An elementary course on the

Mathematical Structure of Classical Dynamics

I. Historical and Analytical Dynamics

Renato Grassini

Dipartimento di Matematica e Applicazioni

Università di Napoli Federico II

Renato Grassini, *An elementary course on the Mathematical Structure of Classical Dynamics, I. Historical and Analytical Dynamics*, 2013.

Mathematics Subject Classification:
70H45, 58F05, 34A09

Keywords:
Differential equations on Euclidean spaces.
Newton equation, d'Alembert equation, Euler equation.
Lagrange equations, Hamilton equations.

Preface

This paper contains a group of lectures illustrating – at an undergraduate level – the mathematical structure of Classical Dynamics, both in its historical (Newtonian-d'Alembertian) formulation and in its analytical (Lagrangian-Hamiltonian) version.

From the empirical point of view, dynamics basically deals with the problem (chapter 1) of 'predicting' all the *motions*, with respect to an uninfluential observer, that are *possible* for a constrained particle system under the action of a given $\delta\acute{\upsilon}\nu\alpha\mu\iota\varsigma$ (force field). In this connection, the first step is to establish a 'time-evolution law' characterizing the *dynamically possible motions*, and the second step is that of carrying out a discussion of the above law in order to obtain a qualitative picture and/or a quantitative determination of the dynamically possible motions.

From the mathematical point of view, [1] the time-evolution law established in Newtonian-d'Alembertian dynamics (chapter 2) will be shown to result in an *implicit differential equation* – d'Alembert equation – expressed and elementarly discussed in the *geometrical* (i.e.*coordinate-free*) formalism of a Euclidean affine space associated with the observer (d'Alembert equation will be obtained from the historical Newton equation for unconstrained systems by taking the possible dynamical effects of the constraints into due consideration, and will then be specialized in the classical Newton and Euler's equations for rigid systems).

A further discussion of the time-evolution law will be carried out by translating d'Alembert equation into a system of *ordinary differential equations* with the aid of the *analytical* (i.e.*coordinate*) formalism traditionally adopted in Lagrangian-Hamiltonian dynamics (chapter 3). The price to pay for the above translation will generally be to give up *global* dynamics (study of the dynamically possible motions in the whole space allowed to the particle system by the constraints) and to confine oneself to *local* dynamics (study of the dynamically possible motions within the region of the above space covered by the arbitrarily chosen system of coordinates).

So what is left is to give a deeper insight into d'Alembert equation with the aim of gaining consciousness of the geometric objects directly allowing a global discussion of the equation itself, and, in this way, to develop a 'geometrical' – rather than 'analytical'– formulation of Lagrangian-Hamiltonian dynamics.

But that is the subject matter for higher courses. [2]

---

[1] The mathematical background is treated in Appendix (chapter 4), whose reading is meant to preceed that of the main text.

[2] For an introductory course, see our *Introduction to the Geometry of Classical Dynamics* 2nd edition (2012) (www.docenti.unina.it).

Contents

Chapter 1

The problem of dynamics

Classical dynamics deals with the 'first' of all theoretical-physical problems, which – in the simplest cases – can be described in both empirical and mathematical terms as follows.

1.1 The data

Here is the list of the data of the problem.

It is based on the following two preliminary concepts.

A *reference space* – 'mathematical extension' of a rigid body (like the Earth, a vehicle, the Moon, a planet, the Sun, a star) meant to be the location of an uninfluential observer – is conceived as a 3-dimensional, Euclidean affine space (usually denoted by $\mathcal{E}_3$ and modelled on a vector space E_3 with Euclidean metric $\cdot$).

Time – ordering events on a graduated scale, meant to be independent of any observer – is conceived as a 1-dimensional, oriented, Euclidean affine space (isometrically identified with the oriented real line $\mathbb{R}$).

1.1.1 Configuration space

'Particle' is synonymous with 'point-like body', i.e. a body whose *position* in the chosen reference space $\mathcal{E}_3$ is conventionally defined as a single point of $\mathcal{E}_3$.

So, for an ordered system of ν particles, a position – or *configuration* – in $\mathcal{E}_3$ will be defined as a single point of the Cartesian power $\mathcal{E} := {\mathcal{E}_3}^{\nu}$ (3ν-dimensional Euclidean affine space, modelled on $E := {E_3}^{\nu}$). [1]

[1] See Appendix, section 4.1.1, Exercise 3. See also Appendix, section 4.2.1, Exercise 10, where the Euclidean metric in E is defined by putting $\mathrm{u} \cdot \mathrm{v} := \sum_{i=1}^{\nu} \mathrm{u}_i \cdot \mathrm{v}_i$, for all $\mathrm{u} = (\mathrm{u}_1, \ldots, \mathrm{u}_\nu)$ and $\mathrm{v} = (\mathrm{v}_1, \ldots, \mathrm{v}_\nu)$ in E.

The given particle system may generally be subject in $\mathcal{E}_3$ to some time-independent, *holonomic* (i.e. positional) *constraints*, owing to which it is virtually allowed to occupy only the *admissible configurations* belonging to a region

$$Q \subset \mathcal{E}$$

Q will be assumed to be a smooth manifold embedded in $\mathcal{E}$.

The above manifold Q and its dimension $n := \dim(Q) \leq 3\nu$ are called *admissible configuration space* and *number of the degrees of freedom*, respectively, of the particle system in $\mathcal{E}_3$.

Generally, Q consists of all the points $\mathrm{p} \in \mathcal{E}$ satisfying some scalar inequalities $\{g_\alpha(\mathrm{p}) > 0\}_{\alpha=1,\ldots,\mu}$, called strict *one-sided constraints*, and/or equalities $\{f_\beta(\mathrm{p}) = 0\}_{\beta=1,\ldots,\kappa<3\nu}$, called *two-sided constraints.* As is known, under suitable hypotheses of regularity (continuity and differentiability) on the g_α's and f_β's, Q is a smooth manifold of $\mathcal{E}$, whose dimension $n = 3\nu - \kappa$ is given by the dimension of the Euclidean environment $\mathcal{E}$ minus the number of the two-sided constraints (in presence of strict one-sided constraints only, Q is an open – and then 3ν-dimensional – manifold of $\mathcal{E}$). [2]

Then, for any $\mathrm{p} \in Q$, the tangent vector space $T_\mathrm{p}Q$ is the set of vectors $\delta\mathrm{p} \in E$ satisfying the scalar equalities $\{d_\mathrm{p} f_\beta\,(\delta\mathrm{p}) = 0\}_{\beta=1,\ldots,\kappa}$. Such vectors can be regarded as *virtual displacements* (starting from $\mathrm{p} \in Q$), i.e. displacements 'virtually' allowed by the constraints, since any 'small' $\delta\mathrm{p} \in T_\mathrm{p}Q$ takes – up to higher order infinitesimals – the point p belonging to Q, i.e. satisfying $g_\alpha(\mathrm{p}) > 0$ and $f_\beta(\mathrm{p}) = 0$, to a point $\mathrm{p}+\delta\mathrm{p}$ still belonging to Q i.e. satisfying $g_\alpha(\mathrm{p}+\delta\mathrm{p}) > 0$ by continuity and $f_\beta(\mathrm{p}+\delta\mathrm{p}) \approx f_\beta(\mathrm{p}) + d_\mathrm{p} f_\beta\,(\delta\mathrm{p}) = 0$ by differentiability. [3]

1.1.2 Mass distribution

The response of the system to any internal or external influence, will generally depend on how 'massive' its particles are, the *mass* of a particle being conceived as a positive scalar quantity. [4]

The *mass distribution* carried by the system will be denoted by

$$m = (m_1, \ldots, m_\nu)$$

[2] See Appendix, section 4.5.2, *Implicit Function Theorem* and *Open subspaces.*

[3] See Appendix, section 4.5.2, Proposition 14 and Exercise 38.

[4] An 'operational' definition of *mass* will follow from the Newtonian theory (see section 2.1.4).

1.1.3 Force field

The 'force' $-\delta\acute{\upsilon}\nu\alpha\mu\iota\varsigma-$ resultant of all the internal and/or external influences acting in $\mathcal{E}_3$ on the particles of the system, is described as a vector-valued mapping

$$\mathrm{F} : \mathbb{R} \times TQ \to E : (t, \mathrm{p}, \mathrm{v}) \mapsto \mathrm{F}(t, \mathrm{p}, \mathrm{v}) = (\mathrm{F}_1(t, \mathrm{p}, \mathrm{v}), \ldots, \mathrm{F}_\nu(t, \mathrm{p}, \mathrm{v}))$$

where $\mathbb{R}$ is meant to be the *time* and $TQ \subset T\mathcal{E}$ the space of the *admissible positions* and *virtual velocities* allowed by the constraints. [5]

Once assigned (on empirical grounds) such a 'law of force' F, it will be said to be the *force field* acting in the reference space $\mathcal{E}_3$ on the particle system. [6]

F will be assumed to be a *smooth mapping*, in the sense of being the restriction of a C^∞ differentiable mapping defined on an open subset of $\mathbb{R}\times T\mathcal{E}$ containing $\mathbb{R} \times TQ$.

If F does not depend on time or just depends on position, it will be said to be a *time-independent* or *positional* force field (and its values will be denoted by $\mathrm{F}(\mathrm{p}, \mathrm{v})$ or $\mathrm{F}(\mathrm{p})$), respectively.

1.2 The question

With reference to the above data, the problem of dynamics will now be stated.

1.2.1 Smooth motions

The unknown of the problem is 'motion' $-\kappa\acute{\iota}\nu\eta\mu\alpha-$ whose matematical description is the concern of *kinematics.*

A *smooth motion* of the particle system in the reference space $\mathcal{E}_3$ is described as a smooth (i.e. C^∞ differentiable) parametrized curve

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t)$$

of the Euclidean space $\mathcal{E}$, establishing a configuration $\mathrm{p}(t)$ at each *time* t of an open *time interval* $I \subset \mathbb{R}$ (if I is closed, γ is meant to be the restriction of a smooth motion defined on an open interval containing I). [7]

[5] For the concepts of *velocity* and *virtual velocity*, see section 1.2.1.

[6] In assigning F, any possible interaction between two particles of the system will be required to obey Newton's principle of action and reaction (see section 2.1.4).

[7] If γ is a constant mapping, it is said to degenerate into a *state of rest*.

Along γ, owing to C^∞ differentiability, the positions

$$\mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t)) \in \mathcal{E}$$

of the particles, their *velocities*

$$\dot{\mathrm{p}}(t) = (\dot{\mathrm{p}}_1(t), \ldots, \dot{\mathrm{p}}_\nu(t)) \in E$$

and their *accelerations*

$$\ddot{\mathrm{p}}(t) = (\ddot{\mathrm{p}}_1(t), \ldots, \ddot{\mathrm{p}}_\nu(t)) \in E$$

(as well as higher-order derivatives) are all differentiable functions of time. [8]

Remark that, along γ, the first derivative $\dot{\mathrm{p}}(t)$ is meant to be the *velocity* of the particle system at time t, since it measures – up to higher order infinitesimals – the displacements

$$\mathrm{p}(t+1) - \mathrm{p}(t) = (\, \mathrm{p}_1(t+1) - \mathrm{p}_1(t), \ldots\ldots, \mathrm{p}_\nu(t+1) - \mathrm{p}_\nu(t)\,)$$

of the particles in the unit time interval $[t, t+1]$. In the same way, the second derivative $\ddot{\mathrm{p}}(t)$ is meant to be the *acceleration* of the particle system at time t, since it measures – up to higher order infinitesimals – the increments of velocity

$$\dot{\mathrm{p}}(t+1) - \dot{\mathrm{p}}(t) = (\, \dot{\mathrm{p}}_1(t+1) - \dot{\mathrm{p}}_1(t), \ldots\ldots, \dot{\mathrm{p}}_\nu(t+1) - \dot{\mathrm{p}}_\nu(t)\,)$$

of the particles in the unit time interval $[t, t+1]$.

Clearly, any vector of E is the velocity (or the acceleration) at a certain time along some smooth motion.

In particular, a smooth motion γ will be said to be an *admissible motion*, if it is allowed by the constraints, i.e. its orbit lies on Q:

$$\mathrm{Im}(\gamma) \subset Q$$

As a consequence, for every $\mathrm{p} \in Q$, any vector $\mathrm{v} \in T_{\mathrm{p}}Q$ can be regarded as a *virtual velocity*, i.e. a velocity 'virtually' allowed by the constraints, since – by the very definition of tangent vector – it is the velocity at a certain time along an admissible motion passing at that time through p.

[8] See Appendix, section 4.4.2, C^∞ *differentiable curves* and Exercise 31.

1.2.2 Dynamically possible motions

Smooth dynamics basically deals with the time-evolution problem – in the unknown γ – expressed by the following question:

“ For the above constrained point-mass system in the chosen reference space, what are the *smooth motions* that are *possible* in presence of the given $\delta\acute{\upsilon}\nu\alpha\mu\iota\varsigma$? ”

1.2.3 Mechanical system

Such motions will be called the *dynamically possible motions* (DPMs) of *mechanical system*

$$\mathcal{S} := (Q, m, \mathrm{F})$$

They will be compared, in the sequel, with the *inertial motions* of $\mathcal{S}$, which would be possible in absence of force ($\mathrm{F} = 0$), i.e. the DPMs of the *associated* mechanical system

$$\mathcal{S}^{(o)} := (Q, m, 0)$$

Chapter 2

Historical dynamics

Here is the historical (Newtonian-d'Alembertian) answer to the 'predictive' question of dynamics.

2.1 Newton

Newton's answer to the problem of dynamics is primarily referred to a particle system with the maximal number $n = 3\nu$ of degrees of freedom, that is, an unconstrained system or one subject at most to 'real' or 'fictitious' strict one-sided constraints (which – as is clear on empirical grounds - cannot produce any 'dynamical' effect). [1]

Therefore, in examining the mathematical structure of Newtonian dynamics, we shall consider a mechanical system $\mathcal{S} = (Q, m, \mathrm{F})$ whose configuration space Q is an open manifold of $\mathcal{E}$ (and whose force field F *does not* include any dynamical constraint effect).

2.1.1 Newton equation

After Newton, a smooth motion of the particle system in $\mathcal{E}_3$ – i.e. a smooth parametrized curve of $\mathcal{E}$ – is a DPM of $\mathcal{S} = (Q, m, \mathrm{F})$, iff it is a solution of

[1] Think, for instance, of an unconstrained particle system, whose possible 'collisions' (say, collisions between any particle of the system and an external one situated at a fixed point $\mathrm{o} \in \mathcal{E}_3$ and – in the case $\nu > 1$ – collisions between any two particles of the system) we want to leave out of our study. In such a case, the configuration space will just be assumed to be the open manifold

$$Q = \{(\mathrm{p}_1, \ldots, \mathrm{p}_\nu) \in \mathcal{E} \mid d(\mathrm{p}_i, \mathrm{o}) > 0\,,\ d(\mathrm{p}_i, \mathrm{p}_j) > 0\,,\ \forall\, i \neq j = 1, \ldots, \nu\}$$

excluding – through the above 'fictitious', strict, one-sided constraints – that, in an admissible configuration of the system, any two particles can ever occupy the same position.

the second-order differential equation on $\mathcal{E}$, reducible to normal form on Q, defined by [2]

$$\begin{aligned} \mathrm{D}_{Newt} &:= \{(t,\mathrm{p},\mathrm{v},\mathrm{a}) \in \mathbb{R}\times T^2\mathcal{E} \mid \mathrm{p}\in Q\,,\; m\,\mathrm{a} = \mathrm{F}(t,\mathrm{p},\mathrm{v})\} \\ &= \{(t,\mathrm{p},\mathrm{v},\mathrm{a}) \in \mathbb{R}\times T^2\mathcal{E} \mid (t,\mathrm{p},\mathrm{v})\in \mathbb{R}\times TQ\,,\; \mathrm{a} = \frac{1}{m}\,\mathrm{F}(t,\mathrm{p},\mathrm{v})\} \\ &= \mathrm{Graph}\left(\frac{1}{m}\,\mathrm{F}\right) \end{aligned}$$

and called *Newton equation* associated with $\mathcal{S}$.

If F is time-independent, we have $\mathrm{D}_{Newt} = \mathbb{R}\times\mathcal{D}_{Newt}$, where

$$\mathcal{D}_{Newt} := \{(\mathrm{p},\mathrm{v},\mathrm{a}) \in T^2\mathcal{E} \mid \mathrm{p}\in Q\,,\; m\,\mathrm{a} = \mathrm{F}(\mathrm{p},\mathrm{v})\}$$

is a time-independent differential equation, which can as well take the place of D_{Newt}. [3]

2.1.2 Newton's law of dynamics

So, a smooth motion $\gamma : t\in I \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t),\ldots,\mathrm{p}_\nu(t)) \in \mathcal{E}$ is a DPM of $\mathcal{S}$, iff it satisfies *Newton's law of dynamics* [4]

$$(t,\mathrm{p}(t),\dot{\mathrm{p}}(t),\ddot{\mathrm{p}}(t)) \in \mathrm{D}_{Newt}$$

which includes both the *admissibility condition*

$$\mathrm{p}(t)\in Q$$

exhibiting the 'kinematical' effect of the constraints (which only allow motions living in the constraint manifold Q), and the classical *Newton's principle of dynamics*

$$m\,\ddot{\mathrm{p}}(t) = \mathrm{F}(t,\mathrm{p}(t),\dot{\mathrm{p}}(t))$$

[2] For any $\mathrm{w} = (\mathrm{w}_1,\ldots,\mathrm{w}_\nu)\in E$, put

$$m\,\mathrm{w} := (m_1\mathrm{w}_1,\ldots,m_\nu\mathrm{w}_\nu)\in E$$

and

$$\frac{1}{m}\,\mathrm{w} := \left(\frac{1}{m_1}\,\mathrm{w}_1,\ldots,\frac{1}{m_\nu}\,\mathrm{w}_\nu\right)\in E$$

Clearly, for any $\mathrm{u} = (\mathrm{u}_1,\ldots,\mathrm{u}_\nu)\in E$, we have $\mathrm{u} = \frac{1}{m}\mathrm{w}$,iff $\mathrm{w} = m\mathrm{u}$.

[3] See Appendix, section 4.6.2, *Second-order differential equations.*

[4] "$\forall\, t\in I$" will generally be understood.

that is,

$$m_i\,\ddot{\mathrm{p}}_i(t) = \mathrm{F}_i(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\ , \quad i = 1, \ldots, \nu$$

Newton's principle of dynamics requires that, along a DPM, the force $\mathrm{F}_i(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$, *active* on the i-th particle of the system, is to 'counterbalance' the *inertial force* $-m_i\,\ddot{\mathrm{p}}_i(t)$ of the particle, that is, the sum of active and inertial forces is to vanish. [5]

Newton's principle of dynamics (in normal form)

$$\ddot{\mathrm{p}}(t) = \frac{1}{m}\,\mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$$

that is,

$$\ddot{\mathrm{p}}_i(t) = \frac{1}{m_i}\,\mathrm{F}_i(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\ , \quad i = 1, \ldots, \nu$$

equivalently requires that, along a DPM, the acceleration $\ddot{\mathrm{p}}_i(t)$, prescribed to the i-th particle of the system by $\mathrm{F}_i(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$, is in norm to be inversely proportional to m_i, which therefore measures the greater or lesser 'inertia' opposed by the particle to the 'accelerating influence' of the force and then deserves the name of *inertial mass*.

Owing to its reducibility to normal form and the smoothness of F, D_{Newt} is *deterministic*, that is, for any choice of Cauchy data

$$(t_o, (\mathrm{p}_o, \mathrm{v}_o)) \in \mathbb{R} \times TQ$$

there exists a unique maximal solution to Cauchy problem

$$(\mathrm{D}_{Newt}, t_o, (\mathrm{p}_o, \mathrm{v}_o))$$

(all of the other solutions to the above problem being just restrictions of the maximal one). [6]

So, once that the position p_o and the velocity v_o of the particle system at an initial time t_o are known, Newton equation proves to be *potentially* able to *determine* uniquely the whole time-evolution of the system, since it *predicts* the existence of a *unique* maximal DPM of $\mathcal{S}$ –say $\gamma : t \in I \mapsto \mathrm{p}(t) \in Q \subset \mathcal{E}$ with $t_o \in I$– satisfying the initial conditions $\mathrm{p}(t_o) = \mathrm{p}_o$ and $\dot{\mathrm{p}}(t_o) = \mathrm{v}_o$.

[5] See also the remark at the end of section 2.1.5 and footnote [8].
[6] See Appendix, section 4.6.2, *Determinism theorem*.

2.1.3 Newton's law of inertia

Owing to the above theory, the inertial motions of $\mathcal{S} = (Q, m, \mathrm{F})$, i.e. the DPMs of $\mathcal{S}^{(o)} = (Q, m, 0)$, are the solutions of the time-independent differential equation

$$\begin{aligned} \mathbb{R} \times \mathcal{D}^{(o)}_{Newt} &= := \mathbb{R} \times \{(\mathrm{p}, \mathrm{v}, \mathrm{a}) \in T^2\mathcal{E} \mid \mathrm{p} \in Q\,,\ m\mathrm{a} = 0\} \\ &= := \mathbb{R} \times \{(\mathrm{p}, \mathrm{v}, \mathrm{a}) \in T^2\mathcal{E} \mid \mathrm{p} \in Q\,,\ \mathrm{a} = 0\} \end{aligned}$$

associated with $\mathcal{S}^{(o)}$ (remark that $\mathcal{D}^{(o)}_{Newt}$ is a *universal* equation, in the sense that it does not depend – and then its solutions will not depend – on the mass distribution of the system).

So $\gamma : t \in I \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t)) \in \mathcal{E}$ is an inertial motion of $\mathcal{S}$, iff it satisfies *Newton's law of inertia*

$$(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathcal{D}^{(o)}_{Newt}$$

which includes both the admissibility condition

$$\mathrm{p}(t) \in Q$$

and the classical *Newton's principle of inertia*

$$\ddot{\mathrm{p}}(t) = 0$$

that is,

$$\ddot{\mathrm{p}}_i(t) = 0\,, \quad i = 1, \ldots, \nu$$

requiring that, along an inertial motion, the acceleration of each particle is to vanish.

Clearly, for any choice of Cauchy data $(t_o, (\mathrm{p}_o, \mathrm{v}_o)) \in \mathbb{R} \times TQ$, the unique maximal solution to Cauchy problem $(\mathcal{D}^{(o)}_{Newt}, t_o, (\mathrm{p}_o, \mathrm{v}_o))$ is given by the restriction $\gamma := \varphi|_I$ of the affine mapping

$$\varphi : \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t) := \mathrm{p}_o + (t - t_o)\,\mathrm{v}_o$$

to the largest open interval $I \ni t_o$ contained in the open subset $\varphi^{-1}(Q) \subset \mathbb{R}$.

Putting $\mathrm{p}_o = (\mathrm{p}_{o1}, \ldots, \mathrm{p}_{o\nu}) \in Q \subset \mathcal{E}$ and $\mathrm{v}_o = (\mathrm{v}_{o1}, \ldots, \mathrm{v}_{o\nu}) \in T_{\mathrm{p}_o}Q = E$, the above solution reads, for all $i = 1, \ldots, \nu$ and $t \in I$,

$$\mathrm{p}_i(t) = \mathrm{p}_{oi} + (t - t_o)\,\mathrm{v}_{oi}$$

If $\mathrm{v}_{oi} = 0$, the inertial motion of the i-th particle degenerates into the state of *rest*

$$\mathrm{p}_i(t) = \mathrm{p}_{oi}$$

If $\mathrm{v}_{oi} \neq 0$, the inertial motion of the i-th particle is *rectilinear*

$$\mathrm{p}_i(t) \in \mathrm{p}_{oi} + \mathrm{Span}\,(\mathrm{v}_{oi})$$

and *uniform*

$$|\dot{\mathrm{p}}_i(t)| = |\mathrm{v}_{oi}| = \text{const.} > 0$$

Remark that, under the action of a force field $\mathrm{F} \neq 0$, the DPMs will be expected to deviate from the inertial motions (obeying Newton's principle of inertia), by virtue of the not all vanishing accelerations prescribed to them by F (through Newton's principle of dynamics). [7]

2.1.4 Newton's law of action and reaction

Finally recall that any kind of physical *interaction* $(\mathrm{f}_1(\mathrm{p}_1, \mathrm{p}_2), \mathrm{f}_2(\mathrm{p}_1, \mathrm{p}_2))$ between two particles is meant to obey, at each admissible configuration $(\mathrm{p}_1, \mathrm{p}_2)$ with $\mathrm{p}_1 \neq \mathrm{p}_2$, *Newton's principle of action and reaction*

$$0 \neq \mathrm{f}_1(\mathrm{p}_1, \mathrm{p}_2) = -\,\mathrm{f}_2(\mathrm{p}_1, \mathrm{p}_2) \in \mathrm{Span}\,(\mathrm{p}_1 - \mathrm{p}_2)$$

As a consequence, if we consider a DPM $\mathrm{p}(t) = (\mathrm{p}_1(t), \mathrm{p}_2(t))$ of two particles in a reference space where they are only subject to a given interaction, from Newton's principle of dynamics

$$m_1\,\ddot{\mathrm{p}}_1(t) = \mathrm{f}_1(\mathrm{p}(t))\ , \quad m_2\,\ddot{\mathrm{p}}_2(t) = \mathrm{f}_2(\mathrm{p}(t))$$

it follows that

$$m_1\,\ddot{\mathrm{p}}_1(t) = -\,m_2\,\ddot{\mathrm{p}}_2(t) \neq 0$$

whence

$$m_1\,|\ddot{\mathrm{p}}_1(t)| = m_2\,|\ddot{\mathrm{p}}_2(t)| \neq 0$$

and then

$$\frac{m_2}{m_1} = \frac{|\ddot{\mathrm{p}}_1(t)|}{|\ddot{\mathrm{p}}_2(t)|}$$

So, once that a *unit mass* particle has been chosen (say $m_1 = 1$), we infer, for any other particle, the 'operational' definition of mass

$$m_2 = \frac{|\ddot{\mathrm{p}}_1(t)|}{|\ddot{\mathrm{p}}_2(t)|}$$

[7] According to a pre-Newtonian conception, a force field would rather prescribe the velocities of the DPMs (through a law that would therefore result in a first-order differential equation) and every inertial motion would then degenerate into a state of rest.

2.1.5 Newton's law of statics

Within the area of dynamics, concerning the DPMs of $\mathcal{S} = (Q, m, \mathrm{F})$, *statics* aims at finding out those DPMs which possibly degenerate into a state of rest.

In other words, the problem of statics, is that of determining the *equilibrium configurations* of $\mathcal{S}$, that is, the configurations $\mathrm{p}_* \in \mathcal{E}$ s.t.

$$\gamma_* : \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}_*(t) := \mathrm{p}_*$$

(i.e. the state of rest at p_*) is a DPM of $\mathcal{S}$.

Now, owing to Newton's law of dynamics, a configuration $\mathrm{p}_* \in \mathcal{E}$ is an equilibrium configuration, iff

$$\mathrm{p}_*(t) \in Q, \quad m\,\ddot{\mathrm{p}}_*(t) = \mathrm{F}(t, \mathrm{p}_*(t), \dot{\mathrm{p}}_*(t))$$

(with $\mathrm{p}_*(t) = \mathrm{p}_*$ and $\dot{\mathrm{p}}_*(t) = \ddot{\mathrm{p}}_*(t) = 0$), that is, iff it satisfies *Newton's law of statics*

$$\mathrm{p}_* \in Q, \quad 0 = \mathrm{F}(t, \mathrm{p}_*, 0)$$

As a consequence, the problem of extracting from Q the equilibrium configurations can be solved by means of *Newton's principle of statics*, stating that an admissible configuration $\mathrm{p}_* \in Q$ is an equilibrium configuration, iff F vanishes at p_*, namely

$$\mathrm{F}(t, \mathrm{p}_*, 0) = 0$$

that is,

$$\mathrm{F}_i(t, \mathrm{p}_*, 0) = 0\ , \quad i = 1, \ldots, \nu$$

Remark that, if F is the resultant of a number of force fields, say $\mathrm{F} = \mathrm{F}' + \mathrm{F}''$, the equilibrium configurations of $\mathcal{S}$ are characterized by Newton's law of statics as those configurations p_* of Q where $\mathrm{F}'(t, \mathrm{p}_*, 0)$ and $\mathrm{F}''(t, \mathrm{p}_*, 0)$ *counterbalance* each other, i.e. their sum vanishes. [8]

2.2 d'Alembert

d'Alembert's answer to the problem of dynamics is generally referred to a particle system with any number $n \leq 3\nu$ of degrees of freedom, that is, a system which may also be subject to two-sided constraints (which – as is clear on empirical grounds – can produce 'dynamical' effects).

[8] In a similar manner, the DPMs of $\mathcal{S}$ are 'statically' characterized by Newton's law of dynamics as those admissible motions $\mathrm{p}(t)$ along which the active force $\mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$ and the inertial force $-m\,\ddot{\mathrm{p}}(t)$ counterbalance each other.

Therefore, in examining the mathematical structure of d'Alembertian dynamics, we shall consider a mechanical system $\mathcal{S} = (Q, m, \mathrm{F})$ whose configuration space Q is an arbitrary smooth manifold of $\mathcal{E}$ (and whose force field F *may* include some known dynamical constraint effect).

2.2.1 d'Alembert equation

After d'Alembert, a smooth motion of the particle system in $\mathcal{E}_3$ –i.e. a smooth parametrized curve of $\mathcal{E}$– is a DPM of $\mathcal{S} = (Q, m, \mathrm{F})$, iff it is a solution of the second-order differential equation in implicit form on $\mathcal{E}$ defined by [9]

$$\begin{aligned}\mathrm{D}_{d'Al} &:= \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \mid (\mathrm{p}, \mathrm{v}) \in TQ,\ \mathrm{F}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a} \in T_{\mathrm{p}}^{\perp}Q\} \\ &= \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \mid (\mathrm{p}, \mathrm{v}) \in TQ, (\mathrm{F}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a}) \cdot \delta\mathrm{p} = 0, \forall \delta\mathrm{p} \in T_{\mathrm{p}}Q\}\end{aligned}$$

and called *d'Alembert equation* associated with $\mathcal{S}$.

If F is time-independent, we have $\mathrm{D}_{d'Al} = \mathbb{R} \times \mathcal{D}_{d'Al}$, where

$$\mathcal{D}_{d'Al} := \{(\mathrm{p}, \mathrm{v}, \mathrm{a}) \in T^2\mathcal{E} \mid (\mathrm{p}, \mathrm{v}) \in TQ,\ \ \mathrm{F}(\mathrm{p}, \mathrm{v}) - m\,\mathrm{a} \in T_{\mathrm{p}}^{\perp}Q\}$$

is a time-independent differential equation, which can as well take the place of $D_{d'Al}$. [10]

Clearly, in absence of two-sided constraints (when Q is an open manifold of $\mathcal{E}$ and hence $T_{\mathrm{p}}^{\perp}Q = E^{\perp} = \{0\}$ for all $\mathrm{p} \in Q$), d'Alembert equation *does not* differ from Newton equation.

2.2.2 d'Alembert's law of dynamics

So, a smooth motion $\gamma : t \in I \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t)) \in \mathcal{E}$ is a DPM of $\mathcal{S}$, iff it satisfies *d'Alembert's law of dynamics* [11]

$$(t, \mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathrm{D}_{d'Al}$$

which includes both the *admissibility condition*

$$\mathrm{p}(t) \in Q$$

[9] $T_{\mathrm{p}}^{\perp}Q$ will denote the orthogonal complement of $T_{\mathrm{p}}Q$ in E (see Appendix, section 4.2.1, *Euclidean metric*).

[10] See Appendix, section 4.6.2, *Second-order differential equations*.

[11] "$\forall\, t \in I$" will generally be understood.

(equivalent to $(\mathrm{p}(t), \dot{\mathrm{p}}(t)) \in TQ$) exhibiting the 'kinematical' effect of the constraints, which only allow motions living in the constraint manifold Q, and the historical *d'Alembert's principle of dynamics*

$$\mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) - m\,\ddot{\mathrm{p}}(t) \in T^{\perp}_{\mathrm{p}(t)}Q$$

requiring that, along a DPM, the sum of active and inertial forces – not necessarily vanishing – is to keep 'orthogonal' to Q.

d'Alembert's principle can be given the mechanical formulation, known as *principle of virtual works*, [12]

$$(\mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) - m\,\ddot{\mathrm{p}}(t)) \cdot \delta\mathrm{p} = 0\ , \quad \forall\, \delta\mathrm{p} \in T_{\mathrm{p}(t)}Q$$

that is,

$$\textstyle\sum_{i=1}^{\nu} \Big(\mathrm{F}_i(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) - m_i\,\ddot{\mathrm{p}}_i(t)\Big) \cdot \delta\mathrm{p}_i =\ , \quad \forall\, \delta\mathrm{p} = (\delta\mathrm{p}_1, \ldots, \delta\mathrm{p}_\nu) \in T_{\mathrm{p}(t)}Q$$

requiring that, along a DPM, the virtual wok of the sum of active and inertial forces is to vanish identically.

d'Alembert's principle is completely Newtonian 'in spirit', since it characterizes the DPMs – among the admissibe motions – as those satisfying

$$\Phi(t) := m\,\ddot{\mathrm{p}}(t) - \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \in T^{\perp}_{\mathrm{p}(t)}Q$$

that is,

$$m\,\ddot{\mathrm{p}}(t) = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) + \Phi(t)$$
$$\Phi(t) \in T^{\perp}_{\mathrm{p}(t)}Q$$

The first of the above two conditions is nothing but Newton's principle of dynamics with a right hand side encompassing the possible 'dynamical' effects of the constraints, particularly an 'unknown' *constraint reaction* $\Phi(t)$, which – according to the second condition – is orthogonal to Q (whereas a possible *constraint friction* tangent to Q, being able to be given a 'known' law as a suitable function of virtual velocity, is meant to be embodied in F).

Recall that the value of F at any $(t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times TQ$ uniquely splits into the sum [13]

$$\mathrm{F}(t, \mathrm{p}, \mathrm{v}) = \mathrm{F}_{\mathrm{tang}}(t, \mathrm{p}, \mathrm{v}) + \mathrm{F}_{\mathrm{orth}}(t, \mathrm{p}, \mathrm{v})$$

[12] For any given 'force vector' $\mathrm{f} \in E$ and any 'virtual displacement' $\delta\mathrm{p} \in T_{\mathrm{p}}Q$, the scalar product $\mathrm{f} \cdot \delta\mathrm{p}$ is called the *virtual work* of f along $\delta\mathrm{p}$.

[13] E is the direct sum of $T_{\mathrm{p}}Q$ and $T^{\perp}_{\mathrm{p}}Q$.

with

$$\mathrm{F}_{\mathrm{tang}}(t, \mathrm{p}, \mathrm{v}) \in T_{\mathrm{p}}Q \ , \quad \mathrm{F}_{\mathrm{orth}}(t, \mathrm{p}, \mathrm{v}) \in T_{\mathrm{p}}^{\perp}Q$$

and remark that the vector component $\mathrm{F}_{\mathrm{orth}}$, orthogonal to the constraint manifold Q, is 'dynamically uninfluential', since d'Alembert's condition

$$\mathrm{F}_{\mathrm{tang}}(t, \mathrm{p}, \mathrm{v}) + \mathrm{F}_{\mathrm{orth}}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a} \in T_{\mathrm{p}}^{\perp}Q$$

(appearing in $\mathrm{D}_{d'Al}$) is obviously equivalent to

$$\mathrm{F}_{\mathrm{tang}}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a} \in T_{\mathrm{p}}^{\perp}Q$$

So only the vector component $\mathrm{F}_{\mathrm{tang}}$, tangent to the constraint manifold Q, is to be thought of as the 'effective force' acting on the system.

$D_{d'Al}$ can be shown to be reducible to normal form on Q. [14]

As a consequence, $D_{d'Al}$ –owing to the smoothness of F– is *deterministic*, that is, for any choice of Cauchy data

$$(t_o, (\mathrm{p}_o, \mathrm{v}_o)) \in \mathbb{R} \times TQ$$

there exists a unique maximal solution to Cauchy problem

$$(\mathrm{D}_{d'Al}, t_o, (\mathrm{p}_o, \mathrm{v}_o))$$

(all of the other solutions to the above problem being just restrictions of the maximal one). [15]

So, once that the position p_o and the velocity v_o of the particle system at an initial time t_o are known, d'Alembert equation proves to be *potentially* able to *determine* uniquely the whole time-evolution of the system, since it *predicts* the existence of a *unique* maximal DPM of $\mathcal{S}$ –say $\gamma : t \in I \mapsto \mathrm{p}(t) \in Q$ with $t_o \in I$– satisfying the initial conditions $\mathrm{p}(t_o) = \mathrm{p}_o$ and $\dot{\mathrm{p}}(t_o) = \mathrm{v}_o$.

A maximal DPM $\gamma : I \to Q$ of $\mathcal{S}$ is said to be *complete* in the *future* and/or in the *past*, if I is upperly and/or lowerly unbounded (i.e. $\sup(I) = +\infty$ and/or $\inf(I) = -\infty$), respectively

In any case of boundedness, γ will exhibit a *time-incompleteness*.

[14] See the *Introduction* quoted in Preface (footnote [2]), where the proof of the above result is shown to give rise to a deep investigation on the geometry underlying d'Alembert equation.

[15] See also Chapter 3, footnote [5].

d'Alembert's inequality

You might want to understand how to generalize the above theory to a mechanical system $\mathcal{S}$ admitting 'non-strict' one-sided constraints $g_\alpha \geq 0$ (as well as two-sided constraints $f_\beta = 0$).

In such a case, the set of configurations allowed by the constraints is given by

$$Q := Q^o \cup \partial Q$$

where

$$Q^o := \{\mathrm{p} \in \mathcal{E} \mid g_\alpha(\mathrm{p}) > 0\,, \quad f_\beta(\mathrm{p}) = 0\,, \ \forall\, \alpha, \beta\}$$

is the *interior* of Q and

$$\begin{aligned} \partial Q := \{\mathrm{p} \in \mathcal{E} \mid \quad & g_{\alpha_i}(\mathrm{p}) > 0 \quad \text{for some values of } \alpha \\ & g_{\alpha_j}(\mathrm{p}) = 0 \quad \text{for the remaining values of } \alpha \\ & f_\beta(\mathrm{p}) = 0 \quad \text{for all the values of } \beta\} \end{aligned}$$

is the *boundary* of Q.

