Mathematical Representation of Empirical Phenomena

A Case Study of 18th-Century Mathematical and Mechanical Concepts

by

Serban Dragulin

B.Sc. Computer Science, Polytechnic University of Bucharest, 2006
B.A. Philosophy, University of Bucharest, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Arts

in

THE FACULTY OF GRADUATE STUDIES
(Philosophy)

The University Of British Columbia
(Vancouver)

December 2011

© Serban Dragulin, 2011
Abstract

I suggest a theory of scientific models which considers that a model is composed of three parts: an idealized system, a theoretical description and a set of mathematical equations. Each component is connected to the previous one by bridge principles, i.e. any assertion, whether justified or simply postulated, which establishes a correspondence between two parts of a model. The goal of my research was to find within 18th century mechanics a set of laws that could be considered mediators between the parts of models. The greater part is dedicated to an analysis of the conceptual developments made by Euler and Lagrange in the 18th century both in mathematics and in mechanics. In the conclusion, I show that the relations between the parts of models are too complex to be expressed in a single, unifying assertion.
# Table of Contents

Abstract ................................................................. ii

Table of Contents ...................................................... iii

List of Figures ........................................................... v

Glossary ................................................................. vi

1 Introduction ............................................................ 1

2 A Textbook Approach .................................................. 9
   2.1 The Simple Linearized Pendulum ................................. 10
      2.1.1 The Damped Linearized Pendulum ......................... 12
   2.2 The Compound Pendulum ........................................... 15
   2.3 The Double Pendulum ............................................ 16
   2.4 Breaking Down the Formulas ..................................... 20
   2.5 The Structure of a Scientific Model ............................ 28
   2.6 Conclusions ..................................................... 32

3 The Status of Science in the Late 17th, Early 18th Centuries ....... 34
   3.1 18th-century Physics ............................................. 34
3.2 The Status of Mathematics Before Euler ........................................ 38
  3.2.1 Newton ........................................................................ 42

4 Foundational Research in the 18th Century: Euler ......................... 50
  4.1 Foundations of Calculus ...................................................... 51
    4.1.1 Eulerian Functions ..................................................... 51
    4.1.2 Infinitely Small Quantities ............................................ 63
    4.1.3 Calculus of Finite Differences (CFD) and Infinitesimal Calculus (CI) ... 70
  4.2 Foundations of Mechanics ................................................... 72

5 Foundational Research in the 18th Century: Lagrange ..................... 82
  5.1 Lagrange’s Criticism of His Predecessors ................................ 84
  5.2 Foundations of Lagrangian Calculus ...................................... 85
  5.3 Euler and Lagrange on the Calculus of Variations .................... 96
    5.3.1 Euler ...................................................................... 98
    5.3.2 Lagrange .................................................................. 101
  5.4 The Project of Mechanics in Lagrange’s Méchanique Analitique ....... 106

6 Conclusions ............................................................................. 130

Bibliography ............................................................................. 139
List of Figures

Figure 2.1  Simple pendulum ................................................................. 10
Figure 2.2  Compound pendulum ......................................................... 15
Figure 2.3  Double pendulum ............................................................... 17
Figure 3.1  Variable quantities of a curve ............................................ 41
Figure 3.2  Curve quadrature in Newton’s *De quadratura* ..................... 48
Figure 4.1  Prop. 14, Ch.2, *Mechanica* ............................................. 80
Figure 5.1  Graphical representation of Proposition III of Euler’s *Methodus* 98
Figure 5.2  Lagrange’s illustration of pvv ............................................. 113
Figure 5.3  The reduction of a curve to polygons ................................... 122
Glossary

**cfd**  Calculus of finite differences

**ci**  Calculus of infinitesimals

**el**  Euler-Lagrange equation

**pcf**  Principle of composition of forces

**pda**  D’Alembert’s Principle

**pl**  Principle of the lever

**pla**  Principle of least action

**pp**  Principle of pulleys

**pvv**  Principle of virtual velocities

**pww**  Principle of virtual work
Chapter 1

Introduction

It has become common philosophical practice to construe models as playing a representational role in science. For instance, Ronald Giere claims that

What is special about models is that they are designed so that elements of the model can be identified with features of the real world. This is what makes it possible to use models to represent aspects of the world. [Giere, 2004, p. 747]

R.I.G Hughes has a weaker approach, suggesting that

if we examine a theoretical model [...] we shall achieve some insight into the kind of representation that it provides. [Hughes, 1997, p. S329]

Daniela Bailer-Jones also supports the representational role of models:

the term “representation” has been introduced to capture how scientific models can be about empirical phenomena. [Bailer-Jones, 2003, p. 60]

while for Mauricio Suarez the aim of a theory of scientific representation is

to lay down the general conditions that such disparate models must meet to carry out a representational function: It does not need to stipulate the conditions for accurate, true, or complete, representation. [Suarez, 2004, p. 767]
Despite its importance for philosophical theories of scientific representation, the concept of model has not yet been clearly defined. In simpler terms, we do not know how to answer the question: “What is a scientific model?” Some properties of models are known and have been thoroughly investigated in the literature. For instance, we know that models can be about classes of empirical phenomena, that they involve processes of abstraction and idealization and that they are developed to explain and predict phenomena. However, summing up these characteristics still will not tell us what a model really is.

The main aim of this paper is to suggest and defend a theory of scientific models as representations. In doing so, I will assume that models have a representational role in science and will try to define models focusing on those properties in virtue of which a model can be considered to represent a class of phenomena. My suggestion is that a model has three components: an idealized system, a theoretical description of this system and mathematical formalisms. The idealized system is defined by a series of sentences expressed in natural language which identify the components of the system and the relations between them. The system is idealized because there need be no perfect real-world correspondent to it, e.g. there are no frictionless pendulums.

The theoretical description is a set of mathematical equations which express the relations between the theoretical entities of the system, e.g. force, kinetic or potential energy, wave amplitudes, and so on. An example would be the equation $\|\vec{F}_f\| = 0$ which simply states that the force of friction, represented by $\vec{F}_f$, is null. By using theoretical terms to describe the idealized system, this layer connects the model with a particular theory. In consequence, the model will necessarily be associated with a theory. This explains the position of Mary Morgan and Margaret Morrison in [Morgan and Morrison, 1999] which views models as mediators between theories and empirical phenomena. The present view also agrees with [Morgan and Morrison, 1999] in that it considers the process of modelling to be independent
of a particular theory. For instance, the model of the simple pendulum is not *deduced* from Newton’s laws of motion, but from the laws of motion in conjunction with a series of independently justified principles and the theoretical description of the idealized system.

Finally, the mathematical layer includes the equations of motion of the system which have been deduced following the procedure outlined above, i.e. by applying mathematical methods and principles to the theoretical description and the laws and principles related to the theory. One example of a mathematical principle, belonging to the calculus of variations, is the Euler-Lagrange equation, \( \frac{\partial L}{\partial q} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \), which is used to obtain the equations of motion of a system. Only when \( L \) is interpreted as the difference of the kinetic and the potential energy of a system can the Euler-Lagrange equation be applied to mechanics. In the following chapters I will focus on the relations between these three layers in general, and on the relation between the mathematical and the theoretical layer in particular. I will not address the other important problems concerning models, that is, “Why and how do models represent phenomena?” but only sketch a possible answer, leaving a more in-depth analysis of this topic for further research.

Considering the definition of scientific model just given, and the aim of the present work, an ideal starting point for investigation will be 18th-century mechanics and calculus. There are four reasons why a historical analysis of the development of 18th-century science is a good way to support the assumption that scientific models have a three-layered structure. First, 18th-century mathematicians moved away from the geometrical procedures of their predecessors. This is witnessed in Euler’s *Introductio in analysin infinitorum* and Lagrange’s *Théorie des fonctions analytiques, Leçons sur le calcul des fonctions*, and *Méchanique Analytique*, as well as in the works of d’Alembert and Johann and Jakob Bernoulli. The emphasis on the analytical expression of a functional relation, disconnected from its previous geometrical representation, allowed for a deductive formulation of mechanical principles. In turn,
these principles were used to represent and solve various problems in mechanics. Finding an answer to the questions “What justified 18th-century mathematicians and physicists to apply the new mathematical methods to physical problems?” and “How did they proceed in this direction?” would provide a good explanation of the connection between the first two layers of a model (the mathematical and the theoretical).

A second reason why an incursion into the history of 18th-century mechanics is relevant to our present purposes is that during this period we see the emergence of new theoretical concepts, such as force, and a reconsideration and systematization of older ones, in particular of the concept of vis viva.

Thirdly, 18th-century scientists addressed foundational issues of both calculus and mechanics, and in doing so they also discussed the role of idealization and abstraction in the scientific process. The interest in foundational issues introduced new concepts to mathematics (functions, derivative, etc.), as well as a higher degree of rigor to a chaotic mathematical domain.

Finally, before this period there was no clear distinction between calculus, mechanics and geometry, and the great achievement of the 18th century was to separate each theory by introducing fundamental algorithms. The domains of these sciences were separated, e.g. it was known what the object study of geometry was, but geometry still employed mechanical concepts. Finding algorithms to solve classes of problems was the main motivation for the separation of the calculus from mechanics during the 18th century and by analyzing the gradual separation of the two domains we will understand better the separation of the mathematical and theoretical layers.

The following chapters are organized following the structure of the presentation above. The first chapter presents a definition of a model, following the algorithm used to deduce the equations of motion of the simple and double pendulum. Conceptually, the chapter is
divided in two parts. In the first part I present the algorithms for deducing the equations of motion, while in the second part I show that the algorithmic procedure follows the three-layered definition of model.

The second chapter presents the status of science before Euler’s Mechanica and Introductio in analysin infinitorum. The aim is to introduce a term of comparison with the 18th-century development of the calculus and mechanics. Here I introduce a taxonomy of the domains of science in the pre-Euler era and discuss the status of experimental practices within these domains. Moreover, in this chapter I analyze, following [Boyer, 1949], [Kline, 1990] and [Grabiner, 2005] the geometrical nature of the differential calculus and mechanics before the ground-breaking work of Euler and Lagrange. As will be shown, 17th-century mathematicians, Newton and Leibniz in particular, had problems proving that the methods of the differential calculus could meet the standards of rigor of Euclidean geometry. The main problem was to show that by using infinitely small quantities, or differentials, the results of the calculus were not approximations of their rigorous geometrical counterparts. The main objections to the differential calculus were those raised in [Berkeley, 2009], reiterated in [Carnot and Browell, 1832] and by Lagrange in [Lagrange, 1806] and [Lagrange, 1797].

