives the maximum entropy compatible with this value of M. Thus the entropies are equal and hence the densities are the same. So we have $z + Z(\Psi(x)) \equiv z' + Z'(\Psi(x))$, which shows that $\mathrm{im}(\Psi)$ is contained in the subspace of vectors $y \in E$ such that $[Z - Z'](y) = z' - z$. □

(16.210) THEOREM: (Hypotheses (16.200).)

a) Let Ω be an open subset of E^*. Let us assume that there exists a function $f \in \mathcal{C}(V)$ such that

$$e^{-Z(\Psi(x))} \leq f(x) \qquad \forall\, x \in V\,,\ \forall\, Z \in \Omega$$

and that the integrals

$$I_0' = \int_V f(x)\,\varphi(dx) \qquad \text{and} \qquad I_1' = \int_V \Psi(x) f(x) \varphi(dx)$$

converge.[360] Then the function $[Z \mapsto I_0]$ is *once differentiable* in Ω, and we may differentiate I_0 under the integral sign:

$$\diamondsuit \qquad \delta I_0 \equiv \int_V -[\delta Z](\Psi(x))\, \mathrm{e}^{-Z(\Psi(x))} \varphi(dx)\,.$$

b) Furthermore, if the integral

$$I_2' = \int_V [\Psi(x) \otimes \Psi(x)] f(x) \varphi(dx)$$

†Editors' note: Meaning that the variance tensor is represented by an invertible matrix.

[360]When this is the case, the integrals I_0 and I_1 are said to converge *uniformly* on Ω.

converges, we may also differentiate I_1 under the integral sign:

$$\heartsuit \qquad \delta I_1 \equiv -\int_V \Psi(x)[\delta Z](\Psi(x))\,\mathrm{e}^{-Z(\Psi(x))}\varphi(dx)\,.$$

PROOF: We will take this result for granted. □

Let us now assume the hypotheses of (16.210a). By comparing $\diamondsuit$ with (16.203) and (16.204) we find that

$$\delta z \equiv -[\delta Z](M)\,, \tag{16.211}$$

which (after identifying E with its bidual $[E^*]^*$) can be written as

$$M = -\frac{\partial z}{\partial Z}\,. \tag{16.212}$$

This last formula can be verified in example (16.207).[361] □

Let us suppose in addition that the hypotheses of (16.210b) are satisfied. Then M is a once differentiable function of Z. From (16.210$\heartsuit$) one can easily deduce the formula

$$[\delta Z](\delta M) = -\frac{1}{I_0}\int_V \left\{ [\delta Z]\left(\Psi(x) - \frac{I_1}{I_0} \right) \right\}^2 e^{-Z(\Psi(x))}\,\varphi(dx)\,, \tag{16.213}$$

whence the inequality

$$[\delta Z](\delta M) \leq 0\,. \tag{16.214}$$

Moreover, *if the image of Ψ is not contained in an affine subspace of dimension less than that of E*, then

$$[\delta Z \neq 0] \quad \Longrightarrow \quad [[\delta Z](\delta M) < 0]\,. \tag{16.215}$$

It obviously follows that $\partial M/\partial Z$ is injective. Since $[Z \mapsto M]$ is injective, we see that $[Z \mapsto M]$ is a diffeomorphism[362] (from the open subset of E^* where the conditions of existence and differentiability are satisfied, to an open subset of E). Thus the maps $[M \mapsto z]$ and $[M \mapsto s]$ are differentiable (at least once). Combining (16.200$\heartsuit$) and (16.211) we immediately find

$$\delta s \equiv Z(\delta M)\,, \tag{16.216}$$

[361] The set of Z where I_0 converges is open and can be covered by open sets Ω_j where the hypotheses of (16.210a) and (16.210b) hold.

[362] In a restricted sense, since we are not sure that derivatives of order ≥ 2 exist. However, the theorems that give invertibility still apply.

which can be written as

(16.217) $$Z \equiv \frac{\partial s}{\partial M}\,.$$

The Gibbs canonical ensemble of a dynamical group

(16.218) HYPOTHESES: Let U be a symplectic manifold and let G be a dynamical group of U.[363] We will assume that

1) U is a Hausdorff manifold,
2) G is a *connected Lie group*,[364]
3) the G action on U possesses a *moment* μ.

Thus there exists a map $\Psi\ [x \mapsto \mu]$ from U to $\mathcal{G}^*$, the dual of the Lie algebra $\mathcal{G}$ of G,[365] such that

$$\sigma(Z_U(x)) \equiv -\mathrm{d}[\mu \cdot Z] \qquad \forall\, Z \in \mathcal{G} \quad (Z \text{ constant})\,,$$

where σ is the Lagrange form on U. □

We know that U has a measure λ — *the Liouville measure* — defined in terms of the *symplectic density* φ (see (16.96)). This measure is absolutely continuous and strictly positive, that is, it satisfies the hypotheses (16.197), and thus every completely continuous probability measure on U possesses a *Liouville entropy* s (which we will abbreviate as *entropy*). In this way we can construct Gibbs measures using the map Ψ, which leads to the following theorem.

(16.219) THEOREM: Under the hypotheses (16.218) we obtain the following results.

1) Let Ω be the largest open subset of $\mathcal{G}$ such that $\forall\, Z \in \Omega$ the integrals

$$\diamondsuit \qquad \begin{aligned} I_0 &= \int_U \mathrm{e}^{-\mu(Z)}\varphi(dx) \\ I_1 &= \int_U \mu\, \mathrm{e}^{-\mu(Z)}\varphi(dx) \\ I_2 &= \int_U \mu \otimes \mu\, \mathrm{e}^{-\mu(Z)}\varphi(dx) \end{aligned}$$

[363] Definition (11.1).

[364] Recall that every dynamical group is a Lie group, which is necessarily Hausdorff (6.12).

[365] Recall that elements of $\mathcal{G}^*$ are called *torsors* of G. Sufficient conditions for the existence of a moment have been given in (11.8).

converge and the maps $Z \mapsto I_0$ and $Z \mapsto I_1$ are once differentiable under the integral sign.[366]

For $Z \in \Omega$ we define

$\heartsuit$ $$z \equiv \log(I_0)$$

$\spadesuit$ $$M \equiv \frac{I_1}{I_0}.$$

2) For $Z \in \Omega$ there exists a probability measure ζ defined by the density

$\clubsuit$ $$e^{-[z+\mu(Z)]}\varphi \qquad \text{(a Gibbs measure).}$$

It satisfies[367]

$\bowtie$ $$\text{average}\,(\mu) = M\,,$$

and its entropy is given by

$\sharp$ $$s = z + M \cdot Z\,.$$

Among all completely continuous probability measures satisfying $\bowtie$, the probability measure $\clubsuit$ is the one which has the largest entropy.

3) The maps $Z \mapsto z$, $Z \mapsto s$, and $Z \mapsto M$ are once differentiable and we have

$\triangle$ $$\frac{\partial z}{\partial Z} \equiv -M\,, \qquad \delta M \cdot \delta Z \leq 0 \qquad \forall\, \delta Z\,.$$

PROOF: The above results are a simple transcription of the results of (16.200–214) up to a change of notation: since the image of Ψ lies in $\mathcal{G}^*$, the dual space of the vector space $\mathcal{G}$, we can identify every element Z of $\mathcal{G}^{**}$ with an element of $\mathcal{G}$ by writing $\mu(Z)$ instead of $Z(\mu)$. □

(16.220) THEOREM: (Continuation of (16.219).) Let us assume furthermore that G acts *effectively* on U, that is,

$$[a \in G,\ \underline{a}_U = \underline{1}_U] \implies [a = \text{the identity element of } G]\,.$$

Then the map $Z \mapsto M$ is injective. Its image Ω^* is an open subset of $\mathcal{G}^*$, the maps $M \mapsto z$, $M \mapsto s$, and $M \mapsto Z$ are once differentiable, and we have

$\triangledown$ $$\frac{\partial s}{\partial M} \equiv Z\,,^{368} \qquad \delta M \cdot \delta Z < 0 \qquad (\forall\, \delta Z \neq 0)\,.$$

[366] In (16.210) we have given sufficient conditions. It is enough that they are satisfied on the elements of an open cover of Ω.

[367] The correct notation is $\zeta(\Psi) = M$. We implicitly assume that we have turned x into a random variable with probability measure ζ. Then formula $\bowtie$ follows from (16.191).

PROOF: Due to theorems (16.209) and (16.215) it suffices to show that the image of Ψ is not contained in an affine subspace of dimension less than that of $\mathcal{G}^*$, that is, that there does not exist $Z \in \mathcal{G}$ and $c \in \mathbf{R}$ such that

$$\mu\, Z \equiv c \qquad (Z \neq 0\,,\ Z \text{ and } c \text{ constants})\,.$$

Using the definition of a moment (see (16.218)), this would imply that we would have $\sigma(Z_U(x)) \equiv 0$, and thus $Z_U(x) = 0$, since σ is injective (9.1). From (6.12) it follows immediately that for every real α, the element $a \equiv \exp(\alpha\, Z_G)(e)$ of G would satisfy $\underline{a}_U \equiv \underline{1}_U$. Hence $a = e$, because of (16.220). This is impossible since $da/d\alpha \equiv Z_G(a) \neq 0$, according to (3.1). □

(16.221) We know that the moment μ of the dynamical group G is defined only *up to an additive constant* (theorem (11.8)). If we replace μ by $\mu - \mu_0$, one can check that the definition (16.219.1) of the open set Ω does not change. Moreover, for $Z \in \Omega$ the variables z, M, and s change according to

$$\diamondsuit \qquad z \to z + \mu_0 Z\,, \qquad M \to M - \mu_0\,, \qquad \text{and} \qquad s \to s\,.$$

We conclude that the probability measure ♣ associated with Z *does not change*, and thus that the set of these probability measures does not depend on the choice of the moment μ. We will call this set of probability measures the *Gibbs canonical ensemble of the dynamical group G.*

PROOF OF THEOREM (16.222) BELOW: Let a be an arbitrary element of G (Fig. 16.III). Since $\underline{a}_U^{-1}$ is a symplectomorphism, the image under $\underline{a}_U^{-1}$ of the Liouville measure λ is equal to λ (Liouville's theorem (16.99)). For $Z \in \Omega$ the integral I_0 can be written as $\lambda(x \mapsto e^{-\Psi(x)Z})$ and then theorem (16.150) shows that

$$\diamondsuit \qquad I_0 = \lambda(x \mapsto e^{-\Psi(\underline{a}_U^{-1}(x))Z})\,.$$

But we know that

$$\Psi(\underline{a}_U^{-1}(x)) \equiv \underline{a}_{\mathcal{G}^*}^{-1}(\Psi(x)) + \theta(a^{-1})\,,$$

where θ is a symplectic cocycle of G (11.17◇). Substituting this into ◇ and defining

$$\heartsuit \qquad Z' = \underline{a}_{\mathcal{G}}(Z)\,,$$

[368] Abuse of notation meaning $\delta s \equiv \delta M \cdot Z$.

we see, using (11.15), that the integral I'_0 corresponding to the value $\heartsuit$ of Z is convergent and that

$$\clubsuit \qquad I'_0 = e^{\theta(a^{-1})Z} I_0 .$$

An analogous computation shows that the integrals I'_1 and I'_2 also converge and that the maps $Z' \mapsto I'_0$ and $Z' \mapsto I'_1$ are differentiable under the integral sign. Using formulæ $\clubsuit$, (16.219), (11.17$\heartsuit$), and (16.86) we obtain the following theorem.

(16.222) THEOREM: (Continuation of (16.219); see Fig. 16.III.) The set Ω is invariant under every transformation $\underline{a}_{\mathcal{G}}$, where $a \mapsto \underline{a}_{\mathcal{G}}$ is the *adjoint representation* of G. The variables (16.219) change according to

$$\begin{aligned} Z &\to \underline{a}_{\mathcal{G}}(Z) \\ z &\to z + \theta(a^{-1})Z \equiv z - \theta(a)\underline{a}_{\mathcal{G}}(Z)\,^{369} \\ s &\to s \\ M &\to \underline{a}_{\mathcal{G}^*}(M) + \theta(a) = \underline{a}_{\mathcal{G}^*_\theta}(M)\,^{370} \\ \zeta &\to \underline{a}^+_U(\zeta)\,.^{371} \end{aligned}$$

□

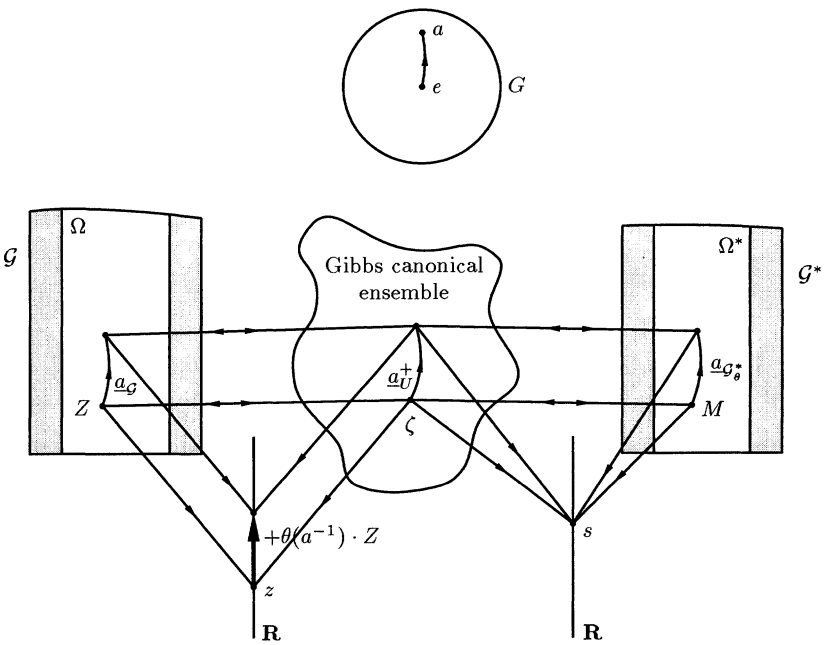


Fig 16.III

Let us state a purely algebraic lemma.

(16.223) LEMMA: Let f be a *symplectic cocycle* of a Lie algebra $\mathcal{G}$, that is, a 2-form on $\mathcal{G}$ satisfying

$$f(Z_1)([Z_2, Z_3]) + f(Z_2)([Z_3, Z_1]) + f(Z_3)([Z_1, Z_2]) \equiv 0 .$$

Let Z be an element of the *kernel* of f and let E be the image of $\mathrm{ad}(Z)$. Then there exists a symmetric tensor g on E defined by

$$g([Z, Z'])(Z'') = f(Z')(Z'') \qquad \forall\, Z' \in \mathcal{G}\,,\ \forall\, Z'' \in E\,. \qquad \square$$

(16.224) THEOREM: (Continuation of (16.222).) Let f be the derivative of θ at the identity element and let us define (notation (6.13))

$$f_Z \equiv f + M \cdot \mathrm{ad} \qquad \forall\, Z \in \Omega\,.$$

Then

a) f_Z is a *symplectic cocycle* of $\mathcal{G}$ which does not depend on the choice of the moment μ of G.

b) $$f_Z(Z) = 0 \qquad \forall\, Z \in \Omega\,.$$

c) There exists a symmetric tensor g defined on the image E of $\mathrm{ad}(Z)$ such that

$\diamondsuit$ $$g([Z, Z'])(Z'') = f_Z(Z')(Z'') \qquad \forall\, Z' \in \mathcal{G}\,,\ \forall\, Z'' \in E\,.$$

Moreover, g satisfies

$\heartsuit$ $$g(Z'')(Z'') \leq 0 \qquad \forall\, Z'' \in E\,.$$

d) If E is not the vector space $\{0\}$ and if G acts *effectively* on U, then g gives E the structure of a *negative Euclidean space*.

PROOF: Since we know that f is a symplectic cocycle (11.33), f_Z is also a symplectic cocycle due to the Jacobi identity.

If we replace μ by $\mu - \mu_0$ (theorem (11.8)), it follows from the formulæ (11.18), (11,16), and (16.221$\diamondsuit$) that f_Z does not change. Hence (a).

---

[369]The cocycle θ, associated with the group G and with the moment μ, has been defined in (11.7).

[370]Notation (11.28).

[371]$\underline{a}_U^+(\zeta)$ is the image under $\underline{a}_U$ of the probability measure ζ (notation (16.86) or (16.149)).

Differentiating the identity

$$[Z \mapsto z](\underline{a}_{\mathcal{G}}(Z)) \equiv z - \theta(a)\,\underline{a}_{\mathcal{G}}(Z)\,,$$

which is an immediate consequence of (16.222), and taking $a = e$, $\delta a = Z'$, and $\delta Z = 0$, it follows, using (6.11), (6.26), and (16.219△), that $M[Z', Z] = -f(Z')(Z)$. Hence (b).

From this the existence of the tensor g satisfying ◇ follows immediately (use lemma (16.223)).

As above let us differentiate the identity

$$[Z \mapsto M](\underline{a}_{\mathcal{G}}(Z)) \equiv \underline{a}_{\mathcal{G}^*}(M) + \theta(a)\,,$$

which is again a consequence of (16.222). We find

$$\frac{\partial M}{\partial Z}(-[Z', Z]) \equiv f(Z') + M \cdot \mathrm{ad}(Z') \equiv f_Z(Z')\,.$$

Writing $\delta Z = [Z, Z'] = Z''$, formula (16.219△) $\delta M \cdot \delta Z \leq 0$ can be written as $f_Z(Z')(Z'') \leq 0$, which is ♡. Similarly, if $Z'' \neq 0$ and if G acts effectively, the inequality $\delta M \cdot \delta Z < 0$ (16.220) shows that $g(Z'')(Z'') < 0$. Thus g is necessarily injective. □

§17. The principles of statistical mechanics

Statistical states

(17.1) Let U be the space of motions of a dynamical system. In classical or relativistic mechanics, the state of the system (that is, the complete description of its history) is characterized by a point x in U.



By a *statistical state* of the system we will mean a *probability measure* μ on the manifold U.[372]

Put another way: in statistical mechanics, the motion x of a dynamical system is a *random variable* (definition (16.183)), whose probability measure defines the *statistical state* of the system.

(17.2) What interpretation should one give to a statistical state? The opinions on this subject are rather diverse.[373] We will tackle this question only after having seen some examples ((17.52) below).

Hypotheses of the kinetic theory of gases

(17.3) The initial problem of statistical mechanics was to study a gas. The "kinetic theory" describes a gas as a *dynamical system* with a space of

[372] One often defines (see for example Mackey [26]) a statistical state by means of a probability measure μ_t on the phase space U_t at time t (definition (12.31)). Furthermore, one postulates a condition (formulated by means of a partial differential equation in the case where the probability measure is completely continuous) which amounts to saying that the *image* $P^+(\mu_t)$ of the probability measure μ_t under the projection P to the space of motions U (Fig. 12.I) is *independent* of t. It is this image μ which we will consider directly.

[373] A statistical state represents:
—an average over numerous repeated experiments;
—a simplification in order to study systems whose space of motions has a very large dimension;
—an evaluation of the influence of a "large" system on a "small" one to which it is weakly coupled;
—a measure of our partial knowledge of the actual state of the system (the theory of information);
—a realistic description of nature (this point of view necessarily has to be compared with the quantum mechanical point of view): see (19.17)).

motions of a large but *finite* dimension. This system then is described, in a microscopic fashion, by means of molecules (they themselves being treated most often as material points or rigid bodies) interacting among themselves and with the walls of the box which is supposed to contain the gas.

(17.4) These interactions can be treated by *scattering theory*. To simplify, let us assume that the forces acting directly on the molecules (electromagnetic field, gravitation) are negligible and that the interactions amount to instantaneous collisions.[374] The symplectomorphisms of scattering theory can then be indexed by time.[375] Let us recall the formalism (15.61).

If x is a motion of the gas which does not involve a collision at time t, we write

(17.5) $$x = A_t(x^*)\,,$$

where x^* denotes the motion that the molecules would have if they were all free, with the positions and velocities that they have effectively at time t. We define

(17.6) $$S_{t'}^t = A_t^{-1} \circ A_{t'} \qquad \forall\, t, t' \in \mathbf{R}\,.$$

The $S_{t'}^t$ are local symplectomorphisms from U^* to U^* (where U^* is the manifold of motions of the system of free molecules) which satisfy the relations (15.67):

(17.7) $$\begin{aligned} S_t^t &= 1_{E_t} \\ S_t^{t'} &= [S_{t'}^t]^{-1} \\ S_{t'}^t \circ S_{t''}^{t'} &< S_{t''}^t\,. \end{aligned}$$

(17.8) Here E_t is the open subset of U^* consisting of the free motions such that, at time t, all molecules are inside the box without touching each other.

(17.9) We know that the manifold of motions of the gas is completely defined when we know the $S_{t'}^t$ (see (15.68)).

The Galilei group acts on the manifold U^* and hence so does the 1-dimensional subgroup of time translations. We denote by $\underline{e}_{U^*}$ the sym-

[374] We essentially assume that there exist *free paths* for the molecules. The hypothesis that the collisions are instantaneous *is not indispensable* for applying scattering theory.

The kinetic theory envisages *also* the case where the mutual interactions between the molecules are not negligible at a distance. Having formulated hypotheses about these interactions, one investigates especially whether one can find equilibrium states with n molecules having — in a certain sense — a limit when n tends to infinity.

[375] See for example the case of the planar mirror (15.110).

plectomorphism $\underline{a}_{U^*}$ with (notation (13.3))

$$a \equiv \begin{bmatrix} 1_{\mathbf{R}^3} & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \qquad e \in \mathbf{R}. \tag{17.10}$$

The map $e \mapsto \underline{e}_{U^*}$ allows us to consider the additive group $\mathbf{R}$ as a dynamical group of U. If x^* is a free motion of the molecules, then $\underline{e}_{U^*}(x^*)$ is the motion which is obtained by a *delay* of e.

(17.11) Let us now assume that the box is *immobile*. We then assume that the $S_t^{t'}$ *commute* with the $\underline{e}_{U^*}$ (that is, the collision laws are not changed by a delay). It follows that time translations form a *symmetry group* (see (15.69) to (15.73)) for the scattering process. $\mathbf{R}$ also acts on the motions of the gas in such a way that

$$\underline{e}_U \circ A_t = A_t \circ \underline{e}_{U^*} \qquad \forall t, e \in \mathbf{R}.$$

$\underline{e}_U(x)$ is evidently the motion of the gas obtained by delaying the motion x by e.

Thus the gas can be considered to be a *conservative system* — in the sense of (12.153). The moment of the group of time translations is simply the *Hamiltonian* h (or if one prefers the *energy*), up to a sign. This can be written as

$$\frac{\partial}{\partial t}[\underline{t}_U(x)] \equiv F(\underline{t}_U(x)), \tag{17.12}$$

where F is the vector field

$$F(x) \equiv -\operatorname{grad} h \tag{17.13}$$

defined by the symplectic gradient of the dynamical variable h.

Equilibria of a conservative system

(17.14) Let us consider such a conservative system. We will call an *equilibrium* of the system any motion x which is *invariant under the group of time translations*:

$$\underline{t}_U(x) = x \qquad \forall t \in \mathbf{R}.$$

We see immediately that the equilibria are the points of U where $\operatorname{grad} h$ vanishes.

Likewise, we define a *statistical equilibrium* of the system to be a *statistical state* μ *which is invariant under the group of time translations*, that is,

$$\underline{t}_U^+(\mu) = \mu \qquad \forall t \in \mathbf{R}.$$

Here $\underline{t}_U^+(\mu)$ denotes the image of the measure μ under the diffeomorphism $\underline{t}_U$, which is again a probability measure on U (16.184). This definition generalizes the preceding one, when we identify the motion x with the probability measure defined by the Dirac measure δ_x.[376]

(17.15) In order to find equilibria, one might imagine taking an arbitrary motion x and letting it "age". This is evidently the same as *advancing* the motion an amount t in time and seeing whether there exists a limit of $\underline{-t}_U(x)$ as t tends to $+\infty$.

Likewise one can ask for a statistical state μ whether there exists a limit μ_∞ of $\underline{-t}_U^+(\mu)$ as t tends to ∞ and whether it is a statistical equilibrium.

But how do we *define* the "limit" μ_∞ of this measure? We can think of the *norm limit*, defined by

$$\| \underline{-t}_U^+(\mu) - \mu_\infty \| \to 0 , \tag{17.16}$$

or of the *weak limit*, defined by

$$\underline{-t}_U^+(\mu)(g) \to \mu_\infty(g) \qquad \forall g \in \mathcal{K}(U) , \tag{17.17}$$

or even of the *weak limit in the mean*, defined by

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T [\underline{-t}_U^+(\mu)(g)]\, dt \to \mu_\infty(g) \qquad \forall g \in \mathcal{K}(U) . \tag{17.18}$$

(17.19) It is easy to show that each of these conditions implies the following. Using Birkhoff's theorem[377] one can give conditions in which the weak limit in the mean exists and is a statistical equilibrium. In that case the mass of the limit μ_∞ is equal to the mass of μ, namely 1.

One can also show, in certain cases, that the *average value of the Hamiltonian* h (assumed to be μ-integrable) is preserved in the limit:

$$\mu_\infty(x \mapsto h) = \mu(x \mapsto h) .$$

(17.20) If μ is a completely continuous measure, one can easily verify that the entropy of $\underline{t}_U^+(\mu)$ is *constant* and equal to that of μ. *But* the entropy of the weak limit μ_∞ might be *different* from that of μ.[378] If μ is not an equilibrium state, one can show — under certain hypotheses — that the

[376] Indeed, $A^+(\delta_x) = \delta_{A(x)}$ for every homeomorphism A.

[377] See Birkhoff [6] and also, Blanc-Lapierre, et al. [7].

[378] In other words, entropy is not continuous in the weak topology, as will be shown in example (17.22).

entropy of μ_∞ is *greater* than that of μ.[379] We will only give a heuristic argument (17.21) and an example (17.22).

(17.21) HEURISTIC ARGUMENT: Let us assume that μ is completely continuous. Writing it in the form $\mu = \mu_0 \times f$ with μ_0 the Liouville measure of U (16.96), the entropy of μ is, by definition, $\mu_0(H \circ f)$, where H is the function defined in (16.197). Let us denote by μ_T the measure defined by

$$\mu_T(g) = \frac{1}{T}\int_0^T [\underline{-t_U^+}](\mu)(g)\,dt \qquad [\forall\, g \in \mathcal{K}(U)]\,.$$

One can easily see (using Liouville's theorem (16.99)) that $\mu_T = \mu_0 \times f_T$, where f_T is the continuous function defined by

$$f_T(x) = \frac{1}{T}\int_0^T f(\underline{t}_U(x))\,dt\,.$$

From this one can deduce the formula

$$\mu_{2T} = \tfrac{1}{2}\,[\mu_T + \underline{-T_U}(\mu_T)]\,.$$

It is immediate that the measures μ_T and $\underline{-T_U}(\mu_T)$ have the same entropy. *If these two measures are not equal, it follows that the entropy of μ_{2T} is greater than that of μ_T.* Indeed, if $\rho\varphi$ and $\rho'\varphi$ are the densities of these two probability measures,[380] the entropy of the half-sum can be written as $\int_U H(\frac{1}{2}(\rho+\rho'))\,\varphi(dx)$. Since H is convex (Fig. 16.II), this entropy is greater than $\int_U \frac{1}{2}\,[H(\rho)+H(\rho')]\,\varphi(dx)$ and thus greater than the half-sum of the entropies (which are equal in the case under consideration). □

(17.22) EXAMPLE: Consider a symplectic manifold U of dimension 2 equipped with a canonical chart $\begin{pmatrix} p \\ q \end{pmatrix} \mapsto x$. We define an action of $\mathbf{R}$ on U by means of the Hamiltonian $h \equiv [p^2+q^2]^2$. Let μ be a probability measure on U: $\mu = \mu_0 \times f$, where μ_0 is the Liouville density and f is a function of the form[381]

$$f(x) \equiv p^+\,F(h)\,.$$

The positive function F should be chosen so that $\text{mass}\,(\mu) = 1$ and so that h is μ integrable (Fig. 17.I).

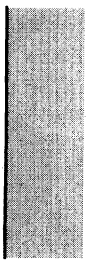
---

[379] To be compared with the famous "H-theorem" of Boltzmann, certainly false as stated, which asserts that the entropy is an increasing function of time when there is no equilibrium.

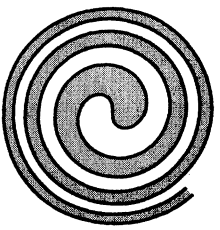
[380] In (17.21) φ will denote the symplectic density of U (16.94).

[381] Notation (16.32): $p^+ = p$ for $p \geq 0$ and $p^+ = 0$ for $p \leq 0$.

Fig 17.I



support of μ



support of $[\underline{-t}]^+(\mu)$

With these assumptions one can easily verify that the *weak limit* of $[\underline{-t_U}]^+(\mu)$ *exists* and that it is the probability measure $\mu_\infty = \mu_0 \times f^*$ where $f^*(x) \equiv h^{1/4}F(h)/\pi$. Moreover, one can show that *the average value* of the energy h is *the same* for μ and μ_∞, and that the entropy of μ_∞ is obtained from the entropy of μ *by adding the positive number* $\log(2\pi/e)$. □

(17.23) If we start from the preconceived idea that the "natural" equilibrium states are those that are spontaneously established by aging, the above remarks lead us to the idea that the equilibria are all the more "natural" if they have a greater entropy. This in turn leads us to formulate the following *principle*.

(17.24) The "*natural*" equilibrium states of a system form the *Gibbs canonical ensemble* of the dynamical group of *time translations*.

Indeed, among those equilibrium states for which the average value of the energy takes a prescribed value, the Gibbs measures are those which have the *largest entropy* (theorem (16.200)).

One certainly can give more refined motivations for this principle.[382] Be that as it may, we are going to study this principle in itself and compare it with experiments. Later on in (17.77) we will propose a modification. □

(17.25) A natural equilibrium state will thus be characterized by an element Z of the Lie algebra of the Lie group $\mathbf{R}$, that is, Z is a *real number*. We will see later on that Z determines the *equilibrium temperature*. Denoting by E the average value of the energy h, the results (16.219) allow us to write the formulæ:

$$\diamondsuit \qquad z = \log \int_U \mathrm{e}^{-Z\,h}\,\varphi(dx)$$

[382] See in particular, Kinchin [21], Blanc-Lapierre et al. [7], Jaynes [19], and Mackey [26].

(17.26)
$$\heartsuit \qquad E = -\frac{\partial z}{\partial Z}$$
$$\clubsuit \qquad s = z + EZ$$
$$\spadesuit \qquad \frac{\partial E}{\partial Z} \leq 0\,,$$

where we have introduced the real variable z, called *Planck's thermodynamic potential*. The density of the equilibrium state is

(17.27)
$$\mathrm{e}^{-[z+Zh]}\,\varphi\,,$$

where φ denotes, as in (17.26), the symplectic density on U.

In the general case, when the motions of the system are not all equilibria, one can apply (16.220). It turns out that the function $Z \mapsto E$ is *strictly decreasing* and that

(17.28)
$$\spadesuit \qquad \frac{\partial E}{\partial Z} < 0$$
$$\bowtie \qquad \frac{\partial s}{\partial E} = Z\,. \qquad \square$$

Applying (16.221) we find that, if we replace the Hamiltonian h by $h - h_0$ (h_0 being an arbitrary constant), then the variables (17.26) undergo the substitution

(17.29)
$$Z \to Z\,, \qquad s \to s\,, \qquad z \to z - h_0 Z\,, \qquad E \to E - h_0\,,$$

and hence the set of natural equilibrium states *does not depend* on the choice of h_0.

As the group of time translations is abelian, theorem (16.222) shows that the elements of this group act on the equilibrium states *without changing them* and without changing the values of the variables z, E, and s. The principle of Gibbs (17.24) indeed defines *equilibrium* states in the sense of (17.14).

The *dimensions* of the quantities associated with equilibria are given by (notation (13.13)):

(17.30)

σ	A
φ	A^N

Z	TA^{-1}
E	AT^{-1}

e^z	A^{-N}
e^s	A^{-N}

,

where $2N$ is the dimension of the symplectic manifold U.

We see that Z *does not have* the same dimension T as an infinitesimal time translation ε (13.13), and that the variables z and s *do not have any dimension*, as long as one has not chosen $A = 1$, that is, as

long as one has not selected a unit of action (as we will do later on in quantum mechanics). On the contrary, the variable E always has the same dimension as energy.

Let us assume that the system is composed of several noninteracting systems. We know that the space of motions U of the composite system is the *symplectic direct product* of the spaces of motions U_j of the component systems (12.146). This implies that we can choose the Hamiltonian of the composite system to be the sum of the Hamiltonians h_j of the component systems. We also know that the Liouville measure on U is the tensor product of the Liouville measures on U_j (16.98). Applying (16.143) and (16.114) we see that the integral I_0 (notation (16.201)) can be written in the form

$$I_0 = \prod_j \int_{U_j} \mathrm{e}^{-Zh_j} \varphi_j(dx_j)\,, \tag{17.31}$$

whence the statement:

(17.32) Every natural equilibrium state of a system *composed of noninteracting systems* is obtained by taking the *tensor product* of natural equilibrium states of the corresponding component systems, all *at the same temperature*.[383] The variables z, E, and s of the composite system are obtained by adding the corresponding variables z_j, E_j, and s_j of the component systems. □

(17.33) We thus can say that in every natural equilibrium of a composite system, the motions of the component systems are *independent random variables* (definition (16.188)). Conversely, if we endow each component system with a natural equilibrium, we can define (by taking the tensor product) an equilibrium of the composite system. This equilibrium *is natural only if the temperatures of the component systems are all equal.*

This fact could make one doubt the physical validity of the Gibbs principle (17.24). But one should note that, experimentally, the existence of an interaction, however weak, prevents the establishment of a nonisothermal equilibrium (for example, the thermal conductivity of a solid plate separating two media). Let us also note that treatises on thermodynamics define the equality of two temperatures by the fact that juxtaposition of two systems in equilibrium is again a system in equilibrium. Besides, this is the working principle of most *thermometers*.

[383] See (17.25)

Ideal gases

(17.34) Suppose we have to study the natural equilibrium of a gas enclosed in a box. Such a study obviously assumes a detailed knowledge of the interactions between the molecules in order to be able to compute the integral (notation (16.201))

(17.35) $$I_0 = \int_U \mathrm{e}^{-Zh}\varphi(dx)\,.$$

Let us choose a time t. A first *approximation* consists of replacing the integral I_0 by an integral taken over the set of motions which do not have a collision at time t,[384] a set which is equal to $\mathrm{im}(A_t)$ (notation (17.5)). Using the fact that A_t is a symplectomorphism (15.64), we thus will replace I_0 by I_0':

(17.36) $$I_0' = \int_{E_t} \mathrm{e}^{-Zh}\varphi_*(dx^*)\,,$$

where φ_* is the symplectic density of the space of *free* motions.

(17.37) A *second approximation* consists of neglecting the proper volume of the molecules, that is to say, more precisely, of replacing E_t by the set E_t' of free motions in which each molecule is inside the box at time t.[385] This approximation is called the *ideal gas* approximation.



In these approximations the probability density, the entropy, the thermodynamic potential, and the average value of the energy are *the same as if the molecules did not have any interaction at all*,[386] their motions being defined by their positions and velocities at the chosen time t. Referring to (17.32), we see that we are brought back[387] to the study of the natural equilibrium of *each molecule assumed to be alone in the box at the given temperature*. Thus there occurs a *statistical decomposition* of the equilibrium state. However, *there is no dynamical decomposition* because the space of motions of the gas is not in any way the direct product of the spaces of motions of isolated molecules.

---

[384]This amounts to assuming that the set of motions for which a collision occurs at time t is contained in an open set of very small probability. If this probability can be made smaller than any given positive number, one says that the set is *negligible*.

[385]One thus considers to be very small the probability that molecules touch or interpenetrate each other at the given time t.

[386]And as a consequence could pass through each other without being deflected.

[387]Either by assuming that the space of free motions of the molecules is the *direct product* of the spaces of motions of each molecule (see (17.32)), or by assuming that the molecules are *indiscernible* and then working on the covering space defined in (15.46).

A monatomic ideal gas

(17.38) In a monatomic ideal gas each molecule is considered as a material point in classical mechanics. The results are the following:

At each time t, the *position* and the *velocity* of each molecule are *independent* random variables.

The probability measure of the velocity $\mathbf{v}$ has a density

$$\left[\frac{mZ}{2\pi}\right]^{3/2} e^{-mZ\|\mathbf{v}\|^2/2}\,\varphi \qquad \text{(the Maxwell distribution of velocities)}, \tag{17.39}$$

where m is the mass and φ is the Euclidean density on $\mathbf{R}^3$.[388]

The probability density of the position $\mathbf{r}$ is[389]

$$\frac{1}{V}\,\varphi\,, \tag{17.40}$$

where φ is still the Euclidean density on $\mathbf{R}^3$ and V is the volume of the box.

An arbitrary ideal gas

(17.41) In an arbitrary ideal gas the size of the molecules is assumed to be small compared to the size of the box (this means essentially that a molecule is "inside" the box if its center of mass is). Using *nonrelativistic* mechanics we know that the space of free motions of a molecule is the *symplectic product* of the space of motions of the center of mass and the space of proper motions (13.24). This leads to the following results.

(17.42) a) The *position*, *velocity*, and the *proper motion* of each molecule are independent random variables.[390]

---

[388] φ coincides with the Lebesgue density (16.85). It is thus a *normal distribution* (see (16.206)). Assuming, as is customary, that m is positive, we see that the set Ω (notation (16.219)) is $\mathbf{R}^+$: Z can take *all positive values*.

[389] This formula assumes that exterior forces, such as the electric field or gravity, are negligible. If gravity is not negligible, one would have found the *barometric distribution* (see (17.110)).

[390] In the case of indiscernible molecules this result is valid on the covering space. The actual probability measures are obtained by taking the image under the projection $x \mapsto \widehat{x}$ (notation (15.45)).

b) The probability measure of velocity and the one of position are the same as in the monatomic case (17.39) and (17.40).[391]
c) The probability measure of the proper motions is the Gibbs measure corresponding to the given value of Z.

An ideal gas thermometer

(17.43) Let us reproduce schematically the principle of an apparatus used for the precise measurement of temperature (the *hydrogen thermometer*). A gas is enclosed in a cylinder equipped with a piston of mass M and weight Mg, which can only move vertically — by means of ideal constraints (Fig. 17.II). In this model the influence of gravity on the molecules themselves is neglected.[392]

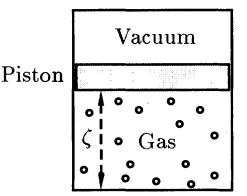


Fig 17.II

Under these conditions, *the positions of the molecules are no longer independent random variables.* But the height ζ of the piston itself is a *random variable*, whose probability density can be calculated as[393]

$$\rho = [ZMg]^{n+1} \mathrm{e}^{-Zmg\zeta} \frac{\zeta^n}{n!}, \tag{17.44}$$

where n denotes the *number of molecules* in the box.

One can immediately deduce that the average value of ζ is given by

$$\text{average}(\zeta) = \frac{n+1}{ZMg}. \tag{17.45}$$

Hence, by introducing the *pressure* exerted by the piston $p = Mg/A$ (A = area of the piston) and the *random variable* V, which is the enclosed

[391]There is a direct experimental verification of the Maxwell distribution (17.39). The frequency of a spectral line emitted by an atom depends apparently on its velocity (the Doppler-Fizeau effect). Thus the spectral lines emitted by a hot gas are broadened with a spectral distribution which is easily deduced from (17.39). This effect is used to measure temperatures high up in the atmosphere and in stars.

[392]Just to simplify the calculations. A more complete study shows that this effect is indeed negligible under standard conditions.