Under the usual hypotheses of regularity on the g_α's and f_β's, the 'space' $V_\mathrm{p}Q$ of the virtual displacements starting from any $\mathrm{p} \in Q$ will be defined as follows.

If $\mathrm{p} \in Q^o$, put

$$V_\mathrm{p}Q := T_\mathrm{p}Q^o$$

with

$$T_\mathrm{p}Q^o = \{\delta\mathrm{p} \in E \mid d_\mathrm{p} f_\beta\,(\delta\mathrm{p}) = 0\,, \ \forall\, \beta\}$$

(each 'small' $\delta\mathrm{p} \in T_\mathrm{p}Q$ 'virtually' takes – up to higher order infinitesimals – the point p belonging to Q^o to a point $\mathrm{p} + \delta\mathrm{p}$ still belonging to Q^o).

If $\mathrm{p} \in \partial Q$, put

$$V_\mathrm{p}Q := T_\mathrm{p}\partial Q \cup V^o_\mathrm{p}\partial Q$$

with

$$T_\mathrm{p}\partial Q = \{\delta\mathrm{p} \in E \mid d_\mathrm{p} g_{\alpha_j}\,(\delta\mathrm{p}) = 0\,, \quad d_\mathrm{p} f_\beta\,(\delta\mathrm{p}) = 0\,, \quad \forall\, \alpha_j, \beta\}$$

(each 'small' $\delta\mathrm{p} \in T_\mathrm{p}\partial Q$ 'virtually' takes – up to higher order infinitesimals – the point p belonging to ∂Q to a point $\mathrm{p} + \delta\mathrm{p}$ still belonging to ∂Q) and

$$V^o_\mathrm{p}\partial Q := \{\delta\mathrm{p} \in E \mid d_\mathrm{p} g_{\alpha_j}\,(\delta\mathrm{p}) > 0\,, \quad d_\mathrm{p} f_\beta\,(\delta\mathrm{p}) = 0\,, \quad \forall\, \alpha_j, \beta\}$$

(each 'small' $\delta\mathrm{p} \in T^o_\mathrm{p}\partial Q$ 'virtually' takes – up to higher order infinitesimals – the point p belonging to ∂Q to a point $\mathrm{p} + \delta\mathrm{p}$ belonging to Q^o).

A virtual displacement $\delta \mathrm{p}$ belonging to $T_{\mathrm{p}}Q^o$ or $T_{\mathrm{p}}\partial Q$ is said to be *reversible*, since the opposite $-\delta \mathrm{p}$ is still a virtual displacement (belonging to $T_{\mathrm{p}}Q^o$ or $T_{\mathrm{p}}\partial Q$, respectively).

A virtual displacement $\delta \mathrm{p}$ belonging to $V^o_{\mathrm{p}}\partial Q$ is said to be *irreversible*, since the opposite $-\delta \mathrm{p}$ is not a virtual displacement.

An admissible motion $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E} \subset \mathcal{E}$, will then be considered a DPM of $\mathcal{S}$, iff it is a solution of the *generalized* d'Alembert equation

$$\mathrm{D}_{d'Al} := \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R}\times T^2\mathcal{E} \mid (\mathrm{p}, \mathrm{v}) \in TQ,\ (\mathrm{F}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a})\cdot\delta\mathrm{p} \leq 0, \forall \delta\mathrm{p} \in V_{\mathrm{p}}Q\}$$

that is, iff it satisfies admissibility condition $\mathrm{p}(t) \in Q$ and d'Alembert's principle in the generalized form of an inequality

$$(\mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) - m\,\ddot{\mathrm{p}}(t)) \cdot \delta\mathrm{p} \leq 0\ , \quad \forall\, \delta\mathrm{p} \in V_{\mathrm{p}(t)}Q$$

The latter indeed corresponds to Newton's principle

$$m\,\ddot{\mathrm{p}}(t) = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) + \Phi(t)$$

completed with the phisically natural condition

$$\Phi(t) \cdot \delta\mathrm{p} \geq 0\ , \quad \forall\, \delta\mathrm{p} \in V_{\mathrm{p}(t)}Q$$

according to which the unknown constraint reaction $\Phi(t)$, when non-null, is bound,
(i) at any $\mathrm{p}(t) \in Q^o$, to be 'orthogonal' to Q^o (its angle with each non-null vector $\delta\mathrm{p} \in T_{\mathrm{p}(t)}Q^o$, tangent to Q^o, being required to be $\pi/2$),
(ii) at any $\mathrm{p}(t) \in \partial Q$, to be 'orthogonal' to ∂Q (its angle with each non-null vector $\delta\mathrm{p} \in T_{\mathrm{p}(t)}\partial Q$, tangent to ∂Q, being required to be $\pi/2$) and to 'point towards' Q^o (its angle with each vector $\delta\mathrm{p} \in V^o_{\mathrm{p}(t)}Q$, pointing towards Q^o, being required to be not greater than $\pi/2$).

In the sequel, we shall only deal with d'Alembert's principle in the special form of an equality.

2.2.3 d'Alembert's law of inertia

Owing to the above theory, the inertial motions of $\mathcal{S} = (Q, m, \mathrm{F})$, i.e. the DPMs of $\mathcal{S}^{(o)} = (Q, m, 0)$, are the solutions of the time-independent differential equation

$$\mathbb{R} \times \mathcal{D}^{(o)}_{d'Al} := \mathbb{R} \times \{(\mathrm{p}, \mathrm{v}, \mathrm{a}) \in T^2\mathcal{E} \mid (\mathrm{p}, \mathrm{v}) \in TQ\ ,\ m\,\mathrm{a} \in T^{\perp}_{\mathrm{p}}Q\}$$

associated with $\mathcal{S}^{(o)}$ (remark that $\mathcal{D}^{(o)}_{d'Al}$ *need not* be a universal equation, since it generally depends –and then its solutions will generally depend– on the mass distribution of the system). [16]

So $\gamma : t \in I \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t)) \in \mathcal{E}$ is an inertial motion of $\mathcal{S}$, iff it satisfies *d'Alembert's law of inertia*

$$(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathcal{D}^{(o)}_{d'Al}$$

which includes both the obvious admissibility condition $\mathrm{p}(t) \in Q$ (equivalent to $(\mathrm{p}(t), \dot{\mathrm{p}}(t)) \in TQ$) and the *geodesic law* [17]

$$m\,\ddot{\mathrm{p}}(t) \in T^{\perp}_{\mathrm{p}(t)}Q$$

requiring that, along an inertial motion, the inertial force field - not necessarily vanishing- is to keep 'orthogonal' to Q.

The geodesic law can be given the mechanical formulation, in terms of virtual works,

$$m\,\ddot{\mathrm{p}}(t) \cdot \delta\mathrm{p} = 0\ , \quad \forall\, \delta\mathrm{p} \in T_{\mathrm{p}(t)}Q$$

that is,

$$\textstyle\sum_{i=1}^{\nu}\, m_i\, \ddot{\mathrm{p}}_i(t) \cdot \delta\mathrm{p}_i = 0\ , \quad \forall\, \delta\mathrm{p} = (\delta\mathrm{p}_1, \ldots, \delta\mathrm{p}_\nu) \in T_{\mathrm{p}(t)}Q$$

requiring that, along an inertial motion, the virtual wok of the inertial forces is to vanish identically.

Remark that, also in presence of a non-vanishing active force field, the DPMs *need not* differ from the inertial motions, since

$$D_{d'Al} = \mathbb{R} \times \mathcal{D}^{(o)}_{d'Al}$$

if –and only if (owing to reducibility to normal form)– F is orthogonal to the constraint manifold, i.e.

$$\mathrm{F}_{tang}(t, \mathrm{p}, \mathrm{v}) = 0\ , \quad \forall\, (t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times TQ$$

2.2.4 d'Alembert's law of energy

d'Alembert equation implies a well known 'energy law' along the DPMs of $\mathcal{S}$, which will now be discussed.

[16] See, for instance, section 2.3.5, *Euler's geodesic equation.*

[17] For the term 'geodesic', see the remark at the end of section 3.1.2 and footnote [8].

Let

$$\begin{aligned}\mathsf{K} : TQ \to \mathbb{R} : (\mathrm{p}, \mathrm{v}) \mapsto \mathsf{K}(\mathrm{p}, \mathrm{v}) : &= \frac{1}{2}\, m\, \mathrm{v} \cdot \mathrm{v} \\ &= \frac{1}{2} \sum_{i=1}^{\nu} m_i\, \mathrm{v}_i \cdot \mathrm{v}_i\end{aligned}$$

be the *kinetic energy* of $m \in \mathcal{S}$, and

$$\begin{aligned}\Pi : \mathbb{R} \times TQ \to \mathbb{R} : (t, \mathrm{p}, \mathrm{v}) \mapsto \Pi(t, \mathrm{p}, \mathrm{v}) : &= \mathrm{F}(t, \mathrm{p}, \mathrm{v}) \cdot \mathrm{v} \\ &= \sum_{i=1}^{\nu} \mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) \cdot \mathrm{v}_i\end{aligned}$$

the *power* of $\mathrm{F} \in \mathcal{S}$.

Then let $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in \mathcal{E}$ be a DPM of $\mathcal{S}$, i.e. a solution of d'Alembert equation.

Owing to the admissibility condition, equivalent to $\mathrm{Im}(\dot{\gamma}) \subset TQ$, we can evaluate K along $\dot{\gamma} = (\mathrm{p}, \dot{\mathrm{p}})$, obtaining the time function $\mathsf{K} \circ \dot{\gamma} = \mathsf{K}(\mathrm{p}, \dot{\mathrm{p}})$.
Owing to the principle of virtual works, which implies $m\, \ddot{\mathrm{p}}(t) \cdot \dot{\mathrm{p}}(t) = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \cdot \dot{\mathrm{p}}(t)$, the derivative of the above function is

$$\begin{aligned}\frac{d}{dt}\Big|_t \mathsf{K}(\mathrm{p}, \dot{\mathrm{p}}) &= \frac{d}{dt}\Big|_t (\frac{1}{2}\, m\, \dot{\mathrm{p}} \cdot \dot{\mathrm{p}}) \\ &= m\, \ddot{\mathrm{p}}(t) \cdot \dot{\mathrm{p}}(t) \\ &= \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \cdot \dot{\mathrm{p}}(t) \\ &= \Pi(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\end{aligned}$$

Hence, for all $t_o, t_1 \in I$,

$$\int_{t_o}^{t_1} \frac{d}{dt}\Big|_t \mathsf{K}(\mathrm{p}, \dot{\mathrm{p}})\, dt = \int_{t_o}^{t_1} \Pi(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\, dt$$

where the quadrature at the left hand side is the variation of $\mathsf{K}(\mathrm{p}, \dot{\mathrm{p}})$ between the limits of integration. [18]
That proves the following *balance law*:

Theorem 1 *Along any DPM , the variations of kinetic energy are counterbalanced by the power correspondingly spent by the force field, i.e.*

$$\mathsf{K}(\mathrm{p}(t_1), \dot{\mathrm{p}}(t_1)) - \mathsf{K}(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) = \int_{t_o}^{t_1} \Pi(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\, dt \ , \ \ \forall\, t_o,\, t_1 \in I$$

[18] See Appendix, section 4.6.1, *Quadrature*.

Some consequences of the above law now follow.

Simple is the case of a *dissipative system* $\mathcal{S}$, that is, one acted upon by a force field F –explicitly depending on velocity– whose power is never positive, namely

$$\Pi(t, \mathrm{p}, 0) = 0 \ , \quad \Pi(t, \mathrm{p}, \mathrm{v}) < 0 \ \text{ if } \ \mathrm{v} \neq 0$$

The name 'dissipative' is due to the fact that Theorem 1 implies the following *dissipation law*:

Theorem 2 *Along any DPM (with non-vanishing velocity) of a dissipative system* $\mathcal{S}$*, the kinetic energy decreases, i.e.*

$$\mathsf{K}(\mathrm{p}(t_1), \dot{\mathrm{p}}(t_1)) < \mathsf{K}(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) \ , \quad \forall \, [t_o \, , t_1] \subset I$$

More interesting is the case of a *conservative system* $\mathcal{S}$, that is, one acted upon by a a positional force field F, whose *virtual work is an exact differential*

$$\mathrm{F}(\mathrm{p}) \cdot |_{T_{\mathrm{p}}Q} = - \, d_{\mathrm{p}}\mathsf{V} \ , \quad \forall \, \mathrm{p} \in Q$$

that is to say, for all $\mathrm{p} \in Q$, the linear form

$$\mathrm{F}(\mathrm{p}) \cdot |_{T_{\mathrm{p}}Q} \in T^*_{\mathrm{p}}Q$$

coincides (up to the sign) with the differential

$$d_{\mathrm{p}}\mathsf{V} \in T^*_{\mathrm{p}}Q$$

of a *potential energy*

$$\mathsf{V} : Q \to \mathbb{R}$$

V is meant to be a *smooth function*, in the sense of being the restriction of a C^∞ differentiable, real-valued function $\tilde{\mathsf{V}}$ defined on an open subset of $\mathcal{E}$ containing Q, and its differential at any $\mathrm{p} \in Q$ is defined through $d_{\mathrm{p}}\tilde{\mathsf{V}} \in E^*$ by putting $d_{\mathrm{p}}\mathsf{V} := (d_{\mathrm{p}}\tilde{\mathsf{V}})|_{T_{\mathrm{p}}Q} \in T^*_{\mathrm{p}}Q$.

Along any admissible motion $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$, the above hypothesis of 'exactness' implies (owing to the chain rule) [19]

$$\begin{aligned}\Pi(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) &= \mathrm{F}(\mathrm{p}(t)) \cdot \dot{\mathrm{p}}(t) \\ &= -\, d_{\mathrm{p}(t)} \mathsf{V}\,(\dot{\mathrm{p}}(t)) \\ &= -\, d_{\mathrm{p}(t)} \tilde{\mathsf{V}}\,(\dot{\mathrm{p}}(t)) \\ &= -\left.\frac{d}{dt}\right|_t \tilde{\mathsf{V}}(\mathrm{p}) \\ &= -\left.\frac{d}{dt}\right|_t \mathsf{V}(\mathrm{p})\end{aligned}$$

Then, if γ is a DPM, Theorem 1 reads

$$\mathsf{K}(\mathrm{p}(t_1), \dot{\mathrm{p}}(t_1)) - \mathsf{K}(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) = \mathsf{V}(\mathrm{p}(t_o)) - \mathsf{V}(\mathrm{p}(t_1))$$

that is,

$$\mathsf{K}(\mathrm{p}(t_1), \dot{\mathrm{p}}(t_1)) + \mathsf{V}(\mathrm{p}(t_1)) = \mathsf{K}(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) + \mathsf{V}(\mathrm{p}(t_o))$$

So the name 'conservative' is due to the fact that, for the *mechanical energy*

$$\mathsf{E} : TQ \to \mathbb{R} : (\mathrm{p}, \mathrm{v}) \mapsto \mathsf{E}(\mathrm{p}, \mathrm{v}) := \mathsf{K}(\mathrm{p}, \mathrm{v}) + \mathsf{V}(\mathrm{p})$$

(kinetic *plus* potential energy), the above theorem states the following *conservation law*:

Theorem 3 *Along any DPM of a conservative system $\mathcal{S}$, the mechanical energy keeps constant, i.e.*

$$E(\mathrm{p}(t_1), \dot{\mathrm{p}}(t_1)) = E(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) \;, \quad \forall\, t_o, t_1 \in I$$

In mathematical terms, the above conservation law just states that E is a first integral of d'Alembert equation.

2.2.5 d'Alembert's law of statics

Recall that a configuration $\mathrm{p}_* \in \mathcal{E}$ is said to be an *equilibrium configuration* of $\mathcal{S} = (Q, m, \mathrm{F})$, if

$$\gamma_* : \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}_*(t) := \mathrm{p}_*$$

(i.e. the state of rest at p_*) is a DPM of $\mathcal{S}$.

[19] See Appendix, section 4.4.1, *Chain rule*.

Then, owing to d'Alembert's law of dynamics, a configuration $\mathrm{p}_* \in Q$ is an equilibrium configuration, iff

$$\mathrm{p}_*(t) \in Q, \quad \mathrm{F}(t, \mathrm{p}_*(t), \dot{\mathrm{p}}_*(t)) - m\, \ddot{\mathrm{p}}_*(t) \in T^{\perp}_{\mathrm{p}_*(t)} Q$$

(with $\mathrm{p}_*(t) = \mathrm{p}_*$ and $\dot{\mathrm{p}}_*(t) = \ddot{\mathrm{p}}_*(t) = 0$), that is, iff it satisfies *d'Alembert's law of statics*

$$\mathrm{p}_* \in Q, \quad \mathrm{F}(t, \mathrm{p}_*, 0) \in T^{\perp}_{\mathrm{p}_*} Q$$

As a consequence, the problem of extracting from Q the equilibrium configurations, can be solved by means of the following *d'Alembert's principle of statics*:

Theorem 4 *An admissible configuration* $\mathrm{p}_* \in Q$ *is an equilibrium configuration, iff* F *keeps orthogonal to* Q *at* p_**, namely*

$$\mathrm{F}(t, \mathrm{p}_*, 0) \in T^{\perp}_{\mathrm{p}_*} Q$$

In terms of virtual work, d'Alembert's law of statics states that $\mathrm{p}_* \in Q$ is an equilibrium configuration of $\mathcal{S}$, iff the virtual work of F vanishes at p_*, namely

$$\mathrm{F}(t, \mathrm{p}_*, 0) \cdot \delta\mathrm{p} = 0 \;, \quad \forall\, \delta\mathrm{p} \in T_{\mathrm{p}_*} Q$$

that is,

$$\textstyle\sum_{i=1}^{\nu} \mathrm{F}_i(t, \mathrm{p}_*, 0) \cdot \delta\mathrm{p}_i = 0 \;, \quad \forall\, \delta\mathrm{p} = (\delta\mathrm{p}_1, \ldots, \delta\mathrm{p}_\nu) \in T_{\mathrm{p}_*} Q$$

Remark that, for a conservative force field, the above law reads

$$\mathrm{F}(\mathrm{p}_*) \cdot |_{T_{\mathrm{p}_*} Q} = 0$$

with

$$\mathrm{F}(\mathrm{p}_*) \cdot |_{T_{\mathrm{p}_*} Q} = -\, d_{\mathrm{p}_*} \mathsf{V}$$

So, in the conservative case, d'Alembert's law of statics takes the following expression:

Theorem 5 *An admissible configuration* $\mathrm{p}_* \in Q$ *is an equilibrium configuration of a conservative system* $\mathcal{S}$ *, iff it is a* singular point *of potential energy* V*, i.e.*

$$d_{\mathrm{p}_*} \mathsf{V} = 0$$

Stability

Clearly, if p_* is an equilibrium configuration, the state of rest γ_* at p_* is the maximal DPM determined by the initial conditions $(\mathrm{p}_*, 0)$, for any 'initial time' t_o.

In qualitative terms, the equilibrium at p_* is said to be 'stable', if a 'suitably small' perturbation of the above initial conditions $(\mathrm{p}_*, 0)$ at any given time t_o determines an 'arbitrarily small' motion about p_* (namely, a maximal DPM, complete in the future, which – after t_o – differs from the state of rest at p_* as little as is required).

In mathematical terms, an equilibrium configuration $\mathrm{p}_* \in Q$ of $\mathcal{S}$ is then said to be *stable*, if , for any two (arbitrarily small) positive numbers δ and ϵ, there exist two (suitably small) positive numbers $\delta_o < \delta$ and $\epsilon_o < \epsilon$ s.t. any choice, at a given time $t_o \in \mathbb{R}$, of initial conditions $\mathrm{p}_o \in Q$ and $\mathrm{v}_o \in T_{\mathrm{p}_o}Q$ satisfying

$$\mathrm{p}_o \in \mathcal{B}^{\delta_o}_{\mathrm{p}_*}\ ,\quad \mathsf{K}(\mathrm{p}_o, \mathrm{v}_o) < \epsilon_o$$

determines a maximal DPM $\gamma : t \in I \mapsto \mathrm{p}(t) \in Q \subset \mathcal{E}$, with $t_o \in I$, satisfying

$$\sup(I) = +\infty$$

and, for all $t > t_o$,

$$\mathrm{p}(t) \in \mathcal{B}^{\delta}_{\mathrm{p}_*}\ ,\quad \mathsf{K}((\mathrm{p}(t), \dot{\mathrm{p}}(t)) < \epsilon$$

Criterions (i.e. sufficient conditions) for the stability or instability of the 'static' solutions of differential equations, are the typical results of 'stability theory'. We shall state (without proof) the following:

Dirichlet's stability criterion If the potential energy V of a conservative system $\mathcal{S}$ admits a strict local minimum at a point $\mathrm{p}_* \in Q$, then such a point is a stable equilibrium configuration

[Note: adding a dissipative force field would not destroy the stability].

Chetaev's instability criterion If the potential energy V of a conservative system $\mathcal{S}$ does not admit a strict local minimum at an equilibrium configuration $\mathrm{p}_* \in Q$ and the absence of minimum can be detected, in a chart ξ of Q with $\mathrm{Im}(\xi) \ni \mathrm{p}_*$, by examining the higher-order partial derivatives of $\mathsf{V} \circ \xi$ at $\xi^{-1}(\mathrm{p}_*)$, then the equilibrium at p_* is unstable.

2.3 Euler

d'Alembert equation will now be referred to a particle system subject to the constraint of rigidity and will be specialized in the classical Newton and Euler's equations of 'rigid body dynamics'.

2.3.1 Rigid body

We shall deal with a 3-dimensional *rigid body*, conceived as an ordered ν-tuple of distinct points

$$\tilde{\mathrm{p}} = (\tilde{\mathrm{p}}_1, \ldots, \tilde{\mathrm{p}}_\nu) \subset \tilde{\mathcal{E}}_3$$

of a *body space* $\tilde{\mathcal{E}}_3$ (3-dimensional, oriented, Euclidean affine space), [20] and assumed to fulfil the '3-dimensionality' condition

$$\mathrm{Span}\{(\tilde{\mathrm{p}}_i - \tilde{\mathrm{p}}_j)_{i,j=1,\ldots,\nu}\} = \tilde{E}_3$$

($\tilde{E}_3$ being the vector space on which $\tilde{\mathcal{E}}_3$ is modelled).

The *center of mass* of the rigid body $\tilde{\mathrm{p}} = (\tilde{\mathrm{p}}_1, \ldots, \tilde{\mathrm{p}}_\nu)$, carrying a *mass-distribution* $m = (m_1, \ldots, m_\nu)$, is the unique point $\tilde{\mathrm{c}} \in \tilde{\mathcal{E}}_3$ satisfying [21]

$$\textstyle\sum_{i=1}^{\nu} m_i(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}}) = 0$$

(i) In a given *reference space* $\mathcal{E}_3$ (modelled on E_3), an *admissible configuration*

$$\mathrm{p} = (\mathrm{p}_1, \ldots, \mathrm{p}_\nu) \subset \mathcal{E}_3$$

of the rigid body −i.e. a configuration allowed by the constraint of rigidity− will be obtained from $\tilde{\mathrm{p}}$ through an orientation-preserving affine isometry

$$\alpha : \tilde{\mathcal{E}}_3 \to \mathcal{E}_3$$

that is, [22]

$$\mathrm{p} = \alpha \circ \tilde{\mathrm{p}}$$

[20] Recall our initial remark (in section 1.1) on a 3-dimensional, oriented, Euclidean affine space meant as a 'mathematical extension' of a rigid body.

[21] Check that the centre of mass is

$$\tilde{\mathrm{c}} = \tilde{\mathrm{o}} + \frac{1}{m_o} \textstyle\sum_{i=1}^{\nu} m_i(\tilde{\mathrm{p}}_i - \tilde{\mathrm{o}})\ , \quad m_o := \sum_{i=1}^{\nu} m_i$$

(the choice of $\tilde{\mathrm{o}} \in \tilde{\mathcal{E}}_3$ is immaterial).

[22] Recall that any ordered ν-tuple, like $\tilde{\mathrm{p}} = (\tilde{\mathrm{p}}_1, \ldots, \tilde{\mathrm{p}}_\nu) \subset \tilde{\mathcal{E}}_3$, can be regarded as a mapping $\tilde{\mathrm{p}} : i \in \{1, \ldots, \nu\} \mapsto \tilde{\mathrm{p}}_i \in \mathcal{E}_3$.

which means [23]

$$p_i = \alpha(\tilde{p}_i)$$

that is,

$$p_i = c + A\,(\tilde{p}_i - \tilde{c}) \qquad (\star)$$

where

$$c := \alpha(\tilde{c}) \in \mathcal{E}_3$$

satisfying

$$\sum_{i=1}^{\nu} m_i(p_i - c) = 0$$

is the *position of the centre of mass* in the configuration p of the rigid body in $\mathcal{E}_3$, [24] and

$$A : \tilde{E}_3 \to E_3$$

is the linear part of α and then an orientation-preserving linear isometry.

A geometric insight on the *configuation space*

$$Q \subset \mathcal{E} := \mathcal{E}_3^{\nu}$$

of the rigid body (set of all the admissible configurations of $\tilde{p}$ in $\mathcal{E}_3$) now follows.

First consider the set $\Im_a(\tilde{\mathcal{E}}_3, \mathcal{E}_3)$ of all the orientation-preserving affine isometries of $\tilde{\mathcal{E}}_3$ onto $\mathcal{E}_3$ and the set $\Im_l(\tilde{E}_3, E_3)$ of all the orientation-preserving linear isometries of $\tilde{E}_3$ onto E_3. As we have seen above, we have an identification

$$\mathcal{E}_3 \times \Im_l(\tilde{E}_3, E_3) = \Im_a(\tilde{\mathcal{E}}_3, \mathcal{E}_3)$$

determined by the centre of mass $\tilde{c} \in \tilde{\mathcal{E}}_3$ through the splitting $(\star)$. Moreover, if $\tilde{\mathcal{E}}_3$ and $\mathcal{E}_3$ are identified with $\mathbb{R}^3$ by means of orthogonal Cartesian systems, we have a further identification

$$\mathbb{R}^3 \times SO(3) = \mathcal{E}_3 \times \Im_l(\tilde{E}_3, E_3)$$

$\mathbb{R}^3 \times SO(3)$ being the group of all the orientation-preserving affine isometries of $\mathbb{R}^3$, [25] which can be shown –through the Implicit Function Theorem– to be a 6-dimensional smooth manifold (embedded in $\mathbb{R}^{12}$).

---

[23] "$\forall\, i = 1, \dots, \nu$" will be understood.

[24] The expression of c is the same as the one given in footnote [21] for $\tilde{c}$ (just remove $\sim$).

[25] Recall that the subgroup $SO(3) \subset GL(3, \mathbb{R})$ of special orthogonal matrixes corresponds to the group of all the orientation-preserving linear isometries of $\mathbb{R}^3$.

Now, owing to the 3-dimensionality of $\tilde{\mathrm{p}}$, the surjective mapping

$$\alpha \in \Im_a(\tilde{\mathcal{E}}_3, \mathcal{E}_3) \longmapsto \mathrm{p} = \alpha \circ \tilde{\mathrm{p}} \in Q$$

can easily be shown to be injective too. As a consequence, through the above bijection, Q inherits the charts of $\mathbb{R}^3 \times SO(3)$ and, through those charts, exhibits the structure of a 6-dimensional smooth manifold (embedded in $\mathcal{E}$).

(ii) Consider now an *admissible motion* of the rigid body in the reference space, say $\gamma : t \in I \subset \mathbb{R} \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t)) \in Q \subset \mathcal{E}$.

The admissibility condition [26]

$$\mathrm{p}(t) = \alpha(t) \circ \tilde{\mathrm{p}}$$

allows γ to be described in terms of a *rigid motion* of $\tilde{\mathcal{E}}_3$ in $\mathcal{E}_3$, i.e.

$$\alpha(t) \in \Im_a(\tilde{\mathcal{E}}_3, \mathcal{E}_3)$$

Hence the splitting

$$\mathrm{p}_i(t) = \mathrm{c}(t) + A(t)\,(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})$$

or, with an arbitrary choice of a point $\mathrm{o} \in \mathcal{E}_3$,

$$\mathrm{p}_i(t) = (\mathrm{o} + A(t)\,(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) + (\mathrm{c}(t) - \mathrm{o})$$

where

$$\mathrm{c}(t) \in \mathcal{E}_3$$

(value of $\alpha(t)$ at $\tilde{\mathrm{c}}$) and

$$A(t) \in \Im_l(\tilde{E}_3, E_3)$$

(linear part of $\alpha(t)$) describe −along the above rigid motion− the motion *of* $\tilde{\mathrm{c}}$ and the motion *around* $\tilde{\mathrm{c}}$ (as if $\tilde{\mathrm{c}}$ were stationary at $\mathrm{o} \in \mathcal{E}_3$), respectively.

The derivative $\dot{\mathrm{p}}(t) = (\dot{\mathrm{p}}_1(t), \ldots, \dot{\mathrm{p}}_\nu(t)) \in T_{\mathrm{p}(t)}Q$ is then given by [27]

---

[26] "$\forall\, t \in I$" will generally be understood.

[27] Recall that the *wedge product* $\wedge : (\mathrm{u}, \mathrm{v}) \in E_3 \times E_3 \mapsto \mathrm{u} \wedge \mathrm{v} \in E_3$ is the bilinear mapping defined as follows:
(i) if (u, v) is a linearly dependent system, then $\mathrm{u} \wedge \mathrm{v} = 0$;
(ii) if (u, v) is a linearly independent system, then

$$|\mathrm{u} \wedge \mathrm{v}| := |\mathrm{u}||\mathrm{v}| \sin \angle(\mathrm{u}, \mathrm{v}) \neq 0\,, \;\; \mathrm{u} \wedge \mathrm{v} \in \big(\mathrm{Span}\,(\mathrm{u}, \mathrm{v})\big)^{\perp}, \;\; (\mathrm{u}, \mathrm{v}, \mathrm{u} \wedge \mathrm{v}) \in \mathrm{Or}(E_3)$$

$\mathrm{Or}(E_3)$ being the orientation given to E_3.

$$\begin{aligned}\dot{\mathrm{p}}_i(t) &= \dot{c}(t) + \dot{A}(t)\,(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}}) = \dot{c}(t) + (\dot{A}(t) \circ A(t)^{-1})\,(\mathrm{p}_i(t) - \mathrm{c}(t)) \\ &= \dot{c}(t) + \omega(t) \wedge (\mathrm{p}_i(t) - \mathrm{c}(t))\end{aligned}$$

where $\dot{c}(t) \in E_3$ is the velocity of the centre of mass and $\omega(t) \in E_3$ is said to be the *angular velocity* of the motion around the centre of mass.

As a consequence, for any $\mathrm{p} \in Q$ (with centre of mass c), a vector $\mathrm{v} = (\mathrm{v}_1, \ldots, \mathrm{v}_\nu) \in T_\mathrm{p}Q$ (tangent to an admissible motion passing through p) can be expressed in the form

$$\mathrm{v}_i = \mathrm{v}_\mathrm{c} + \mathrm{w} \wedge (\mathrm{p}_i - \mathrm{c}) \qquad (\star\star)$$

with $(\mathrm{v}_\mathrm{c}, \mathrm{w}) \in E_3 \times E_3$.

That amounts to saying $T_\mathrm{p}Q \subset V_\mathrm{p}$ –where V_p denotes the 6-dimensional image of the injective linear mapping

$$(\mathrm{v}_\mathrm{c}, \mathrm{w}) \in E_3 \times E_3 \longmapsto \mathrm{v} = (\mathrm{v}_1, \ldots, \mathrm{v}_\nu) \in E := E_3^\nu$$

defined by $(\star\star)$– and then, for dimensional reasons, $T_\mathrm{p}Q = V_\mathrm{p}$.

So the above mapping $(\star\star)$ determines a canonical isomorphism

$$E_3 \times E_3 \to T_\mathrm{p}Q$$

2.3.2 Cardinal equations

We shall now consider a *rigid system*

$$\mathcal{R} = (Q, m, \mathrm{F})$$

i.e. a mechanical system modelling (in a reference space $\mathcal{E}_3$) a 3-dimensional rigid body acted upon by a force field F.

In an orthonormal basis $(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3) \in \mathrm{Or}(E_3)$, the components of $\mathrm{w} = \mathrm{u} \wedge \mathrm{v}$ are given (for any $i = 1, 2, 3$) by

$$w^i = u^{i+1} v^{i+2} - v^{i+1} u^{i+2}$$

(where $i+1$ and $i+2$ are obtained from i by circular permutations of $(1, 2, 3)$).

Also recall that there exists one and only one vector $\omega(t) \in E_3$ s.t.

$$\omega(t) \wedge = \dot{A}(t) \circ A(t)^{-1}$$

where the derivative $\dot{A}(t) : \tilde{E}_3 \to E_3$ of $A(t) \in \Im_l(\tilde{E}_3, E_3)$ can be defined through the previously mentioned identification of $\Im_l(\tilde{E}_3, E_3)$ with $SO(3)$.

The DPMs of $\mathcal{R}$ are the solutions of d'Alembert differential equation

$$\begin{aligned}\mathrm{D}_{d'Al} := \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \quad | \quad & (\mathrm{p}, \mathrm{v}) \in TQ\,, \\ & (\mathrm{F}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a}) \cdot \delta\mathrm{p} = 0\,,\ \forall\,\delta\mathrm{p} \in T_{\mathrm{p}}Q\}\end{aligned}$$

which, in the present case, can be expressed as follows.

Focus on 'd'Alembert's condition'

$$(\mathrm{F}(t, \mathrm{p}, \mathrm{v}) - m\,\mathrm{a}) \cdot \delta\mathrm{p} = 0\,,\ \forall\,\delta\mathrm{p} \in T_{\mathrm{p}}Q$$

(with $(\mathrm{p}, \mathrm{v}) \in TQ$), i.e.

$$\textstyle\sum_{i=1}^{\nu} (\mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - m_i\,\mathrm{a}_i) \cdot \delta\mathrm{p}_i\ ,\quad \forall\,\delta\mathrm{p} = (\delta\mathrm{p}_1, \ldots, \delta\mathrm{p}_\nu) \in T_{\mathrm{p}}Q$$

Recall that any $\delta\mathrm{p} \in T_{\mathrm{p}}Q$ is the image of a unique $(\delta\mathrm{c}, \varphi) \in E_3 \times E_3$ through canonical isomorphism ($\star\star$), i.e.

$$\delta\mathrm{p}_i = \delta\mathrm{c} + \varphi \wedge (\mathrm{p}_i - \mathrm{c})$$

(where c is the center of mass at configuration p).

As a consequence, d'Alembert's condition reads

$$\begin{aligned}&\textstyle\sum_{i=1}^{\nu} (\mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - m_i\,\mathrm{a}_i) \cdot \delta\mathrm{c}\, + \\ &\textstyle\sum_{i=1}^{\nu} (\mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - m_i\,\mathrm{a}_i) \cdot (\varphi \wedge (\mathrm{p}_i - \mathrm{c})) = 0\ ,\quad \forall\,(\delta\mathrm{c}, \varphi) \in E_3 \times E_3\end{aligned}$$

that is, [28]

$$\begin{aligned}&\Big(\textstyle\sum_{i=1}^{\nu} \mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - m_i\,\mathrm{a}_i\Big) \cdot \delta\mathrm{c}\, + \\ &\Big(\textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge (\mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - m_i\,\mathrm{a}_i)\Big) \cdot \varphi = 0\ ,\quad \forall\,(\delta\mathrm{c}, \varphi) \in E_3 \times E_3\end{aligned}$$

and then it is equivalent to the couple of conditions

$$\textstyle\sum_{i=1}^{\nu} \mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - \sum_{i=1}^{\nu} m_i\mathrm{a}_i = 0$$

$$\textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) - \sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i\mathrm{a}_i = 0$$

If we put

$$\begin{aligned}\mathrm{R}^{(\mathrm{F})} \ &:\ \mathbb{R} \times TQ \to E_3 \\ &:\ (t, \mathrm{p}, \mathrm{v}) \longmapsto \mathrm{R}^{(\mathrm{F})}(t, \mathrm{p}, \mathrm{v}) := \textstyle\sum_{i=1}^{\nu} \mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) \\ \mathrm{T}^{(\mathrm{F})} \ &:\ \mathbb{R} \times TQ \to E_3 \\ &:\ (t, \mathrm{p}, \mathrm{v}) \longmapsto \mathrm{T}^{(\mathrm{F})}(t, \mathrm{p}, \mathrm{v}) := \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_i(t, \mathrm{p}, \mathrm{v})\end{aligned}$$

---

[28] Recall that, for all $\mathrm{u}_1, \mathrm{u}_2, \mathrm{u}_3 \in E_3$, $\mathrm{u}_1 \cdot (\mathrm{u}_2 \wedge \mathrm{u}_3) = \mathrm{u}_2 \cdot (\mathrm{u}_3 \wedge \mathrm{u}_1) = \mathrm{u}_3 \cdot (\mathrm{u}_1 \wedge \mathrm{u}_2)$.

(*resultant* and *torque* −relative to the centre of mass− of $\mathrm{F}(t, \mathrm{p}, \mathrm{v})$) and

$$\begin{aligned} \mathrm{R}^{(m)} \;&:\; TQ \times E \to E_3 \\ &:\; (\mathrm{p}, \mathrm{v}, \mathrm{a}) \longmapsto \mathrm{R}^{(m)}(\mathrm{p}, \mathrm{v}, \mathrm{a}) := -\textstyle\sum_{i=1}^{\nu} m_i \mathrm{a}_i \\ \mathrm{T}^{(m)} \;&:\; TQ \times E \to E_3 \\ &:\; (\mathrm{p}, \mathrm{v}, \mathrm{a}) \longmapsto \mathrm{T}^{(m)}(\mathrm{p}, \mathrm{v}, \mathrm{a}) := -\textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \mathrm{a}_i \end{aligned}$$

(*resultant* and *torque* −relative to the centre of mass− of $\mathrm{F}^{(m)}(\mathrm{p}, \mathrm{v}, \mathrm{a}) := -m\mathrm{a}$), we can state the following result:

Theorem 6 *For the rigid system* $\mathcal{R}$ *, d'Alembert differential equation takes the expression*

$$\begin{aligned} \mathrm{D}_{d'Al} := \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \quad | \quad & (\mathrm{p}, \mathrm{v}) \in TQ\,, \\ & -\mathrm{R}^{(m)}(\mathrm{p}, \mathrm{v}, \mathrm{a}) = \mathrm{R}^{(\mathrm{F})}(t, \mathrm{p}, \mathrm{v}) \\ & -\mathrm{T}^{(m)}(\mathrm{p}, \mathrm{v}, \mathrm{a}) = \mathrm{T}^{(\mathrm{F})}(t, \mathrm{p}, \mathrm{v})\} \end{aligned}$$

The solutions of $\mathrm{D}_{d'Al}$ are then the admissible motions

$$\gamma : I \subset \mathbb{R} \to Q \subset \mathcal{E} : t \mapsto \mathrm{p} = \mathrm{p}(t)$$

satisfying

$$-\mathrm{R}^{(m)}(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) = \mathrm{R}^{(\mathrm{F})}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$$

$$-\mathrm{T}^{(m)}(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) = \mathrm{T}^{(\mathrm{F})}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$$

So we have obtained the following: [29]

Theorem 7 *The DPMs of* $\mathcal{R}$ *are the admissible motions* $\gamma : I \to Q$ *satisfying the* cardinal equations

$$-\mathrm{R}^{(m)} \circ \ddot{\gamma} = \mathrm{R}^{(\mathrm{F})} \circ (\tau, \dot{\gamma})$$

$$-\mathrm{T}^{(m)} \circ \ddot{\gamma} = \mathrm{T}^{(\mathrm{F})} \circ (\tau, \dot{\gamma})$$

which require that, along a DPM, the resultant and the torque of the inertial force field are to counterbalance the resultant and the torque, respectively, of the active force field.