With minor differences, all accounts of the differential calculus before Lagrange’s work followed the same pattern. The differential calculus is concerned with the relative rates of change of quantities. In other words, suppose that \( y \) is a function of \( x \), then the problem is to calculate, for an infinitesimal increment \( \omega \) of \( x \), the ratio between \( \omega \) and the corresponding increment of \( y \). For instance, if \( y = x^2 \), then for \( x + \omega \), \( y \) will be \( x^2 + 2x\omega + \omega^2 \). The difference between the value of \( y \) for \( x \) and its value for \( x + \omega \) will be \( 2x\omega + \omega^2 \). The ratio \( \frac{dy}{dx} \) will be \( 2x + \omega \). At this point, the approaches to the differential calculus consider \( \omega \) a negligible quantity, 0, or as tending towards 0. [Berkeley, 2009] pointed out that if \( \omega \) is negligible, then the calculus is not rigorous, for these minor errors could add up in an unpredictable manner.
so that the end result of the methods of differential calculus will be uncertain. If \( \omega = 0 \) then \( dx = \omega = 0 \), so that calculating \( \frac{dy}{dx} \) will inevitably require a division by zero. If \( \omega \to 0 \), then it is not clear whether \( \omega \) is an assignable or unassignable quantity. These issues were partly solved by Lagrange’s algebraic approach to the foundations of calculus, as pointed out by Craig Fraser in [Fraser, 1987] and argued in chapter 5.

The third chapter investigates Euler’s contributions to the foundations of the calculus and mechanics. An important concept in Euler’s work, presented in *Introductio in analysin infinitorum*, was that of function. To appreciate the magnitude of this breakthrough, I present briefly the history of the concept of function, following [Youschkevitch, 1976], [Boyer, 1949], [Kline, 1990]. The history of the concept of function will be better understood if we assume the distinction drawn by Giovanni Ferraro [Ferraro, 2000] between *functional relations* and *functional expressions*. The concept of functional relation is nothing else than the perceived correlation between two or more quantities. However, a functional relation can be expressed as a table of values, as a geometrical curve, or as an analytical formula.

Euler’s breakthrough was to define functions analytically. A function was no longer just a curve expressing the relation between an abscissa and an ordinate, but a mathematical formula expressing the relation between two or more variables\(^1\). This was a step forward in separating the calculus from its geometrical origins, but Euler’s mathematical methods still relied on infinitely small quantities and as such they were still open to Berkeley’s objections.

Moreover, the application of the differential calculus to Euler’s mechanics, as presented in *Mechanica*, partly depends on an analogy between the mathematical concept of infinitesimal and Euler’s infinitely small point of mass. Given the conceptual problems of infinitesimals, using them to represent motion would be unjustified. However, there is more to Euler’s mechanics than this. I agree with Dieter Suisky’s claim in [Suisky, 2009] that the foundation

\(^1\)This will be explained in greater detail in chapters 2 and 3.
of Euler’s mechanics is the distinction between internal and external principles, i.e. between rest and motion as essential characteristics of bodies and forces as causes of change. To be more precise, Euler wanted to construct his mechanics on \textit{a priori} principles. These principles were the essential properties of bodies deduced through a purely rational process.

He found that the fundamental property of bodies is impenetrability. From this follow the other three essential properties: extent, mobility and persistence. In this sense, impenetrability is an \textit{internal principle} of mechanics. The external principle is that of force. For Euler, forces are \textit{necessarily} external and the only causes of motion, since it is in the nature of bodies to remain at rest. From this distinction Euler deduces Newton’s equations of motion. In this sense, Euler’s mechanics can be viewed as an attempt to provide an \textit{a priori} justification of Newtonian mechanics.

The fourth chapter is dedicated to Lagrange’s work on the foundations of the calculus and mechanics. As Craig Fraser observes in [Fraser, 1985], [Fraser, 1987], Lagrange’s views on the foundations of the differential calculus and, implicitly, of the calculus of variations, can be divided into two periods. The first period begins with Lagrange’s letters to Euler in the 1750s and ends with his \textit{Méchanique Analitique} in 1788. During this period Lagrange assumed the views of his predecessors on the differential calculus and developed the axiomatic mechanics of \textit{Méchanique Analitique}. In his second period, Lagrange constructed the calculus upon a firm algebraic basis, which did not appeal to infinitesimals or limits. He was also the first to introduce functions as the primary object of the differential calculus, and to replace differentials (i.e. infinitesimals) with derivative functions. Although the \textit{Méchanique Analitique} belongs to Lagrange’s “infinitesimal period”, he did not reject its findings. The reason for this was that for Lagrange although the methods using infinitesimals were problematic, they nevertheless yielded accurate results, and this accuracy could be proved by applying methods that did not rely on infinitesimals, i.e. his calculus of derivatives. More importantly, in the
preface to the second edition of the *Méchanique Analitique*, published after his algebraic turn, he points out that he used infinitesimals for *ease* of calculation, and that the results could be justified by algebraic means.

This justifies Helmut Pulte’s claim in [Pulte, 1998] that Lagrange was an instrumentalist. The claim is further strengthened by Lagrange’s own approach to mechanics. The *Méchanique Analitique* contains no geometric constructions, but also no philosophical reflections on the nature of space, time, mass, etc. These are not the basic concepts of his mechanics. Instead, he founds statics on a set of principles he considers self-evident and which he translates into mathematical equations, and then reduces dynamics to statics by proving d’Alembert’s principle. The fundamental principle of mechanics is the principle of virtual velocities (henceforth *pvv*) which Lagrange proves in a not too satisfactory way as argued in [Pulte, 1998].

In this chapter I also discuss the deduction of the Euler-Lagrange equation, both in Euler’s *Methodus Inveniendi* [Euler, 1743] and in Lagrange’s subsequent letters to Euler. This is a particularly important discussion for two reasons. First, because the Euler-Lagrange equation is a mathematical principle used to derive the equations of motion for any dynamical system and, second, because the Euler-Lagrange equations were the result of Euler and Lagrange’s efforts to justify the principle of least action (*pla*). This principle was the basis of Lagrange’s mechanical thought prior to the *Méchanique Analitique*, as expressed in his treatise, *Application de la méthode exposée dans le mémoire précédent à la solution de différentes problèmes de dynamique* (1760) [Lagrange, 1867a].
This chapter is an analysis of a textbook approach to what can be called the pendulum family of models. This is not intended as a presentation of all models within the family, but only as an exposition of the mathematical tools currently used in dynamic systems theory for some members of this hierarchy of models.

The taxonomy of pendulum models is constructed following two criteria. The first is structural and regards the components of the dynamical system. Thus, the root of the hierarchy is the simple pendulum composed of a bob attached to a massless rod which, in turn, is connected to a frictionless fulcrum. By connecting via another rod a second bob to the first, the resulting model will be of a double pendulum. If we take into consideration the mass of the rod, then the resulting model is that of the compound pendulum. The second criterion is given by the degree of abstraction of environmental factors. Thus, if, for any of the previous models, we consider the friction of the fulcrum, we obtain the damped version of each of the previous models, a total of three different models.

However, it should not be understood that this family of models includes replicas of actual pendulums, i.e. that take into account every single physical attribute of the real world.
system, such as colour, volume, etc. Each model defines what it deems as being relevant for the questions it wants to answer. Insofar as some properties of the real system do not influence the predicted outcome, they are ignored by the mathematical representation of the system. There is another constraint on mathematical representation, apart from the aforementioned criterion of relevance, dictated by mathematics itself. Just as one cannot use a pickaxe to bring down the Statue of liberty, the mathematical apparatus proves incapable of handling all possible scenarios with the same ease and elegance. More precisely, if $\theta$ is bigger than $\sin \theta$, then the resulting differential equations will be non-linear. In order to avoid complications, I will address only those cases in which the relation $\sin \theta \approx \theta$ holds.

2.1 The Simple Linearized Pendulum

The simplest model of a pendulum consists of a point of mass attached to a massless rigid rod which is itself tied to a frictionless fulcrum, moving without air resistance. The first step

![Figure 2.1: Illustration of the simple pendulum](image)

is to identify and label the components of the system. Thus, the mass is $m$, the length of the rod is $l$, and the angle between the rod and the normal is $\theta$. $\theta$ is the variable property of the system, whose equation of variation must be found. The other two components are constants.
The second step is to identify the forces acting on the elements of the system. These are the gravitational force \( G \) and the tension \( T \) in the rod acting on the bob. Using the parallelogram rule, we break up \( G \) into a component parallel with the string \((mg \cos \theta)\) and another component perpendicular to the string \((mg \sin \theta)\). Since the bob does not detach from the rod, nor does it move upwards, bending or compressing the rod, we can conclude that \( T = mg \cos \theta \). Therefore, the only force responsible for the bob’s movement is \(-mg \sin \theta\).

The minus sign indicates the fact that the force points in the negative direction of the \( x \)-axis when \( \theta > 0 \). Next we calculate the bob’s acceleration from Newton’s second law of mechanics, \( F = ma \).\(^1\) From \( mg \sin \theta = ma \) we find the acceleration \( a = g \sin \theta \). At the same time, the acceleration is \( a = \frac{d^2 s}{dt^2} \), where \( s \) is the arc starting from the lowest point of the bob’s trajectory to the point of angle \( \theta \). The relation between \( s \) and \( \theta \) is \( s = l \theta \), which means that \( a = \frac{d^2 s}{dt^2} = l \frac{d^2 \theta}{dt^2} = -g \sin \theta \). This gives the equation of motion:

\[
\frac{d^2 \theta}{dt^2} + \frac{g}{l} \sin \theta = 0
\]  

(2.1)

Since we are dealing only with small oscillations, i.e. \( \sin \theta \approx \theta \), we can linearize the previous equation thus

\[
\frac{d^2 \theta}{dt^2} + \frac{g}{l} \theta = 0
\]  

(2.2)

which has the solution

\[
\theta = \theta_0 \sin \omega t + \phi_0
\]  

(2.3)

where \( \omega = \sqrt{\frac{g}{l}} \) is the angular frequency, \( \theta_0 \) is the angular amplitude and \( \phi_0 \) is the initial angle of the pendulum. Since the fulcrum is frictionless and there is no friction with the air, the pendulum will swing indefinitely with the period \( T = 2\pi \sqrt{\frac{L}{g}} \). How the solution of

\(^1\)In the next chapters it will be shown that this is not Newton’s formulation, but an 18th century expression.
the equation of motion was obtained, i.e. how to find the solution of a second-order linear differential equation, will be dealt with in the next section. For now the focus is on the steps that must be followed in order to find the equations of the main parameters of the system.

2.1.1 The Damped Linearized Pendulum

How do the above equations change if we take into account air friction? For low velocities the drag force is directly proportional to the first power of the velocity. This suggestion seems to model realistically the behaviour of a damped pendulum with small angular amplitudes, i.e. we can approximate $\sin \theta$ with $\theta$. The formula of the drag force is

$$F_D = -2\gamma \frac{d\theta}{dt}$$

(2.4)

where $\gamma$ is a constant determined by experiment. Since we are still considering the low oscillations case, we obtain the equation of motion of the damped pendulum by adding $F_D$ to the linearized version of (2.1):

$$\frac{d^2\theta}{dt^2} + 2\gamma \frac{d\theta}{dt} + \frac{g}{l} \theta = 0$$

(2.5)

(2.5) is the result of the principle of composition of forces. Since the only force responsible for movement is the tangential component of $\vec{G}$, and the force of friction impedes motion, then the movement of the bob will be the result of the composition of these two forces.