[393]The Lebesgue density on $\mathbf{R}^+$ is implicit.

volume ($V \equiv A\zeta$), we obtain

$$p \times \text{average}(V) = \frac{n+1}{Z} \,. \tag{17.46}$$

[394] □

Now experiments show that real gases — when their density is very low[395] — satisfy to a very high degree the law of Mariotte-Gay-Lussac-Avogadro:

$$pV = n\,kT\,, \tag{17.47}$$

where T defines the *absolute temperature* and k is *Boltzmann's constant*

$$k = 1.3804 \cdot 10^{-16} \text{ erg/degree Kelvin.} \tag{17.48}$$

□

Since the ratio $n/(n+1)$ is very close to 1 in experiments, comparing (17.46) with (17.47) gives the relation

$$Z = \frac{1}{kT} \tag{17.49}$$

between the usual *absolute temperature* T measured by the thermometer and the parameter Z. But we still have to understand why the random variable V in an experiment can be confused with its average value.

It is easy to deduce from (17.44) that the average value of ζ^2 is

$$\text{average}(\zeta^2) = \frac{[n+1][n+2]}{[ZMg]^2} \,, \tag{17.50}$$

from which follows the *variance* of ζ (definition (16.196)). The square root of the *variance*, called the *mean quadratic error*, or *standard deviation* of ζ, equals

$$\Delta\zeta = \frac{\text{average}(\zeta)}{\sqrt{n+1}} \,. \tag{17.51}$$

This formula shows that $\Delta\zeta/\text{average}(\zeta)$ is *very small* under usual circumstances (of the order of 10^{-11} for a liter of hydrogen at ordinary temperature). Moreover, one can see that the graph of the probability measure (17.44) possesses a *very sharp peak*. The value (17.51) of $\Delta\zeta$ gives an estimate of the *size* of this peak (Fig. 17.III only gives a qualitative impression of it).

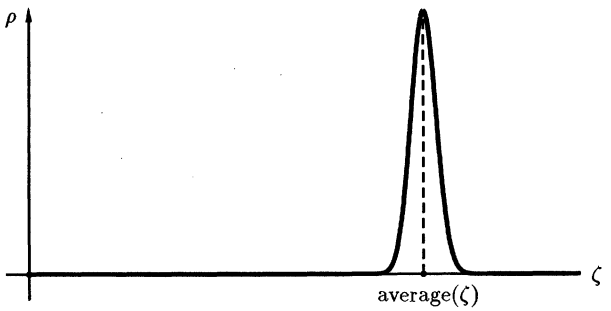
(17.52) We thus have to admit, as a rule how to interpret statistical mechanics, that a random variable whose probability measure is highly concentrated can be confused — experimentally — with its average value.

[394] Recall that Z is necessarily positive (17.39) and thus p too.

[395] The approximations in the theory of an ideal gas evidently assume that the molecules are well separated from each other — and thus that the gas is rarefied.

Mathematically one can prove that the probability measure of the variable ζ/average ζ converges to the Dirac measure δ_1 *in the sense of the weak topology* (17.17), *when* n *tends to infinity.* □

Fig 17.III



Heat and work

We know that the entropy s of a system can be expressed as a function of the energy E and that we have (17.28⋈)

$$\frac{\partial s}{\partial E} = Z\,. \tag{17.53}$$

(17.54) It turns out that Z is always positive under conditions we know of (see (17.39) and (17.49)). We thus can express E as a function of s and, if need be, of a variable u describing the characteristics of the system — u varying over a manifold, to fix ideas. If $\begin{pmatrix} u \\ s \end{pmatrix} \mapsto E$ is differentiable, we can write (because of (17.53))

$$\delta E \equiv \frac{\delta s}{Z} + \varpi(\delta u)\,, \tag{17.55}$$

where ϖ is a 1-form (it too is called the *thermodynamic potential*!). Using (17.49) we can write

$$\delta E = T\,\delta S + \varpi(\delta u)\,, \qquad \text{with } S \equiv k\,s\,. \tag{17.56}$$

In general, it is S that is called "entropy" in thermodynamics; it coincides with s if we choose the unit of temperature so that the Boltzmann constant equals 1.

If we let the equilibrium evolve as a function of a parameter and if we denote by δ the associated derivation, formula (17.56) is generally written as

$$\delta E = \delta Q + \delta W \tag{17.57}$$

with

(17.58) $$\delta Q = T\,\delta S \qquad \text{and} \qquad \delta W = \varpi(\delta u)\,.$$

The variable Q is called *heat* and the variable W *work*. Note that Q and W *are not functions of* S *and* u, but are *action integrals* (in the sense of the calculus of variations) and depend on the chosen path of evolution.

(17.59) Several kinds of evolution deserve to be examined separately. *Adiabatic* evolution: S = constant, no exchange of heat, one can take $W = E$. *Isothermal* evolution: T = constant, one can take $Q = TS$ and $W = E - TS$. *Heating* or *refrigeration*: u = constant, one can take $Q = E$; it is this case where the formulæ (17.26) are valid.

(17.60) The formulæ (17.57) and (17.58) represent *the two principles of thermodynamics* ((i) the equivalence of heat and work and (ii) dQ/T is a "total differential"). Let us give a first example of a thermodynamic calculation.

On the manifold swept out by the pair $\begin{pmatrix} S \\ u \end{pmatrix}$, Poincaré's theorem $\mathrm{d}\mathrm{d}E = 0$ gives

$$\diamondsuit \qquad \frac{\delta T}{T}\,\delta' Q - \frac{\delta' T}{T}\,\delta Q + \delta[\varpi(\delta' u)] - \delta'[\varpi(\delta u)] - \varpi([\delta, \delta']\,u) \equiv 0$$

(the *generalized Maxwell* relation). Choosing $\delta' T$ = constant, $\delta' u = 0$, $\delta T = 0$, and δu independent of T we obtain the Clausius-Clapeyron relation†

$$\heartsuit \qquad \delta Q = -T\,\frac{\partial \varpi}{\partial T}(\delta u)\,,$$

which is valid for every infinitesimal isothermal transformation described by δu. Experimentally this relation is well satisfied, for example in the case of liquefaction of a gas.

Specific heat

Let us consider a gas enclosed in a container of adjustable volume V. We can take V as the parameter u. In this case formula (17.56) becomes

(17.61) $$\delta E = T\,\delta S - p\,\delta V\,,$$

where p is called the *pressure*.[396]

†Editors' note: This is a generalized form. In example (17.61) the variable u is the volume V and $\varpi(\delta u)$ is written as $-p\,\delta V$. One then recovers the more familiar form of the Clausius-Clapeyron equation: $\delta Q = T\,\frac{\partial p}{\partial T}\,\delta V$.

[396]It coincides with the result of the measurement described in (17.43), provided we assume a priori that the variance of the volume is negligible (we have calculated this variance in the case of an ideal gas).

(17.62) The quantity $\partial Q/\partial T$ is called the *heat capacity* of the gas. It obviously depends on the chosen path of evolution.

In particular, the heat capacity *at constant volume* is

$$c \equiv T\,\frac{\partial S}{\partial T} \equiv -k\,Z^2\,\frac{\partial E}{\partial Z} \equiv -k\,Z\,\frac{\partial s}{\partial Z}\,. \tag{17.63}$$

Inequality (17.28♠) shows that c is *always positive*.

(17.64) For an evolution at *constant pressure*, we can obviously take $W \equiv -p\,V$, whence $Q \equiv E + p\,V$ (the variable $E + p\,V$ is called the *enthalpy*). We will denote the corresponding heat capacity by C.

EXAMPLE: Let us assume[397] that Planck's thermodynamic potential of a fluid can be written as

$$z \equiv f(Z) + g(V)\,. \tag{17.65}$$

Using the formulæ (17.26), (17.61), and (17.58) one can easily obtain the *energy*

$$E = -f'(Z)\,, \tag{17.66}$$

the *entropy*

$$s = f(Z) - Z\,f'(Z) + g(V)\,, \tag{17.67}$$

the *pressure*

$$p = \frac{g'(V)}{Z}\,, \tag{17.68}$$

and the *heat capacities*

$$c = k\,Z^2\,f''(Z)\,, \qquad C = k\left[Z^2 f''(Z) - \frac{(g'(V))^2}{g''(V)}\right]. \tag{17.69}$$

□

A calculation shows that the thermodynamic potential of an *ideal gas* is of the form (17.65) with

$$\begin{cases} f(Z) = -\dfrac{3n}{2}\log(Z) + \sum\limits_j [\frac{3}{2}\,\log(2\pi\,m_j) + z_j] \\ g(V) = n\log(V)\,, \end{cases} \tag{17.70}$$

[397]It is in this way that an *ideal fluid* is *defined* in thermodynamic treatises. Or, equivalently, by the fact that the energy is independent of the volume (formula (17.66), the experiment of Joule).

where the m_j are the masses of the n molecules and the z_j are the thermodynamic potentials of their proper motions.[398]

We see that the heat capacity c (at constant volume) *depends on the model chosen for the molecules.* One can easily prove the following relations — which are independent of the model:

$$c_m \geq \tfrac{3}{2}\, k \tag{17.71}$$

and

$$C_m - c_m = k\,,^{399} \tag{17.72}$$

where c_m and C_m denote the *specific heats per molecule* (the quotients of c and C by the number n of molecules). Consequently, for the ratio γ of the specific heats we obtain

$$\gamma = \frac{C}{c} = 1 + \frac{k}{c_m} \leq \frac{5}{3}\,. \tag{17.73}$$

The relations (17.71), (17.72), and (17.73) are well verified by real fluids that are not too much compressed.

The next table gives the values of z, c_m, and γ in various models.[400]

(17.74)

Model	z	c_m	γ
material point	0	$3\,k/2$	5/3
particle with spin	0	$3\,k/2$	5/3
rectilinear solid	$\log T$ + constant	$5\,k/2$	7/5
nonrectilinear solid	$\frac{3}{2}\log T$ + constant	$3\,k$	4/3
two points attracted by a force proportional to r^p	$\left[\frac{3}{2} + \frac{3}{p+1}\right] \log T$ + constant	$\left[3 + \frac{3}{p+1}\right] k$	$< 4/3$

Experiments show, for monatomic gases, that γ is close to 5/3. The model of a molecule consisting of a material point or a particle with spin is thus acceptable. For a bi-atomic gas, γ is a decreasing function of T, which generally remains between 1.6 and 1.3. We see that the above models do not fit, since they give a constant value for c_m and γ.

[398] See (17.42). If the molecules are indiscernible, one should diminish $f(Z)$ by the constant $\log n!$, which does not change p, c, or C.

[399] The relation of Robert von Mayer.

[400] One should note that the existence of spin does not change the results, as the alleged "principle of equipartition of energy" might lead us to believe.

The difficulty in finding a suitable molecular model has been one of the important motivations for quantum mechanics.

(17.75) The specific heat of a *solid* can be calculated immediately by assuming that its constituent atoms oscillate around an equilibrium position. In the linearized approximation (see (12.150)) the space of motions is a vector space of dimension $6n$ (where n is the number of atoms, considered as material points), and the Hamiltonian is a positive quadratic form. In such a case the probability measure of the natural equilibrium is a *normal distribution*. Formula (16.207) gives immediately $z \equiv -3n \log(Z) +$ constant, and thus we obtain the atomic specific heat $c_{at} = 3k$. Agreement with experiment is good if the temperature is not too low (*the rule of Dulong and Petit*).

Covariant statistical mechanics

(17.76) The group of time translations (7.9) is a subgroup of the Galilei group, but, as a simple calculation shows, *it is not a normal subgroup*. If a dynamical system is *conservative* in an inertial reference frame, it follows that it *need not be conservative in another*. Thus the formulation of the principle of Gibbs (17.24) has to be generalized to become compatible with Galilean relativity.

We propose the following principle.

(17.77) If a dynamical system is invariant under a Lie subgroup G' of the Galilei group, then the natural equilibria of the system form the Gibbs ensemble of the dynamical group G'.

Let $\mathcal{G}'$ be the Lie algebra of G'. We know that $\mathcal{G}'$ is a Lie subalgebra of the Lie algebra $\mathcal{G}$ of G. An equilibrium of the system is characterized by an element Z of $\mathcal{G}'$ and thus of $\mathcal{G}$. Using notation (13.4) we thus can write

(17.78)
$$Z = \begin{bmatrix} j(\omega) & \beta & \gamma \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix}.$$

Z varies over the set Ω defined in (16.219) and to each Z is associated an element M of $(\mathcal{G}')^*$, which is the average value of the moment μ. We now can apply the formulæ (16.219) and (16.220) which generalize the (17.79) thermodynamic relations (17.26), (17.27), and (17.28). We see that it is Z (17.78) which generalizes the notion of "temperature". The isothermal theorem (17.32) generalizes immediately: the equilibrium of a system composed of several noninteracting parts is obtained by attributing to

each component an equilibrium, *all corresponding to the same value of* Z. The entropy s, the Planck potential z, and the average moment M are *additive*.

(17.80) Just as in the classical case, we will interpret Z by means of a "thermometer" made up of a *connected bubble of an ideal gas*. We again apply scattering theory to the free motion of the molecules and we assume that the scattering process admits G' as a *symmetry group* (definition (15.69)). By applying the theory of symmetry groups (15.73), the approximation of ideal gases (17.36), the barycentric decomposition ((13.14) to (13.36)), and by using definitions (16.219) and (16.221) of the Gibbs ensemble, we see that the motion x of the *center of mass* of each molecule (considered as a material point of mass m) and the *proper motion* x' of each molecule are *independent* random variables, having probability densities

$$C\mathrm{e}^{-\psi(x)\cdot Z}\,\varphi \tag{17.81}$$

and

$$C'\mathrm{e}^{-\psi'(x')\cdot Z}\,\varphi'\,, \tag{17.82}$$

respectively, where $\psi(x)$ and $\psi'(x')$ are the moments (in the dual $\mathcal{G}^*$ of the Lie algebra of the Galilei group), φ and φ' the corresponding symplectic densities, and C and C' constants which depend on the bubble. □

In order to interpret formula (17.81), we are going to perform some changes of variables.

Using Noether's theorem, we can write $\psi(x) \equiv \psi(y)$ (Fig. 12.III), where $y \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \end{pmatrix}$ lies in the evolution space V of the *center of mass* and where ψ is the map

$$\psi(y) \equiv m\{\mathbf{r}\times\mathbf{v},\ \mathbf{r}-\mathbf{v}\,t,\ \mathbf{v},\ \tfrac{1}{2}\|\mathbf{v}\|^2\} \qquad \text{(formula (13.36))}. \tag{17.83}$$

Since the Galilei group acts on V, we can *define*

$$y_0 \equiv \begin{pmatrix} t_0 \\ \mathbf{r}_0 \\ \mathbf{v}_0 \end{pmatrix} \equiv \exp\left(-\frac{t}{\varepsilon}\,Z_V\right)(y)\,. \tag{17.84}$$

Moreover we have

$$\begin{pmatrix} \mathbf{r}_0 \\ t_0 \\ 1 \end{pmatrix} = \exp\left\{-\begin{bmatrix} j(\boldsymbol{\omega}t/\varepsilon) & \boldsymbol{\beta}\,t/\varepsilon & \boldsymbol{\gamma}\,t/\varepsilon \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix}\right\}\cdot\begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \tag{17.85}$$

(theorem (6.12)), which in particular allows us to verify that

$$t_0 \equiv 0\,. \tag{17.86}$$

On the other hand, formula (12.116) gives

$$\begin{pmatrix} \mathbf{v}_0 \\ 1 \\ 0 \end{pmatrix} = \exp\{-\frac{t}{\varepsilon} Z\} \begin{pmatrix} \mathbf{v} \\ 1 \\ 0 \end{pmatrix}. \tag{17.87}$$

Writing $a \equiv \exp(tZ/\varepsilon)$, it follows that

$$\begin{aligned} \psi(y)\cdot Z &\equiv \psi(\underline{a}_V(y_0))\cdot Z \equiv [\underline{a}_{\mathcal{G}^*}\psi(y_0) + \theta(a)]\cdot Z && (11.17\diamondsuit) \\ &\equiv \psi(y_0)\underline{a^{-1}}_{\mathcal{G}}(Z) + \theta(a)\cdot Z && (11.15)\,. \end{aligned}$$

Applying the integral formula (11.22♠) and the fact that the cocycle θ is symplectic, we see that $\theta(a)\cdot Z \equiv 0$. Moreover, one can easily see that $\underline{a^{-1}}_{\mathcal{G}}(Z) \equiv Z$, and thus

$$\psi(y)\cdot Z \equiv \psi(y_0)\cdot Z\,. \tag{17.88}$$

Using (17.83), (17.86), and (12.122) we obtain

$$\psi(y)\cdot Z \equiv -m\varepsilon \left[\tfrac{1}{2}\,\|\mathbf{v}_0 - \mathbf{U}^*\|^2 + f(\mathbf{r}_0)\right] \tag{17.89}$$

with

$$\mathbf{U}^* \equiv \frac{\boldsymbol{\omega}}{\varepsilon} \times \mathbf{r}_0 + \frac{\boldsymbol{\gamma}}{\varepsilon} \tag{17.90}$$

and

$$f(\mathbf{r}_0) \equiv \langle \mathbf{r}_0, \frac{\boldsymbol{\beta}}{\varepsilon}\rangle - \tfrac{1}{2}\,\|\mathbf{U}^*\|^2\,. \tag{17.91}$$

Formula (17.85) shows that

$$\mathbf{r}_0 \equiv A(\mathbf{r}+\mathbf{R}) \qquad \left[A = \exp\left(j\left(\frac{-\boldsymbol{\omega}\,t}{\varepsilon}\right)\right) \in \mathrm{SO}(3),\ \mathbf{R}\in\mathbf{R}^3\right], \tag{17.92}$$

where A and $\mathbf{R}$ are functions of t. One can interpret (17.92) as saying that $\mathbf{r}_0$ is the position of the molecule in a *new reference frame*, which in general is *accelerated* (it suffices to realize that it is the same as formula (12.59)). We now define

$$\mathbf{r}^* \equiv \mathbf{r}_0, \qquad \mathbf{v}^* \equiv \text{the relative velocity with respect to the new reference frame,}^{401} \tag{17.93}$$

and

$$\boldsymbol{\omega}^* = \frac{\boldsymbol{\omega}}{\varepsilon}, \qquad \boldsymbol{\beta}^* = \frac{\boldsymbol{\beta}}{\varepsilon}, \qquad \boldsymbol{\gamma}^* = \frac{\boldsymbol{\gamma}}{\varepsilon}\,. \tag{17.94}$$

[401] Do not confuse $\mathbf{v}^*$ and $\mathbf{v}_0$.

We then obtain

$$\mathbf{U}^* \equiv \boldsymbol{\omega}^* \times \mathbf{r}^* + \boldsymbol{\gamma}^*, \qquad f(\mathbf{r}^*) \equiv \langle \boldsymbol{\beta}^*, \mathbf{r}^* \rangle - \tfrac{1}{2} \|\mathbf{U}^*\|^2 . \tag{17.95}$$

The *velocity* $\mathbf{U}$ and the *drag acceleration* $\boldsymbol{\Gamma}$ (12.66) are given by

$$\begin{pmatrix} \mathbf{U} \\ 1 \\ 0 \end{pmatrix} \equiv \exp(\frac{t}{\varepsilon} Z) \cdot \begin{pmatrix} \mathbf{U}^* \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{\Gamma} \\ 0 \\ 0 \end{pmatrix} \equiv \exp(\frac{t}{\varepsilon} Z) \cdot \begin{pmatrix} \boldsymbol{\omega}^* \times \mathbf{U}^* + \boldsymbol{\beta}^* \\ 0 \\ 0 \end{pmatrix} . \tag{17.96}$$

Using (17.87) and (17.96), the *relative velocity* (formula (12.67))

$$\mathbf{v}^* = \exp(j(-\boldsymbol{\omega}' t))[\mathbf{v} - \mathbf{U}^*] , \tag{17.97}$$

can be written as

$$\mathbf{v}^* = \mathbf{v}_0 - \mathbf{U}^* . \tag{17.98}$$

On the other hand, the formulæ (12.44) and (12.67) show that in the new reference frame there appears a force, called the *force of inertia*,[402] given by

$$\mathbf{F}^* \equiv \mathbf{E}^* - \mathbf{B}^* \times \mathbf{v}^* , \tag{17.99}$$

with

$$\mathbf{E}^* \equiv -m \, [\boldsymbol{\omega}^* \times \mathbf{U}^* + \boldsymbol{\beta}^*] \tag{17.100}$$

and

$$\mathbf{B}^* \equiv 2m \, \boldsymbol{\omega}^* . \tag{17.101}$$

It is immediate that these forces can be derived from a *potential* v^* and a *vector potential* $\mathbf{A}^*$ (definition (12.91)) given by

$$v^* = m \, f(\mathbf{r}^*) \qquad \mathbf{A}^* = m \, \boldsymbol{\omega}^* \times \mathbf{r}^* , \tag{17.102}$$

which allows us to construct the *Hamiltonian* (formula (12.93))

$$h^* = \tfrac{1}{2} m \, \|\mathbf{v}^*\|^2 + m \, f(\mathbf{r}^*) . \tag{17.103}$$

Comparing formula (17.103) with (17.89) and taking (17.98) into account, we thus can write the probability density (17.81) of the center of mass of a molecule as

$$C \mathrm{e}^{\varepsilon h^*} \varphi^* , \tag{17.104}$$

[402] The two terms in expression (17.99) of $\mathbf{F}^*$ are called the *centrifugal force* and the *Coriolis force*.

where φ^* denotes the symplectic density in the new reference frame (since the change of reference frame preserves the symplectic form σ (12.68); it also preserves the symplectic density). We thus have proven the following theorem.

(17.105) THEOREM: (Notation (17.78), (17.94).) The natural equilibrium of an ideal gas (in the general sense (17.77)) corresponding to a given value of Z appears as a natural equilibrium (in the classical sense (17.24)) with absolute temperature

$$\diamondsuit \qquad T = -\frac{1}{k\varepsilon} \qquad (k = \text{the Boltzmann constant})$$

in the reference frame in which the position $\mathbf{r}^*$ is given by

$$\heartsuit \qquad \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \equiv \exp\left\{ t \begin{bmatrix} j(\boldsymbol{\omega}^*) & \boldsymbol{\beta}^* & \boldsymbol{\gamma}^* \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right\} \cdot \begin{pmatrix} \mathbf{r}^* \\ 0 \\ 1 \end{pmatrix}.$$ [403]

(17.106) REMARK: In making the change of variables (17.84) we have implicitly assumed that $\varepsilon \neq 0$. The interpretation (17.105$\diamondsuit$) shows that ε is *negative* under usual conditions.

(17.107) REMARK: If a system is decomposed into noninteracting subsystems, the equilibrium is obtained by assigning the same value of Z to the components (17.79). Thus the subsystems not only have *the same absolute temperature* T (17.105$\diamondsuit$), but also *the same reference frame of apparent equilibrium* (17.105$\heartsuit$) and (17.94).

(17.108) REMARK: We have assumed for simplicity that there are no forces acting directly on the molecules (for example, terrestrial attraction in case of the atmosphere). In the general case, it suffices to add the corresponding potential to the Hamiltonian h^*.

Examples

(17.109) THE WIND ($\boldsymbol{\omega} = 0$, $\boldsymbol{\beta} = 0$).

With these assumptions we find

$$\mathbf{r} \equiv \mathbf{r}^* + \boldsymbol{\gamma}^* t\,, \qquad f(\mathbf{r}^*) = \text{constant}\,.$$

[403] This formula is the equation of space-time orbits of the group of dimension 1, whose Lie algebra is generated by Z (these orbits are indexed by $\mathbf{r}^*$).

The new reference frame is an *inertial* reference frame. In this frame the probability of presence† is constant and the distribution of velocities is Maxwellian (17.39). The vector $\boldsymbol{\gamma}^*$ can be interpreted, in the old reference frame, as the velocity of the wind. □

(17.110) AN ACCELERATING ROCKET ($\boldsymbol{\omega} = 0, \boldsymbol{\gamma} = 0$).

As the matrix which appears in (17.105♡) is nilpotent (its cube is zero), its exponential can be calculated immediately using the power series expansion (6.17). One obtains

$$\diamondsuit \qquad \mathbf{r} = \mathbf{r}^* + \tfrac{1}{2}\boldsymbol{\beta}^* t^2 .$$

Thus the new reference frame is *uniformly accelerating* with acceleration $\boldsymbol{\beta}^*$ (one might think of the cockpit of a rocket ship whose functioning is controlled ad hoc). Formula (17.100) shows that there prevails an *artificial gravity* $-\boldsymbol{\beta}^*$.[404] The internal potential is $v^* = m\langle \boldsymbol{\beta}^*, \mathbf{r}^* \rangle$, and thus the probability density of presence is proportional to

$$\exp\left(-\frac{m}{kT}\langle \boldsymbol{\beta}^*, \mathbf{r}^* \rangle\right).$$

We thus find the barometric distribution, obtained by assuming that terrestrial gravity is constant and equal to $-\boldsymbol{\beta}^*$. □

(17.111) A CENTRIFUGE ($\boldsymbol{\beta} = 0,\ \boldsymbol{\gamma} = 0$).

With these assumptions we find

$$\mathbf{r} \equiv \exp(j(\boldsymbol{\omega}^* t))\, \mathbf{r}^* .$$

The new reference frame is thus *uniformly rotating*, where $\boldsymbol{\omega}^*$ is the *angular velocity vector*.[405]

The probability of presence of the gas is proportional to

$$\exp\left(\frac{m}{2kT}\|\boldsymbol{\omega}^* \times \mathbf{r}^*\|^2\right). \tag{17.112}$$

The appearance of m in the above expression shows — in the case of an inhomogeneous gas — that *the relative concentration of the various constituents varies with the distance to the axis of rotation*, according to a formula which can be deduced immediately from (17.112). This effect

†Editors' note: Meaning the probability density of position.

[404]This example has been chosen by Einstein to illustrate the principle of equivalence between the force of gravity and that of inertia.

[405]Formula (17.107) shows that the centrifuge must turn with *the same angular velocity* as the gas and must have *the same temperature*.

is well verified experimentally; it is used for the enrichment of uranium. □

Let us now apply formula (17.82) to the above case, assuming that the gas consists of *particles with spin* (for example, an ionized hydrogen plasma). Using (13.35) and (14.50) we see that the "element of probability" of the unit vector $\mathbf{u}$ defining the direction of the spin is

(17.113) $$C \exp\left(\frac{s}{kT}\langle \mathbf{u}, \boldsymbol{\omega}^* \rangle\right) \varphi(d\mathbf{u}),$$

where C is a constant, s the spin, and $\varphi(d\mathbf{u})$ *the element of solid angle* (16.174). It follows that the most likely orientation of the spin is $\boldsymbol{\omega}^*$. The ratio of its probability to that of the opposite orientation is $\exp(2s \,\|\, \boldsymbol{\omega}^* \| / kT)$. The experimental verification of this result seems to be difficult because this number is very close to 1 under the usual circumstances. □

The statistical equilibrium of an isolated system

(17.114) It seems natural, a priori, to apply principle (17.77) to an isolated system. But this attempt comes up against a fundamental difficulty: such equilibria *do not exist.*

One can arrive at this conclusion using a purely algebraic method, *which thus is valid whatever the model of the system.* Formula (b) of theorem (16.224) can be written as

(17.115) $$m\, f_0(Z) + M \cdot \mathrm{ad}(Z) = 0,$$

where m is the *mass* of the system, f_0 the symplectic cocycle defined in (12.130), Z the element of the Lie algebra defining the equilibrium, and M the average value of the Galilean moment. Now one can verify that equation (17.115) imposes conditions on Z which are *not* satisfied on *any open set.* And this is incompatible with the hypotheses (16.219) defining the Gibbs ensemble.

(17.116) Another way to arrive at the above conclusion uses the *barycentric decomposition* (13.24). If the system is in a natural equilibrium, its center of mass (put on a par with a material point endowed with all the mass of the system) is also. However, one can check directly that a free material point does not have a natural equilibrium under the action of the Galilei group.[406] □

[406] Furthermore, we know that such equilibria would define equilibria for ideal fluids.

(17.117) It is clear that the difficulty presented above is of a *cosmological* type: the Galilean universe is unsuitable for having a statistical equilibrium. And although Galilean cosmology is rather naive, nevertheless observation confirms this fact: *the universe is expanding.*

It is thus natural to avoid this cosmological problem by investigating whether the *proper motion*[407] on its own can have an equilibrium in conformity with principle (17.77). We know that the Galilei group acts on the space of proper motions by the intermediary of the group G' consisting of the matrices

$$\begin{bmatrix} A & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \qquad A \in \mathrm{SO}(3)\,,\ e \in \mathbf{R} \tag{17.118}$$

(see (13.30)). This leads us to postulate that every natural equilibrium of the space of proper motions is defined by an element

$$Z = \begin{bmatrix} j(\boldsymbol{\omega}) & 0 & 0 \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \tag{17.119}$$

of the Lie algebra of G'.

(17.120) Thus we find ourselves in the case studied in (17.111). This implies that there exists a reference frame (uniformly rotating about an axis passing through the center of mass 0 with angular velocity vector $\boldsymbol{\omega}^* = \boldsymbol{\omega}/\varepsilon$) in which the system appears to be in a usual equilibrium with temperature $T = -1/k\varepsilon$.

Here formula (16.224b) can be written as

$$\boldsymbol{\sigma} \times \boldsymbol{\omega}^* = 0\,,$$

where $\boldsymbol{\sigma}$ is the average value of the proper angular momentum (13.35). Thus there exists a number I such that

$$\boldsymbol{\sigma} = I\,\boldsymbol{\omega}^*\,. \tag{17.121}$$

This property is described in classical language by saying that the axis of rotation is a *principal axis of inertia* of the system and that I is the *moment of inertia* about this axis.

But experiments, as well as calculations, show that ideal fluids do not possess equilibria (with or without acceleration of the type (17.105)), *except if they are enclosed in a box.*

[407] Definition (13.26).

Furthermore, inequality (16.224♡) becomes simply

(17.122) $$I > 0\,.$$

Observations actually confirm these last conclusions (rotation of celestial bodies), whereas various phenomena (radioactivity, radiation) do not admit a thermal equilibrium.

Relativistic statistical mechanics

Obviously we obtain a relativistic statement by replacing the Galilei group by the Poincaré group in (17.77):

(17.123) If a dynamical system is invariant under a Lie subgroup G' of the Poincaré group G, then the natural equilibria of the system form the Gibbs ensemble of the dynamical group G'.

This principle seems to agree in a satisfactory way with experiments. We are going to study some of its consequences, so let us recall some notation.

The element of the Lie algebra $\mathcal{G}$ of G which describes the equilibrium will be denoted (as in (13.54)) by

(17.124) $$Z = \begin{bmatrix} \Lambda & \Gamma \\ 0 & 0 \end{bmatrix},$$

where Γ is a vector in Minkowski space E_4 and Λ is a skew Hermitian linear map on this space ($\overline{\Lambda} = -\Lambda$).

(17.125) The barycentric decomposition does not exist in relativistic mechanics. However, by using theorem (16.224) and the fact that the symplectic cohomology of the Poincaré group is zero (13.73), one can show that *an isolated system has no natural equilibria* — except in the case, devoid of any physical interpretation, where the Poincaré group leaves invariant every motion of the system.[408]

We thus have to restrict ourselves to a subgroup G' strictly smaller than G. We will now examine some examples.

[408]Of course one might call such a system "vacuum" and consider that it is a natural equilibrium for every Z. This remark ceases to be purely linguistic in quantum mechanics.

A relativistic ideal gas

Let us consider an ideal gas *consisting of material points* (particles without spin).

(17.126) First we assume that $\Lambda = 0$ (notation (17.124)). One can show that the integral I_0 (notation (16.201)) is convergent only if Γ is a *past pointing* vector (the mass of the particles is assumed to be *positive*). We thus can choose an inertial reference frame $\mathcal{R}$ with respect to which the vector Γ is of the form $\mathcal{R}\begin{pmatrix}0\\ \varepsilon\end{pmatrix}$ with $\varepsilon < 0$ (notation (13.55)). It follows that Z is an infinitesimal time translation, which, as in (17.105), allows us to interpret ε by the formula

$$\varepsilon = -\frac{1}{kT}\,, \tag{17.127}$$

where T is the absolute temperature and k is Boltzmann's constant.

Let us describe a particle by its linear momentum $\mathbf{p}$ (13.79) and its position $\mathbf{r}$ (at $t = 0$, in the reference frame $\mathcal{R}$). Due to (13.60) and description (14.24) of particles without spin, we find the elementary probability (for a single particle) to be

$$\exp\left(-z - \frac{\sqrt{\|\mathbf{p}\|^2 + m^2}}{kT}\right) \varphi(d\mathbf{p})\,\varphi(d\mathbf{r})\,, \tag{17.128}$$

where φ is the Euclidean density and z is Planck's thermodynamic potential. Formula (16.203) gives

$$z = \log\left(4\pi V m^3 F\left(\frac{m}{kT}\right)\right)\,, \tag{17.129}$$

where V is the volume of the box containing the gas (*measured in the reference frame* $\mathcal{R}$) and F is the function defined by

$$F(x) = \int_0^\infty \mathrm{e}^{-x\cosh\varphi}\,[\sinh\varphi]^2 \cosh\varphi\, d\varphi\,; \tag{17.130}$$

it can also be described by

$$F(x) = \frac{K_2(x)}{x}\,, \tag{17.131}$$

where K_2 denotes a modified Bessel function.[409]

From (17.128) and (17.129) one can deduce the distributions of velocities and of compressibility, which replace the nonrelativistic distributions

[409] See Watson [38].

of Maxwell and Mariotte-Gay-Lussac. The average velocity[410] is zero, that is, the gas seems to be at rest in the reference frame $\mathcal{R}$.

(17.132) The function F can be approximated using formulæ that are valid at low temperature.[411] One then can show that the difference between the relativistic and nonrelativistic distributions are negligible as long as the quantity m/kT is large compared to unity. This is the case in all known examples (for hydrogen, m/k is close to 10^{13} degrees Kelvin, whereas the temperature in the center of the hottest stars seems to be less than 10^{10} degrees Kelvin). □

By changing the inertial reference frame, one can observe a "relativistic wind". The temperature T relative to the proper reference frame is obviously a relativistic invariant, defined by the formula

(17.133) $$T = \frac{1}{k\sqrt{\langle \Gamma, \Gamma \rangle}}.$$

How then can one define the "temperature T' relative to the new reference frame?" If we really want a definition which only involves the temporal component of the vector Γ and which reduces to the proper temperature in the case of the proper reference frame, we are led to the formula of Planck

(17.134) $$T' = T\sqrt{1 - v^2} \qquad (v = \text{the velocity of the "wind"}).$$

This formula has been contested recently by various authors. It seems to us that this dispute is a bit academic, since formula (17.134) is a mere definition, and a useless one at that. It suffices to note that the
(17.135) *temperature vector* Γ[412] defines at the same time the *proper temperature*, the *velocity of the wind*, and the *direction of the time flow*.

In the general case when $\Lambda \neq 0$, the gas can only be located at those points X of space-time E_4 where the vector

(17.136) $$\Gamma_X \equiv Z_{E_4}(X) \equiv \Lambda \cdot X + \Gamma$$

is *past pointing*. In a neighborhood of the point X, the velocity distribution is the same as in the preceding case with Γ_X as temperature vector.

---

[410]The velocity is expressed as a function of $\mathbf{p}$ by $\mathbf{v} = \frac{\mathbf{p}}{\sqrt{\|\mathbf{p}\|^2 + m^2}}$ (14.27).

[411]See Synge [35].

[412]The temperature vector defined by Van Dantzig, Bergmann, and Costa de Beauregard equals $-k\Gamma$.

We can describe for instance a *relativistic centrifuge* by choosing $\boldsymbol{\beta} = \boldsymbol{\gamma} = 0$ in definition (13.56) of Z. This gives us the temperature vector

$$\Gamma_X = \mathcal{R}\begin{pmatrix} \boldsymbol{\omega} \times \mathbf{r} \\ \varepsilon \end{pmatrix}, \tag{17.137}$$

which shows that the velocity of the wind is

$$\mathbf{v} = \boldsymbol{\omega}' \times \mathbf{r} \qquad (\boldsymbol{\omega}' = \frac{\boldsymbol{\omega}}{\varepsilon}). \tag{17.138}$$

We thus recover the field of velocities of a *nonrelativistic* rotating rigid body.[413]

Formula (17.137) indicates at the same time the existence of a new effect, namely, the *proper* temperature T of the fluid *is not constant*. It is given, as a function of its value T_0 on the axis of rotation, by

$$T = \frac{T_0}{\sqrt{1 - \|\mathbf{v}\|^2}}. \tag{17.139}$$

Statistical equilibria of photons

(17.140) Let us study a massless particle (14.29) in a reference frame $\mathcal{R}$. We know how to define the trajectory of the particle relative to $\mathcal{R}$ (14.33). The point $X = \begin{pmatrix} \mathbf{r} \\ 0 \end{pmatrix}$ on this trajectory defines the "position" $\mathbf{r}$ of the particle at time $t = 0$. One can specify every motion x of the particle by $\mathbf{r}$ and the *linear momentum* $\mathbf{p}$ (13.79). Applying (13.59), (13.87), (13.90), (14.29♡, ♯, ♠), and (14.33) we find

$$I \equiv \mathcal{R}\begin{pmatrix} \eta\,\mathbf{p} \\ \|\mathbf{p}\| \end{pmatrix} \tag{17.141}$$

$$J \equiv \frac{1}{2\|\mathbf{p}\|^2}\,\mathcal{R}\begin{pmatrix} \eta\,\mathbf{p} \\ -\|\mathbf{p}\| \end{pmatrix} \tag{17.142}$$

$$\Omega \equiv \mathcal{R}\begin{pmatrix} j\left(\frac{\eta}{\|\mathbf{p}\|}\mathbf{p}\right) & 0 \\ 0 & 0 \end{pmatrix}\mathcal{R}^{-1} \tag{17.143}$$

$$P \equiv \mathcal{R}\begin{pmatrix} \mathbf{p} \\ \eta\,\|\mathbf{p}\| \end{pmatrix}. \tag{17.144}$$

[413] Recall that there does not exist a relativistic definition of an accelerated rigid body (see (13.69)).

Substituting the above in (14.29♣) gives

$$\sigma(\delta x)(\delta' x) \equiv \langle \delta\mathbf{p}, \delta'\mathbf{r}\rangle - \langle \delta'\mathbf{p}, \delta\mathbf{r}\rangle - \frac{\chi s\eta}{\|\mathbf{p}\|^3}\langle \mathbf{p}, \delta\mathbf{p}\times\delta'\mathbf{p}\rangle\,, \tag{17.145}$$

where η is the sign of the energy, χ the helicity, and s the spin.

To calculate the symplectic density φ (16.94), it is convenient to note that — for every basis S — $\varphi(S)$ is the square root of the determinant of the matrix with components σ_{jk} of the Lagrange form σ with respect to the basis S. Using (17.145) we find

$$\varphi(dx) \equiv \varphi(d\mathbf{p})\,\varphi(d\mathbf{r})\,, \tag{17.146}$$

where the φ on the right hand side describes the Euclidean density on $\mathbf{R}^3$. We see that (17.146) is the same as if $\begin{pmatrix}\mathbf{p}\\ \mathbf{r}\end{pmatrix} \mapsto x$ were a canonical chart (which it is not), and that the numbers χ, s, and η do not intervene in this formula.

Notice that the energy E of the particle is given by Einstein's formula

$$E = \eta\,\|\mathbf{p}\|\,, \tag{17.147}$$

which follows immediately by comparing (13.59) and (17.144), or from (14.34).

We now can study the statistical equilibrium of such a particle at temperature T. Employing principle (17.123) and using the notation (17.124), (17.126), and (17.127) we find the "elementary probability" to be

$$\exp\left(-z - \frac{\eta\,\|\mathbf{p}\|}{kT}\right)\varphi(d\mathbf{p})\,\varphi(d\mathbf{r}), \tag{17.148}$$

where z is the thermodynamic potential which we will calculate later on in (17.155). We see that the integral converges only if the particle is in a *box* of finite volume V and if the numbers η and T have *the same sign*.[414] Since we assume of course that $T > 0$, it follows that the existence of a thermodynamic equilibrium *excludes the existence of particles with negative energy*. On the other hand, the helicity is arbitrary, and the particles can be polarized to the right as well as to the left.