[29] $\tau : t \mapsto \tau(t) := t$ will denote the identity mapping of any open interval of $\mathbb{R}$.

Then the geodesic law takes the form

$$\mathrm{R}^{(m)} \circ \ddot{\gamma} = 0$$
$$\mathrm{T}^{(m)} \circ \ddot{\gamma} = 0$$

requiring that, along an inertial motion of $\mathcal{R}$, the resultant and the torque of the inertial force field are to vanish.

Finally the law of statics takes the form

$$\mathrm{R}^{(\mathrm{F})} \circ (\tau, \dot{\gamma}_*) = 0$$
$$\mathrm{T}^{(\mathrm{F})} \circ (\tau, \dot{\gamma}_*) = 0$$

(where $\gamma_* : t \in \mathbb{R} \mapsto \mathrm{p}_*(t) := \mathrm{p}_* \in Q$), stating that a point $\mathrm{p}_* \in Q$ is an equilibrium configuration of $\mathcal{R}$, iff the resultant and the torque of the active force field vanish at p_*.

Remark At any $(t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times TQ$, the value of the force field F_i active on the i-th particle of the rigid body can be thought of as splitting into the sum

$$\mathrm{F}_i(t, \mathrm{p}, \mathrm{v}) = \mathrm{F}_i^{(\mathrm{int})}(\mathrm{p}) + \mathrm{F}_i^{(\mathrm{ext})}(t, \mathrm{p}, \mathrm{v})$$

where

$$\mathrm{F}_i^{(\mathrm{int})}(\mathrm{p}) = \textstyle\sum_{j=1}^{\nu} \mathrm{F}_{ij}(\mathrm{p})$$

with

$$\mathrm{F}_{ij}(\mathrm{p}) = -\mathrm{F}_{ji}(\mathrm{p}) \ , \quad (\mathrm{p}_i - \mathrm{p}_j) \wedge \mathrm{F}_{ij}(\mathrm{p}) = 0$$

is due to the possible *interactions* $\{\mathrm{F}_{ij}(\mathrm{p})\}_{j=1,\ldots,\nu}$ −obeying the principle of action and reaction− between the i-th particle and the other particles of the rigid body, whereas $\mathrm{F}_i^{(\mathrm{ext})}(t, \mathrm{p}, \mathrm{v})$ is due to *external* influences (if any).

Remark that , owing to the principle of action and reaction, the resultant and the torque of the interactions vanish, since

$$\begin{aligned}
\textstyle\sum_{i=1}^{\nu} \mathrm{F}_i^{(\mathrm{int})}(\mathrm{p}) &= \textstyle\sum_{i,j=1}^{\nu} \mathrm{F}_{ij}(\mathrm{p}) \\
&= \textstyle\sum_{i<j=1}^{\nu} \mathrm{F}_{ij}(\mathrm{p}) + \mathrm{F}_{ji}(\mathrm{p}) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_i^{(\mathrm{int})}(\mathrm{p}) &= \textstyle\sum_{i,j=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_{ij}(\mathrm{p}) \\
&= \textstyle\sum_{i<j=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_{ij}(\mathrm{p}) + (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_{ji}(\mathrm{p}) \\
&= \textstyle\sum_{i<j=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge \mathrm{F}_{ij}(\mathrm{p}) + (\mathrm{c} - \mathrm{p}_j) \wedge \mathrm{F}_{ij}(\mathrm{p}) \\
&= \textstyle\sum_{i<j=1}^{\nu} (\mathrm{p}_i - \mathrm{p}_j) \wedge \mathrm{F}_{ij}(\mathrm{p}) \\
&= 0
\end{aligned}$$

As a consequence, the interactions turn out to be dynamically uninfluential for a rigid system $\mathcal{R}$, since no trace of them is left in the cardinal equations (characterizing, among the admissible motions, the DPMs of $\mathcal{R}$).

2.3.3 Balance equations

The left hand sides of the cardinal equations will now be shown to be 'exact derivatives'.

Indeed, along an admissible motion $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$, we have

$$\begin{aligned} -\mathrm{R}^{(m)} \circ \ddot{\gamma} &= -\mathrm{R}^{(m)}(\mathrm{p}, \dot{\mathrm{p}}, \ddot{\mathrm{p}}) \\ &= \textstyle\sum_{i=1}^{\nu} m_i \ddot{\mathrm{p}}_i \\ &= \frac{d}{dt} \textstyle\sum_{i=1}^{\nu} m_i \dot{\mathrm{p}}_i \end{aligned}$$

and

$$\begin{aligned} -\mathrm{T}^{(m)} \circ \ddot{\gamma} &= -\mathrm{T}^{(m)}(\mathrm{p}, \dot{\mathrm{p}}, \ddot{\mathrm{p}}) \\ &= \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \ddot{\mathrm{p}}_i \\ &= \frac{d}{dt} \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \dot{\mathrm{p}}_i - \textstyle\sum_{i=1}^{\nu} (\dot{\mathrm{p}}_i - \dot{\mathrm{c}}) \wedge m_i \dot{\mathrm{p}}_i \\ &= \frac{d}{dt} \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \dot{\mathrm{p}}_i \end{aligned}$$

since [30]

$$\begin{aligned} -\textstyle\sum_{i=1}^{\nu} (\dot{\mathrm{p}}_i - \dot{\mathrm{c}}) \wedge m_i \dot{\mathrm{p}}_i &= -\textstyle\sum_{i=1}^{\nu} \dot{\mathrm{p}}_i \wedge m_i \dot{\mathrm{p}}_i + \dot{\mathrm{c}} \wedge \textstyle\sum_{i=1}^{\nu} m_i \dot{\mathrm{p}}_i \\ &= \dot{\mathrm{c}} \wedge \textstyle\sum_{i=1}^{\nu} m_i (\dot{\mathrm{c}} + \omega \wedge (\mathrm{p}_i - \mathrm{c})) \\ &= \dot{\mathrm{c}} \wedge \Big(\big(\textstyle\sum_{i=1}^{\nu} m_i \big) \dot{\mathrm{c}} + \omega \wedge \big(\textstyle\sum_{i=1}^{\nu} m_i (\mathrm{p}_i - \mathrm{c}) \big) \Big) \\ &= \dot{\mathrm{c}} \wedge \big(\textstyle\sum_{i=1}^{\nu} m_i \big) \dot{\mathrm{c}} \\ &= 0 \end{aligned}$$

If we put

$$\Lambda : TQ \to E_3 : (\mathrm{p}, \mathrm{v}) \mapsto \Lambda(\mathrm{p}, \mathrm{v}) := \textstyle\sum_{i=1}^{\nu} m_i \mathrm{v}_i$$

(*linear momentum* of $m \in \mathcal{R}$) and

$$\Omega : TQ \to E_3 : (\mathrm{p}, \mathrm{v}) \mapsto \Omega(\mathrm{p}, \mathrm{v}) := \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \mathrm{v}_i$$

(*angular momentum* of $m \in \mathcal{R}$) the above results read

$$-\mathrm{R}^{(m)} \circ \ddot{\gamma} = \frac{d}{dt} (\Lambda \circ \dot{\gamma})$$

and

$$-\mathrm{T}^{(m)} \circ \ddot{\gamma} = \frac{d}{dt} (\Omega \circ \dot{\gamma})$$

Then, as to the dynamics of $\mathcal{R}$, we can rephrase Theorem 7 as follows:

[30] Recall that $\dot{\mathrm{p}}_i = \dot{\mathrm{c}} + \omega \wedge (\mathrm{p}_i - \mathrm{c})$ and $\sum_{i=1}^{\nu} m_i (\mathrm{p}_i - \mathrm{c}) = 0$. Also recall that the wedge product of a linearly dependent couple of vectors, vanishes.

Theorem 8 *The DPMs of $\mathcal{R}$ are the admissible motions $\gamma : I \to Q$ satisfying the* balance equations

$$\frac{d}{dt}\,(\Lambda \circ \dot{\gamma}) = \mathrm{R}^{(\mathrm{F})} \circ (\tau, \dot{\gamma})$$
$$\frac{d}{dt}\,(\Omega \circ \dot{\gamma}) = \mathrm{T}^{(\mathrm{F})} \circ (\tau, \dot{\gamma})$$

which require that, along a DPM, the variations of linear and agular momentum are to be balanced by the resultant and the torque, respectively, of the active force field.

2.3.4 Newton and Euler's equations

The balance equations, whose unknown is an admissible motion of $\mathcal{R}$, will now be expressed in terms of new unknowns – motion of the centre of mass and motion around the centre of mass – characterizing the old one.

To this end, we have to re-examine the quantities appearing in the balance equations.

As to Λ, we have – for any $(\mathrm{p}, \mathrm{v}) \in TQ$, characterized by $(\mathrm{c}, A, \mathrm{v}_\mathrm{c}, \mathrm{w})$ – [31]

$$\begin{aligned}\Lambda(\mathrm{p}, \mathrm{v}) &= \textstyle\sum_{i=1}^{\nu} m_i \mathrm{v}_i \\ &= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{v}_\mathrm{c} + \mathrm{w} \wedge (\mathrm{p}_i - \mathrm{c})) \\ &= (\textstyle\sum_{i=1}^{\nu} m_i)\mathrm{v}_\mathrm{c} + \mathrm{w} \wedge (\textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i - \mathrm{c})) \\ &= m_o \mathrm{v}_\mathrm{c}\end{aligned}$$

where

$$m_o := \textstyle\sum_{i=1}^{\nu} m_i$$

is the *total mass* of $\mathcal{R}$.

So, along an admissible motion $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$, whose tangent lifting $\dot{\gamma} = (\mathrm{p}, \dot{\mathrm{p}})$ is characterized by the time functions $(\mathrm{c}, A, \dot{\mathrm{c}}, \omega)$, we obtain

$$\begin{aligned}\Lambda \circ \dot{\gamma} &= \Lambda(\mathrm{p}, \dot{\mathrm{p}}) \\ &= m_o\, \dot{\mathrm{c}}\end{aligned}$$

As to Ω, we have – for any $(\mathrm{p}, \mathrm{v}) \in TQ$, characterized by $(\mathrm{c}, A, \mathrm{v}_\mathrm{c}, \mathrm{w})$ –

$$\begin{aligned}\Omega(\mathrm{p}, \mathrm{v}) &= \textstyle\sum_{i=1}^{\nu} (\mathrm{p}_i - \mathrm{c}) \wedge m_i \mathrm{v}_i \\ &= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i - \mathrm{c}) \wedge (\mathrm{v}_\mathrm{c} + \mathrm{w} \wedge (\mathrm{p}_i - \mathrm{c})) \\ &= (\textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i - \mathrm{c})) \wedge \mathrm{v}_\mathrm{c} + \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i - \mathrm{c}) \wedge (\mathrm{w} \wedge (\mathrm{p}_i - \mathrm{c})) \\ &= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i - \mathrm{c}) \wedge (\mathrm{w} \wedge (\mathrm{p}_i - \mathrm{c}))\end{aligned}$$

[31] See section 2.3.1.

and then the linear form $\Omega(\mathrm{p},\mathrm{v})\cdot$ takes any vector $\mathrm{u} \in E_3$ to

$$\begin{aligned}
\Omega(\mathrm{p},\mathrm{v})\cdot \mathrm{u} &= \Big(\textstyle\sum_{i=1}^{\nu} m_i(\mathrm{p}_i-\mathrm{c})\wedge(\mathrm{w}\wedge(\mathrm{p}_i-\mathrm{c}))\Big)\cdot \mathrm{u}\\
&= \textstyle\sum_{i=1}^{\nu} m_i\,\mathrm{u}\cdot\Big((\mathrm{p}_i-\mathrm{c})\wedge(\mathrm{w}\wedge(\mathrm{p}_i-\mathrm{c}))\Big)\\
&= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{w}\wedge(\mathrm{p}_i-\mathrm{c}))\cdot(\mathrm{u}\wedge(\mathrm{p}_i-\mathrm{c}))\\
&= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{w}\wedge A(\tilde{\mathrm{p}}_i-\tilde{\mathrm{c}}))\cdot(\mathrm{u}\wedge A(\tilde{\mathrm{p}}_i-\tilde{\mathrm{c}}))\\
&=: \mathbb{I}_A(\mathrm{w},\mathrm{u})\\
&= \mathbf{I}_A\mathrm{w}\cdot\mathrm{u}
\end{aligned}$$

where the mapping $\mathbb{I}_A : E_3 \times E_3 \to \mathbb{R}$ above defined is a positive definite, symmetric, bilinear form on E_3 and the *tensor of inertia* –relative to the centre of mass– $\mathbf{I}_A : E_3 \to E_3$ is the linear isomorphism which takes any vector $\mathrm{w} \in E_3$ to the unique vector $\mathbf{I}_A\mathrm{w} \in E_3$ s.t. $\mathbf{I}_A\mathrm{w}\cdot \in E_3^*$ coincides with $\mathbb{I}_A(\mathrm{w},\quad) \in E_3^*$, i.e. $\mathbf{I}_A\mathrm{w}\cdot\mathrm{u} = \mathbb{I}_A(\mathrm{w},\mathrm{u})$ for all $\mathrm{u} \in E_3$. [32]

From the above result, i.e. $\Omega(\mathrm{p},\mathrm{v})\cdot = \mathbf{I}_A\mathrm{w}\,\cdot$, it follows that

$$\Omega(\mathrm{p},\mathrm{v}) = \mathbf{I}_A\mathrm{w}$$

So, along an admissible motion $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$, whose tangent lifting $\dot\gamma = (\mathrm{p},\dot{\mathrm{p}})$ is characterized by the time functions $(\mathrm{c}, A, \dot{\mathrm{c}}, \omega)$, we obtain

$$\begin{aligned}
\Omega\circ\dot\gamma &= \Omega(\mathrm{p},\dot{\mathrm{p}})\\
&= \mathbf{I}_A\omega
\end{aligned}$$

As to R^{F} and T^{F}, they may depend on on $t \in \mathbb{R}$ and $(\mathrm{p},\mathrm{v}) \in TQ$, the latter being in turn expressed in function of $(\mathrm{c}, A, \mathrm{v_c}, \mathrm{w})$ and therefore we shall write

$$\begin{aligned}
\mathrm{R}^{\mathrm{F}}(t,\mathrm{p},\mathrm{v}) &= \mathrm{R}^{\mathrm{F}}(t,\mathrm{c},A,\mathrm{v_c},\mathrm{w})\\
\mathrm{T}^{\mathrm{F}}(t,\mathrm{p},\mathrm{v}) &= \mathrm{T}^{\mathrm{F}}(t,\mathrm{c},A,\mathrm{v_c},\mathrm{w})
\end{aligned}$$

So, along an admissible motion $\gamma : t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$, whose tangent lifting $\dot\gamma = (\mathrm{p},\dot{\mathrm{p}})$ is characterized by the time functions $(\mathrm{c}, A, \dot{\mathrm{c}}, \omega)$, we shall write

$$\begin{aligned}
\mathrm{R}^{\mathrm{F}}\circ(\tau,\dot\gamma) &= \mathrm{R}^{\mathrm{F}}(\tau,\mathrm{c},A,\dot{\mathrm{c}},\omega)\\
\mathrm{T}^{\mathrm{F}}\circ(\tau,\dot\gamma) &= \mathrm{T}^{\mathrm{F}}(\tau,\mathrm{c},A,\dot{\mathrm{c}},\omega)
\end{aligned}$$

Then we can rephrase Theorem 8 as follows:

[32] Recall that the linear mapping $\mathrm{u} \in E_3 \mapsto \mathrm{u}\cdot \in E_3^*$ is an isomorphism (see Appendix, section 4.2.1, *Euclidean metric*).

Theorem 9 *The DPMs of $\mathcal{R}$ are the admissible motions $t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in Q$ characterized by the motions $t \in I \mapsto \mathrm{c} = \mathrm{c}(t) \in \mathcal{E}_3$ of the centre of mass and the motions $t \in I \mapsto A = A(t) \in \Im_l(\tilde{E}_3, E_3)$ around the centre of mass satisfying* Newton and Euler's equations

$$m_o\, \ddot{\mathrm{c}} = \mathrm{R}^{\mathrm{F}}(\tau, \mathrm{c}, A, \dot{\mathrm{c}}, \omega)$$
$$\frac{d}{dt}\,(\mathbf{I}_A \omega) = \mathrm{T}^{(\mathrm{F})}(\tau, \mathrm{c}, A, \dot{\mathrm{c}}, \omega)$$

Remark that, if

$$\begin{aligned} \mathrm{R}^{\mathrm{F}} &= \mathrm{R}^{\mathrm{F}}(t, \mathrm{c}, \mathrm{v_c}) \\ \mathrm{T}^{\mathrm{F}} &= \mathrm{T}^{\mathrm{F}}(t, A, \mathrm{w}) \end{aligned}$$

then, along the DPMs of $\mathcal{R}$, the motions of the centre of mass are not affected by those around the centre of mass, and vice versa, since the system of Newton and Euler's equations splits into two mutually independent equations with *separated variables*, namely *Newton's equation*

$$m_o\, \ddot{\mathrm{c}} = \mathrm{R}^{\mathrm{F}}(\tau, \mathrm{c}, \dot{\mathrm{c}})$$

in the unknown $\mathrm{c} = \mathrm{c}(t) \in \mathcal{E}_3$, and *Euler's equation*

$$\frac{d}{dt}\,(\mathbf{I}_A \omega) = \mathrm{T}^{(\mathrm{F})}(\tau, A, \omega)$$

in the unknown $A = A(t) \in \Im_l(\tilde{E}_3, E_3)$.

On the one hand, the possible motions of the centre of mass – solutions of Newton's equation – coincide with the DPMs of the mechanical system modelling a (fictitious) unconstrained point-like body of mass m_o acted upon by the force field R^{F} in $\mathcal{E}_3$.

On the other hand, the possible motions around the centre of mass – solutions of Euler's equation – coincide with the DPMs of the mechanical system which would model the given rigid body under the (fictitious) hypothesis of its centre of mass being constrained to stand stationary in $\mathcal{E}_3$.

2.3.5 Euler's equation

We shall now give an insight into Euler's equation, meant as the equation governing the DPMs of a mechanical system $\mathcal{R}_{\tilde{o}}$ modelling a rigid body under the hypoyhesis of a point $\tilde{\mathrm{o}}$ of its body space $\tilde{\mathcal{E}}_3$ (not necessarily the centre of mass) being constrained to stay stationary at a given point o of the reference space $\mathcal{E}_3$. [33]

[33] In such a case, a rigid motion $\mathrm{p}_i(t) = \mathrm{o} + A(t)\,(\tilde{\mathrm{p}}_i - \tilde{\mathrm{o}})$ is determined only by $A(t)$ and its velocity $\dot{\mathrm{p}}_i(t) = \omega(t) \wedge \big(\mathrm{p}_i(t) - \mathrm{o}\big)$ is determined only by $\omega(t)$. The DPMs of $\mathcal{R}_{\tilde{o}}$ are

(i) The left hand side of Euler's equation is the time derivative of the angular momentum $\mathbf{I}_A\omega$, which depends on time through a rigid motion $A = A(t)$ around the stationary point and its angular velocity $\omega = \omega(t)$.

In order to evaluate such a derivative, it is convenient to introduce the tensor of inertia –relative to the stationary point– in the body space, i.e. the linear isomorphism $\tilde{\mathbf{I}} : \tilde{E}_3 \to \tilde{E}_3$ which takes any vector $\tilde{\mathrm{w}} \in \tilde{E}_3$ to the unique vector $\tilde{\mathbf{I}}\tilde{\mathrm{w}}$ s.t. $\tilde{\mathbf{I}}\tilde{\mathrm{w}} \cdot \tilde{\mathrm{u}} = \sum_{i=1}^{\nu} m_i(\tilde{\mathrm{w}} \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{o}})) \cdot (\tilde{\mathrm{u}} \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{o}}))$ for all $\tilde{\mathrm{u}} \in \tilde{E}_3$.

Let $A \in \Im_l(\tilde{E}_3, E_3)$.

For all $\tilde{\mathrm{w}}, \tilde{\mathrm{u}} \in \tilde{E}_3$ and $\mathrm{w} = A(\tilde{\mathrm{w}}),\ \mathrm{u} = A(\tilde{\mathrm{u}}) \in E_3$, we have [34]

$$\begin{aligned}
\mathbf{I}_A\mathrm{w} \cdot \mathrm{u} &= \textstyle\sum_{i=1}^{\nu} m_i(\mathrm{w} \wedge A(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \cdot (\mathrm{u} \wedge A(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \\
&= \textstyle\sum_{i=1}^{\nu} m_i(A(\tilde{\mathrm{w}}) \wedge A(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \cdot (A(\tilde{\mathrm{u}}) \wedge A(\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \\
&= \textstyle\sum_{i=1}^{\nu} m_i A(\tilde{\mathrm{w}} \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \cdot A(\tilde{\mathrm{u}} \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \\
&= \textstyle\sum_{i=1}^{\nu} m_i(\tilde{\mathrm{w}} \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \cdot (\tilde{\mathrm{u}}) \wedge (\tilde{\mathrm{p}}_i - \tilde{\mathrm{c}})) \\
&= \tilde{\mathbf{I}}\tilde{\mathrm{w}} \cdot \tilde{\mathrm{u}} \\
&= A(\tilde{\mathbf{I}}\tilde{\mathrm{w}}) \cdot A(\tilde{\mathrm{u}}) \\
&= A(\tilde{\mathbf{I}}\tilde{\mathrm{w}}) \cdot \mathrm{u}
\end{aligned}$$

Hence

$$\mathbf{I}_A\mathrm{w} = A(\tilde{\mathbf{I}}\tilde{\mathrm{w}})$$

whenever

$$\mathrm{w} = A(\tilde{\mathrm{w}})$$

Now, let $A = A(t)$ be a rigid motion around the stationary point and $\omega = \omega(t)$ its angular velocity.

If we introduce the time function $\tilde{\omega}(t) := A(t)^{-1}(\omega(t))$, from

$$\omega = A(\tilde{\omega})$$

we obtain, by time derivation, [35]

$$\begin{aligned}
\dot{\omega} &= A(\dot{\tilde{\omega}}) + \dot{A}(\tilde{\omega}) \\
&= A(\dot{\tilde{\omega}}) + (\dot{A} \circ A^{-1} \circ A)\,(\tilde{\omega}) \\
&= A(\dot{\tilde{\omega}}) + \omega \wedge A(\tilde{\omega}) \\
&= A(\dot{\tilde{\omega}}) + \omega \wedge \omega \\
&= A(\dot{\tilde{\omega}})
\end{aligned}$$

then determined by the solutions of Euler's equation.

[34] Recall that A preserves both the scalar and the wedge product, i.e. $\tilde{\mathrm{v}}_1 \cdot \tilde{\mathrm{v}}_2 = A(\tilde{\mathrm{v}}_1) \cdot A(\tilde{\mathrm{v}}_2)$ and $A(\tilde{\mathrm{v}}_1 \wedge \tilde{\mathrm{v}}_2) = A(\tilde{\mathrm{v}}_1) \wedge A(\tilde{\mathrm{v}}_2)$ for all $\tilde{\mathrm{v}}_1, \tilde{\mathrm{v}}_2 \in \tilde{E}_3$.

[35] Remark that the formula below implies: $\dot{\omega} = 0$, iff $\dot{\tilde{\omega}} = 0$ (i.e. ω keeps constant, iff so does $\tilde{\omega}$).

As a consequence,

$$\begin{aligned}\frac{d}{dt}\left(\mathbf{I}_A\omega\right) &= \frac{d}{dt}A(\tilde{\mathbf{I}}\tilde{\omega})\\ &= A(\tilde{\mathbf{I}}\tilde{\omega})^{\boldsymbol{\cdot}} + \dot{A}(\tilde{\mathbf{I}}\tilde{\omega})\\ &= A(\tilde{\mathbf{I}}\dot{\tilde{\omega}}) + (\dot{A}\circ A^{-1}\circ A)\,(\tilde{\mathbf{I}}\tilde{\omega})\\ &= A(\tilde{\mathbf{I}}\dot{\tilde{\omega}}) + \omega\wedge A(\tilde{\mathbf{I}}\tilde{\omega})\\ &= A(\tilde{\mathbf{I}}\dot{\tilde{\omega}}) + A(\tilde{\omega})\wedge A(\tilde{\mathbf{I}}\tilde{\omega})\end{aligned}$$

From the last two equalities, we obtain

$$\frac{d}{dt}\left(\mathbf{I}_A\omega\right) = \mathbf{I}_A\dot{\omega} + \omega\wedge\mathbf{I}_A\omega$$

and

$$\frac{d}{dt}\left(\mathbf{I}_A\omega\right) = A(\tilde{\mathbf{I}}\dot{\tilde{\omega}} + \tilde{\omega}\wedge\tilde{\mathbf{I}}\tilde{\omega})$$

respectively.

So we can state the following result:

Theorem 10 *Euler's equation, characterizing the DPMs of* $\mathcal{R}_{\tilde{o}}$ *, reads*

$$\mathbf{I}_A\dot{\omega} + \omega\wedge\mathbf{I}_A\omega = \mathrm{T}^{(\mathrm{F})}(\tau, A, \omega)$$

or

$$A(\tilde{\mathbf{I}}\dot{\tilde{\omega}} + \tilde{\omega}\wedge\tilde{\mathbf{I}}\tilde{\omega}) = \mathrm{T}^{(\mathrm{F})}(\tau, A, A(\tilde{\omega}))$$

Euler's equation can also be given a scalar (rather than vector) formulation by orderly equalling the components of its left and right hand sides in a 'moving' basis of E_3 .

To this end, we recall that the tensor of inertia $\tilde{\mathbf{I}}$, owing to its symmetry and positive definiteness, admits at most three distinct (positive) eigenvalues –which will be denoted by (I_h) or (I^h) with $h = 1,2,3$– and admits at least one orthonormal basis of eigenvectors $(\tilde{\mathrm{e}}_h)$, with the given orientation in $\tilde{E}_3$, orderly corresponding to the above eigenvalues, i.e. [36]

$$\tilde{\mathbf{I}}\,\tilde{\mathrm{e}}_h = I_h\tilde{\mathrm{e}}_h$$

[36] Just recall that a real number I_h is an eigenvalue of $\tilde{\mathbf{I}}$, iff there exists a non-zero vector $\tilde{\mathrm{e}}_h$, called eigenvector of $\tilde{\mathbf{I}}$ corresponding to I_h, s.t. $\tilde{\mathbf{I}}\,\tilde{\mathrm{e}}_h = I_h\tilde{\mathrm{e}}_h$ (see Appendix, section 4.2.1, *Eigenvalues and eigenvectors*).

Now remark that any $A \in \Im_l(\tilde{E}_3, E_3)$ takes the basis of eigenvectors $(\tilde{\mathrm{e}}_h)$ of $\tilde{\mathbf{I}}$ to a basis of eigenvectors

$$\mathrm{e}_h = A(\tilde{\mathrm{e}}_h)$$

of $\mathbf{I}_A$ corresponding to the same eigenvalues, since

$$\begin{aligned} \mathbf{I}_A \mathrm{e}_h &= A(\tilde{\mathbf{I}}\tilde{\mathrm{e}}_h) \\ &= A(I_h \tilde{\mathrm{e}}_h) \\ &= I_h A(\tilde{\mathrm{e}}_h) \\ &= I_h \mathrm{e}_h \end{aligned}$$

Then, if A depends on time (rigid motion around the stationary point), the above basis of eigenvectors (e_h) depends on time as well (*moving principal basis*), and, by time derivation, we obtain the *Poisson's formulae*

$$\begin{aligned} \dot{\mathrm{e}}_h &= \dot{A}(\tilde{\mathrm{e}}_h) \\ &= (\dot{A} \circ A^{-1} \circ A)(\tilde{\mathrm{e}}_h) \\ &= \omega \wedge A(\tilde{\mathrm{e}}_h) \\ &= \omega \wedge \mathrm{e}_h \end{aligned}$$

If the angular velocity of the above rigid motion is expressed, in the moving basis, by

$$\omega = \omega^h \mathrm{e}_h$$

its time derivative is expressed by

$$\begin{aligned} \dot{\omega} &= \dot{\omega}^h \mathrm{e}_h + \omega^h \dot{\mathrm{e}}_h \\ &= \dot{\omega}^h \mathrm{e}_h + \omega^h (\omega \wedge \mathrm{e}_h) \\ &= \dot{\omega}^h \mathrm{e}_h + \omega \wedge \omega^h \mathrm{e}_h \\ &= \dot{\omega}^h \mathrm{e}_h + \omega \wedge \omega \\ &= \dot{\omega}^h \mathrm{e}_h \end{aligned}$$

As a consequence,

$$\begin{aligned} \mathbf{I}_A \dot{\omega} &= \mathbf{I}_A(\dot{\omega}^h \mathrm{e}_h) \\ &= \dot{\omega}^h (\mathbf{I}_A \mathrm{e}_h) \\ &= \dot{\omega}^h (I_h \mathrm{e}_h) \\ &= (I^h \dot{\omega}^h) \mathrm{e}_h \end{aligned}$$

and, in the same way,

$$\mathbf{I}_A \omega = (I^h \omega^h) \mathrm{e}_h$$

whence

$$\omega \wedge \mathbf{I}_A \omega = -((I^{h+1} - I^{h+2})\omega^{h+1}\omega^{h+2})\mathrm{e}_h$$

So, if we put [37]

$$\mathrm{T}^{(\mathrm{F})} = T^h \mathrm{e}_h$$

(i.e. $T^h := \mathrm{T}^{(\mathrm{F})} \cdot \mathrm{e}_h$), we obtain the following reformulation of Euler's equation:

Theorem 11 *Euler's equation reads*

$$I^h \dot{\omega}^h - (I^{h+1} - I^{h+2})\omega^{h+1}\omega^{h+2} = T^h \quad (h = 1, 2, 3)$$

Gyroscopic effects

A meaningful example of qualitative analysis of the solutions of Euler's equation, carried out with the aid of the above scalar formulation, will now be shown.

If

$$I^1 = I^2$$

i.e. the tensor of inertia admits no more than two distinct eigenvalues, the rigid body is said to have a *gyroscopic structure* around the axis

$$\tilde{\mathcal{G}} = \tilde{\mathrm{o}} + \mathrm{Span}(\tilde{\mathrm{e}}_3) \subset \tilde{\mathcal{E}}_3$$

called *gyroscopic axis*; the above *Euler's scalar equations* then take the form

$$\begin{aligned} &I(\dot{\omega}^1 - \omega^2\omega^3) + I^3\omega^2\omega^3 = T^1 \\ &I(\dot{\omega}^2 + \omega^3\omega^1) - I^3\omega^3\omega^1 = T^2 \\ &I^3\dot{\omega}^3 = T^3 \end{aligned}$$

where we have put $I := I^1 = I^2$.

Now let

$$T^3 = 0$$

i.e. $T^{(\mathrm{F})}$ is orthogonal, at every admissible configuration, to the corresponding position of the gyroscopic axis; [38] in such a case, along any solution of Euler's

[37] Dependence on (τ, A, ω) will be understood.

[38] Check that every 'configuration' $\alpha = (\mathrm{o}, A)$ of the body space in the reference space, takes $\tilde{\mathcal{G}}$ to the position $\mathcal{G} = \mathrm{o} + \mathrm{Span}(\mathrm{e}_3) \subset \mathcal{E}_3$ with $\mathrm{e}_3 = A(\tilde{\mathrm{e}}_3)$.

scalar equations, the angular velocity $\omega = \omega(t)$ exhibits a constant ‘gyroscopic component’, say

$$\omega^3 = r_o$$

If $\omega = \omega(t)$ corresponds to an initial condition $\omega(t_o) = \omega_o$ satisfying

$$\omega_o^1 = 0\,, \quad \omega_o^2 = 0\,, \quad \omega_o^3 = r_o \gg 0$$

(strong initial angular velocity around the initial position of the gyroscopic axis), then – at least in a small neighbourhood of t_o – the ratio

$$\frac{1}{r_o}\,\omega = \left(\frac{\omega^1}{r_o}\,\mathrm{e}_1 + \frac{\omega^2}{r_o}\,\mathrm{e}_2\right) + \mathrm{e}_3$$

keeps very close to e_3, so that we can accept the approximation

$$\frac{1}{r_o}\,\omega = \mathrm{e}_3$$

whence

$$\begin{aligned} \frac{1}{r_o}\,\dot{\omega} &= \dot{\mathrm{e}}_3 \\ &= \omega \wedge \mathrm{e}_3 \end{aligned}$$

or, in terms of components,

$$\frac{1}{r_o}\,\dot{\omega}^1 = \omega^2\,, \quad \frac{1}{r_o}\,\dot{\omega}^2 = -\omega^1$$

Euler’s scalar equations then imply

$$I^3\omega^2 r_o = T^1\,, \quad -I^3\omega^1 r_o = T^2$$

that is,

$$\omega^2 = \frac{1}{I^3 r_o} T^1\,, \quad -\omega^1 = \frac{1}{I^3 r_o} T^2$$

or, in vector formalism,

$$\omega \wedge \mathrm{e}_3 = \frac{1}{I^3 r_o}\,\mathrm{T}^{(\mathrm{F})}$$

If $\mathrm{T}^{(\mathrm{F})} \neq 0$, the above result shows the following *gyroscopic effects*: on the one hand, the *tenaciousness of the gyroscopic axis* (whose ‘displacements’ $\dot{\mathrm{e}}_3 = \omega \wedge \mathrm{e}_3$, in any unit time interval, are inversely proportional to the large value of r_o and are then very small in norm) and, on the other hand, the *tendency of the gyroscopic axis to parallelism with the torque* (since the above ‘displacements’, whenever perceptible, are parallel to $\mathrm{T}^{(\mathrm{F})}$).

Euler's geodesic equation

We shall now discuss the Cauchy problem set up by *Euler's geodesic equation* (characterizing the inertial motions of $\mathcal{R}_{\tilde{o}}$)

$$\frac{d}{dt}(\mathbf{I}_A\omega) = 0 \qquad (\circ)$$

or

$$\mathbf{I}_A\dot{\omega} + \omega \wedge \mathbf{I}_A\omega = 0 \qquad (\circ)$$

or

$$\tilde{\mathbf{I}}\dot{\tilde{\omega}} + \tilde{\omega} \wedge \tilde{\mathbf{I}}\tilde{\omega} = 0 \qquad (\tilde{\circ})$$

or

$$I^h\dot{\omega}^h - (I^{h+1} - I^{h+2})\omega^{h+1}\omega^{h+2} = 0 \quad (h = 1, 2, 3) \qquad (\hat{\circ})$$

and initial conditions

$$A(t_o) = A_o\,, \;\; \omega(t_o) = \omega_o$$

Initial conditions with

$$\omega_o = 0$$

obviously determine a 'static' maximal solution

$$\omega = 0$$

namely, the state of rest $A = A_o$ at the initial configuration.

Initial conditions with

$$\omega_o \neq 0$$

could then be expected to determine a 'uniform' maximal solution

$$\omega = \omega_o$$

namely, the uniform circular motion of rotation around the stationary position $\mathcal{P}_o = \mathrm{o} + \mathrm{Span}(\omega_o) \subset \mathcal{E}_3$ of the axis $\tilde{\mathcal{P}}_o = \tilde{\mathrm{o}} + \mathrm{Span}(\tilde{\omega}_o) \subset \tilde{\mathcal{E}}_3$ with

$$\tilde{\omega}_o := A_o^{-1}(\omega_o)$$

However, remark that the constant function $\omega = \omega_o$ satisfies ($\circ$), iff the constant function $\tilde{\omega} = \tilde{\omega}_o$ satisfies ($\tilde{\circ}$), i.e.

$$\tilde{\omega}_o \wedge \tilde{\mathbf{I}}\tilde{\omega}_o = 0$$

that is, iff $\tilde{\omega}_o$ is an eigenvector of $\tilde{\mathbf{I}}$.

Therefore, a *principal axis* –i.e. an axis $\tilde{\mathcal{P}}_o = \tilde{\mathrm{o}} + \mathrm{Span}(\tilde{\omega}_o) \subset \tilde{\mathcal{E}}_3$ whose direction is that of an eigenvector $\tilde{\omega}_o$ of $\tilde{\mathbf{I}}$– is the only kind of axis around which uniform circular motions of rotation are possible by inertia, and then it will be thought of as a 'symmetry axis' for the mass distribution of the rigid body around $\tilde{\mathrm{o}}$ (described by the tensor of inertia).

So, if the body is given an initial angular velocity ω_o around a non-principal axis (that is, $\tilde{\omega}_o$ is not an eigenvector of $\tilde{\mathbf{I}}$), then the asymmetry of the mass distribution with respect to such an axis will be thought of as the cause of the fact that, along the corresponding inertial motion, the angular velocity does not keep its initial value ω_o.