Let us solve this equation and find the variation in time of $\theta$. The first step in solving a second-order linear differential equation is to substitute the function with $e^{rt}$ or, in our case, $e^{rt}$ since the derivatives are relative to the variable $t$. The choice of $e^{rt}$ as a provisional solution is not arbitrary. The general form of a second-order linear equation is $a \frac{d^2y}{dx^2} + b \frac{dy}{dx} + c = 0$,
where $a$, $b$ and $c$ are constants and $a \neq 0$ ($a$ could be 0, but this case does not concern us). The derivative of the function $y = e^{rx}$ is a constant multiple of the function itself, i.e. $\frac{dy}{dx} = re^{rx}$. If we replace $y$ with $e^{rx}$ we get

$$ar^2e^{rx} + bre^{rx} + c = 0 \quad (2.6)$$

$$e^{rx}(ar^2 + br + c) = 0 \quad (2.7)$$

Since $e^{rx}$ is never 0, it follows that $e^{rx}$ is a solution to our equation only if $r$ is a root of $ar^2 + br + c = 0$. This way, we reduced our initial second-order differential equation to a simple quadratic equation usually called the auxiliary or characteristic equation. The next step is to find the roots of this equation and determine the solutions of the initial differential equation. We know that $e^{rx}$ is a solution of the equation but since the characteristic equation has two roots, the solutions will be $e^{r_1x}$ and $e^{r_2x}$. The general solution of the equation is $y(x) = c_1y_1(x) + c_2y_2x$.

Returning to the equation of motion of the simple damped pendulum, a similar replacement as above gives us the characteristic equation with its two roots

$$r^2 + 2\gamma r + g/l = 0 \quad (2.8)$$

$$r_1 = \frac{1}{2}(-2\gamma + \sqrt{4\gamma^2 - 4g/l}) = -\gamma + \sqrt{\gamma^2 - 4\omega_0^2} \quad (2.9)$$

$$r_2 = -\gamma - \sqrt{\gamma^2 - 4\omega_0^2} \quad (2.10)$$

where $\omega_0 = \sqrt{g/l}$. Depending on the values of $\gamma$ and $r$, we can identify three distinct cases.

First, when $\gamma^2 - \omega_0^2 < 0$ or $\gamma^2 < \omega_0^2$, the pendulum is said to be underdamped, i.e. there is small friction with elements of the environment. The roots of the characteristic equation

\footnote{This is a mathematical theorem, but its proof is beyond our purposes.}
in this case are \( r_1 = -\gamma + i\sqrt{\omega_0^2 - \gamma^2} = -\gamma + i\omega \) and \( r_2 = -\gamma + i\omega \) where \( \omega^2 = \omega_0^2 - \gamma^2 \) for convenience.

So the general solution of the equation of motion in this case is

\[
\theta(t) = c_1 e^{(-\gamma + i\omega)t} + c_2 e^{(-\gamma - i\omega)t} \tag{2.11}
\]

where \( c_1 \) and \( c_2 \) are arbitrary.\(^3\) We can further simplify this equation using Euler’s formula:

\[
e^{ix} = \cos x + i \sin x:
\]

\[
\theta(t) = c_1 e^{-\gamma t} e^{i\omega t} + c_2 e^{-\gamma t} e^{-i\omega t} \tag{2.12}
\]

\[
\theta(t) = e^{-\gamma t} (c_1 e^{i\omega t} + c_2 e^{-i\omega t}) = e^{-\gamma t} (c_1 (\cos \omega t + i \sin \omega t) + c_2 (\cos \omega t - i \sin \omega t)) \tag{2.13}
\]

\[
\theta(t) = e^{-\gamma t} (c_1 + c_2) \cos \omega t + ie^{-\gamma t} (c_1 - c_2) \sin \omega t \tag{2.14}
\]

But we are interested in real solutions, mainly because it is hard to interpret a complex one. Since the general solution of the differential equation is a linear sum of linearly independent solutions it follows that \((c_1 - c_2) \sin(\omega t)\) and \((c_1 + c_2) \cos \omega t\) are also solutions of the second-order linear equation. Let us define \( C = c_1 - c_2 \), then a solution is \( \theta(t) = Ce^{-\gamma t} \cos \omega t + \phi \), where \( C \) is the initial amplitude and \( \phi \) is the initial phase shift.

If \( \gamma^2 = \omega_0^2 \) then the pendulum’s motion is critically damped and \( r_1 = r_2 = r = -\gamma \). In this case the mathematics dictates that the general solutions is

\[
\theta(t) = e^{-\gamma t} [A + Bt] \tag{2.15}
\]

\(^3\)The mathematical solution takes these constants to be arbitrary. The physical interpretation, however, is different.
where the two constants are, as usual, determined by the initial conditions. Finally, the third possibility is $\gamma^2 > \omega^2$ in which case pendulum has an overdamped motion described by

$$\theta(t) = e^{-\gamma t}[Ae^{\sqrt{\gamma^2 - \omega_0^2}t} + Be^{-\sqrt{\gamma^2 - \omega_0^2}t}]$$ (2.16)

### 2.2 The Compound Pendulum

The mathematical description of the compound pendulum includes the concept of moment of inertia ($I_p$) about the pivot point. The moment of inertia is generally viewed as the rotational equivalent of the mass. It is expressed in $kg \times m^2$ and depends both on the mass and the shape of the rotating object. For instance, a cylinder with mass $m$ and radius $r$ rotating around its axis has the moment of inertia on this axis $I_z = \frac{mr^2}{2}$. On the axes $x$ and $y$, in the plane perpendicular on the cylinder’s axis, the moment of inertia is $I_x = I_y = \frac{1}{12}m(3r^2 + h^2)$, where $h$ is the height of the cylinder. A simple circular hoop rotating around the z-axis perpendicular on its center has the moments of inertia: $I_z = mr^2$; $I_x = I_y = mr^2/2$. 

![Diagram of the compound pendulum](image)
In general, the moment of inertia of a given object rotating around an axis is the sum
of the moments of inertia of every particle of the object, i.e. \( I = \sum_{i=1}^{nm_i} r_i^2 \). The formula
\( m_i r_i \) entails that particles situated closer to the axis of rotation contribute less to the overall
result than those further away. This then captures the idea of mass distribution relative to a
certain axis. Obviously, this is not a workable formula; we cannot determine the masses and
the number of particles that make up an object. But we have a mathematical workaround.
If we assume that the object is divided into infinitely small volumes having elementary mass
\( dm \), then the moment of inertia is \( I = \int r^2 dm \). This allows us to compute the formulas in
the previous examples.

Having clarified the notion of moment of inertia, the equation of motion of the compound
pendulum is
\[
I \frac{d^2 \theta}{dt^2} + mgl \sin \theta = 0 \tag{2.17}
\]
which for small angular displacements becomes \( I \frac{d^2 \theta}{dt^2} + mgl \theta = 0 \). The period of the compound
pendulum is \( T = \frac{1}{2\pi} \sqrt{\frac{I}{mgl}} \).

### 2.3 The Double Pendulum

We start with the simple linearized pendulum and tie a second bob to the first one. There is
no friction in the fulcrums and no air friction. Because the system is complex and the motion
of each bob depends on the motion of the other, a Newtonian approach would be overly
difficult, since the values of the forces that acted upon each bob would be interdependent. If
the force generated by \( m_1 \) and acting on \( m_2 \) is \( F_{12} \) and the force generated by \( m_2 \) and acting
on \( m_1 \) is \( F_{21} \), then, when \( m_2 \) is at the point of maximum height \( F_{21} \) will have a determinate
magnitude \( x \). However, when \( m_2 \) begins to descend, the magnitude of the force \( F_{21} \) will
increase proportionally with the velocity of $m_2$. Moreover, if $F_{21}$ increases, then so will the velocity of $m_1$. In consequence, the magnitude of $F_{12}$ will also increase and further accelerate $m_2$. Therefore, in order to derive the equations of motion following the same steps as in the case of the simple pendulum, we would have to find the relation between $F_{21}$ and $F_{12}$. Instead, I will use the Lagrangian Equation. The coordinates of the two bobs are:

$$x_1 = l_1 \sin \theta_1$$  

$$y_1 = -l_1 \cos \theta_1$$  

$$x_2 = x_1 + l_2 \sin \theta_2$$  

$$y_2 = -y_1 + l_2 \cos \theta_2$$

Since the energy of the system is conserved, we can use the following formula

$$L = K - P$$

where $K$ is the total kinetic energy of the system, $P$ is the total potential energy and $L$ is the Lagrangian function of the system. Therefore, we must find the equations of the two
energies. The potential energies, as given by the potential equation, of the two bobs are:

\[ P_1 = m_1 g y_1 = m_1 g l_1 \cos \theta_1 \] (2.23)
\[ P_2 = m_2 g y_2 = m_2 g (l_2 \cos \theta_2 + l_1 \cos \theta_1) \] (2.24)

The general equation of kinetic energy is \( K = \frac{m v^2}{2} \). By rewriting \( v^2 = \dot{x}^2 + \dot{y}^2 \), where \( \dot{x} \) and \( \dot{y} \) are the derivatives of the two coordinates with respect to time, we find the equation of the velocity

\[ \dot{x}_1 = l_1^2 \dot{\theta}_1^2 \cos^2 \theta_1 \] (2.25)
\[ \dot{y}_1 = l_1^2 \dot{\theta}_1^2 \sin^2 \theta_1 \] (2.26)
\[ \dot{x}_2 = l_2^2 \dot{\theta}_2^2 \cos^2 \theta_2 + l_1^2 \dot{\theta}_1^2 \cos \theta_1 \cos \theta_2 \] (2.27)
\[ \dot{y}_2 = l_2^2 \dot{\theta}_2^2 \sin^2 \theta_2 + l_1^2 \dot{\theta}_1^2 \sin \theta_1 \sin \theta_2 \] (2.28)
\[ K_1 = \frac{m_1}{2} l_1^2 \dot{\theta}_1^2 \] (2.29)
\[ K_2 = \frac{m_2}{2} (l_2^2 \dot{\theta}_2^2 + l_1^2 \dot{\theta}_1^2 + 2 l_1 l_2 \dot{\theta}_2 \dot{\theta}_1 \cos (\theta_1 - \theta_2)) \] (2.30)

Finally, we can calculate the Lagrangian

\[ L = (K_1 + K_2) - (P_1 + P_2) \] (2.31)
\[ L = \frac{m_1}{2} l_1^2 \dot{\theta}_1^2 + \frac{m_2}{2} (l_2^2 \dot{\theta}_2^2 + l_1^2 \dot{\theta}_1^2 + 2 l_1 l_2 \dot{\theta}_2 \dot{\theta}_1 \cos (\theta_1 - \theta_2)) + m_1 g l_1 \cos \theta_1 + m_2 g l_2 \cos \theta_2 + m_2 g l_1 \cos \theta_1 \] (2.32)
This result is used in the Euler-Lagrange equation which, for the first bob, has the following form:

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}_1} \right) = \frac{\partial L}{\partial \theta_1} \tag{2.33}
\]