(17.149)



If we now replace η by 1, we can easily deduce from (17.148) the probability density of the energy as[415]

$$\frac{1}{2[kT]^3}\,\mathrm{e}^{-E/kT}\,E^2\,dE \qquad (E>0)\,. \tag{17.150}$$

□

[414] Note the consistency of this result with (14.77): time reversal, which changes the sign of the energy, also changes the sign of the temperature (17.135).

[415] The energy E and the velocity $\mathbf{p}/E$ are *independent* random variables (the velocity being distributed uniformly on the unit sphere).

(17.151) Let us try to describe the radiation inside a box kept at temperature T (*black body radiation*) as a thermodynamic equilibrium of photons, as discussed above. As in the case of an ideal gas, we can treat this problem using scattering theory, which allows us to ignore the details of the interactions between the photons and the walls of the box. But there is an essential difference with an ideal gas, namely, one cannot pump light into the box, that is, *the number of photons in the box is given by nature.*

(17.152) According to which probability distribution? We will hypothesize that *light* is a dynamical system whose manifold of motions U_Φ is *composite* in the sense of (16.8). U_Φ will contain not only the manifold U_1 (of dimension 6) of motions of a single photon (14.29), but also all manifolds U_N (of dimension $6N$) describing the motions of N photons, which were studied in (15.45).

We even include — by hypothesis — the manifold U_0 (of dimension 0) consisting of a single point and describing the *photon vacuum.* We will call the composite manifold U_Φ the *Fock manifold.*

(17.153) Since each pure manifold U_N is symplectic, we know how to define its Liouville measure (for $N = 0$ we choose the measure defined in (16.70)). We define the Liouville measure μ on U_Φ by (notation (16.100))

$$\mu_{U_N} = \text{the Liouville measure of } U_N \qquad (\forall N)\,,$$

which allows us to apply the Gibbs principle (17.24) or (17.123). Taking into account that the manifold U_1 has two connected components corresponding to the two values of the helicity, we find

(17.154) $$I_0 = \sum_{N=0}^{\infty} \frac{[16\pi\, V k^3 T^3]^N}{N!}\,.$$

Hence — recognizing the power series of the exponential function — we get

(17.155) $$z = 16\pi\, V[kT]^3\,.$$

The formulæ (17.26) allow us to deduce from (17.155) that the average value $\tilde{E}$ of the *total energy of the photons contained in the box* is

(17.156) $$\tilde{E} = -\frac{\partial z}{\partial[1/kT]} = 48\pi\, V[kT]^4$$

and that the *entropy* is

(17.157) $$s = z + \frac{\tilde{E}}{kT} = 64\pi\, V[kT]^3\,.$$

As in (17.54) we can calculate the average energy as a function of the entropy and the volume

$$\tilde{E} = \frac{3}{16}\,\pi^{-1/3}\,s^{4/3}\,V^{-1/3}\,. \tag{17.158}$$

Hence

$$\delta Q = kT\,\delta s \qquad\qquad \delta W = -16\pi\,[kT]^4\,\delta V\,, \tag{17.159}$$

where Q is the heat and W the work (formulæ (17.57), (17.58)). Thus we see *radiation pressure* appearing (compare with (17.61)), which is given by

$$p = 16\pi\,[kT]^4 = \frac{1}{3}\,\frac{\tilde{E}}{V}\,. \tag{17.160}$$

Note that the *number N of photons is a random variable*. The probability associated to each value of N is given by[416]

$$p_N = \frac{z^N}{N!}\,\mathrm{e}^{-z}\,. \tag{17.161}$$

It follows that the average value of N is given by

$$\widetilde{N} = z\,. \tag{17.162}$$

One can also calculate the probability $p_{n,n'}$, namely that there are $n + n'$ photons, of which n are polarized to the left and n' to the right:

$$p_{n,n'} = \left[\frac{z}{2}\right]^{n+n'} \frac{\mathrm{e}^{-z}}{n!\,(n')!}\,. \tag{17.163}$$

Referring back to (17.30), we see that the above formulæ are not dimensionally homogeneous, that is, they depend on the chosen unit of action. Since one can recognize (17.156) as the Stefan-Boltzmann law

$$\frac{\tilde{E}}{VT^4} = \text{constant}\,, \tag{17.164}$$

we can use the experimental value of Stefan's constant to choose the unit of action, which gives the value

$$6.45 \cdot 10^{-27}\,\mathrm{g\ cm^2\ sec^{-1}}\,. \tag{17.165}$$

With this choice, formula (17.160), which gives the radiation pressure, also agrees with experiments.

[416]The probability distribution (17.161) is called a *Poisson* distribution.

Thus we might presume that the other formulæ above are valid as well and we might calculate for example the average number of photons contained in a box of one liter at room temperature (300° Kelvin):

(17.166) $$\widetilde{N} = z \simeq 5 \cdot 10^{11} ,$$

the probability that all these photons are polarized to the left:

(17.167) $$p = \mathrm{e}^{-z/2} \simeq 10^{-10^{11}} ,$$

etcetera.

(17.168) Unfortunately formula (17.150), which gives the energy distribution of the individual photons,[417] is not in agreement with experiment, which follows very closely the law proposed by Planck. Figure (17.IV) gives the probability density of the variable $u = E/kT$ for the statistics described above ($\frac{1}{2}\, u^2 \mathrm{e}^{-u}$) and for Planck's law $\left(\dfrac{15}{\pi^4}\, \dfrac{u^3}{[\mathrm{e}^u - 1]}\right)$.[418]

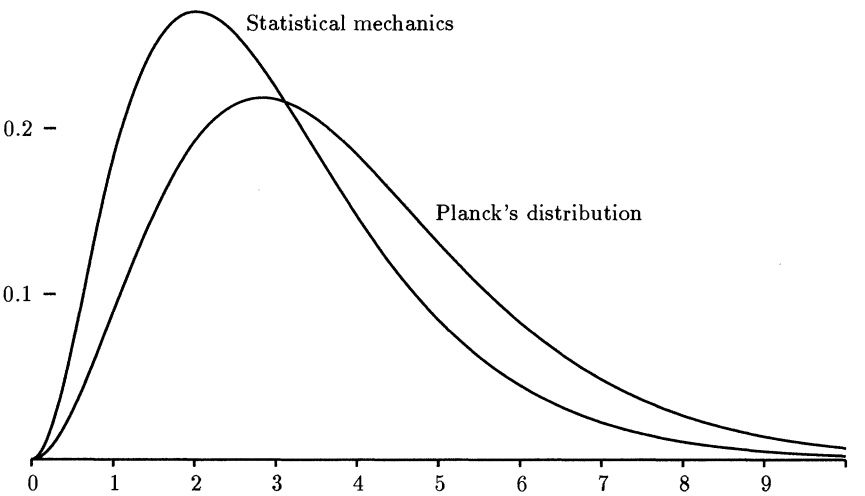


Fig 17.IV

[417]The photoelectric effect gives, in principle, a way to measure directly the energy of a photon. Quantum mechanics allows it to be measured indirectly by means of the formula $E = h\nu$, where h is Planck's constant and ν is the frequency of the associated light. It follows that figure (17.IV) is a *spectrum*.

[418]When Planck proposed this formula (December 1900), the notion of a photon did not yet exist (it was invented by Einstein in 1905). Up to that time physics only proposed hybrid models for black body radiation, using at the same time classical statistical mechanics and Maxwell's equations (which represent already quantized photons). For instance, the formula of Rayleigh-Jeans which leads to a divergent integral ("the ultraviolet catastrophe").

(17.169) As in the case of ideal gases or solids, comparing statistical mechanics with experiments (without calling into question the classical models proposed for the particles)[419] thus shows the necessity of a quantum mechanical treatment of phenomena.

[419]If we would not have postulated the *indiscernibility of photons* in the sense of (15.47), the term $N!$ would not have appeared in the denominator of (17.154). This would have lead to replacing the Stefan-Boltzmann law (17.156) by $\widetilde{E} = \frac{48\pi V[kT]^4}{1-16\pi V[kT]^3}$, which is completely unacceptable because it leads to an infinite energy for temperatures of the order of $1\,^\circ$ K (when $V = 1$ cm^3).

Chapter V

A Method of Quantization

§18. Geometric quantization

Prequantum manifolds †

(18.1) We will call the multiplicative group of complex numbers of modulus 1 a *torus* and we will denote it by T. T is a compact Lie group of dimension 1 whose Lie algebra is the set of purely imaginary complex numbers. For $z \in T$ there exists a real number s such that $z = \exp(is)$, where s is defined modulo an integer multiple of 2π. □

Let Y be a Hausdorff manifold and let $\xi \mapsto \varpi$ be a differentiable field of 1-forms on Y (Fig. 18.I). If the following properties are satisfied

(18.2)
$$\diamondsuit \qquad \dim(\ker(\mathrm{d}\varpi)) \equiv 1$$
$$\heartsuit \qquad \dim(\ker(\varpi) \cap \ker(\mathrm{d}\varpi)) \equiv 0\,,$$

the field $\xi \mapsto \varpi$ is said to define a *contact structure* on Y.

If we define

(18.3)
$$\sigma \equiv \mathrm{d}\varpi\,,$$

then σ is a 2-form whose rank, as we know, is even (8.9). Condition (18.4) (18.2$\diamondsuit$) then implies that *the dimension of Y is odd* (at every point ξ).

(18.5) Since $\dim(\ker(\sigma))$ is constant, Y is a *presymplectic* manifold (9.8); its *leaves* are curves (lines of force).

Now let us suppose that these curves are closed.[420] Using the calculus of variations applied to the Lagrangian $\varpi\,\dot{\xi}$, one can show that the action integral

(18.6)
$$a = \oint \varpi\,\frac{d\xi}{ds}\,ds\,,$$

integrated over the leaf passing through ξ_0, satisfies $\partial a/\partial\xi_0 = 0$. Thus it is constant on every connected component of Y (15.71), (15.72). We will say that Y is a prequantum manifold if $a \equiv 2\pi$.

We are now going to formulate the definition of a prequantum manifold in a more precise and suitable way by introducing the *vector* $\mathrm{i}_Y(\xi)$,

† Editors' note: The original French version called them "quantum manifolds". In consultation with the author we decided to change it to "prequantum manifold" in order to adhere to standard terminology and in order to avoid confusion with quantum groups and related theories

[420] In the usual sense of the word, that is, homeomorphic to T.

defined (using the conditions (18.2)) by the two linear equations

$$\begin{aligned} \sigma(\mathrm{i}_Y(\xi)) &\equiv 0 \\ \varpi(\mathrm{i}_Y(\xi)) &\equiv 1\,. \end{aligned} \tag{18.7}$$

The first equation shows that $\mathrm{i}_Y(\xi)$ is tangent to the line of force through ξ; the second equation specifies its magnitude.

Using a canonical chart (10.9), one can show that i_Y is a *differentiable* vector field. If we select a point ξ_0 on V and consider the point $\xi \equiv \exp(s\,\mathrm{i}_Y)(\xi_0)$ (notation (3.1)), then the map $s \mapsto \xi$ satisfies $\partial\xi/\partial s \equiv \mathrm{i}_Y(\xi)$. From this we deduce that ξ varies over the line of force passing through ξ_0. Using this parameter s to calculate the action integral (18.6), we see that we have made a "complete revolution" when $s = 2\pi$. In other words, the curve $s \mapsto \xi$ is *periodic* of period 2π. We recapitulate this discussion as follows.

(18.8) A Hausdorff manifold Y will be called a *prequantum manifold* if

a) There exists a differentiable field of 1-forms $\xi \mapsto \varpi$ on Y which defines a contact structure (18.2) on Y, that is,

$\diamondsuit$ $\qquad \dim(\ker\sigma)) \equiv 1 \qquad [\sigma \equiv \mathrm{d}\varpi]$

$\heartsuit$ $\qquad \dim(\ker(\varpi) \cap \ker(\sigma)) \equiv 0\,.$

b) The torus T *acts* on Y (6.4) in such a way that [421]

$\bowtie$ $\qquad \underline{z}_Y(\xi) = \xi \iff z = 1 \qquad [z \in T]$

$\clubsuit$ $\qquad \sigma(\mathrm{i}_Y(\xi)) \equiv 0$

$\spadesuit$ $\qquad \varpi(\mathrm{i}_Y(\xi)) \equiv 1\,.$ □

One should also note the formula

$$z = \mathrm{e}^{is} \in T \implies \underline{z}_Y = \exp(s\,\mathrm{i}_Y) \tag{18.9}$$

and the fact that *the structure of a prequantum manifold is entirely defined by the manifold structure of Y and the 1-form ϖ*, because the action of T follows from ϖ by the formulæ (18.8$\clubsuit$), (18.8$\spadesuit$), and (18.9).

Using the fact the the lines of force (which are the *orbits* of T) are compact, one can show that the characteristic foliation of Y is *sectionable*. Consequently, the set U of lines of force admits the structure of a *symplectic manifold* (theorem (9.10)) whose Lagrange form σ is defined by

$$\sigma(\delta x)(\delta' x) \equiv \sigma(\delta\xi)(\delta'\xi)\,, \tag{18.10}$$

[421] For the notation $\underline{z}_Y$: see (6.4).

where $\xi \mapsto x$ denotes the projection of Y onto U (Fig. 18.I).

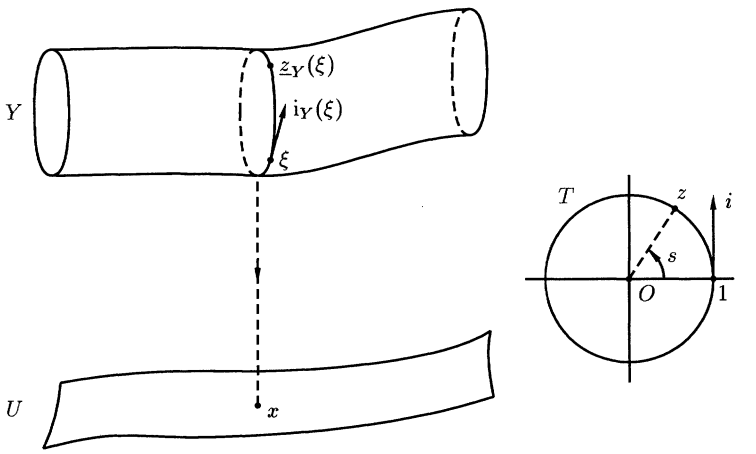


Fig 18.I

(18.11) We will say that U is the *base* of the prequantum manifold Y;[422] the dimension of U is *one less* than the dimension of Y.

(18.12) One can show that the manifold U is Hausdorff and that the map $\xi \mapsto x$ is proper (16.26). More precisely, the inverse image under $\xi \mapsto x$ of a compact subset of U is a compact subset of Y.

Prequantization† of a symplectic manifold

(18.13) Conversely, let U be a *Hausdorff symplectic manifold*. We will call *prequantization of* U the construction of a prequantum manifold Y and a *symplectomorphism* A between the base of Y and U (Fig. 18.II). If we write $x \equiv A(\text{orbit}(\xi))$ and $P = [\xi \mapsto x]$, then

Y is a prequantum manifold

P is a differentiable map from Y to U

(18.14) $\ker(D(P)(\xi))$ is generated by $i_Y(\xi)$ $\qquad \forall \xi \in Y$

$x \in U \implies P^{-1}(x)$ is an orbit of T (acting on Y)

$\sigma_Y(\delta\xi)(\delta'\xi) \equiv \sigma_U(\delta x)(\delta' x) \qquad [x \equiv P(\xi)]$.

[422] Y is called a *principal fiber bundle* with *structure group* T and *base* U. But here we will not use the general theory of fiber bundles.

†Editors' note: As for prequantum manifolds, the original French speaks of "quantization". We have changed to "prequantization" in order to adhere to standard terminology and in order to be consistent with the change to prequantum manifold.

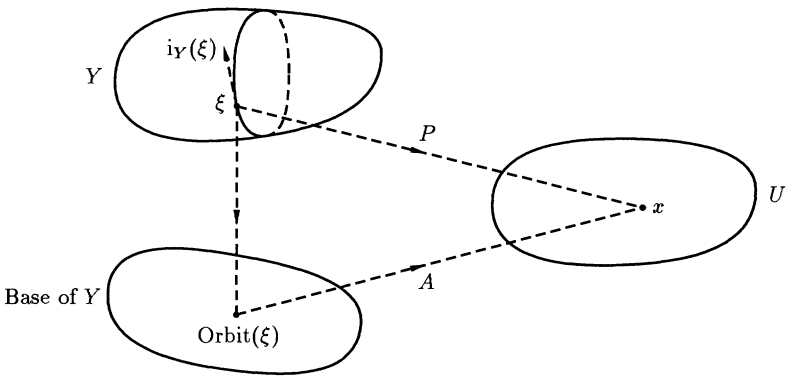


Fig 18.II

(18.15) On the other hand, if we have constructed a prequantum manifold Y and a map P satisfying (18.14), it follows that there exists a symplectomorphism A from the base of Y onto U such that $P(\xi) \equiv A(\text{orbit}(\xi))$. Thus, *prequantizing a Hausdorff symplectic manifold U is the same as constructing a pair (Y, P) satisfying the axioms* (18.14).

(18.16) When the above problem has a solution, we will say the U is *prequantizable*. Note that every symplectic manifold which is symplectomorphic to a prequantizable manifold is prequantizable. □

(18.17) It is clear that *every open set U_0 in a prequantizable manifold U is prequantizable*. In fact, Y_0, the inverse image of U_0 under P, and the map P_0, the restriction of P to Y_0, provide a prequantization of U_0. □

(18.18) If a manifold U is the union of disjoint open sets, one thus can prequantize U by separately prequantizing these open sets. In particular, if U is a *composite* manifold (16.8), one should separately prequantize the open sets where U has a fixed dimension (16.9).

(18.19) THEOREM: Let U be a Hausdorff symplectic manifold and let $\{U_j\}$ be an open cover of U such that [423]

$$\begin{cases} \text{every } U_j \text{ is simply connected;} \\ \text{the Lagrange form } \sigma \text{ on } U \text{ admits a potential } \varpi_j \text{ on } U_j. \end{cases}$$

If U is *prequantizable*, there exist differentiable maps $x \mapsto z_{jk}$ from each of the (nonempty) $U_j \cap U_k$ to the torus T such that

$$\diamondsuit \qquad \varpi_k(\delta x) - \varpi_j(\delta x) \equiv \frac{\delta z_{jk}}{i\, z_{jk}} \qquad \forall\ \delta x$$

[423] It suffices to take an open cover by cells (see (1.47), (1.54), and (4.37)).

and

$$\begin{array}{lll} & x \in U_j & \Longrightarrow \quad z_{jj} = 1 \\ \heartsuit & x \in U_j \cap U_k & \Longrightarrow \quad z_{jk} z_{kj} = 1 \\ & x \in U_j \cap U_k \cap U_\ell & \Longrightarrow \quad z_{jk} z_{k\ell} z_{\ell j} = 1 \,. \end{array}$$

PROOF: Let us only sketch how to construct z_{jk}. In the open set Y_j, the inverse image of U_j under the projection $\xi \mapsto x$, one can check that the 1-form ϖ'_j defined by

$$\varpi'_j(\delta\xi) \equiv \varpi_j(\delta x) - \varpi(\delta\xi) \tag{18.20}$$

has zero exterior derivative and that its kernel has the same dimension as U_j. From the fact that U_j is simply connected, one can deduce that the leaves of the characteristic foliation of ϖ'_j are transverse sections of Y_j. Let Φ_j be one of them. For $x \in U_j \cap U_k$ we can define z_{jk} by

$$\underline{z_{jk}}(\xi_j) \equiv \xi_k \,, \tag{18.21}$$

where ξ_j and ξ_k are the points of Φ_j and Φ_k that project to x. □

On the other hand, the above conditions *suffice* for U to be prequantizable. More generally,[424] we have

(18.22) THEOREM: Let U be a *Hausdorff symplectic manifold* and $\{U_j\}$ an *open cover* of U. Assume that there exist potentials ϖ_j for the 2-form σ defined on U_j, and differentiable maps $x \mapsto z_{jk}$ from $U_j \cap U_k$ to T satisfying (18.19$\diamondsuit$ and $\heartsuit$). Then U is prequantizable.

PROOF: One can verify that the relation $\sim$ defined on the set of triples

$$\begin{pmatrix} x \\ j \\ z \end{pmatrix} \qquad x \in U_j \,,\ z \in T \,, \tag{18.23}$$

by

$$\begin{pmatrix} x \\ j \\ z \end{pmatrix} \sim \begin{pmatrix} x' \\ j' \\ z' \end{pmatrix} \quad \Longleftrightarrow \quad x' = x \,,\ \frac{z'}{z} = z_{j'j} \tag{18.24}$$

is an equivalence relation (due to (18.19$\heartsuit$)). If we call $A_j\begin{pmatrix} x \\ z \end{pmatrix}$ the *equivalence class* of $\begin{pmatrix} x \\ j \\ z \end{pmatrix}$ under the relation $\sim$, one can verify that the set Y

[424] The U_j are no longer supposed to be simply connected.

of equivalence classes admits the structure of a Hausdorff manifold such that the maps A_j are diffeomorphisms from $U_j \times T$ to Y. On account of equation (18.19$\diamondsuit$), we can define a field $\xi \mapsto \varpi$ of 1-forms on Y by

$$\varpi(\delta\xi) \equiv \frac{\delta z}{iz} + \varpi_j(\delta x) \qquad \left[\xi \equiv A_j \begin{pmatrix} x \\ z \end{pmatrix}\right]. \tag{18.25}$$

We define an action of the torus T on Y by

$$\underline{z'}_Y \left(A_j \begin{pmatrix} x \\ z \end{pmatrix} \right) \equiv A_j \begin{pmatrix} x \\ z'z \end{pmatrix}. \tag{18.26}$$

One can check that these definitions satisfy the axioms (18.8) and thus that Y is a *prequantum manifold*. Finally note that Y and the map P defined by

$$P \left(A_j \begin{pmatrix} x \\ z \end{pmatrix} \right) \equiv x \tag{18.27}$$

satisfy the axioms (18.14). □

We thus have an algorithm to recognize whether a manifold is prequantizable, and to carry out its prequantization. We now treat some examples.

(18.28) Let us start by noting that the conditions for prequantizability (18.19) are not invariant if we multiply the 2-form σ by an arbitrary scalar. The formulæ (13.13) of dimensional analysis applied to examples borrowed from dynamics show that prequantization depends on the chosen unit of *action*. We will choose the value (denoted by $\hbar$)

$$1.0544\,10^{-27} \text{ g cm}^2 \text{ sec}^{-1}, \tag{18.29}$$

which is *Planck's constant* divided by 2π.[425] A physical interpretation of this choice will be given later on. □

Prequantization of a symplectic manifold admitting a potential

If U is a symplectic manifold admitting a potential, that is (9.14) if there exists a 1-form ϖ_U on U such that $\sigma_U \equiv \mathrm{d}\varpi_U$, then we can apply theorem (18.22) with the (open) cover $\{U_j\}$ consisting of the single set U. This immediately gives

[425] This is the same as saying that the integral a (18.16) equals Planck's constant.

(18.30) THEOREM: Every Hausdorff symplectic manifold U *admitting a potential* is *prequantizable*. Moreover, one can take the prequantum manifold Y to be the *direct product* $U \times T$ with

$$\begin{cases} \varpi\left(\delta\begin{pmatrix} x \\ z \end{pmatrix}\right) \equiv \dfrac{\delta z}{iz} + \varpi_U(\delta x)\,^{426} \\ \underline{z'}_Y\begin{pmatrix} x \\ z \end{pmatrix} \equiv \begin{pmatrix} x \\ z'z \end{pmatrix} \\ P\begin{pmatrix} x \\ z \end{pmatrix} \equiv x\,. \end{cases}$$

□

(18.31) This result applies especially to the case when U is a *symplectic vector space* (9.4)[427] (and thus in particular to the space of motions of a *nonrelativistic material point* (see (13.22)) or of a *harmonic oscillator* (12.150)), as well as to the *Kepler problem* and more generally to the *n-body problem*, since we know a potential (12.156) for these problems.

Prequantization of a sphere S_2

In (11.40) we have seen that the sphere S_2 is equipped with a symplectic structure if we define (λ being a positive constant)

$$\sigma(\delta\mathbf{x})(\delta'\mathbf{x}) = \lambda\langle\mathbf{x}, \delta\mathbf{x} \times \delta'\mathbf{x}\rangle \qquad [\mathbf{x} \in \mathbf{R}^3\,,\ \|\mathbf{x}\| = 1]\,. \tag{18.32}$$

The two charts F_+ and F_- defined in (1.11) have as images the open sets U_+ and U_- (S_2 without its north and south pole, respectively), which are cells (since F_+ and F_- are defined on $\mathbf{R}^2$). On $U_\pm$ we can define potentials $\varpi_\pm$ of σ by

$$\varpi_\pm(\delta\mathbf{x}) = \pm 2\lambda\frac{\alpha\,\delta\beta - \beta\,\delta\alpha}{1 + \alpha^2 + \beta^2} \qquad \left[\mathbf{x} \equiv F_\pm\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right]. \tag{18.33}$$

We now apply the rule (18.19◇). *If* the sphere is prequantizable, there exists a map $x \mapsto z$[428] from $U_+ \cap U_-$ to T such that

$$\varpi_+(\delta\mathbf{x}) - \varpi_-(\delta\mathbf{x}) = \frac{\delta z}{iz}\,. \tag{18.34}$$

[426] Note that ϖ *is not* the inverse image of ϖ_U under the projection P.

[427] If x varies over a symplectic vector space then $\varpi = \frac{1}{2}\sigma(x)$ is indeed a potential for σ.

[428] We write z for z_{-+}.

On the *equator* $\alpha \equiv \cos\varphi$, $\beta \equiv \sin\varphi$, and F_+ and F_- coincide. From (18.33) and (18.34) it then follows that

$$2\lambda\,\delta\varphi = \frac{\delta z}{iz},$$

and hence $z = z_0\,e^{2i\lambda\varphi}$. Since z takes the same value for $\varphi = 0$ and $\varphi = 2\pi$, we see that

(18.35) The sphere S_2 is not prequantizable if 2λ is not an integer (notation (18.32)).

(18.36) In passing, this result shows that *the sphere is not a (symplectic) manifold admitting a potential.* If it were, it would be, because of (18.30), prequantizable for every value of λ.

(18.37) We now are going to show that S_2 is indeed prequantizable if 2λ is an integer, not using theorem (18.22), but by a direct method (we will see later on why these procedures are equivalent).

Let us denote by Y the 3-dimensional sphere S_3, seen as a submanifold of $\mathbf{C}^2$ defined by the equation (notation (6.53m))

$$\overline{\xi}\xi \equiv 1 \qquad \left[\,\xi \equiv \begin{bmatrix} {}^1z \\ {}^2z \end{bmatrix},\ \overline{\xi} \equiv \begin{bmatrix} {}^1\overline{z} & {}^2\overline{z} \end{bmatrix};\ {}^1z, {}^2z \in \mathbf{C}\,\right] \tag{18.38}$$

equipped with the *real* 1-form ϖ:

$$\varpi(\delta\xi) \equiv \frac{\overline{\xi}\,\delta\xi}{i}. \tag{18.39}$$

One can easily verify that the axioms (18.8) of prequantum manifolds are satisfied with

$$\sigma(\delta\xi)(\delta'\xi) \equiv 2\,\mathrm{Re}\left(\frac{\delta\overline{\xi}\,\delta'\xi}{i}\right) \qquad [\mathrm{Re} = \text{real part}], \tag{18.40}$$

$$\underline{z}_Y(\xi) \equiv \xi z \qquad z \in T, \tag{18.41}$$

and

$$\mathrm{i}_Y(\xi) \equiv \xi\, i. \tag{18.42}$$

Let us now introduce the *Pauli spin matrices*

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{18.43}$$

and let us define

$${}^j[P(\xi)] \equiv \overline{\xi} \cdot \sigma_j \cdot \xi \qquad [j = 1,2,3]. \tag{18.44}$$

In this way we obtain a map P from Y onto the sphere S_2.[429] One can verify that (18.14) holds if we define the symplectic structure on S_2 by formula (18.32) with $\lambda = \frac{1}{2}$. □

(18.45) In this way we have prequantized the sphere for $\lambda = \frac{1}{2}$; the prequantization when 2λ is an arbitrary integer will follow from (18.47). □

Prequantization by "fusion"[430]

(18.46) Let (Y, P) be a prequantization of a symplectic manifold U (notation (18.15)) and let n be a positive or negative integer (Fig. 18.III).

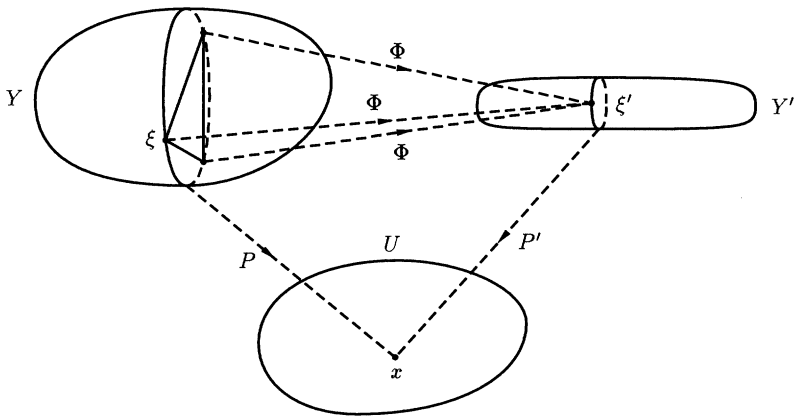


Fig 18.III

The set G of $\underbrace{e^{2\pi i p/n}}_{Y}$ $(p = 1, 2, \ldots, |n|)$ is a *discrete group* of diffeomorphisms of Y (1.44). If $\xi' \equiv \Phi(\xi)$ is the orbit of ξ under G,[431] we know that the quotient $Y' = Y/G$ admits the structure of a manifold and that the triple (Y, G, Φ) defines a *covering space* of Y' (1.45).

One can check that the 1-form $n\varpi$ is the inverse image under Φ of a 1-form ϖ' on Y', that ϖ' gives Y' the structure of a prequantum manifold,

---

[429] By writing $\xi = \begin{pmatrix} a+bi \\ c+di \end{pmatrix}$, we find that $P(\xi) = \begin{pmatrix} 2(ac+bd) \\ 2(ad-bc) \\ a^2+b^2-c^2-d^2 \end{pmatrix}$, which allows us to make a direct verification of (18.14). For a more sophisticated argument, one should refer to the general theory of spinors (see for example Souriau [32], §44.)

[430] We borrow this term from Louis de Broglie (see for example Visconti [37]). There indeed exists an analogy between the construction of de Broglie and (18.46), though not an obvious one.

[431] $\Phi(\xi)$ can be seen as the set of vertices of a "regular polygon" with $|n|$ sides. These vertices are thus *identified* when passing to the quotient.

that the map P can be factored as $P' \circ \Phi$, and that the pair (Y', P') is a prequantization of the manifold U *equipped with the symplectic structure obtained by replacing the Lagrange form* σ_U *by* $n\,\sigma_U$. We thus obtain:

(18.47) If a symplectic manifold U is prequantizable, it remains prequantizable if we multiply its Lagrange form by a nonzero integer.

(18.48) EXAMPLE: The sphere S_2 — *for* $\lambda = 1$ — is prequantized by identifying two antipodal points, ξ and $-\xi$, of the manifold Y (18.38). The manifold Y' thus obtained is called the *real projective space of dimension* 3.

Prequantization of a direct product (Fig. 18.IV)

Let U_j be symplectic manifolds $(j = 1, 2, \ldots, n)$, let (Y_j, P_j) be a prequantization of U_j and let ϖ_j denote the prequantum 1-form on Y_j. On the direct product Y of the Y_j we define a 1-form ϖ by

$$\varpi(\delta\xi) \equiv \sum_j \varpi_j(\delta\xi_j) \qquad \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}. \tag{18.49}$$

It can be verified that the *characteristic foliation* of the 1-form ϖ associates to each ξ the $(n-1)$-dimensional vector space defined by

$$\delta\xi_j = \lambda_j \, \mathrm{i}_{Y_j}(\xi_j) \quad \text{(do not sum)} \qquad [\lambda_j \in \mathbf{R}\,,\ \sum_j \lambda_j = 0]\,. \tag{18.50}$$

Thus we can describe the corresponding leaf as the orbit of the group action

$$\xi_j \to \underline{z_j}_{Y_j}(\xi_j) \qquad [z_j \in T\,, \quad z_1 z_2 \cdots z_n = 1]\,. \tag{18.51}$$

It follows that this foliation is sectionable and that the orbit ξ' of ξ under the action (18.51) varies over the quotient manifold Y'. We know that ϖ is the inverse image of a 1-form ϖ' on Y' (5.25). If P' denotes the map $\xi' \mapsto x \equiv \begin{pmatrix} P_1(\xi_1) \\ \vdots \\ P_n(\xi_n) \end{pmatrix}$, one can verify that (Y', P') *prequantizes the* (18.52) *direct product* $U = U_1 \times \cdots \times U_n$ (equipped with its product symplectic structure (9.7)).

(18.53) Conversely, *if a direct product of symplectic manifolds is prequantizable, then so is each of its factors.* This is a particular case of the following theorem.

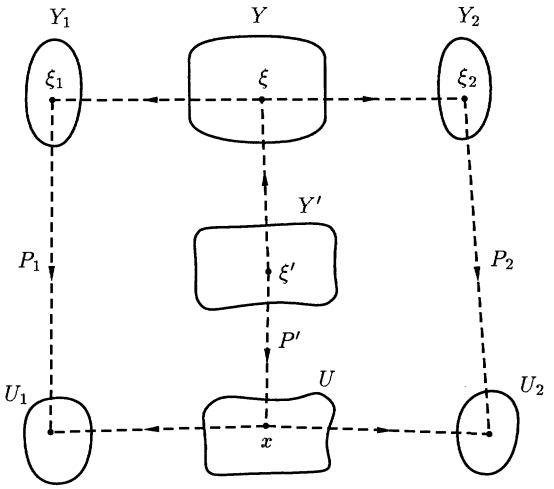


Fig 18.IV

(18.54) THEOREM: Let U be a prequantizable manifold. If U_0 is a submanifold of U such that

U_0 is defined by an equation (hypothesis (1.40)), and
U_0 is symplectic for the induced structure (9.29),

then U_0 is prequantizable.

PROOF: Let $A(x) = y_0$ be the defining equation of U_0 (notation (1.40)) and let (Y, P) be a prequantization of U. One can verify that $A(P(\xi)) = y_0$ is the defining equation of a submanifold Y_0 of Y. The inverse image ϖ_0 of ϖ under the inclusion of Y_0 in Y is a 1-form. One can verify that ϖ_0 gives Y_0 the structure of a prequantum manifold (the T-orbits in Y_0 are the T-orbits in Y), and that (Y_0, P_0) is a prequantization of U_0, with $P_0 = P \circ 1_{Y_0}$. □

(18.55) EXAMPLE: It follows from formula (14.48) that the space of motions of a *nonrelativistic particle with spin* is the symplectic direct product of the space of motions of a material point (prequantizable according to (18.31)) and a sphere S_2 (with the symplectic structure (18.32), with $\lambda = -s$). The result (18.37) and the preceding construction (18.52) show that:

(18.56) A nonrelativistic particle with spin is prequantizable if (and only if) its spin is of the form

$$\diamondsuit \qquad s = \frac{n}{2}, \qquad n \text{ an integer.}^{432} \qquad \square$$

Because of convention (18.29), since spin has the dimension of an action A ((13.13) and (14.50)), formula (18.56♢) can also be written as

$$s = n\frac{\hbar}{2}, \qquad n \text{ an integer.} \tag{18.57}$$

We are now going to extend the above results to *relativistic* particles.

(18.58) THEOREM: A relativistic particle without spin is prequantizable (whatever its mass).

PROOF: Indeed, according to (14.24♣) and (14.24e), we see that the Lagrange form on the space of motions U admits a potential $\varpi_U(\delta x) = m\overline{X}\,\delta I$. It now suffices to apply (18.30). □

Prequantization of a relativistic particle with spin $\frac{1}{2}$

We are going to use *Dirac spinors*; they can be defined as elements ψ of a *complex 4-dimensional Euclidean space* $\mathcal{E}$ (definition (6.51)) having a basis $\mathcal{S}$ such that

$$\overline{\mathcal{S}} \cdot \mathcal{S} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{bmatrix}. \tag{18.59}$$

It will be convenient to write

$$\psi \equiv \mathcal{S}\begin{pmatrix} \psi' \\ \psi'' \end{pmatrix}, \tag{18.60}$$

where ψ' and ψ'' are elements of $\mathbf{C}^2$ (called *half-spinors* or *two-component spinors*). Formula (18.59) can then be written as

$$\overline{\psi}\psi \equiv \mathrm{Re}\,(\overline{\psi'} \cdot \psi''). \tag{18.61}$$

Let us introduce the five operators

$$\begin{aligned} \gamma_j &= \mathcal{S}\begin{bmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{bmatrix}\mathcal{S}^{-1} \qquad (j = 1,2,3;\ \sigma_j = \text{the Pauli spin matrices (18.43)}) \\ \gamma_4 &= \mathcal{S}\begin{bmatrix} 0 & \underline{1} \\ \underline{1} & 0 \end{bmatrix}\mathcal{S}^{-1} \qquad (18.62) \\ \gamma_5 &= \mathcal{S}\begin{bmatrix} -\underline{i} & 0 \\ 0 & +\underline{i} \end{bmatrix}\mathcal{S}^{-1}. \end{aligned}$$

[432] As in the case of the sphere (18.36), it follows that the manifold of motions *does not admit a potential*.

Let us also choose a Lorentz frame $\mathcal{R}$ (13.41) for the Minkowski space E and let us define

$$\gamma(\mathcal{R}(x)) = \sum_{j=1}^{4} \gamma_j \, {}^j x \qquad \forall x \in \mathbf{R}^4 , \tag{18.63}$$

which associates to every vector $X = \mathcal{R}(x)$ of E a linear operator $\gamma(X)$ on $\mathcal{E}$. One can verify that

$$\left.\begin{aligned} &\gamma(X)^2 = \overline{X} \cdot X \, 1_{\mathcal{E}} \\ &\gamma(X) \cdot \gamma_5 + \gamma_5 \cdot \gamma(X) = 0 \\ &\overline{\gamma(X)} = \gamma(X) \\ &\gamma_5^2 = -1_{\mathcal{E}} \\ &\overline{\gamma}_5 = \gamma_5 \end{aligned}\right\} \qquad \forall X \in E , \tag{18.64}$$

that the relations

$$\overline{\psi}\psi \equiv 1 \qquad \text{and} \qquad \overline{\psi}\,\gamma_5\,\psi \equiv 0 \tag{18.65}$$

are equivalent to

$$\overline{\psi'}\psi'' = 1 , \tag{18.66}$$

and that they are the defining equations (in the sense of (1.40)) of a 6-dimensional submanifold Σ of $\mathcal{E}$. One can also verify that the vectors I and J in E defined by [433]

$$\left.\begin{aligned} &\overline{I} \cdot \delta X = \overline{\psi} \cdot \gamma(\delta X) \cdot \psi \\ &\overline{J} \cdot \delta X = i\,\overline{\psi} \cdot \gamma(\delta X) \cdot \gamma_5 \cdot \psi \end{aligned}\right\} \qquad \forall \delta X \in E \tag{18.67}$$

are solutions of the equations

$$\begin{aligned} \overline{I} \cdot I &= 1 \qquad\qquad I \text{ a future pointing vector } [434] \\ \overline{I} \cdot J &= 0 \\ \overline{J} \cdot J &= -1 , \end{aligned} \tag{18.68}$$

provided ψ belongs to Σ.