A complete qualitative discussion of the Cauchy problem, including the case of $\tilde{\omega}_o$ not being an eigenvector of $\tilde{\mathbf{I}}$, can easily be carried out under the hypothesis of gyroscopic structure

$$I^1 = I^2 =: I$$

In such a case, along the maximal solution $A = A(t)$ of Euler's geodesic equation, the angular velocity $\omega = \omega(t)$ exhibits a constant gyroscopic component

$$\omega^3 = \omega_o^3 =: r_o$$

(owing to the third of scalar equations ($\hat{\mathrm{o}}$)) and then the conservation law of angular momentum

$$\mathbf{I}_A \omega = \mathbf{I}_{A_o} \omega_o =: \Omega_o$$

(following from ($\circ$)) is expressed, in a moving principal basis, by

$$I(\omega^1 \mathrm{e}_1 + \omega^2 \mathrm{e}_2) + I^3 r_o \mathrm{e}_3 = \Omega_o$$

From the above first integrals, we deduce

$$\begin{aligned} \omega &= (\omega^1 \mathrm{e}_1 + \omega^2 \mathrm{e}_2) + \omega^3 \mathrm{e}_3 \\ &= \left(-\frac{I^3}{I} r_o \mathrm{e}_3 + \frac{1}{I} \Omega_o \right) + r_o \mathrm{e}_3 \\ &= \left(1 - \frac{I^3}{I} \right) r_o \mathrm{e}_3 + \frac{1}{I} \Omega_o \end{aligned}$$

If $\tilde{\omega}_o$ is not an eigenvector of $\tilde{\mathbf{I}}$, the above solution is said to be a *regular precession*, composed of a *proper* uniform rotation $\omega_1 := \left(1 - \frac{I^3}{I}\right) r_o \mathrm{e}_3$ around the moving gyroscopic axis $\mathcal{G} = \mathrm{o} + \mathrm{Span}(\mathrm{e}_3)$ and a *precessional* uniform rotation $\omega_2 := \frac{1}{I}\Omega_o$ around the precessional axis $\mathcal{P} = \mathrm{o} + \mathrm{Span}(\Omega_o)$.

Chapter 3

Analytical dynamics

From *historical* dynamics we shall now pass on to *analytical* (Lagrangian-Hamiltonian) dynamics, by translating d'Alembert equation into a system of ordinary differential equations with the aid of the coordinate formalism on the admissible configuration space Q. To that end, the price to pay will generally be to give up *global dynamics* (study of the DPMs on the whole space Q) and to confine oneself to a *local dynamics* (study of the DPMs within a coordinate domain of Q).

3.1 Lagrange

Let $\mathcal{S} := (Q, m, \mathrm{F})$ be a mechanical system. Among the solutions of equation $\mathrm{D}_{d'Al}$ associated with $\mathcal{S}$ (i. e. the DPMs of $\mathcal{S}$), those living in the coordinate domain of a (local) chart of Q will be given a coordinate characterization.

3.1.1 Lagrange differential equation

Let

$$\xi : W \subset \mathbb{R}^n \to \mathcal{U} \subset Q$$

be a (local) chart of Q (giving the points of $\mathcal{U}$ n-tuples of coordinates, called *Lagrangian coordinates*).

Then let

$$t \in I \mapsto \mathrm{p} = \mathrm{p}(t) \in \mathcal{U}$$

a smooth motion living in the coordinate domain $\mathcal{U}$ and

$$t \in I \mapsto q = q(t) = \xi^{-1}(\mathrm{p}(t)) \in W$$

its *coordinate expression* in ξ.

As is known, such an admissible motion is a solution of $\mathrm{D}_{d'Al}$, iff it satisfies the principle of virtual works, requiring, for all $t \in I$,

$$m\,\ddot{\mathrm{p}}(t) \cdot \delta\mathrm{p} = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \cdot \delta\mathrm{p}\ , \quad \forall\, \delta\mathrm{p} \in T_{\mathrm{p}(t)}Q = \mathrm{Im}\, d_{q(t)}\xi$$

that is,

$$m\,\ddot{\mathrm{p}}(t) \cdot d_{q(t)}\xi\,(\delta q) = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \cdot d_{q(t)}\xi\,(\delta q)\ , \quad \forall\, \delta q \in \mathbb{R}^n$$

or

$$m\,\ddot{\mathrm{p}}(t) \cdot \frac{\partial \mathrm{p}}{\partial q^h}\Big|_{q(t)} \delta q^h = \mathrm{F}(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \cdot \frac{\partial \mathrm{p}}{\partial q^h}\Big|_{q(t)} \delta q^h\ , \quad \forall\, \delta q \in \mathbb{R}^n$$

As a consequence, $\mathrm{p} = \mathrm{p}(t) \in \mathcal{U}$ is a solution of $\mathrm{D}_{d'Al}$, iff it satisfies the system of *Lagrange equations* [1]

$$m\,\ddot{\mathrm{p}} \cdot \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \;=\; \mathrm{F}(\tau, \mathrm{p}, \dot{\mathrm{p}}) \cdot \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q$$

(h = 1, ..., n), orderly equalling –along the motion– the *Lagrangian components* of the inertial forces (up to the sign) and those of the active forces, respectively.

The left and right hand sides of Lagrange equations will now be expressed in terms of the unknown $q = q(t) \in W$, by means of the coordinate expression K of the kinetic energy, that is, for all $(q, v) \in W \times \mathbb{R}^n$,

$$K(q, v) := \mathsf{K}(\xi(q), d_q\xi\,(v))$$

and the coordinate expression $F = (F_h)_{h=1,\dots,n}$ of the Lagrangian components of F, that is, for all $(t, q, v) \in \mathbb{R} \times W \times \mathbb{R}^n$,

$$F_h(t, q, v) := \mathrm{F}(t, \xi(q), d_q\xi\,(v)) \cdot \frac{\partial p}{\partial q^h}\Big|_q$$

[1] $\tau : t \mapsto \tau(t) := t$ will denote the inclusion mapping of any open interval into $\mathbb{R}$.

As to the left hand side, we have [2]

$$
\begin{aligned}
m\,\ddot{\mathrm{p}}\cdot\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q &= \frac{d}{dt}\left(m\,\dot{\mathrm{p}}\cdot\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q\right) - m\,\dot{\mathrm{p}}\cdot\frac{d}{dt}\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \\
&= \frac{d}{dt}\left(m\,\dot{\mathrm{p}}\cdot\frac{\partial \dot{\mathrm{p}}}{\partial v^h}\Big|_{(q,v)}\right) - m\,\dot{\mathrm{p}}\cdot\frac{\partial \dot{\mathrm{p}}}{\partial q^h}\Big|_{(q,v)} \\
&= \frac{d}{dt}\frac{\partial}{\partial v^h}\Big|_{(q,v)}\left(\frac{1}{2}\,m\,\dot{\mathrm{p}}\cdot\dot{\mathrm{p}}\right) - \frac{\partial}{\partial q^h}\Big|_{(q,v)}\left(\frac{1}{2}\,m\,\dot{\mathrm{p}}\cdot\dot{\mathrm{p}}\right) \\
&= \frac{d}{dt}\frac{\partial K}{\partial v^h}\Big|_{(q,v)} - \frac{\partial K}{\partial q^h}\Big|_{(q,v)}
\end{aligned}
$$

As to the right hand side, we have

$$
\begin{aligned}
\mathrm{F}(\tau,\mathrm{p},\dot{\mathrm{p}})\cdot\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q &= \mathrm{F}(\tau,\xi(q),d_q\xi\,(v))\cdot\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \\
&= F_h(\tau,q,v)
\end{aligned}
$$

So Lagrange equations read

$$
\frac{d}{dt}\frac{\partial K}{\partial v^h}\Big|_{(q,v)} - \frac{\partial K}{\partial q^h}\Big|_{(q,v)} = F_h(\tau,q,v)
$$

that is,

$$
\frac{\partial^2 K}{\partial v^k \partial v^h}\Big|_{(q,v)}\dot{v}^k + \frac{\partial^2 K}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial K}{\partial q^h}\Big|_{(q,v)} = F_h(t,q,v)
$$

with $v^h = \dot{q}^h$.

---

[2] Preliminarly recall that

$$
\dot{\mathrm{p}} = d_q\xi\,(v) = \frac{\partial \mathrm{p}}{\partial q^k}\Big|_q v^k
$$

is expressed in ξ as a function of $2n$ arguments (q,v), where $q = q(t)$ is the coordinate expression of the motion and $v = \dot{q} = \dot{q}(t)$ its first-order derivative (see Appendix, section 4.4.2, *Coordinate expression*).

Hence, for all $h = 1,\dots,n$,

$$
\frac{\partial \dot{\mathrm{p}}}{\partial v^h}\Big|_{(q,v)} = \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q\,, \quad \frac{\partial \dot{\mathrm{p}}}{\partial q^h}\Big|_{(q,v)} = \frac{\partial^2 \mathrm{p}}{\partial q^h \partial q^k}\Big|_q v^k = \frac{d}{dt}\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q
$$

Lagrange equations are then the conditions characterizing the solutions $q = q(t) \in W$ of the second-order differential-algebraic equation in implicit form on $\mathbb{R}^n$ given by

$$\begin{aligned} \mathrm{D}_{Lagr} \quad = \quad & \{(t, q, v, a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid q \in W, \\ & \frac{\partial^2 K}{\partial v^k \partial v^h}\Big|_{(q,v)} a^k + \frac{\partial^2 K}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial K}{\partial q^h}\Big|_{(q,v)} = F_h(t, q, v), \\ & \forall\, h = 1, \ldots, n\,\} \end{aligned}$$

Lagrange equations and equations

$$\dot{q}^h = v^h$$

can as well be seen as the conditions characterizing the solutions $(q, v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ of the first-order differential-algebraic equation in implicit form on $\mathbb{R}^{2n}$ given by

$$\begin{aligned} \mathrm{D}'_{Lagr} \quad = \quad & \{(t; q, v; u, a) \in \mathbb{R} \times T\mathbb{R}^{2n} \mid (q, v) \in W \times \mathbb{R}^n, \\ & u^h = v^h, \\ & \frac{\partial^2 K}{\partial v^k \partial v^h}\Big|_{(q,v)} a^k + \frac{\partial^2 K}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial K}{\partial q^h}\Big|_{(q,v)} = F_h(t, q, v), \\ & \forall\, h = 1, \ldots, n\,\} \end{aligned}$$

D_{Lagr} (or D'_{Lagr}) is called *Lagrange differential equation* (locally) associated with (K, F).

The above considerations prove the following

Theorem 12 *A smooth motion* $\mathrm{p} = \mathrm{p}(t) \in \mathcal{U}$ *is a solution of* $D_{d'Al}$, *iff its coordinate expression* $q = q(t) \in W$ *is a solution of* D_{Lagr} *or, equivalently,* [3] *the first projection of a solution* $(q, v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ *of* D'_{Lagr}.

D_{Lagr} (resp. D'_{Lagr}) is reducible to normal form on W (resp. $W \times \mathbb{R}^n$) – and then deterministic – owing to the non-singularity of matrix $\left(\frac{\partial^2 K}{\partial v^k \partial v^h}\Big|_{(q,v)}\right)$, which can be checked as follows.

[3] See Appendix, section 4.6.2, *First-order reformulation.*

At any

$$\mathrm{p} = \xi(q) \in \mathcal{U} \subset Q\,, \quad \mathrm{v} = d_q\xi(v) = v^h \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \in T_{\mathrm{p}}Q$$

the value of the kinetic energy is

$$\mathsf{K}(\mathrm{p}, \mathrm{v}) = \frac{1}{2}\, m\, \mathrm{v} \cdot \mathrm{v} = \frac{1}{2} \left(m\, \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \cdot \frac{\partial \mathrm{p}}{\partial q^k}\Big|_q \right) v^h v^k$$

Putting

$$g_{hk}(q) := m\, \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \cdot \frac{\partial \mathrm{p}}{\partial q^k}\Big|_q$$

we obtain the coordinate expression

$$K(q, v) = \frac{1}{2}\, g_{hk}(q) v^h v^k$$

whence

$$\frac{\partial K}{\partial v^h}\Big|_{(q,v)} = g_{hk}(q) v^k$$

and

$$\frac{\partial^2 K}{\partial v^k \partial v^h}\Big|_{(q,v)} = g_{hk}(q)$$

Now remark that $\mathsf{K}(\mathrm{p}, \cdot\,)$ is the semisquare of the norm of the Euclidean metric defined in $T_{\mathrm{p}}Q$ by

$$g_{\mathrm{p}} : T_{\mathrm{p}}Q \times T_{\mathrm{p}}Q \to \mathbb{R} : (\mathrm{v}, \mathrm{w}) \mapsto g_{\mathrm{p}}(\mathrm{v}, \mathrm{w}) := m\, \mathrm{v} \cdot \mathrm{w}$$

whose non-degenerateness

$$g_{\mathrm{p}}(\mathrm{v}\,,\ \cdot\ ) = 0, \quad \text{iff} \ \ \mathrm{v} = 0$$

in tems of the components

$$g_{\mathrm{p}} \left(\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q, \frac{\partial \mathrm{p}}{\partial q^k}\Big|_q \right) = g_{hk}(q)$$

is expressed by

$$g_{hk}(q) v^k = 0, \ h = 1, \ldots, n\,, \quad \text{iff} \ \ (v^k)_{k=1,\ldots,n} = 0$$

Hence the non-singularity of matrix $(g_{kh}(q))$.

As to global dynamics, from Theorem 12 we infer what follows.

Let $\gamma : t \in I \mapsto \mathrm{p}(t) \in Q$ be an admissible motion.

If ξ is a chart of Q near a point $\mathrm{p}(t_*) \in \mathrm{Im}(\gamma)$, i.e. $\mathrm{p}(t_*) \in \mathcal{U} := \mathrm{Im}(\xi)$, then, by continuity, there exists an open subinterval $I_* \ni t_*$ of I s.t. $\mathrm{Im}(\gamma|_{I_*}) \subset \mathcal{U}$ and $\xi^{-1} \circ \gamma|_{I_*}$ will be said to be a (local) coordinate expression of γ near $\mathrm{p}(t_*)$.

Now, if γ is a solution of $\mathrm{D}_{d'Al}$, i.e. $\mathrm{Graph}(\ddot{\gamma}) \subset \mathrm{D}_{d'Al}$, then, for each point $\mathrm{p}(t_*) \in \mathrm{Im}(\gamma)$, we have $\mathrm{Graph}(\ddot{\gamma}|_{I_*}) \subset \mathrm{D}_{d'Al}$, that is, $\gamma|_{I_*}$ is a solution of $\mathrm{D}_{d'Al}$. [4] As a consequence, owing to Theorem 12, the (local) coordinate expression $\xi^{-1} \circ \gamma|_{I_*}$ of γ near $\mathrm{p}(t_*)$ is a solution of D_{Lagr}.

Conversely, if, near each point $\mathrm{p}(t_*) \in \mathrm{Im}(\gamma)$, γ admits a (local) coordinate expression $\xi^{-1} \circ \gamma|_{I_*}$ which is a solution of D_{Lagr}, then, owing to Theorem 12, $\gamma|_{I_*}$ is a solution of $\mathrm{D}_{d'Al}$, i.e $\mathrm{Graph}(\ddot{\gamma}|_{I_*}) \subset \mathrm{D}_{d'Al}$, whence $(t_*, \ddot{\gamma}(t_*)) \in \mathrm{D}_{d'Al}$. Owing to the arbitrariness of t_* in I, that means $\mathrm{Graph}(\ddot{\gamma}) \subset \mathrm{D}_{d'Al}$, that is, γ is a solution of $\mathrm{D}_{d'Al}$.

So we have proved the following

Theorem 13 *An admissible motion is a solution of* $\mathrm{D}_{d'Al}$*, iff, near each point of its orbit, it admits a (local) coordinate expression which is a solution of* D_{Lagr}. [5]

3.1.2 Euler-Lagrange differential equation

The dynamics of a *conservative* system $\mathcal{S}$ can be given a formulation where the kinetic energy K and the virtual work of F merge into a unique object, namely the *Lagrangian function*

$$\mathsf{L} : TQ \to \mathbb{R} : (\mathrm{p}, \mathrm{v}) \mapsto \mathsf{L}(\mathrm{p}, \mathrm{v}) := \mathsf{K}(\mathrm{p}, \mathrm{v}) - \mathsf{V}(\mathrm{p})$$

(kinetic *minus* potential energy), whose (local) coordinate expression in a chart $\xi : W \to \mathcal{U}$ of Q is given, for all $(q, v) \in W \times \mathbb{R}^n$, by

$$L(q, v) := K(q, v) - V(q)$$

with

$$V(q) := \mathsf{V}(\xi(q))$$

[4] Recall that, owing to the local character of derivatives, $\ddot{\gamma}|_{I_*}$ is the second-order tangent lift of $\gamma|_{I_*}$.

[5] Owing to the determinism of D_{Lagr}, Theorem 13 implies that any Cauchy problem $\big(\mathrm{D}_{d'Al}, t_o, (\mathrm{p}_o, \mathrm{v}_o)\big)$, with $(\mathrm{p}_o, \mathrm{v}_o) \in TQ$, admits of solutions and they locally agree (i.e. any two of them coincide on a neighbourhood of $t_o \in \mathbb{R}$). Hence one could infer that they are all restrictions of a unique maximal solution.

Indeed, the Lagrangian components components of F in ξ are [6]

$$\begin{aligned} F_h(q) &:= \mathrm{F}(\xi(q)) \cdot \frac{\partial \mathrm{p}}{\partial q^h}\Big|_q \\ &= -\, d_{\xi(q)}\mathsf{V}\Big(\frac{\partial \mathrm{p}}{\partial q^h}\Big|_q\Big) \\ &= -\,(d_{\xi(q)}\mathsf{V} \circ d_q\xi)\,(\delta_h) \\ &= -\, d_q V\,(\delta_h) \\ &= -\,\frac{\partial V}{\partial q^h}\Big|_q \end{aligned}$$

As a consequence, Lagrange equations read

$$\frac{d}{dt}\frac{\partial K}{\partial v^h}\Big|_{(q,v)} - \frac{\partial K}{\partial q^h}\Big|_{(q,v)} = -\frac{\partial V}{\partial q^h}\Big|_q$$

or

$$\frac{d}{dt}\frac{\partial (K-V)}{\partial v^h}\Big|_{(q,v)} - \frac{\partial (K-V)}{\partial q^h}\Big|_{(q,v)} = 0$$

The above equations do not differ from *Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial v^h}\Big|_{(q,v)} - \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = 0$$

that is,

$$\frac{\partial^2 L}{\partial v^k \partial v^h}\Big|_{(q,v)} \dot v^k + \frac{\partial^2 L}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = 0$$

with $v^h = \dot q^h$.

Euler-Lagrange equations are the conditions characterizing the solutions $q = q(t) \in W$ of the time-indepent, second-order, differential-algebraic equation in implicit form on $\mathbb{R}^n$ given by

$$\begin{aligned} \mathcal{D}_{Eul-Lagr} \;=\; & \{(q,v,a) \in T^2\mathbb{R}^n \mid q \in W, \\ & \frac{\partial^2 L}{\partial v^k \partial v^h}\Big|_{(q,v)} a^k + \frac{\partial^2 L}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = 0, \\ & \forall\, h = 1, \ldots, n\,\} \end{aligned}$$

[6] Recall the chain rule for $V = \tilde{\mathsf{V}} \circ \xi = \mathsf{V} \circ \xi$, i.e. $d_q V = d_{\xi(q)}\tilde{\mathsf{V}} \circ d_q\xi = d_{\xi(q)}\mathsf{V} \circ d_q\xi$. Also recall that (δ_h) denotes the natural basis of $\mathbb{R}^n$.

Euler-Lagrange equations and equations

$$\dot{q}^h = v^h$$

can as well be seen as the conditions characterizing the solutions $(q, v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ of the time-indepent, first-order, differential-algebraic equation in *implicit form* on $\mathbb{R}^{2n}$ given by

$$\begin{aligned}
\mathcal{D}'_{Eul-Lagr} \;=\; & \{(q, v; u, a) \in T\mathbb{R}^{2n} \mid (q, v) \in W \times \mathbb{R}^n, \\
& u^h = v^h, \\
& \frac{\partial^2 L}{\partial v^k \partial v^h}\Big|_{(q,v)} a^k + \frac{\partial^2 L}{\partial q^k \partial v^h}\Big|_{(q,v)} v^k - \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = 0, \\
& \forall\, h = 1, \ldots, n\,\}
\end{aligned}$$

$\mathcal{D}_{Eul-Lagr}$ (or $\mathcal{D}'_{Eul-Lagr}$) is called *Euler-Lagrange differential equation* locally associated with L.

The above considerations prove that, for a conservative system, Theorem 12 can be reformulated as follows.

Theorem 14 *A smooth motion* $\mathrm{p} = \mathrm{p}(t) \in \mathcal{U}$ *is a solution of* $\mathcal{D}_{d'Al}$, *iff its coordinate expression* $q = q(t) \in W$ *is a solution of* $\mathcal{D}_{Eul-Lagr}$ *or, equivalently, the first projection of a solution* $(q, v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ *of* $\mathcal{D}'_{Eul-Lagr}$.

Clearly, $\mathcal{D}_{Eul-Lagr}$ (resp. $\mathcal{D}'_{Eul-Lagr}$) is reducible to normal form on W (resp. $W \times \mathbb{R}^n$) owing to the non-singularity of matrix $\left(\frac{\partial^2 L}{\partial v^k \partial v^h}\Big|_{(q,v)} \right)$, which does not differ from the analogous matrix of partial derivatives of K.

A kind of dynamical problem which – owing to its local character – can naturally be treated in coordinate formalism, is that of determining the 'small motions' that are dynamically possible around a stable equilibrium configuration.

For instance, let $\mathrm{p}_* = \xi(0) \subset Q$ be a point where

$$V(0) = 0\,, \quad \frac{\partial V}{\partial q^h}\Big|_0 = 0$$

and the symmetric bilinear form ω on $T_{\mathrm{p}_*}Q$ of components (in ξ)

$$\omega_{hk} := \omega\left(\frac{\partial \mathrm{p}}{\partial q^h}\Big|_0, \frac{\partial \mathrm{p}}{\partial q^k}\Big|_0 \right) := \frac{1}{2} \frac{\partial^2 V}{\partial q^h \partial q^k}\Big|_0$$

is positive definite. Owing to Dirichlet's criterion, p_* is a stable equilibrium configuration.

Along a dynamically possible small motion around p_*, the coordinate expression $q = q(t)$ and its derivative $v = v(t)$ can be treated as first-order infinitesimals, and therefore can be thought of as satisfying 'linearized' Euler-Lagrange equations (where all of the higher-order infinitesimals have been cancelled).

In order to work out such linearized equations, it will suffice to approximate the Lagrangian function to the second order (its derivatives will then be first order), i.e.

$$L(q,v) = \frac{1}{2}(g_{hk}\, v^h v^k - \omega_{hk}\, q^h q^k)$$

(where $g_{hk} := g_{hk}(0)$ are the components of g_{p_*}), or

$$L(q,v) = \frac{1}{2}(\delta_{hk}\, v^h v^k - \omega_h^2\, \delta_{hk}\, q^h q^k)$$

in *normal coordinates* at p_* (to which there corresponds, in $T_{\mathrm{p}_*}Q$, a g_{p_*}-orthonormal basis of eigenvectors of ω, whose positive eigenvalues are denoted by ω_h^2 or $(\omega^h)^2$). [7]

The linear Euler-Lagrange equations associated with the above approximated Lagrangian are the *harmonic oscillators*

$$\ddot{q}^h(t) + (\omega^h)^2 q^h(t) = 0$$

which characterize a dynamically possible small motion as a composition $q(t) = q^h(t)\,\delta_h$ *of harmonic oscillations*

$$q^h(t) = r^h \cos(\omega^h t + \varphi^h)$$

with *frequencies* $\omega_h/2\pi$, where $\omega_h := \sqrt{\omega_h^2}$ (the non-negative *amplitudes* r^h and the *phases* φ^h being uniquely determined by initial conditions).

The harmonic oscillation with the lowest frequency is said to be the *fundamental note*, all of the others are called *higher harmonics*.

As to the global dynamics of a conservative system, from Theorem 13 and Theorem 14 we infer what follows.

Theorem 15 *An admissible motion is a solution of* $\mathcal{D}_{d'Al}$*, iff it is a* geodesic curve *of* (Q, L)*, i.e. a smooth curve of* Q *admitting, near each point of its orbit, a (local) coordinate expression which is a solution of the equation* $\mathcal{D}_{Eul-Lagr}$ *locally associated with* L.

That is exactly what happens to the inertial motions of any mechanical system $\mathcal{S}$, which just turn out to be the geodesic curves of (Q, K). [8]

[7] See Appendix, section 4.2.1, *Eigenvalues and eigenvectors*.

[8] If Q is a spherical surface (like the surface of the Earth) and K is the kinetic energy

3.2 Hamilton

Through a suitable diffeomorphism, Euler-Lagrange equations will now be transformed into the celebrated 'Hamilton equations', which are of basic importance both in mathematical physics (for their being the core of all the modern developments in analytical and geometrical dynamics) and in theoretical physics (for their being the canonical springboard to 'quantization').

3.2.1 Legendre transformation

With reference to the coordinate formalism adopted for the local dynamics of a conservative system, we call (local) *Legendre transformation* the mapping

$$\mathcal{L} : W \times \mathbb{R}^n \to W \times \mathbb{R}^n : (q, v) \mapsto \mathcal{L}(q, v) = (q, p)$$

which leaves the Lagrangian coordinates $q \in W$ invariant and takes the *generalized velocities* $v = (v^h) \in \mathbb{R}^n$ onto the *kinetic momenta* $p = (p_h) \in \mathbb{R}^n$ given by

$$\begin{aligned} p_h &:= \left.\frac{\partial L}{\partial v^h}\right|_{(q,v)} \\ &= \left.\frac{\partial K}{\partial v^h}\right|_{(q,v)} \\ &= g_{hk}(q) v^k \end{aligned}$$

Owing to the non-singularity of matrix $(g_{hk}(q))$, Legendre transformation is a diffeomorphism and its inverse

$$\mathcal{L}^{-1} : W \times \mathbb{R}^n \to W \times \mathbb{R}^n : (q, p) \mapsto \mathcal{L}^{-1}(q, p) = (q, v)$$

will be denoted by

$$v^h = v^h(q, p)$$

Legendre diffeomorphism transforms the (local) mechanical energy

$$(q, v) \in W \times \mathbb{R}^n \mapsto E(q, v) = K(q, v) + V(q) \in \mathbb{R}$$

into the (local) *Hamiltonian function* given by [9]

---

of a unit mass (characterizing the Euclidean metric induced by E on each $T_{\mathrm{p}}Q$), the orbit of a non-degenerate geodesic ($\gamma\varepsilon\acute{\omega}\delta\eta\varsigma$) curve –literally, terrestial curve– can be shown to be a great circle of the sphere (like the meridians or other similar orbits on the Earth).

[9] Remark that $2\,K(q, v) = g_{hk}(q) v^h v^k = \left.\frac{\partial K}{\partial v^k}\right|_{(q,v)} v^k = \left.\frac{\partial L}{\partial v^k}\right|_{(q,v)} v^k = p_k v^k$.

$$\begin{aligned} H : (q,p) \in W \times \mathbb{R}^n \mapsto H(q,p) &:= E(\mathcal{L}^{-1}(q,p)) \\ &= E(q,v) \\ &= K(q,v) + V(q) \\ &= 2\,K(q,v) - L(q,v) \\ &= p_k\, v^k - L(q,v) \in \mathbb{R} \end{aligned}$$

whose partial derivatives are then given by

$$\frac{\partial H}{\partial q^h}\Big|_{(q,p)} = p_k\,\frac{\partial v^k}{\partial q^h}\Big|_{(q,p)} - \frac{\partial L}{\partial v^k}\Big|_{(q,v)}\,\frac{\partial v^k}{\partial q^h}\Big|_{(q,p)} - \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = -\frac{\partial L}{\partial q^h}\Big|_{(q,v)}$$

and

$$\frac{\partial H}{\partial p_h}\Big|_{(q,p)} = v^h + p_k\,\frac{\partial v^k}{\partial p_h}\Big|_{(q,p)} - \frac{\partial L}{\partial v^k}\Big|_{(q,v)}\,\frac{\partial v^k}{\partial p_h}\Big|_{(q,p)} = v^h$$

3.2.2 Hamilton differential equation

Legendre transformation shifts the problem of integrating $\mathcal{D}_{Eul-Lagr}$, onto the problem of integrating a normal differential equation associated with H, as follows.

Consider a smooth curve $t \in I \mapsto (q,v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ and its Legendre transformed $t \in I \mapsto (q,p) = (q(t), p(t)) \in W \times \mathbb{R}^n$.

The fact that the above curves (q,v) and (q,p) correspond to each other through $\mathcal{L}$, means

$$p_h = \frac{\partial L}{\partial v^h}\Big|_{(q,v)}, \quad v^h = v^h(q,p)$$

and then, along such curves, we have [10]

$$v^h = \frac{\partial H}{\partial p_h}\Big|_{(q,p)}, \quad \frac{\partial L}{\partial q^h}\Big|_{(q,v)} = -\frac{\partial H}{\partial q^h}\Big|_{(q,p)}$$

[10] See the derivatives of H evaluated at the end of section 3.2.1.

As a consequence, the curve (q, v) satisfies the system of Euler-Lagrange equations

$$\dot{q}^h = v^h \,, \qquad \frac{d}{dt}\frac{\partial L}{\partial v^h}\Big|_{(q,v)} = \frac{\partial L}{\partial q^h}\Big|_{(q,v)}$$

iff its Legendre transformed (q, p) satisfies the system of *Hamilton equations*

$$\dot{q}^h = \frac{\partial H}{\partial p_h}\Big|_{(q,p)} \,, \qquad \dot{p}_h = -\frac{\partial H}{\partial q^h}\Big|_{(q,p)}$$

Hamilton equations are the conditions characterizing the solutions $(q, p) = (q(t), p(t)) \in W \times \mathbb{R}^n$ of the time-independent, first-order, differential-algebraic equation in *normal form* on $\mathbb{R}^{2n}$ given by the graph

$$\begin{aligned} \mathcal{D}'_{Ham} &= \operatorname{Graph}(\Gamma_H) \\ &= \{(q, p; v, w) \in T\mathbb{R}^{2n} \mid (q, p) \in W \times \mathbb{R}^n \,,\ (v.w) = \Gamma_H(q, p)\} \end{aligned}$$

of the (local) *Hamiltonian vector field*

$$\Gamma_H : (q, p) \in W \times \mathbb{R}^n \subset \mathbb{R}^{2n} \mapsto \Gamma_H(q, p) := \left(\frac{\partial H}{\partial p_h}\Big|_{(q,p)} , -\frac{\partial H}{\partial q^h}\Big|_{(q,p)}\right) \in \mathbb{R}^{2n}$$

that is,

$$\begin{aligned} \mathcal{D}'_{Ham} = \{&(q, p; v, w) \in T\mathbb{R}^{2n} \mid (q, p) \in W \times \mathbb{R}^n, \\ &v^h = \frac{\partial H}{\partial p_h}\Big|_{(q,p)} ,\ w_h = -\frac{\partial H}{\partial q^h}\Big|_{(q,p)} \\ &\forall\, h = 1, \ldots, n\,\} \end{aligned}$$

$\mathcal{D}'_{Ham}$ is called *Hamilton differential equation* associated with H.

The above considerations prove the following

Theorem 16 *A smooth curve* $(q, v) = (q(t), v(t)) \in W \times \mathbb{R}^n$ *ia s solution of* $\mathcal{D}'_{Eul-Lagr}$, *iff its Legendre transformed* $(q, p) = (q(t), p(t)) \in W \times \mathbb{R}^n$ *is a solution of* $\mathcal{D}'_{Ham}$.

Clearly, along a solution $(q, p) = (q(t), v(t)) \in W \times \mathbb{R}^n$ of $\mathcal{D}'_{Ham}$, we have

$$\begin{aligned} \frac{d}{dt}\, H(q, p) &= \frac{\partial H}{\partial q^h}\Big|_{(q,p)} \dot{q}^h + \frac{\partial H}{\partial p_h}\Big|_{(q,p)} \dot{p}_h \\ &= \frac{\partial H}{\partial q^h}\Big|_{(q,p)} \frac{\partial H}{\partial p_h}\Big|_{(q,p)} - \frac{\partial H}{\partial p_h}\Big|_{(q,p)} \frac{\partial H}{\partial q^h}\Big|_{(q,p)} \\ &= 0 \end{aligned}$$

whence, by a quadrature, we obtain the following result (which corresponds to the conservation law of mechanical energy):

Theorem 17 *Along any solution* $(q,p) = (q(t),p(t)) \in W \times \mathbb{R}^n \subset \mathbb{R}^{2n}$ *of* $\mathcal{D}'_{Ham}$ *, determined by initial conditions* $q(t_o) = q_o$ *and* $p(t_o) = p_o$ *, the Hamiltonian function keeps constant, that is,* $H(q,p) = c_o$ *or* $(q,p) \in H^{-1}(c_o)$ *with* $c_o := H(q_o,p_o)$ *.*

The above conservation law just states that H is a first integral of $\mathcal{D}'_{Ham}$, that is, each solution of Hamilton equations lies on one of the *level subspaces* $\{H^{-1}(c_o)\}_{c_o \in \mathbb{R}}$ of H.

Remark that, for any $c_o \in \mathbb{R}$ and $(q_o,p_o) \in H^{-1}(c_o)$, the vector $\Gamma_H(q_o,p_o)$ is *tangent* to $H^{-1}(c_o)$ at (q_o,p_o) and points towards the *future*, since it coincides with the time-derivative $(\dot{q}(t_o),\dot{p}(t_o))$ of a solution of $\mathcal{D}'_{Ham}$ lying on $H^{-1}(c_o)$.

The level subspaces of H and the time-orientation of Γ_H can be useful devices for a qualitative analysis of the solutions of Hamilton equations.

A classical example now follows.

Example In the case $n = 1$, the *level lines* $\{H^{-1}(c_o)\}_{c_o \in \mathbb{R}}$, time-oriented by Γ_H, directly provide a qualitative *portrait* of the solutions of Hamilton equation.

Consider, for instance, the typical Hamiltonian function (arising in quite a number of concrete dynamical problems) defined by

$$(q,p) \in W \times \mathbb{R} \mapsto H(q,p) := \frac{1}{2\mu}\, p^2 + V(q) \in \mathbb{R}$$

where W is an open interval of $\mathbb{R}$, $\mu > 0$ and $V : W \to \mathbb{R}$.

Each level line $H^{-1}(c_o)$ is then described by the algebraic equation

$$p = \pm\sqrt{2\mu(c_o - V(q))}$$

and is time-oriented by the tangent vectors

$$\Gamma_H(q,p) = \left(\frac{1}{\mu}\, p,\ -\frac{dV}{dq}\Big|_q\right)$$

A qualitative description of $H^{-1}(c_o)$ can directly be obtained from the graph of V, since, owing to the above algebraic equation,

(i) $H^{-1}(c_o)$ is contained in $\{(q,p) \in W \times \mathbb{R} \mid V(q) \le c_o\}$

(ii) $H^{-1}(c_o) \ni (q,0)$, iff $V(q) = c_o$

(iii) $H^{-1}(c_o)$ generally consists of two branches, symmetric to each other with respect to the 'q-axis', along which $|p|$ will be increasing or decreasing according to whether $V(q)$ is decreasing or increasing, respectively.

Moreover, a solution $t \in I \mapsto (q,p) = (q(t), p(t)) \in W \times \mathbb{R}$ of Hamilton equations

$$\dot{q} = \frac{1}{\mu} p \, , \quad \dot{p} = -\frac{dV}{dq}\Big|_q$$

lying on $H^{-1}(c_o)$

$$p = \pm\sqrt{2\mu(c_o - V(q))}$$

satisfies

$$\dot{q} = \pm\sqrt{\frac{2}{\mu}(c_o - V(q))}$$

So, as far as $V(q) < c_o$, one has $\dot{q} \neq 0$ and then $q = q(t)$ is invertible with inverse $t = t(q)$ satisfying

$$\Big(\frac{dt}{dq}\Big)_q = \Big(\frac{dq}{dt}\Big)^{-1}_{t=t(q)} = \pm\frac{1}{\sqrt{\frac{2}{\mu}(c_o - V(q))}}$$

whence, by a quadrature, we obtain the time interval

$$t(q_1) - t(q_o) = \pm\int_{q_o}^{q_1} \frac{1}{\sqrt{\frac{2}{\mu}(c_o - V(q))}}\, dq$$

corresponding to the 'travel' from q_o to q_1.

Hence information about the completeness or incompleteness (in the past and in the future) of the solution, can be obtained.

Chapter 4

Appendix: mathematical methods

This Appendix provides the mathematical background for our course. [1]

Historical (Newtonian-d'Alembertian) dynamics basically rests on 'intrinsic' techniques of differential calculus in the context of Euclidean geometry, whereas analytical (Lagrangian-Hamiltonian) dynamics traditionally adopts the 'coordinate' techniques of real analysis. We put great emphasis on the intrinsic techniques, stressing, on the one hand, the geometric structures on which they are based and recovering, on the other hand, the coordinate techniques as a specialization of the above general calculus.

So we start from *affine geometry* (section 4.1), where the structure of an 'affine space' is shown to synthesize all of the linear properties of Euclidean geometry (i.e. properties concerning the connection between 'points' and 'vectors'). Then we continue with *metric geometry* (section 4.2), where the structure of a 'Euclidean affine space' is shown to allow the treatment of the metric properties as well (i.e. properties concerning kinds of measure such as 'length' and 'angle', 'distance' and 'nearness'). The metric notion of 'nearness' in a Euclidean space naturally leads to *topology* (section 4.3), on which the fundamental concepts of 'limit' and 'continuity' of a function are based. Then *differentiation* (section 4.4) on Euclidean spaces arises –in the spirit of the infinitesimal analysis– as a process of linear approximation, which aims at describing the behaviour of a function, in an 'infinitesimal neighbourhood' of a point of its domain, by means of a linear mapping (its 'differential' at that point). The above calculus in turn provide the tools for further geometric investigation, concerning the differential-topological properties of *manifolds* (section 4.5), meant as a kind of 'well-behaved' subspaces of a Euclidean space, including familiar 'loci' such as smooth curves and surfaces which 'infinitesi-

[1] Fundamentals of vector algebra and real analysis are the only prerequisites.

mally' resemble straight lines and planes. A 'manifold-like' approach is finally sketched to *differential equations* on a Euclidean space (section 4.6), which will prove (in the main text) to be the typical mathematical model of a 'time-evolution law' in Classical Dynamics.

4.1 Affine geometry

The structural treatment of the *linear* concepts of Euclidean geometry (i.e. concepts concerning the connection between 'points' and 'vectors') gives rise to the category of affine spaces, whose objects and morphisms will now be introduced.