The component derivatives are

\[
\frac{\partial L}{\partial \theta_1} = m_1 l_1^2 \dot{\theta}_1 + m_2 l_1^2 \dot{\theta}_1 + m_2 l_1 l_2 \dot{\theta}_2 \cos (\theta_1 - \theta_2) \tag{2.34}
\]

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}_1} \right) = (m_1 + m_2) l_1^2 \ddot{\theta}_1 + m_2 l_1 l_2 \ddot{\theta}_2 \cos (\theta_1 - \theta_2) - m_2 l_1 l_2 \dot{\theta}_2 \sin (\theta_1 - \theta_2)(\dot{\theta}_1 - \dot{\theta}_2) \tag{2.35}
\]

\[
\frac{\partial L}{\partial \theta_1} = -l_1 g (m_1 + m_2) \sin \theta_1 - m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin (\theta_1 - \theta_2) \tag{2.36}
\]

Thus, (2.33) for the first bob becomes

\[
(m_1 + m_2) l_1^2 \ddot{\theta}_1 + m_2 l_1 l_2 \ddot{\theta}_2 \cos (\theta_1 - \theta_2) + m_2 l_1 l_2 \dot{\theta}_2^2 \sin (\theta_1 - \theta_2) + l_1 g (m_1 + m_2) \sin \theta_1 = 0 \tag{2.37}
\]

which, divided by \( l_1 \) is

\[
(m_1 + m_2) l_1 \ddot{\theta}_1 + m_2 l_2 \ddot{\theta}_2 \cos (\theta_1 - \theta_2) + m_2 l_2 \dot{\theta}_2^2 \sin (\theta_1 - \theta_2) + g (m_1 + m_2) \sin \theta_1 = 0 \tag{2.38}
\]

Through a similar process we get the equation of motion for the second bob:

\[
m_2 l_2 \ddot{\theta}_2 + m_2 l_1 \ddot{\theta}_1 \cos (\theta_1 - \theta_2) - m_2 l_1 \dot{\theta}_1^2 \sin (\theta_1 - \theta_2) + m_2 g \sin \theta_2 = 0 \tag{2.39}
\]
2.4 Breaking Down the Formulas

It can be observed from the previous section that there is a structural similarity between the algorithms used to deduce the equations of motion. Consider the case of the simple linearized pendulum. The first step was to describe the overall structure of the system by individuating its components and the relations between them. Thus, we have a simple bob connected to a fixed point by a massless rod. Here we also establish the relations of dependency between the components of the system. If we consider $l$ and $m$ variable, then the angle $\theta$ will be a function of time, of $l$ and $m$.

The next step is to identify the forces acting on the system. Obviously, the first force to be taken into account is gravity, $\vec{G}$. Since the bob’s trajectory is an arc, it follows that there is a centripetal force acting on the bob and oriented towards the fulcrum, $\vec{T}$, and a force with the same magnitude as $\vec{T}$ but pulling in the opposite direction. This force is the vertical component of $\vec{G}$, $\vec{G}_y$ and is obtained by applying the principle of decomposition of forces. Since this principle is derived from the principle of composition of forces (henceforth PCF) which is fundamental for Lagrangian statics, I will present the latter:

... a body that is set in uniform motion following two different direction at the same time, necessarily moves along the diagonal of the parallelogram whose sides it would travel separately in virtue of each force acting independently.\(^4\)

The PCF should not be confused with the principle of composition of motion. The latter principle was originally proved by Galileo, but the composition of forces is a 17th-century development. The proof of this principle, as given by d’Alembert in *Opuscules* rests on two principles: (i) two forces acting on an object are equivalent to a single force dividing the angle between the first two in two equal angles and (ii) multiplying the two forces with the

\(^4\)“...un corps qu’on fait mouvoir uniformément suivant deux directions différentes à la fois, parcourt nécessairement la diagonale du parallélogramme dont il eut parcouru séparément les côtés en vertu de chacun des deux mouvements.” [Lagrange, 1853, p. 10]
same factor will yield a proportional resultant force. These two principles are not part of Newtonian mechanics, but independently justified principles that use the theoretical concepts of Newton’s theory (i.e. forces)

Returning to the pendulum, we observe that the component of $\vec{G}$ whose direction is the tangent to the arc of the trajectory is the force which makes the bob move, $\|\vec{G}\| = mg \sin \theta$, and it is this force which allows us to find the acceleration of the bob, $a = g \sin \theta$.

The introduction of forces allows us to rephrase the constrains of the system. Thus, the requirement that the system be frictionless can be expressed now mathematically as $\|\vec{F}_f\| = 0$, where $\vec{F}_f$ is the force of friction acting on various components of the system.

The next step of the deductive process leading to the equation of motion is to use the definition of the acceleration as the second derivative of space relative to time $a = \frac{d^2s}{dt^2}$ and the geometrical relation between the angle of displacement, the length of the arc and the length of the rod, $s = l\theta$. By making the appropriate substitutions we obtain the equation of motion.

The solutions of the equation are obtained by applying purely mathematical methods which are reinterpreted.

In summary, the algorithm by means of which we obtain the equation of motion of a system follows these general steps:

1. Describe the structure of the system and individuate the elements of the structure
2. Construct the theoretical description of the system in this case, individuate the forces
3. Calculate the theoretical magnitudes (forces, acceleration)
4. Deduce the dynamic, i.e. the equation of motion.

The final result of the first step is a set of sentences in natural language which define an abstract system. I say “define” because even though the simple pendulum resembles an actual
physical system, it does so only in an idealized manner. After all, there are no frictionless pendulums and no pendulums with massless rods.

Rather than construing the description given in step (1) as an incomplete description of the structure of a physical system which ignores some features of the actual system, I suggest that this description is a definite description of a postulated system, an idealized entity embroidered within a story world. More precisely, the description in step (1) is not an idealization or an approximation of a real system, but the definition of a fiction which may resemble the real system.

The problem however, is that if the model is about an idealized system, then we have no reason to believe that it would apply to the real thing, in this case, an actual pendulum. I will return to this later on, but for now it suffices to point out this fundamental assumption, that the equations of motion describe the behavior of an idealized system.

The natural-language description of the system must individuate the relevant properties of the system, i.e. those properties that play a causal role in determining a measurable outcome. There are two kinds of properties: those that remain constant and those that vary in time. It is the variation in time of the latter that we are trying to capture in functional form, although the final mathematical expression will also depend on the constant properties. The variable properties, e.g. $\theta$ denoted by variables in the mathematical form, may depend on invariable physical properties of the system, e.g. $g$, or on properties that might vary, e.g. $l$, $m$. As we have seen in the previous sections, the solutions to the equations of motion depend on these quantities, although there we considered $l$ and $m$ constant. However, the generality of the mathematical representation of the behavior in time of $\theta$ is given by the fact that we can replace $l$ and $m$ by any value. In this way the equation applies to any pendulum.

Therefore, in this case mathematical functions represent more than just the variation in time of the angle of displacement. They also represent the relation between the “provision-
ally” constant quantities of the system \((l, m)\) and those quantities that change in time. It follows that before introducing functions as mathematical representations of relations of interdependency between quantities, we must first have the mathematical concepts of variable and constant.

As it will be shown in chapter 4, Euler’s *Introductio in analysin infinitorum* respects this order, introducing variables and constants before functions. More importantly, in the preface of *Institutiones calculi differentialis* he gives the intuitive basis for introducing functions:

In order that this difference between constant quantities and variables might be clearly illustrated, let us consider a shot fired from a cannon with a charge of gunpowder. This example seems to be especially appropriate to clarify this matter. There are many quantities involved here: First, there is the quantity of gunpowder; then, the angle of elevation of the cannon above the horizon; third, the distance travelled by the shot; and, fourth, the length of time the shot is in the air... Hence, if we always keep the same quantity of powder, the elevation of the barrel will vary continuously with the distance travelled and the shots duration of time in the air. In this case, the amount of powder, or the force of the explosion, will be the constant quantity... Those quantities that depend on others in this way, namely, those that undergo a change when others change, are called *functions* of these quantities.

This definition applies rather widely and includes all ways in which one quantity can be determined by others. Hence, if \(x\) designates the variable quantity, all other quantities that in any way depend on \(x\) or are determined by it are called its functions. [Euler, 2000, Preface]

This passage resembles the first step of the present algorithm, in that it individuates those quantities that vary and those that remain constant and, more importantly, determines the relations of dependency between them. I will call these relations *functional relations* to distinguish them from *functional expressions*. A functional relation is any perceived dependency between two physical quantities. On the other hand, a functional expression is any means of representing the functional relation. In Euler’s case, the functional expression is an analytic formula, but a functional relation can be represented through other means: graphs, tables, geometric constructions, etc.

An important thing to observe is that functional relations obtain only between quantities.
This is true both for Euler’s definition of function and the modern one. In Euler’s example, the variable quantity is the amount of gunpowder, while the “quantity of space”, i.e. distance, through which the cannonball passes and the “quantity of time” it spends in the air are functions of the amount of powder.

Applying Euler’s reasoning to the case of the simple pendulum, we consider $l$ and $m$ to be constant and the angle $\theta$ a function of these two. But $l$, $m$ and $\theta$ are quantifiable and in virtue of this they can be represented as variables. As it will be shown in chapter 4, Euler understood variables differently than we do today, and these differences are reflected in his concept of function as well.

But functions as representations of a relation of dependency between quantities are static representations in that they express the value of the dependent quantity for particular values of the variables. Derivatives, on the other hand, are dynamical representations of these dependencies. In more concise terms, the derivative represents the rate at which the primitive function changes relative to the variation of one of its variables. The expression “change in time of $\theta$” becomes, in mathematical terms, $\frac{d\theta}{dt}$ which also represents “change at a time” or the “instantaneous angular velocity”.

Variations could also be represented as ratios of differences, without introducing the concept of derivatives, e.g. $\frac{f(x_2)-f(x_1)}{x_2-x_1}$. The accuracy of this representation is given by the value of the difference $x_2-x_1$. For instance, consider the function $\sin x$ and $x_2 = \pi$, $x_1 = -\pi$. The variation of the function is $\sin \pi - \sin -\pi = 0 - 0 = 0$. In consequence, the ratio $\frac{\sin x_2-\sin x_1}{x_2-x_1}$ will also be 0. This would lead to the wrong conclusion that on the interval $[-\pi, \pi]$ the function $\sin x$ does not vary at all. A more accurate representation would be obtained by making $x_2$ approach $x_1$ or vice versa. An exact representation would be obtained only if the length of the interval $[x_1, x_2]$ is considered infinitely small. Of course, this assumption is problematic, since a perfectly accurate representation could be obtained only if the length of the interval
is 0. However, if this is the case, then there would be no variation of the variable at all. This problem and others pertaining to infinitely small quantities will be treated in greater detail in chapters 3, 4 and 5.