(18.69) Conversely, if I and J satisfy (18.68), then there exist spinors ψ in Σ such that (18.67) holds. Moreover, these spinors form an *orbit of the torus* acting on Σ according to

$$z_\Sigma(\psi) \equiv \psi z \qquad z \in T . \tag{18.70}$$

[433] The operators $\gamma(\delta X)$ and $i\gamma(\delta X)\gamma_5$ are hermitian and thus the right hand sides of the formulæ (18.67) are real.

[434] The direction of the future is that of the fourth vector in $\mathcal{R}$.

As ψ varies over Σ, one can verify the following identity

$$\mathrm{Re}\,(i\,\delta\overline{\psi}\cdot\delta'\psi) = \tfrac{1}{4}\,\mathrm{Tr}\,(\delta\Omega\cdot\Omega\cdot\delta'\Omega)\,, \tag{18.71}$$

where, as in (14.4♠), we define

$$\Omega \equiv j(I)(J)\,. \tag{18.72}$$

Thus, if we define

$$y \equiv \begin{pmatrix} X \\ I \\ J \end{pmatrix} \qquad \eta \equiv \begin{pmatrix} X \\ \psi \end{pmatrix}, \tag{18.73}$$

we see (by comparing with (14.4c)) that (18.67) defines a differentiable map $\eta \mapsto y$ from the *direct product* $W = E \times \Sigma$ onto the *evolution space* V *of a particle with spin* [(14.4c), (14.19), (Fig. 18.V)]. If we define a 1-form ϖ_W on W by

$$\varpi_W(\delta\eta) \equiv -i\,\overline{\psi}\,\delta\psi - m\,\overline{I}\,\delta X\,, \tag{18.74}$$

where m is a nonzero number, one can see that its exterior derivative $\sigma_W = \mathrm{d}\varpi_W$ is given by

$$\sigma_W(\delta\eta)(\delta'\eta) = -\tfrac{1}{2}\,\mathrm{Tr}(\delta\Omega\,\Omega\,\delta'\Omega) + m\,[\overline{\delta X}\,\delta' I - \overline{\delta' X}\,\delta I]\,. \tag{18.75}$$

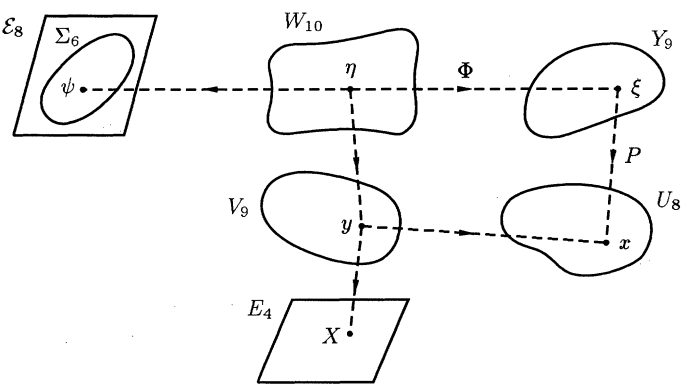


Fig 18.V

In the right hand side of (18.75) one should recognize the Lagrange form (14.4♣) of *a particle with spin of mass* m *and spin* $s = \frac{1}{2}$. Thus we have

$$\sigma_W(\delta\eta)(\delta'\eta) = \sigma(\delta x)(\delta' x)\,, \tag{18.76}$$

where x denotes the motion of the particle defined by the initial condition y. From the preceding one can deduce the leaves of the characteristic foliation of ϖ_W; they are the orbits of the additive group $\mathbf{R}$ acting on W according to

$$\lambda_W \begin{pmatrix} X \\ \psi \end{pmatrix} = \begin{pmatrix} X + I\,\lambda \\ \psi\, e^{im\lambda} \end{pmatrix}. \tag{18.77}$$

This foliation is sectionable and thus, if we define

$$\xi \equiv \Phi(\eta) \equiv \text{orbit of } \eta \text{ according to (18.77)}, \tag{18.78}$$

ξ varies over a manifold Y such that Φ is a differentiable map from W to Y. The 1-form ϖ_W, which is an integral invariant, passes to the quotient Y and defines a 1-form ϖ on Y:

$$\varpi(\delta\xi) \equiv \varpi_W(\delta\eta). \tag{18.79}$$

Finally, one can verify that

(18.80) *the 1-form ϖ gives Y the structure of a prequantum manifold and (Y, P) $[P = (\xi \mapsto x)]$ is a prequantization of U.* □

(18.81) THEOREM: The relativistic particle with spin (model (14.4)) is prequantizable if and only if its spin satisfies

$$\diamondsuit \qquad s = n\frac{\hbar}{2}, \qquad n \text{ an integer.}$$

PROOF: (a) If $\diamondsuit$ is satisfied, then the particle is prequantizable because of (18.80) and (18.47).

(b) If the particle is prequantizable, one can check that the space U_0 of motions having a given trajectory (14.4e) satisfies the conditions (18.54), and thus it must be prequantizable. Moreover, one can check that U_0 is symplectomorphic to the sphere S_2, having the symplectic structure (18.32) with $\lambda = s$. Condition $\diamondsuit$ then follows from (18.35). □

Prequantization of a massless particle

In order to find a prequantization condition, we can use the parameters $\mathbf{p}$, $\mathbf{r}$ defined in statistical mechanics (17.140). If we take $\|\mathbf{p}\| \equiv 1$, $\mathbf{r} \equiv 0$ in (17.145), then we find a manifold U_0 (as in the proof of (18.81)), which is prequantizable if the particle is, and which is symplectomorphic to a sphere. As before we deduce the necessary condition

$$s = n\frac{\hbar}{2}, \qquad n \text{ an integer,} \tag{18.82}$$

where this time s denotes the spin in the sense of (14.29◇).

Massless, spin $\frac{1}{2}$

Again we use a spinor ψ which varies over the manifold Σ (18.65) defined by

$$\overline{\psi}\,\psi \equiv 1 \qquad\qquad \overline{\psi}\,\gamma_5\,\psi \equiv 0\,, \tag{18.83}$$

but this time we define I and J by

$$\begin{aligned} \overline{I}\cdot\delta X &\equiv \overline{\psi}\,\gamma(\delta X)\,\frac{1+i\,\gamma_5}{\sqrt{2}}\psi \\ \overline{J}\cdot\delta X &\equiv -\overline{\psi}\,\gamma(\delta X)\,\frac{1-i\,\gamma_5}{\sqrt{2}}\psi\,. \end{aligned} \tag{18.84}$$

One can check the formulæ[435]

$$\begin{aligned} \overline{I}\cdot I &\equiv 0 \qquad\qquad I \text{ a future pointing vector} \\ \overline{I}\cdot J &\equiv -1 \\ \overline{J}\cdot J &\equiv 0 \end{aligned} \tag{18.85}$$

and

$$\mathrm{Re}\,(i\,\delta\overline{\psi}\,\delta'\psi) \equiv \tfrac{1}{4}\,\mathrm{Tr}\,(\delta\Omega\cdot\Omega\cdot\delta'\Omega) \tag{18.86}$$

with, as before,

$$\Omega \equiv j(I)(J)\,. \tag{18.87}$$

If we define

$$y \equiv \begin{pmatrix} X \\ I \\ J \end{pmatrix} \qquad\qquad \eta \equiv \begin{pmatrix} X \\ \psi \end{pmatrix} \tag{18.88}$$

and

$$\varpi_W(\delta\eta) \equiv -\chi i\,\overline{\psi}\,\delta\psi - \theta\,\overline{I}\,\delta X \qquad\qquad (\chi = \pm 1\,,\ \theta = \pm 1)\,, \tag{18.89}$$

one can see that $\eta \mapsto y$ is a map from $W = E \times \Sigma$ onto the manifold V defined in (14.29c), and that the exterior derivative $\sigma_W = \mathrm{d}\varpi_W$ satisfies

$$\sigma_W(\delta\eta)(\delta'\eta) \equiv \sigma(\delta x)(\delta' x)\,, \tag{18.90}$$

where x denotes the motion (characterized by y) of a *massless particle with spin* $s = \frac{1}{2}$, *helicity* χ, and having θ as the *sign of the energy* (14.29).

[435] Just perform the substitutions $I \mapsto \frac{I-J}{\sqrt{2}}$, $J \mapsto \frac{I+J}{\sqrt{2}}$ in the formulæ with nonzero mass.

Using the notation $X \equiv \mathcal{R}\begin{pmatrix}\mathbf{r}\\ t\end{pmatrix}$, $\psi \equiv \mathcal{S}\begin{pmatrix}\psi'\\ \psi''\end{pmatrix}$ [(13.40), (18.60)], (18.89) can also be written as

$$\varpi_W(\delta\eta) \equiv \chi\,\mathrm{Re}\,(-i\,\overline{\psi''}\,\delta\psi') - \frac{\theta}{\sqrt{2}}\left[\overline{\psi'}\,\psi'\,\delta t - \overline{\psi'}\,\sigma_j\,\psi'\,\delta\mathbf{r}^j\right]. \tag{18.91}$$

One can easily deduce from (18.91) that the characteristic foliation of ϖ_W is made up of 3-*dimensional affine manifolds*, submanifolds of $E \times \mathcal{E}$, on which ψ' is constant (they are thus much easier to describe than their projections on V, which are given in (14.29e)). They have a common transverse section Y defined (see (17.140)) by the equations

$$\begin{cases} t = 0 \\ \mathbf{g}_X = 0 \end{cases} \qquad \begin{array}{l}(X \text{ belongs to the trajectory relative to the}\\ \text{frame } \mathcal{R}, \text{ in the sense of (14.33)}).\end{array} \tag{18.92}$$

One can see that Y is a 7-dimensional manifold which can be described by the variable

$$\xi \equiv \begin{pmatrix}\psi'\\ \mathbf{r}\end{pmatrix} \qquad \psi' \neq 0\,, \tag{18.93}$$

and that we then have

$$\begin{aligned} \psi'' &\equiv \frac{\psi'}{\|\psi'\|^2} \\ {}^j\mathbf{p} &= \frac{\theta}{\sqrt{2}}\,\overline{\psi'}\,\sigma_j\,\psi' \qquad (\mathbf{p} = \text{linear momentum}) \\ E &= \frac{\theta}{\sqrt{2}}\,\overline{\psi'}\,\psi' \qquad (E = \text{energy}). \end{aligned} \tag{18.94}$$

The 1-form ϖ induced by ϖ_W on Y is given by

$$\varpi(\delta\xi) \equiv \chi\,\mathrm{Re}\,(-i\,\overline{\psi''}\,\delta\psi') + \langle\mathbf{p},\,\delta\mathbf{r}\rangle\,. \tag{18.95}$$

(18.96) Comparing this with (17.145), one can verify that Y *prequantizes* the particle. Obviously this is the same as defining Y as the *quotient* of W by the characteristic foliation of ϖ_W. □

Massless, spin 1

(18.97) One can proceed to the case of spin 1 and more generally to spin $n/2$ (n an integer) by means of the method of fusion (18.46). However, because of its importance (for photons) we will explicitly prequantize the case $s = 1$.

We will use the variable

$$\eta \equiv \begin{pmatrix} S \\ P \\ X \end{pmatrix}, \tag{18.98}$$

where S denotes an orthonormal basis of a negative subspace of dimension 2:

$$S \equiv [S_1 \; S_2] \qquad \overline{S} \cdot S \equiv \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \tag{18.99}$$

such that (see (14.17))

$$\Omega \equiv S \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \overline{S}\,. \tag{18.100}$$

The energy-momentum P satisfies (θ being the sign of the energy):

$$\overline{P} \cdot P = 0 \qquad \theta P = \text{future pointing} \tag{18.101}$$

and

$$\overline{P} \cdot S_1 = 0 \qquad \overline{P} \cdot S_2 = 0\,. \tag{18.102}$$

The *equations* (18.99) through (18.102) restrict the variable η to a 10-dimensional manifold W, on which we can define a 1-form ϖ_W by

$$\varpi_W(\delta\eta) \equiv \chi\, \overline{S_2}\, \delta S_1 - \overline{P}\, \delta X\,. \tag{18.103}$$

One can easily verify that the 3-dimensional Lie group of matrices

$$a \equiv \begin{bmatrix} 1 & 0 & 0 & \alpha \\ 0 & 1 & 0 & \beta \\ \beta & -\alpha & 1 & \gamma \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \alpha, \beta, \gamma \in \mathbf{R}\,, \tag{18.104}$$

(sometimes called the *Heisenberg group*) acts on W according to

$$\underline{a}_W \begin{pmatrix} S \\ P \\ X \end{pmatrix} \equiv \begin{pmatrix} S + P[\beta - \alpha] \\ P \\ X + S \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + P\gamma \end{pmatrix}, \tag{18.105}$$

and that its orbits (which are 3-dimensional) are the leaves of the characteristic foliation of the 1-form ϖ_W. This foliation turns out to be sectionable (the equations (18.92) give a transverse section for all the leaves).

(18.106) The 1-form ϖ_W induces on the quotient Y (of dimension 7) a 1-form ϖ which turns Y into a prequantum manifold. And thus Y prequantizes the massless particle with spin 1, helicity χ, having θ as the sign of its energy. □

Planck manifolds

(18.107) Let Y be a prequantum manifold. We define a *Planck manifold* to be any submanifold Φ of Y on which the 1-form ϖ vanishes:

$$[\delta\xi \text{ tangent to } \Phi \text{ at } \xi] \implies [\varpi(\delta\xi) = 0]\,. \qquad \square$$

This is equivalent to saying that the *inverse image* of ϖ under the *injective immersion* of Φ in Y is zero. Applying theorem (4.36), it follows that the inverse image of $\sigma \equiv \mathrm{d}\varpi$ under the injective immersion is also zero:

(18.108) $$[\delta\xi,\, \delta'\xi \text{ tangent to } \Phi \text{ at } \xi] \implies [\sigma(\delta\xi)(\delta'\xi) = 0]\,,$$

(18.109) that is, every Planck manifold is *isotropic* (definition (9.33)). □

(18.110) Let us assume that (Y, P) is a prequantization of a symplectic manifold U and that Φ is a Planck submanifold of Y. One can verify immediately that the tangent map of the restriction of P to Φ, that is, $D(P \circ 1_\Phi)(\xi)$, is injective. It follows that locally (that is, in every open set in some open cover of Φ), $P \cdot 1_\Phi$ is injective (theorem (1.35)).

If $P \circ 1_\Phi$ is *globally* injective, $P \circ 1_\Phi$ is an *injective immersion* (1.36) of Φ in U. Its image F is a submanifold of U (theorem (1.38)) diffeomorphic to Φ. Comparing (18.108) and (18.14) shows that F is *isotropic*. We will
(18.111) say that Φ is a *Planck lift of the isotropic manifold* F. □

Conversely, if F is an isotropic submanifold of U, to find a Planck lift of it, it suffices to note that we obtain a *foliation of the inverse image of* F *under* P by means of the equations

(18.112) $$\delta x \text{ is tangent to } F\,, \qquad \varpi(\delta\xi) = 0\,,$$

and that the desired Planck manifold must be an integral manifold of this foliation. Thus it suffices to see whether P is injective on the leaves. □

Likewise, if we know an *isotropic foliation* $x \mapsto H$ of U, then the equations

(18.113) $$[\delta\xi \in \mathcal{H}] \iff [\delta x \in H\,,\ \varpi(\delta\xi) = 0]$$

define a foliation $\xi \mapsto \mathcal{H}$ of Y. The leaves of $\mathcal{H}$ are, at least locally, Planck lifts of the leaves of H. □

Quantomorphisms

DEFINITION: Let Y and Y' be manifolds made prequantum by the 1-forms ϖ and ϖ', respectively, and let α be a diffeomorphism from Y *onto* Y' (Fig. 18.VI). We will say that α is a *quantomorphism* if the image of ϖ under α is the 1-form ϖ':

(18.114) $$[\xi' \equiv \alpha(\xi)] \implies [\varpi'(\delta\xi') \equiv \varpi(\delta\xi)]\,.$$ □

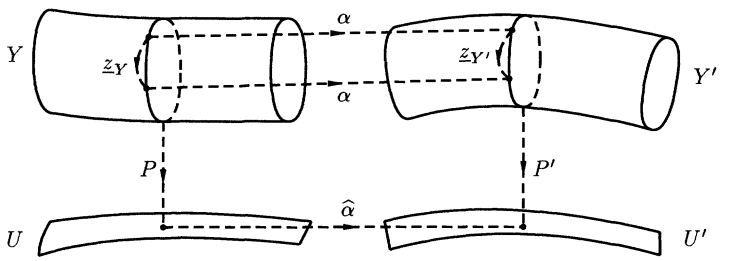


Fig 18.VI

As their name indicates, *quantomorphisms are isomorphisms of the structure of prequantum manifolds.* On account of (18.8,9), one can see that (18.114) implies

(18.115) $$\alpha(\underline{z}_Y(\xi)) \equiv \underline{z}_{Y'}(\alpha(\xi)) \qquad [z \in T]\,.$$

If Y and Y' prequantize the symplectic manifolds U and U' (by means of the projections P and P'), formula (18.115) shows that α *descends* to U and U', that is, there exists a map $\widehat{\alpha}$ defined by

(18.116) $$\widehat{\alpha}(P(\xi)) \equiv P'(\alpha(\xi))\,;$$

one can check that $\widehat{\alpha}$ is a *symplectomorphism* from U *onto* U'.

(18.117) In particular let us assume that $U = U'$ and that there exists a quantomorphism α such that $\widehat{\alpha} = 1_U$. If this is the case we will say that (Y, P) and (Y', P') are *equivalent prequantizations* of the symplectic manifold

U (it is indeed sufficient, in that case, to *identify* ξ with $\alpha(\xi)$ in order that the prequantizations be the same).

(18.118) We will say that a manifold U is *monoprequantizable* if it is prequantizable and if all its prequantizations are equivalent.[436] □

(18.119) Let (Y, P) be a prequantization of a symplectic manifold U. Let us denote by $\mathrm{Quant}(Y)$ the set of quantomorphisms from Y to itself and by P^* the map $\alpha \to \widehat{\alpha}$ (18.116) which assigns to an element $\alpha \in \mathrm{Quant}(Y)$ the induced symplectomorphism $\widehat{\alpha}$ on U (Fig. 18.VII):

(18.120) $$P^*(\alpha)(P(\xi)) = P(\alpha(\xi)) \qquad [\forall\, \xi \in Y]\,.$$

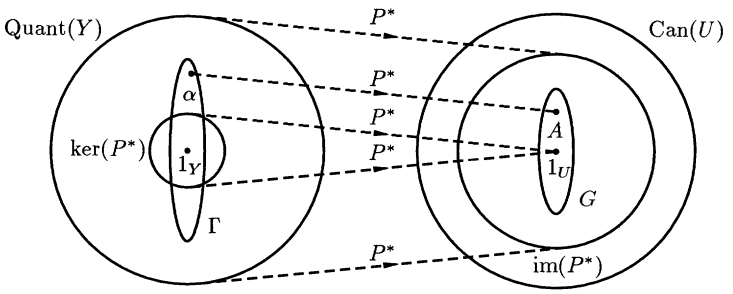


Fig 18.VII

(18.121) It is obvious that $\mathrm{Quant}(Y)$ is a *group* and that P^* is a homomorphism from this group to the group $\mathrm{Can}(U)$ of symplectomorphisms of U:

$$P^*(\alpha \circ \alpha') \equiv P^*(\alpha) \circ P^*(\alpha')\,.$$

(18.122) The *kernel* $\ker(P^*) = (P^*)^{-1}(1_U)$ is a *normal subgroup* of $\mathrm{Quant}(Y)$. If U is *connected*, one can verify that $\ker(P^*)$ contains only the maps $\underline{z}_Y$ where z varies over the torus.

(18.123) If $A = P^*(\alpha)$, we will say that the quantomorphism α is a *lift* of A, and that the symplectomorphism A is *liftable* (to Y). The image $\mathrm{im}(P^*)$ of liftable symplectomorphisms is a *subgroup* of $\mathrm{Can}(U)$. □

(18.124) If Γ is a *group of quantomorphisms* of Y, its image G under P^* is a *group of symplectomorphisms* of U. We then will say that Γ is a *lift* of G and that G is *liftable group*. Γ is called an *isomorphic lift* if P^* is injective on Γ, that is, if $\Gamma \cap \ker(P^*)$ only contains 1_Y. In that case P^* induces an *isomorphism* from Γ onto G.

[436] For example, (18.130), (18.134), and (18.135) below.

(18.125) In order that a group of symplectomorphisms of U be liftable, it is necessary and sufficient that it be contained in $\mathrm{im}(P^*)$. *However, it does not necessarily admit an isomorphic lift.*

(18.126) THEOREM: If Γ is an isomorphic lift of a group G, then the statements (a) and (b) are equivalent.

a) Γ' is another isomorphic lift of G.

b) There exists a homomorphism K from G to $\ker(P^*)$ such that

$$[\gamma' \in \Gamma'] \iff [\exists\, \gamma \in \Gamma :\ \gamma' = K(P^*(\gamma)) \circ \gamma]\,. \qquad \square$$

If in addition U is *connected*, it follows, using (18.122), that K can be written in the form

(18.127) $$K(g) = \underline{\chi(g)}_Y \qquad [\forall\, g \in G]\,,$$

where χ is a homomorphism from G to the torus T, that is, a *character* of G. □

(18.128) *Thus different isomorphic lifts of G are in 1–1 correspondence with the characters of G.*

(18.129) If U is *monoprequantizable*, *every symplectomorphism A of U is liftable.* Indeed, one can check that $(Y, A \circ P)$ is a prequantization of U, and thus *equivalent* to (Y, P), and that the quantomorphism α which provides the passage from one prequantization to the other is a lift of A.

Homotopy and prequantization

(18.130) THEOREM: Let U be a *prequantizable* and *simply connected* manifold (1.53). Then U is *monoprequantizable.*

PROOF: The reader can find a proof of this theorem in the reference below.[437] It can also be deduced from the general theory of homology of fiber bundles. □

(18.131) THEOREM: Let (U, G, F) be a *covering space* (definition (1.42)) of a symplectic manifold U' (see (Fig. 18.VIII)).

1) If U' admits a prequantization (Y', P'), we obtain a prequantization (Y, P) of U by the following construction:

[437] See Souriau [33].

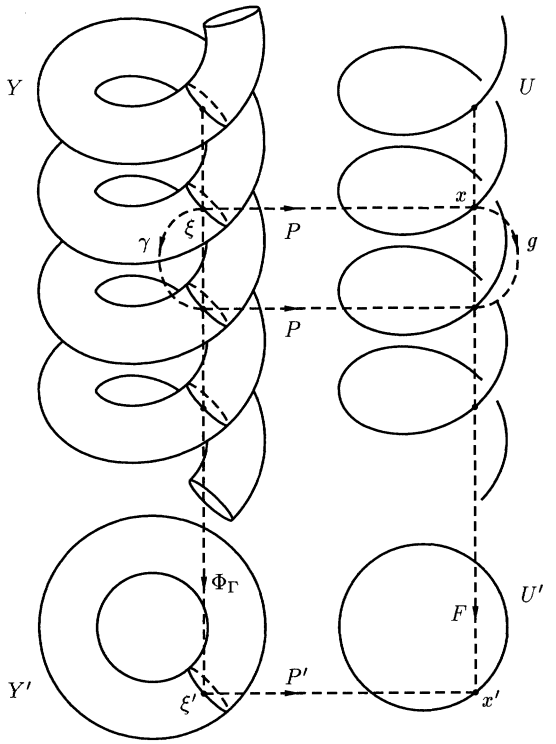


Fig 18.VIII

$$\diamondsuit \quad \left[\begin{array}{rcl} Y & = & \left\{ \begin{array}{l} \text{the submanifold of } Y' \times U, \text{ described by the} \\ \text{variable } \xi \equiv \begin{pmatrix} \xi' \\ x \end{pmatrix} \text{ and defined by the } \textit{equation} \\ P'(\xi') \equiv F(x) \qquad [\xi' \in Y',\ x \in U]^{\dagger} \end{array} \right. \\ \varpi_Y(\delta\xi) & \equiv & \varpi_{Y'}(\delta\xi') \\ P(\xi) & \equiv & x\,. \end{array} \right.$$

Moreover, the group G admits an *isomorphic lift* Γ (18.124), namely the set of γ such that

$$\heartsuit \qquad \gamma \begin{pmatrix} \xi' \\ x \end{pmatrix} \equiv \begin{pmatrix} \xi' \\ g(x) \end{pmatrix} \qquad [g \in G]\,.$$

2) Conversely, let us assume that U admits a prequantization (Y, P) and that the group G *admits an isomorphic lift* Γ on Y. Then Γ is

[†]Editors' note: This construction commonly is called the pull back of the bundle $P' : Y' \mapsto U'$ by the map $F : U \mapsto U'$.

a *discrete group* of permutations of Y. We obtain a prequantization (Y', P') of U' by the following construction:

$$\clubsuit \quad \left[\begin{array}{rcl} Y' & = & \left\{ \begin{array}{l} \text{the } \textit{quotient manifold} \text{ of } Y \text{ by the group } \Gamma \text{ (def-} \\ \text{inition (1.45))} \end{array} \right. \\ \varpi_{Y'}(\delta\xi') & \equiv & \varpi_Y(\delta\xi) \qquad [\xi \equiv \text{the } \Gamma\text{-orbit of } \xi] \\ P'(\xi') & \equiv & F(x)\,. \end{array} \right.$$

PROOF: Elementary. □

(18.132) Theorems (18.130) and (18.131) allow us to *count the number of nonisomorphic prequantizations of a given prequantizable manifold* U'. Let us assume that U' is connected (using (18.18), we can immediately reduce the problem to this case), and let (U, G, F) be a *universal covering space* of U' (1.55). To each prequantization (Y', P') of U', theorem (18.131.1) associates a prequantization (Y, P) of U such that G admits an isomorphic lift Γ. Since U is simply connected (1.55), theorem (18.130) shows that the prequantization of U is unique up to equivalence. One can easily verify that two equivalent prequantizations of U' lead to the same lift Γ of G on Y. Thus the prequantizations of U' obtained using (18.131.2) with the aid of different lifts of G are *pairwise inequivalent, and every prequantization of* Y' *is equivalent to one of them.* Applying (18.128) we thus obtain the following result.

(18.133)  A prequantizable connected manifold U' admits as many *inequivalent* prequantizations as there are distinct *characters* of its *homotopy group* G. □

In particular, we may complete (18.130) by the statement:

(18.134) In order that a prequantizable manifold be monoprequantizable, it is necessary and sufficient that the homotopy group of each of its connected components not have any characters other than the trivial one ($\chi(g) \equiv 1$). □

(18.135) EXAMPLES: One can check that the *sphere* S_2, that *every symplectic vector space*, and that *every space of motions of a free particle* (with or without spin, relativistic or not, massless or not) is *simply connected.* These manifolds are thus *monoprequantizable* whenever they are prequantizable. Consequently, the divers means we have employed to prequantize them is of no importance, since any other method would have  lead to an equivalent prequantization.[438] □

[438] If we had prequantized the relativistic particle with spin $\frac{1}{2}$ by algorithm (18.22),

Systems of elementary particles

Let us consider a system of n noninteracting particles.

(18.136) If these particles are *different*, we know that the space of motions is the direct product $U_1 \times \cdots \times U_n$ of the respective spaces of motions of the individual particles (15.37). This is a prequantizable space (18.52) (provided twice each spin is an integer) and is simply connected (1.60). Thus it is *monoprequantizable*. □

If the particles are *identical*, the space U' of motions of the system has a covering space (U, S_n, F), where U is an open subset of

$$U_1 \times \cdots \times U_n = [U_1]^n,$$

and where S_n denotes the group of permutations of the n particles (see (15.46)). U is prequantizable (18.17). One can show that U is simply connected and thus that S_n is the homotopy group of U'. Knowing that the group S_n (called the symmetric group of degree n) has exactly two characters χ_+ and χ_-:

(18.137)
$$\begin{cases} \chi_+(P) = 1 \text{ for every } P \in S_n \\ \chi_-(P) = \text{the parity of } P,^{439} \end{cases}$$

we see that:

(18.138) A system of n identical particles without interactions ($n \geq 2$) admits exactly two inequivalent prequantizations,[440] associated (by (18.132), (18.133)) to the characters χ_+ and χ_- of the symmetric group S_n. We shall call them the prequantization of *Bose-Einstein* and *Fermi-Dirac*, respectively. □

Infinitesimal quantomorphisms

(18.139) Let (Y, P) be a prequantization of a symplectic manifold U, and let u be a *dynamical variable* on U, that is, $x \mapsto u$ is a differentiable scalar field on U. We will associate to u *a vector field* $\underset{u}{\delta}$ *on* Y, $\xi \mapsto \underset{u}{\delta}\xi$, defined by

---

the result would have been *independent of the chosen (open) cover*, and *equivalent to the one we obtained using spinors* (18.80).

[439] χ_- is defined by the property [$\chi_-(P) = -1$ if P is an exchange of two particles], since every permutation is a product of exchanges and because χ_- is a character.

[440] As before, prequantization is only possible if *twice the spin is an integer*. But geometry *does not provide the relation between the spin and the character* χ as suggested by experiments.

the equations

$$\begin{aligned} \varpi(\underset{u}{\delta}\xi) &\equiv u \\ \sigma(\underset{u}{\delta}\xi) &\equiv -\mathrm{d}u\,. \end{aligned} \tag{18.140}$$

The axioms (18.2) show that this system of linear equations in $\underset{u}{\delta}\xi$ has *a unique solution*. Writing $x \equiv P(\xi)$ one can check that

$$\underset{u}{\delta} x = \operatorname{grad} u\,, \tag{18.141}$$

where grad denotes the *symplectic gradient* on U. *The notation* $\underset{u}{\delta}$ *is compatible with the one we already used in* (9.20).

Let us note the formulæ

$$\underset{u+u'}{\delta}\xi \equiv \underset{u}{\delta}\xi + \underset{u'}{\delta}\xi\,, \qquad \underset{su}{\delta}\xi \equiv s\underset{u}{\delta}\xi \qquad (s = \text{constant})\,. \tag{18.142}$$

(18.143) THEOREM: Let $\xi \mapsto \delta\xi$ be a differentiable vector field on a prequantum manifold Y (hypotheses (18.139)). In order that the Lie derivative $\delta_L\varpi$ be zero, it is necessary and sufficient that there exists a dynamical variable u on U such that

$$\delta\xi \equiv \underset{u}{\delta}\xi\,.$$

PROOF: Applying Cartan's formula (4.39), we see that $\delta_L\varpi \equiv 0$ becomes

$$\diamondsuit \qquad \sigma(\delta\xi) + \mathrm{d}\,[\varpi(\delta\xi)] \equiv 0\,,$$

and this equality is indeed satisfied if $\delta = \underset{u}{\delta}$ (18.140).

Conversely, if $\diamondsuit$ is satisfied, and if we *define* $u \equiv \varpi(\delta\xi)$, then we have $\sigma(\delta\xi) \equiv -\mathrm{d}u$. Hence (notation (18.7))

$$[\mathrm{d}u]\,(\mathrm{i}_Y(\xi)) \equiv -\sigma(\delta\xi)(\mathrm{i}_Y(\xi)) \equiv 0\,,$$

and thus u is a *first integral* (5.22) of the vector field $\xi \mapsto \mathrm{i}_Y(\xi)$. Thus there exists a differentiable field $x \mapsto u$ on U ((5.21), (18.15)). Because the equations (18.140) are satisfied, we indeed have $\delta\xi \equiv \underset{u}{\delta}\xi$. □

From this result we deduce the formula

$$[\underset{u}{\delta}, \underset{v}{\delta}]\,\xi \equiv \underset{[u,v]}{\delta}\xi \qquad \left(\begin{array}{lcl} [u,v] & = & \text{Poisson bracket} \\ [\underset{u}{\delta}, \underset{v}{\delta}] & = & \text{Lie bracket} \end{array}\right), \tag{18.144}$$

which generalizes (9.25) (exercise!).

(18.145) DEFINITION: Let (Y, P) be a prequantization of a symplectic manifold U and let f be a differentiable vector field on Y. We will say that f is an *infinitesimal quantomorphism* if

$$\clubsuit \qquad \exp(sf) \in \mathrm{Quant}(Y) \qquad [\forall\, s \in \mathbf{R}]\,.$$

(18.145) THEOREM: In order that f be an infinitesimal quantomorphism, it is necessary and sufficient that

$\diamondsuit$ there exists a dynamical variable u, defined on U, such that

$$f(\xi) \equiv \underset{u}{\delta}\xi \qquad (18.140)\,;$$

$\heartsuit$ there exists a number $\alpha > 0$ such that $\exp(\alpha \widehat{f})$ is defined on all of U, where $\widehat{f}$ denotes the field $x \mapsto \underset{u}{\delta} x$ on U.[†]

PROOF: Theorem (3.7), which serves as the definition of the Lie derivative, shows that the vector field f is an infinitesimal quantomorphism if and (only if)

$\bowtie$ $\qquad \exp(sf)$ is defined on Y for every s

$$\spadesuit \qquad \delta_L \varpi \equiv 0 \qquad [\delta\xi \equiv f(\xi)]\,.$$

Theorem (3.10) shows that we can replace $\bowtie$ by:

$\triangle$ $\qquad$ there exists an $\alpha > 0$ such that $\exp(\alpha f)$ is defined on Y.

Theorem (18.143) shows the equivalence between $\diamondsuit$ and $\spadesuit$. Thus only the equivalence of $\heartsuit$ and $\triangle$ remains to be shown.

We obviously have

$$\frac{d}{ds}\left[P(\exp(s\,f)(\xi))\right] \equiv \widehat{f}(x) \qquad \text{provided } \exp(s\,f)(\xi) \text{ exists,}$$

which shows that $\triangle \Rightarrow \heartsuit$.

Conversely, assume $\heartsuit$. Let ξ_0 be a point of Y and let $x_0 = P(\xi_0)$. Using the properties of *compactness* (16.16), (16.22), we can decompose the interval $[0, \alpha]$ into a finite number of subintervals such that the image of each subinterval under the map $s \mapsto \exp(s\,\widehat{f})(x_0)$ is contained in a single simply connected open set U_j (see (18.19)), on which σ admits a potential ϖ_j.

[†]Editors' note: This condition is usually formulated by saying that the vector field $\widehat{f}$ is complete.

Theorem (18.130) shows that there exists a diffeomorphism A_j from $T \times U_j$ to the open set Y_j over U_j such that

$$P\left(A_j \begin{pmatrix} z \\ x \end{pmatrix}\right) \equiv x \qquad \varpi\left(\delta\left[A_j \begin{pmatrix} z \\ x \end{pmatrix}\right]\right) \equiv \frac{\delta z}{i\, z} + \varpi_j(\delta x)\,.$$

One can check that the map

$$s \mapsto \xi \equiv A_j \begin{pmatrix} x \\ e^{i\varphi} \end{pmatrix} \qquad \left[x \equiv \exp(s\, f)(x_0),\ \ \varphi = \int \varpi_j \cdot \frac{dx}{ds}\, ds\right]$$

is a solution of the differential equation $\frac{d\xi}{ds} \equiv f(\xi)$, *for every choice of a primitive* φ. Pasting together the solutions on the subintervals defined above, we can easily construct a solution defined on $[0, \alpha]$ which equals ξ_0 for $s = 0$. Theorem (3.2) shows that this solution equals $\exp(s\, f)(\xi_0)$. Hence $\triangle$ follows since ξ is arbitrary. □

(18.146) We thus see that if f is an infinitesimal quantomorphism, then the map

$$\underline{s}_Y(\xi) = \exp(s\, f)(\xi)$$

defines an action by quantomorphisms of the additive group $\mathbf{R}$ *on* Y.

EXAMPLE: If we take $u \equiv 1$, we have $\underset{u}{\delta} x \equiv 0$. Because condition (18.145♡) is satisfied, $\xi \mapsto \underset{1}{\delta}\xi$ is an infinitesimal quantomorphism. Furthermore, comparing (18.140) and (18.7) shows that

(18.147) $$\underset{1}{\delta} \equiv \mathrm{i}_Y(\xi)\,.$$

Quantization of dynamical groups

(18.148) Let (Y, P) be a prequantization of the symplectic manifold U and let G be a Lie group *which acts on* Y *by quantomorphisms*.

Since the maps $\underline{a}_Y$ $(a \in G)$ preserve the 1-form ϖ, they also preserve its exterior derivative σ (4.36). G is thus a *dynamical group* of the presymplectic manifold Y (11.1). If we define (notation (18.119))

(18.149) $$\underline{a}_U = P^*(\underline{a}_Y) \qquad \forall\, a \in G\,,$$

which also can be written as

(18.150) $$\underline{a}_U(P(\xi)) \equiv P(\underline{a}_Y(\xi))\,,$$

it follows from (11.5) that the map

$$a \mapsto \underline{a}_U$$

turns G into a *dynamical group of the symplectic manifold* U.

(18.151) In such a case we will say that G is a *quantizable dynamical group* for the prequantization (Y, P) of U, and that the map $a \mapsto \underline{a}_Y$ is a *quantization* of G. □

It is clear that the set of maps $\underline{a}_Y$ is a *lift* of the set of maps $\underline{a}_U$ — a set which is thus a *liftable group*. We will see that it is not sufficient that a dynamical group be liftable in order that it be quantizable.

Since G must act on Y preserving the 1-form ϖ, we know (11.10) that we obtain a moment μ of G by writing

$$\mu \cdot Z \equiv \varpi(Z_Y(\xi)) \qquad \forall\, Z \in \mathcal{G}\,, \tag{18.152}$$

where $\mathcal{G}$ is the Lie algebra of G, and that the associated symplectic cocycle θ is zero (11.21). Comparing the above equation with (18.140) we note that

$$\underset{\mu\cdot Z}{\delta}\xi \equiv Z_Y(\xi)\,. \tag{18.153}$$

The *generalized version of Noether's theorem* (11.12) shows that μ depends on ξ only via x, and that μ is also a moment of G *seen as a dynamical group of* U. It is immediate that the cocycle associated with this latter moment is also zero, that is,

$$\psi(\underline{a}_U(x)) \equiv \underline{a}_{\mathcal{G}^*}(\psi(x)) \qquad \psi = [x \mapsto \mu]\,, \quad a \in G\,. \tag{18.154}$$

Using the terminology of §11, we thus see that:

(18.155) If a dynamical group of a symplectic manifold is *quantizable, then its symplectic cohomology is zero.* □

We will see in (18.167) that this necessary condition *is not sufficient.*

(18.156) In the case that a dynamical group G is *liftable but not quantizable*, it might happen that one can find a Lie group G' acting on Y by quantomorphisms and providing a lift of G. Thus for $a \in G$ there would exist an element a' of G' such that

$$\underline{a}_U = \underline{a'}_U \qquad [\underline{a'}_U(P(\xi)) \equiv P(\underline{a'}_Y(\xi))]\,. \tag{18.157}$$

If in addition G acts *effectively* (16.220), the above equation defines a map Φ

$$a \equiv \Phi(a')\,, \tag{18.158}$$

which is *a homomorphism from G' onto G*. One says that G' is an *extension* of G. In this way we find a *quantizable extension* G' of the dynamical group G.

EXAMPLE: Let E be a symplectic vector space. Seen as an *additive group*, E acts on itself by translations

$$\underline{a}_E(x) = a + x \qquad \forall\, a, x \in E\,, \tag{18.159}$$

whence

$$Z_E(x) = Z \qquad \forall\, x, Z \in E\,. \tag{18.160}$$

We thus obtain a dynamical group of E having the moment

$$\mu \equiv \sigma(x)\,. \tag{18.161}$$

The associated cocycle θ is equal to σ and thus its cohomology class is not zero.

E admits a prequantization (Y, P):

$$\left\{\begin{array}{rcll} Y & = & E \times T & \\ \varpi\,\delta\begin{pmatrix} x \\ z \end{pmatrix} & \equiv & \dfrac{\delta z}{i\,z} + \frac{1}{2}\,\sigma(x)(\delta x) & [z \in T,\ x \in E] \\ P\begin{pmatrix} x \\ z \end{pmatrix} & \equiv & x & \end{array}\right. \tag{18.162}$$

(see (18.30), (18.31)). Because E is simply connected, it is monoprequantizable (18.130). Thus the dynamical group E is liftable (18.129), but is not quantizable, since its cohomology is not zero (18.155).