4.1.1 Affine spaces

In the category of affine spaces, the objects –sets equipped with affine structure– are defined and behave as follows.

Affine structure

An *affine space* $\mathcal{E}$, modelled on a vector space E, [2] is a non-empty set on which the vector space acts freely and transitively as an Abelian group of operators.

Such a 'compact' definition will now be illustrated.

The above mentioned *action* of E on $\mathcal{E}$ is a *plus mapping*

$$+ : \mathcal{E} \times E \longrightarrow \mathcal{E} : (\mathrm{p}, \mathrm{v}) \mapsto \mathrm{q} = \mathrm{p} + \mathrm{v}$$

which is required to satisfy –for all $\mathrm{p} \in \mathcal{E}$ and $\mathrm{v}, \mathrm{w} \in E$– the following properties: [3]

1. $(\mathrm{p} + \mathrm{v}) + \mathrm{w} = \mathrm{p} + (\mathrm{v} + \mathrm{w})$, $\mathrm{p} + 0 = \mathrm{p}$
2. $\mathrm{v} \neq 0 \implies \mathrm{p} + \mathrm{v} \neq \mathrm{p}$
3. $\mathcal{E} = \mathrm{p} + E := \{\mathrm{p} + \mathrm{v}\}_{\mathrm{v} \in E}$

[2] We shall only consider finite-dimensional vector spaces on the field $\mathbb{R}$ of real numbers (also called *scalars*). In the sequel, the symbol 0 will denote both the zero of $\mathbb{R}$ and the zero of any vector space.

[3] Such properties are naturally suggested by empirical geometry, where any vector v (represented by an oriented segment with an arbitrary origin) takes its origin p to its end-point $\mathrm{p} + \mathrm{v}$ and the sum of vectors is defined by the parallelogram rule.

The *dimension* of $\mathcal{E}$ is defined by putting

$$\dim(\mathcal{E}) := \dim(E)$$

We shall always assume

$$m := \dim(\mathcal{E}) > 0$$

(otherwise, $\mathcal{E}$ would degenerate into a singleton).

$\mathcal{E}$ is said to be *oriented*, if so is E. [4]

Some comments, motivating the 'action' terminology above introduced, will now follow.

Translations

Each vector $\mathrm{v} \in E$ acts on $\mathcal{E}$ as an operator – called *translation* – through the mapping

$$+\mathrm{v} : \mathcal{E} \longrightarrow \mathcal{E} : \mathrm{p} \mapsto \mathrm{q} = \mathrm{p} + \mathrm{v}$$

Owing to property 1, the composition of any two translations $+\mathrm{v}$ and $+\mathrm{w}$ is the translation $+(\mathrm{v}+\mathrm{w})$ (and is therefore commutative, as well as associative) and $+0$ is the neutral element $\mathrm{id}_{\mathcal{E}}$ (as a consequence, any translation $+\mathrm{v}$ turns out to be a bijection, whose inverse is $+(-\mathrm{v})$). [5] So the set of translations, equipped with the operation of composition, is an Abelian group.

E is then said to act on $\mathcal{E}$ as an *Abelian group of operators*.

Owing to property 2, a non-null translation does not admit fixed points.

In such a case, the action of E on $\mathcal{E}$ is said to be *free*.

Owing to property 3, the *orbit* $\mathrm{p}+E$ of any point $\mathrm{p} \in \mathcal{E}$ under the action of E coincides with the whole space $\mathcal{E}$.

In such a case, the action of E on $\mathcal{E}$ is said to be *transitive*.

The affine space $\mathcal{E}$ is said to be *modelled* on the vector space E (in the sense that $\mathcal{E}$ exhibits an 'affinity' with E), owing to the following

---

[4] Recall that two bases Φ, Φ' of E (meant as linear isomorphisms of $\mathbb{R}^m$ onto E) are said to have the same orientation, if the linear automorphism $\Phi^{-1} \circ \Phi'$ of $\mathbb{R}^m$ has a positive determinant. E is then said to be oriented, if it is equipped with an orientation, i.e. one of the two equivalence classes determined by the above equivalence relation in the set of its bases. Any $\mathbb{R}^n$, meant as a vector space, will be thought of as equipped with the orientation determined by its natural basis.

[5] We shall put $\mathrm{p} - \mathrm{v} := \mathrm{p} + (-\mathrm{v})$.

Proposition 1 *Any 'origin'* $\mathrm{p} \in \mathcal{E}$ *determines a bijection of* E *onto* $\mathcal{E}$ *, given by*

$$\mathrm{p}+ : E \longrightarrow \mathcal{E} : \mathrm{v} \mapsto \mathrm{q} = \mathrm{p} + \mathrm{v}$$

Proof: Injectivity follows from properties 1 and 2, since $\mathrm{p} + \mathrm{v} = \mathrm{p} + \mathrm{w}$ implies $(\mathrm{p}+\mathrm{v})-\mathrm{w} = (\mathrm{p}+\mathrm{w})-\mathrm{w}$, that is, $\mathrm{p}+(\mathrm{v}-\mathrm{w}) = \mathrm{p}+(\mathrm{w}-\mathrm{w}) = \mathrm{p}+0 = \mathrm{p}$ and then $\mathrm{v} = \mathrm{w}$. Surjectivity follows from property 3, which, for any point $\mathrm{q} \in \mathcal{E}$, guarantees the existence of a vector $\mathrm{v} \in E$ s.t. $\mathrm{q} = \mathrm{p} + \mathrm{v}$. □

That amounts to saying that, for any two points p and q of $\mathcal{E}$, there exists a unique vector v of E s.t.

$$\mathrm{q} = \mathrm{p} + \mathrm{v}$$

Such a vector, *linking* p *to* q , will be denoted by

$$\mathrm{v} = \mathrm{q} - \mathrm{p}$$

So, to the affine plus mapping, there corresponds the *minus mapping*

$$- : \mathcal{E} \times \mathcal{E} \to E : (\mathrm{p}, \mathrm{q}) \mapsto \mathrm{v} = \mathrm{q} - \mathrm{p}$$

which satisfies the following properties:

Proposition 2
(1) *For any* $\mathrm{p} \in \mathcal{E}$ *, the mapping*

$$-\,\mathrm{p} : \mathcal{E} \to E : \mathrm{q} \mapsto \mathrm{v} = \mathrm{q} - \mathrm{p}$$

is bijective.
(2) *For any triplet* $\mathrm{o}, \mathrm{p}, \mathrm{q} \in \mathcal{E}$ *,*

$$\mathrm{q} - \mathrm{p} = (\mathrm{q} - \mathrm{o}) + (\mathrm{o} - \mathrm{p})$$

Proof: The first property is due to the fact that

$$-\mathrm{p} = (\mathrm{p}\,+)^{-1}$$

The second property is due to

$$\mathrm{p} + ((\mathrm{q} - \mathrm{o}) + (\mathrm{o} - \mathrm{p})) = (\mathrm{p} + (\mathrm{o} - \mathrm{p})) + (\mathrm{q} - \mathrm{o}) = \mathrm{o} + (\mathrm{q} - \mathrm{o}) = \mathrm{q}$$

□

Exercise 1 *Any minus mapping satisfying properties* (1) *and* (2)*, corresponds to one, and only one, plus mapping defining an affine structure.*

Exercise 2 *For all* $\mathrm{p}, \mathrm{q} \in \mathcal{E}$ *and* $\mathrm{v} \in E$ *, the following simple 'algebraic' rules hold:*

$$\mathrm{p} - \mathrm{p} = 0\,, \quad \mathrm{q} - \mathrm{p} = -(\mathrm{p} - \mathrm{q})\,, \quad (\mathrm{q} + \mathrm{v}) - \mathrm{p} = (\mathrm{q} - \mathrm{p}) + \mathrm{v} = \mathrm{q} - (\mathrm{p} - \mathrm{v})$$

Product affine space

Exercise 3 *The Cartesian product*

$$\mathcal{E} := \mathcal{E}_1 \times \cdots \times \mathcal{E}_\nu$$

of affine spaces $\mathcal{E}_1, \dots, \mathcal{E}_\nu$ *–modelled on vector spaces* $E_1, \dots, E_\nu$ *, respectively– turns into an affine space modelled on the product vector space or direct sum*

$$E := E_1 \times \cdots \times E_\nu$$

(whose dimension is the sum of the dimensions), if the latter is let to act on the former by putting

$$\mathrm{p} + \mathrm{v} := (\mathrm{p}_1 + \mathrm{v}_1, \dots, \mathrm{p}_\nu + \mathrm{v}_\nu)\,, \quad \forall\, \mathrm{p} = (\mathrm{p}_1, \dots, \mathrm{p}_\nu) \in \mathcal{E}\,,\ \mathrm{v} = (\mathrm{v}_1, \dots, \mathrm{v}_\nu) \in E$$

whence

$$\mathrm{q} - \mathrm{p} = (\mathrm{q}_1 - \mathrm{p}_2, \dots, \mathrm{q}_\nu - \mathrm{p}_\nu)\,, \quad \forall\, \mathrm{p} = (\mathrm{p}_1, \dots, \mathrm{p}_\nu) \in \mathcal{E}\,,\ \mathrm{q} = (\mathrm{q}_1, \dots, \mathrm{q}_\nu) \in \mathcal{E}$$

Vector affine space

Exercise 4 *A vector space V turns into an affine space modelled on itself, if the* vector space $E := V$ *is thought of as acting on the* set $\mathcal{E} := V$ *through the sum operation of* V. [6]

4.1.2 Affine morphisms

In the category of affine spaces, the morphisms –mappings 'compatible' with the structure ($\mu\, o\, \varrho\, \varphi\, \acute{\eta}$) of the objects– are defined and behave as follows.

Affine mappings

Let $\mathcal{A}$ and $\mathcal{E}$ be affine spaces (modelled on vector spaces A and E, respectively).

An *affine morphism* –or *affine mapping*– of $\mathcal{A}$ in $\mathcal{E}$, [7] is a mapping $f : \mathcal{A} \to \mathcal{E}$ satisfying, for one (and then every) $x_o \in \mathcal{A}$ and all $v \in A$,

$$f(x_o + v) = f(x_o) + F(v)$$

where $F : A \to E$ is a (uniquely determined) linear mapping, called the *linear part* of f. [8]

---

[6] In particular, for all $n \geq 1$, $\mathbb{R}^n$ can be viewed both as a vector space and as an affine space.

[7] We say *affine endomorphism*, if $\mathcal{A} = \mathcal{E}$.

[8] The linear mappings are the morphisms of the category of vector (or linear) spaces.

Remarkable are the following two characterizations of the affine mappings:

Exercise 5
(i) *$f : \mathcal{A} \to \mathcal{E}$ is an affine mapping with linear part $F : A \to E$, iff, for any given $x_o \in \mathcal{A}$ and all $x \in \mathcal{A}$,*

$$f(x) - f(x_o) = F(x - x_o)$$

(i.e. the increment $f(x) - f(x_o)$ of f is a linear function of the increment $x - x_o$ given to the value x_o of its argument).
(ii) *$f : \mathcal{A} \to \mathcal{E}$ is an affine mapping with linear part $F : A \to E$, iff, for all $x \in \mathcal{A}$,*

$$f(x) = \mathrm{p}_o + F(x - x_o)$$

with $x_o \in \mathcal{A}$ and $\mathrm{p}_o \in \mathcal{E}$ (i.e. f is determined by its value $f(x_o) = \mathrm{p}_o$ at a given point x_o plus a linear function F of the increment given to the value x_o of its argument).

Simple properties of affine mappings :

Exercise 6
(i) *If f and g are two composable affine mappings with linear parts F and G, respectively, then the composite $g \circ f$ is an affine mapping with linear part $G \circ F$.*
(ii) *An affine mapping is injective (or surjective), iff so is its linear part.*

Now check the following examples of affine mappings:

Exercise 7
(i) *A* constant mapping *between affine spaces is an affine mapping, whose linear part vanishes.*
(ii) *A* linear mapping *between vector spaces is an affine mapping, coinciding with its own linear part.*
(iii) *The* projection *of $\mathcal{E} := \mathcal{E}_1 \times \cdots \times \mathcal{E}_\nu$ onto its i-th factor $\mathcal{E}_i$ is an affine mapping, whose linear part is the projection of $E := E_1 \times \cdots \times E_\nu$ onto its i-th factor E_i.*
(iv) *An* affine endomorphism of $\mathbb{R}^n$ *is a mapping*

$$f : x \in \mathbb{R}^n \mapsto y = y_o + a\,(x - x_o) \in \mathbb{R}^n$$

with $y_o \in \mathbb{R}^n$ and $a \in gl(n, \mathbb{R})$. [9]

[9] $gl(n, \mathbb{R})$ denotes the set of all the $n \times n$ real matrixes (i.e. the linear endomorphisms of $\mathbb{R}^n$) and $a\,(x - x_o)$ denotes 'row by column' multiplication.

Affine isomorphisms

An *affine isomorphism* of $\mathcal{A}$ onto $\mathcal{E}$, [10] is a bijective affine mapping $f : \mathcal{A} \to \mathcal{E}$ (an affine isomorphism between oriented affine spaces is said to *preserve the orientation*, if so does its linear part). [11]

Exercise 8
(i) *An affine mapping is an isomorphism, iff so is its linear part.*
(ii) *If* f *is an affine isomorphism with linear part* F, *then* f^{-1} *is an affine isomorphism with linear part* F^{-1}.

Examples of affine isomorphisms:

Exercise 9
(i) *For any* $\mathrm{p} \in \mathcal{E}$, $\mathrm{p}+ : E \to \mathcal{E}$ *(and then* $-\mathrm{p} : \mathcal{E} \to E$*) is an affine isomorphism, whose linear part is the identity mapping* id_E.
(ii) *For any* $\mathrm{v} \in E$, $+\mathrm{v} : \mathcal{E} \to \mathcal{E}$ *is an affine automorphism, whose linear part is the identity mapping* id_E.
(iii) *An affine automorphism of* $\mathbb{R}^n$ *is a mapping*

$$f : x \in \mathbb{R}^n \mapsto y = y_o + a\,(x - x_o) \in \mathbb{R}^n$$

with $y_o \in \mathbb{R}^n$ *and* $a \in Gl(n, \mathbb{R})$. [12]

Cartesian coordinates

In an m-dimensional affine space $\mathcal{E}$ (modelled on E), a *Cartesian system* is an affine isomorphism

$$\phi : \mathbb{R}^m \to \mathcal{E}$$

It is completely determined by an origin in $\mathcal{E}$ and a basis of E, namely

$$\mathrm{o} = \phi(0)$$

and, if $\Phi : \mathbb{R}^m \to E$ denotes the linear part of ϕ, [13]

$$\mathrm{e}_i = \Phi(\delta_i), \quad i = 1, \ldots, m$$

---

[10] We say *affine automorphism*, if $\mathcal{A} = \mathcal{E}$.

[11] A linear mapping $F : A \to E$ between oriented vector spaces is said to preserve the orientation, if it takes the orientation of A onto that of E.

[12] $Gl(n, \mathbb{R}) \subset gl(n, \mathbb{R})$ denotes the group of all the non-singular $n \times n$ real matrixes (i.e. the linear automorphisms of $\mathbb{R}^n$).

[13] $\delta = (\delta_i)_{i=1,\ldots,m}$ – with $\delta_i = (\delta_i^j)_{j=1,\ldots,m}$ (δ_i^j being Kronecker's symbol) – is the natural basis of $\mathbb{R}^m$.

With reference to such a Cartesian system, any point

$$\mathrm{p} \in \mathcal{E}$$

can be epressed as an invertible affine function [14]

$$\begin{aligned} \mathrm{p} &= \phi(x) \\ &= \phi(0) + \Phi(x) \\ &= \phi(0) + \Phi(x^i \delta_i) \\ &= \phi(0) + x^i \Phi(\delta_i) \\ &= \mathrm{o} + x^i \mathrm{e}_i \end{aligned}$$

of its *Cartesian coordinates*

$$(x_i)_{i=1,\ldots,m} = x = \phi^{-1}(\mathrm{p}) \in \mathbb{R}^m$$

4.2 Metric geometry

A complete study of Euclidean geometry, including the structural treatment of its *metric* concepts (i.e. concepts concerning kinds of measure $-\mu \acute{\epsilon} \tau \varrho o \nu-$ such as 'length' or 'distance' and 'angle') gives rise to the more special category of Euclidean affine spaces, whose objects and morphisms will now be presented.

4.2.1 Euclidean affine spaces

In the category of Euclidean affine spaces, the objects –sets equipped with Euclidean affine structure– are defined and behave as follows.

Euclidean metric

A *Euclidean affine space* is an affine space $\mathcal{E}$ modelled on a Euclidean vector space E, i.e. a vector space equipped with a *Euclidean metric*.

We recall that a Euclidean metric on E is a positive definite, symmetric, bilinear form, i.e. a mapping

$$g : E \times E \to \mathbb{R} : (\mathrm{u}, \mathrm{v}) \mapsto g(\mathrm{u}, \mathrm{v})$$

satisfying, for all $\mathrm{u}, \mathrm{v}, \mathrm{w} \in E$ and $a, b \in \mathbb{R}$,

[14] Summation symbol over an index will be understood, whenever the index appears both in upper and in lower position.

$$g(\mathrm{u},\mathrm{v}) = g(\mathrm{v},\mathrm{u})$$

(*symmetry*),

$$g(\mathrm{u}, a\,\mathrm{v} + b\,\mathrm{w}) = a\,g(\mathrm{u},\mathrm{v}) + b\,g(\mathrm{u},\mathrm{w})$$

(*bilinearity*),

$$\mathrm{u} \neq 0 \Rightarrow g(\mathrm{u},\mathrm{u}) > 0$$

(*positive definiteness*).

The last property implies

$$g(\mathrm{u},\mathrm{v}) = 0\,,\ \forall\,\mathrm{v} \in E \Longrightarrow \mathrm{u} = 0$$

(*non-degenerateness*). [15]

As a consequence, the correspondence law

$$\mathrm{u} \mapsto g(\mathrm{u},\ \ )$$

is a linear mapping of E in E^* [16] with a trivial kernel and then, for dimensional reasons, is a (linear) isomorphism.

For any two vectors $\mathrm{u},\mathrm{v} \in E$, the real number

$$\mathrm{u}\cdot\mathrm{v} := g(\mathrm{u},\mathrm{v})$$

will be called *the scalar product* of u and v.

With the above notation for the scalar product, the properties of symmetry, bilinearity, positive definiteness and non-degenerateness read

$$\mathrm{u}\cdot\mathrm{v} = \mathrm{v}\cdot\mathrm{u}$$

$$\mathrm{u}\cdot(a\,\mathrm{v} + b\,\mathrm{w}) = a\,\mathrm{u}\cdot\mathrm{v} + b\,\mathrm{u}\cdot\mathrm{w}$$

$$\mathrm{u} \neq 0 \Rightarrow \mathrm{u}\cdot\mathrm{u} > 0$$

$$\mathrm{u}\cdot\mathrm{v} = 0\,,\ \forall\,\mathrm{v} \in E \Longleftrightarrow \mathrm{u} = 0$$

and, putting $\mathrm{u}\cdot\ := g(\mathrm{u},\ \ )$, the induced linear isomorphism of E onto E^* reads

$$\mathrm{u} \in E \mapsto \mathrm{u}\cdot\ \in E^*$$

---

[15] The converse $\mathrm{u} = 0 \Longrightarrow g(\mathrm{u},\mathrm{v}) = 0\,,\ \forall\,\mathrm{v} \in E$ holds true owing to bilinearity.

[16] Recall that E^* denotes the set of all the linear mappings of E in $\mathbb{R}$, called *linear forms* on E (e.g., for any $\mathrm{v} \in E$,

$$g(\mathrm{u},\ \ ) : \mathrm{v} \in E \mapsto g(\mathrm{u},\mathrm{v}) \in \mathbb{R}$$

is a linear form on E). As E^* can be endowed with natural operations of sum and multiplication by scalars, it turns out to exhibit the structure of a vector space, called the *dual space* of E, whose dimension is the same as that of E.

The main metric concept arising from the scalar product is that of *Euclidean norm* or *Euclidean length* $|\mathrm{u}|$ of a vector $\mathrm{u} \in E$, defined by putting

$$|\mathrm{u}| := \sqrt{\mathrm{u}^2}$$

with

$$\mathrm{u}^2 := \mathrm{u} \cdot \mathrm{u}$$

The scalar product itself can be expressed in terms of the *quadratic form* defined by the semisquare of the norm $\frac{1}{2}|\ \ |^2$, owing to the *polarization identity*

$$\mathrm{u} \cdot \mathrm{v} = 1/2\,(|\mathrm{u}+\mathrm{v}|^2 - |\mathrm{u}|^2 - |\mathrm{v}|^2|)$$

A *normal* vector u is one of unit length, i.e. $|\mathrm{u}| = 1$. [17]

Through the above definition, E is given the structure of a *normed vector space*, i.e. it is endowed with a *norm*

$$|\ \ | : E \to \mathbb{R} : \mathrm{u} \mapsto |\mathrm{u}|$$

satisfying, for all $\mathrm{u}, \mathrm{v} \in E$ and $a \in \mathbb{R}$,

$$\begin{aligned} &\mathrm{u} = 0 \Rightarrow |\mathrm{u}| = 0\,, \quad \mathrm{u} \neq 0 \Rightarrow |\mathrm{u}| > 0 \\ &|a\mathrm{u}| = |a||\mathrm{u}| \\ &|\mathrm{u}+\mathrm{v}| \leq |\mathrm{u}| + |\mathrm{v}| \end{aligned}$$

The above *triangle inequality* is equivalent (owing to the polarization identity) to *Schwarz inequality*

$$|\mathrm{u} \cdot \mathrm{v}| \leq |\mathrm{u}||\mathrm{v}|$$

which allows the metric concept of *angle* $\angle(\mathrm{u}, \mathrm{v}) \in [0, \pi]$ between two non-null vectors $\mathrm{u}, \mathrm{v} \in E$ to be defined by putting

$$\angle(\mathrm{u}, \mathrm{v}) := \arccos \frac{\mathrm{u} \cdot \mathrm{v}}{|\mathrm{u}||\mathrm{v}|}$$

The *orthogonality* condition $\angle(\mathrm{u}, \mathrm{v}) = \frac{\pi}{2}$ can therefore be expressed in the form $\mathrm{u} \cdot \mathrm{v} = 0$ (recall that, owing to non-degenerateness, the latter condition $\mathrm{u} \cdot \mathrm{v} = 0$ is fulfilled for all $\mathrm{v} \in E$, iff $\mathrm{u} = 0$).

The *orthogonal complement* of a vector subspace W of E is then the vector subspace given by $W^\perp := \{\mathrm{v} \in E \,|\, \mathrm{v} \cdot \mathrm{w} = 0\,,\ \forall\, \mathrm{w} \in W\}$ and fulfilling [18] $E = W \oplus W^\perp$ (in particular, owing to non-degenerateness, $E^\perp = \{0\}$).

[17] Under the hypothesis $\dim(E) > 0$, we obtain a normal vector $\mathrm{u} := \frac{\mathrm{v}}{|\mathrm{v}|}$ from any non-null vector $\mathrm{v} \in E$.

[18] Recall that $\oplus$ denotes the direct sum of vector spaces.

Exercise 10

(i) *If* $E_1, \ldots, E_\nu$ *are Euclidean vector spaces, so is* $E = E_1 \times \cdots \times E_\nu$ *by putting, for all* $\mathrm{u} = (\mathrm{u}_1, \ldots, \mathrm{u}_\nu)$ *and* $\mathrm{v} = (\mathrm{v}_1, \ldots, \mathrm{v}_\nu)$ *in* E *,*

$$\mathrm{u} \cdot \mathrm{v} = \sum_{i=1}^{\nu} \mathrm{u}_i \cdot \mathrm{v}_i$$

whence the norm

$$|\mathrm{u}| = \sqrt{\sum_{i=1}^{\nu} {\mathrm{u}_i}^2}$$

(ii) *A scalar product is naturally defined in* $\mathbb{R}$ *by the ordinary multiplication of real numbers; hence the natural scalar product in any Cartesian power* $\mathbb{R}^n$ *, given, for all* $u = (u^1, \ldots, u^n)$ *and* $v = (v^1, \ldots, v^n)$ *in* R^n *, by*

$$u \cdot v = \sum_{i=1}^{n} u^i \cdot v^i$$

with norm

$$|u| = \sqrt{\sum_{i=1}^{n} u^{i\,2}}$$

Eigenvalues and eigenvectors

Any basis of E can be trasformed into a (g-)*orthonormal* basis $(\mathrm{e}_h)_{h=1,\ldots,m}$, i.e. one where g is characterized by the diagonal matrix of *components* given by [19]

$$g_{hk} := \mathrm{g}(\mathrm{e}_h, \mathrm{e}_k) = \mathrm{e}_h \cdot \mathrm{e}_k = \delta_{hk}$$

A simultaneous diagonalization can be shown to happen to any linear endomorphism

$$F : E \to E$$

defining a bilinear form

$$\mathbb{F} : E \times E \to \mathbb{R} : (\mathrm{u}, \mathrm{v}) \mapsto \mathbb{F}(\mathrm{u}, \mathrm{v}) := F(\mathrm{u}) \cdot \mathrm{v}$$

symmetric and positive definite.

[19] Check that the natural basis of $\mathbb{R}^m$ is orthonormal (with respect to the natural Euclidean metric defined in Exercise 10 (ii)).

To this end, recall that a real number F_h is said to be an *eigenvalue* of F, iff there exists a non-zero vector e_h, called *eigenvector* corresponding to F_h, s.t. $F(\mathrm{e}_h) = F_h\,\mathrm{e}_h$. Then recall that, owing to symmetry and positive definiteness, F admits at most m distinct (positive) eigenvalues $(F_h)_{h=1,\dots,m}$ and that there exists a g-orthonormal basis $(\mathrm{e}_h)_{h=1,\dots,m}$ of eigenvectors of F orderly corresponding to the above eigenvalues. In such a basis, F is characterized by the diagonal matrix of components given by

$$F_{hk} := \mathbb{F}(\mathrm{e}_h, \mathrm{e}_k) = F(\mathrm{e}_h) \cdot \mathrm{e}_k = F_h\,\mathrm{e}_h \cdot \mathrm{e}_k = F_h\,\delta_{hk}$$

Euclidean distance

The Euclidean length in E naturally leads to the *Euclidean distance* $d(\mathrm{p}, \mathrm{q})$ from a point $\mathrm{p} \in \mathcal{E}$ to another $\mathrm{q} \in \mathcal{E}$, defined by putting

$$d(\mathrm{p}, \mathrm{q}) := |\mathrm{q} - \mathrm{p}|$$

So, owing to the properties of the norm, $\mathcal{E}$ is given the structure of a *metric space*, i.e. it is endowed with a distance function

$$d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} : (\mathrm{p}, \mathrm{q}) \mapsto d(\mathrm{p}, \mathrm{q})$$

satisfying, for all $\mathrm{o}, \mathrm{p}, \mathrm{q} \in \mathcal{E}$,

$$\mathrm{p} = \mathrm{q} \;\Rightarrow\; d(\mathrm{p}, \mathrm{q}) = 0\,, \quad \mathrm{p} \neq \mathrm{q} \;\Rightarrow\; d(\mathrm{p}, \mathrm{q}) > 0$$

$$d(\mathrm{q}, \mathrm{p}) = d(\mathrm{p}, \mathrm{q})$$

$$d(\mathrm{p}, \mathrm{q}) \leq d(\mathrm{p}, \mathrm{o}) + d(\mathrm{o}, \mathrm{q})$$

4.2.2 Euclidean affine morphisms

In the category of Euclidean affine spaces, the morphisms – mappings compatible with both the affine and the metric structure of the objects – are defined and behave as follows.

Euclidean affine mappings

Let $\mathcal{A}$ and $\mathcal{E}$ be Euclidean affine spaces (modelled on Euclidean vector spaces A and E, respectively).

A *Euclidean affine morphism* –or *Euclidean affine mapping*– of $\mathcal{A}$ in $\mathcal{E}$ is an affine mapping $f : \mathcal{A} \to \mathcal{E}$ which *preserves the metric*, i.e.

$$d(f(x), f(y)) = d(x, y) \ , \quad \forall\, x, y \in \mathcal{A}$$

Proposition 3 *An affine mapping* $f : \mathcal{A} \to \mathcal{E}$ *preserves the metric, iff its linear part* $F : A \to E$ preserves the norm*, i.e.* $|F(v)| = |v|, \forall\, v \in A$ *.*

Proof: Just note that the metric preserving condition reads $|f(y) - f(x)| = |y - x|$, that is, $|F(y - x)| = |y - x|$ for all $x, y \in \mathcal{A}$, whence our claim. [20] $\square$

Exercise 11 *A Euclidean affine mapping is injective.*

Affine isometries

A *Euclidean affine isomorphism* –or *affine isometry*– of $\mathcal{A}$ onto $\mathcal{E}$ is a bijective Euclidean affine mapping $f : \mathcal{A} \to \mathcal{E}$.

Exercise 12
(i) *An affine mapping is an isometry, iff so is its linear part.* [21]
(ii) *A Euclidean affine mapping is an isometry, iff its domain and codomain have the same dimension.*
(iii) *If an affine mapping is an isometry, so is its inverse.*

Exercise 13
(i) *A translation of* $\mathcal{E}$ *is an affine isometry.*
(ii) *An affine isometry of* $\mathbb{R}^n$ *is a mapping*

$$f : x \in \mathbb{R}^n \mapsto y = y_o + a\,(x - x_o) \in \mathbb{R}^n$$

with $y_o \in \mathbb{R}^n$ *and* $a \in O(n)$ *.* [22]

---

[20] Recall that, owing to the polarization identity, F preserves the norm, iff it *preserves the scalar product* (i.e. $F(u) \cdot F(v) = u \cdot v, \forall\, u, v \in A$).

[21] A linear isometry, between Euclidean vector spaces, is a norm preserving, linear isomorphism.

[22] $O(n) \subset Gl(n, \mathbb{R})$ denotes the group of all the orthogonal $n \times n$ real matrixes (i.e. the linear isometries of $\mathbb{R}^n$).

Orthogonal Cartesian coordinates

In an m-dimensional Euclidean affine space $\mathcal{E}$ (modelled on E), an *orthogonal Cartesian system* –giving *orthogonal Cartesian coordinates* to the points of $\mathcal{E}$– is an affine isometry

$$\phi : \mathbb{R}^m \to \mathcal{E}$$

(with linear part $\Phi : \mathbb{R}^m \to E$), which just corresponds to an origin $\mathrm{o} = \phi(0)$ and an *orthonormal* basis $\mathrm{e}_i = \Phi(\delta_i)$ $(i = 1, \ldots, m)$.

Equipollence

The primitive concepts of empirical geometry, where an attached vector is thought of as an 'oriented segment' and a free vector is conceived as a complete class of 'equipollent' oriented segments, can all be recovered from the structure of a Euclidean affine space $\mathcal{E}$.

Define an *oriented segment* $\overrightarrow{\mathrm{p}_0\mathrm{p}_1}$, with end-points $(\mathrm{p}_0, \mathrm{p}_1) \in \mathcal{E} \times \mathcal{E}$, by putting

$$\overrightarrow{\mathrm{p}_0\mathrm{p}_1} := \{\mathrm{p}(t) = \mathrm{p}_0 + t\,(\mathrm{p}_1 - \mathrm{p}_0)\}_{t\in[0,1]}$$

Clearly, $\mathrm{p}_0 = \mathrm{p}(0)$ and $\mathrm{p}_1 = \mathrm{p}(1)$ belong to $\overrightarrow{\mathrm{p}_0\mathrm{p}_1}$.

If $\mathrm{p}_0 = \mathrm{p}_1$, the oriented segment degenerates into a singleton.

If $\mathrm{p}_0 \neq \mathrm{p}_1$, the non-degenerate oriented segment lies on the *straight line* [23] obtained by translating p_0 along the direction of $\mathrm{p}_1 - \mathrm{p}_0$, [24] i.e.

$$\begin{aligned} \overrightarrow{\mathrm{p}_0\mathrm{p}_1} &\subset \mathrm{p}_0 + \mathrm{Span}(\mathrm{p}_1 - \mathrm{p}_0) \\ &= \{\mathrm{p}(t) = \mathrm{p}_0 + t\,(\mathrm{p}_1 - \mathrm{p}_0)\}_{t\in\mathbb{R}} \end{aligned}$$

and exhibits the ordering defined by putting

$$\mathrm{p}(t_0) \prec \mathrm{p}(t_1)$$

whenever

$$t_0 < t_1$$

that is,

$$\mathrm{p}(t_1) - \mathrm{p}(t_0) = (t_1 - t_0)(\mathrm{p}_1 - \mathrm{p}_0) \in \mathrm{Span}^+(\mathrm{p}_1 - \mathrm{p}_0)$$

[23] Straight lines are a particular case of 'affine subspaces' (see section 4.5.1).

[24] Recall that the *direction* of a non-zero vector $\mathrm{v} \in E$ is the 1-dimensional vector subspace of E spanned by v, i.e. $\mathrm{Span}(\mathrm{v}) := \{t\,\mathrm{v}\}_{t\in\mathbb{R}}$. The *orientation* determined by v on its own direction is then given by $\mathrm{Span}^+(\mathrm{v}) := \{t\,\mathrm{v}\}_{t\in\mathbb{R}^+}$ (where $\mathbb{R}^+ := (o, +\infty)$).

Now it appears quite natural to call $\mathrm{Span}^+(p_1 - p_0)$ the *oriented direction* of $\overrightarrow{p_0p_1}$ and $d(p_0, p_1) = |p_1 - p_0|$ the *length* of $\overrightarrow{p_0p_1}$.

Then, in the set

$$\Sigma := \{\overrightarrow{p_0p_1}\}_{(p_0,p_1)\in\mathcal{E}\times\mathcal{E}}$$

of all the oriented segments of $\mathcal{E}$, *equipollence* is the classical equivalence relation defined by putting

$$\overrightarrow{p_0p_1} \sim \overrightarrow{q_0q_1}$$

whenever both the oriented segments degenerate into singletons or otherwise have the same oriented direction and the same length, i.e.

$$\begin{array}{rcl} \mathrm{Span}^+(p_1 - p_0) & = & \mathrm{Span}^+(q_1 - q_0) \\ d(p_0, p_1) & = & d(q_0, q_1) \end{array}$$

Proposition 4 $\quad \overrightarrow{p_0p_1} \sim \overrightarrow{q_0q_1} \iff p_1 - p_0 = q_1 - q_0\,.$

Proof: The statement is trivial for degenerate oriented segments.

In the non-degenerate case, the first of the two equipollence condition, i.e. $\mathrm{Span}^+(p_1 - p_0) = \mathrm{Span}^+(q_1 - q_0)$, reads

$$p_1 - p_0 = t\,(q_1 - q_0)\,, \quad t > 0$$

and then the second condition, i.e. $|p_1 - p_0| = |q_1 - q_0|$, reads

$$t = 1$$

So the above two conditions amount to saying $p_1 - p_0 = q_1 - q_0$. □

The above proposition allows the quotient $\Sigma_{/\sim}$ to be taken into E by the mapping [25]

$$\tau : \Sigma_{/\sim} \to E : \ [\overrightarrow{p_0p_1}] \mapsto p_1 - p_0$$

Proposition 5 $\quad$ *τ is a bijection*

[25] Remark that, for any given $p_o \in \mathcal{E}$, every equivalence class $[\overrightarrow{q_0q_1}] \in \Sigma_{/\sim}$ can be expressed in the form $[\overrightarrow{q_0q_1}] = [\overrightarrow{p_0p_1}]$, with $p_1 = p_o + (q_1 - q_o)$.

Proof: On the one hand, if $\tau([\overrightarrow{p_0p_1}]) = \tau([\overrightarrow{q_0q_1}])$, i.e. $p_1 - p_0 = q_1 - q_0$, then $\overrightarrow{p_0p_1} \sim \overrightarrow{q_0q_1}$, i.e. $[\overrightarrow{p_0p_1}] = [\overrightarrow{q_0q_1}]$ (τ is injective).

On the other hand, if – for any $v \in E$ – we consider p_0 and $p_1 := p_0 + v \in \mathcal{E}$, we obtain $\tau([\overrightarrow{p_0p_1}] = p_1 - p_0 = v$ (τ is surjective). □

τ can be regarded as a canonical linear isometry which identifies the space $\Sigma_{/\sim}$ of *free* vectors with the space E of translations, if the former is given the Euclidean vector space structure defined by the *parallelogram rule*

$$\begin{aligned} [\overrightarrow{p_0p_1}] + [\overrightarrow{p_1p_2}] &:= \tau^{-1}((p_1 - p_0) + (p_2 - p_1)) \\ &= \tau^{-1}(p_2 - p_0) \\ &= [\overrightarrow{p_0p_2}] \end{aligned}$$

the *dilation rule*

$$\begin{aligned} t\,[\overrightarrow{p_0p_1}] &:= \tau^{-1}(t\,(p_1 - p_0)) \\ &= [\overrightarrow{p_0\,p(t)}] \end{aligned}$$

– with $p(t) = p_0 + t\,(p_1 - p_0) \subset p_0 + \mathrm{Span}(p_1 - p_0)$ – and the *scalar product*

$$[\overrightarrow{p_0p_1}] \cdot [\overrightarrow{p_1p_2}] := (p_1 - p_0) \cdot (p_2 - p_1)$$

In the same way, the bijection

$$\tau_{p_0} : \{p_0\} \times E \to \Sigma_{p_0} : \ (p_0, v) \mapsto \overrightarrow{p_0\,p_0 + v}$$

can be regarded as a canonical linear isometry which identifies the space $\{p_0\} \times E$ of translations starting from a given point p_0 with the space

$$\Sigma_{p_0} := \{\overrightarrow{p_0p_1}\}_{p_1 \in \mathcal{E}}$$

of the vectors *attached* at that point.

Notice the following sequence of canonical isometries

$$\Sigma_{/\sim} \to E \to \{p_0\} \times E \to \Sigma_{p_0}$$

$$[\overrightarrow{p_0p_1}] \mapsto v = p_1 - p_o \mapsto (p_o, v) \mapsto \overrightarrow{p_0\,p_0 + v} = \overrightarrow{p_0p_1}$$

4.3 Topological structures

'Calculus' will provide a set of methods for analyzing the behaviour of functions on Euclidean affine spaces. Such methods rest on the fundamental concepts of 'limit' and 'continuity', based on the 'topological structure' of a Euclidean affine space, which will now be focused and naturally generalized.