It might be objected that derivatives are functions themselves and so it follows that functions can be used to represent both functional relations and their variations. This is true, however, in the previous paragraph I chose to ignore this aspect in order to distinguish between two represented objects: a relation between quantities and the relative variation of two interdependent quantities. Moreover, from a historical perspective the introduction of derivatives as functions is attributed to the last years of Lagrange’s career. Throughout the 18th century, relative variations of functions were represented as ratios of differentials or, which amounts to the same thing, of infinitely small quantities. As it will be argued in chapter 3, reliance on these quantities was problematic, especially in the absence of a formal definition of limits. This problem was solved by Cauchy early in the 19th century, an event which also established the firm foundations of calculus.

Another reason for presenting derivatives as ratios of infinitesimals rather than functions is that the concept of infinitely small quantity was used in the deduction of the equation of motion of the compound pendulum in 2.2. There we assumed that the oscillating object was divided into infinitely small volumes having an infinitely small mass $dm$. The motivation for introducing this concept was given by the impossibility of calculating the total moment of inertia of the oscillating object, i.e. $I = \sum_{i=1}^{n} m_i r_i^2$. This finitesimal approach is the reverse of the way in which I introduced ratios of infinitesimals in the previous paragraph. Now, instead of subtracting a value of a function from another, we compute the sum of all values possible for $m$, the mass of a particle, in the discrete set \{$m_1$, $m_2$, \ldots, $m_n$\}. Obviously, we cannot find out the value of this sum, since we do not know the values of $m_1$, $m_2$, etc. Instead, we assume that the body is divided into infinitely small volumes and apply the rules
of the integral calculus. This method has its roots in Euler’s *Institutiones calculi differentialis*, where he considered the differential and integral calculus special cases of the calculus of finite differences and sums, respectively. To the modern reader, the justification of this method is given by the theory of limits. However, at the time Euler wrote the *Institutiones*, the notion of limit had not been mathematized and was still intuitive. One question that will be addressed in chapter 4 is whether he was justified in assuming that the differential calculus is a special case of the calculus of finite differences.

Returning to the algorithm for the deduction of the equations of motion it can be observed that, in step (2) we introduced theoretical terms in the description of the described structure. In the case of the simple pendulum this implies individuating the forces acting on the components of the system and calculating their magnitudes. This step describes in theoretical terms the abstract system resulted after (1), by individuating the forces and their magnitudes. Using this description and the laws of the overarching theory we deduce the function representing the dependency on time of the angle of displacement. In formulating the laws of mechanics there are two fundamental mathematical concepts: function and derivative. It is because we define velocity mathematically as infinitesimal variations of space in infinitesimal variations of time that we can further define the acceleration as variation of velocity in time.

The concept of function and further on that of derivative and partial derivative are central to how we represent the (idealized) system.

However, the laws of mechanics and the theoretical description are usually insufficient. We also need boundary conditions that represent the physical constraints of the system. In the case of the simple pendulum the boundary condition is $T = mg \sin \theta$. Here the boundary condition does not have any particular influence on the solution, but in different scenarios, the boundary conditions are indispensable to the elimination of variables from the system of equations that results after applying the theoretical laws and principles. The boundary
conditions are essentially mathematical equations which fulfil a double role. On one hand, they are the mathematical expression of the physical limits of the model. This is the case with \( T = mg \sin \theta \) which, as previously mentioned, expresses the physical requirement that the bob’s trajectory is an arc.

On the other hand, boundary conditions are mathematical constraints on the functions used to represent the behavior of the system. For instance, in the simple case of a quantum potential well, the boundary conditions are that the first derivative of the wave function and the wave function itself are smooth and finite at the boundary points, i.e. the coordinates of the potential well. This is due to the dependency of Schrödinger’s equation on the second derivative of the wave function which must be finite at all points in order to be intelligible. This is possible only if the first derivative of the wave function and the wave function are continuous and finite at all points. Thus, in this case the boundary conditions are mathematical in nature. The boundary conditions are a distinct part of the theoretical description of the system, although they might not be explicitly stated. In general, the mathematical boundary conditions, such as continuity of some functions, are tacitly assumed and invoked when we have to choose between several solutions or when we have to eliminate some unknown variables. The final two steps are less interesting conceptually, as they are just applications of mathematical methods to the equations represented in the previous steps.

The equations of motion of the double pendulum are deduced following the same algorithm. We begin by laying down the structural description of the system and labelling its components. The constant properties of the system are \( l_1, l_2, m_1 \) and \( m_2 \). The variable properties whose evolution in time must be represented mathematically are \( \theta_1 \) and \( \theta_2 \).

The theoretical description constructed in the second step makes use of the concepts of kinetic and potential energy. In the case of the simple pendulum this description introduced the concept of force which allowed us to apply Newton’s laws. The description given is
summed up either by the equation \( L = K - P \) or by \( H = K + P \) where the values of \( K \) and \( P \) are calculated in the third step of the algorithm by summing the individual energies of each bob, i.e. \( K = K_1 + K_2, \ P = P_1 + P_2 \).

Deducing the equations of motion of the double pendulum is not as straightforward as deducing the equations for the simple pendulum. In addition to the usual mathematical methods, the Euler-Lagrange equation was used to find the analytical expression of the variation of the two angles \( \theta_1 \) and \( \theta_2 \). The physical significance of the Euler-Lagrange equation will be analyzed in chapter 5. For now, it suffices to say that unlike the principle of the composition/decomposition of forces which is introduced by the physical theory, the Euler-Lagrange equation(EL) is mathematical in nature. Both the principle of composition and the EL equation belong to what I will call bridge principles which connect the three components of a model: the idealized system, the theoretical description and the mathematical representation.

### 2.5 The Structure of a Scientific Model

The idealized system postulated by the model is described by a set of sentences expressed in natural language. The mathematical layer connects with the theoretical description via mathematical bridge principles. In turn, the theoretical description connects with the idealized system via a second kind of bridge principles, which depend on the theory within which we are working. By bridge principle I understand any assertion, whether justified or simply postulated, which establishes a correspondence between the elements of one layer and those of another, e.g. between functions and theoretical concepts, or between concepts and elements of the idealized system.

The mathematical component is the set of equations and their solutions which represent
the evolution in time of the relevant properties of the idealized system. The relation of representation, however, depends on the other two layers of the model. In the absence of the two other layers the variables and constants of the equations of motion would have no meaning; they would be nothing more than mathematical expressions. To understand the importance of layers (2) and (3) we must take a closer look at the role of mathematics in general in the proposed account of models.

It can easily be observed from the deductions of the equations of motion presented in this chapter, that mathematics plays a part both in describing the system from a theoretical point of view at the second layer, and in representing the relevant properties of the system. However, the equations of motion and their solutions are obtained through purely mathematical methods and this does not explain why they represent the system. To find an answer, we should turn to the third step of the algorithm presented in the previous section, i.e. calculating the magnitudes of the forces, because this is the point in which the mathematization of the represented phenomenon occurs.

Here, theoretical entities which previously have been labelled, e.g. the force of friction by $\|\vec{F}_f\|$, are represented as mathematical functions, e.g. $f(x, y) = xy$, where $x$ is interpreted as the mass of an object and $y$ is the body’s acceleration. As shown in the previous section, the fundamental concepts of mathematical representation, as identified in the pendulum example, are, in logical order, variables, constants, functions, infinitesimals and ratios of infinitesimals. Therefore, to understand how the equations at the mathematical layer function, we must see in greater detail the relation between these concepts and the entities they represent.

To illustrate this relation I will draw upon an analogy with art\textsuperscript{5} and consider these mathematical concepts equivalent to the oils and brushes used to create an artistic representation.

\textsuperscript{5}Whether such an analogy is permitted or not is subject to debate. However, I will follow Bas van Fraassen [Van Fraassen, 2008] which argues that this analogy is permissible and informative.
These concepts are the building blocks of mathematical representations, just as the oils and brushes are prerequisites for constructing an artistic representation. Of course, it should not be understood that the two kinds of representation are created following the same constraints, i.e. that a physicist simply arranges symbols on a sheet of paper just as a painter brushes a canvas.\(^6\)

First and foremost, the symbols must be organized following the syntactical rules of mathematics. Second, the symbols represent particular instantiations of the general concepts.\(^7\) For instance, \(dx\) is the infinitely small increment of a particular variable \(x\) and \(\sin x\) does not represent the general concept of function, but only an instantiation of it. Symbols represent these instances of mathematical concepts which, in combination, represent empirical phenomena. The representational relation between symbols and instances of concepts is established through an act of fiat, or an “initial baptism” to use Kripke’s terminology. As such, whether we use roman or Greek letters to represent a function is simply a matter of convention. Likewise, whether variables are represented by letters from \(x\) to the end of the alphabet or by letters from \(a\) on, is again a matter of convention, provided of course, that we can distinguish between variables, constants and functions.

The instances of the general concepts, not their symbols, are analogous to the brush strokes on canvas. Just as the combinations of oils of different colors make up the features and shades of an idealized representation of Socrates’ face in Jacques-Louis David’s Death of Socrates, so do the combinations of the particular instances of mathematical concepts\(^8\)

\(^6\)There are numerous constraints on the process of constructing a scientific representation. For more details see [Van Fraassen, 2008],[Morgan and Morrison, 1999],[Bueno, 2000]. In the following chapters I will focus only on those imposed by the degree of mathematical advancement.

\(^7\)The relation between the general concept and such an instantiation can be regarded as the same relation between genus and species.

\(^8\)I take this expression to refer especially to the aforementioned concepts, i.e. variable, constant, function, ratios of infinitesimals and infinitesimals themselves, although any other notion is acceptable, e.g. matrices, groups, sets, etc.
combine to construct a part of the theoretical description.

This analogy can be further expanded to include the second layer of the model, i.e. the theoretical description, by considering the theoretical entities and the relations between them analogous to parts of the painting. For instance, the sum of brushstrokes that ultimately represent Socrates in David’s painting are, in this example, analogous to \( T = mg \sin \theta \), which is a part of the theoretical description of the system. Ultimately, the combination of all individual parts of the description is analogous to the painting and just as the painting represents a particular event, so does the theoretical description as a whole represent a phenomenon. Moreover, David’s painting represents an idealized version of the actual event, not the actual event, just as the first two layers of a model represent an idealized phenomenon.

However, if we say that this system is *postulated* and that its description is connected via bridge principles to its mathematical representation, then we must still explain how this idealized system connects with its real counterpart. In other words, what makes the model a representation of *that* particular physical system? My suggestion is that the relation between the idealized and the real world system is that of *iterative correlation*.\(^9\)

Iterative correlation is the gradual adjustment of the model’s components to “fit the empirical data” prompted by inconsistencies between the predictions made by the model and the measured quantities of the real world system. Given the previously suggested structure of the model, there are two possible sources of inconsistency. First, it might be the case that the postulated system does not resemble the relevant aspects of the physical system. The criteria of relevance are determined on a case-by-case basis, but there is one characteristic which is essential to all criteria: the structural description of the postulated system must express the dependent quantities as variables and the constant quantities as constants. In other

\(^9\)This is just a brief sketch of the concept of iterative correlation, as the focus of the present chapter is to clarify the relations between model-layers.
words, the variable properties of the idealized system must match the variable properties of
the real world counterpart, and the same goes for the constant ones. For instance, if the real
system is a double pendulum, but the model considers that one of the masses is fixed, then
the resulting model will not give accurate predictions, for it will not capture the structure of
the real system.