It is easy to find all lifts of the translations of E. One can check that all of them can be obtained as (left) translations of Y, provided we give Y the *structure of a Lie group* with multiplication

$$\begin{pmatrix} x \\ z \end{pmatrix} \times \begin{pmatrix} x' \\ z' \end{pmatrix} \equiv \begin{pmatrix} x + x' \\ z z'\, \mathrm{e}^{-i\,\sigma(x)(x')/2} \end{pmatrix}. \tag{18.163}$$

This group is called the *Weyl group*.†

It is clear that P is a homomorphism from Y onto E making the Weyl group into a *quantizable extension* of E. One should note that it is a *nonabelian* extension of the abelian group E. □

†Editors' note: Sometimes called the Heisenberg group. This group is not to be confused with the Heisenberg group defined in (18.104).

EXAMPLE: SO(3) is a dynamical group of the sphere S_2, having the symplectic structure (characterized by a positive number λ)

$$\sigma(\delta x)(\delta' x) = \lambda\langle x, \delta x \times \delta' x\rangle \qquad [x \in \mathbf{R}^3,\ \|x\| \equiv 1]\,, \tag{18.164}$$

(see (11.40)).

S_2 is prequantizable if 2λ is an integer (18.45). If SO(3) is quantizable, it has a moment μ defined by (18.152) and thus it satisfies (18.154). Formula (11.17$^\sharp_\sharp$) defines μ on the derived Lie algebra of the Lie algebra SO(3), which is the Lie algebra of SO(3) itself. We find that

$$\mu(j(Z)) \equiv -\lambda\,\langle x, Z\rangle \qquad [\forall\, x \in S_2,\ \forall\, Z \in \mathbf{R}^3]\,. \tag{18.165}$$

Formula (18.152) then gives the infinitesimal quantomorphisms $[j(Z)]_Y$ of the group. Applying (6.12c) it follows that we must have

$$\exp(j(Z)_Y) = 1_Y \qquad \text{for } \|Z\| = 2\pi\,. \tag{18.166}$$

(18.167) *If 2λ is odd, then the above equation is not satisfied, and thus the group* SO(3) *is not quantizable, even though its cohomology is zero.* However, it is liftable because S_2 is simply connected.

Let us discuss the case $\lambda = \frac{1}{2}$ with the prequantization (18.38, 39, 44). The computation above gives

$$j(Z)_Y(\xi) = -\frac{i}{2}\gamma(Z)\xi \qquad [\forall\, \xi \in Y] \tag{18.168}$$

with γ defined by

$$\gamma(Z) = \sum_{j=1}^{3} \sigma_j\, {}^j Z \qquad [\forall Z \in \mathbf{R}^3]\,, \tag{18.169}$$

where the σ_j are the Pauli spin matrices (18.43). From this we deduce

$$\exp(j(Z)_Y) = \cos\left(\frac{\|Z\|}{2}\right) - i\,\sin\left(\frac{\|Z\|}{2}\right)\gamma\left(\frac{Z}{\|Z\|}\right), \tag{18.170}$$

a quantity which equals -1 for $\|Z\| = 2\pi$. However, the result of this calculation is not just negative: it shows that the vector fields $j(Z)_Y$ defined by (18.168) are infinitesimal quantomorphisms and that their exponentials (18.170) *generate the group* SU(2), acting by quantomorphisms on Y (which, we recall, is the sphere $\|\xi\| \equiv 1$ in $\mathbf{C}^2$) according to

$$\underline{a}_Y(\xi) \equiv a\xi \qquad [a \in \mathrm{SU}(2),\ \xi \in Y]\,. \tag{18.171}$$

Thus SU(2) is a *quantizable extension* of SO(3). One can easily compute the homomorphism Φ defined in (18.158):

$$\gamma(\Phi(a)\cdot Z) \equiv a\cdot\gamma(Z)\cdot\overline{a} \qquad [\forall\, a\in \mathrm{SU}(2)\,,\ \forall\, Z\in \mathbf{R}^3]\,. \tag{18.172}$$

It is the one which has been discovered by Hamilton and which allows one to express a rotation in terms of *quaternions.*[441] Its kernel is $a=\pm 1$. □

(18.173) EXAMPLE: An analogous phenomenon occurs in the case of the restricted Poincaré group G, which by construction is a dynamical group of the space of motions of a relativistic particle. G is quantizable only if the spin of the particle is zero or an integer. In the case $s=\frac{1}{2}$, we can find a quantizable extension G' of G as follows. The complex special linear group on $\mathbf{C}^2$ (denoted $\mathrm{Sl}(2,\mathbf{C})$) acts on the space $\mathcal{E}$ of Dirac spinors and on Minkowski space E in such a way that (notation (18.59), (18.63)):

$$\underline{a}_{\mathcal{E}} \equiv \mathcal{S}\cdot\begin{bmatrix} a & 0\\ 0 & \overline{a^{-1}}\end{bmatrix}\cdot\mathcal{S}^{-1} \qquad a\in \mathrm{Sl}(2,\mathbf{C})\,,\ Y\in E\,. \tag{18.174}$$

$$\underline{a}_{\mathcal{E}}\cdot\gamma(Y)\cdot\underline{a^{-1}}_{\mathcal{E}} \equiv \gamma(\underline{a}_E(Y)) \tag{18.175}$$

One can check that $a\mapsto \underline{a}_E$ is a homomorphism from $\mathrm{Sl}(2,\mathbf{C})$ *onto the connected Lorentz group*, whose kernel has 2 elements ($a=\pm 1$).

The Lie group G' is defined as the product $\mathrm{Sl}(2,\mathbf{C})\times E$ with the multiplication

$$\begin{pmatrix} a\\ C\end{pmatrix}\times\begin{pmatrix} a'\\ C'\end{pmatrix} \equiv \begin{pmatrix} aa'\\ C+\underline{a}_E C'\end{pmatrix} \qquad \begin{bmatrix} a,a'\in \mathrm{Sl}(2,\mathbf{C})\\ C,C'\in E\end{bmatrix}. \tag{18.176}$$

The map

$$\begin{pmatrix} a\\ C\end{pmatrix}\mapsto\begin{pmatrix} \underline{a}_E & C\\ 0 & 1\end{pmatrix} \tag{18.177}$$

is a homomorphism from G' onto the restricted Poincaré group G. G' acts on the manifold W (18.73) by

$$\underline{\begin{pmatrix} a\\ C\end{pmatrix}}_{W}\begin{pmatrix} X\\ \psi\end{pmatrix} \equiv \begin{pmatrix} \underline{a}_E\cdot X+C\\ \underline{a}_E\cdot\psi\end{pmatrix} \qquad \begin{bmatrix} X\in E\\ \psi\in\Sigma\ (18.65)\end{bmatrix}. \tag{18.178}$$

[441] Quaternions can be thought of as the set of complex 2×2 matrices a satisfying $a+\overline{a}=\mathrm{Tr}(a)\cdot 1$, with a norm defined by $\|a\|^2=\det(a)$. SU(2) is *the unit sphere in the real 4-dimensional space of quaternions.*

The Hilbert space of a prequantum manifold

(18.179) Let Y be a prequantum manifold. We will denote by $\mathcal{H}(Y)$ the set of *differentiable functions* Ψ *on* Y *with compact support satisfying the condition*

$$\Psi(\underline{z}_Y(\xi)) = z\,\Psi(\xi) \qquad \forall\, z \in T\,,\ \forall\, \xi \in Y\,. \tag{$\Diamond$}$$

$\mathcal{H}(Y)$ is a *complex vector space.* □

(18.180) Let (Y, P) be a prequantization of a symplectic manifold U. For $\Psi, \Psi' \in \mathcal{H}(Y)$ one can verify immediately that the product

$$\overline{\Psi(\xi)} \cdot \Psi'(\xi)$$

is constant on the orbits of the torus. Thus there exists a map F:

$$F(x) \equiv \overline{\Psi(\xi)} \cdot \Psi'(\xi) \qquad [x \equiv P(\xi)]$$

defined on U. It is easily verified that F is differentiable (hence continuous) and that its support is compact. Thus we can apply the *Liouville measure* μ on U to it. We obtain a complex number $\mu(F)$,[442] which we will denote by $\langle \Psi, \Psi' \rangle$, that is, using the notation of (16.144, 145),

$$\langle \Psi, \Psi' \rangle = \int_U \overline{\Psi(\xi)} \cdot \Psi'(\xi)\, \varphi(dx) \qquad \forall\, \Psi, \Psi' \in \mathcal{H}(Y)\,, \tag{18.181}$$

where φ denotes the symplectic density (16.94) on U.

(18.182) DEFINITION: A *pre-Hilbert* space is a complex vector space E such that

$$\begin{cases} \text{there exists a map } \Psi \mapsto \overline{\Psi} \text{ from } E \text{ to its dual, satisfying} \\ \overline{\Psi'} \cdot \Psi = \overline{\overline{\Psi} \cdot \Psi'}^{\,443} \qquad \forall\, \Psi, \Psi' \in E\,, \\ \Psi \neq 0 \implies \overline{\Psi} \cdot \Psi > 0\,. \end{cases}$$ □

(18.183) It follows immediately from the properties of positive measures that $\mathcal{H}(Y)$ becomes a *pre-Hilbert space* if we define

$$\overline{\Psi} \cdot \Psi \equiv \langle \Psi, \Psi' \rangle \qquad [\Psi, \Psi' \in \mathcal{H}(Y)]\,. \tag{$\Diamond$}$$ □

(18.184) THEOREM/DEFINITION:[444] Let E be a *pre-Hilbert space.*

[442] We have $\mu(F) = \mu(F_1) + i\,\mu(F_2)$, where $F(x) \equiv F_1(x) + i\,F_2(x)$ with F_1, F_2 being real valued (see (16.144)).

[443] The right hand side of the equation is the *complex conjugate* of the complex number $\overline{\Psi} \cdot \Psi'$.

[444] The proof can be found in Naimark [27].

1) E becomes a *normed* vector space (16.105) if we define

$$\|\Psi\| = \sqrt{\overline{\Psi} \cdot \Psi} \qquad \forall\, \Psi \in E\,.$$

2) $\Psi \mapsto \overline{\Psi}$ is an *injective antilinear* map[445] from E to its *topological dual* (16.108) E' satisfying

$$\|\,\overline{\Psi}\,\| \equiv \|\Psi\|\,.^{446}$$

3) In order that E be complete[447] it is necessary and sufficient that the map $\Psi \mapsto \overline{\Psi}$ be a bijection from E *onto* E' (Riesz' theorem). In such a case E is called a *Hilbert space*; in particular it is a Banach space.[447]
4) Every pre-Hilbert space E can be *completed*, that is, *embedded in a Hilbert space* $\widehat{E}$ so that the operations $\begin{pmatrix}\Psi\\ \Psi'\end{pmatrix} \mapsto \Psi + \Psi'$, $\begin{pmatrix}\Psi\\ z\end{pmatrix} \mapsto \Psi z$, and $\begin{pmatrix}\Psi\\ \Psi'\end{pmatrix} \mapsto \overline{\Psi} \cdot \Psi'$ of E are induced from those of $\widehat{E}$, and so that E is *dense* in $\widehat{E}$.[448]

(18.185) DEFINITION: A *C^*-algebra* is a complex Banach algebra A[449] equipped with an antilinear map[450] $a \mapsto a^*$ from A to itself such that:[451]

$$\bowtie \qquad \begin{cases} [a^*]^* &= a \\ [a \times b]^* &= b^* \times a^* \qquad \forall\, a, b \in A\,. \\ \|a\| &= \|a^*\| = \sqrt{\|a^* \times a\|} \end{cases}$$

(18.185) THEOREM: Let E be a Hilbert space, let $B(E)$ be the set of *bounded* linear maps (16.107) from E to itself, and let a be an element of $B(E)$. Then there exists an element a^* of $B(E)$ defined by

$$\diamondsuit \qquad \overline{a^*(\Psi)} = \overline{\Psi} \circ a \qquad \forall\, \Psi \in E\,.$$

---

[445] A map F is *antilinear* if it is linear over the real numbers and satisfies $F(i\,\Psi) \equiv -i\,F(\Psi)$.

[446] $\|\,\overline{\Psi}\,\|$ is the norm of $\overline{\Psi}$ in the sense of (16.107$\diamondsuit$). The formula $\|\overline{\Psi}\| = \|\Psi\|$ thus contains the *Cauchy-Schwarz-Bunyakowski* inequality $|\overline{\Psi} \cdot \Psi'| \leq \|\Psi\| \times \|\Psi'\|$.

[447] Definition (16.106)

[448] A subset H of a normed vector space E is *dense* if every nonempty ball of E contains points of H. In case E is *already complete*, we have $\widehat{E} = E$.

[449] Recall definition (16.160, 161): A is a complex Banach space (16.106) with a map $\begin{pmatrix}a\\ b\end{pmatrix} \mapsto a \times b$, which is *associative* and *bilinear* such that $\|a \times b\| \leq \|a\|\,\|b\|$ $(\forall\, a, b \in A)$.

[450] Definition (18.184.2): $[a + b]^* \equiv a^* + b^*$, $[s\,a]^* \equiv \overline{s}\,a^*$ if s is complex.

[451] A C^*-algebra is thus an algebra having an involution (16.165, 166).

The map $*$ turns $B(E)$ into a C^*-algebra.

PROOF: We already know that $B(E)$ is a complex Banach space (16.107). It is easy to verify that the product (composition) turns it into a Banach algebra. The existence of a map $a \mapsto a^*$ from $B(E)$ to itself defined by $\Diamond$ is a simple consequence of (18.184), and in particular of Riesz' theorem. The verification of the axioms $\bowtie$ of a C^*-algebra is a nice little exercise.[452] □

(18.186) Let A be a C^*-algebra and let $a \in A$. The element a^* is called the *adjoint* of a, a is called *self-adjoint* if $a^* = a$, and a is called *normal* if $a \times a^* = a^* \times a$. □

A C^*-algebra A may have a *unit*, that is, an element 1_A (necessarily unique) such that

(18.187) $$1_A \times a = a \times 1_A = a \qquad \forall a \in A\,.$$[453] □

In such a case, a is called *unitary* if

(18.188) $$a^* \times a = a \times a^* = 1_A\,.$$

The unitary elements form a group under the multiplication $\times$ (called the *unitary group* of A), having 1_A as an identity element and $a \mapsto a^*$ as inverse. □

Let E be a pre-Hilbert space, let a be a *bounded* linear operator mapping E to E, and let $\widehat{E}$ be the completion of E (18.184).

(18.189) It is easy to see that a has a *unique extension $\widehat{a}$ which belongs to $B(\widehat{E})$*. It is defined by continuity: if a sequence Ψ_n of elements of E converges, in $\widehat{E}$, to Ψ ($\Psi \in \widehat{E}$), then

(18.190) $$\widehat{a}(\Psi) = \lim_{n\to\infty} a(\Psi_n)\,.$$

The map $a \mapsto \widehat{a}$ is a *homomorphism of normed algebras* (it is linear and satisfies $\widehat{a \times b} \equiv \widehat{a} \times \widehat{b}$, $\|\widehat{a}\| \equiv \|a\|$). One can check that[454]

(18.191) $$[\widehat{a} \text{ is } \textit{unitary}] \iff \left[\begin{array}{l} a \text{ is an } \textit{isometry}\text{: } \|a(\Psi)\| = \|\Psi\| \quad [\forall\, \Psi \in E] \\ \mathrm{im}(a) \text{ is } \textit{dense} \text{ in } E \end{array}\right].$$

[452] The solution can be found in Naimark [27]

[453] Which happens obviously in the case $A = B(E)$.

[454] One might use the identity $\overline{\Psi} \cdot \Psi' \equiv \frac{1}{4}\left[\|\Psi + \Psi'\|^2 - i\,\|\Psi + i\,\Psi'\|^2 - \|\Psi - \Psi'\|^2 + i\,\|\Psi - i\,\Psi'\|^2\right]$.

(18.192) DEFINITION: Let E be a pre-Hilbert space. A linear map A from E to itself is *hermitian* if it satisfies

$$\triangle \qquad \overline{A(\Psi)} = \overline{\Psi} \circ A \qquad \forall\, \Psi \in E\,.$$

Let (Y, P) be a prequantization of a symplectic manifold U. To every dynamical variable u defined on U, we can associate an operator $\widehat{u}$ on $\mathcal{H}(Y)$ defined by[455]

$$\bowtie \qquad \widehat{u}(\Psi)(\xi) \equiv -i \underset{u}{\delta}[\Psi(\xi)] \qquad \forall\, \Psi \in \mathcal{H}(Y)\,.$$

(18.192) THEOREM:

♠ $\widehat{u}$ is a *hermitian* operator.

♡ The map $u \mapsto \widehat{u}$ is *linear* and *injective*.[456]

$$\diamondsuit \qquad \widehat{1} \;=\; 1_{\mathcal{H}(Y)}$$

$$\clubsuit \qquad \widehat{u} \circ \widehat{u'} - \widehat{u'} \circ \widehat{u} \;=\; -i\,\widehat{[u, u']_P}\,.^{457}$$

PROOF: Let K be the support, which is compact, of an element Ψ of $\mathcal{H}(Y)$. In the open set complementary to K (16.19), $\Psi(\xi)$ is identically zero, thus also $\underset{u}{\delta}[\Psi(\xi)]$. Consequently, the set of ξ where $\underset{u}{\delta}[\Psi(\xi)] \neq 0$ is contained in K and hence is relatively compact (16.25). Its closure, which is the support of $\widehat{u}(\Psi)$, is thus *compact*.

One can easily verify that the field $\xi \mapsto \underset{u}{\delta}\xi$ is invariant under the transformations $\underset{Y}{z}$ $(z \in T)$. Differentiating the identity

$$\Psi(\underset{Y}{z}(\xi)) \equiv z\,\Psi(\xi)$$

satisfied by Ψ (18.179),[458] we deduce that $\widehat{u}(\Psi)$ satisfies the same identity. Thus $\widehat{u}$ maps $\mathcal{H}(Y)$ indeed into $\mathcal{H}(Y)$. The linearity of this map is trivial to verify.

Let Ψ, Ψ' be elements of $\mathcal{H}(Y)$ and let us define

$$\alpha \equiv \overline{\Psi(\xi)} \cdot \Psi'(\xi)\,.$$

---

[455] Notation $\underset{u}{\delta}$: see (18.140).

[456] From the real vector space of all dynamical variables on U to the real vector space of hermitian operators on $\mathcal{H}(Y)$.

[457] $[u, u']_P$ denotes the *Poisson bracket* (9.22) of two arbitrary dynamical variables u and u'.

[458] And using (18.144) and (18.147)

Applying the derivation $\underset{u}{\delta}$, one can see that

$$\underset{u}{\delta}\,\alpha \equiv i\,\left[\overline{\Psi(\xi)}\cdot\widehat{u}(\Psi')(\xi) - \overline{\widehat{u}(\Psi)(\xi)}\cdot\Psi'(\xi)\right],$$

whence, using definition (18.181),

$$\int_U \underset{u}{\delta}\,\alpha\,\varphi(dx) = i\,\left[\overline{\Psi}\cdot\widehat{u}(\Psi') - \overline{\widehat{u}(\Psi)}\cdot\Psi'\right].$$

Using a partition of unity (theorem (16.49)), we can write α in the form $\sum_j \alpha_j$, where the α_j are differentiable and have their support each inside the image of a canonical chart. Thus we find (9.23a)

$$\sharp \qquad \int_U \underset{u}{\delta}\,\alpha\,\varphi(dx) = \sum_j \int\cdots\int [u,\alpha_j]_P\, dp_1\ldots dp_n\; d^1q\ldots d^nq\,.$$

Applying formula (10.7), which gives the Poisson bracket in canonical coordinates, we find that

$$[u,\alpha_j]_P = \sum_{k=1}^{n} \frac{\partial}{\partial\, {}^k q}\left[\alpha_j \frac{\partial u}{\partial p_k}\right] - \frac{\partial}{\partial p^k}\left[\alpha_j \frac{\partial u}{\partial\, {}^k q}\right].$$

Integrating by parts and noting that α_j vanishes outside a compact set, it follows that the integrals appearing on the right hand side of equation $\sharp$ are *zero*. Thus we have

$$\overline{\Psi}\cdot\widehat{u}(\Psi') = \overline{\widehat{u}(\Psi)}\cdot\Psi'\,,$$

which shows that $\widehat{u}$ is indeed *hermitian* (18.192$\triangle$), because Ψ and Ψ' are arbitrary.

The linearity of the map $u \mapsto \widehat{u}$ follows immediately from (18.142). To show that it is injective, it is thus sufficient to construct (u being a dynamical variable which is not identically zero) an element Ψ_0 of $\mathcal{H}(Y)$ and a point ξ_0 of Y such that

$$\widehat{u}(\Psi_0)(\xi_0) \neq 0\,.$$

Such a construction is not difficult if U is a manifold admitting a potential with the prequantization (18.30). The construction can be reduced to this case by choosing a cell U_0 of U on which u is not identically zero, and then using theorem (18.130).

Finally, the verification of the formulæ $\diamondsuit$ and $\clubsuit$ is immediate, using (18.147) and (18.179) for the first and (18.144) and (2.45) for the second. $\square$

EXAMPLE: Consider a symplectic vector space equipped with a global canonical chart (see (10.3)). Since this space is prequantizable (18.31), it follows from (18.192) that the operators $\widehat{p}_j$, $\widehat{q}_k$ associated to the canonical coordinates are hermitian and satisfy

$$
\begin{array}{lll}
[\widehat{p}_j, \widehat{p}_k]_- = 0 \qquad [\widehat{q}_j, \widehat{q}_k]_- = 0 & & \forall\, j,k \\
[\widehat{p}_j, \widehat{q}_k]_- = 0 & & j \neq k \\
[\widehat{p}_j, \widehat{q}_j]_- = -i\,\underline{1} & & \forall\, j\,,
\end{array}
\tag{18.193}
$$

where the notation $[a,b]_-$ means the *commutator* $ab - ba$. □

As we can see in this last example, the operators $\widehat{u}$ are not always *bounded*, which poses enormous difficulties for their study. Von Neumann has shown that, if a is a hermitian operator satisfying certain conditions (*which are not always satisfied by the operators* $\widehat{u}$), then we can define the exponentials $\exp(is\,\widehat{a})$, s real, which are bounded operators.

In analogy with the finite dimensional case, we expect that the following identity holds:

$$
\frac{d}{ds}\left[\exp(is\,\widehat{u})(\Psi)(\xi)\right] = i\,\widehat{u}\left(\left(\exp(is\,\widehat{u})(\Psi)\right)(\xi)\right). \tag{18.194}
$$

And indeed this equality is valid if the field $f\,[\xi \mapsto \underset{u}{\delta}\,\xi]$ is an *infinitesimal quantomorphism* (18.145) and if we define

$$
\exp(is\,\widehat{u})(\Psi)(\xi) \equiv \Psi\left(\exp(s\,f)(\xi)\right). \tag{18.195}
$$

This suggests to let the *quantomorphisms act directly*, without taking the detour of infinitesimal quantomorphisms and exponentiation. This is possible because of the following theorem.

(18.196) THEOREM: Let Y be a prequantum manifold.

1) If a is a quantomorphism of Y and if we define

$$
f(a)(\Psi) = \Psi \circ a^{-1} \qquad \forall\, \Psi \in \mathcal{H}(Y)\,,
$$

then $f(a)$ is a *surjective isometric linear* map from $\mathcal{H}(Y)$ to itself.

2) If we denote by $\underline{a}$ the bounded extension $\widehat{f(a)}$ (18.189) of $f(a)$ on the Hilbert space $H(Y)$, which denotes the completion of $\mathcal{H}(Y)$, then the map $a \mapsto \underline{a}$ is an *injective homomorphism* from $\mathrm{Quant}(Y)$ *to the group of unitary operators* on $H(Y)$.[459]

[459] One also says: a *faithful unitary representation* of $\mathrm{Quant}(Y)$ on $H(Y)$.

PROOF: One can check that the support of $f(a)(\Psi)$ is the image under a of the support of Ψ, and thus is compact (16.22). Since a commutes with the action of the torus (18.115), $f(a)(\Psi)$ satisfies the identity (18.179$\diamondsuit$) as does Ψ. Thus $f(a)$ is a map from $\mathcal{H}(Y)$ to $\mathcal{H}(Y)$. One can see immediately that f is a homomorphism and that $f(a)$ is linear and surjective (because it has $f(a^{-1})$ as its inverse). Liouville's theorem (16.99) shows that $f(a)$ is an isometry (definition (18.191)), hence of norm 1 and thus bounded. It follows from (18.189 to 191) that $f(a)$ has a unique bounded extension $\underline{a}$ on $H(Y)$, that $\underline{a}$ is unitary, and that $a \mapsto \underline{a}$ is a homomorphism. To show that this homomorphism is injective, it suffices to find, for $a \neq 1_Y$, a $\Psi \in \mathcal{H}(Y)$ such that $f(a)(\Psi) \neq \Psi$. As in the proof of (18.192), one can obtain such a Ψ by an elementary construction and using theorem (18.130). □

(18.197) If Γ is a *group of quantomorphisms* of Y, the restriction to Γ of the map $a \mapsto \underline{a}$ obviously defines a *unitary representation* of Γ. Applying this remark to the examples (18.159 to 18.178), we obtain representations of groups that play an essential role in quantum mechanics.

§19. Quantization of dynamical systems

The correspondence principle

(19.1) Let us call a *classical system* a dynamical system that conforms to the principles (13.1) or (13.70). A classical system is thus characterized, in relativistic mechanics as well as in Galilean mechanics, by a space of motions which is a symplectic manifold.

(19.2) Experiments show that this description is insufficient in certain cases (for example, specific heats at low temperature or black body radiation). There exist *phenomena*, called *quantum*, of which a common characteristic is that Planck's constant h plays a role.

(19.3) Thus one considers that true mechanics must be "quantum" and should contain classical mechanics as an *approximation*, which should be valid when h can be treated as being *very small*.

In fact, the relation between quantum mechanics and classical mechanics does not seem to reduce to a simple process of approximation: "the formulation of the fundamental principles of quantum mechanics ... is fundamentally impossible without the intervention of classical mechanics".[460]

(19.4) From this point of view, the *classical* description of a real system (thus *quantum* a priori) is a *partial* description, which can serve as an *approximation* when one considers Planck's constant to be small.



Thus there has to exist a classical system corresponding to every quantum mechanical system.[461] If we assume this *correspondence principle*, it is legitimate to start with the classical description of a system in order to construct its quantum mechanical description. This is what one calls the *quantization* of the classical system.[462]

---

[460] See Landau and Lifschitz [25].

[461] Provided, of course, one does not arbitrarily restrict the above definition of classical systems (for example by refusing to consider particles with spin and only allowing systems of material points).

[462] Note well the difference with the case of relativistic and Galilean mechanics. Although certain relativistic systems, as we have seen, have an approximate Galilean description, there is no *correspondence* which associates a *single* Galilean system to such a relativistic system (see especially (14.54, 55)). And hence, vice versa, neither does there exist a "relativization" algorithm.

(19.5) In this section we are going to investigate whether the geometrical construction we have developed in §18 — precisely under the name of "prequantization" — can play a role in this correspondence process. However, we will content ourselves with examining various cases as to how this geometric quantization leads to the usual structures of quantum mechanics.

Let us formulate the initial hypothesis:

(19.6) The space of classical motions of a dynamical system is a *prequantizable symplectic manifold.*[463]

(19.7) A first consequence of this hypothesis concerns the *spin* of particles. We have seen in detail in §18 that the spin must be of the form $n\hbar/2$ (n an integer) for the space of motions to be prequantizable. This is indeed what experiments show (the observed values of n range from 0 to more than 10). In the case of *massless* particles, one encounters the values $n = 1$ (neutrinos) and $n = 2$ (photons).[464] □

(19.8) Let us note that the principle (19.6) does not prohibit the description of a system composed of several noninteracting subsystems. The classical description is constructed by means of a direct product of symplectic manifolds (13.1.II), (13.70.II), which is thus prequantizable (theorem (18.52)). □

(19.9) If we study a mechanical system using *scattering theory*, we know that the scattering symplectomorphisms S_j^k (15.65) define the symplectic structure of an open subset $\tilde{U}'$ of the space $\tilde{U}$ of scattered motions (15.68).[465] Since $\tilde{U}$ must be prequantizable, $\tilde{U}'$ must be too (18.17). This imposes further conditions on the S_j^k, which supplement the relations (15.67). □

(19.10) In other physical systems for which we know a classical model (for example, systems of indistinguishable particles, certain atoms or molecules, or solids treated as harmonic oscillators), *one can see that condition* (19.6) *is in fact satisfied* (see (18.138) and (18.31)).

State vectors and observables

(19.11) The initial hypothesis (19.6) presupposes the existence of a prequantiza-

[463]Using the convention (18.28) which amounts to choosing the constant $\hbar$ as the unit of measurement for those quantities having the dimension of an action.

[464]The conjectured *graviton* would correspond to $n = 4$.

[465]$\tilde{U}'$ is the space of *unconstrained* motions.

tion (Y, P) of the space of motions U of an arbitrary classical system. We are now going to let the *prequantum manifold Y itself* enter the scene. We will call it the "prequantum manifold of the system".[466]

(19.12) DEFINITION: Let Y be the prequantum manifold of a dynamical system. By *state vectors* of the system we will mean those elements Ψ of the pre-Hilbert space $\mathcal{H}(Y)$ that satisfy

$$\diamondsuit \qquad \overline{\Psi} \cdot \Psi = 1 \qquad \text{(notation (18.183))}. \qquad \square$$

Looking back at definition (18.181), we see that $\diamondsuit$ can be written as

$$\int_U |\Psi(\xi)|^2 \, \varphi(dx) = 1 \,, \tag{19.13}$$

where φ denotes the *symplectic density* on U. This definition amounts to saying that the function

$$\rho(x) \equiv |\Psi(\xi)|^2 \tag{19.14}$$

defines a *probability measure*[467] *on the space of motions*, that is, a *state in the sense of statistical mechanics* (17.1). This makes possible a *probabilistic interpretation* for each state vector Ψ: the probability that the motion of the system belongs to an open subset Ω of U equals the integral[468,469]

$$\int_\Omega |\Psi(\xi)|^2 \, \varphi(dx) \,. \tag{19.15}$$

---

[466]We assume that nature defines the prequantization (Y, P) *up to equivalence* (18.117). And indeed, in the case of a system of indistinguishable particles, which admits two inequivalent prequantizations (18.138), experiments show that it is the Bose-Einstein prequantization which should be chosen if the spin is an integer and the Fermi-Dirac prequantization if the spin is a half-integer.

[467]We mean, of course, the completely continuous measure having $\rho(x)\varphi$ as its *density*.

[468]See (16.179), (16.124), and (16.136).

[469]If, in the case of a particle for example, we choose a time t_0 and an open set A in $\mathbf{R}^3$, the probability that the particle is in A at time t_0 equals

$$\int_{\Omega_{t_0,A}} |\Psi(\xi)|^2 \, \varphi(dx) \,,$$

where $\Omega_{t_0,A}$ is *the set of motions of the particle for which its trajectory* (in the sense of (14.4e)) *intersects the hypersurface $t = t_0$ in a point of A*. We thus obtain the *spatial density $f(\mathbf{r})$ of the probability of presence* by taking the image (in the sense of (16.149)) of the probability measure (19.14) under the map $x \mapsto \mathbf{r}$ (defined at time t_0 by the equations of the trajectory) and by writing this image in the form $\lambda \times f$, where λ denotes the *Euclidean measure* (16.94) on $\mathbf{R}^3$.

The *average value* of a dynamical variable u equals[470]

$$\int_U u\,|\Psi(\xi)|^2\,\varphi(dx)\,. \tag{19.16}$$

(19.17) But these rules are no longer applicable when one examines more closely the physical concept of average value (19.207) or state (19.215). The principal utility of the probabilistic interpretation is thus to show that statistical mechanics, as well as classical mechanics, is an approximation to quantum mechanics. □

On the other hand, we know that we can associate *to each dynamical variable* u an operator $\widehat{u}$ (18.192) such that

♠ $\widehat{u}$ is a hermitian operator on $\mathcal{H}(Y)$;

♡ the map $u \mapsto \widehat{u}$ is linear (over the real numbers);

(19.18) ♢ $$\widehat{1} = \underline{1}_{\mathcal{H}(Y)}\,;$$

if u and v are dynamical variables, we have

♣ $$\widehat{u} \circ \widehat{v} - \widehat{v} \circ \widehat{u} = -i\,\widehat{[u,v]}_P\,.$$ □

(19.19) It is thus natural to identify these operators $\widehat{u}$ with *observables* as Dirac has defined them,[471] since the results (19.18) have been chosen by him as the *principles* of quantum mechanics.

(19.20) EXAMPLE: Let us consider a conservative dynamical system, and let us choose (as in (12.32)) canonical coordinates p_j, ${}^jq$ for the *phase space at time* t. Every function $u(p_1, \ldots, p_N, {}^1q, \ldots, {}^Nq)$ can be seen as a dynamical variable *depending on* t. Hamilton's equations (12.111, 112) and the formula of Poisson (10.7) show that

$$\frac{du}{dt} = [H, u]_P\,, \tag{19.21}$$

where H is the Hamiltonian. Applying (19.18♣) and *taking the liberty, allowed because of* (19.18♡), *of differentiating under the hat*, we find

$$\frac{d}{dt}\,\widehat{u} = i\,[\widehat{H} \circ \widehat{u} - \widehat{u} \circ \widehat{H}]\,, \tag{19.22}$$



a formula also postulated by Dirac, as the *equation of motion*.

In the particular case of the canonical coordinates p_j, ${}^jq$ themselves,

[470] See (16.191).

[471] See Dirac [11] and [12].

we obtain the relations

$$\begin{aligned} &\hat{p}_j \circ \hat{p}_k - \hat{p}_k \circ \hat{p}_j = 0 \qquad {}^j\hat{q} \circ {}^k\hat{q} - {}^k\hat{q} \circ {}^j\hat{q} = 0 && \forall\, j,k \\ &\hat{p}_j \circ {}^k\hat{q} - {}^k\hat{q} \circ \hat{p}_j = 0 && j \neq k \\ &\hat{p}_j \circ {}^j\hat{q} - {}^j\hat{q} \circ \hat{p}_j = -i \cdot 1_{\mathcal{H}(Y)} && \forall\, j \\ &\frac{d\hat{p}_j}{dt} = i\,[\hat{H} \circ \hat{p}_j - \hat{p}_j \circ \hat{H}] \qquad \frac{d^j\hat{q}}{dt} = i\,[\hat{H} \circ {}^j\hat{q} - {}^j\hat{q} \circ \hat{H}] && \forall\, j\,. \end{aligned} \tag{19.23}$$

It is the set of relations (19.23) chosen by Pauli and Heisenberg as the principles of quantum mechanics.

The formulation of Planck's condition

(19.24) As its name suggests, one might think that it is a *state vector* which describes "that which happens" during an experiment. Hypothesizing this, how do we determine it?

We will formulate, under the name of *Planck's condition*, a property of state vectors. Later on we will examine cases in which it seems to hold.

(19.25) Let us assume that an *isotropic foliation* of the space of motions U has been given. We have seen (18.113) how we can, in favorable circumstances, lift this foliation to a *Planck foliation* of the prequantum manifold Y. We will say that a state vector Ψ of the system satisfies *Planck's condition* (relative to the foliation) if Ψ is a first integral, that is, if $\Psi(\xi)$ is *constant on every leaf*. □

(19.26) Let us assume that a certain number of dynamical variables u_j ($j = 1, 2, \ldots, p$) are in *involution*, that is, their mutual Poisson brackets are zero:

$$[u_j, u_k]_P \equiv 0 \qquad \forall\, j,k\,. \tag{19.27}$$

Thus, as we have already observed (theorem (9.37)), if the equations

$$u_j = \text{constant} \tag{19.28}$$

define connected submanifolds of constant dimension of the space of motions, then these manifolds are leaves of a *coisotropic foliation*. At each point x of U, the vector space orthogonal to the leaf passing through x is the *kernel* of the 2-form σ induced on the leaf (9.35). Thus on each coisotropic leaf (and hence on all of U) this kernel determines an *isotropic*

(19.29) *foliation* (9.8). One can easily see that the corresponding *Planck lift* is the foliation $\xi \mapsto H$ of Y defined as follows:

(19.30) H is the vector space generated by the vectors

$$\underset{u_j}{\delta}\,\xi - u_j\,\underset{1}{\delta}\,\xi\,. \qquad \square$$

The corresponding Planck condition thus will be satisfied by a state vector Ψ if

(19.31) $$\underset{u_j}{\delta}\,\Psi(\xi) \equiv u_j\,\underset{1}{\delta}\,\Psi(\xi) \qquad\qquad \forall\, j\,.$$

Taking (18.179$\diamondsuit$) and (18.192$\bowtie, \diamondsuit$) into account, Planck's condition takes the following form.

(19.32) If $u \equiv f(u_1, \ldots, u_p)$, we have $[\hat{u}(\Psi)](\xi) \equiv u \times \Psi(\xi)$. $\square$

(19.33) The set of differentiable functions Ψ on Y satisfying the above condition obviously form a complex vector space (the *Planck space*). If v is a dynamical variable and if there exist functions f_j such that

(19.34) $$[v, u_j]_P \equiv f_j(u_1, \ldots, u_p) \qquad\qquad [j = 1, \ldots, p]\,,$$

the reader can verify (by applying (18.144)) that *the operator $\hat{v}$ leaves the Planck space invariant.*

Stationary states

Let us consider a conservative system. Its energy E is a dynamical variable and the manifolds $E =$ constant define a coisotropic foliation of U.

(19.35) We will say that a state vector Ψ of the system is *stationary* if it satisfies the corresponding Planck condition:

$\diamondsuit$ $$\hat{E}(\Psi) = E \times \Psi\,. \qquad \square$$

(19.36) Let us write this condition on the evolution space, traditionally considered to be the direct product of the phase space with time.

We know that $\hat{E}(\Psi) = -i\,\underset{E}{\delta}\,\Psi$ (18.192) and that $\underset{E}{\delta}\,\Psi = -\frac{\partial \Psi}{\partial t}$ (see (12.152)). Thus (19.35) can be written as

$$i\,\frac{\partial \Psi}{\partial t} = E \times \Psi\,,$$

or as

$$\Psi = \Psi_0\,\mathrm{e}^{-i\,Et}\,,$$

where Ψ_0 is a function "independent of time". Thus Ψ depends harmonically on t with period $T = 2\pi/E$. Taking into account the dimensions of the units of energy and time (13.13), we have (18.29)

$$\frac{ET}{2\pi} = \hbar\,,$$

or, introducing the frequency $\nu = 1/T$ and Planck's constant $h = 2\pi\hbar$,

(19.37) $$E = h\nu\,.$$



It is exactly this formula which Planck took as a hypothesis in his original work on the theory of quanta. It allows — at least formally — to translate in terms of *frequencies* results established in terms of *energy*; for example, the Doppler effect (which we have treated in (15.113)) or the spectroscopic interpretation of figure 17.IV. This law of Planck is used twice in Bohr's law of radiation:

(19.38) $$E' - E'' = h\nu\,,$$

where E' and E'' are the energies of the system in two given *stationary* states and where ν is the frequency of the emitted (or absorbed) photon in the transition from one state to the other.

The formation of wave equations

(19.39) Let U be the space of motions of a dynamical system and let (Y, P) be its prequantization.

Let G be a dynamical group of U having a moment μ. Let us assume that G has a *normal abelian Lie subgroup* $\tilde{G}$ and that the symplectic cohomology *of the dynamical group* $\tilde{G}$ is zero. Let $\mathcal{G}$ and $\tilde{\mathcal{G}}$ be the Lie algebras of G and $\tilde{G}$, respectively, and let f be the symplectic cocycle of $\mathcal{G}$ defined in (11.17).