4.3.1 Topological spaces

A Euclidean affine space naturally exhibits a 'topological structure', expressing the properties of 'nearness' in a locus ($\tau\acute{o}\pi o\varsigma$) like the 'neighbourhood' of a point.

Topology

In a Euclidean affine space $\mathcal{E}$, endowed with its Euclidean distance d, the points of an *open ball*

$$\mathcal{B}^r_{\mathrm{p}_o} := \{\mathrm{p} \in \mathcal{E} \mid d(\mathrm{p}_o, \mathrm{p}) < r\}$$

centred at any given point $\mathrm{p}_o \in \mathcal{E}$, with a suitably small radius $r > 0$, can be thought of as 'near' to p_o.

Exercise 14 *For all* $\mathrm{p}_o \in \mathcal{E}$ *and* $r > 0$, $\mathcal{B}^r_{\mathrm{p}_o} - \{\mathrm{p}_o\} \neq \emptyset$. [26]

More generally, the points of a subset obtained from the union of (suitably small) open balls, one (at least) of which containing p_o, can be thought of as 'near' to p_o as well.

Exercise 15 *If* $\mathcal{U}_{\mathrm{p}_o} \subset \mathcal{E}$ *is union of open balls, one (at least) of which containing* p_o, *then there exists a radius* $r > 0$ *s.t.* $\mathcal{B}^r_{\mathrm{p}_o} \subset \mathcal{U}_{\mathrm{p}_o}$.

So 'nearness' leads us to focus on the open balls, which are in turn a *basis* of – in the sense that they generate, via set-theoretical union – the *Euclidean topology* of $\mathcal{E}$, i.e. the collection $\tau_{\mathcal{E}}$ containing all the subsets of $\mathcal{E}$ obtained from the union of open balls (and then the open balls themselves) plus the empty subset. [27]

Exercise 16 $\tau_{\mathcal{E}}$ *satisfies*
(1) $\emptyset, \mathcal{E} \in \tau_{\mathcal{E}}$
(2) $\{\mathcal{U}_\alpha\} \subset \tau_{\mathcal{E}} \Rightarrow \bigcup_\alpha \mathcal{U}_\alpha \in \tau_{\mathcal{E}}$
(3) $\{\mathcal{U}_1, \mathcal{U}_2\} \subset \tau_{\mathcal{E}} \Rightarrow \mathcal{U}_1 \cap \mathcal{U}_2 \in \tau_{\mathcal{E}}$
and
(4) *For any two distinct points* $\mathrm{p}_o \neq \mathrm{p}_1$ *of* $\mathcal{E}$, *there exists a subset* $\mathcal{U}_{\mathrm{p}_o} \in \tau_{\mathcal{E}}$ *containing* p_o *and a subset* $\mathcal{U}_{\mathrm{p}_1} \in \tau_{\mathcal{E}}$ *containing* p_1 *s.t.* $\mathcal{U}_{\mathrm{p}_o} \cap \mathcal{U}_{\mathrm{p}_1} = \emptyset$.

---

[26] Such a property is not fulfilled in any metric space. Think, for instance, of a non-empty set $\mathcal{E}$ with the distance function defined by $d(\mathrm{p}, \mathrm{q}) = 0$ or 1 according to whether $\mathrm{p} = \mathrm{q}$ or $\mathrm{p} \neq \mathrm{q}$, respectively, and notice that, for any positive $r < 1$ and any point $\mathrm{p}_o \in \mathcal{E}$, $\mathcal{B}^r_{\mathrm{p}_o} = \{\mathrm{p}_o\}$.

[27] If $\mathcal{E} = \mathcal{E}_1 \times \cdots \times \mathcal{E}_\nu$, the Cartesian products $\{\mathcal{U}_1 \times \cdots \times \mathcal{U}_\nu\}_{\mathcal{U}_1 \in \tau_{\mathcal{E}_1}, \ldots, \mathcal{U}_\nu \in \tau_{\mathcal{E}_\nu}}$ could be shown to be another basis of $\tau_{\mathcal{E}}$.

The above properties of Euclidean topology show the way for the transition to 'general topology'.

On any non-empty set X, a collection of subsets τ_X containing the trivial ones

1. $\emptyset, X \in \tau_X$

and satisfying – with respect to the set-theoretical operations of union and intersection – the 'closure' properties

2. $\{\mathcal{V}_\alpha\} \subset \tau_X \Rightarrow \bigcup_\alpha \mathcal{V}_\alpha \in \tau_X$

and

3. $\{\mathcal{V}_1, \mathcal{V}_2\} \subset \tau_X \Rightarrow \mathcal{V}_1 \cap \mathcal{V}_2 \in \tau_X$

is said to be a *topology* on X.

A *Hausdorff topology* is one satisfying the 'separability' property

4. For any two distinct points $x_o \neq x_1$ of X, there exists a subset $\mathcal{V}_{x_o} \in \tau_X$ containing x_o and a subset $\mathcal{V}_{x_1} \in \tau_X$ containing x_1 s.t. $\mathcal{V}_{x_o} \cap \mathcal{V}_{x_1} = \emptyset$.

Once equipped with a topology τ_X, X is said to be a *topological space.*

The subsets of X belonging to τ_X and their complements, are then called the *open subsets* and the *closed subsets* of X, respectively.

An open subset $\mathcal{V}_{x_o} \in \tau_X$, containing a given point $x_o \in X$, is said to be an *open neighbourhood* of x_o in the topological space X.

A *Hausdorff topological space* is one carrying a Hausdorff topology.

Among the subsets of a topological space X, the open subsets can be characterized as follows:

Exercise 17 *$S \subset X$ is an open subset, iff each $x_o \in S$ is an* internal point *of S, i.e there exists an open neighbourhood $\mathcal{V}_{x_o}$ of x_o in X entirely contained in S.*

Any subset S of X will be thought of as a *topological subspace* of X with the following 'inherited' topology:

Exercise 18 *The collection of subsets of $S \subset X$ given by*

$$\tau_S := \{S \cap \mathcal{V}\}_{\mathcal{V} \in \tau_X}$$

is a topology on S, called subspace topology *(if S is an open subset of X, then τ_S is the collection of all the open subsets of X contained in S).*

A point $x_o \in S$ is said to be an *isolated point* of S, if

$$\{x_o\} \in \tau_S$$

As a consequence, a point $x_o \in S$ is a *non-isolated point* of S, iff $\{x_o\}$ does not belong to τ_S, i.e.

$$\mathcal{W}_{x_o} - \{x_o\} \neq \emptyset$$

for all the open neighbourhoods $\mathcal{W}_{x_o} \in \tau_S$ (of x_o in S), that is to say,

$$(S \cap \mathcal{V}_{x_o}) - \{x_o\} \neq \emptyset$$

for all the open neighbourhoods $\mathcal{V}_{x_o} \in \tau_X$ (of x_o in X).

In the latter form, the above property can as well be referred to any point $x_o \in X$, whether belonging to S or not. If $x_o \in X$ satisfies such a property, it is said to be an *accumulation point* for S.

Exercise 19 *Let S be a subset of a Euclidean affine space (the latter being endowed with its Euclidean topology). If $x_o \in S$ is an internal point of S, then it is both a non-isolated point of S and an accumulation point for $S - \{x_o\}$.*

Limit

Let $\mathcal{A}$ and $\mathcal{E}$ be Euclidean affine spaces (modelled on vector spaces A and E, respectively), $f : S \subset \mathcal{A} \to \mathcal{E}$ a mapping defined on a subset S of $\mathcal{A}$ and $x_o \in \mathcal{A}$ an accumulation point for S.

After Cauchy, a point $\mathrm{p}_o \in \mathcal{E}$ is said to be a *limit* of f at x_o, in symbols $\mathrm{p}_o = \lim_{x \to x_o} f(x)$, if

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 : x \in S,\ 0 < |x - x_o| < \delta \implies |f(x) - \mathrm{p}_o| < \epsilon$$

Exercise 20 *Let $f : S \subset \mathcal{A} \to \mathcal{E}$, $\psi : S \subset \mathcal{A} \to E$ and $h = f + \psi : x \in S \subset \mathcal{A} \mapsto h(x) = f(x) + \psi(x) \in \mathcal{E}$. If any two of the above mappings admit limits at x_o, so does the third and*

$$\lim_{x \to x_o} h(x) = \lim_{x \to x_o} f(x) + \lim_{x \to x_o} \psi(x)$$

Cauchy's definition of limit clearly exhibits a topological nature, since it amounts to saying

$$\forall\, \mathcal{B}^{\epsilon}_{\mathrm{p}_o} \in \tau_{\mathcal{E}},\ \exists\, \mathcal{B}^{\delta}_{x_o} \in \tau_{\mathcal{A}} : f\Big((S \cap \mathcal{B}^{\delta}_{x_o}) - \{x_o\}\Big) \subset \mathcal{B}^{\epsilon}_{\mathrm{p}_o}$$

that is, owing to Exercise 15,

$$\forall\, \mathcal{U}_{\mathrm{p}_o} \in \tau_{\mathcal{E}},\ \exists\, \mathcal{U}_{x_o} \in \tau_{\mathcal{A}} : f\Big((S \cap \mathcal{U}_{x_o}) - \{x_o\}\Big) \subset \mathcal{U}_{\mathrm{p}_o}$$

In the latter form, the definition of limit can be extended to general topological spaces.

Let X and Y be topological spaces (with topologies τ_X and τ_Y, respectively), $f : S \subset X \to Y$ a mapping defined on a subset S of X and $x_o \in X$ an accumulation point for S. A point $y_o \in Y$ will be said to be a *limit* of f at x_o, in symbols $y_o = \lim_{x \to x_o} f(x)$, if

$$\forall \mathcal{V}_{y_o} \in \tau_Y, \exists \mathcal{V}_{x_o} \in \tau_X : f\Big((S \cap \mathcal{V}_{x_o}) - \{x_o\}\Big) \subset \mathcal{V}_{y_o}$$

(remark that the above is a non-trivial requirement, owing to the 'accumulation' hypothesis $(S \cap \mathcal{V}_{x_o}) - \{x_o\} \neq \emptyset$ for all the open neighbourhoods $\mathcal{V}_{x_o} \in \tau_X$).

Proposition 6 *If Y is a Hausdorff topological space, there exists at most a unique limit of $f : S \subset X \to Y$ at a point $x_o \in X$ of accumulation for S.*

Proof: (Reductio ad absurdum)

If there existed two distinct limits y_o, y_1 in Y, they could be separated by two disjoint open neighbourhoods $\mathcal{V}_{y_o}, \mathcal{V}_{y_1}$ and then –by the very definition of limit– there would exist two open neighbourhoods $\mathcal{V}^{(o)}_{x_o}, \mathcal{V}^{(1)}_{x_o}$ in τ_X s.t.

$$f\Big((S \cap \mathcal{V}^{(o)}_{x_o}) - \{x_o\}\Big) \subset \mathcal{V}_{y_o}$$

and

$$f\Big((S \cap \mathcal{V}^{(1)}_{x_o}) - \{x_o\}\Big) \subset \mathcal{V}_{y_1}$$

whence, for $\mathcal{V}_{x_o} := \mathcal{V}^{(o)}_{x_o} \cap \mathcal{V}^{(1)}_{x_o} \in \tau_X$,

$$\emptyset \neq f\Big((S \cap \mathcal{V}_{x_o}) - \{x_o\}\Big) \subset \mathcal{V}_{y_o} \cap \mathcal{V}_{y_1} = \emptyset$$

which is manifestly inconsistent. □

Remark that, if $x_o \in S$ is a non-isolated point of S, then the above definition of limit reads $\forall \mathcal{V}_{y_o} \in \tau_Y, \exists \mathcal{W}_{x_o} \in \tau_S : f\Big(\mathcal{W}_{x_o} - \{x_o\}\Big) \subset \mathcal{V}_{y_o}$.

So, in the case $S = X$, we have that $y_o \in Y$ is a limit of $f : X \to Y$ at a non-isolated point $x_o \in X$, if

$$\forall \mathcal{V}_{y_o} \in \tau_Y, \exists \mathcal{V}_{x_o} \in \tau_X : f\Big(\mathcal{V}_{x_o} - \{x_o\}\Big) \subset \mathcal{V}_{y_o}$$

4.3.2 Topological morphisms

In the category of topological spaces, the morphisms are mappings exhibiting the familiar character of 'continuity'.

Continuous mappings

Let X and Y be topological spaces (with topologies τ_X and τ_Y, respectively).

As is well known, a mapping $f : X \to Y$ is said to be *continuous* at a non-isolated point $x_o \in X$, if

$$f(x_o) = \lim_{x \to x_o} f(x)$$

If $x_o \in X$ is an isolated point, f is still said to be continuous at x_o.

The above definition of continuity at a point, can be re-expressed as follows:

Proposition 7 *$f : X \to Y$ is continuous at $x_o \in X$, iff*

$$\forall \mathcal{V}_{f(x_o)} \in \tau_Y, \exists \mathcal{V}_{x_o} \in \tau_X : f(\mathcal{V}_{x_o}) \subset \mathcal{V}_{f(x_o)}$$

Proof: If x_o is a non-isolated point of X, f is continuous at x_o, iff, for any $\mathcal{V}_{f(x_o)} \in \tau_Y$, there exists a $\mathcal{V}_{x_o} \in \tau_X$ satisfying condition

$$f(\mathcal{V}_{x_o} - \{x_o\}) \subset \mathcal{V}_{f(x_o)}$$

which, owing to $f(x_o) \in \mathcal{V}_{f(x_o)}$, amounts to saying

$$f(\mathcal{V}_{x_o}) \subset \mathcal{V}_{f(x_o)}$$

If x_o is an isolated point of X, the above condition is trivially fulfilled by choosing $\mathcal{V}_{x_o} = \{x_o\} \in \tau_X$. □

A *topological morphism* of X in Y is a *continuous mapping* $f : X \to Y$, i.e. one that is continuous at each point of X.

Exercise 21 *An affine mapping $f : \mathcal{A} \to \mathcal{E}$ between Euclidean affine spaces, is continuous.* [28]

[28] Make use of the fact that, for any non-vanishing linear mapping $F : A \to E$ between Euclidean vector spaces, there exists a real number $|F| > 0$ s.t. $|F(v)| \leq |F||v|$ for all $v \in A$.

An important characterization (and then an equivalent definition) of continuous mappings is the following:

Proposition 8 *$f : X \to Y$ is a continuous mapping, iff $f^{-1}(\mathcal{V}) \in \tau_X$ for all $\mathcal{V} \in \tau_Y$.*

Proof: Let f be continuous. For any $\mathcal{V} \in \tau_Y$, $f^{-1}(\mathcal{V})$ is either empty or non-empty. In the former case, $f^{-1}(\mathcal{V}) = \emptyset \in \tau_X$. In the latter case, each point $x_o \in f^{-1}(\mathcal{V})$ is an internal point of $f^{-1}(\mathcal{V})$, since there exists, by continuity, a $\mathcal{V}_{x_o} \in \tau_X$ s.t. $f(\mathcal{V}_{x_o}) \subset \mathcal{V}$, i.e. $\mathcal{V}_{x_o} \subset f^{-1}(\mathcal{V})$, whence $f^{-1}(\mathcal{V}) \in \tau_X$.

Conversely, let $f^{-1}(\mathcal{V}) \in \tau_X$ for all $\mathcal{V} \in \tau_Y$. For every $x_o \in X$ and $\mathcal{V}_{f(x_o)} \in \tau_Y$, there exists $\mathcal{V}_{x_o} := f^{-1}(\mathcal{V}_{f(x_o)}) \in \tau_X$ fulfilling $f(\mathcal{V}_{x_o}) \subset \mathcal{V}_{f(x_o)}$. □

Owing to the above proposition, the following properties can easily be proved: [29]

Exercise 22
(i) *The composition of continuos mappings is continuous.*
(ii) *The inclusion mapping $\iota_X : X \hookrightarrow \tilde{X}$ is continuous.*
(iii) *If $\tilde{f} : \tilde{X} \to Y$ is continuous and $X \subset \tilde{X}$, then the restricted mapping $f : X \to Y$, i.e. $f := \tilde{f}|_X := \tilde{f} \circ \iota_X$, is continuous.*
(iv) *If $\tilde{f} : X \to \tilde{Y}$ is continuous and $\tilde{f}(X) \subset Y \subset \tilde{Y}$, then the induced mapping $f : X \to Y$, s.t. $\tilde{f} = \iota_Y \circ f$, is continuous.*
(v) *If $\tilde{f} : \tilde{X} \to \tilde{Y}$ is continuous, $X \subset \tilde{X}$ and $\tilde{f}(X) \subset Y \subset \tilde{Y}$, then the induced mapping $f : X \to Y$, s.t. $\tilde{f}|_X = \iota_Y \circ f$, is continuous.*
(vi) *A continuous mapping $\tilde{f} : \tilde{X} \to \tilde{Y}$ takes a connected subspace $X \subset \tilde{X}$ onto a connected subspace $Y := \tilde{f}(X) \subset \tilde{Y}$.* [30]

Homeomorphisms

A *topological isomorphism* of X onto Y, called *homeomorphism*, is an invertible continuous mapping $f : X \to Y$ with continuous inverse $f^{-1} : Y \to X$.

Exercise 23 *A homeomorphism $f : X \to Y$ is an open mapping* [31] *with open inverse $f^{-1} : Y \to X$, and then it induces the bijection of τ_X onto τ_Y given by $\mathcal{U} \in \tau_X \mapsto f(\mathcal{U}) \in \tau_Y$ and inverted by $\mathcal{V} \in \tau_Y \mapsto f^{-1}(\mathcal{V}) \in \tau_Y$.*

[29] $X, Y, \tilde{X}, \tilde{Y}$ will denote topological spaces. If $X \subset \tilde{X}$ and/or $Y \subset \tilde{Y}$, X and Y will be meant to be topological subspaces of $\tilde{X}$ and $\tilde{Y}$, respectively.

[30] A topological space is said to be *connected*, if it *is not* union of two disjoint open subsets (of its own topology).

[31] An *open mapping* $f : X \to Y$ is one s.t. $f(\mathcal{U}) \in \tau_Y$ for all $\mathcal{U} \in \tau_X$.

4.4 Differential calculus

We shall now introduce the fundamentals of differential calculus on Euclidean affine spaces.

4.4.1 Differentiable mappings

Differentiation is a process of 'infinitesimal linearization', which – if applicable – allows a mapping between Euclidean affine spaces, restricted to an 'arbitrarily small' open subset of its domain of definition, to be replaced, up to 'higher-order infinitesimals', by the restriction of a suitable affine mapping.

Differentiability

Let $\mathcal{A}$ and $\mathcal{E}$ be Euclidean affine spaces (modelled on vector spaces A and E, respectively).

A mapping $f : S \subset \mathcal{A} \to \mathcal{E}$ defined on a subset S of $\mathcal{A}$, is said to be *differentiable* at a point $x_o \in S$, if

(i) x_o is an internal point of S, [32]

(ii) there exists an affine mapping $g : \mathcal{A} \to \mathcal{E}$, such that the *difference* $\varphi := f - g_{|S} : x \in S \mapsto \varphi(x) := f(x) - g(x) \in E$ vanishes at x_o, i.e.

$$\varphi(x_o) = 0$$

and is *higher-order infinitesimal* at x_o, i.e.

$$\lim_{x \to x_o} \frac{\varphi(x)}{|x - x_o|} = 0$$

Exercise 24 *The above higher-order infinitesimality condition implies the ordinary infinitesimality condition* $\lim_{x \to x_o} \varphi(x) = 0$.

If f is differentiable at x_o, the affine mapping g is uniquely determined, since – owing to (ii) – its value at x_o is given by $f(x_o)$ and its linear part $d_{x_o} f : A \to E$, called *differential* of f at x_o, will be shown to take the value

$$d_{x_o} f\,(u) = \lim_{a \to 0} \frac{f(x_o + au) - f(x_o)}{a} \qquad (*)$$

at any $u \in A$ such that $|u| = 1$ [33] (where a is meant to vary in $(-r, r) - \{0\}$ with $r > 0$ s.t. $\mathcal{B}^r_{x_o} \subset S$).

[32] Recall that an internal point $x_o \in S$ is a non-isolated point of S and an accumulation point for $S - \{x_o\}$.

[33] The above values completely determine the differential, since (owing to linearity) $d_{x_o} f$ vanishes at the zero vector of A and, at any non-null vector $v = |v| u \in A$ (with $|u| = 1$), takes the value $d_{x_o} f\,(v) = |v|\, d_{x_o} f(u)$.

The meaning of differentiability is the following.

If f is differentiable at x_o, then – in an 'arbitrarily small' open ball $\mathcal{B}^r_{x_o}$ contained in S – the difference φ between the given mapping f and the affine mapping g is negligible

$$\varphi_{|\mathcal{B}^r_{x_o}} = f_{|\mathcal{B}^r_{x_o}} - g_{|\mathcal{B}^r_{x_o}} \approx 0$$

So, for all $x \in \mathcal{B}^r_{x_o} \subset S$, one has – up to higher-order infinitesimals –

$$f(x) \approx g(x)$$

that is,

$$f(x) \approx f(x_o) + d_{x_o}f\,(x - x_o)$$

or

$$f(x) - f(x_o) \approx d_{x_o}f\,(x - x_o)$$

(the increment of f is a linear function – up to higher-order infinitesimals – of the increment given to the value x_o of its argument).

That expresses the announced process of 'infinitesimal linearization'.

A mapping $f : S \subset \mathcal{A} \to \mathcal{E}$ differentiable at each point of its domain S (which then turns out to be an open subset of $\mathcal{A}$), is said to be a *differentiable mapping*.

Exercise 25 *An affine mapping* $f : \mathcal{A} \to \mathcal{E}$, *with linear part* $F : A \to E$, *is differentiable at each point* $x_o \in \mathcal{A}$ *and* $d_{x_o}f = F$ *(in particular* $d_{x_o}f = 0$, *if* f *is constant).*

Differentiability and derivability

The right hand side of $(*)$ is called *derivative* of f at x_o along the oriented direction $\mathrm{Span}^+(u)$. [34]

So equality $(*)$, which will now be proved, means that differentiability implies derivability:

[34] See footnote [24].

Proposition 9 *If f is differentiable at x_o, then the derivative of f at x_o along any oriented direction* $\mathrm{Span}^+(u)$ *exists and equals $d_{x_o}f\,(u)$.*

Proof: On the one hand, for any $a \in (-r, r) - \{0\}$, we have

$$\begin{aligned}
\left|\frac{f(x_o+au)-f(x_o)}{a} - d_{x_o}f\,(u)\right| &= \left|\frac{(f(x_o+au)-f(x_o)) - d_{x_o}f\,(au)}{a}\right| \\
&= \left|\frac{f(x_o+au)-(f(x_o)+d_{x_o}f\,(au))}{a}\right| \\
&= \left|\frac{f(x_o+au)-g(x_o+au)}{a}\right| \\
&= \left|\frac{\varphi(x_o+au)}{a}\right| \\
&= \frac{|\varphi(x_o+au)|}{|a|} \\
&= \frac{|\varphi(x_o+au)|}{|(x_o+au)-x_o|}
\end{aligned}$$

On the other hand, as φ is higher-order infinitesimal at x_o, for any $\epsilon > 0$ there exists a suitably small $\delta > 0$, say $\delta < r$, such that, whenever $0 < |(x_o+au)-x_o| < \delta$, we have

$$\frac{|\varphi(x_o+au)|}{|(x_o+au)-x_o|} < \epsilon$$

So, for any $\epsilon > 0$ there exists a $\delta > 0$, with $\delta < r$, such that, whenever $0 < |a| < \delta$, we have

$$\left|\frac{f(x_o+au)-f(x_o)}{a} - d_{x_o}f\,(u)\right| < \epsilon$$

That proves our claim. □

Differentiability and continuity

Differentiability also implies continuity:

Proposition 10 *If f is differentiable at x_o, then it is continuous at x_o.*

Proof: As x_o is a non-isolated point of S, we have to prove that

$$f(x_o) = \lim_{x \to x_o} f(x)$$

To this end, recall that g is (affine and then) continuous, and φ is (higher-order) infinitesimal at x_o, i.e.

$$\lim_{x \to x_o} g(x) = g(x_o) = f(x_o)\,, \qquad \lim_{x \to x_o} \varphi(x) = 0$$

As $f(x) = g(x) + \varphi(x)$ for all $x \in S$, we have

$$\lim_{x \to x_o} f(x) = \lim_{x \to x_o} g(x) + \lim_{x \to x_o} \varphi(x)$$

Hence our claim. □

Rules of differentiation

Now we give the list (but not the proof) of the basic rules of differentiation:

Local character

If (and only if) $f : S \subset \mathcal{A} \to \mathcal{E}$ is differentiable at $x_o \in S$, so is its restriction $f|_{\mathcal{V}_{x_o}} : \mathcal{V}_{x_o} \subset \mathcal{A} \to \mathcal{E}$ to an open neighbourhood $\mathcal{V}_{x_o} \subset S$ of x_o in $\mathcal{A}$ and

$$d_{x_o}(f|_{\mathcal{V}_{x_o}}) = d_{x_o} f$$

Additivity

If $f : S \subset \mathcal{A} \to \mathcal{E}$ and $\psi : S \subset \mathcal{A} \to E$ are differentiable at $x_o \in S$, so is their sum $h = f + \psi : S \subset \mathcal{A} \to \mathcal{E}$ and [35]

$$d_{x_o}(f + \psi) = d_{x_o} f + d_{x_o} \psi$$

Leibniz rule

If $f : S \subset \mathcal{A} \to \mathbb{R}$ and $h : S \subset \mathcal{A} \to \mathbb{R}$ are differentiable at $x_o \in S$, so is their product $fh : S \subset \mathcal{A} \to \mathbb{R}$ and [36]

$$d_{x_o}(fh) = f(x_o)\, d_{x_o} h + h(x_o)\, d_{x_o} f$$

[35] Linear mappings like $d_{x_o} f : A \to E$ and $d_{x_o} \psi : A \to E$, can naturally be summed.

[36] Recall that $d_{x_o} f$, $d_{x_o} h$ and $d_{x_o}(fh)$ belong to the dual vector space A^*.

Remark that the rule extends to other kinds of product between vector-valued mappings.

Consider, e.g., the affine mapping $\varphi : (\mathrm{p}, \mathrm{q}) \in \mathcal{E} \times \mathcal{E} \mapsto \varphi(\mathrm{p}, \mathrm{q}) := \mathrm{q} - \mathrm{p} \in E$, whose linear part is $\Phi : (\mathrm{u}, \mathrm{v}) \in E \times E \mapsto \Phi(\mathrm{u}, \mathrm{v}) := \mathrm{v} - \mathrm{u} \in E$, and the scalar product $d^2 = \varphi \cdot \varphi : (\mathrm{p}, \mathrm{q}) \in \mathcal{E} \times \mathcal{E} \mapsto \varphi(\mathrm{p}, \mathrm{q}) \cdot \varphi(\mathrm{p}, \mathrm{q}) = (\mathrm{q} - \mathrm{p}) \cdot (\mathrm{q} - \mathrm{p}) = |\mathrm{q} - \mathrm{p}|^2 = d^2(\mathrm{p}, \mathrm{q}) \in \mathbb{R}$ (square of distance), whose differential at any $(\mathrm{p}_o, \mathrm{q}_o) \in \mathcal{E} \times \mathcal{E}$ is $d_{(\mathrm{p}_o, \mathrm{q}_o)} d^2 = 2\varphi(\mathrm{p}_o, \mathrm{q}_o) \cdot d_{(\mathrm{p}_o, \mathrm{q}_o)} \varphi = 2(\mathrm{q}_o - \mathrm{p}_o) \cdot \Phi$.

Chain rule

If $f : S \subset \mathcal{A} \to \mathcal{H}$ is differentiable at $x_o \in S$ and $h : U \subset \mathcal{H} \to \mathcal{E}$ is differentiable at $f(x_o) \in f(S) \subset U$, so is their composite $h \circ f : S \subset \mathcal{A} \to \mathcal{E}$ at x_o and

$$d_{x_o}(h \circ f) = d_{f(x_o)}h \circ d_{x_o}f$$

Exercise 26 *If (and only if)*

$$\begin{aligned} f = (f_1, \ldots, f_\nu) \quad &: \quad S \subset \mathcal{A} \to \mathcal{E} := \mathcal{E}_1 \times \cdots \times \mathcal{E}_\nu \\ &: \quad x \mapsto f(x) = (f_1(x), \ldots, f_\nu(x)) \end{aligned}$$

is differentiable at $x_o \in S$*, so is each projection* $f_i : S \subset \mathcal{A} \to \mathcal{E}_i$ *and its differential at* x_o *is given by*

$$\begin{aligned} d_{x_o}f = (d_{x_o}f_1, \ldots, d_{x_o}f_\nu) \quad &: \quad A \to E := E_1 \times \cdots \times E_\nu \\ &: \quad v \mapsto d_{x_o}f(v) = (d_{x_o}f_1(v), \ldots, d_{x_o}f_\nu(v)) \end{aligned}$$

Hint: Let $pr_i : \mathcal{E} \to \mathcal{E}_i$ (resp. $\mathrm{pr}_i : E \to E_i$) be the projection of $\mathcal{E}$ onto its i-th factor $\mathcal{E}_i$ (resp. the projection of E onto its i-th factor E_i). Then apply the chain rule to $f_i = pr_i \circ f$, obtaining $d_{x_o}f_i = d_{f(x_o)}pr_i \circ d_{x_o}f = \mathrm{pr}_i \circ d_{x_o}f$. □

Diffeomorphisms

A *diffeomorphism* is a bijection $f : S \to M$ between two open subsets $S \subset \mathcal{A}$ and $M \subset \mathcal{E}$ s.t. $\iota_M \circ f : S \to M \hookrightarrow \mathcal{E}$ is differentiable together with $\iota_S \circ f^{-1} : M \to S \hookrightarrow \mathcal{A}$.

Exercise 27 *If* $f : S \to M$ *is a diffeomorphism, then it is a homeomorphism.*

We shall identify f and f^{-1} with $\iota_M \circ f$ and $\iota_S \circ f^{-1}$, respectively.

Proposition 11 *If* $f : S \to M$ *is a diffeomorphism, then its differential* $d_{x_o}f$ *at any* $x_o \in S$ *is a linear isomorphism and*

$$(d_{x_o}f)^{-1} = d_{f(x_o)}f^{-1}$$

Proof: Owing to the rules of differentiation, from $f^{-1} \circ f = \iota_S$ and $f \circ f^{-1} = \iota_M$ we deduce

$$d_{f(x_o)}f^{-1} \circ d_{x_o}f = d_{x_o}(f^{-1} \circ f) = d_{x_o}\iota_S = d_{x_o}\iota_{\mathcal{A}} = \iota_A$$

and

$$d_{x_o}f \circ d_{f(x_o)}f^{-1} = d_{f(x_o)}(f \circ f^{-1}) = d_{f(x_o)}\iota_M = d_{f(x_o)}\iota_{\mathcal{E}} = \iota_E$$

whence our claim. □

Exercise 28 *If there exists a diffeomorphism* $f : S \to M$ *between two open subsets* $S \subset \mathcal{A}$ *and* $M \subset \mathcal{E}$*, then* $\dim(\mathcal{A}) = \dim(\mathcal{E})$.

C^∞ ***differentiability***

We shall now introduce differentiability of higher order.

We start with a kind of mapping

$$\xi : W \subset \mathbb{R}^n \to \mathcal{E} : q \mapsto \mathrm{p} = \xi(q)$$

defined on an open subset W of $\mathbb{R}^n$ [37] and taking its values in a Euclidean affine space $\mathcal{E}$ (ξ is said to be a *parametrization* of $\mathrm{Im}(\xi) \subset \mathcal{E}$).

If ξ is differentiable, its differential at any $q = (q^1, \ldots, q^n) \in W$ is a linear mapping [38]

$$\begin{aligned} d_q\xi : \mathbb{R}^n \to E : \delta q = \delta q^h \delta_h \mapsto \delta \mathrm{p} &= d_q\xi\,(\delta q) \\ &= (d_q\xi\,(\delta_h))\delta q^h \\ &= \left.\frac{\partial p}{\partial q^h}\right|_q \delta q^h \quad \approx \quad \xi(q + \delta q) - \xi(q) \end{aligned}$$

whose image

$$\mathrm{Im}(d_q\xi) = \mathrm{Span}\left(\left.\frac{\partial p}{\partial q^h}\right|_q\right)_{h=1,\ldots,n}$$

is spanned by the directional derivatives

$$\left.\frac{\partial \mathrm{p}}{\partial q^h}\right|_q := \lim_{a\to 0} \frac{\xi(q + a\,\delta_h) - \xi(q)}{a} = d_q\xi\,(\delta_h)$$

which, if q is let to vary in W, determine the n-tuple of mappings

$$\frac{\partial \mathrm{p}}{\partial q^h} : W \subset \mathbb{R}^n \to E : q \mapsto \left.\frac{\partial \mathrm{p}}{\partial q^h}\right|_q$$

called *first-order partial derivatives* of ξ.

The above mappings, if differentiable, give rise to the $n \times n$ matrix of *second-order partial derivatives* of ξ, i.e.

$$\frac{\partial^2 \mathrm{p}}{\partial q^k \partial q^h} : W \subset \mathbb{R}^n \to E : q \mapsto \left.\frac{\partial^2 \mathrm{p}}{\partial q^k \partial q^h}\right|_q := \left.\frac{\partial}{\partial q^k}\right|_q \frac{\partial \mathrm{p}}{\partial q^h}$$

By iterating such a procedure, under further hypotheses of differentiability, one can obtain higher-order partial derivatives of ξ.

---

[37] The case $n = 1$ will be treated in detail in the next section 4.4.2.

[38] Recall that $(\delta_1, \ldots, \delta_n)$ denotes the natural basis of $\mathbb{R}^n$.

ξ will be said to be a C^∞ *differentiable* mapping, if it is differentiable and admits differentiable partial derivatives of any order (clearly, a C^∞ mapping is continuous and admits continuous partial derivatives of any order). [39]

Elementary operations (such as restriction to open subsets, sum, product, composition) preserve C^∞ differentiability.

Owing to the local character of differentiability, ξ is C^∞ differentiable iff so is its restiction to an open neighbourhood of each point of its domain.

Exercise 29 *An affine mapping*

$$\xi : \mathbb{R}^n \to \mathcal{E} : q \mapsto \mathrm{p} = \mathrm{o} + q^h \, \mathrm{e}_h$$

(where $\mathrm{o} := \xi(0)$ *and* $\mathrm{e}_h = \Xi(\delta_h)$ *,* $\Xi : \mathbb{R}^n \to E$ *being the linear part of* ξ *) is* C^∞ *differentiable, with constant first-order derivatives*

$$\frac{\partial \mathrm{p}}{\partial q^h} = \mathrm{e}_h$$

and vanishing higher-order derivatives.

We shall now consider a more general kind of mapping

$$f : S \subset \mathcal{A} \to \mathcal{E}$$

defined on an open subset of an arbitrary, n-dimensional, Euclidean affine space $\mathcal{A}$.

Choose a Cartesian system ϕ in $\mathcal{A}$ (affine isomorphism of $\mathbb{R}^n$ onto $\mathcal{A}$) and put $W := \phi^{-1}(S)$ (open subset of $\mathbb{R}^n$).

f will be said to be C^∞ *differentiable*, if so is $f \circ \phi|_W$.

The following exercise is to show that the C^∞ differentiability of f does not depend on the choice of ϕ.

Exercise 30 *Let* ϕ *,* ϕ' *be two Cartesian systems in* $\mathcal{A}$ *and put* $W := \phi^{-1}(S)$ *,* $W' := \phi'^{-1}(S)$ *. If* $f \circ \phi|_W$ *is* C^∞*, so is* $f \circ \phi'|_{W'}$

4.4.2 Differentiable curves

We shall now specialize our considerations to functions of one real variables (with values in an arbitrary Euclidean affine space), whose role is crucial in the next sections (as well as in the main text).

[39] If a mapping is C^∞, then – owing to a classical result of Analysis – the matrix of its second-order partial derivatives is symmetric.

C^∞ ***differentiable parametrized curves***

Let

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p} = \mathrm{p}(t)$$

be a *parametrized curve* of a Euclidean affine space $\mathcal{E}$, defined on an open interval I (connected open subset of $\mathbb{R}$).

If γ is continuous, the *orbit* $\mathrm{Im}(\gamma)$ described by γ is a connected subspace of $\mathcal{E}$.

If γ is differentiable, its differential at any $t \in I$ is a linear mapping

$$\begin{aligned} d_t\gamma : \mathbb{R} \to E : dt \mapsto d\mathrm{p} &= d_t\gamma\,(dt) \\ &= (d_t\gamma\,(1))dt \\ &= \left.\frac{d\mathrm{p}}{dt}\right|_t dt \quad \approx \quad \mathrm{p}(t+dt) - \mathrm{p}(t) \end{aligned}$$

whose image

$$\mathrm{Im}(d_t\gamma) = \mathrm{Span}\left(\left.\frac{d\mathrm{p}}{dt}\right|_t\right)$$

is spanned by the derivative

$$\left.\frac{d\mathrm{p}}{dt}\right|_t := \lim_{a\to 0} \frac{\mathrm{p}(t+a) - \mathrm{p}(t)}{a} = d_t\gamma\,(1)$$

which, if t is let to vary in I, determines the parametrized curve of E

$$\frac{d\mathrm{p}}{dt} : I \to E : t \mapsto \left.\frac{d\mathrm{p}}{dt}\right|_t =: \dot{\mathrm{p}}(t)$$

called *first-order derivative* of γ.

The above parametrized curve, if differentiable, gives rise –at any $t \in I$– to

$$\left.\frac{d^2\mathrm{p}}{dt^2}\right|_t := \left.\frac{d\dot{\mathrm{p}}}{dt}\right|_t := \lim_{a\to 0} \frac{\dot{\mathrm{p}}(t+a) - \dot{\mathrm{p}}(t)}{a}$$

which in turn determines the parametrized curve of E

$$\frac{d^2\mathrm{p}}{dt^2} : I \to E : t \mapsto \left.\frac{d^2\mathrm{p}}{dt^2}\right|_t =: \ddot{\mathrm{p}}(t)$$

called *second-order derivative* of γ.

By iterating such a procedure, under further hypotheses of differentiability, one can obtain higher-order derivatives of γ.

So γ is a C^∞ differentiable (parametrized) curve, if it is differentiable and admits differentiable derivatives of any order.