Second, the postulated system may correctly identify the constant and the variable prop-
nerties of the real world system and the dependencies between the latter, but the mathematical
layer may not express these dependencies accurately. For instance, the relation between the
angle of displacement and the length of the rod of a simple pendulum is \( \frac{d^2\theta}{dt^2} + \frac{g}{l}\theta = 0 \). Now, if
instead of this equation the model suggests that this relation is expressed by \( \frac{d^2\theta}{dt^2} + \left(\frac{g}{l}\right)^{-1}\theta = 0 \),
then the predictions made by the model will not match the measured values. Thus, even
though the idealized system postulated by the model is isomorphic with the physical system,
the mathematical representation of the relations between the different magnitudes does not
correspond with reality. In this case, the mathematical representation is modified so that the
predicted values match the measured ones. In this sense, iterative correlation is less a definite
relation between a model and the real world and more an intentional process of “adjusting”
the model to reality.

Although a great deal more could and should be said about the notion of correlation, an
in-depth analysis of it is not the purpose of the present work. Instead, the focus will be to
clarify the relations between the three components of the model, in particular, the notion of
bridge principles mediating between the mathematical layer and the postulated system.

2.6 Conclusions

To recapitulate, the algorithms used to deduce the equations of motion of the simple and
double pendulum suggest the following structure of what we call a model:
1. Underlying postulated idealized system (e.g. \(m_1, m_2, l_1, l_2, \theta_1, \theta_2\))

2. Theoretical description of the system

3. Mathematical formalisms

Therefore, the answer to the ontological question “What is a scientific model?” is a triplet composed of an idealized physical system, a theoretical description and a set of mathematical equations representing, mathematically, the behavior of this system. The three components are connected by bridge principles, which I have loosely defined as any assertion, whether justified or simply postulated, which establishes a correspondence between the elements of one layer and those of another. To understand the significance of each layer, we must take a closer look at the concepts specific to each layer and at the ways in which they are connected. This will be the task of the following chapters/
Chapter 3

The Status of Science in the Late 17th, Early 18th Centuries

This chapter is a brief overview of the status of science before Euler and Lagrange. The goal is to illustrate the difficulties the two had to face and, in doing this, to present a basis of comparison for their achievements. Thus, in the first section of this chapter I will try to distinguish between the various types of mathematical and mechanical inquiries developed in the 18th century. The second section will focus on the development of mechanics in connection with the development of calculus in mathematics and the separation of the latter from geometry.

3.1 18th-century Physics

The scientific landscape of the 18th century bears little resemblance with the clearly delineated domains of present-day scientific inquiry. During this period, physics was not an experimental science and its domain bore little resemblance with the one attributed to it today. Eighteenth century physics was a part of natural philosophy concerned with the gen-
eral principles of the nature of matter and change. As such, its domain included topics from chemistry and biology to the human sciences.

Moreover, physics was not a mathematized discipline. Mathematics played a more important role in optics, astronomy, harmonics, and any other discipline which lay between mathematics and physics. During the 16th century, these intermediary sciences, together with mechanics, progressed more than physics itself. Physics became more oriented towards the empirical with the development of academic institutions throughout Europe.

Two such institutions were the Académie Royale des Sciences in Paris and the Royal Academy in London, both founded in the 1660s. The aim of these new institutions and of those similar to them created throughout Europe was to advance knowledge, unlike universities, whose goal was to pass on accumulated knowledge. [Home, 2008] distinguishes three groups whose main concern was physics, in 18th-century France. First, there were university professors whose main concern was the individuation of first causes and the certainty of this knowledge. Their inquiries were made from an Aristotelian standpoint following a non-mathematical and verbal expository method. The second group comprised les physicien of the Académie Royale des Sciences. Their aim was to expand knowledge through empirical means backed by a mechanistic theory. Finally, there were the Cartesians, who were allowed to join the Académie in 1699, after being denied access for almost forty years for their “dogmatic” views. Their method was also experimental, and their theory mechanistic; like the members of the first group, their aim was certain knowledge of final causes.

Around the same time, in the 1660s, a tradition of public speaking on experimental demonstrations was established by Jacques Rohault, and a similar course was introduced into the University of Paris in the 1690s, headed by Pierre Poliniére. The lectures of the course were published in 1709 and represented “the first major infiltration of an experimental outlook into the French educational system” [Home, 2008, p. 356].
Things were somewhat different across the channel. The meetings of the Royal Society of London were opportunities to demonstrate experiments that previously had been conducted privately and to discuss their consequences. Courses focusing on experiments were introduced at Oxford in 1704 by John Keill and at Cambridge in the following years. James Hodgson and Francis Hauksbee also presented courses on experimental demonstrations, which were a way of advertising Hauksbee’s instruments. By the end of the 18th century, physics had become a largely experimental science, whose main subjects were the constitution of matter, heat theory, electrostatics and optics [Grattan-Guinness, 1990].

What is the place of mechanics in this heterogeneous mix of scientific domains? Before answering this question, we must distinguish for sake of clarity between mechanics, Newtonian mechanics and physics. Following [Grattan-Guinness, 1990] I will consider mechanics to refer to “the study of the rest and motion of bodies (which themselves are construed widely to include point-masses, extended solid and fluid objects, instruments, machines, frameworks, and constructs) under the action of normal physical forces (and so excluding, for example, heat, electrical, and magnetic actions)” [Grattan-Guinness, 1990, p. 315]. The main topics of mechanics thus defined are equilibrium and motion, questions regarding inertia, force, energy, momentum, collisions, vibration, rotation and oscillation. Newtonian mechanics was a branch of mechanics, a tradition in which all investigations have as a basis the gravitation law and Newton’s three laws of motion. Mechanics in general and, implicitly, Newtonian mechanics were mathematized. This feature distinguishes these two domains from physics, which did not benefit from the contemporary development of mathematical tools. Although it addressed topics on the constitution of matter, just as mechanics did, physics had a non-mathematical approach which set it apart from mechanics.

Within mechanics there are two other traditions: variational and energy mechanics. I will not go into details regarding the history of Newtonian mechanics, as the latter two traditions
are more relevant for the subject at hand. Moreover, the interaction with these two traditions gave Newtonian mechanics the form it has today, one example being the modern-day expression of Newton’s second law as $F = ma$, a form popularized by Euler in the eighteenth century. A second reason for focusing on variational and energy mechanics is that the scientists working within these traditions contributed greatly to the advancement of mathematical techniques and to their applications to mechanics. These techniques were presented as principles and were attributed different roles within mechanics by their creators. The most important of these principles, all of which will be discussed further on in more detail, were d’Alembert’s principle (PDA), which reduced dynamics to statics, the principle of virtual velocities (PVV), the principle of virtual work (PVW), the principle of least action (PLA), the principle of momentum, the principle of the lever (PL), etc.

An example of the heterogeneity of the 18th-century scientific environment is the introduction of the concept of energy. The development of the concept of energy was mainly the result of pressures from the domain of engineering. Although we encounter in the works of Euler and Lagrange, for instance, formulas that modern-day physics uses to express kinetic ($1/2mv^2$) and potential energy (e.g. $mgh$), in the eighteenth century these expressions had a different meaning altogether. For instance, Lagrange took them to be purely instrumental devices, useful for calculations. In fact, the concept of energy emerged only early in the 19th century, while the distinction between kinetic and potential energy was introduced even later by Thomson and Tait. For an 18th-century scientist the familiar formula $1/2mv^2$ would mean, at most, half of the *vis viva*, or *force vive*. These mathematical expressions got their meaning from the domain of engineering. In *Du Calcul de l’Effet des Machines* (1829), Gaspard Coriolis introduced the concept of work as the main magnitude of the theory of moving engines, and he changed the expression of living force from $mv^2$ to the more familiar form of $1/2mv^2$. His definition of work as $\int Pds$, where $P$ is the force acting in the direction of the
displacement $ds$, had been around, in one form or another, in the works of Louis Navier, John
Smeaton, Jean Charles Borda and others which had been using similar formulas to calculate
the productivity of water wheels and steam engines, usually in terms of the amount of coal
required to raise a known weight to a fixed height. For instance, for Jean Victor Poncelet,
in “Course de mécanique appliqué aux machines” (1829), force vive is work when the only
resistance is the body’s inertia[Grattan-Guinness, 1990].

The disputes between the three traditions of mechanics (Newtonian, variational and
energy-based) arose mainly on the matter of the logical priority of the principles one tra-
dition held to be foundational for mechanics. The competition did not regard the validity
of the principles themselves, as all traditions supported the same principles, although they
justified them differently. Instead, the main question was which principles can be considered
sufficient to deduce the principles of the opposing traditions. Of course, there were some re-
searchers that can be considered “mavericks” that cannot be placed within a tradition. One
example is Euler who supported P.L.A Maupertuis’ principle of least action as being quite
general, but who subscribed to other traditions when investigating problems in subdomains
of mechanics such as hydrodynamics, machines, etc.

3.2 The Status of Mathematics Before Euler

The status of 18th-century calculus is quite accurately described by Giovanni Ferraro as “a
nonrigorous corpus of manipulative techniques, which succeeded in anticipating certain mod-
ern results thanks to a series of lucky circumstances and fortuitous cases.”[Ferraro, 2000,
p. 107] This series of fortunate events lead to the development of new branches of mathemat-
ics, such as differential equations, infinite series, differential geometry, calculus of variations,
etc., to the introduction of new types of functions (hyperbolic, trigonometric, functions of
several variables) and, through Euler and Lagrange, to the separation of the calculus from geometry.

During this period, more important than this development of the calculus, is the growing interest in its foundations, beginning with its creators, Newton and Leibniz, in the late 17th century and ending with Lagrange’s *Leçons*. Although historians of mathematics [Boyer, 1949], Kline [1990], Ferraro [2004] seem to agree that the interest in the foundation of calculus in the 18th century was deliberate, on a closer inspection, this view seems misguided. In [Grabiner, 2005] the reason is that Lagrange is the first mathematician of his time to make foundational issues subjects of research in their own right. However, following Grabiner, Lagrange’s predecessors touched this subject for presentation purposes only, not because they viewed the topic of foundations of the calculus as particularly important:

Solving problems was important, not proofs about the concepts used in solving them. The attitude of most of these mathematicians, implicit in their usual choice of problems and methods, may be summed up in a remark attributed to d’Alembert: “Go on, go on; the faith will come to you.” [Grabiner, 2005, p. 23]

There are two reasons [Grabiner, 2005] why foundational issues were addressed in the 18th century prior to Lagrange. The first is that the calculus was a new domain of mathematics so most works on the subject had to begin with a presentation of the basic concepts. Second, these works were also written with pedagogical intentions: “...these expositions were intended as textbooks for a growing readership.” [Grabiner, 2005, p. 23] The audience included both members of the scientific community and the general public. The increasing number of scientific journals and societies were instrumental to organizing and expanding the scientific community, while the success of Newtonian physics lead to a growing public interest in science “prompting both mathematicians and philosophers to explain the calculus to laymen” [Grabiner, 2005, p. 24]. It can be said, therefore, that Lagrange’s predecessors
touched upon foundational issues only to make their exposition of the subject matter, be it calculus or mechanics, clearer. However, Lagrange recognized the importance of discussing the basic concepts, rather than simply defining them. This is why in 1784, at his suggestion, the Berlin Academy offered a prize for an adequate solution to the problem of foundations of the calculus. The competition was ultimately won by Simon L’Huilier, but the Academy was not satisfied. None of the participants justified the results of the calculus. More important than simply recognizing the importance of foundational research is that Lagrange “derived his existing major results from his foundation”[Grabiner, 2005, p. 37].