From the formulæ (11.17♯), (9.22), and (11.9) we easily obtain

(19.40) $$[\mu\cdot Z, \mu\cdot Z']_P \equiv \mu\cdot[Z, Z'] + f(Z)(Z') \qquad \forall\, Z, Z' \in \mathcal{G}\,.$$

This shows that if we choose a basis $Z_1, \ldots, Z_p$ for $\tilde{\mathcal{G}}$, then the dynamical variables

(19.41) $$u_j \equiv \mu\cdot Z_j$$

are *in involution*.

(19.42) We can then apply construction (19.26) and define a Planck space (by means of (19.32)). If Z is a fixed element of $\mathcal{G}$, equation (19.40)

shows that $[\mu \cdot Z, \mu \cdot Z_j]_P$ is a function of $u_1, \ldots, u_p$, because $[Z, Z_j]$ is a fixed element of $\tilde{\mathcal{G}}$ (since $\tilde{G}$ is a normal subgroup of G, $\tilde{\mathcal{G}}$ is an ideal of $\mathcal{G}$) and because $f(Z)(Z_j)$ is constant.

We see that condition (19.34) is satisfied. Hence *for every* $Z \in \mathcal{G}$, *the observable* $\widehat{\mu \cdot Z}$ *leaves the Planck space invariant.* □

(19.43) This construction can be applied *to every isolated dynamical system* by taking for G either the Poincaré group or the Galilei group (according to whether the treatment is relativistic or not), and for $\tilde{G}$ the group of *space-time translations*. We can then choose for the functions u_j the *energy* and the three components of *linear momentum*. The Planck space so obtained consists of stationary state vectors (in the sense of (19.35)), whatever the choice of an inertial reference frame.[472]

Thus this space will be invariant under the observables $\widehat{\mu \cdot Z}$ associated to various elements Z of $\mathcal{G}$ (relativistic moments or not). □

(19.44) We are now going to treat various cases of *elementary particles*. In order to give a spatio-temporal interpretation of the results, we will use — in addition to the space of motions U and the associated prequantum manifold Y — the evolution space V (Fig. 19.I) and the manifold W, the *fibered product* of V and Y, which we have already met in the special case of figure 18.V.

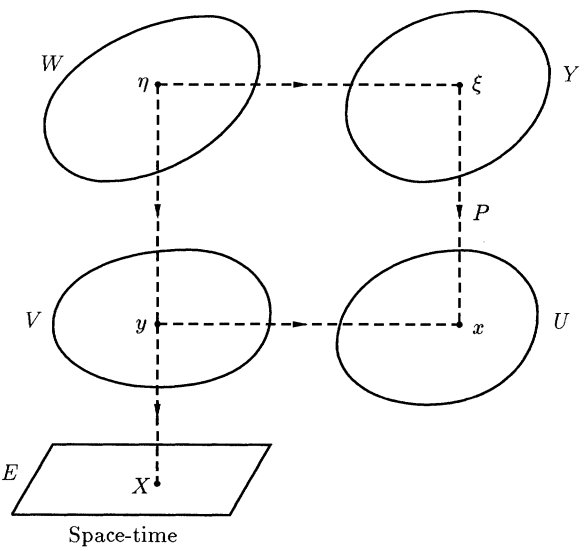


Fig 19.I

[472] We thus obtain a *spatial Planck relation* $\|\mathbf{p}\| = h/\lambda$, where λ is the *wavelength*. In this way, for a nonrelativistic photon, the formulæ $\mathbf{p} = k\,\mathbf{u}$ (14.53♡) and $s = \hbar$ give $\lambda = 2\pi\, s/k$ (see (15.89)).

The nonrelativistic material point

(19.45) The evolution space V is described by the triple $y \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \end{pmatrix}$ (12.26). Let us define W as the *direct product* of V and the torus T, described by the variable

$$\eta \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \\ \zeta \end{pmatrix} \qquad \zeta \in T\,, \tag{19.46}$$

and let us consider the 1-form ϖ on W

$$\varpi(\delta\eta) \equiv \langle m\,\mathbf{v},\, \delta\mathbf{r}\rangle - \tfrac{1}{2}\,m\,v^2\,\delta t + \frac{\delta\zeta}{i\,\zeta} \qquad (v = \|\mathbf{v}\|)\,. \tag{19.47}$$

One can easily determine the *characteristic foliation* of ϖ:

$$[\delta\eta \in \ker(\varpi)\cap\ker(\mathrm{d}\varpi)] \iff [\delta\mathbf{v} \equiv 0,\ \delta\mathbf{r} \equiv \mathbf{v}\,\delta t,\ \frac{\delta\zeta}{i\,\zeta} \equiv -m\frac{v^2}{2}\,\delta t]\,. \tag{19.48}$$

Hence the *quotient manifold* Y is described by the variable

$$\xi \equiv \begin{pmatrix} \boldsymbol{\rho} \\ \mathbf{v} \\ z \end{pmatrix} \qquad \begin{aligned} \boldsymbol{\rho} &\equiv \mathbf{r} - \mathbf{v}\,t \\ z &\equiv \zeta\,\mathrm{e}^{imv^2\,t/2}\,. \end{aligned} \tag{19.49}$$

The 1-form ϖ passes naturally to the quotient (theorem (5.21)) giving

$$\varpi(\delta\xi) \equiv \varpi(\delta\eta) \equiv \langle m\,\mathbf{v},\, \delta\boldsymbol{\rho}\rangle + \frac{\delta z}{i\,z}\,. \tag{19.50}$$

A computation gives the exterior derivative $\sigma \equiv \mathrm{d}\varpi$:

$$\begin{aligned} \sigma(\delta\eta)(\delta'\eta) &\equiv \langle m\,\delta\mathbf{v},\, \delta'\mathbf{r} - \mathbf{v}\,\delta t'\rangle - \langle m\,\delta'\mathbf{v},\, \delta\mathbf{r} - \mathbf{v}\,\delta t\rangle \\ &\equiv \langle m\,\delta\mathbf{v},\, \delta'\boldsymbol{\rho}\rangle - \langle m\,\delta'\mathbf{v},\, \delta\boldsymbol{\rho}\rangle\,, \end{aligned} \tag{19.51}$$

which one should recognize as the *Lagrange form* of the free particle of mass m (12.40). The space of motions U is described by the variable

$$x \equiv \begin{pmatrix} \boldsymbol{\rho} \\ \mathbf{v} \end{pmatrix}. \tag{19.52}$$

We see that σ is defined on the four manifolds W, V, Y, and U, and that (Y, P) $[x \equiv P(\xi)]$ *prequantizes* the system. We know that this system is monoprequantizable (18.135).

Now let

$$Z = \begin{bmatrix} j(\boldsymbol{\omega}) & \boldsymbol{\beta} & \boldsymbol{\gamma} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix} \tag{19.53}$$

be an element of the Lie algebra $\mathcal{G}$ of the Galilei group G (notation (12.118)). We know that G acts on V preserving σ, and that the vector $\delta y \equiv Z_V(y)$ is given by (12.119)

$$\begin{aligned} \delta t &\equiv \varepsilon \\ \delta \mathbf{r} &\equiv \boldsymbol{\omega} \times \mathbf{r} + \boldsymbol{\beta} t + \boldsymbol{\gamma} \\ \delta \mathbf{v} &\equiv \boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\beta}\,, \end{aligned} \tag{19.54}$$

and thus

$$\delta x \equiv \begin{pmatrix} \delta \boldsymbol{\rho} \\ \delta \mathbf{v} \end{pmatrix} \equiv \begin{pmatrix} \boldsymbol{\omega} \times \boldsymbol{\rho} + \boldsymbol{\gamma} - \mathbf{v}\varepsilon \\ \boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\beta} \end{pmatrix}. \tag{19.55}$$

We know (11.9) that

$$\delta x \equiv \underset{\mu \cdot Z}{\delta} x \tag{19.56}$$

(notation (9.20)), where μ denotes the *Galilean moment*; it follows that $x \mapsto \delta x$ is an *infinitesimal symplectomorphism.*

Let us recall the formulæ (12.123) and (12.122):

$$\mu \equiv \{\mathbf{l}, \mathbf{g}, \mathbf{p}, E\} \qquad \mu \cdot Z \equiv \langle \mathbf{l}, \boldsymbol{\omega} \rangle - \langle \mathbf{g}, \boldsymbol{\beta} \rangle + \langle \mathbf{p}, \boldsymbol{\gamma} \rangle - E\varepsilon\,. \tag{19.57}$$

It follows from (12.125) that we can *choose* the following moment μ:

$$\begin{aligned} \mathbf{l} &\equiv m\,\mathbf{r} \times \mathbf{v} \equiv m\,\boldsymbol{\rho} \times \mathbf{v} \\ \mathbf{g} &\equiv m\,[\mathbf{r} - \mathbf{v}\,t] \equiv m\,\boldsymbol{\rho} \\ \mathbf{p} &\equiv m\,\mathbf{v} \\ E &\equiv \tfrac{1}{2}\,m\,v^2\,. \end{aligned} \tag{19.58}$$

(19.59) The Galilei group is *liftable* to Y (18.129), *but it is not quantizable*, because its symplectic cohomology is not zero (18.155). We can, however, lift the vector field $x \mapsto \delta x$ to Y by defining

$$\varpi(\delta \xi) \equiv \mu \cdot Z\,. \tag{19.60}$$

The formulæ (19.55, 57, 58) immediately give

$$\frac{\delta z}{i\,z} \equiv \tfrac{1}{2}\,m v^2 \varepsilon - m\,\langle \boldsymbol{\rho}, \boldsymbol{\beta} \rangle\,. \tag{19.61}$$

(19.62) Moreover, we have $\delta \xi \equiv \underset{\mu \cdot Z}{\delta} \xi$ (notation (18.140)). Theorem (18.145) and a closer look at the differential equation $d\xi/dt \equiv \delta \xi$ show that the map $\xi \mapsto \delta \xi$ is an *infinitesimal quantomorphism.*

The formulæ (19.54, 61, 49) allow us to define a vector field $\eta \mapsto \delta\eta$ on W for which we immediately find that

$$\frac{\delta\zeta}{i\zeta} \equiv -m\,\langle \boldsymbol{\beta}, \mathbf{r}\rangle\,. \tag{19.63}$$

□

If $\tilde{\mathcal{G}}$ is the Lie algebra of the group $\tilde{G}$ of space-time translations, we have

$$[Z \in \tilde{\mathcal{G}}] \iff [\boldsymbol{\omega} = \boldsymbol{\beta} = 0]\,. \tag{19.64}$$

The *Planck space* defined in (19.43) is the set of differentiable functions Ψ on Y satisfying

$$\begin{cases} \underset{1}{\delta}\,\Psi(\xi) \equiv i\,\Psi(\xi) \\ \underset{\mu\cdot Z}{\delta}\,\Psi(\xi) \equiv i\,\mu\cdot Z\Psi(\xi) \qquad\qquad \forall Z \in \tilde{\mathcal{G}}\,. \end{cases} \tag{19.65}$$

The preceding formulæ allow us to solve these equations without difficulty:

$$\Psi(\xi) \equiv \overline{\psi} z \mathrm{e}^{im\,\langle \mathbf{v}, \boldsymbol{\rho}\rangle} \equiv \overline{\psi}\zeta \mathrm{e}^{im[\langle \mathbf{v}, \mathbf{r}\rangle - v^2 t/2]}\,, \tag{19.66}$$

where ψ is a complex valued differentiable function *of* $\mathbf{v}$ *only*.

According to (19.42), this Planck space is invariant under the operators $\widehat{\mu\cdot Z} \equiv -i\,\underset{\mu\cdot Z}{\delta}$ $(\forall Z \in \mathcal{G})$. Indeed a computation shows that the effect of applying the operator $\widehat{\mu\cdot Z}$ on Ψ amounts to replacing ψ by the function

$$i\frac{\partial\psi}{\partial\mathbf{v}}(\boldsymbol{\omega}\times\mathbf{v} + \boldsymbol{\beta}) + m[\langle \mathbf{v}, \boldsymbol{\gamma}\rangle - v^2\varepsilon/2]\,\psi\,, \tag{19.67}$$

which is indeed a function of $\mathbf{v}$ only, since $\boldsymbol{\omega}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and ε are constants.

(19.68) EXAMPLE: Let τ be an instant of *time*. If we choose

$$\boldsymbol{\omega} = 0\,, \quad \boldsymbol{\beta} = \begin{pmatrix} -1/m \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} \tau/m \\ 0 \\ 0 \end{pmatrix}, \quad \varepsilon = 0\,,$$

it follows that $\mu\cdot Z$ is equal to the first component r_1 of the position vector $\mathbf{r}$ of the particle at time τ. The associated operator $\hat{r}_1$ (the "position operator") changes ψ into

$$-\frac{i}{m}\frac{\partial\psi}{\partial v_1} + \tau v_1\,\psi\,. \tag{16.69}$$

□

(19.70) It is easy to verify that the elements of this Planck space (19.66) *are not* state vectors, because their support is not compact (see (19.241)).†

(19.71) If the support of the map $\mathbf{v} \mapsto \psi$ is compact, we can introduce the variable (compare (19.66))

$$\tilde{\psi} \equiv \int_{\mathbf{R}^3} \psi \, \mathrm{e}^{-im\,[(\mathbf{v},\mathbf{r})-v^2 t/2]} \, \varphi(d\mathbf{v}) \,, \tag{19.72}$$

where φ is the Euclidean density.

It follows that $\tilde{\psi}$ is a differentiable function of $\mathbf{r}$ and t. Differentiating under the integral sign shows that $\tilde{\psi}$ satisfies the partial differential equation

$$\frac{1}{2m} \triangle \tilde{\psi} - i \frac{\partial \tilde{\psi}}{\partial t} \equiv 0 \,. \tag{19.73}$$

(19.74) It is this equation which Schrödinger has proposed a priori to describe the nonrelativistic free particle in quantum mechanics. The validity of this equation has been confirmed experimentally many times. The variable $\tilde{\psi}$ is called the *wave function* of the particle.[473]

As the theory of the Fourier transform shows, the passage from the function $\mathbf{v} \mapsto \psi$ to the function $\begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} \mapsto \tilde{\psi}$[474] is injective, and thus we can describe the operators $\widehat{\mu \cdot Z}$ by their action on $\tilde{\psi}$. Using the formulæ (19.57), (19.67), (19.72), and (19.69) we obtain the following list:

$$\left\{ \begin{array}{ll} \widehat{p}_j : & i \dfrac{\partial}{\partial r_j} \\ \widehat{r}_j : & r_j - \dfrac{i}{m}[t-\tau]\dfrac{\partial}{\partial r_j} \\ \widehat{E} : & -i \dfrac{\partial}{\partial t} \,. \end{array} \right. \tag{19.75}$$

(19.76) The reader can verify directly that the "position operators" $\widehat{r}_j$ preserve the solutions of the Schrödinger equation. This is not the case for the operators of multiplication by r_j, with which they are confused if one does not distinguish the times t and τ. As for $\widehat{p}_j$ and $\widehat{E}$, they are the usual operators of quantum mechanics.

---

†Editors' note: This particular problem has been solved by the introduction of either half-densities, half-forms, or polarizers. See Guillemin and Sternberg [16] and Duval [14].

[473]One obtains an equivalent quantization by changing i into $-i$ in all of the above formulæ. Thus one can also write the Schrödinger equation in the form $\frac{1}{2m}\triangle\tilde{\psi} + i\frac{\partial\tilde{\psi}}{\partial t} \equiv 0$.

[474]Or even $\mathbf{r} \mapsto \tilde{\psi}$, if t has been fixed.

One should also note that the Schrödinger equation expresses, on $\tilde{\psi}$, the relation

$$\widehat{E} = \frac{1}{2m}\left[\widehat{p}_1^2 + \widehat{p}_2^2 + \widehat{p}_3^2\right], \tag{19.77}$$

giving the Hamiltonian operator $\widehat{E}$ as a function of the linear momentum operators $\widehat{p}_j$. This relation is an immediate consequence of the relation

$$E = \frac{1}{2m}\left[p_1^2 + p_2^2 + p_3^2\right] \qquad \text{(compare (19.58))} \tag{19.78}$$

and of theorem (19.32). It was formulated *a priori* in the original conception of the Schrödinger equation. □

The Galilei group (12.115) lifts to Y and W as the group of maps

$$\begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \\ \zeta \end{pmatrix} \mapsto \begin{pmatrix} t+e \\ A(\mathbf{r}) + \mathbf{b}t + \mathbf{c} \\ A(\mathbf{v}) + \mathbf{b} \\ \zeta \mathrm{e}^{-im\left[\langle \mathbf{b}, A(\mathbf{r})\rangle + \mathbf{b}^2 t/2 + \text{constant}\right]} \end{pmatrix}, \tag{19.79}$$

which gives a *quantizable extension* (18.156) of dimension 11 of the Galilei group.

(19.80) The reader can check directly that this extension preserves the Planck space, by computing its action on the wave function $\tilde{\psi}$.[475]

The relativistic material point

(19.81) Carrying over the preceding calculations to the case of a relativistic material point leads to the following results (Fig. 19.I still applies):

$$\begin{cases} y \equiv \begin{pmatrix} X \\ I \end{pmatrix} & \overline{I}\cdot I \equiv 1\,,\ I \text{ future pointing (notation (14.24))}. \\ \eta \equiv \begin{pmatrix} X \\ I \\ \zeta \end{pmatrix} & \zeta \in T \\ \xi \equiv \begin{pmatrix} \Xi \\ I \\ z \end{pmatrix} & \Xi \equiv [1 - I\cdot\overline{I}]X \qquad \overline{I}\cdot\Xi \equiv 0\,, \\ & z \equiv \zeta\, \mathrm{e}^{-im\overline{I}\cdot X} \end{cases} \tag{19.82}$$

[475] V. Bargmann has shown the cohomological origin of the failure of naïve attempts to define the action of the Galilei group on wave functions (see (19.59)).

$$\varpi(\delta\eta) \equiv \varpi(\delta\xi) \equiv -m\overline{I}\cdot\delta X + \frac{\delta\zeta}{i\,\zeta} \equiv m\ \overline{\Xi}\cdot\delta I + \frac{\delta z}{i\,z}\,, \tag{19.83}$$

and

$$\sigma(\delta x)(\delta' x) \equiv m[\overline{\delta X}\cdot\delta' I - \overline{\delta' X}\cdot\delta I]\,. \tag{19.84}$$

If we denote by $Z = \begin{bmatrix} \Lambda & \Gamma \\ 0 & 0 \end{bmatrix}$ (notation (13.54)) an element of the Lie algebra $\mathcal{G}$ of the Poincaré group G, the associated derivation δ is given by

$$\begin{aligned} \delta X &\equiv \Lambda X + \Gamma & \qquad \delta\Xi &\equiv \Lambda\Xi + \left[1 - I\cdot\overline{I}\right]\Gamma \\ \delta I &\equiv \Lambda\cdot I \qquad \delta\zeta \equiv 0 & \frac{\delta z}{i\,z} &\equiv -m\,\overline{I}\cdot\Gamma\,, \end{aligned} \tag{19.85}$$

and the moment μ of the Poincaré group by

$$\mu\cdot Z \equiv m\left[\overline{X}\cdot\Lambda\cdot I - \overline{I}\cdot\Gamma\right]. \tag{19.86}$$

The Lie algebra of the group of space-time translations is obtained by taking $\Lambda = 0$. The associated Planck space is the set of all functions Ψ such that

$$\Psi(\xi) \equiv \overline{\psi}z \equiv \overline{\psi}\zeta e^{-im\overline{I}\cdot X}\,, \tag{19.87}$$

where ψ is a complex valued function of I which is differentiable on the manifold H defined by the equations

$$\overline{I}\cdot I \equiv 1\,, \qquad I \text{ future pointing.}^{476} \tag{19.88}$$

The operator $\widehat{\mu\cdot Z}$ acts on the Planck space by changing ψ into

$$i\frac{\partial\psi}{\partial I}(\Lambda\cdot I) - m\,\overline{I}\cdot\Gamma\psi\,. \tag{19.89}$$

If the support of the function $I \mapsto \psi$ is compact, we can define $\tilde{\psi}$ by

$$\tilde{\psi} = \int_H \psi\, e^{im\,\overline{I}\cdot X}\,\varphi(dI)\,, \tag{19.90}$$

where φ denotes the Euclidean density on H.[477] $\tilde{\psi}$ is a differentiable complex valued function of X satisfying the *Klein-Gordon equation*

$$\Box\tilde{\psi} + m^2\,\tilde{\psi} \equiv 0\,. \tag{19.91}$$

[476] $\overline{I}\cdot I = 1$ is the equation of a manifold with two connected components ("the two-sheeted hyperboloid"). H is one of the components; the other, H', is its opposite.

[477] The tangent space to H at I is orthogonal to I. It is a 3-dimensional negative Euclidean space.

Note that we can also obtain a solution of the Klein-Gordon equation by taking the integral (19.90) over the manifold H', the opposite of H, which is described by the unit vectors that are *past pointing*. Observing that

$$\varphi(dI) = \frac{\varphi(d\mathbf{p})}{m^2 E} \qquad \begin{cases} E = \sqrt{m^2 + p^2} & (14.28) \\ \varphi(d\mathbf{p}) = \text{the volume element of } \mathbf{R}^3\,, \end{cases} \tag{19.92}$$

we can write (19.90) in the form

$$\tilde{\psi} \equiv \int_{\mathbf{R}^3} \psi \, \frac{e^{i\,[-\langle \mathbf{p},\mathbf{r}\rangle + t\sqrt{p^2+m^2}]}}{m^2\sqrt{m^2+p^2}}\, \varphi(d\mathbf{p})\,. \tag{19.93}$$

It follows that for every value of $\mathbf{r}$ the frequency spectrum of $\tilde{\psi}$ is positive (it starts at $\nu = m/2\pi = m/h$), whereas it will be negative on H'. This is the reason why $\tilde{\psi}$ is called a *positive frequency solution* of the Klein-Gordon equation. (19.94)

And it is indeed by these solutions that one usually represents, in quantum mechanics, particles without spin.

The nonrelativistic particle with spin $\frac{1}{2}$

As we have seen in (18.55), we can prequantize the nonrelativistic particle with spin $\frac{1}{2}$ by means of the direct product algorithm (18.52), using the prequantization (18.45) of the sphere S_2 and the prequantization of the particle without spin. This leads to the following results. (19.95)

$$\begin{cases} \eta \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \\ \boldsymbol{\zeta} \end{pmatrix} & \boldsymbol{\zeta} \in \mathbf{C}^2\,, \quad \|\boldsymbol{\zeta}\| \equiv 1\,, \\ y \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \\ \mathbf{u} \end{pmatrix} & u_j = \overline{\boldsymbol{\zeta}}\,\sigma_j\,\boldsymbol{\zeta}\,, \quad \|\mathbf{u}\| \equiv 1 \quad \text{(notation (18.43))}\,, \\ \xi \equiv \begin{pmatrix} \boldsymbol{\rho} \\ \mathbf{v} \\ \mathbf{z} \end{pmatrix} & \boldsymbol{\rho} \equiv \mathbf{r} - \mathbf{v}t \\ & \mathbf{z} \equiv \boldsymbol{\zeta}\, e^{im\,v^2 t/2}\,, \quad \mathbf{z} \in \mathbf{C}^2\,, \quad \|\mathbf{z}\| \equiv 1\,, \end{cases} \tag{19.96}$$

and

$$\begin{aligned} \varpi(\delta\eta) &\equiv \langle m\,\mathbf{v},\, \delta\mathbf{r}\rangle - \tfrac{1}{2}\,mv^2\,\delta t + \operatorname{Re}\left(\frac{\overline{\boldsymbol{\zeta}}\,\delta\boldsymbol{\zeta}}{i}\right) \\ &\equiv m\,\langle \mathbf{v},\, \delta\boldsymbol{\rho}\rangle + \operatorname{Re}\left(\frac{\overline{\mathbf{z}}\,\delta\mathbf{z}}{i}\right). \end{aligned} \tag{19.97}$$

Let E_1 be the set of functions Ψ such that

$$\Psi(\xi) \equiv \overline{\psi} \cdot \mathbf{z}\, e^{im\langle \mathbf{v}, \boldsymbol{\rho}\rangle} \equiv \overline{\psi} \cdot \boldsymbol{\zeta}\, e^{im[\langle \mathbf{v}, \mathbf{r}\rangle - v^2 t/2]}, \tag{19.98}$$

where ψ is a $\mathbf{C}^2$-valued differentiable function of $\mathbf{v}$ only. It is easy to verify the conditions (19.65), and thus E_1 is a *vector subspace of the Planck space*.

(19.99) One can check that E_1 is invariant under the operators $\widehat{\mu \cdot Z}$, where Z belongs to the Lie algebra of the Galilei group. This applies in particular to the operators "angular momentum", "linear momentum", "position", and "energy", which the reader can compute himself. As in the case of a particle without spin, we can associate to Ψ the variable

$$\widetilde{\psi} \equiv \int_{\mathbf{R}^3} \psi\, e^{-im[\langle \mathbf{v}, \mathbf{r}\rangle - v^2 t/2]}\, \varphi(d\mathbf{v}) \qquad \text{(see (19.72))}. \tag{19.100}$$

(19.101) $\widetilde{\psi}$ is an element of $\mathbf{C}^2$ *whose two components are solutions of the Schrödinger equation*. One should recognize the description of the electron as proposed by *Pauli*.

(19.102) As the theorem of barycentric decomposition (13.34) shows, we can make the group SO(3) act on U in addition to the Galilei group (this action consists of turning the spin vector $\mathbf{u}/2$ without modifying the trajectory). The corresponding moments are simply the components of the spin. Let us choose the third component $\frac{1}{2} u_3 \equiv \frac{1}{2} \overline{\mathbf{z}} \cdot \sigma_3 \cdot \mathbf{z}$. One easily finds

$$\underset{u_3/2}{\delta} \boldsymbol{\rho} \equiv 0, \qquad \underset{u_3/2}{\delta} \mathbf{v} \equiv 0, \qquad \underset{u_3/2}{\delta} \mathbf{z} \equiv \tfrac{1}{2} \sigma_3 \mathbf{z},$$

which shows that $\underset{u_3/2}{\delta}$ *also preserves the space* E_1. Using (19.98) and (19.100) it follows that the associated observable $\hat{u}_3/2$ maps ψ to $-\frac{1}{2}\sigma_3\psi$ and $\widetilde{\psi}$ to $-\frac{1}{2}\sigma_3\widetilde{\psi}$. Referring back to formula (18.43) for the Pauli matrices, we see that this operator has eigenvalues $-\frac{1}{2}$ and $+\frac{1}{2}$ (called "spin down" and "spin up"), and that the corresponding eigenvectors are those ψ (or $\widetilde{\psi}$) that have only one nonzero component. This is exactly the way the spin operators in quantum mechanics behave.

The space E_1 *is not the full Planck space*. One can check that the conditions (19.65) are also satisfied by Ψ of the form

$$\Psi(\xi) \equiv T(\mathbf{z})(\mathbf{z})(\overline{\mathbf{z}})\, e^{im\langle \mathbf{v}, \boldsymbol{\rho}\rangle}, \tag{19.103}$$

where T is a trilinear map depending on $\mathbf{v}$ only.[478] These solutions do not seem to have a physical interpretation. We see that the Planck condition

[478] And more generally, for Ψ in the space E_p $(p = 1, 2, \ldots)$:

is not sufficient for selecting those "state vectors" which are observed in nature.†

The relativistic particle with spin $\frac{1}{2}$

(19.104) Using the prequantization (18.80) we find:

$$\eta \equiv \begin{pmatrix} X \\ \zeta \end{pmatrix}, \qquad \varpi(\delta\eta) \equiv \mathrm{Re}(-i\,\overline{\zeta}\delta\zeta) - m\,\overline{\zeta}\gamma(\delta X)\zeta\,, \tag{19.105}$$

where X is a point of space-time and ζ a Dirac spinor (18.59) satisfying (notation (18.62, 63))

$$\overline{\zeta}\cdot\zeta \equiv 1 \qquad\qquad \overline{\zeta}\gamma_5\zeta \equiv 0\,. \tag{19.106}$$

Let us recall that there exists a unit future pointing vector I, which is a function of ζ, such that (18.67, 68)

$$\overline{\zeta}\cdot\gamma(\delta X)\cdot\zeta \equiv \overline{I}\cdot\delta X\,. \tag{19.107}$$

One can verify that the Planck conditions hold on the space E_1 of functions Ψ of the form

$$\Psi(\xi) \equiv \overline{\psi}\cdot\zeta\,\mathrm{e}^{-im\overline{I}\cdot X}\,, \tag{19.108}$$

where ψ is a Dirac spinor which is a *function of I only*.

Note that the above relation does not completely define the map $I \mapsto \psi$. Indeed, from (19.106) and (19.107) one can deduce the relation

$$\gamma(I)\zeta \equiv \zeta\,. \tag{19.109}$$

Since $\gamma(I)^2 = 1$ (18.64), we see that $\frac{1+\gamma(I)}{2}$ and $\frac{1-\gamma(I)}{2}$ are *supplementary projections*. Decomposing ψ according to the (two) associated subspaces, and taking into account that $\gamma(I)$ is hermitian (18.64), it follows that the second "component" *does not contribute* to the expression (19.108).

◇ $\qquad \Psi(\xi) \equiv T(\mathbf{z})\ldots(\mathbf{z})(\overline{\mathbf{z}})\ldots(\overline{\mathbf{z}})\,\mathrm{e}^{im\,\langle \mathbf{v},\,\boldsymbol{\rho}\rangle}\,,$

where the tensor T is $2p-1$ times complex linear with p arguments $\mathbf{z}$ and $p-1$ arguments $\overline{\mathbf{z}}$. Note that these spaces are not disjoint, because of the relation $\overline{\mathbf{z}}\cdot\mathbf{z} \equiv 1$. We can make them disjoint by assuming that the coefficients of T are such that ◇ is a *harmonic* function of $\mathbf{z}$ on $\mathbf{C}^2$.

†Editors' note: This problem has been solved by the introduction of complex polarizations, that is, complex Planck "foliations" (19.25). These will "automatically" exclude the spaces E_p for $p > 1$ (see also (19.121), (19.134)). However, for complex polarizations no group theoretic interpretation in the sense of (19.39–43) seems to be known.

Thus we may assume that ψ reduces to its first "component", a condition which can be written as

(19.110) $$\psi \equiv \gamma(I)\,\psi\,.^{479}$$

With this assumption the correspondence between Ψ and $I \mapsto \psi$ becomes injective. In case the function $I \mapsto \psi$ has compact support (on the manifold H (19.88) over which I varies), we can define a variable $\tilde{\psi}$ by

(19.111) $$\tilde{\psi} \equiv \int_H \psi\, e^{im\overline{I}\cdot X}\,\varphi(dI) \qquad \text{(compare with (19.90))}\,.$$

Differentiating under the integral sign with respect to X, we find

(19.112) $$\delta\tilde{\psi} \equiv im \int_H \overline{I}\cdot\delta X\psi\, e^{im\overline{I}\cdot X}\,\varphi(dI)\,.$$

Taking (19.110) into account, we deduce the *Dirac equation*[480]

(19.113) $$\sum_{\mu=1}^{4} \gamma^\mu \partial_\mu \tilde{\psi} - im\,\tilde{\psi} \equiv 0\,.^{481}$$

(19.114) Note that we would have found a second family of solutions of the Dirac equation by replacing H in (19.112) with the opposite manifold H' (19.88). The given solutions above have only positive frequencies (see (19.94)). They are exactly the positive frequency solutions of the Dirac equation which one usually uses for the quantum mechanical description of a particle with spin $\frac{1}{2}$.[482]

(19.115) In (18.174, 175) we have given two linear representations $M \mapsto \underline{M}_{\mathcal{E}}$ and $M \mapsto M_E$ of the group $\mathrm{Sl}(2,\mathbf{C})$ on the space $\mathcal{E}$ of spinors and on space-time E, respectively. We have constructed in (18.176) a quantizable extension G' of the Poincaré group, whose elements

(19.116) $$a \equiv \begin{pmatrix} M \\ C \end{pmatrix} \qquad \begin{matrix} M \in \mathrm{Sl}(2,\mathbf{C}) \\ C \in E \end{matrix}$$

act on space-time according to

(19.117) $$\underline{a}_E(X) \equiv \underline{M}_E X + C$$

[479] In a reference frame in which the velocity is *small*, $\gamma(I)$ is *approximately equal* to γ_4. Looking at expression (18.62) for γ_4, we see that the half-spinors ψ' and ψ'' (18.60) which form ψ are *almost equal*. This allows us to make a connection with the nonrelativistic theory above which uses only a "two-component spinor".

[480] $\gamma^\mu = \sum_\nu g^{\mu\nu}\gamma_\nu$ and thus $\gamma^j = -\gamma_j$ $(j = 1,2,3)$ and $\gamma^4 = \gamma_4$.

[481] One can easily deduce from this equation that the components of $\tilde{\psi}$ are solutions of the Klein-Gordon equation. This can also be deduced directly from (19.111).

[482] See Kastler [20].

and on the manifold W according to

$$\underline{a}_W \begin{pmatrix} X \\ \zeta \end{pmatrix} \equiv \begin{pmatrix} \underline{a}_E(X) \\ \underline{M}_{\mathcal{E}}\,\zeta \end{pmatrix}. \tag{19.118}$$

This group leaves the space E_1 invariant. A computation indeed shows that the action of a on E_1 consists of replacing the function $I \mapsto \psi$ by the function $I^* \mapsto \psi^*$ where

$$I^* \equiv \underline{M}_E I \qquad\qquad \psi^* \equiv \underline{M}_{\mathcal{E}} \psi\, \mathrm{e}^{-im\overline{I}^*\cdot C}, \tag{19.119}$$

and consequently $X \mapsto \tilde{\psi}$ by $X^* \mapsto \tilde{\psi}^*$ where

$$X^* \equiv \underline{a}_E(X) \qquad\qquad \tilde{\psi}^* \equiv M_{\mathcal{E}}\,\tilde{\psi}. \tag{19.120}$$



Thus it is the group G' (and not the Poincaré group G) which acts on the space of solutions of the Dirac equation. This classical paradox of quantum mechanics gets a natural explanation here.[483]

(19.121) As in the nonrelativistic case (19.103), we can construct a whole family of subspaces $E_2, E_3, \ldots$ of the Planck space, disjoint from E_1 and all invariant under the group G'.†

The massless particle with spin $\frac{1}{2}$

(19.122) Referring to the prequantization (18.96) we define

$$\eta \equiv \begin{pmatrix} X \\ \zeta \end{pmatrix}, \tag{19.123}$$

where ζ is still a Dirac spinor satisfying[484]

$$\overline{\zeta}\,\zeta \equiv 1 \qquad\qquad \overline{\zeta}\,\gamma_5\,\zeta \equiv 0. \tag{19.124}$$

The future pointing lightlike vector I defined by (18.84)

$$\overline{I}\cdot\delta X \equiv \overline{\zeta}\gamma(\delta X)\,\frac{1+i\,\gamma_5}{\sqrt{2}}\,\zeta \tag{19.125}$$

coincides with the energy-momentum vector P (we assume that the sign of the energy is positive).

---

[483] Note the difference with the case (19.80): it is not the cohomology which prevents G from acting, but the structure of the prequantum manifold Y.

† See the editors' note after (19.103).

[484] This ζ has been denoted as ψ in (18.83).

The prequantum 1-form ϖ is defined by

$$\varpi(\delta\eta) \equiv \varpi(\delta\xi) \equiv \chi \operatorname{Re}(-i\,\boldsymbol{\zeta}\delta\boldsymbol{\zeta}) - \overline{I}\cdot\delta X\,, \tag{19.126}$$

where $\chi = \pm 1$ denotes the *helicity*.

One can easily find the Planck space. Let $\boldsymbol{\psi}$ be a spinor which is a differentiable function of I on the "light cone" K defined by

$$\overline{I}\cdot I \equiv 0 \qquad I \text{ future pointing}\,, \tag{19.127}$$

and which satisfies

$$\gamma(I)\,\boldsymbol{\psi} \equiv 0 \qquad \text{and} \qquad \gamma_5\boldsymbol{\psi} \equiv \chi i\,\boldsymbol{\psi}\,. \tag{19.128}$$

Then the Ψ defined as

$$\Psi(\xi) \equiv \begin{cases} \overline{\boldsymbol{\psi}}\cdot\boldsymbol{\zeta}\mathrm{e}^{-i\,\overline{I}\cdot X} & \text{if } \chi = 1 \\ \overline{\boldsymbol{\zeta}}\cdot\boldsymbol{\psi}\mathrm{e}^{-i\,\overline{I}\cdot X} & \text{if } \chi = -1 \end{cases} \tag{19.129}$$

form the Planck space.

On the light cone K there exists a field of densities φ which is invariant under the Lorentz group and which can be defined by

$$\varphi(dI) \equiv \frac{\varphi(d\mathbf{p})}{E} \qquad (E \equiv \|\mathbf{p}\|;\ \text{compare with (19.92)})\,. \tag{19.130}$$

As before we suppose that the support of the function $I \mapsto \boldsymbol{\psi}$ is compact, and we define

$$\tilde{\boldsymbol{\psi}} \equiv \int_K \boldsymbol{\psi}\,\mathrm{e}^{i\,\overline{I}\cdot X}\,\varphi(dI)\,. \tag{19.131}$$

Using (19.128), it follows that $\tilde{\boldsymbol{\psi}}$ is a positive frequency solution of the system of equations

$$\gamma^\mu\,\partial_\mu\tilde{\boldsymbol{\psi}} \equiv 0\,^{485} \qquad \gamma_5\,\tilde{\boldsymbol{\psi}} \equiv \chi i\,\tilde{\boldsymbol{\psi}}\,.^{486} \tag{19.132}$$

(19.133) Since the experiments suggested by Lee and Yang on the "nonconservation of parity", it is indeed the positive frequency solutions that are used to describe *neutrinos* (where the two helicities are in fact observed).

---

[485]This equation implies the wave equation $\Box\tilde{\boldsymbol{\psi}} \equiv 0$.

[486]This equation expresses the fact that one of the (two) half-spinors in $\tilde{\boldsymbol{\psi}}$ (18.60) is zero. This is why one says that $\tilde{\boldsymbol{\psi}}$ is a two-component spinor.

(19.134) Note that the difficulties encountered in (19.103) and (19.121) do not occur: here the Planck condition is sufficient to find the right Planck space and the right wave equations.[487]

The massless particle with spin 1

(19.135) We will discuss the case where the *sign of the energy* is $+1$ and where the particle can be *polarized to the right as well as to the left* (definition (14.28)). This allows the *orthochronous* Poincaré group (which has two components) to act as a dynamical group.

Using (14.77) and (18.18), it follows from the prequantization (18.106) in the connected case that we can define

(19.136)
$$\eta \equiv \begin{pmatrix} S_1 \\ S_2 \\ P \\ X \\ \chi \end{pmatrix} \qquad \text{and} \qquad \varpi(\delta\eta) \equiv \chi\, \overline{S}_2\, \delta S_1 - \overline{P} \cdot \delta X\,,$$

where $\chi = \pm 1$ is the helicity of the particle, X a point of space-time E, P the energy-momentum vector on the light cone K (19.27) defined by

(19.137)
$$\overline{P} \cdot P \equiv 0 \qquad P \text{ future pointing,}$$

and where the vectors S_1 and S_2 satisfy

(19.138)
$$\overline{S}_1 \cdot S_1 \equiv \overline{S}_2 \cdot S_2 \equiv -1 \qquad \overline{P} \cdot S_1 \equiv \overline{P} \cdot S_2 \equiv \overline{S}_1 \cdot S_2 \equiv 0\,.$$

Contrary to the case of spin $\frac{1}{2}$, the Poincaré group is *quantizable* (in the sense of (18.151)). We obtain this quantization by defining, for

$$a \equiv \begin{bmatrix} L & C \\ 0 & 1 \end{bmatrix} \qquad \text{(notation (13.51), } L \text{ orthochronous),}$$

(19.139)
$$\underline{a}_W \begin{pmatrix} S_1 \\ S_2 \\ P \\ X \\ \chi \end{pmatrix} \equiv \begin{pmatrix} LS_1 \\ LS_2\, \chi_s(a) \\ LP \\ LX + C \\ \chi_s(a)\, \chi \end{pmatrix},$$

[487] This phenomenon can be linked to the following observation. The dimension of the isotropic foliations equals 3 in both cases. If the mass is zero, the dimension of the space of motions is 6 and the characteristic foliation is *self-orthogonal* (8.10). Thus the functions u_j defining the foliation (19.32) form what is called a "maximal set of observables". This is not the case if the mass is nonzero, because then the dimension of the space of motions is 8.

where $\chi_s(a)$ denotes the *spatial character* of a (14.73).