Exercise 31 *If (and only if)*

$$\begin{array}{rcl}\gamma = (\gamma_1, \ldots, \gamma_\nu) & : & I \to \mathcal{E} := \mathcal{E}_1 \times \cdots \times \mathcal{E}_\nu \\ & : & t \mapsto \mathrm{p}(t) = (\mathrm{p}_1(t), \ldots, \mathrm{p}_\nu(t))\end{array}$$

is (C^∞) *differentiable, so is each projection* $\gamma_i : t \in I \mapsto \mathrm{p}_i(t) \in \mathcal{E}_i$ *and, for any* $t \in I$ *,*

$$\dot{\mathrm{p}}(t) = (\dot{\mathrm{p}}_1(t), \ldots, \dot{\mathrm{p}}_\nu(t)) \in E := E_1 \times \cdots E_\nu$$

(and so on for higher-order derivatives).

In particular, if

$$q(t) = (q^1(t), \ldots, q^n(t)) = q^h(t)\, \delta_h \in \mathbb{R}^n$$

is a (C^∞) *differentiable curve of* $\mathbb{R}^n$ *, then*

$$\dot{q}(t) = (\dot{q}^1(t), \ldots, \dot{q}^n(t)) = \dot{q}^h(t)\, \delta_h \in \mathbb{R}^n$$

(and so on for higher-order derivatives).

Reparametrization

Let

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t)$$

be a (C^∞) differentiable curve of $\mathcal{E}$, obtained from another one

$$\chi : J \subset \mathbb{R} \to \mathcal{E} : s \mapsto \mathrm{p}(s)$$

through a (C^∞) differentiable *reparametrization*

$$\sigma : I \subset \mathbb{R} \to \mathbb{R} : t \mapsto s(t)$$

with $\mathrm{Im}(\sigma) \subset J$, that is,

$$\gamma = \chi \circ \sigma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t) = \mathrm{p}(s(t))$$

Clearly, $\mathrm{Im}(\gamma) \subset \mathrm{Im}(\chi)$.

Moreover, owing to the chain rule,

$$\begin{array}{rcl}\dot{\mathrm{p}}(t) & = & d_t\gamma\,(1) \\ & = & (d_t(\chi \circ \sigma))(1) \\ & = & (d_{s(t)}\chi \circ d_t\sigma)(1) \\ & = & d_{s(t)}\chi\,(d_t\sigma\,(1)) \\ & = & d_{s(t)}\chi\,(\dot{s}(t))\end{array}$$

or

$$\begin{array}{rcl}\dot{\mathrm{p}}(t) & = & (d_{s(t)}\chi\,(1))\dot{s}(t) \\ & = & \left.\dfrac{d\mathrm{p}}{ds}\right|_{s(t)} \dot{s}(t)\end{array}$$

Coordinate expression

Let

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t)$$

be a parametrized curve of $\mathcal{E}$ such that, near each point $\mathrm{p}(t_*) \in \mathrm{Im}(\gamma)$, there exists a piece of orbit

$$\mathrm{Im}(\gamma|_{I_*}) \subset \mathcal{U}$$

(image of an open subinterval $I_* \subset I$ containing t_*) which lives in a subset

$$\mathcal{U} = \mathrm{Im}(\xi) \subset \mathcal{E}$$

parametrized by an injective (C^∞) differentiable mapping

$$\xi : W \subset \mathbb{R}^n \to \mathcal{E} : q \mapsto \xi(q)$$

and whose *coordinate expression* (through $\xi^{-1} : \mathcal{U} \to W \subset \mathbb{R}^n$)

$$\gamma_{_\xi} := \xi^{-1} \circ \gamma|_{I_*} : I_* \to \mathbb{R}^n : t \mapsto q(t) := \xi^{-1}(\mathrm{p}(t))$$

is (C^∞) differentiable (hence

$$\gamma|_{I_*} = \xi \circ \gamma_{_\xi} : I_* \to \mathcal{E} : t \mapsto \mathrm{p}(t) = \xi(q(t))$$

is (C^∞) differentiable as well).

Under the above hypothesis, γ is (C^∞) differentiable and, for all $t_* \in I_*$,

$$\begin{aligned}
\dot{\mathrm{p}}(t_*) &= d_{t_*}\gamma(1) \\
&= d_{t_*}\gamma|_{I_*}(1) \\
&= (d_{t_*}(\xi \circ \gamma_{_\xi}))(1) \\
&= (d_{q(t_*)}\xi \circ d_{t_*}\gamma_{_\xi})(1) \\
&= d_{q(t_*)}\xi\,(d_{t_*}\gamma_{_\xi}\,(1)) \\
&= d_{q(t_*)}\xi\,(\dot{q}(t_*))
\end{aligned}$$

or (recalling that $\dot{q}(t_*) = \dot{q}^h(t_*)\,\delta_h$)

$$\begin{aligned}
\dot{\mathrm{p}}(t_*) &= d_{q(t_*)}\xi\,(\dot{q}^h(t_*)\,\delta_h) \\
&= (d_{q(t_*)}\xi\,(\delta_h))\,\dot{q}^h(t_*) \\
&= \left.\frac{\partial \mathrm{p}}{\partial q^h}\right|_{q(t_*)} \dot{q}^h(t_*)
\end{aligned}$$

4.5 Embedded manifolds

We shall now be concerned with the differential-topological properties of 'embedded manifolds', meant as a kind of well behaved subspaces of a Euclidean affine space, including familiar 'loci' such as smooth curves and surfaces (which 'infinitesimally' resemble straight lines and planes).

4.5.1 Flat manifolds

We start with the study of 'flat manifolds' (like straight lines and planes).

Affine subspaces

Let $\mathcal{E}$ be a Euclidean affine space (modelled on a vector space E).

A non-empty subset $\mathcal{A} \subset \mathcal{E}$ is said to be an n-dimensional *flat manifold* embedded in $\mathcal{E}$ (n being a positive integer), if it is the orbit of a point $\mathrm{p}_o \in \mathcal{E}$ under the affine action of an n-dimensional vector subspace $A \subset E$, i.e.

$$\mathcal{A} = \mathrm{p}_o + A := \{\mathrm{p}_o + \mathrm{v}\}_{\mathrm{v} \in A}$$

(clearly, $\mathrm{p}_o \in \mathcal{A}$ and $\mathcal{A} = \mathrm{p} + A$ for all $\mathrm{p} \in \mathcal{A}$).

A flat manifold $\mathcal{A} \subset \mathcal{E}$ is usually called an n-dimensional *affine subspace* of $\mathcal{E}$, owing to to the following fact:

Exercise 32 *A non-empty subset $\mathcal{A} \subset \mathcal{E}$ inherits the structure of an affine space, modelled on a vector subspace $A \subset E$, iff $\mathcal{A} = \mathrm{p}_o + A$ for some $\mathrm{p}_o \in \mathcal{E}$.*

Hint: The above 'inheritance' means that, if $+$ denotes the affine structure of $\mathcal{E}$, the restriction $+|_{\mathcal{A} \times A}$ takes values in $\mathcal{A}$ and defines an affine action of A on $\mathcal{A}$. □

Special names are adopted in the following cases: $\mathcal{A}$ is said to be a *straight line*, a *plane* or a *hyperplane*, according to whether $\dim(\mathcal{A}) = 1$, $\dim(\mathcal{A}) = 2$ or $\dim(\mathcal{A}) = \dim(\mathcal{E}) - 1 > 0$, respectively.

Exercise 33
(i) $\dim(\mathcal{A}) \leq \dim(\mathcal{E})$, *the equality holding iff* $\mathcal{A} = \mathcal{E}$.
(ii) $\mathcal{A}$ *is naturally Euclidean.*
(iii) *In $\mathcal{A}$, meant both as a Euclidean affine space of its own and as a topological subspace of $\mathcal{E}$, the Euclidean topology coincides with the subspace topology.*

Hint: As to (iii), remark that any open ball of $\mathcal{A}$ is the intersection of $\mathcal{A}$ with the open ball of $\mathcal{E}$ with the same centre and the same radius. □

Affine parametrizations

An n-dimensional affine subspace $\mathcal{A}$ can *globally* be given an affine parametrization, as follows.

Let $\phi : \mathbb{R}^n \to \mathcal{A}$ be a Cartesian system in $\mathcal{A}$, with linear part $\Phi : \mathbb{R}^n \to A$.

As $\iota_{\mathcal{A}} : \mathcal{A} \hookrightarrow \mathcal{E}$ is an injective affine mapping (with injective linear part $\iota_A : A \hookrightarrow E$), the composite

$$\xi := \iota_{\mathcal{A}} \circ \phi : \mathbb{R}^n \to \mathcal{E}$$

is an injective affine mapping of $\mathbb{R}^n$ into $\mathcal{E}$, satisfying

$$\mathrm{Im}(\xi) = \mathcal{A}$$

(clearly

$$\Xi := \iota_A \circ \Phi : \mathbb{R}^n \to E$$

satisfying

$$A = \mathrm{Im}(\Xi)$$

is the injective linear part of ξ).

The following exercise is to show that the affine subspaces are the only subsets of $\mathcal{E}$ that can be parametrized by means of an injective affine mapping:

Exercise 34 *If $\xi : \mathbb{R}^n \to \mathcal{E}$ is an injective affine mapping, with linear part $\Xi : \mathbb{R}^n \to E$, then $\mathrm{Im}(\xi) \subset \mathcal{E}$ is an n-dimensional affine subspace modelled on $\mathrm{Im}(\Xi) \subset E$.*

Such a kind of *affine parametrization* –which characterizes the affine subspaces– exhibits the following differential-topological properties:

Proposition 12 *If ξ is an affine parametrization of $\mathcal{A}$,* [40] *then*
(1) *$\xi : \mathbb{R}^n \to \mathcal{A}$ is a homeomorphism;*
(2) *$\xi : \mathbb{R}^n \to \mathcal{E}$ is C^∞ differentiable, with injective differential $d_q\xi : \mathbb{R}^n \to E$ at each $q \in \mathbb{R}^n$.*

Proof: Recall that
(1) $\xi : \mathbb{R}^n \to \mathcal{A}$ is an affine isomorphism and therefore a homeomorphism;
(2) $\xi : \mathbb{R}^n \to \mathcal{E}$ is affine and then C^∞, with injective differential $d_q\xi = \Xi$ at each $q \in \mathbb{R}^n$. □

[40] The mapping induced by $\xi : \mathbb{R}^n \to \mathcal{E}$ onto its own image $\mathcal{A}$ will be denoted by $\xi : \mathbb{R}^n \to \mathcal{A}$. Moreover $\mathcal{A}$ will be meant to be equipped with its subspace topology (as to which, recall Exercise 33 (iii)).

4.5.2 Smooth manifolds

We shall now pass on to the study of 'smooth manifolds' in a Euclidean environment, conceived as 'loci' which admit (local) parametrizations exhibiting the same differential-topological properties as those of the (global) parametrizations of the affine subspaces.

Manifold structure

A non-empty subset $Q \subset \mathcal{E}$ will be said to be an n-dimensional *smooth manifold* embedded in $\mathcal{E}$ (n being a positive integer), if each point of Q is contained in an open subset [41]

$$\mathcal{U} = \mathrm{Im}(\xi) \in \tau_Q$$

parametrized by a mapping

$$\xi : W \subset \mathbb{R}^n \to \mathcal{E}$$

defined on an open subset W of $\mathbb{R}^n$ and satisfying the following differential-topological properties: [42]

1. $\xi : W \to \mathcal{U}$ is a homeomorphism;
2. $\xi : W \to \mathcal{E}$ is C^∞ differentiable, with injective differential $d_q\xi : \mathbb{R}^n \to E$ at each $q \in W$.

Each point $\mathrm{p} \in \mathcal{U}$ is then given a unique n-tuple of *coordinates* $q = (q^1, \ldots, q^n) = \xi^{-1}(\mathrm{p}) \in W$ by the n-dimensional *chart* $\xi^{-1} : \mathcal{U} \to W$, [43] which is said to be *local* or *global* according to whether its *coordinate domain* $\mathcal{U}$ is strictly contained in Q or coincides with Q, respectively. [44]

Any collection of n-dimensional charts whose coordinate domains cover the whole manifold, is said to be an *atlas*.

Special names are adopted for Q in the following cases: Q is said to be a smooth *curve*, *surface* or *hypersurface*, according to whether $\dim(Q) = 1$, $\dim(Q) = 2$ or $\dim(Q) = \dim(\mathcal{E}) - 1$, respectively.

---

[41] τ_Q will denote the subspace topology of Q.

[42] The mapping induced by $\xi : W \to \mathcal{E}$ onto its own image $\mathcal{U}$, will be denoted by $\xi : W \to \mathcal{U}$.

[43] In the sequel, the name 'chart' will often be referred to the parametrization ξ as well.

[44] For instance, a flat manifold admits global charts, determined by its affine parametrizations.

Exercise 35 *Prove that* $\dim(Q) \leq \dim(\mathcal{E})$.

Hint: Recall that $\mathrm{Im}(d_q\xi)$ is a vector subspace of E, whose dimension is given by $\dim(\mathrm{Im}(d_q\xi)) = \dim(\mathbb{R}^n) - \dim(\mathrm{Ker}(d_q\xi))$ and then, owing to the above property 2, $\dim(\mathrm{Im}(d_q\xi)) = n = \dim(Q)$. □

Exercise 36 *A smooth manifold embedded in an affine subspace* $\mathcal{A} \subset \mathcal{E}$ *is a smooth manifold of* $\mathcal{E}$.

Locally Euclidean topology

From a topological point of view, Q is a *locally Euclidean* subspace of $\mathcal{E}$, in the sense that it is covered, owing to the *topological property* 1, by open subsets (namely, the coordinate domains of its charts) homeomorphic to open subsets of Euclidean space $\mathbb{R}^n$.

Smoothness

From a differential point of view, Q is a *smooth* subspace of $\mathcal{E}$, in the sense that it admits, owing to the *smoothness property* 2, an n-dimensional 'tangent vector space' at each one of its points, as will now be shown.

Smooth parametrized curves and tangent vectors

Remark that 'tangency' is an 'infinitesimal' concept, linked to derivation as follows.

First consider a *smooth* (i.e. C^∞) parametrized curve of $\mathcal{E}$, say

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \mathrm{p}(t)$$

The vector $\mathrm{v} = \dot{\mathrm{p}}(t) \in E$, obtained from γ by derivation at any $t \in I$, is said to be *tangent* at $\mathrm{p} = \mathrm{p}(t)$ to γ. [45]

[45] Think of the representation of $\dot{\mathrm{p}}(t) := \lim_{a\to 0} \frac{1}{a}\big(\mathrm{p}(t+a) - \mathrm{p}(t)\big)$ –when it does not vanish– as an oriented segment attached at $\mathrm{p}(t)$, obtained via limit from a secant segment attached at the same point (and pointing towards the 'future' determined by the increasing 'time' t).

Now consider a *smooth* parametrized curve

$$\gamma : I \subset \mathbb{R} \to Q$$

of Q, that is, a smooth curve of $\mathcal{E}$ s.t.

$$\mathrm{Im}(\gamma) \subset Q$$

A vector of E tangent to a smooth parametrized curve γ of Q at a point p of its orbit, will then be said to be *tangent* at p to Q.

Tangent vector spaces

For any $\mathrm{p} \in Q$, the set $T_{\mathrm{p}}Q$ of all the vectors of E tangent at p to Q −i.e. tangent at p to the smooth parametrized curves of Q passing through p− will be called the *tangent vector space* of Q at p.

The name is due to the following

Proposition 13 $T_{\mathrm{p}}Q$ *is a vector subspace of* E *and* $\dim(T_{\mathrm{p}}Q) = \dim(Q)$.

Proof: Let $\xi : W \to \mathcal{U}$ be a chart of Q near $\mathrm{p} = \xi(q) \in \mathcal{U}$.

Recall that, owing to property 2, $\mathrm{Im}(d_q\xi) \subset E$ is a vector subspace of dimension $n = \dim(Q)$.

So our claim will be established by proving the following

Lemma $(\star)$ $\mathrm{Im}(d_q\xi) = T_{\mathrm{p}}Q$

Let $\mathrm{v} \in T_{\mathrm{p}}Q$, i.e.

$$\mathrm{v} = \dot{\mathrm{p}}(t_*)$$

for some smooth parametrized curve $\gamma : t \in I \mapsto \mathrm{p}(t) \in Q$ passing, at $t_* \in I$, through $\mathrm{p}(t_*) = \mathrm{p}$.

As $\mathcal{U}$ is an open neighbourhood of $\mathrm{p}(t_*) = \mathrm{p}$ in Q, there exists (by continuity) an open interval $I_* \subset I$ containing t_* s.t. $\gamma(I_*) \subset \mathcal{U}$. Then γ can be given the coordinate expression

$$\gamma_{\xi} := \xi^{-1} \circ \gamma|_{I_*} : I_* \to \mathbb{R}^n : t \mapsto q(t) := \xi^{-1}(\mathrm{p}(t))$$

(which could be shown to be C^∞) satisfying $q(t_*) = \xi^{-1}(\mathrm{p}(t_*)) = \xi^{-1}(\mathrm{p}) = q$.

By derivation at t_* of γ or, equivalently, of

$$\gamma|_{I_*} = \xi \circ \gamma_{\xi} : I_* \to Q : t \mapsto \mathrm{p}(t) = \xi(q(t))$$

we obtain [46]

$$\mathrm{v} = \dot{\mathrm{p}}(t_*) = d_{q(t_*)}\xi\,(\dot{q}(t_*)) = d_q\xi\,(\dot{q}(t_*))$$

Hence $\mathrm{v} \in \mathrm{Im}(d_q\xi)$.

Conversely, let $\mathrm{v} \in \mathrm{Im}(d_q\xi)$, i.e.

$$\mathrm{v} = d_q\xi\,(v)$$

for some $v \in \mathbb{R}^n$.

Choose $t_* \in \mathbb{R}$ and consider the affine mapping

$$\lambda : \mathbb{R} \to \mathbb{R}^n : t \mapsto q(t) := q + v\,(t - t_*)$$

with linear part $\Lambda : dt \in \mathbb{R} \mapsto v\,dt \in \mathbb{R}^n$, satisfying $q(t_*) = q \in W$ and $\dot{q}(t_*) = d_{t_*}\lambda\,(1) = \Lambda(1) = v$.

As W is an open neighbourhood of $q(t_*) = q$ in $\mathbb{R}^n$, there exists (by continuity) an open interval I_* containig t_* s.t. $\lambda(I_*) \subset W$ and then $\gamma_\xi := \lambda|_{I_*}$ is the C^∞ coordinate expression of

$$\gamma := \xi \circ \gamma_\xi : I_* \to Q : t \mapsto \mathrm{p}(t) = \xi(q(t))$$

Clearly, γ is a smooth parametrized curve of Q passing through

$$\mathrm{p}(t_*) = \xi(q(t_*)) = \xi(q) = \mathrm{p}$$

with derivative

$$\dot{\mathrm{p}}(t_*) = d_{q(t_*)}\xi\,(\dot{q}(t_*)) = d_q\xi\,(v) = \mathrm{v}$$

Hence $\mathrm{v} \in T_{\mathrm{p}}Q$. □

Exercise 37 *If ξ is a chart near $\mathrm{p} = \xi(q)$, then $d_q\xi : \mathbb{R}^n \to \mathrm{Im}(d_q\xi) = T_{\mathrm{p}}Q$ is the linear isomorphism which takes any n-tuple $v \in \mathbb{R}^n$ onto the vector $\mathrm{v} \in T_{\mathrm{p}}Q$ with components v in the basis*

$$\left(\frac{\partial \mathrm{p}}{\partial q^1}\Big|_q, \ldots, \frac{\partial \mathrm{p}}{\partial q^n}\Big|_q\right)$$

[46] See section 4.4.2, *Coordinate expression.*

Exercise 38
(i) *Coordinate domains*

If Q *is an n-dimensional smooth manifold, so is any coordinate domain* $\mathcal{U} \subset Q$ *and –for all* $\mathrm{p} \in \mathcal{U}$ *–*

$$T_{\mathrm{p}}\mathcal{U} = T_{\mathrm{p}}Q$$

(ii) *Flat manifolds*

If $\mathcal{A} \subset \mathcal{E}$ *is a flat manifold, modelled on* $A \subset E$*, then, for all* $\mathrm{p} \in \mathcal{A}$*,*

$$T_{\mathrm{p}}\,\mathcal{A} = A$$

(iii) *Open manifolds*

An open subset $M \subset \mathcal{E}$ *is a smooth manifold of maximal dimension, i.e.* $\dim(M) = \dim(\mathcal{E})$*, and –for all* $\mathrm{p} \in M$ *–*

$$T_{\mathrm{p}}M = E$$

Hint: (iii) A Cartesian system of $\mathcal{E}$, say $\phi : \mathbb{R}^m \to \mathcal{E}$ with $m := \dim(\mathcal{E})$, once restricted to the open subset $W := \phi^{-1}(M) \subset \mathbb{R}^m$, determines an m-dimensional global chart $\xi := \phi|_W : W \to \mathcal{E}$ onto M. □

Implicit Function Theorem

Most important is the following geometric version of the well known 'Implicit function theorem' of Analysis, concerning the manifold structure of 'loci', in a Euclidean space of dimension $m > 1$, described by means of algebraic equations.

Proposition 14 *Let*

$$g = (g_1, \ldots, g_\mu) : \mathcal{E} \to \mathbb{R}^\mu : \mathrm{p} \mapsto g(\mathrm{p}) = (g_1(\mathrm{p}), \ldots, g_\mu(\mathrm{p}))$$

be a continuous mapping, and

$$f = (f_1, \ldots, f_\kappa) : M \subset \mathcal{E} \to \mathbb{R}^\kappa : \mathrm{p} \longmapsto f(\mathrm{p}) = (f_1(\mathrm{p}), \ldots, f_\kappa(\mathrm{p}))$$

a differentiable map, defined on the open subset $M = g^{-1}(\mathbb{R}^+)^\mu$ *(where* $\mathbb{R}^+ := (0, +\infty)$*) and taking values in* $\mathbb{R}^\kappa$ *(where* $\kappa < m := \dim(\mathcal{E})$*), with surjective differential* $d_{\mathrm{p}}f : E \to \mathbb{R}^\kappa$ *at each point* p *of*

$$\begin{aligned} Q &:= f^{-1}(0) \\ &= \{\mathrm{p} \in \mathcal{E} \mid g_1(\mathrm{p}) > 0, \ldots, g_\mu(\mathrm{p}) > 0,\ f_1(\mathrm{p}) = 0, \ldots, f_\kappa(\mathrm{p}) = 0\} \end{aligned}$$

Q *is then an n-dimensional smooth manifold with*

$$n := m - \kappa$$

Proof: The above mentioned theorem of Analysis states that, for any $\mathrm{p} \in Q$, the fact of $d_{\mathrm{p}}f$ being surjective implies the existence of an open neighbourhood $\mathcal{U}$ of p in Q, an open subset W of $\mathbb{R}^n$ and a κ-tuple of real-valued functions $\varphi^1 : W \to \mathbb{R}, \dots, \varphi^\kappa : W \to \mathbb{R}$ such that

$$\mathcal{U} = \mathrm{Im}(\xi)$$

where $\xi : W \to \mathcal{E}$ is an n-dimensional chart on Q defined (with a suitable choice of a Cartesian systen $\varphi : \mathbb{R}^m \to \mathcal{E}$) by

$$(q^1, \dots, q^n) \stackrel{\xi}{\longmapsto} \varphi(\, x^1, \dots, x^n, x^{n+1}, \dots, x^{n+\kappa}\,)$$

with

$$x^1 = q^1, \dots\dots, x^n = q^n$$

and

$$x^{n+1} = \varphi^1(q^1, \dots, q^n), \dots\dots, x^{n+\kappa} = \varphi^\kappa(q^1, \dots, q^n)$$

That proves our claim. □

As a consequence,

Proposition 15 *If* $Q = f^{-1}(0)$, *then – for all* $\mathrm{p} \in Q$ –

$$T_{\mathrm{p}}Q = \mathrm{Ker}(d_{\mathrm{p}}f)$$

Proof: Let ξ be an n-dimensional chart of Q near $\mathrm{p} = \xi(q)$. As ξ is composable with f and $f \circ \xi = 0$, we have $d_{\mathrm{p}}f \circ d_q\xi = d_q(f \circ \xi) = 0$, whence

$$\mathrm{Im}(d_q\xi) \subset \mathrm{Ker}(d_{\mathrm{p}}f)$$

Moreover, the surjectivity of $d_{\mathrm{p}}f$ and the injectivity of $d_q\xi$ imply the dimensional result

$$\dim(\mathrm{Ker}(d_{\mathrm{p}}f)) = \dim(E) - \dim(\mathrm{Im}(d_{\mathrm{p}}f)) = m - \kappa = n = \dim(\mathrm{Im}(d_q\xi))$$

As a consequence,

$$\mathrm{Im}(d_q\xi) = \mathrm{Ker}(d_{\mathrm{p}}f)$$

Owing to Lemma ($\star$), that proves our claim. □

Exercise 39
(i) $\mathrm{Ker}(d_{\mathrm{p}}f) = \{\delta\mathrm{p} \in E \mid d_{\mathrm{p}}f_1(\delta\mathrm{p}) = 0, \dots, d_{\mathrm{p}}f_\kappa(\delta\mathrm{p}) = 0\}$.
(ii) *For any* $\mathrm{p} \in Q$ *(i.e.* $f(\mathrm{p}) = 0$*), if* $\delta\mathrm{p} \in T_{\mathrm{p}}Q$ *(i.e.* $d_{\mathrm{p}}f(\delta\mathrm{p}) = 0$*) and* $\mathrm{p} + \delta\mathrm{p} \in M$, *then* $f(\mathrm{p} + \delta\mathrm{p}) \approx 0$ *(i.e., up to higher order infinitesimals,* $\mathrm{p} + \delta\mathrm{p} \in Q$*).*

Examples of manifolds

Some examples of manifolds, arising from the geometric version of the Implicit function theorem, will now be given.

Hypersphere

In a Euclidean affine space $\mathcal{E}$ with $\dim(\mathcal{E}) > 1$, the *hypersphere* [47] of centre $\mathrm{o} \in \mathcal{E}$ and radius $r > 0$, is defined by

$$Q := \{\mathrm{p} \in \mathcal{E} \mid d(\mathrm{o}, \mathrm{p}) = r\}$$

Q can also be described as

$$Q = f^{-1}(0)$$

with

$$f : \mathcal{E} \to \mathbb{R} : \mathrm{p} \mapsto f(\mathrm{p}) := (\mathrm{p} - \mathrm{o}) \cdot (\mathrm{p} - \mathrm{o}) - r^2$$

As

$$f = \psi \cdot \psi - r^2$$

(where $\psi := -\mathrm{o} : \mathrm{p} \in \mathcal{E} \mapsto \psi(\mathrm{p}) := (\mathrm{p} - \mathrm{o}) \in E$ is an affine mapping with linear part id_E and r^2 is thought of as a constant function on $\mathcal{E}$), f is differentiable and its differential $d_\mathrm{p} f : E \to \mathbb{R}$ at any $\mathrm{p} \in \mathcal{E}$ is

$$d_\mathrm{p} f = 2\, \psi(\mathrm{p}) \cdot d_\mathrm{p}\psi = 2\, (\mathrm{p} - \mathrm{o}) \cdot \mathrm{id}_E$$

that is, for all $\mathrm{v} \in E$,

$$d_\mathrm{p} f(\mathrm{v}) = 2\, (\mathrm{p} - \mathrm{o}) \cdot \mathrm{v}$$

At each $\mathrm{p} \in Q$, where $\mathrm{p} - \mathrm{o} \neq 0$, $d_\mathrm{p} f$ does not vanish identically, that is to say, the dimension of $\mathrm{Im}(d_\mathrm{p} f) \subset \mathbb{R}$ is non-null, whence

$$\mathrm{Im}(d_\mathrm{p} f) = \mathbb{R}$$

which shows the surjectivity of $d_\mathrm{p} f$.

As a consequence, owing to Propositions 14 and 15, Q is a smooth manifold with

$$\dim(Q) = \dim(\mathcal{E}) - 1$$

and, at each $\mathrm{p} \in Q$, [48]

$$\begin{aligned} T_\mathrm{p} Q &= \mathrm{Ker}(d_\mathrm{p} f) \\ &= \{\mathrm{v} \in E \mid (\mathrm{p} - \mathrm{o}) \cdot \mathrm{v} = 0\} \\ &= \mathrm{Span}^\perp(\mathrm{p} - \mathrm{o}) \end{aligned}$$

[47] A hypersphere of $\mathcal{E}$ is just a *sphere* or a *circle*, if $\dim(\mathcal{E})$ is 3 or 2, respectively.
[48] Recall that $^\perp$ denotes the orthogonal complement in E.

Hypercylinder

In a Euclidean affine space $\mathcal{E}$ with $\dim(\mathcal{E}) > 1$, the *hypercylinder* of axis $\mathcal{A} = \mathrm{o} + \mathrm{Span}\,(\mathrm{u}) \subset \mathcal{E}$ (with $|\mathrm{u}| = 1$) and radius $r > 0$ is defined by

$$Q := \{\mathrm{p} \in \mathcal{E} \mid d(\mathrm{p}^*, \mathrm{p}) = r\}$$

where $\mathrm{p}^* := \mathrm{o} + (\,(\mathrm{p} - \mathrm{o}) \cdot \mathrm{u}\,)\mathrm{u}$ is the *orthogonal projection* of $\mathrm{p} \in \mathcal{E}$ onto $\mathcal{A}$.

Exercise 40 $\mathrm{p}^* := \mathrm{o} + (\,(\mathrm{p} - \mathrm{o}) \cdot \mathrm{u}\,)\mathrm{u} = \mathrm{p} - [(\mathrm{p} - \mathrm{o}) - (\,(\mathrm{p} - \mathrm{o}) \cdot \mathrm{u}\,)\mathrm{u}]$ *is the unique point of* $\mathcal{E}$ *such that* $\mathrm{p}^* \in \mathcal{A} = \mathrm{o} + \mathrm{Span}(\mathrm{u})$ *and* $\mathrm{p} - \mathrm{p}^* \in \mathrm{Span}^{\perp}(\mathrm{u})$.

Q can also be described as

$$Q = f^{-1}(0)$$

with

$$f : \mathcal{E} \to \mathbb{R} : \mathrm{p} \mapsto f(\mathrm{p}) := (\mathrm{p} - \mathrm{p}^*) \cdot (\mathrm{p} - \mathrm{p}^*) - r^2$$

As

$$f = \psi \cdot \psi - r^2$$

(where $\psi : \mathrm{p} \in \mathcal{E} \mapsto \psi(\mathrm{p}) := \mathrm{p} - \mathrm{p}^* = (\mathrm{p} - \mathrm{o}) - (\,(\mathrm{p} - \mathrm{o}) \cdot \mathrm{u}\,)\mathrm{u} \in E$ is an affine mapping with linear part $\Psi : \mathrm{v} \in E \mapsto \Psi(\mathrm{v}) := \mathrm{v} - (\,\mathrm{v} \cdot \mathrm{u}\,)\mathrm{u}$ and r^2 is thought of as a constant function on $\mathcal{E}$), f is differentiable and its differential $d_{\mathrm{p}} f : E \to \mathbb{R}$ at any $\mathrm{p} \in \mathcal{E}$ is

$$d_{\mathrm{p}} f = 2\,\psi(\mathrm{p}) \cdot d_{\mathrm{p}}\psi = 2\,(\mathrm{p} - \mathrm{p}^*) \cdot \Psi$$

that is, for all $\mathrm{v} \in E$,

$$d_{\mathrm{p}} f(\mathrm{v}) = 2\,(\mathrm{p} - \mathrm{p}^*) \cdot \Psi(\mathrm{v}) = 2\,(\mathrm{p} - \mathrm{p}^*) \cdot (\,\mathrm{v} - (\,\mathrm{v} \cdot \mathrm{u}\,)\mathrm{u}\,) = 2\,(\mathrm{p} - \mathrm{p}^*) \cdot \mathrm{v}$$

At each $\mathrm{p} \in Q$, where $\mathrm{p} - \mathrm{p}^* \neq 0$, $d_{\mathrm{p}} f$ does not vanish identically, that is to say, the dimension of $\mathrm{Im}(d_{\mathrm{p}} f) \subset \mathbb{R}$ is non-null, whence

$$\mathrm{Im}(d_{\mathrm{p}} f) = \mathbb{R}$$

which shows the surjectivity of $d_{\mathrm{p}} f$.

As a consequence, owing to Propositions 14 and 15, Q is a smooth manifold with

$$\dim(Q) = \dim(\mathcal{E}) - 1$$

and, at each $\mathrm{p} \in Q$,

$$\begin{aligned} T_{\mathrm{p}} Q &= \mathrm{Ker}(d_{\mathrm{p}} f) \\ &= \{\mathrm{v} \in E \mid (\mathrm{p} - \mathrm{p}^*) \cdot \mathrm{v} = 0\} \\ &= \mathrm{Span}^{\perp}(\mathrm{p} - \mathrm{p}^*) \end{aligned}$$

Half-hypercone

In a Euclidean affine space $\mathcal{E}$ with $\dim(\mathcal{E}) > 1$, the *half-hypercone* of vertex $\mathrm{o} \in \mathcal{E}$, axis $\mathcal{A} = \mathrm{o} + \mathrm{Span}\,(\mathrm{u}) \subset \mathcal{E}$ (with $|\mathrm{u}| = 1$) and contained in the open half-space $M := \{\mathrm{p} \in \mathcal{E} \mid (\mathrm{p}-\mathrm{o})\cdot \mathrm{u} > 0\}$ is defined by

$$
\begin{aligned}
Q : \;&=\; \left\{\mathrm{p} \in M \mid \angle((\mathrm{p}-\mathrm{o}),\mathrm{u}) = \frac{\pi}{4}\right\} \\
&=\; \left\{\mathrm{p} \in M \mid \arccos\frac{(\mathrm{p}-\mathrm{o})\cdot \mathrm{u}}{|\mathrm{p}-\mathrm{o}|} = \frac{\pi}{4}\right\} \\
&=\; \left\{\mathrm{p} \in M \mid \frac{(\mathrm{p}-\mathrm{o})\cdot \mathrm{u}}{|\mathrm{p}-\mathrm{o}|} = \cos\frac{\pi}{4}\right\} \\
&=\; \left\{\mathrm{p} \in M \mid \frac{(\mathrm{p}-\mathrm{o})\cdot \mathrm{u}}{|\mathrm{p}-\mathrm{o}|} = \frac{\sqrt{2}}{2}\right\} \\
&=\; \left\{\mathrm{p} \in M \mid (\mathrm{p}-\mathrm{o})\cdot(\mathrm{p}-\mathrm{o}) - 2\left((\mathrm{p}-\mathrm{o})\cdot \mathrm{u}\right)^2 = 0\right\}
\end{aligned}
$$

Q can also be described as

$$Q = f^{-1}(0)$$

with

$$f : M \to \mathbb{R} : \mathrm{p} \mapsto f(\mathrm{p}) := \langle(\mathrm{p}-\mathrm{o})|(\mathrm{p}-\mathrm{o})\rangle$$

where

$$\langle\quad|\quad\rangle : E \times E \to \mathbb{R} : (\mathrm{v},\mathrm{w}) \mapsto \langle \mathrm{v}|\mathrm{w}\rangle := \mathrm{v}\cdot\mathrm{w} - 2\,(\mathrm{v}\cdot\mathrm{u})(\mathrm{w}\cdot\mathrm{u})$$

Exercise 41 *Tha above mapping* $\langle\quad|\quad\rangle$ *is a* Minkowskian metric *on the vector space E, that is to say, a non-degenerate, symmetric, bilinear form admitting, in a suitable basis* $(\mathrm{e}_1,\dots,\mathrm{e}_{m-1},\mathrm{e}_m)$ *of* E*, the* signature $(1,\dots,1,-1)$*, i.e.* $\langle \mathrm{e}_i|\mathrm{e}_j\rangle = 0$ *for* $i \neq j$*,* $\langle \mathrm{e}_i|\mathrm{e}_i\rangle = 1$ *for all* $i = 1,\dots,m-1$ *and* $\langle \mathrm{e}_m|\mathrm{e}_m\rangle = -1$*.* [49]

Hint: In the Euclidean vector space E, choose an orthonormal basis $(\mathrm{e}_1,\dots,\mathrm{e}_{m-1},\mathrm{e}_m)$ with $\mathrm{e}_m = \mathrm{u}$. □

[49] A vector space E equipped with a Minkowskian metric is said to be a *Minkowski vector space* and an affine space $\mathcal{E}$ modelled on E a *Minkowski affine space.* $Q = f^{-1}(0)$ is then said to be a *Minkowski half-cone* in $\mathcal{E}$.

As

$$f = \langle \psi | \psi \rangle$$

(where $\psi := -\mathrm{o}|_M$ is the restriction to M of the affine mapping $-\mathrm{o}$, with linear part id_E), f is differentiable and its differential $d_{\mathrm{p}} f : E \to \mathbb{R}$ at any $\mathrm{p} \in M$ is

$$d_{\mathrm{p}} f = 2 \, \langle \psi(\mathrm{p}) | d_{\mathrm{p}} \psi \rangle = 2 \, \langle (\mathrm{p} - \mathrm{o}) | \mathrm{id}_E \rangle$$

that is, for all $\mathrm{v} \in E$,

$$d_{\mathrm{p}} f(\mathrm{v}) = 2 \, \langle (\mathrm{p} - \mathrm{o}) | \mathrm{v} \rangle$$

As $\mathrm{p} - \mathrm{o} \neq 0$ for all $\mathrm{p} \in M$ (and then for all $\mathrm{p} \in Q$), $d_{\mathrm{p}} f$ does not vanish identically, that is to say, the dimension of $\mathrm{Im}(d_{\mathrm{p}} f) \subset \mathbb{R}$ is non-null, whence

$$\mathrm{Im}(d_{\mathrm{p}} f) = \mathbb{R}$$

which shows the surjectivity of $d_{\mathrm{p}} f$.

As a consequence, owing to Propositions 14 and 15, Q is a smooth manifold with

$$\dim(Q) = \dim(\mathcal{E}) - 1$$

and, at each $\mathrm{p} \in Q$, [50]

$$\begin{aligned} T_{\mathrm{p}} Q &= \mathrm{Ker}(d_{\mathrm{p}} f) \\ &= \{ \mathrm{v} \in E \mid \langle (\mathrm{p} - \mathrm{o}) | \mathrm{v} \rangle = 0 \} \\ &= \mathrm{Span}^{\top} (\mathrm{p} - \mathrm{o}) \end{aligned}$$

Tangent bundle

The *tangent bundle* of a smooth manifold $Q \subset \mathcal{E}$ is the disjoint union of its tangent vector spaces, i.e.

$$TQ = \bigcup_{\mathrm{p} \in Q} (\{\mathrm{p}\} \times T_{\mathrm{p}} Q) \subset \mathcal{E} \times E$$

So, for any $(\mathrm{p}, \mathrm{v}) \in \mathcal{E} \times E$, we have

$$(\mathrm{p}, \mathrm{v}) \in TQ \iff \mathrm{p} \in Q \,, \; \mathrm{v} \in T_{\mathrm{p}} Q$$

[50] $^{\top}$ will denote the orthogonal complement defined by the Minkowskian metric in E.