Nevertheless, even if 18th-century mathematicians touched upon foundational issues only for presentation purposes, an understanding of the concepts they considered fundamental for the calculus is important for two reasons:

1. The calculus loses its geometrical interpretation, and as a result, it is possible to apply analytical tools to other kinds of problems.

2. The explicit definition of fundamental mathematical concepts and methods made possible their correlation with their mechanical counterparts.

To understand why these points are relevant, it must be understood that 17th-century mathematical analysis was a “corpus of analytical tools (algebraic equations and operations, later the differential and the rules of the calculus) for the study of geometric objects, namely curved lines.”[Bos, 1974] The geometrical quantities of mathematical interest, the conceptual precursors of the concept of variable, were defined relative to points on a curve. The relations between these geometrical variables were expressed by equations, after Descartes’ introduction of algebraic methods to geometry, or by descriptions in prose of the geometrical method of constructing the curve, in those cases in which the introduction of an equation was not possible. These descriptions were complemented by graphical methods which “played a
role in the early calculus that would later be filled by the function concept” [Fraser, 2008, p. 309]. The tangent (τ) of a curve at a point equals the length of the subtangent (σ) at that point. The area under the curve was given by an integral. Calculating the curvature at a point required the determination of the radius of the curvature. The first mathematician that emphasized the importance of decoupling analysis from geometry was Euler, although the extent to which he achieved this project is debatable [Kline, 1990; Bell [1992], Boyer [1949], Fraser [2008]. Thus, in Methodus Inveniendi Lineas Curvas, he notes that it is possible to give a purely analytical solution of the problem of maxima and minima, i.e. finding in a class of curves the curve which maximizes or minimizes a geometrical property. This implied that instead of looking for a curve, the focus shifted towards finding from a class of equations an equation of $x$ and $y$ such that the quantity of interest is an extremum. An in-depth analysis of Euler’s contributions to the calculus will be given later. For now it suffices to point out that the scientific progress of the 18th century depended on this departure from geometry, which was, essentially, a process of abstraction enabling the application of mathematical tools to a wider domain of problems.

This brings us to the second point, i.e. the relation between the fundamental mathematical concepts and mechanical quantities. Obviously, the nature of this relation depends on what are considered to be the fundamental concepts of the calculus and of mechanics. As it
will be shown, this was no easy matter to settle, since during the period under investigation there were almost as many “foundations of the calculus” as there were mathematicians. To illustrate this point, as well as to introduce some of the key terms that will be used in the later chapters, I will give a brief overview of the 17th-century predecessors of Euler and Lagrange.

3.2.1 Newton

The two relevant features of 18th-century calculus, i.e. the separation from geometry and correlation of mathematical concepts with mechanical ones, are hard to separate. The main reason is that the justification of the foundational concepts and methods of the calculus relied on 17th-century geometrical intuitions for their justification. This tendency is present, to a certain extent, in Euler’s work as well, and is definitively removed by Lagrange. The relation between geometry and the foundations of the calculus is clearer in the works of Newton and Leibniz on the fundamental principles of the calculus.

The dispute between Newtonians and Leibnizians over the precedence in the invention of the calculus is not relevant to our present concerns. More important are their different sets of fundamental concepts and the justification given for the methods of the calculus. Therefore, the order in which I will present these should not be interpreted as siding with a particular party in the dispute.

Newton’s View on Mechanics and Geometry

There are three theories that play a role in Newton’s foundations of the calculus:

1. infinitesimal quantities
2. fluxions
3. prime and ultimate ratios.

The nature of the relation between the three is disputable. In [Boyer, 1949] they are viewed as different interpretations of the calculus occurring in Newton’s works and presented as equivalent in the *Principia*. In [Kitcher, 2007] they are sets of concepts that played a determinate role in Newton’s foundations of the calculus:

..the theory of fluxions yielded the heuristic methods of the calculus. Those methods were to be justified rigorously by the theory of ultimate ratios. The theory of infinitesimals was to abbreviate the rigorous proof, and Newton thought that he had shown the abbreviation to be permissible. Rather than competing for the same position, the three theories were designed for quite distinct tasks.[Kitcher, 2007, p. 34]

[Suisky, 2009] considers that Newton views infinitesimal quantities as the geometrical foundation of the calculus, while fluxions are mechanical representations of the process of generating curves. In spite of these differences, all three views have in common the recognition of the geometrical roots of Newtonian calculus.

In the preface of the *Principia* Newton notes that the term “mechanics” is an abbreviation of the term “practical mechanics” which denotes the “practice of artificers” and it lacks precision, due to the imperfection of the artificers themselves. Therefore, the difference between mechanics and geometry is given by the complete accuracy of the latter and the imprecision of the former. However, the root of the distinction is not inherent in the “art of mechanics” since the lack of accuracy is introduced by the one practicing the art. In particular, *rational* mechanics is a perfectly rigorous abstraction of practical mechanics, which studies the mathematical laws governing natural phenomena.

More importantly, for Newton geometry is *founded upon* (rational) mechanics:

..for the description of right lines and circles, upon which geometry is founded, belongs to mechanics. Geometry does not teach us to draw these lines, but requires them to be drawn. For it requires that the learner should first be taught to describe these accurately, before he enters upon geometry; then it shows how by these operations problems may be solved. To describe right lines and circles are problems, but
not geometrical problems. The solution of these problems is required from mechanics, and by geometry
the use of them, when so solved, is shown. ... Therefore geometry is founded in mechanical practice, and
is nothing but that part of universal Mechanics which accurately proposes and demonstrates the art of
measuring. [Newton et al., 1729][Preface]

The objects of geometry are, therefore, the curves and figures constructed by mechanics. These figures are postulated entities which represent the starting point of geometrical calculations. Geometry is about curves, tangents, etc., but the description of these figures is the task of mechanics. This means that in order to solve geometrical problems, e.g. maxima and minima, tangents to a curve, etc., one must know how the figures had been kinematically generated. Thus, rational mechanics is just as precise as geometry, but it precedes geometry, as it is a pre-condition for the existence of geometry.

Newton clearly expresses his intent to provide a general theory of the kinematic generation of curves, based on any natural force:

But since the manual arts are chiefly employed in the moving of bodies, it happens that geometry is commonly referred to their magnitude, and mechanics to their motion. In this sense rational mechanics will be the science of motions resulting from any forces whatsoever, and of the forces required to produce any motions, accurately proposed and demonstrated.[Newton et al., 1729][Preface]

Geometry can be applied to natural phenomena precisely because it has its roots in mechanical practice. Since mechanics depends on the notion of artificer, e.g. technician, God, etc., which introduces a certain amount of inaccuracy, the difference between practical and rational mechanics can also be expressed in terms of perfect and imperfect artificers. Rational mechanics is the science of motion of a perfect artificer, therefore it can provide exact descriptions of curves, thus being able to provide the exactness required for geometrical proofs. The question we must answer now is how does mechanics describe the curves in an exact way?
The Theory of Fluxions

For Newton, the theory of fluxions would be the right answer to the previous question. Newton’s theory of fluxions reflects the kinematic generation of curves and in his *De Methodis Serierum et Fluxionum* he makes clear that his method of fluxions is applied to geometrical quantities generated by motion: lines generated by the motion of points, surfaces generated by the motion of lines, etc. The basic concepts of the theory of fluxions are *fluents*, *fluxions* and *moments*:

1. Fluents represent the quantity of the flow and are labelled $x$, $y$, $z$, etc.

2. Fluxions represent the instantaneous velocity of the flow and are labelled $\dot{x}$, $\dot{y}$, $\dot{z}$, etc.

3. Moments of fluents are the infinitely small quantities by which the fluents are incremented in an infinitely small time interval and are labelled $o\dot{x}$, $o\dot{y}$, $o\dot{z}$, etc. A fluent is *incremented* by its moment.

The dependence on the time element $o$ ensures the continuous generation of any curve, by modelling the generative process on the intuitive continuous flow of time.

Based on these concepts, the subject of the fluxionary calculus is twofold:

Now it remains, that for an Illustration of the Analytick Art, I should give some Specimens of Problems, especially such as the nature of Curves will supply. But first it may be observed, that all the difficulties of these may be reduced to these two Problems only, which I shall propose concerning a Space described by local Motion, any how accelerated or retarded. I. The Length of the Space described being continually (that is, at all Times) given; to find the Velocity of the Motion at any Time proposed. II. The Velocity of Motion being continually given; to find the Length of the Space described at any Time proposed. [Newton, 1736, Ch.1 §55-58]

This way, the problems of the “analytical art” (i.e. calculus) are reduced to two fundamental problems:
1. the relation between fluents being given, to calculate the relation between their fluxions

2. the relation between fluxions being given, to calculate the relation between fluents.

As an example of Newton’s solution to the first problem, let us take the general equation

\[ x^3 - ax^2 + axy - y^3 = 0 \] (3.1)

which expresses the relation holding between the fluents \( x \) and \( y \), at all times. The moments \( o\dot{x} \) and \( o\dot{y} \) are infinitely small time increments of the fluents, so the equation will also hold for \( x + o\dot{x} \) and \( y + o\dot{y} \), thus obtaining

\[(x + o\dot{x})^3 - a(x + o\dot{x})^2 + a(x + o\dot{x})(y + o\dot{y}) - (y + o\dot{y})^3 = 0 \] (3.2)

By applying the binomial theorem we obtain

\[
\begin{align*}
  x^3 &+ 3x^2 o\dot{x} + 3x^2 o\dot{x}^2 + 3x^2 o^2 \dot{x} + x^3 o^3 \\
  - ax^2 &- 2axo\dot{x} - ax^2 o^2 \\
  + axy &+ a\dot{x}o\dot{y} + a\dot{y}o\dot{x} + a\dot{x}o\dot{y}o^2 \\
  - y^3 &- 3y^2 o\dot{y} - 3y^2 o^2 \dot{y} - y^3 o^3 = 0
\end{align*}
\] (3.3)

After subtracting (3.1) to find the difference, the terms not containing \( o \) cancel out, then we divide everything by \( o \). The terms still containing \( o \) will be neglected, as \( o \) is infinitely small. The end result is

\[ 3\dot{x}x^2 - 2a\dot{x}x + \dot{a}xy + a\dot{y}x - 3\dot{y}y^2 = 0 \] (3.4)

Thus, starting from the relation between fluents we obtained the relation between fluxions.
Infinitely Small Quantities

In his *De analysi per aequationes numero terminorum infinitas*, written in 1669 and published in 1711, Newton employed the notion of infinitely small quantities and the binomial theorem to find the relation between the abscissa $x$ and the ordinate $y$ of a curve, given the area under the curve. An application of this is presented in [Boyer, 1949, p. 191]. Let the area be $z = \left(\frac{n}{m+n}\right) ax^{\frac{m+n}{n}}$. If we increase the abscissa by the infinitesimal quantity $o$, the area will be increased by $oy$, the area of a rectangle of width $o$. Therefore, $z + oy = \left(\frac{n}{m+n}\right) a(x + o)^{\frac{m+n}{n}}$.