A computation shows[488] that the Planck space is the set of Ψ given as

$$\Psi(\xi) \equiv \overline{A}[S_1 + i\chi S_2]\, e^{-\overline{P}\cdot X}, \tag{19.140}$$

where A is an element of $E_{\mathbf{C}}$, *the complexification of the Euclidean space* E, that is,

$$A \equiv A' + iA'', \qquad \overline{A} = \overline{A'} - i\,\overline{A''}, \qquad A', A'' \in E, \tag{19.141}$$

(19.142) such that A is a differentiable function of P on the light cone K and such that $\overline{P}\cdot A \equiv 0$.

It should be noted that (19.140) defines A only up to a *"gauge" transformation*:

$$A \mapsto A + P f(P), \tag{19.143}$$

where f is a differentiable complex valued function on K. □

The quantized Poincaré group (19.139) acts on this Planck space by the transformation

$$[P \mapsto A] \to [L^{-1}P \mapsto L\cdot A\, e^{-i\overline{P}\cdot C}]. \tag{19.144}$$ □

Assuming that the support of $P \mapsto A$ is compact, we can define

$$\tilde{A} \equiv \int_K A\, e^{i\overline{P}\cdot X}\, \varphi(dP), \tag{19.145}$$

---

[488]First of all one can easily verify that necessarily $\Psi(\xi) \equiv \varphi\, e^{-i\overline{P}\cdot X}$, where φ is a complex valued function of S_1, S_2, P, and χ which is differentiable on the manifold defined by the equations (19.137, 138). Moreover, this Ψ is invariant under the transformations

$$S_1 \mapsto S_1 + P\alpha, \qquad S_2 \mapsto S_2 + P\beta, \qquad P \mapsto P,\ \chi \mapsto \chi \qquad \text{(compare (18.105))}$$

and such that

$$\diamondsuit \qquad [\delta S_1 = -S_2\chi, \quad \delta S_2 = S_1\chi, \quad \delta P = 0] \implies [\delta\varphi = i\varphi].$$

For a fixed P, φ depends only on $\widehat{S}_1$, $\widehat{S}_2$, and χ (for every vector Z orthogonal to P we denote by $\widehat{Z}$ its orbit under the group of translations parallel to P. Because P is lightlike, the $\widehat{Z}$ form a negative 2-dimensional Euclidean space $\mathcal{E}_2$, for which $[\widehat{S}_1\ \widehat{S}_2]$ is an orthonormal basis). Because of $\diamondsuit$, one can check that φ can be written as $\Gamma_\chi(\widehat{S}_1 + i\chi\widehat{S}_2)$, where Γ_1 and Γ_{-1} are elements of $\mathcal{E}_2^*$, or equivalently as $\varphi \equiv [\overline{B'} - i\overline{B}'][S_1 + i\chi S_2]$, where B' and B'' are vectors in $\mathcal{E}_2$ defined by this relation. Thus we can *choose* two vectors A' and A'' of E orthogonal to P such that $B' = \widehat{A}'$ and $B'' = \widehat{A}''$. Since such a choice can be made globally and differentiably (for instance by setting the temporal components of A' and A'' equal to 0), we can take $A \equiv A' + iA''$.

where φ is the density on the light cone K given in (19.130). It follows easily (using (19.117) and (19.122)) that $\tilde{A}$ is a vector in $E_{\mathbf{C}}$ and a differentiable function of X satisfying the *wave equation*

$$\Box \tilde{A} \equiv 0 \tag{19.146}$$

and the Lorentz equation

$$\partial_\mu \tilde{A}^\mu \equiv 0\,. \tag{19.147}$$

The gauge transformations (19.143) act on $\tilde{A}$ according to

$$\tilde{A}_\mu \mapsto \tilde{A}_\mu + \partial_\mu u\,, \tag{19.148}$$

where u is defined by

$$u \equiv -i \int_K f(P) \mathrm{e}^{i\overline{P}\cdot X}\, \varphi(dP)\,. \tag{19.149}$$

The function u is also a solution of the wave equation

$$\Box u \equiv 0\,. \tag{19.150}$$

(19.151) As in the preceding examples, we see that $\tilde{A}$ (and u) are *positive frequency* solutions of the wave equation. And in quantum mechanics a *photon* is indeed represented by a complex vector field $X \mapsto \tilde{A}$ with the properties that it is a positive frequency solution of the *wave equation*, that it satisfies the *Lorentz condition* (19.147), and that it is defined up to a "*gauge transformation of the second kind*" (19.148).[489] □

To determine how the quantized Poincaré group (19.139) acts on $X \mapsto \tilde{A}$, it suffices to substitute (19.144) into (19.145). Remembering that the measure defined by φ is invariant under the Lorentz group (19.130), we obtain the transformation

$$[X \mapsto \tilde{A}] \mapsto [\underline{a^{-1}}_E(X) \mapsto L \cdot \tilde{A}]\,, \tag{19.152}$$

that is, the quantized orthochronous Poincaré group acts on $X \mapsto \tilde{A}$ as on a *vector field*.[490] □

One can avoid gauge transformations, if one wants to, by introducing the complex 2-form F, a function of P, defined by

$$F_{\mu\nu} \equiv i\,[P_\mu A_\nu - P_\nu A_\mu]\,. \tag{19.153}$$

[489] See Kastler [20].

[490] Formula (2.38). It acts separately on the real and imaginary parts of the complex vector $\tilde{A}$.

F is obviously invariant under gauge transformations and it *characterizes* Ψ by the formula

$$-iP_\mu \Psi(\xi) \equiv \sum_\nu \overline{F}_{\mu\nu} \cdot [S_1^\nu + i\chi S_2^\nu] \, e^{i\overline{P}\cdot X} . \tag{19.154}$$

This F satisfies the system of equations

$$\begin{cases} \sum_\mu P_\mu F^{\mu\nu} \equiv 0 \,^{491} \\ P_\mu F_{\nu\rho} + P_\nu F_{\rho\mu} + P_\rho F_{\mu\nu} \equiv 0 . \end{cases} \tag{19.155}$$

If we define

$$\widetilde{F}_{\mu\nu} \equiv \int_K F_{\mu\nu} \, e^{i\overline{P}\cdot X} \, \varphi(dP) , \tag{19.156}$$

it follows that

$$\widetilde{F}_{\mu\nu} \equiv \partial_\mu \widetilde{A}_\nu - \partial_\nu \widetilde{A}_\mu \tag{19.157}$$

and that

$$\begin{cases} \sum_\mu \partial_\mu \widetilde{F}^{\mu\nu} \equiv 0 \\ \partial_\mu \widetilde{F}_{\nu\rho} + \partial_\nu \widetilde{F}_{\rho\mu} + \partial_\rho \widetilde{F}_{\mu\nu} \equiv 0 . \end{cases} \tag{19.158}$$

(19.159) Thus the $\widetilde{F}_{\mu\nu}$ are the components of a complex 2-form (the "electromagnetic field" form) which has the 1-form with components $\widetilde{A}_\mu$ as a *potential* in the sense of (9.14) ($\widetilde{A}$ is called the "electromagnetic potential"). The $\widetilde{F}_{\mu\nu}$ satisfy the *Maxwell equations* (19.158).[492] □

We can also describe the Planck space by means of real solutions $\widehat{A}$, $\widehat{F}$ of the wave equations (19.146, 147, 148, 150, 157, 158). It obviously suffices to define

$$\begin{cases} \widehat{A} \equiv \mathrm{Re}(\widetilde{A}) \\ \widehat{F} \equiv \mathrm{Re}(\widetilde{F}) . \end{cases} \tag{19.160}$$

These functions then satisfy the classical equations of electromagnetism in vacuum.

(19.161) At first sight it might seem that these real solutions do not completely define the state vector Ψ. Writing out in detail the components $\widehat{A}_\mu = \frac{1}{2}[\widetilde{A}_\mu + \overline{\widetilde{A}}_\mu]$ and $\widehat{F}_{\mu\nu} = \frac{1}{2}[\widetilde{F}_{\mu\nu} + \overline{\widetilde{F}}_{\mu\nu}]$, and taking (19.145) and (19.156) into account, we see however that the real fields contain positive and negative frequencies at the same time. Hence it is sufficient to *filter out*

---

[491] $F^{\mu\nu} \equiv \sum_{\mu',\nu'} g^{\mu\mu'} g^{\nu\nu'} F_{\mu'\nu'}$.

[492] Exercise: by introducing the vectors $\mathbf{E}$ and $\mathbf{B}$ (in $\mathbf{C}^3$) defined by $\widetilde{F}(\delta X)(\delta' X) \equiv \langle \delta\mathbf{r}\, \delta' t - \delta'\mathbf{r}\, \delta t, \mathbf{E}\rangle + \langle \delta\mathbf{r} \times \delta'\mathbf{r}, \mathbf{B}\rangle$, rewrite the Maxwell equations (19.158) in a 3-dimensional form and give them their usual appearance. Deduce from these equations the wave equation $\Box \widetilde{F} \equiv 0$.

the negative frequencies (in the usual sense of applied harmonic analysis) to reconstitute the complex fields and thus Ψ.

Assemblées of particles†

(19.162) Let U be the space of motions of a particle of a given type.[493] We will describe an assemblée of such particles (that is, a systems consisting of an indeterminate number of indistinguishable particles) by its space of motions, which we will call the *Fock manifold*[494] and denote by U_Φ.

(19.163) As a set, U_Φ is simply the set of *finite subsets* of U. Thus U_Φ decomposes into a disjoint union of manifolds U_n (each describing the subset consisting of n motions, that is, a system of n particles). Thus U_Φ is a *composite* manifold (16.8, 9, 10).

U_0 is the manifold of dimension 0 containing the single element $\emptyset$ (the empty subset of U). U_1 is canonically identified with U. The structure of a symplectic manifold for U_n ($n \geq 2$) has already been given in (15.45). Let us recall the construction.

If S_n denotes the group of permutations of the integers $1, 2, \ldots, n$ ("the symmetric group of degree n"), we can define an action of S_n on the direct product U^n by

(19.164)
$$\underline{\Pi^{-1}}_{U^n} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \equiv \begin{pmatrix} x_{\Pi(1)} \\ \vdots \\ x_{\Pi(n)} \end{pmatrix} \qquad \Pi \in S_n\,, \quad x_1, \ldots, x_n \in U\,.$$

Let us denote by

(19.165)
$$x \equiv \Phi_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

the orbit of $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ under the group S_n (Fig. 19.II).

---

†Editors' note: In the French original the same word assemblée was used with an explanatory note saying that it is "franglais". Since no adequate English word exists, we decided to leave it untranslated.

[493]It does not necessarily concern *elementary* particles, but a priori arbitrary systems (for example, α particles), whose mutual interactions are assumed to be negligible.

[494]We have already described it, in the case of photons, when discussing black body radiation (17.152).

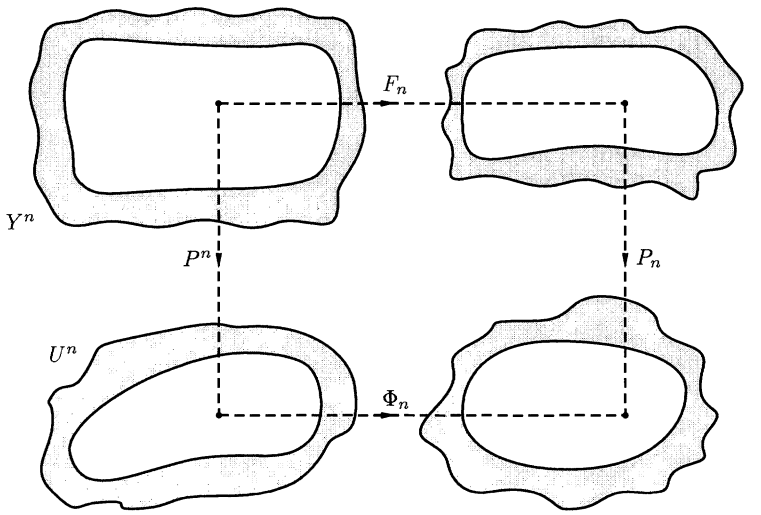


Fig 19.II

(19.166) If we denote by $\tilde{U}^n$ the set of $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ such that the x_j are *all different*, then $\tilde{U}^n$ is *open* in U^n and is invariant under the action of S_n. Since S_n is a *discrete group of diffeomorphisms* (1.45), we can give the quotient $U_n = \tilde{U}^n/S_n$, which is the image of $\tilde{U}^n$ under Φ_n, the structure of a manifold such that the map Φ_n is differentiable. It is this manifold U_n which constitutes the space of motions of a system of n identical particles.[495] The inverse image of the Lagrange form on U_n under Φ_n is the Lagrange form on $\tilde{U}^n$ (which is the one induced by the Lagrange form on U^n, which in turn is defined in (9.6)).

(19.167) Let us note that $(\tilde{U}^n, S_n, \Phi_n)$ is a *covering space* of U_n. In the case of elementary particles which we have studied before, $\tilde{U}^n$ is simply connected, and thus this covering space is *universal* (1.55). □

Let (Y, P) be a prequantization of U. To prequantize the Fock manifold U_Φ, it suffices to prequantize separately each U_n (18.18).

The prequantization (Y_0, P_0) of U_0 is immediate:

(19.168) $$Y_0 = T\,, \qquad P_0(z) \equiv \emptyset \qquad (z \in T)\,, \qquad \varpi(\delta z) \equiv \frac{\delta z}{i\,z}\,.$$

The prequantization (Y_1, P_1) of U_1 is supposed to be known already

[495] Taking the quotient by the group maps — from a set theoretic point of view — an n-tuple of x_j to the *set* of x_js. This is the mathematical formulation of the indistinguishability of the particles.

($Y_1 = Y$, $P_1 = P$). As for U_n ($n \geq 2$), it admits *two inequivalent prequantizations* associated with the two characters χ_+ and χ_- of the group S_n (18.138). If χ denotes one of these characters, we can describe the prequantization (Y_n, P_n) as follows.

Let $\widetilde{Y}^n$ be the set of $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in Y^n$ such that the $P(\xi_j)$ are all distinct; it can be seen as the inverse image of $\widetilde{U}^n$ under the map P^n given by

$$P^n \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \equiv \begin{pmatrix} P(\xi_1) \\ \vdots \\ P(\xi_n) \end{pmatrix} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{19.169}$$

On Y^n, or on $\widetilde{Y}^n$, we define an equivalence relation $\sim$ by

$$\left[\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \sim \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_n \end{pmatrix} \right] \iff \begin{bmatrix} \exists \Pi \in S_n, \ \exists z_1, \ldots, z_n \in T \\ z_1 \cdots z_n = \chi(\Pi) \\ \xi'_j = \underset{\sim}{z_j}_Y(\xi_{\Pi(j)}) \qquad \forall j \end{bmatrix}. \tag{19.170}$$

If we denote by

$$\xi \equiv F_n \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \tag{19.171}$$

the equivalence class of $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ according to the equivalence relation $\sim$, then the manifold Y_n is the image of $\widetilde{Y}^n$ under F_n (Fig. 19.II), and the map P_n is defined by

$$x \equiv P_n(\xi)\,, \tag{19.172}$$

that is, by the commutativity of the diagram in figure 19.II:

$$P_n \circ F_n = \Phi_n \circ P^n\,. \tag{19.173}$$

The prequantum 1-form ϖ on Y_n is given by

$$\varpi(\delta\xi) \equiv \sum_j \varpi_j(\delta\xi_j)\,, \tag{19.174}$$

where ϖ_j is the prequantum 1-form on Y evaluated at the point ξ_j.

(19.175) To finish the prequantization of U_Φ we have to choose the character χ for each n. There are arguments suggesting that it should always be

the same, whatever n is.[496] Thus we will treat the two cases $\chi \equiv \chi_+$ ($\chi \equiv \chi_-$, respectively) by saying that the particles are *bosons* (*fermions*, respectively). In fact, these are the only ones encountered in nature.[497] Finally, the prequantization (Y_Φ, P_Φ) of U_Φ is defined by

(19.176)
$$\begin{cases} Y_\Phi = \text{the (disjoint) union of the } Y_n \ (n = 0, 1, 2, \ldots) \\ \qquad\quad \text{defined above,} \\ P_\Phi = \text{the common extension of the } P_n. \end{cases}$$

Creation and annihilation operators

(19.177) We are going to study the pre-Hilbert space $\mathcal{H}(Y_\Phi)$ associated with the prequantum manifold Y_Φ (definitions (18.179, 183)), and we are going to construct on this space some operators having remarkable algebraic properties.

(19.178) Let Ψ be an element of $\mathcal{H}(Y_n)$ $(n = 0, 1, 2 \ldots)$. We will *identify* Ψ *with its extension by* 0 *to all* $Y_{n'}$ $(\forall\, n' \neq n)$. This extension is an element of $\mathcal{H}(Y_\Phi)$. □

(19.179) THEOREM: Let Ψ be an element of $\mathcal{H}(Y_\Phi)$. Then Ψ can be written in a unique way as a finite sum

$$\Psi = \Psi_0 + \Psi_1 + \cdots + \Psi_m \,,$$

where each Ψ_n belongs to $\mathcal{H}(Y_n)$.[498] We obtain this decomposition by letting Ψ_n be the restriction of Ψ to Y_n.

PROOF: Since the support of Ψ is compact (18.179), it admits a finite subcover coming from the open cover $(Y_0, Y_1, \ldots)$. The rest of the argument is trivial. □

(19.180) THEOREM: $\mathcal{H}(Y_0)$ is a 1-dimensional vector space consisting of (complex) multiples of Ω defined by

$$\Omega(z) \equiv z \,.$$

[496] If not, the behavior of a system would depend on the number of particles of the same type likely to exist in the rest of the universe.

[497] In the case of nonelementary particles, it could happen that $\widetilde{U}^n$ is not simply connected. Then the homotopy group of U_n no longer equals S_n and could have more than two characters. This would lead to prequantizations of a new type.

[498] $\mathcal{H}(Y_\Phi)$ is called the *direct sum* of the $\mathcal{H}(Y_n)$.

Ω, seen as an element of $\mathcal{H}(Y_\Phi)$ (by (19.178)), is a state vector.[499] It is called the *vacuum*.

(19.181) THEOREM: For $\Psi \in \mathcal{H}(Y_n)$ $[n \geq 2]$ the function $\hat{\Psi} = \Psi \circ F_n$ (Fig. 19.II) satisfies the relations

$$\diamondsuit \quad \begin{cases} \hat{\Psi} \text{ is a differentiable function on } \tilde{Y}^n \text{ with compact support} \\ \hat{\Psi}\begin{pmatrix} \underline{z_1}(\xi_1) \\ \vdots \\ \underline{z_n}(\xi_n) \end{pmatrix} \equiv z_1 \times \cdots \times z_n \cdot \hat{\Psi}\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & z_j \in T\,,\ \xi_j \in Y \\ \hat{\Psi}\begin{pmatrix} \xi_{\Pi(1)} \\ \vdots \\ \xi_{\Pi(n)} \end{pmatrix} \equiv \chi(\Pi)\,\hat{\Psi}\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}^{500} & \Pi \in S_n\,,\ \xi_j \in Y\,. \end{cases}$$

Conversely, if $\hat{\Psi}$ satisfies $\diamondsuit$, then there exists a unique element Ψ of $\mathcal{H}(Y_n)$ such that

$$\Psi \circ F_n = \hat{\Psi}\,.$$

PROOF: The verification is left to the good care of the reader. One should use (19.170), which allows one in particular to show that F_n is *proper*, and (18.51). □

(19.182) THEOREM/DEFINITION:

I) Let ξ be a point of Y. There exists a linear operator $a^-(\xi)$ on $\mathcal{H}(Y_\Phi)$, which maps each $\mathcal{H}(Y_n)$ to $\mathcal{H}(Y_{n-1})$ and $\mathcal{H}(Y_0)$ to zero, defined by[501]

$$\heartsuit \quad \begin{cases} (a) \quad a^-(\xi)(\Psi)(z) \equiv z\,\Psi(\xi) & z \in Y_0\,(=T) \\ (b) \quad a^{-1}(\xi)(\Psi)(F_n\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}) \equiv \Psi(F_{n+1}\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \\ \xi \end{pmatrix}) & \Psi \in \mathcal{H}(Y_\Phi) \\ & n = 1, 2, 3 \ldots \end{cases}$$

$a^-(\xi)$ is called the *annihilation operator* associated to ξ.

II) Let θ be *an element of* $\mathcal{H}(Y)$. There exists a linear operator $a^+(\theta)$

[499] With the convention (already used in (17.153)) that the Liouville measure of $U_0 = \{\emptyset\}$ has mass 1.

[500] Thus $\hat{\Psi}$ is *symmetric* in the case of bosons and *antisymmetric* in the case of fermions.

[501] We define $F_1(\xi_1) \equiv \xi_1$.

on $\mathcal{H}(Y_\Phi)$, which maps each $\mathcal{H}(Y_n)$ to $\mathcal{H}(Y_{n+1})$, defined by

♣
$$\begin{cases} (a) \quad a^+(\theta)(\Psi)(z) \equiv 0 & \Psi \in \mathcal{H}(Y_\Phi) \\ (b) \quad a^+(\theta)(\Psi)(\xi_1) \equiv \theta(\xi_1)\Psi(1) & z \in Y_0 \ (= T) \\ & \xi_1 \in Y_1 \\ (c) \quad a^+(\theta)(\Psi)(F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}) \equiv & n = 2, 3, \dots \\ \equiv \frac{1}{(n-1)!}\sum_{\Pi\in S_n}\chi(\Pi)\,\Psi(F_{n-1}\begin{pmatrix}\xi_{\Pi(1)}\\ \vdots \\ \xi_{\Pi(n-1)}\end{pmatrix})\,\theta(\xi_{\Pi(n)}) & \end{cases}$$

$a^+(\theta)$ is called the *creation operator* associated to θ; a^+ is bilinear.

III) The operators a^+ and a^- satisfy the following identities.

♠
$$\begin{cases} (a) \quad a^-(\xi)(\Omega) \equiv 0 & \xi, \xi' \in Y \quad \theta, \theta' \in \mathcal{H}(Y) \\ (b) \quad a^-(\xi)(\theta) \equiv \theta(\xi)\,\Omega & \Omega = \text{vacuum (19.180)} \\ (c) \quad a^+(\theta)(\Omega) \equiv \theta & \varepsilon = \begin{cases} +1 \quad \text{(bosons)} \\ -1 \ \text{(fermions)} \end{cases} \\ (d) \quad a^-(\xi)\circ a^-(\xi') - \varepsilon\, a^-(\xi')\circ a^-(\xi) \equiv 0 & \\ (e) \quad a^+(\theta)\circ a^+(\theta') - \varepsilon\, a^+(\theta')\circ a^+(\theta) \equiv 0 & \\ (f) \quad a^-(\xi)\circ a^+(\theta) - \varepsilon\, a^+(\theta)\circ a^-(\xi) \equiv \underline{\theta(\xi)}\,.^{502} & \end{cases}$$

PROOF: To verify that $a^-(\xi)(\Psi)$ and $a^+(\theta)(\Psi)$ defined by ♡ and ♣ are indeed elements of $\mathcal{H}(Y_\Phi)$, it suffices to apply criterion (19.181). The remaining assertions of the theorem can be established by long but elementary verifications. It is useful to establish first the formula

(19.183)
$$\begin{aligned} a^+(\theta)(\Psi)(F_{n+1}\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n \\ \xi\end{pmatrix}) \equiv \Psi\circ F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}\theta(\xi) + \\ + \varepsilon\left[\Psi\circ F_n\begin{pmatrix}\xi\\ \xi_2\\ \vdots \\ \xi_n\end{pmatrix}\theta(\xi_1) + \cdots + \Psi\circ F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_{n-1}\\ \xi\end{pmatrix}\theta(\xi_n)\right], \end{aligned}$$

[502]For a scalar s we denote by $\underline{s}$ the operator of multiplication by s (here on the complex vector space $\mathcal{H}(Y_\Phi)$).

which simplifies the verification of (♠f). Similarly, the formula

$$\Psi(F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix}) \equiv [a^-(\xi_1)\circ\cdots\circ a^-(\xi_n)](\Psi)(1) \tag{19.184}$$

immediately gives (♠d). To verify (♠e), one can show, using (♠f), that the operator $A = a^+(\theta)\circ a^+(\theta') - \varepsilon\, a^+(\theta')\circ a^+(\theta)$ commutes or anticommutes with all $a^-(\xi)$. Using (19.184) to compute $A(\Psi)(F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix})$, one can see that this number is zero. □

Note that we can construct an element Ψ of $\mathcal{H}(Y_n)$ by defining

$$\Psi = \left[a^+(\Psi_n)\circ a^+(\Psi_{n-1})\circ\cdots\circ a^+(\Psi_1)\right](\Omega) \qquad [\Psi_j \in \mathcal{H}(Y)]\,. \tag{19.185}$$

The values of the above Ψ on Y_n can be calculated using formula (19.183). We obtain

$$\Psi\circ F_n\begin{pmatrix}\xi_1\\ \vdots \\ \xi_n\end{pmatrix} = \sum_{\Pi\in S_n}\chi(\Pi)\,\Psi_1(\xi_{\Pi(1)})\cdots\Psi_n(\xi_{\Pi(n)})\,. \tag{19.186}$$

(19.187) If Ψ (19.185) is the state vector of an assemblée of particles, one says that *the assemblée consists of n particles, the first being in the state* Ψ_1, *the second in the state* Ψ_2, etc.

In the case of *fermions*, one can see that the quantity (19.186) equals the determinant

$$\det\left(\begin{bmatrix}\Psi_1(\xi_1) & \cdots & \Psi_1(\xi_n)\\ \cdots & \cdots & \cdots \\ \Psi_n(\xi_1) & \cdots & \Psi_n(\xi_n)\end{bmatrix}\right). \tag{19.188}$$

This shows that Ψ is identically zero (and thus is incapable of being a state vector) if the $\Psi_1,\ldots,\Psi_n$ are not *all different*. This is *Pauli exclusion principle*.

(19.189) On the contrary, *bosons* can *all* be *in the same state*.[503]

Let us now study the pre-Hilbert space structure of $\mathcal{H}(Y_\Phi)$. One can check trivially that

$$[\Psi\in\mathcal{H}(Y_n)\,,\quad \Psi'\in\mathcal{H}(Y_{n'})\,,\quad n\neq n'] \implies [\overline{\Psi}\cdot\Psi' = 0]\,, \tag{19.190}$$

[503]It seems that this is the case for *superfluid helium* and also for *coherent light* (which explains why we can describe it, as if it were composed of a single photon, as a solution of the Maxwell equations (19.160)).

that is, the spaces $\mathcal{H}(Y_n)$ are *pairwise orthogonal*. To compute the inner product of $\Psi, \Psi' \in \mathcal{H}(Y_\Phi)$, it thus suffices to use the direct sum decomposition (19.179)

$$\Psi = \sum_n \Psi_n \qquad \Psi' = \sum_n \Psi'_n \tag{19.191}$$

to obtain

$$\overline{\Psi} \cdot \Psi' = \sum_n \overline{\Psi}_n \cdot \Psi'_n \,, \tag{19.192}$$

where the above sum is finite.

By definition (18.181) we have

$$\overline{\Psi}_n \cdot \Psi'_n \equiv \int_{U_n} \overline{\Psi_n(\xi)} \cdot \Psi'_n(\xi)\, \varphi(dx) \,. \tag{19.193}$$

To compute the above integral, let us use the covering space $(\tilde{U}^n, S_n, P_n)$ of U_n (19.167). The Liouville density on $\tilde{U}^n$ equals the tensor product $\varphi \otimes \cdots \otimes \varphi$ (theorem (16.98)). Its image under the map P_n equals $n!$ times the Liouville density on U_n (because S_n has $n!$ elements). Thus (19.193) is transformed into [504]

$$\overline{\Psi}_n \cdot \Psi'_n = \frac{1}{n!} \int_{\tilde{U}^n} \overline{\Psi_n \circ F_n \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}} \cdot \Psi'_n \circ F_n \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \varphi(dx_1) \ldots \varphi(dx_n) \,. \tag{19.194}$$

Using this formula, the fact that $\tilde{U}^n$ is *dense* in U^n, and by performing some transformations on the integral, the reader will be able to establish the formula

$$\langle \Psi,\, a^+(\theta)(\Psi') \rangle \equiv \langle a^-(\theta)(\Psi), \Psi' \rangle \qquad \theta \in \mathcal{H}(Y)\,, \Psi, \Psi' \in \mathcal{H}(Y_\Phi)\,, \tag{19.195}$$

where we have defined

$$a^-(\theta)(\Psi)(\xi_\Phi) = \int_U \overline{\theta(\xi)} \cdot a^-(\xi)(\Psi)(\xi_\Phi)\, \varphi(dx) \qquad \forall\, \xi_\Phi \in Y_\Phi \,. \tag{19.196}$$

It follows that the linear operators $a^+(\theta)$ and $a^-(\theta)$ are each other's *transpose* (6.35). Using theorem (19.182), the reader can immediately establish the following list of formulæ.

(19.197) THEOREM: $a^-(\theta)$ is a linear operator on $\mathcal{H}(Y_\Phi)$ which maps each $\mathcal{H}(Y_n)$ to $\mathcal{H}(Y_{n-1})$. $a^-(\theta)$ is called the *annihilator* ("of a particle in the state θ"). a^- is an *anti-linear* operator (18.184) having the following properties.

[504] One could, if wanted, suppress the term $1/n!$ in front of the integral sign by working with $\Psi_n/\sqrt{n!}$ instead of with Ψ_n. This introduces factors $\sqrt{n+1}$ and $1/\sqrt{n+1}$ in the definition of the creation and annihilation operators, respectively. These complications, although useless, are traditional.

(a) $a^-(\theta) = a^+(\theta)^{*\dagger}$
(b) $a^+(\theta) = a^-(\theta)^{*\dagger}$
(c) $a^-(\theta)(\Omega) \equiv 0$
(d) $a^-(\theta)(\theta') \equiv \langle \theta, \theta' \rangle \Omega$
(e) $a^-(\theta) \circ a^-(\theta') - \varepsilon\, a^-(\theta') \circ a^-(\theta) \equiv 0$
(f) $a^+(\theta) \circ a^+(\theta') - \varepsilon\, a^+(\theta') \circ a^+(\theta) \equiv 0$
(g) $a^-(\theta) \circ a^+(\theta') - \varepsilon\, a^+(\theta') \circ a^-(\theta) \equiv \underline{\langle \theta, \theta' \rangle}$ □

Ω = vacuum
$\theta, \theta' \in \mathcal{H}(Y)$
$\varepsilon = \begin{cases} +1 & \text{(bosons)} \\ -1 & \text{(fermions)} \end{cases}$

In the case of fermions, one should note that the formulæ (e, f, g) show that

(19.198) $$\left[a^+(\theta) + a^-(\theta)\right]^2 \equiv \underline{\|\theta\|^2},$$

and that formula (g) shows that the creation operators $a^+(\theta)$ and annihilation operators $a^-(\theta)$ [505] are bounded; more precisely

(19.199) $$\|a^+(\theta)\| \leq \|\theta\|, \qquad \|a^-(\theta)\| \leq \|\theta\|.$$

(19.200) It is this formalism of creation and annihilation operators that is sometimes called "second quantization", and which lies at the foundation of quantum field theory.

Quantum states

Let U be the space of motions of a dynamical system and let Y be its prequantum manifold.

(19.201) We have defined the pre-Hilbert space $\mathcal{H}(Y)$ (18.179) and in (18.196) we have given a *unitary representation* $a \mapsto \underline{a}$ of the group Quant(Y) (of quantomorphisms of Y) on the Hilbert space $H(Y)$, the completion of $\mathcal{H}(Y)$.

A *state vector* of the system, in the sense of (19.12), is an element Ψ of $\mathcal{H}(Y)$, and thus of $H(Y)$, satisfying

(19.202) $$\overline{\Psi} \cdot \Psi = 1.$$

†Editors' note: The * indicates the involution defined in (18.185◇). Strictly speaking the use of * is not justified because we do not know whether $a^+(\theta)$ and/or $a^-(\theta)$ is bounded; this equation should thus be read in the sense of (19.195). And indeed, for bosons one can show that $a^+(\theta)$ is unbounded (and hence $a^-(\theta)$ too). On the other hand, for fermions these operators are bounded (19.199), and thus in that case the use of * is justified.

[505] But not the annihilation operators $a^-(\xi)$ associated to points ξ of U (definition (19.182♡)).

In (19.14) we have seen that we can associate to a state vector a *statistical state* of the system, in which the *average value* of an arbitrary dynamical variable u is given by

$$\int_U u\,|\Psi(\xi)|^2\,\varphi(dx)\,. \tag{19.203}$$

If the state vector Ψ satisfies the Planck condition defined by means of u, we know (19.32) that

$$[\hat{u}(\Psi)](\xi) \equiv u \times \Psi(\xi)\,, \tag{19.204}$$

where $\hat{u}$ denotes the *observable* associated to u. The operator $\hat{u}$, which maps $\mathcal{H}(Y)$ to $\mathcal{H}(Y)$, is hermitian. From (19.203, 204) and (18.181) it follows immediately that the average value of u can be written as

$$\langle \Psi,\, \hat{u}(\Psi) \rangle \tag{19.205}$$

or (18.183) as

$$\text{average}\,(u) = \overline{\Psi} \cdot \hat{u} \cdot \Psi\,. \tag{19.206}$$

(19.207) 

In quantum mechanics, it is formula (19.206) which is adopted as the *definition of average value*. Thus Planck's condition appears as a sufficient condition in order that the *statistical* and *quantum mechanical* averages of u coincide. One thus could give up this Planck condition together with the statistical interpretation.

We are going to transform formula (19.206) in case u generates an infinitesimal quantomorphism according to (18.145). Let f be this infinitesimal quantomorphism:

$$f(\xi) = \underset{u}{\delta}\,\xi\,.$$

By definition (18.192) of $\hat{u}$ we have

$$\hat{u}(\Psi)(\xi) \equiv -i\,\underset{u}{\delta}\,[\Psi(\xi)] \equiv \lim_{s \mapsto 0} \frac{\Psi(\exp(sf)(\xi)) - \Psi(\xi)}{i\,s}\,.$$

Using the fact that $\exp(sf)$ is a quantomorphism (18.145), that the support of Ψ is compact, and that u and Ψ are differentiable, we deduce that [506]

$$\int_U \overline{\Psi(\xi)} \cdot \hat{u}(\Psi)(\xi)\,\varphi(dx) = \lim_{s \mapsto 0} \int_U \overline{\Psi(\xi)} \cdot \frac{\Psi(\exp(sf)(\xi)) - \Psi(\xi)}{i\,s}\,\varphi(dx)\,.$$

[506] The computation of the integral depending on s can be reduced to the computation of a finite sum of integrals over $\mathbf{R}^{2n}$ (where $2n$ is the dimension of U) whose integrands converge uniformly (use theorem (16.49)).

We thus obtain the following theorem.

(19.208) THEOREM: Let u be a dynamical variable such that the vector field

$$f(\xi) \equiv \underset{u}{\delta} \xi$$

is an infinitesimal quantomorphism. If Ψ is a state vector, the quantum mechanical average (19.207) of u in the state Ψ is given by

$$\diamondsuit \qquad \text{average}\,(u) = \frac{1}{i}\frac{d}{ds}\left[\overline{\Psi}\cdot \underline{\exp(sf)}\cdot \Psi\right]_{s=0},$$

where $\underline{\exp(sf)}$ is the operator associated to the quantomorphism $\exp(sf)$ by the unitary representation (18.196). □

(19.209) Let us recall that the condition imposed on u (generating an infinitesimal quantomorphism) can take the form (18.145♡), in which the prequantization of U does not play a role, and which in particular is satisfied

when u is a function with compact support;
when u is a moment of a dynamical group (even nonquantizable).

In statistical mechanics a state is completely determined once we know how to compute the average value of functions with compact support.[507] Analogously, we will define the *quantum state* corresponding to the vector Ψ to be the *function*

(19.210) $$a \mapsto \overline{\Psi}\cdot \underline{a^{-1}}\cdot \Psi\,,$$

which is defined on the group of quantomorphisms of Y. According to formula (19.208♢), this function allows us to compute the quantum mechanical average value of dynamical variables with compact support.[508] We note that the *state* does not change if we dephase the *vector* Ψ (that is, if we multiply it by an element of T).

(19.211) DEFINITION: Let F be a complex valued function on a group G. F is called *of positive type* if for an arbitrary finite collection $a_1, a_2, \ldots, a_p$ of elements of G, the $p \times p$ matrix M with entries

$$^{j}M_k = F(a_j^{-1} \times a_k)$$

is *positive semidefinite*.[509]

---

[507] Because a state is a *measure* on U.

[508] As well as those coming from a (dynamical) moment. One should note that the variables which do not correspond to infinitesimal quantomorphisms lead to difficulties in the usual interpretation of quantum mechanics.

[509] That is, $\overline{Z}\cdot M\cdot Z \geq 0\ \forall\, Z \in \mathbf{C}^p$. One can show that this is equivalent to $\overline{M} = M$ and spectrum$(M) \geq 0$. Thus necessarily $\det(M) \geq 0$ (see Souriau [31]).

(19.211) THEOREM: If $a \mapsto \underline{a}$ is a unitary representation of G on a Hilbert space E, then for $\Psi \in E$ the function

$$a \mapsto \overline{\Psi} \cdot \underline{a^{-1}} \cdot \Psi$$

is *of positive type.*

PROOF: Let $Z = \begin{pmatrix} {}^1z \\ \vdots \\ {}^pz \end{pmatrix}$ be a vector in $\mathbf{C}^p$, let $a_1, \ldots, a_p$ be elements of G, and let M be the $p \times p$ matrix with entries ${}^jM_k = \overline{\Psi} \cdot \underline{a_k^{-1} \times a_j} \cdot \Psi$. A direct computation gives

$$\overline{Z} \cdot M \cdot Z = \| \sum_j \overline{{}^jz}\, \underline{a_j}(\Psi) \|^2 \geq 0 . \qquad \square$$

Let us now indicate some obvious properties of functions F of positive type.

(19.212)
$$\begin{cases} F(a^{-1}) = \overline{F(a)} & \forall\, a \in G \\ |F(a)| \leq F(e) & e = \text{the identity element of } G \end{cases}$$

(take $p = 2$, and choose $a_1 = e$ and $a_2 = a$). It follows that functions of positive type are *bounded.*

(19.213) The set of all functions of positive type on G form a *convex cone*, that it, *every linear combination with positive coefficients of functions of positive type is again a function of positive type.*

(19.214) THEOREM: Let Y be the prequantum manifold of a dynamical system, let Ψ be a state vector of the system, and let F be the corresponding "state" function (19.210), namely

$$F(a) = \overline{\Psi} \cdot \underline{a^{-1}} \cdot \Psi \qquad \forall\, a \in \mathrm{Quant}(Y) .$$

Then

$$\diamondsuit \begin{cases} \text{(I)} & F \text{ is a function of positive type on the group } \mathrm{Quant}(Y) \\ \text{(II)} & F(\underline{z_Y} \circ a) \equiv z\, F(a) \qquad \forall\, z \in T\,,\ \forall\, a \in \mathrm{Quant}(Y) \\ \text{(III)} & F(\underline{1_Y}) = 1 . \end{cases}$$

PROOF: We know that F is of positive type (theorem (19.211)). By writing $F(a)$ in the form $\overline{\Psi} \cdot [\Psi \circ a]$ (definition (18.196)), we see that (II) and (III) transcribe the axioms $\Psi(\underline{z_Y}(\xi)) \equiv z\, \Psi(\xi)$ and $\overline{\Psi} \cdot \Psi = 1$, respectively, of state vectors (19.12). $\square$



We now are going to generalize the notions of *state* and *average value* by taking the results of the previous theorem as axioms.