Exercise 42
(i) *Flat manifolds*
The tangent bundle of a flat manifold $\mathcal{A} \subset \mathcal{E}$ *(affine subspace modelled on a vector subspace* $A \subset E$ *) is trivial, i.e.*

$$T\mathcal{A} = \mathcal{A} \times A$$

(ii) *Open manifolds*
The tangent bundle of an open manifold $M \subset \mathcal{E}$ *is trivial, i.e.*

$$TM = M \times E$$

In particular, the tangent bundle of the open manifold $M := \mathcal{E}$ is the Euclidean affine space (modelled on $E \times E$)

$$T\mathcal{E} = \mathcal{E} \times E$$

Hence, for any smooth manifold $Q \subset \mathcal{E}$,

$$TQ \subset T\mathcal{E}$$

4.6 Differential equations

A 'manifold-like' approach to differential equations on Euclidean affine spaces, will now be sketched.

4.6.1 First-order differential equations in implicit form

The kind of problem giving rise to 'first-order differential equations', will now be discussed.

First-order tangent lift

Let

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \gamma(t) = \mathrm{p}(t)$$

be a smooth parametrized curve of $\mathcal{E}$ (Euclidean affine space modelled on E).
Its *first-order tangent lift*

$$\dot{\gamma} : I \subset \mathbb{R} \to T\mathcal{E} : t \mapsto \dot{\gamma}(t) := (\mathrm{p}(t), \dot{\mathrm{p}}(t))$$

(which simultaneously describes $\mathrm{p}(t)$ and its first-order derivative $\dot{\mathrm{p}}(t)$) lives in the *first-order tangent bundle* $T\mathcal{E} = \mathcal{E} \times E$, i.e.

$$\mathrm{Im}(\dot{\gamma}) \subset T\mathcal{E}$$

The *graph* of $\dot{\gamma}$ is then the subset of the *first-order jet bundle* $\mathbb{R} \times T\mathcal{E}$ given by

$$\mathrm{Graph}(\dot{\gamma}) := \{(t, \dot{\gamma}(t))\}_{t \in I} = \{(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))\}_{t \in I} \subset \mathbb{R} \times T\mathcal{E}$$

The following exercise is to show that $\mathbb{R} \times T\mathcal{E}$ (resp. $T\mathcal{E}$) is entirely 'swept' by the graphs (resp. the orbits) of the first-order tangent lifts.

Exercise 43 *For any* $(t_o, \mathrm{p}_o, \mathrm{v}_o) \in \mathbb{R} \times T\mathcal{E}$ *, there exists a smooth parametrized curve* γ *of* $\mathcal{E}$ *s.t.* $(t_o, \mathrm{p}_o, \mathrm{v}_o) \in \mathrm{Graph}(\dot{\gamma})$ *(and then* $(\mathrm{p}_o, \mathrm{v}_o) \in \mathrm{Im}(\dot{\gamma})$ *).*

First-order differential equations

As a consequence, if a region – generally, a manifold –

$$\mathrm{D} \subset \mathbb{R} \times T\mathcal{E}$$

of the first-order jet bundle is assigned, the problem may arise of determining the smooth parametrized curves of $\mathcal{E}$, say $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$, s.t.

$$\mathrm{Graph}(\dot{\gamma}) \subset \mathrm{D} \tag{$\diamond$}$$

that is,

$$(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \in \mathrm{D}\,, \quad \forall\, t \in I \tag{$\diamond$}$$

Such a 'manifold' D – or the problem $(\diamond)$ itself, in the unknown γ – is called a 'time-dependent' *first-order differential equation in implicit form* on $\mathcal{E}$ (the space where the unknown lives).

If D is decomposable into a Cartesian product

$$\mathrm{D} = \mathbb{R} \times \mathcal{D}\,, \quad \mathcal{D} \subset T\mathcal{E}$$

then condition $(\diamond)$ is equivalent to

$$\mathrm{Im}(\dot{\gamma}) \subset \mathcal{D} \tag{$\circ$}$$

that is,

$$(\mathrm{p}(t), \dot{\mathrm{p}}(t)) \in \mathcal{D}\,, \quad \forall\, t \in I \tag{$\circ$}$$

In such a case, equation D (or $(\diamond)$) can as well be replaced by $\mathcal{D}$ (or $(\circ)$), which is said to be a 'time-independent' first-order differential equation in implicit form on $\mathcal{E}$.

The attributes 'time-dependent' and 'time-independent' are motivated in the following

Exercise 44 *Assigning* $\mathrm{D} \subset \mathbb{R}\times T\mathcal{E}$ *is the same as assigning a 'time-dependent' family* $\{\mathcal{D}_t\}_{t\in\mathbb{R}}$ *of subsets* $\mathcal{D}_t \subset T\mathcal{E}$ *through*

$$(\mathrm{p},\mathrm{v}) \in \mathcal{D}_t \Longleftrightarrow (t,\mathrm{p},\mathrm{v}) \in \mathrm{D}$$

(so $(\diamond)$ *means* $(\mathrm{p}(t),\dot{\mathrm{p}}(t)) \in \mathcal{D}_t\,,\ \forall\, t \in I$*). Condition*

$$\mathrm{D} = \mathbb{R}\times\mathcal{D}, \quad \mathcal{D} \subset T\mathcal{E}$$

is then equivalent to 'time-independence'

$$\mathcal{D}_t = \mathcal{D}, \ \ \forall t \in \mathbb{R}$$

Integrability and integration

A smooth parametrized curve γ of $\mathcal{E}$ satisfying $(\diamond)$ –or $(\circ)$ in the time-independent case– is said to be a *solution* of the equation.

A *maximal* solution is one that is *not* restriction of any other solution.

The problem of establishing the existence of solutions for a differential equation and the problem of determining (or, at least, discussing) such solutions, are called *integrability* and *integration*, respectively. [51]

Constraint manifolds

A smooth manifold $Q \subset \mathcal{E}$ is said to be a *constraint manifold* for the differential equation D, if

$$(t,\mathrm{p},\mathrm{v}) \in \mathrm{D} \Longrightarrow \mathrm{p} \in Q$$

The name is due to the fact that, in such a case, each solution γ of D is bound to live in Q, i.e.

$$\mathrm{Graph}(\dot{\gamma}) \subset \mathrm{D} \Longrightarrow \mathrm{Im}(\gamma) \subset Q$$

(whence $\mathrm{Im}(\dot{\gamma}) \subset TQ$).

First integrals

A *first integrals* of D is a function

$$f : Q \to V$$

[51] The names arise from the well known result of Analysis linking the operation of definite integration (or quadrature) with the probem of determining the primitives of a real-valued function (see 'Quadrature' at the end of the present section).

(Q being a constraint manifold for D and V any kind of space) which keeps constant along each solution γ of D, that is,

$$\mathrm{Graph}(\dot{\gamma}) \subset \mathrm{D} \Longrightarrow \mathrm{Im}(\gamma) \subset f^{-1}(c_o)$$

for some value of the constant $c_o \in \mathrm{Im}(f) \subset V$.

Each solution of D therefore lives in a *level subspace* $f^{-1}(c_o) \subset Q$ of f and the geometric properties of such subspaces may provide qualitative information on the behaviour of the solutions (which makes first integrals a very important tool of integration).

Normal form

A first-order differential equation $\mathrm{D} \subset \mathbb{R} \times T\mathcal{E}$ is said to be reducible to *normal form* on an *open manifold* $Q \subset \mathcal{E}$, if Q is a constraint manifold for D and, for any $(t, \mathrm{p}) \in \mathbb{R} \times Q$, there exists a unique $\mathrm{v} = \Gamma(t, \mathrm{p}) \in E$ s.t. $(t, \mathrm{p}, \mathrm{v}) \in \mathrm{D}$.

Such an equation can therefore be expressed as the graph of a 'time-dependent' mapping

$$\Gamma : \mathbb{R} \times Q \to E : (t, \mathrm{p}) \mapsto \Gamma(t, \mathrm{p})$$

that is,

$$\begin{aligned} \mathrm{D} &= \{(t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times T\mathcal{E} \mid (t, \mathrm{p}) \in \mathbb{R} \times Q \,,\ \mathrm{v} = \Gamma(t, \mathrm{p})\} \\ &= \{(t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times T\mathcal{E} \mid \mathrm{p} \in Q \,,\ \mathrm{v} = \Gamma(t, \mathrm{p})\} \end{aligned}$$

A solution of D is then any smooth $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$ satisfying, for all $t \in I$, $(t, \mathrm{p}(t), \dot{\mathrm{p}}(t)) \in \mathrm{D}$, i.e.

$$\mathrm{p}(t) \in Q \,, \quad \dot{\mathrm{p}}(t) = \Gamma(t, \mathrm{p}(t))$$

Exercise 45 *The graph* D *of* $\Gamma : \mathbb{R} \times Q \to E$ *satisfies condition* $\mathrm{D} = \mathbb{R} \times \mathcal{D}$, *iff* $\Gamma|_{\mathbb{R}\times\{\mathrm{p}\}} = \mathrm{const.} =: \Gamma(\mathrm{p})$ *for all* $\mathrm{p} \in Q$. *In such a case,* $\mathcal{D}$ *itself is reducible to normal form on* Q, *in the sense that it can be expressed as the graph of the above 'time-independent' mapping* $\Gamma : Q \to E$, *i.e.*

$$\mathcal{D} = \{(\mathrm{p}, \mathrm{v}) \in T\mathcal{E} \mid \mathrm{p} \in Q \,,\ \mathrm{v} = \Gamma(\mathrm{p})\}$$

A solution of D *(or* $\mathcal{D}$*) is then any smooth* $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$ *satisfying, for all* $t \in I$, $(\mathrm{p}(t), \dot{\mathrm{p}}(t)) \in \mathcal{D}$, *i.e.*

$$\mathrm{p}(t) \in Q \,, \quad \dot{\mathrm{p}}(t) = \Gamma(\mathrm{p}(t))$$

The integrability of a first-order differential equation in normal form D , is established – under the hypothesis of *smoothnes* (i.e. C^∞ differentiability) on Γ – by the following historical theorem of Analysis: [52]

Determinism theorem

For any choice of *Cauchy data*

$$(t_o\,, \mathrm{p}_o) \in \mathbb{R} \times Q$$

there *exists* a *unique* maximal solution to *Cauchy problem*

$$(\mathrm{D}, t_o\,, \mathrm{p}_o)$$

that is to say, a unique smooth parametrized curve $\gamma : t \in I \subset \mathbb{R} \mapsto \mathrm{p}(t) \in \mathcal{E}$, defined on an open interval $I \ni t_o$, which is a maximal solution of equation D and satisfies the *initial condition*

$$\mathrm{p}(t_o) = \mathrm{p}_o$$

(all of the other solutions to the above problem being just restrictions of the maximal one).

Remark The definition of reducibility to normal form on Q and the determinism theorem can be extended to the case of Q being any smooth manifold of $\mathcal{E}$. [53]

Ordinary first-order differential equations

From the above approach, we shall soon recover the *ordinary* first-order differential equations of elementary Analysis.

A *first-order differential-algebraic equation* on $\mathcal{E} := \mathbb{R}^m$ is the set of all the zeros of a κ-tuple of 'time-dependent' real-valued functions

$$f : \mathbb{R} \times T\mathcal{W} \to \mathbb{R}^\kappa$$

defined on the first-order jet bundle bundle $\mathbb{R} \times T\mathcal{W} = \mathbb{R} \times (\mathcal{W} \times \mathrm{R}^m)$ of an open manifold $\mathcal{W} \subset \mathbb{R}^m$, that is,

$$\begin{aligned} \mathrm{D} &= \{(t, x, y) \in \mathbb{R} \times T\mathbb{R}^m \mid (t, x, y) \in \mathbb{R} \times T\mathcal{W}\,,\ f(t, x, y) = 0\} \\ &= \{(t, x, y) \in \mathbb{R} \times T\mathbb{R}^m \mid x \in \mathcal{W}\,,\ f(t, x, y) = 0\} \end{aligned}$$

[52] The theorem of Analysis concerns ordinary first-order differential equations in normal form (see the next subsection), to which our case can be reduced with the aid of a global chart (Cartesian coordinates) on the open manifold Q (see Exercise 46).

[53] See the *Introduction* quoted in Preface (footnote [2]).

A solution of D is any smooth parametrized curve $t \in I \mapsto x(t) \in \mathbb{R}^m$ satisfying, for all $t \in I$, $(t, x(t), \dot{x}(t)) \in \mathrm{D}$, i.e.

$$x(t) \in \mathcal{W}\,, \quad f(t, x(t), \dot{x}(t)) = 0$$

The latter condition, which corresponds to a κ-tuple of scalar equalities, is usually said to be a *system of* κ *ordinary, first-order, differential equations in the* m*-tuple of unknown functions* $x(t)$.

D is reducible to normal form on $\mathcal{W}$, if it can be expressed as the graph of a 'time-dependent' mapping $X : \mathbb{R} \times \mathcal{W} \to \mathbb{R}^m$, that is,

$$\begin{aligned} \mathrm{D} &= \{(t, x, y) \in \mathbb{R} \times T\mathbb{R}^m \mid (t, x) \in \mathbb{R} \times \mathcal{W}\,,\ y = X(t, x)\} \\ &= \{(t, x, y) \in \mathbb{R} \times T\mathbb{R}^m \mid x \in \mathcal{W}\,,\ y = X(t, x)\} \end{aligned}$$

(in other words, $f(t, x, y) = 0$ is a system of algebraic equations which, for any $(t, x) \in \mathbb{R} \times \mathcal{W}$, admits a unique solution $y = X(t, x) \in \mathbb{R}^m$).

In such a case, the conditions characterizing the solutions of D read, for all $t \in I$,

$$x(t) \in \mathcal{W}\,, \quad \dot{x}(t) = X(t, x(t))$$

The latter condition is also called a (time-dependent) *dynamical system.* [54]

Exercise 46 *Prove that the problem of integrating a first-order differential equation in normal form on an open manifold of an* m*-dimensional affine Euclidean space* $\mathcal{E}$ *can be reduced, by means of Cartesian coordinates, to the problem of integrating a dynamical system on an open manifold of* $\mathbb{R}^m$.

If

$$f : T\mathcal{W} \to \mathbb{R}^\kappa$$

is a κ-tuple of 'time-independent' real-valued functions, the set of its zeros is a 'time-indepent' equation

$$\begin{aligned} \mathcal{D} &= \{(x, y) \in T\mathbb{R}^m \mid (x, y) \in T\mathcal{W}\,,\ f(x, y) = 0\} \\ &= \{(x, y) \in T\mathbb{R}^m \mid x \in \mathcal{W}\,,\ f(x, y) = 0\} \end{aligned}$$

A solution of $\mathcal{D}$ is then any smooth parametrized curve $t \in I \mapsto x(t) \in \mathbb{R}^m$ satisfying, for all $t \in I$, $(x(t), \dot{x}(t)) \in \mathcal{D}$, i.e.

$$x(t) \in \mathcal{W}\,, \quad f(x(t), \dot{x}(t)) = 0$$

The above $\mathcal{D}$ is reducible to normal form on $\mathcal{W}$, iff it can be expressed as the graph of a 'time-independent' mapping $X : \mathcal{W} \to \mathbb{R}^m$, that is,

$$\mathcal{D} = \{(x, y) \in T\mathbb{R}^m \mid x \in \mathcal{W}\,,\ y = X(x)\}$$

In such a case, the conditions characterizing the solutions of $\mathcal{D}$ read, for all $t \in I$,

$$x(t) \in \mathcal{W}\,, \quad \dot{x}(t) = X(x(t))$$

[54] The name arises from the local formulation of the law of Classical Dymamics, which results in a *special kind* of system of ordinary first-order differential equations in normal form.

Linear differential equations

A time-independent equation $\mathcal{D}$ in normal form (with $\mathcal{W} = \mathbb{R}^m$) is a *linear differential equation*, if $X : \mathbb{R}^m \to \mathbb{R}^m$ is a linear mapping.

In such a case, we have $X(x) = \alpha\, x$ with $\alpha \in gl(m, \mathbb{R})$ for all $x \in \mathbb{R}^m$, and then a solution of $\mathcal{D}$ is any smooth parametrized curve $t \in I \mapsto x(t) \in \mathbb{R}^m$ satisfying, for all $t \in I$,

$$\dot{x}(t) = \alpha\, x(t)$$

Exercise 47 *The maximal solution of Cauchy problem* $(\mathcal{D}, t_o, x_o)$*, for any choice of Cauchy data* $(t_o, x_o) \in \mathbb{R} \times \mathbb{R}^m$*, is given by*

$$t \in \mathbb{R} \mapsto x(t) = \Big(exp\,(t - t_o)\alpha\Big) x_o \in \mathbb{R}^m$$

where the exponential *is the sum of the convergent series* [55]

$$exp\,(t - t_o)\alpha \;:= \sum_{n=0}^{\infty} \frac{1}{n!} \Big((t - t_o)\alpha\Big)^n \;\in\; gl(m, \mathbb{R})$$

Quadrature

Once given a C^∞ differentiable function $g : \mathbb{R} \to \mathbb{R}$, consider the first-order differential equation D in normal form on $\mathbb{R}$ expressed as the graph of

$$X : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (t, x) \mapsto X(t, x) := g(t)$$

that is,

$$\mathrm{D} = \{(t, x, y) \in \mathbb{R} \times T\mathbb{R} \mid y = g(t)\}$$

Then a smooth function $t \in I \mapsto x(t) \in \mathbb{R}$ is a solution of D, iff it is a primitive of $g_{|I}$, that is, for all $t \in I$,

$$\dot{x}(t) = g(t)$$

According to a well known theorem of Analysis, for any $(t_o, x_o) \in \mathbb{R} \times \mathbb{R}$, the maximal solution of Cauchy problem (D, t_o, x_o) corresponds to a quadrature (definite integration), namely −for all $t \in \mathbb{R}$−

$$x(t) = x_o + \int_{t_o}^{t} g(s) ds$$

[55] In the sum below, the first term −for $n = 0$− is meant to be the identity matrix.

4.6.2 Second-order differential equations in implicit form

The above theory can naturally be extended to the 'second-order'.

Second-order tangent lift

Let

$$\gamma : I \subset \mathbb{R} \to \mathcal{E} : t \mapsto \gamma(t) = \mathrm{p}(t)$$

be a smooth parametrized curve of $\mathcal{E}$ (Euclidean affine space modelled on E).

Recall that the first-order tangent lift of γ is

$$\dot{\gamma} : I \subset \mathbb{R} \to T\mathcal{E} : t \mapsto \dot{\gamma}(t) = (\mathrm{p}(t), \dot{\mathrm{p}}(t))$$

Then the *second-order tangent lift of* γ, i.e. the first-order tangent lift of $\dot{\gamma}$, is

$$\ddot{\gamma} : I \subset \mathbb{R} \to T(T\mathcal{E}) : t \mapsto \ddot{\gamma}(t) := (\, \mathrm{p}(t), \dot{\mathrm{p}}(t);\ \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)\,)$$

Remark that $\ddot{\gamma}$ is bound to live in a special region of $T(T\mathcal{E}) = T(\mathcal{E} \times E) = (\mathcal{E} \times E) \times (E \times E)$, namely

$$\mathrm{Im}(\ddot{\gamma}) \subset T^2\mathcal{E}$$

with

$$T^2\mathcal{E} := \{(\mathrm{p}, \mathrm{v};\ \mathrm{u}, \mathrm{a}) \in T(T\mathcal{E}) \mid \mathrm{u} = \mathrm{v}\}$$

The *second-order tangent bundle* $T^2\mathcal{E}$ can naturally be expressed in the form

$$T^2\mathcal{E} = \mathcal{E} \times E \times E$$

by identifying

$$(\mathrm{p}, \mathrm{v};\ \mathrm{v}, \mathrm{a}) = (\mathrm{p}, \mathrm{v}, \mathrm{a})$$

As a consequence, the values of $\ddot{\gamma}$ can be denoted by

$$\ddot{\gamma}(t) = (\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t))$$

(they simultaneously describe $\mathrm{p}(t)$, its first-order derivative $\dot{\mathrm{p}}(t)$ and its second-order derivative $\ddot{\mathrm{p}}(t)$).

The *graph* of $\ddot{\gamma}$ is then the subset of the *second-order jet bundle* $\mathbb{R} \times T^2\mathcal{E}$ given by

$$\mathrm{Graph}(\ddot{\gamma}) := \{(t, \ddot{\gamma}(t))\}_{t\in I} = \{(t, \mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t))\}_{t\in I} \subset \mathbb{R} \times T^2\mathcal{E}$$

The following exercise is to show that $\mathbb{R} \times T^2\mathcal{E}$ (resp. $T^2\mathcal{E}$) is entirely 'swept' by the graphs (resp. the orbits) of the second-order tangent lifts.

Exercise 48 *For any* $(t_o, \mathrm{p}_o, \mathrm{v}_o, \mathrm{a}_o) \in \mathbb{R} \times T^2\mathcal{E}$ *, there exists a smooth parametrized curve* γ *of* $\mathcal{E}$ *s.t.* $(t_o, \mathrm{p}_o, \mathrm{v}_o, \mathrm{a}_o) \in \mathrm{Graph}(\ddot{\gamma})$ *(and then* $(\mathrm{p}_o, \mathrm{v}_o, \mathrm{a}_o) \in \mathrm{Im}(\ddot{\gamma})$*).*

Second-order differential equations

As a consequence, if a region – generally, a manifold –

$$\mathrm{D} \subset \mathbb{R} \times T^2\mathcal{E}$$

is assigned, the problem may arise of determining the smooth parametrized curves of $\mathcal{E}$, say $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$, s.t.

$$\mathrm{Graph}(\ddot{\gamma}) \subset \mathrm{D} \qquad (\diamond\diamond)$$

that is,

$$(t, \mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathrm{D}\,, \quad \forall\, t \in I \qquad (\diamond\diamond)$$

Such a 'manifold' D – or the problem $(\diamond\diamond)$ itself, in the unknown γ – is said to be a 'time-dependent' *second-order differential equation in implicit form* on $\mathcal{E}$ (the space where the unknown lives).

If D is decomposable into a Cartesian product

$$\mathrm{D} = \mathbb{R} \times \mathcal{D}$$

with

$$\mathcal{D} \subset T^2\mathcal{E}$$

then condition $(\diamond\diamond)$ is equivalent to

$$\mathrm{Im}(\ddot{\gamma}) \subset \mathcal{D} \qquad (\circ\circ)$$

that is,

$$(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathcal{D}\,, \quad \forall\, t \in I \qquad (\circ\circ)$$

In such a case, equation D (or $(\diamond\diamond)$) can as well be replaced by $\mathcal{D}$ (or $(\circ\circ)$), which is said to be a 'time-independent' second-order differential equation in implicit form on $\mathcal{E}$. [56]

Integrability and integration

A smooth parametrized curve γ of $\mathcal{E}$ satisfying $(\diamond\diamond)$ – or $(\circ\circ)$ in the time-independent case – is said to be a *solution* of the equation.

A *maximal* solution is one that is *not* restriction of any other solution.

The problem of establishing the existence of solutions for a differential equation and the problem of determining (or, at least, discussing) such solutions are called *integrability* and *integration*, respectively. [57]

[56] Exercise 44 on 'time-(in)dependence' can trivially be extended to the second-order.
[57] See 'Quadratures' in the present section.

Constraint manifolds

A smooth manifold $Q \subset \mathcal{E}$ is said to be a *constraint manifold* for the differential equation D, if

$$(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathrm{D} \Longrightarrow \mathrm{p} \in Q$$

The name is due to the fact that, in such a case, all of the solutions of D are bound to live in Q, i.e.

$$\mathrm{Graph}(\ddot{\gamma}) \subset \mathrm{D} \Longrightarrow \mathrm{Im}(\gamma) \subset Q$$

(whence $\mathrm{Im}(\dot{\gamma}) \subset TQ$).

First integrals

A *first integral* of D is a function

$$f : TQ \to V$$

(Q being a constraint manifold for D and V any kind of space) which keeps constant along the first-order tangent lift of each solution γ of D, i.e.

$$\mathrm{Graph}(\ddot{\gamma}) \subset \mathrm{D} \Longrightarrow \mathrm{Im}(\dot{\gamma}) \subset f^{-1}(c_o)$$

for some value of the constant $c_o \in \mathrm{Im}(f) \subset V$.

The first-order tangent lift of each solution of D therefore lives in a *level subspace* $f^{-1}(c_o) \subset TQ$ of f, and the geometric properties of such level subspaces may provide qualitative information on the behaviour of the solutions.

Normal form

A second-order differential equation $\mathrm{D} \subset \mathbb{R} \times T^2\mathcal{E}$ is said to be reducible to *normal form* on an *open manifold* $Q \subset \mathcal{E}$, if Q is a constraint manifold for D and, for any $(t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times TQ$, there exists a unique $\mathrm{a} = \Gamma(t, \mathrm{p}, \mathrm{v}) \in E$ s.t. $(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathrm{D}$.

Such an equation can therefore be expressed as the graph of a 'time-dependent' mapping

$$\Gamma : \mathbb{R} \times TQ \to E : (t, \mathrm{p}, \mathrm{v}) \mapsto \Gamma(t, \mathrm{p}, \mathrm{v})$$

that is,

$$\begin{aligned} \mathrm{D} &= \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \mid (t, \mathrm{p}, \mathrm{v}) \in \mathbb{R} \times TQ\ ,\ \mathrm{a} = \Gamma(t, \mathrm{p}, \mathrm{v})\} \\ &= \{(t, \mathrm{p}, \mathrm{v}, \mathrm{a}) \in \mathbb{R} \times T^2\mathcal{E} \mid \mathrm{p} \in Q\ ,\ \mathrm{a} = \Gamma(t, \mathrm{p}, \mathrm{v})\} \end{aligned}$$

A solution of D is then any smooth $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$ satisfying, for all $t \in I$, $(t, \mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathrm{D}$, i.e.

$$\mathrm{p}(t) \in Q\,, \quad \ddot{\mathrm{p}}(t) = \Gamma(t, \mathrm{p}(t), \dot{\mathrm{p}}(t))$$

Exercise 49 *The graph* D *of* $\Gamma : \mathbb{R} \times TQ \to E$ *satisfies condition* $\mathrm{D} = \mathbb{R} \times \mathcal{D}$ *, iff* $\Gamma|_{\mathbb{R}\times\{(\mathrm{p},\mathrm{v})\}} = \text{const.} =: \Gamma(\mathrm{p},\mathrm{v})$ *for all* $(\mathrm{p},\mathrm{v}) \in TQ$ *. In such a case,* $\mathcal{D}$ *itself is reducible to normal form on* Q *, in the sense that it can be expressed as the graph of the above 'time-independent' mapping* $\Gamma : TQ \to E$ *, i.e.*

$$\begin{aligned}\mathcal{D} &= \{(\mathrm{p},\mathrm{v},\mathrm{a}) \in T^2\mathcal{E} \mid (\mathrm{p},\mathrm{v}) \in TQ\ ,\ \mathrm{a} = \Gamma(\mathrm{p},\mathrm{v})\} \\ &= \{(\mathrm{p},\mathrm{v},\mathrm{a}) \in T^2\mathcal{E} \mid \mathrm{p} \in Q\ ,\ \mathrm{a} = \Gamma(\mathrm{p},\mathrm{v})\}\end{aligned}$$

A solution of D *(or* $\mathcal{D}$ *) is then any smooth* $\gamma : t \in I \mapsto \mathrm{p}(t) \in \mathcal{E}$ *satisfying, for all* $t \in I$ *,* $(\mathrm{p}(t), \dot{\mathrm{p}}(t), \ddot{\mathrm{p}}(t)) \in \mathcal{D}$ *, i.e.*

$$\mathrm{p}(t) \in Q\,, \quad \dot{\mathrm{p}}(t) = \Gamma(\mathrm{p}(t), \dot{\mathrm{p}}(t))$$

The integrability of a second-order differential equation in normal form D, is still established – under the hypothesis of *smoothness* (i.e. C^∞ differentiability) on Γ – by the determinism theorem, which takes now the following expression:

Determinism theorem

For any choice of *Cauchy data*

$$(t_o\,, (\mathrm{p}_o\,, \mathrm{v}_o)) \in \mathbb{R} \times TQ$$

there *exists* a *unique* maximal solution to *Cauchy problem*

$$(\mathrm{D}, t_o\,, (\mathrm{p}_o\,, \mathrm{v}_o))$$

that is to say, a unique smooth parametrized curve $\gamma : t \in I \subset \mathbb{R} \mapsto \mathrm{p}(t) \in \mathcal{E}$, defined on an open interval $I \ni t_o$, which is a maximal solution of equation D and satisfies the *initial conditions*

$$(\mathrm{p}(t_o), \dot{\mathrm{p}}(t_o)) = (\mathrm{p}_o\,, \mathrm{v}_o)$$

(all of the other solutions to the above problem being just restrictions of the maximal one).

Remark The definition of reducibility to normal form on Q and the determinism theorem can be extended to the case of Q being any smooth manifold of $\mathcal{E}$. [58]

[58] See the *Introduction* quoted in Preface (footnote [2]).

Ordinary second-order differential equations

From the above approach, we shall now recover the *ordinary* second-order differential equations of elementary Analysis.

A *second-order differential-algebraic equation* on $\mathcal{E} := \mathbb{R}^n$ is the set of all the zeros of a κ-tuple of 'time-dependent' real-valued functions

$$f : \mathbb{R} \times T^2W \to \mathbb{R}^\kappa$$

defined on the second-order jet bundle $\mathbb{R} \times T^2W := \mathbb{R} \times (W \times \mathbb{R}^n \times \mathbb{R}^n)$ of an open manifold $W \subset \mathbb{R}^n$, that is,

$$\begin{aligned} \mathrm{D} &= \{(t,q,v,a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid (t,q,v,a) \in \mathbb{R} \times T^2W \,,\ f(t,q,v,a) = 0\} \\ &= \{(t,q,v,a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid q \in W \,,\ f(t,q,v,a) = 0\} \end{aligned}$$

A solution of D is any smooth parametrized curve $t \in I \mapsto q(t) \in \mathbb{R}^n$ satisfying, for all $t \in I$, $(t, q(t), \dot{q}(t), \ddot{q}(t)) \in \mathrm{D}$, i.e.

$$q(t) \in W \,, \quad f(t, q(t), \dot{q}(t), \ddot{q}(t)) = 0$$

The latter condition, which corresponds to a κ-tuple of scalar equalities, is usually said to be a *system of κ ordinary, second-order,differential equations in the n-tuple of unknown functions $q(t)$.*

D is reducible to normal form on W, if it can be expressed as the graph of a 'time-dependent' mapping $X : \mathbb{R} \times TW \to \mathbb{R}^n$, that is,

$$\begin{aligned} \mathrm{D} &= \{(t,q,v,a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid (t,q,v) \in \mathbb{R} \times TW \,,\ a = X(t,q,v)\} \\ &= \{(t,q,v,a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid q \in W \,,\ a = X(t,q,v)\} \end{aligned}$$

(in other words, $f(t,q,v,a) = 0$ is a system of algebraic equations which admits – for any $(t,q,v) \in \mathbb{R} \times TW$ – a unique solution $a = X(t,q,v) \in \mathbb{R}^n$).

In such a case, the conditions characterizing the solutions of D read, for all $t \in I$,

$$q(t) \in W \,, \quad \ddot{q}(t) = X(t, q(t), \dot{q}(t))$$

If

$$f : T^2W \to \mathbb{R}^\kappa$$

is a κ-tuple of 'time-independent' real-valued functions, the set of its zeros is a time-independent equation

$$\begin{aligned} \mathcal{D} &:= \{(q,v,a) \in T^2\mathbb{R}^n \mid (q,v,a) \in T^2W \,,\ f(q,v,a) = 0\} \\ &= \{(q,v,a) \in T^2\mathbb{R}^n \mid q \in W \,,\ f(q,v,a) = 0\} \end{aligned}$$

A solution of $\mathcal{D}$ is then any smooth parametrized curve $t \in I \mapsto q(t) \in \mathbb{R}^n$ satisfying, for all $t \in I$,

$$q(t) \in W\,, \quad f(q(t), \dot{q}(t), \ddot{q}(t)) = 0$$

The above $\mathcal{D}$ is reducible to normal form on W, iff it can be expressed as the graph of a 'time-independent' mapping $X : TW \to \mathbb{R}^n$, that is ,

$$\begin{aligned}\mathcal{D} &= \{(q,v,a) \in T^2\mathbb{R}^n \mid (q,v) \in TW\,,\ a = X(q,v)\} \\ &= \{(q,v,a) \in T^2\mathbb{R}^n \mid q \in W\,,\ a = X(q,v)\}\end{aligned}$$

In such a case, the conditions characterizing the solutions of $\mathcal{D}$ read, for all $t \in I$,

$$q(t) \in W\,, \quad \ddot{q}(t) = X(q(t), \dot{q}(t))$$

Linear differential equations

A time-independent equation $\mathcal{D}$ in normal form (with $W = \mathbb{R}^n$) is a *linear differential equation*, if $X : T\mathbb{R}^n = \mathbb{R}^{2n} \to \mathbb{R}^n$ is a linear mapping.

In such a case, we have $X(q,v) = \alpha\, q + \beta\, v$ with $\alpha, \beta \in gl(n, \mathbb{R})$ for all $(q,v) \in \mathbb{R}^{2n}$, and then a solution of $\mathcal{D}$ is any smooth parametrized curve $t \in I \mapsto q(t) \in \mathbb{R}^n$ satisfying, for all $t \in I$,

$$\ddot{q}(t) = \alpha\, q(t) + \beta\, \dot{q}(t)$$

After a first-order reformulation, the above linear equation can be integrated through the 'exponential' method. [59]

Quadratures

Once given a C^∞ differentiable function $g : \mathbb{R} \to \mathbb{R}$, consider the second-order differential equation D in normal form on $\mathbb{R}$ expressed as the graph of

$$X : \mathbb{R} \times T\mathbb{R} \to \mathbb{R} : (t,q,v) \mapsto X(t,q,v) := g(t)$$

that is,

$$\mathrm{D} := \{(t,q,v,a) \in \mathbb{R} \times T^2\mathbb{R} \mid a = g(t)\}$$

Then a smooth function $t \in I \mapsto q(t) \in \mathbb{R}$ is a solution of D, iff, for all $t \in I$,

$$\ddot{q}(t) = g(t)$$

[59] See *First-order reformulation* at the end of the present section and *Linear differential equations* in section 4.6.1.

For any $(t_o, q_o, v_o) \in \mathbb{R}\times T\mathbb{R}$, the maximal solution of Cauchy problem $(\mathrm{D}, t_o, q_o, v_o)$ corresponds to a double quadrature, namely –for all $t \in \mathbb{R}$–

$$\dot{q}(t) = f(t) := v_o + \int_{t_o}^{t} g(s)\, ds$$

and

$$q(t) = q_o + \int_{t_o}^{t} f(s)\, ds$$

First-order reformulation

A second-order equation can be regarded as a *special kind* of first-order equation, as follows.

Consider a second-order differential-algebraic equation on $\mathbb{R}^n$ given by

$$\mathrm{D} = \{(t, q, v, a) \in \mathbb{R} \times T^2\mathbb{R}^n \mid q \in W\,,\ f(t, q, v, a) = 0\}$$

whose solutions $t \in I \mapsto q(t) \in \mathbb{R}^n$ are characterized by condition (for all $t \in I$)

$$(t, q(t), \dot{q}(t), \ddot{q}(t)) \in \mathrm{D}$$

that is,

$$q(t) \in W\,, \quad f(t, q(t), \dot{q}(t), \ddot{q}(t)) = 0$$

Turning back to the original definition of second-order tangent bundle, that is, $T^2\mathbb{R}^n := \{(q, v;\, u, a) \in T(T\mathbb{R}^n) \mid u = v\}$, D can be re-described as the first-order differential-algebraic equation on $T\mathbb{R}^n = \mathbb{R}^{2n}$ given by

$$\mathrm{D}' = \{(t;\, q, v;\, u, a) \in \mathbb{R} \times T\mathbb{R}^{2n} \mid q \in W\,,\ u = v\,,\ f(t, q, v, a) = 0\}$$

whose solutions $t \in I \mapsto x(t) = (q(t), v(t)) \in \mathbb{R}^{2n}$ of D' are characterized by condition (for all $t \in I$)

$$(t, x(t), \dot{x}(t)) = (t;\, q(t), v(t);\, \dot{q}(t), \dot{v}(t)) \in \mathrm{D}'$$

that is,

$$q(t) \in W\,, \quad \dot{q}(t) = v(t)\,, \quad f(t, q(t), v(t), \dot{v}(t)) = 0$$

Clearly, $q(t)$ is a solution of D, iff it is the first projection of a solution $x(t)$ of D' (and $x(t)$ is a solution of D', iff it is the tangent lift of a solution $q(t)$ of D).

So the problem of integrating a second-order equation D on $\mathbb{R}^n$ shifts to the problem of integrating its first-order reformulation D' on $T\mathbb{R}^n = \mathbb{R}^{2n}$.

Exercise 50 *In the case of reducibility to normal form, if* $\mathrm{D} \subset \mathbb{R} \times T^2\mathbb{R}^n$ *is the graph of*

$$X : \mathbb{R} \times \mathcal{W} \to \mathbb{R}^n$$

($\mathcal{W} = TW$*, open manifold of* $\mathbb{R}^{2n}$*, being the tangent bundle of an open manifold* W *of* $\mathbb{R}^n$ *), then* $\mathrm{D}' \subset \mathbb{R} \times T\mathbb{R}^{2n}$ *is the graph of*

$$X' = (pr_3, X) : \mathbb{R} \times \mathcal{W} \to \mathbb{R}^{2n}$$

($pr_3 : \mathbb{R} \times \mathcal{W} \to \mathbb{R}^n$ *being the projection of* $\mathbb{R} \times \mathcal{W} = \mathbb{R} \times W \times \mathbb{R}^n$ *onto its third factor).*