(19.215) DEFINITION: By a *state* of a dynamical system we will mean any function F that satisfies the axioms (19.214◇).

We define the *average value* of a dynamical variable u (to which is associated the infinitesimal quantomorphism $f : \xi \mapsto \underset{u}{\delta}\xi$) in a state F by the formula

♡ $$\text{average}\,(u) = -i\,\frac{d}{ds}\,[F(\exp(sf))]_{s=0}\,,$$

provided the indicated derivative exists. □

(19.216) One can easily verify that the average value of a dynamical variable is *real* (apply (19.212)), and that the average value of a constant function equals this constant.

(19.217) THEOREM: The set of states of a dynamical system is *convex*.

PROOF: Trivial (see (19.213)). □

(19.218) Thus, if $F_1, \ldots, F_n$ are states and if $p_1, \ldots, p_n$ are positive numbers whose sum is 1, then $F = \sum_j p_j F_j$ is also a state. The state F is called a *mixture* of the states F_j.

In the special case where the F_j are derived from state vectors Ψ_j, it follows that the average value of u in the mixed state F can be written in the form [510]

(19.219) $$\text{average}\,(u) = \text{Tr}(\widehat{u} \cdot D)\,,$$

with

(19.220) $$D = \sum_j p_j\, \Psi_j \cdot \overline{\Psi}_j\,.$$

D is a positive [511] bounded operator of finite rank and of trace 1. D is called a *density operator*.

(19.221) It has been noted a long time ago [512] that in quantum mechanics it



is necessary to describe certain physical states by a density operator D together with the rule (19.219). This is especially the case in *quantum chemistry*, where the method of *molecular orbitals* is a means of calculating the density operator representing the state of a molecule at absolute

[510] The trace operator Tr is linear on the space of operators of finite rank from a vector space E to itself. It has the property $\text{Tr}(V \cdot \Gamma) \equiv \Gamma \cdot V$, $V \in E$, $\Gamma \in E^*$, which characterizes it completely.

[511] An element of a C^*-algebra is called positive if it is the square of a self-adjoint operator. Thus it is itself self-adjoint.

[512] By L. Landau and F. Bloch in 1927.

zero (in the *Hartree-Fock approximation*). We see that the generalization of the idea of a state, initially defined by means of a single vector, is necessary to take reality into account.

Let us now give some other examples of states that are compatible with definition (19.215).

(19.222) THEOREM: Let $j \mapsto F_j$ be a map from a Hausdorff manifold I to the space of states of a dynamical system such that $j \mapsto F_j(a)$ is continuous for every $a \in \mathrm{Quant}(Y)$. Let μ be a *probability measure* (16.178) on I. Then there exists a state F defined by

$$F(a) \equiv \mu(j \mapsto F_j(a)) .$$

PROOF: From (19.212) and (19.214.III) we obtain $|F_j(a)| \leq 1$ $(\forall j \; \forall a)$. Consequently, the average value $F(a)$ exists for every a. Trivially F satisfies (19.214.II, III). As for condition (I), it follows (using theorem (16.194)) from the fact that the positive semidefinite matrices of a given order form a *closed convex* set. □

(19.223) In particular, if we consider an *arbitrary* collection of states F_j and numbers $p_j \geq 0$ such that $\sum_j p_j = 1$, then there exists a state F defined by

$$F(a) = \sum_j p_j F_j(a)$$

(it suffices to consider the set of indices as a manifold of dimension 0, see (16.182)). It follows that theorem (19.222) generalizes the notion (19.218) of a mixture of states to an *infinite* mixture.

(19.224) THEOREM: If F_j is a sequence of states such that the limit

$$F(a) = \lim_j F_j(a)$$

exists for every $a \in \mathrm{Quant}(Y)$, then F is a state.

PROOF: Trivial. □

(19.225) THEOREM: If Ψ is a vector in the Hilbert space $H(Y)$ (19.201) such that $\overline{\Psi} \cdot \Psi = 1$, then the function $a \mapsto \overline{\Psi} \cdot a^{-1} \cdot \Psi$ is a state.

PROOF: Trivial. □

(19.226) Obviously it is possible that the definition of states (19.215) proposed above is too general and that it might be a good idea to restrict it by means of supplementary conditions, such as continuity conditions. □

Let G be a Lie group acting by quantomorphisms on the prequantum manifold Y of a system. For every state F of the system the induced function F_G on G given by

$$F_G(a) = F(\underline{a}_Y) \qquad \forall\, a \in G \tag{19.227}$$

is evidently, like the function F, of *positive type*.



If we assume that F_G is *continuous*, one can prove [513] that there exists a *continuous unitary representation* [514] $a \mapsto \underline{a}_E$ of G on a Hilbert space E and an *element* Ψ *of* E such that

$$F_G(a) = \overline{\Psi} \cdot \underline{a^{-1}}_E \cdot \Psi \qquad \forall\, a \in G\,. \tag{19.228}$$

It is obvious that

$$\overline{\Psi} \cdot \Psi = 1\,. \tag{19.229}$$

Conversely, if we have a unitary representation $a \mapsto \underline{a}_E$ of G, an element Ψ of E such that $\overline{\Psi} \cdot \Psi = 1$, and if G acts *effectively* on Y (definition (16.220)), we could define a function F_G of positive type on $\underline{G}_Y$ (which is a subgroup of $\mathrm{Quant}(Y)$) by

$$F_G(\underline{a}_Y) \equiv \overline{\Psi} \cdot \underline{a^{-1}}_E \cdot \Psi\,. \tag{19.230}$$

If F_G is extendible to a state F on the whole group $\mathrm{Quant}(Y)$, we could say that the representation $a \mapsto \underline{a}_E$ and the vector Ψ give a *partial description* of the state F. This description would be sufficient for computing the average values of the moments

$$\mu \cdot Z \equiv \varpi(Z_Y(\xi))\,.^{515} \tag{19.231}$$

We know that $\underset{\mu\cdot Z}{\delta}\xi \equiv Z_Y(\xi)$ (18.153). Formulæ (19.215♡), (6.12c), and (19.230) then give

$$\text{average}\,(\mu \cdot Z) = i\,\frac{d}{ds}\left[\overline{\Psi} \cdot \underline{\exp(sZ_G)(e)}_E \cdot \Psi\right]_{s=0}, \tag{19.232}$$

where e is the identity element of G.

(19.233) EXAMPLE: Let us consider a *nonrelativistic particle without spin* of mass m. We have seen how to associate an element Ψ of the Planck space to each differentiable function $\mathbf{v} \mapsto \psi$ on $\mathbf{R}^3$ with compact support (formula

[513] See Dixmier [13], §13.

[514] That is, for every $X \in E$, the map $a \mapsto \underline{a}_E \cdot X$ is continuous (using the manifold topology for G and the norm topology for E).

[515] See (18.152).

(19.66)). We have also seen that these vectors Ψ *are not* state vectors (19.70). But we can turn the set $\mathcal{E}$ of these vectors into a *pre-Hilbert space* by defining[516]

$$\overline{\Psi} \cdot \Psi = C \int_{\mathbf{R}^3} |\psi|^2 \varphi(d\mathbf{v}) , \tag{19.234}$$

where φ is the Euclidean density and C is a positive constant.

We have seen that the Galilei group has an 11-dimensional quantizable extension G (19.79), which, in addition, is its largest lift. We know that this group acts on $\mathcal{E}$ (19.80). One can verify that it acts isometrically (for the pre-Hilbert structure (19.234)). This gives, by extension, a *unitary representation* of G on the completion $\mathcal{E}'$ of $\mathcal{E}$ (theorem (18.191)).

(19.235) Given an element Ψ of $\mathcal{E}'$, we can then ask the question whether the function F_G defined by formula (19.230) is extendible to a state on the entire group of quantomorphisms. We will only make some remarks concerning this conjecture.

In order that F_G be extendible to a state, it is necessary that $F_G(1) = 1$ and (since G contains the elements $\underline{z}_Y$) that

$$F_G(\underline{z}_Y) \equiv z \qquad \text{(axioms (19.214}\diamondsuit\text{)} . \tag{19.236}$$

In fact, these two conditions are satisfied as soon as $\overline{\Psi} \cdot \Psi = 1$.

The theory of the Fourier transform shows that we can choose the constant C in (19.234) so that

$$\overline{\Psi} \cdot \Psi = \int_{\mathbf{R}^3} |\tilde{\psi}|^2 \varphi(d\mathbf{r}) , \tag{19.237}$$

where $\tilde{\psi}$ is the *wave function* (19.74). The value of the integral can be calculated at an arbitrary time t.

If conjecture (19.235) is true, formula (19.232) allows us to compute the average value of the moments of G by means of the wave function. For example, for the *energy* E we find the formula

$$\begin{aligned} \text{average}(E) &= C \int_{\mathbf{R}^3} \frac{mv^2}{2} |\psi|^2 \varphi(d\mathbf{v}) \\ &= \frac{1}{2m} \int_{\mathbf{R}^3} \left\{ \left| \frac{\partial \tilde{\psi}}{\partial r_1} \right|^2 + \left| \frac{\partial \tilde{\psi}}{\partial r_2} \right|^2 + \left| \frac{\partial \tilde{\psi}}{\partial r_3} \right|^2 \right\} \varphi(d\mathbf{r}) , \end{aligned} \tag{19.238}$$

[516] Knowing the *complex* vector space structure of $\mathcal{E}$ and the *norm* $\overline{\Psi} \cdot \Psi$, we know the *inner product* $\overline{\Psi} \cdot \Psi'$ (see the footnote in (18.191)).

where the integral on the right hand side is taken at an arbitrary time. The formulæ (19.237, 238) are generally adopted a priori.[517] □

(19.239) EXAMPLE: We have constructed above (19.81 to 19.134) Planck spaces (or subspaces) for various kinds of elementary particles. On each of these spaces the lift of the Poincaré (or Galilei) group acts.

The reader will be able to find *in each case* a pre-Hilbert space structure which is invariant under the group. In the case of the photon, one could take

(19.240) $$\overline{\Psi} \cdot \Psi = -\int_K \overline{A} \cdot A \, \varphi(dP) \qquad \text{(notation (19.145))}.$$

Note that this quantity is invariant under gauge transformations (19.143) and that it is indeed positive for $\Psi \neq 0$.

A study along the lines of the nonrelativistic particle without spin is now possible. □

(19.241) The above remarks allow us to envisage a solution to the difficulties in the interpretation of wave functions (see (19.70), (19.103)).†

[517] If we define $\widetilde{\psi} \equiv \sqrt{\frac{\rho}{m}}\, e^{im\alpha}$, then the formulæ (19.237, 238) can be written as

$$m \equiv \int_{\mathbf{R}^3} \rho \, \varphi(d\mathbf{r})$$

and

$$\text{average}\,(E) = \int_{\mathbf{R}^3} \left[\frac{1}{2m^2} \text{grad}^2 \sqrt{\rho} + \frac{\rho}{2} \text{grad}^2 \alpha \right] \varphi(d\mathbf{r}) .$$

These formulæ suggest a curious analogy with a *fluid* of density ρ, having a *density of elastic energy* (of capillary type) $\frac{1}{2m^2}\text{grad}^2\sqrt{\rho}$ and a velocity $\text{grad}\,\alpha$. Indeed, one can check that *the equations of irrotational motions of such a fluid* (the equation of continuity and the Euler equation) *reduce to the Schrödinger equation*.

† Editors' note: For more recent developments in geometric quantization we refer the reader to the books by Guillemin and Sternberg [16], Hurt [17], Śniatycki [30], Simms and Woodhouse [29], and Woodhouse [39], and references therein.

Bibliography

[3] H. Bacry, H. Ruegg, J.-M. Souriau, *Dynamical groups and spherical potentials in classical mechanics*, Commun. Math. Phys. **3**, (1966), 323–333.

[4] V. Bargmann, *On unitary ray representations of continuous groups*, Ann. of Math. **59**, (1954), 1–46.

[5] V. Bargmann, L. Michel, and V.L. Telegdi, *Precession of the polarization of particles moving in a homogeneous magnetic field*, Phys. Rev. Lett. **2**, (1959), 435–436.

[6] G.D. Birkhoff, *Dynamical Systems*, American Mathematical Society, Providence, 1927.

[7] A. Blanc-Lapierre, P. Casal, and A. Tortrat, *Méthodes mathématiques de la mécanique statistique*, Masson, Paris, 1959.

[8] N. Bourbaki, *Espaces vectoriels topologiques*, Chapitres I–II, Hermann, Paris, 1953.

[9] N. Bourbaki, *Intégration*, Chapitres I–IV, V, Hermann, Paris, 1952 (chapitres I–IV), 1967 (chapitre V).

[10] N. Bourbaki, *Topologie générale*, Chapitres I–II, Hermann, Paris, 1951.

[11] P.A.M. Dirac, *The fundamental equations of quantum mechanics*, Proc. Roy. Soc. A **109**, (1926), 642–653.

[12] P.A.M. Dirac, *The principles of quantum mechanics*, Oxford University Press, Oxford, 1930. Fourth revised edition, 1967.

[13] J. Dixmier, *Les C^*-algèbres et leurs représentations*, Gauthier-Villars, Paris, 1964 = *C^*-algebras / Representations of algebras / Representations of Groups*, North Holland, Amsterdam, 1977.

[14] C. Duval, *On the polarizers of compact semisimple Lie groups. Applications*, Ann. Inst. Henri Poincaré A **34** (1981), 95–115.

[15] F. Gallissot, *Les formes extérieures en mécanique*, Ann. Inst. Fourier **4**, (1952), 145–297.

[16] V. Guillemin and S. Sternberg, *Geometric Asymptotics*, American Mathematical Society, Providence, 1977.

[17] N.E. Hurt, *Geometric Quantization in Action*, Reidel, Dordrecht, 1983.

[18] N. Jacobson, *Lie algebras*, Interscience Publishers (Wiley), New York, 1962 = reprint: Dover Publishing Company, New York, 1979.

[19] E.T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. **106**, (1957), 620–630.

[20] D. Kastler, *Introduction à l'électrodynamique quantique*, Dunod, Paris, 1961.

[21] A.I. Kinchin, *Mathematical Foundations of Statistical Mechanics*, Dover, New York, 1949.

[22] J.-L. Lagrange, *Mémoires de la première Classe de l'Institut pour 1808.*

[23] J.-L. Lagrange, *Mécanique Analytique*, tome II, (posthumous edition) = reprint: Édition complète réunissant les notes de la Troisième édition revue, corrigée et annotée par Joseph Bertrand et de la Quatrième édition publiée sous la direction de Gaston Darboux, A. Blanchard, Paris, 1965.

[24] L.D. Landau and E.M. Lifschitz, *Mécanique*, 3^{e} édition, MIR, Moscou, 1969 = *Mechanics*, Pergamon Press, Oxford, 1976.

[25] L.D. Landau and E.M. Lifschitz, *Mécanique Quantique*, 2^{e} édition, MIR, Moscou, 1967 = *Quantum Mechanics*, third edition, Pergamon Press, Oxford, 1977.

[26] G. Mackey, *Mathematical Foundations of Quantum Mechanics*, Benjamin, New York, 1963.

[27] M.A. Naimark, *Normed Rings*, Noordhoff, Groningen, 1964.

[28] L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1966.

[29] D.J. Simms and N.M.J. Woodhouse, *Lectures on Geometric Quantization*, Lecture Notes in Physics **53**, Springer-Verlag, Berlin, 1997.

[30] J. Śniatycki, *Geometric Quantization and Quantum Mechanics*, Springer-Verlag, New York, 1980.

[31] J.-M. Souriau, *Calcul linéaire*, Tome premier & Tome second, Presses Universitaires de France, Paris, 1964 & 1965 = reprint: *Calcul linéaire*, Jacques Gabay, Paris, 1992.

[32] J.-M. Souriau, *Géométrie et relativité*, Hermann, Paris, 1964.

[33] J.-M. Souriau, *Géométrie de l'espace de phases, calcul des variations et mécanique quantique*, Tirage ronéotypé, Faculté des Sciences de Marseille, 1965.

[34] R.F. Streater, Technical Report # 707, University of Maryland, Department of Physics, 1967.

[35] J. Synge, *The relativistic gas*, North Holland, Amsterdam, 1957.

[36] L.H. Thomas, *The kinematics of an electron with spin*, Phil. Mag. **3**, (1927), 1–22.

[37] A. Visconti, *Théorie quantique des champs*, Gauthier-Villars, Paris, 1961.

[38] G.N. Watson, *A Treatise on the Theory of Bessel Functions*, second edition, Cambridge University Press, Cambridge, 1958.

[39] N.M.J. Woodhouse *Geometric Quantization*, second edition, Oxford University Press, Oxford, 1991.

Index

$*$-algebra, 253
A-transportable measure, 249
E-coboundary of G, 108
E-coboundary of $\mathcal{G}$, 110
E-cochain of G, 107
E-cochain of $\mathcal{G}$, 110
E-cocycle of G, 107
E-cocycle of $\mathcal{G}$, 110
N-body problem, 122
λ entropy, 260
μ-integrable, 246
μ-measurable set, 244
μ-proper map, 249
p-form, 35
p^{th}-order moment, 258
C^*-algebra, 340
$\mathcal{G}$-module, 110
4-momentum, 169

absolute
  convergence, 242
  temperature, 282
  value, 228
action, 156
  integral, 62
  on a manifold, 46
addition of random variables, 257
additive group, 45
adiabatic evolution, 284
adjoint
  operator, 341
  representation, 50
admissible coordinate system, 5
algebra, 253
  of continuous functions, 231
analytic continuation, 212
angle of
  incidence, 217
  reflection, 217
angular
  characteristic, 213
  eikonal, 213
  momentum, 148
  velocity vector, 292
annihilation operator, 374
annihilator, 377
anomalous magnetic moment, 200
anti-canonical similitude, 94
antichronous
  element, 173
  transformation, 192
antilinear map, 340
area of the n-sphere, 245
Aristotle group, 168
artificial gravity, 292
assemblée of particles, 370
asynchronous evolution space, 201
atlas, 4
average value of
  a dynamical variable, 382
  the energy, 385
axioms of
  relativistic mechanics, 167
  symplectic mechanics, 154

Bacry, H., 150
ball, 3
Banach
  algebra, 253
    having an involution, 253
  space, 242
Bargmann Michel Telegdi equation, 200
Bargmann, V., 147, 200

barometric distribution, 292
barycenter, 156
  of a relativistic system, 188
base of prequantum manifold, 309
basis, 4
Birkhoff's theorem, 274
Birkhoff, G.D., 274
black body radiation, 300
Blanc-Lapierre, A., 274, 276
blondels, 254
Bohr's law of radiation, 352
Boltzmann's
  constant, 282
  theorem, 275
boolean variable, 19
Borel-Lebesgue theorem, 225
Bose-Einstein prequantization, 331
bosons, 373
bounded
  linear map, 242
  measure, 243
  scattering source, 210
  set, 225
Bourbaki, N., 6, 226, 229, 230, 241, 245, 259
brilliance, 254
broadening of spectral lines, 281
bundle projection, 9

calculus of variations, 62
canonical
  basis, 75
  chart of a
    Lie group, 51
    symplectic manifold, 90
  coordinates, 91
  equations, 65
  similitude, 94
  transformation, 93
    infinitesimal, 97
Cartan
  decomposition, 59
  form, 66
    of the Lagrangian, 66
  theorem, 37
Casal, P., 274, 276
Casimir, 172
Cauchy-Schwarz inequality, 340
cell, 14
center of mass, 148
  decomposition, 157
central force, 150
centrifugal force, 290
centrifuge, 292
change of reference frame, 133
characteristic
  of an equation, 89
  foliation, 44
chart, 5
classical
  eikonal function, 213
  mass, 121
  mechanics, 121
  stationary state, 152
  system, 346
Clausius-Clapeyron relation, 284
closed set, 224
closure of a set, 225
coadjoint representation, 104
coboundary, 108
coherent light, 376
cohomologous E-cocyles
  of G, 108
  of $\mathcal{G}$, 110
cohomology of a
  dynamical group, 104
  Lie algebra, 108
  Lie group, 107
cohomology space of a
  Lie algebra representation, 110
  Lie group representation, 108
coisotropic
  submanifold, 88
  subspace, 73

color of a nonrelativistic photon, 187
commutator, 47
commuting derivations, 26
compact set, 225
compatible coordinate systems, 4
complete normed space, 242
completely continuous measure, 237
complex
  conjugate, 339
  Euclidean space, 55
  vector space, 54
components of
  a tensor, 34
  exterior derivative, 36
composite
  manifold, 224
  system, 149
composition law, 45
conformal transformation, 95
connected
  component, 15
  set, 15
conservative
  scattering source, 211
  system, 152
constant, 18
constrained
  motion, 207
  system, 122
constraint, 122
  equations, 122
  force, 123
contact structure, 307
continuous mapping, 6
contraction operator, 31
contragredient representation, 104
controlled system, 126
convergent integral, 246
convex set, 258
convolution
  algebra, 253
  of two measures, 253
coordinate system, 3, 223
coordinates, 3
Coriolis force, 290
correspondence principle, 346
Coulomb field, 121
covariant
  field, 31
  operator, 31
  tensor, 34
    antisymmetric, 35
covector, 35
covering space, 13
creation operator, 375
curl, 137
curvature of a trajectory, 196

Darboux's theorem, 91
dense subspace, 340
density, 235
  operator, 382
derivation, 19
derivative of exp, 29
derived algebra, 48
diffeomorphism, 8
differentiable
  covariant field, 33
  map, 4
    between manifolds, 6
  vector field, 24
differential
  equation, 27
  form, 35
dimension of a
  coordinate system, 223
  manifold, 5
  manifold at a point, 223
dimensional equations, 156
Dirac
  electron, 363
  equation, 363
  measure, 235

spinors, 318
Dirac, P., 349
direct
image of a covariant field, 32
product
manifold, 5
of symplectic manifolds, 82
sum, 373
direction of time flow, 297
discrete
group of diffeomorphisms, 14
topological space, 224
disjoint union, 15
displacements, 124
divergence, 137
Dixmier, J., 384
Doppler effect, 188
due to reflection, 217
Doppler-Fizeau effect, 281
double refraction, 211
drag
acceleration, 134
velocity, 134
dual
representation, 104
space, 228
Duval, C., 357
dynamical
group, 100
variable, 85
in involution, 88

eccentric anomaly, 121
eccentricity, 121
effective group action, 266
Einstein summation convention, 34
elastic collision, 219
electric
charge, 138, 194
field, 138
electromagnetic
field, 138, 196
form, 369
potential, 369
electrostatic forces, 203
element of
area, 255
solid angle, 255
elementary
dynamical system, 173
particles, 173
system, 173
embedding, 11
empty map, 4
energy, 149, 169
energy-momentum vector, 169
enthalpy, 285
entropy, 265
of a probability measure, 260
equality of action and reaction, 148
equation of motion, 126
equilibrium, 273
temperature, 276
equipartition of energy, 286
equivalent
prequantizations, 326
vector fields, 24
Euclidean
density, 239
norm, 3
space, 55
Euler's identity, 64
Euler-Lagrange equations, 63
Euler-Poinsot problem, 162
event, 133
evolution space, 127
exponential mapping, 27
extension of a map, 207
exterior
derivative, 36
form, 35

Fermat's principle, 213
Fermi-Dirac prequantization, 331

fermions, 373
ferromagnetism, 148
field of
    covariant tensors, 34
    densities, 235
    vector spaces, 38
first
    homotopy group, 16
    integral, 44
flux, 254
Fock manifold, 300, 370
focus, 121
foliated manifold, 38
foliation, 38
    defined by a Lie subalgebra, 51
    preserving map, 41
force, 121
    of inertia, 290
free
    path, 272
    rigid body, 162
Frobenius' theorem, 39
front of a mirror, 216
full Poincaré group, 189
function, 19
    of bounded variation, 234
    of positive type, 380
fundamental
    group, 16
    vector field, 47
future pointing, 173

Galilean relativity, 135
Galilei
    cohomology, 146
    group, 135
Gallisot, F., 130
gauge, 139
    transformation, 367
Gaussian normal measure, 262
general linear group, 45
generalized Maxwell relation, 284
geometric optics, 212
Gibbs
    canonical ensemble, 267
    measure, 266
    principle, 276
Gram matrix, 59
graph of a field, 23
gravitational acceleration, 138
graviton, 347
group
    algebra, 254
    character, 328
    extension, 336
Guillemin, V., 357, 386
gyromagnetic ratio, 199

Haar measure, 239
Hahn-Banach theorem, 242
half spinors, 318
Hamilton's
    equations, 65
    principle, 140
Hamiltonian
    action, 140
    function, 65
    variational problem, 65
harmonic oscillator, 151
Hartree-Fock approximation, 383
Hausdorff space, 27
heat, 284
    capacity, 285
heating, 284
Heisenberg group, 324, 336
helicity, 182
    of a nonrelativistic massless particle, 187
hermitian, 57
    map, 342
Hilbert space, 340
holonomic constraints, 132
homogeneous function, 63
Hurt, N., 386
hydrogen

atom, 204
thermometer, 281
hyperbolic space, 60
hyperfine structure, 203

ideal
constraints, 124
fluid, 285
gas, 279
approximation, 279
nonholonomic constraint, 132
identical particles, 201
identity element, 45
illumination, 255
image of a
differential equation, 28
measure, 249
variational problem, 68
vector field, 23
immersion, 11
in involution, 88
independent
random variables, 257
variable, 18
index of inertia, 60
indiscernible particles, 202
induced
measure, 241
symplectic form, 87
inertial reference frame, 135
infinitesimal
canonical
similitude, 99
transformation, 97
displacement, 124
quantomorphism, 333
inhomogeneous Lorentz group, 164
initial condition, 126
injective, 3
injective immersion, 11
instantaneous axis of rotation, 134
integrability condition, 39
integrable
map, 248
integral
invariant, 43
manifold, 38
integration on a manifold, 236
interaction, 202
potential, 202
interior product, 35
intrinsic angular momentum, 180
invariant
density, 238
Lagrangian, 68
inverse image of a
p-form, 37
covariant field, 31
set, 6
vector field, 23
inverse of an element, 45
inversion of
space, 189
time, 189
isolated dynamical system, 135
isometry, 341
isomorphic
dynamical systems, 155
lift, 327
isothermal evolution, 284
isotropic
submanifold, 88
subspace, 73
isotropy group, 53

Jacobi identity, 25
Jacobi's
equation, 64
manifold, 89
Jacobson, N., 111
Jaynes, E.T., 276
Joule's experiment, 285

Kastler, D., 363, 368
Kepler's
equation, 122

laws, 122
kernel of a tensor field, 38
Killing form, 110
Kinchin, A.I., 276
kinetic
energy, 149
theory of gases, 271
Klein-Gordon equation, 359

Lagrange form, 81
Lagrange's theorem, 213
Lagrange, J., 130
Lagrangian, 66
extrema, 63
manifold, 89
subspace, 77
Landé g-factor, 199
Landau, L.D., 140, 151, 346
latus rectum, 121
law of
friction, 123
Gay-Lussac, 282
induction, 138
leaf of
a foliation, 40
characteristic foliation, 82
Lebesgue
density, 237
measure, 234
left
invariant measure, 239
translation, 46
Lie
algebra, 48
representation, 110
bracket, 25, 47
derivative, 33
group, 45
subalgebra, 51
subgroup, 51
Lifschitz, E.M., 140, 151, 346
lift of a
group of symplectomorphisms, 327
map, 9
quantomorphism, 327
liftable symplectomorphism, 327
light
cone, 365
ray, 212
lightlike vector, 173
line of force, 40
linear
momentum, 148
oscillator, 151
representation, 46
linearization, 151
Liouville
entropy, 265
measure, 240
Liouville's theorem, 240
liquefaction, 284
locally
compact set, 226
equivalent vector fields, 24
logarithm, 54
Lorentz
contraction, 167
equation, 368
frame, 163
group, 164
moment, 170
momentum, 169
lower bound, 227
lumens, 254
luminance, 254
lux, 255

Mach's principle, 135
Mackey, G., 271, 276
magnetic
Hamiltonian, 198
induction, 138
moment, 198
monopoles, 138

resonance, 198
magnetostatic forces, 148, 203
magnets, 148
manifold, 5, 223
admitting a potential, 84
of tangent vectors, 10
quotient by a foliation, 42
structure, 5
Mariotte-Gay-Lussac distribution, 297
mass of a
bounded measure, 244
relativistic particle, 180
system, 156
massless particle, 182
material point, 121
Maurer-Cartan form, 101
Maxwell's
distribution, 280
equations, 138, 369
principle, 137
reciprocity law, 137
mean
quadratic error, 282
value, 258
measure, 228
of a set, 245
on a manifold, 230
Michel, L., 200
Michelson-Morley experiment, 166
Minkowski space, 163
mirage, 217
mixture of states, 382
model of a dynamical system, 155
modular function, 239
molecular orbitals, 382
molecule, 272
moment, 101
map, 101
of Galilei group, 145
of inertia, 294
momentum
map, 101
operator, 357
monoquantizable manifold, 327
motion, 127
beginning, 129
bounded in time, 129
eternal, 129
finishing, 129
Keplerian, 143
multilinear operator, 34
multiple integral, 235
multiplication, 45

Naimark, M.A., 226, 339, 341
natural equilibrium state, 276
negative Euclidean space, 57
negligible set, 279
neighborhood, 10
neutrino, 193
Newton's
constant, 122
equations, 121
Noether's theorem, 68
nondated motion, 152
nonholonomic constraints, 132
noninteracting particles, 200
nonrelativistic particle with spin, 186
norm, 241
limit, 274
normal
magnetic moment, 200
operator, 341
subgroup, 52
normed vector space, 241
null vector, 173

observables, 349
open
cover, 225
map, 10
set, 3
subset of a manifold, 6

orbit of a group action, 49
orbital
  elements, 143
  momentum, 162
orthochronous
  element, 173
  group, 192
orthogonal, 73
  group, 58
  susrface, 41
orthonormal basis, 59

parity reversal, 189
partial
  derivative, 21
  order, 227
particle
  without spin, 180
partition, 49
past pointing, 173
path, 15
Pauli
  electron, 361
  exclusion principle, 376
  spin matrices, 314
perfectly diffusing, 255
perturbation theory, 143
Pfaffian, 171
phase space, 84
  at time t_0, 127
photographic image, 255
photometric state, 254
photometry, 254
photon, 184
  vacuum, 300
planar mirror, 214
Planck
  foliation, 350
  lift, 325, 351
  manifold, 325
  space, 351
Planck's
  condition, 350
  constant, 302
  formula, 352
  law of radiation, 302
  thermodynamic potential, 277
Poincaré
  group, 16, 164
  lemma, 37
Poisson
  bracket, 86
  distribution, 301
Poisson's theorem, 142
polarization, 171
  of the photon, 183
position, 133
  operator, 357
positive
  element, 227
  Euclidean space, 57
  frequency solution, 360
  linear form, 228
  semidefinite matrix, 380
potential
  energy, 149
  function, 139
  of a 2-form, 84
pre-Hilbert space, 339
precession of spin, 198
prequantizable manifold, 310
prequantization, 309
  by fusion, 315
prequantum manifold, 308
pressure, 281
presymplectic manifold, 82
principal
  axis of inertia, 294
  fiber bundle, 309
principle of
  inertia, 148
  conservation of energy, 149
  equality of action and reaction, 148
  Galilean relativity, 135

virtual work, 124
principles of thermodynamics, 284
probability, 256
measure, 256
product
covering space, 17
of a measure with a continuous function, 231
projection map
to a leaf, 42
proper
acceleration, 196
angular momentum, 162
map, 226
motion, 161
temperature, 297
proposition, 19
pull back, 31, 37
pure manifold, 224
push forward of a
covariant field, 32
density, 237
measure, 235
vector field, 23

quantizable
dynamical group, 335
group extension, 336
quantization of a
classical system, 346
dynamical group, 335
dynamical system, 346
quantization of spin, 361
quantomorphism, 326
quantum mechanical
mean value, 379
probability, 348
stationary state, 351
quantum state, 380
quaternions, 338
quotient manifold, 14
by a foliation, 42

radiation pressure, 301
random variable, 256
rank of presymplectic form, 83
Rayleigh-Jeans formula, 302
reference frame, 133
reflection, 215
refrigeration, 284
regular variational problem, 67
relation of Robert von Mayer, 286
relatively compact set, 226
relativistic
center of mass, 188
centrifuge, 298
ideal gas, 296
interaction, 204
mass, 174
material point, 180
mechanics, 167
moment, 169
particle
with spin, 174
particle without spin, 180
wind, 297
restricted
Lorentz group, 164
Poincaré group, 164
resultant
of forces, 125
of torques, 125
Riesz
space, 227
subspace, 228
right
invariant Haar measure, 239
invariant vector field, 50
translation, 49
rigid body, 123
rocket, 292
rotary birefringence, 211
rotation, 133
Ruegg, H., 150
rule of Dulong and Petit, 287

scalar, 35
  field, 31
  product, 57
  spin, 180
scattered
  motions, 205
  system, 205
scattering
  source, 205
  theory, 205
Schrödinger's equation, 357
Schwartz, L., 230
second quantization, 378
section of a
  foliation, 42
  tangent bundle, 23
sectionable foliation, 42
self orthogonal, 74
self-adjoint operator, 341
semisimple
  Lie algebra, 110
  Lie group, 110
sign of the energy, 182
Simms, D., 386
simply connected, 16
simultaneity, 166
skew hermitian, 57
slowing down of clocks, 167
Sniatycki, J., 386
Souriau, J.-M., 150, 167, 170, 258, 259, 262, 315, 328, 380
space inversion, 189
space of
  linear forms, 228
  motions, 129
space-time, 133
spacelike vector, 173
spatial character, 191
special
  linear group, 54
  orthogonal group, 58
  unitary group, 58
specific heat, 284
spectrum, 302
sphere, 3
spin, 174
  of a nonrelativistic massless particle, 187
stabilizer of a
  Lie subalgebra, 52
  point, 53
standard deviation, 282
state
  function, 381
  of a dynamical system, 382
  vector, 348
static scattering source, 211
statistical
  equilibrium, 273
  state, 271
Stefan's constant, 301
Stefan-Boltzmann law, 301
step function, 209
Stern-Gerlach experiment, 198
Sternberg, S., 357, 386
Stieltjes integral, 234
Streater, R.F., xxxiv
strictly positive density, 260
submanifold, 11
sum of two vector spaces, 38
summable function, 246
superfluid helium, 376
support of a
  function, 229
  measure, 241
  scattering source, 205
symmetric
  group, 201
  relation, 73
symmetry of a scattering source, 208
symplectic
  cocycle of G, 111
  cohomology space of G, 112

density, 239
gradient, 85
group, 77
manifold, 81
matrix, 78
mechanics, 154
vector space, 76
symplectomorphic manifolds, 93
symplectomorphism, 93
local, 93
synchronous evolution space, 200
Synge, J., 297
system of
forces, 101
material points, 122

tachyon, 185
tangent
bundle, 9
map, 7
space, 7
vector, 7
Telegdi, V.L., 200
temperature vector, 297
temporal character, 192
tensor field, 35
tensor product of measures, 232
theory of elasticity, 167
thermodynamic
entropy, 283
potential, 283
Thomas, L.H., 200
time reversal, 189
timelike vector, 173
topological
dual space, 242
space, 6
torsor, 101
Tortrat, A., 274, 276
torus, 307
total
force, 148
mass, 148
torque, 148
trajectory of a
massless particle, 184
particle, 176
transmission, 215
transpose, 56
transverse
intersection, 42
section, 42
two component spinors, 318

ultraviolet catastrophe, 302
unconstrained motion, 207
uniform photometric state, 254
unimodular, 239
unit, 341
unitary group, 58
of a C^* algebra, 341
universal covering space, 16
upper bound, 227
Urysohn's lemma, 230

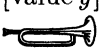
vacuum, 295, 374
value of a variable, 18
variable, 18
variance, 282
tensor, 260
variation, 62
of constants, 143
vector
field, 19
lightlike, 173
future pointing, 173
past pointing, 173
potential, 139
product, 133
spacelike, 173
spin, 180
timelike, 173
velocity, 121
compatible with constraints, 123
of light, 163

vibrations, 151
virtual
    displacements, 124
    work, 124
Visconti, A., 315
volume of the n-ball, 245

Watson, G.N., 296
wave
    equation, 368
    front, 183
    function, 357
wavelength, 353
    of a photon, 213
weak limit, 274
    in the mean, 274
Weyl group, 336
Whitehead's lemma, 111
wind velocity, 297
Woodhouse, N., 386
work, 284

List of notation

notation	term	page
$\mathrm{Can}(V)$	group of canonical transformations on V	94
$\mathcal{C}(V)$	space of continuous functions on V	229
curl	curl	137
d	exterior derivative	36
δ	derivation	19
δ_L	Lie derivative	34
∂_j	partial derivative	21
$\overset{\delta}{u}$	derivation associated to a dynamical variable	85
D_x	tangent space at the point x	7
$D(A)(x)$	tangent map of A at x	7
A^D, V^D	tangent bundle	9
dom	domain of a map	3
div	divergence	137
exp	exponential map	27
$g(\ )(\ )$	real inner product	55
$\mathrm{Gl}(E)$	the general linear group on E	45
$\mathrm{Gl}(n)$	the general linear group on $\mathbf{R}^n$	59
$\mathrm{Gl}(n, \mathbf{C})$	the general linear group on $\mathbf{C}^n$	59
grad	symplectic gradient	85
$\mathcal{H}(Y)$	Hilbert space on Y	339
inf	infimum	227
$j(\ )(\ )$	vector product in 3-dimensions	133
$j(\ )(\ )(\ )$	vector product in 4-dimensions	170
$\mathcal{K}(V)$	space of continuous functions with compact support on V	229
ker	kernel	67
$\mathrm{L}(E)$	linear maps from E to itself	45
$\mathrm{L}^1(V)$	space of measures on a manifold V	245
mass( )	mass of a measure	244
$\mu = \{\mathbf{l}, \mathbf{g}, \mathbf{p}, E\}$	Galilean torsor	146
$\mu = \{M, P\}$	relativistic torsor	165
$\mathrm{Log}(A)$	logarithm of the linear map A	54
$\mathrm{O}(E)$	orthogonal group of E	58
$\mathrm{O}(p, q)$	orthogonal group of index q	60
orth( )	symplectic orthogonal	73
p_j, ${}^j q$	canonical coordinates	91
pf	paffian	171

notation	term	page
$\mathrm{Quant}(Y)$	group of quantomorphisms of Y	327
Re	the real part of	318
sup	supremum	227
$\mathrm{Sl}(E)$	special linear group of E	54
$\mathrm{SO}(E)$	special orthogonal group of E	58
$\mathrm{Sp}(E)$	symplectic group of E	77
$\mathrm{SU}(E)$	special unitary group of E	58
$\mathrm{U}(E)$	the unitary group of E	58
$\mathrm{U}(p,q)$	the unitary group of index q	60
σ, $\sigma(\ )(\ )$	the Lagrange form	130
$\sigma_1, \sigma_2, \sigma_3$	Pauli matrices	337
im	image of a map	3
vol	volume form	170
$x \mapsto y$	a mapping taking x to y	19
A_+	image under A	23
A_-	inverse image under A	31
$[\,,\,]$	Lie bracket	25
$[\,,\,]_P$	Poisson bracket	86
$\underline{a}_V$	action of a on the manifold V	46
Z_V	fundamental vector field	47
$\widehat{x}(a)$	is the same as $\underline{a}_V(x)$	46
$\overline{A}$	the transpose of A	56
$*$	a special linear operator in dimension 4	170
$\otimes$	tensor product of measures	232
$X^+, X^-, \lvert X\rvert$	operators in a Riesz space	228
1_E	identity map on E	6
E_θ	modified coadjoint action	111
[value y]	value of a variable y	18
	a symbol indicating something important to the reader	v
	a symbol warning the reader	v