The next steps are identical to the ones in the method of fluxions: apply the binomial theorem, divide by $o$ then ignore the terms containing $o$. The final result, the equation of the curve, will be $y = ax^{\frac{m}{n}}$.

This method resembles the one presented in *De methodis*, in that it rests on the infinitely small increments. However, both the theory of fluxions and the theory of infinitely small quantities in *De analysi*, presuppose the division by an infinitesimal quantity, then the elimination of the terms containing this quantity. We are thus faced with a dilemma:

1. If the infinitely small quantity $o$ is zero, then we cannot divide the equation by $o$.

2. If $o$ is not zero, then by ignoring the terms containing it, we are approximating. However, geometry requires an exact definition of figures, as observed at the beginning of this section.

Newton on Ultimate Ratios

The problems of method of fluxions and infinitely small quantities were avoided in the *Principia* and the *Treatise on Quadrature* (1693) by the use of ultimate ratios combined with the method of expansion in infinite series. In the introduction of the *Treatise on Quadrature* Newton returns to the theory of fluxions:
Fluxions are very nearly as the Augments of the Fluents generated in equal but very small Particles of Time, and, to speak accurately, they are in the first Ratio of the nascent Augments; but they may be expounded by any Lines which are proportional to them. [Newton and Stewart, 1745, p. 2]

As noted in the previous subsection, the results obtained through the methods of fluxions and infinitely small quantities were obtained by what seemed to be an approximation. Here, by relating fluxions to geometrical concepts, Newton is trying to recover the exactitude required for a proper foundation of the calculus. His proof is also geometrical. Consider the figure

3.2. Newton considers that if the areas ABC and ABDG are described by the ordinates BC, BD, then the ratio of their fluxions, i.e. the velocity of variation of areas, will be equal to the ratio of the fluxions of the two ordinates. We move BC to the new position bc, construct the parallelogram BCE\(b\) and the tangent VTH at point C which intersects bc and BA in points T and V respectively. Cc, Bb and Ec represent the increment of the curve AC, of the abscissa AB and of the ordinate BC. Now, the sides of the triangle CET “are in the first Ratio of these Augments considered as nascent, therefore the Fluxions of AB, BC and AC are as the Sides CE, ET and CT of that Triangle CET, and may be expounded by these same Sides, or, which is the same thing, by the Sides of the Triangle VBC, which is Similar to the Triangle
CET.” [Newton and Stewart, 1745, p. 2]

The same relations of proportionality can be obtained by taking the ultimate ratios of the fluxions. By moving the ordinate $bc$ over $BC$, the points $C$ and $c$ will coincide, as $CK$ will coincide with the tangent $CH$. The triangle $C Ec$ will be similar to CET, so its sides will stand in the same relation as the sides of CET are, which is the same relation as the one existent between the fluxions of $AB$, $BC$ and $AC$. To the modern reader, this procedure resembles passing to the limit of the ratio of two quantities. This interpretation is supported by Newton’s statement in the Principia (Scholium of Lemma XI):

> For those ultimate ratios with which quantities vanish, are not truly the ratios of ultimate quantities, but limits towards which the ratios of quantities, decreasing without limit, do always converge; and to which they approach nearer than by any given difference, but never go beyond, nor in effect attain to, till the quantities are diminished in infinitum. [Newton et al., 1729, p. 39]

Although Boyer views the three different foundations of Newtonian calculus as equivalent, the outline above suggests something different. The method of fluxions and the use of infinitely small quantities are inaccurate, therefore they cannot be used by mechanics for the construction of geometrical figures. The reason is that the relation Newton envisaged between rational mechanics and geometry requires that the figures constructed by mechanics be accurately described (see subsection 3.2.1). This problem is avoided by the introduction of ultimate (geometrical) ratios to justify the application of the method of fluxions. Instead of focusing on infinitely small quantities, this “new” method of fluxions considers their ratios taken to the limit. However, this also marks the return to geometry, as these ratios, as we have seen, represent ratios of geometrical, not arithmetical, magnitudes.

49
Chapter 4

Foundational Research in the 18th Century: Euler

Euler’s first book was the *Mechanica sive motus scientia analytice exposita* (1736) in which he presented the foundations of mechanics and developed a set of algorithms for solving classes of mechanical problems. The analytical tools used in the *Mechanica* were developed and justified (in print) only later on in Euler’s career, in particular in *Introductio in analysin infinitorum* (1748) and *Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serierum* (1755). Euler’s work on the foundations of the calculus, presented in *Introductio in analysin infinitorum* and *Institutiones calculi differentialis*, had to be developed in parallel to that on the foundations of mechanics, since the analytical solutions to mechanical problems could not be justified in the absence of a solid foundation of the calculus itself. Therefore, securing the legitimacy of the calculus as a reliable mathematical discipline has epistemic precedence over its application to mechanics. This is why I will begin with an analysis of Euler’s work on the foundation of the calculus then investigate the connection between the
fundamental mathematical concepts and their mechanical counterparts.

4.1 Foundations of Calculus

Three ideas lie at the basis of Euler’s calculus:

1. Functions as analytical expressions.

2. The nature of infinitely small quantities.

3. Differential calculus as a particular case of the calculus of finite differences.

We examine each in turn

4.1.1 Eulerian Functions

There are several modern definitions of functions. The following are representative:

At the present time the word “function” is used broadly to mean any determinate correspondence between two classes of objects. [Taylor, 1983, p. 2]

A function F is a set of ordered pairs \((x, y)\), no two of which have the same first member. That is, if \((x, y) \in F\) and \((x, z) \in F\) then \(y = z\). [Apostol, 1981, p. 34]

Although the definitions differ in form, essentially they say the same thing: \(f\) is a function defined on the set \(A\) with values in set \(B\), \(f : A \to B\), iff for every element \(x \in A\) there is a unique value \(f(x) \in B\). This means that it is impossible to have for one value \(x \in A\) both \(f(x) = y, y \in B\) and \(f(x) = z\) with \(z \in B\). The correspondence between \(x\) and \(f(x)\) is given in the form of an analytical expression, e.g. \(f(x) = \sin x + e^x\).

Euler’s definition of function in Introductio in analysin infinitorum is very different from the one just presented, but at that time it was a giant leap forward for mathematics. To
be precise, the concept of dependency between quantities, which we called *functional relation*,¹ had been present in human scientific thought since Babylonian times, but the ways in which this dependency was represented, the *functional expression*, varied throughout history.² [Youschkevitch, 1976] identifies three periods in the evolution of the concept of functional relation:

1. **Antiquity** - when the notion of variable quantity did not exist so the study of the functional relation focused only on particular instances of interdependency between measured quantities; during this period the function was usually expressed as a table of particular values.

2. **The Middle Ages (14th-16th century)** - when variable quantities are expressed in geometrical and mechanical terms, but the functional relation is expressed verbally, or graphically.

3. **The Modern Period (16th-18th century)** - when the introduction of analytical expressions of functional relations and the method of expressing functions as sums of infinite power series became common.

The astronomical observations of the Egyptians and the Babylonians were represented as tables of values, which “grasped” the idea of dependence between two or more quantities, but could not do so in a general manner. The concept of variable, fundamental for an abstract representation of functional relations, was lacking, and as a result the relations between

¹The adjective “functional” is introduced to distinguish between correlation and causation. If two quantities are correlated, then the variation of one them does not necessarily result in a variation of the other. Only quantities whose variation necessarily entail the variation of other quantities can be said to stand in a functional relation with the latter quantities. Mathematically, a relation is defined as a set of ordered pairs, but not all sets of ordered pairs can be defined as functions (see [Apostol, 1981, p. 34]).

²The distinction between a functional relation and the way in which this relation is expressed is drawn in [Youschkevitch, 1976] and applied to Euler’s notion of function in [Ferraro, 2000].
the interdependent quantities could only be expressed for particular cases. An example of a tabular representation of a function can be found in Ptolemy’s *Almagest*, which contains “numerous astronomical tables of other quantities, equivalent to rational functions and, also, the simplest irrational functions of the sine” [Youschkevitch, 1976, p. 40] The graphical representation of functional relations took the form of verbal descriptions of methods for constructing curves based on given values of curve parameters, e.g. abscissa, ordinate, subtangent, tangent, normal, etc. For an in-depth analysis of these two periods see Bell [1992], Boyer [1949], Kline [1990] and Youschkevitch [1976]. More relevant to our present purposes is the last period. Here, through the efforts of Euler and Lagrange, but also of Johann, Jakob and Daniel Bernoulli, the concept of functional relation becomes intimately related to that of analytical expression.

In the *Introductio in analysin infinitorum* Euler begins by defining variables and constants:

_A constant quantity is a determined quantity which always keeps the same value._ [Euler, 1990, §1] _A variable quantity is one which is not determined or is universal, which can take on any value._ [Euler, 1990, §2]

But what exactly is a universal quantity? Euler’s own remarks suggest that the relation between the concept of variable and that of quantity is similar to the relation between genus and species [Euler, 1990, §2]. Thus, by variable Euler means the essence of quantities, that set of features that all quantities have in common. The essential property of quantities, as Euler expresses in *Institutiones calculi differentialis* is the capability of being incremented or decremented: “every quantity can be increased or decreased by its own nature indefinitely” [Euler, 2000, Preface]. Modern mathematics has a formal understanding of the concept of variable, that is, a variable is a symbol which stands for a particular element of a set. In contrast, a universal, or abstract, quantity such as the Eulerian variable can “take on” any
value whatsoever:

Since all determined values can be expressed as numbers, a variable quantity takes on all possible numbers (all numbers of all types). [Euler, 1990, §2]

The expression “takes on” should not be understood from the point of view of modern mathematics. When we say that a variable takes on any value from a set, we understand that the symbol representing the variable quantity has a determinate value and this value belongs to the given set. In modern mathematics a variable is necessarily a determinate quantity and the generality of the concept of variable is formal in nature. To illustrate this, consider a simple function defined on the set of natural numbers: \( f : \mathbb{N} \to \mathbb{N}, f(x) = 2x + 1 \).

We could define \( f \), as in one of the modern definitions presented above, as the set of ordered pairs \((0, 1), (1, 3), (2, 5), (3, 7), \ldots\). However, the expression \( 2x + 1 \) is a formal generalization of this definition. It is formal, because rather than enumerating the particular values of \( f \) for every input value, we replace the symbol of the input value with the symbol \( x \) which, from then on, is treated as a determinate value. In other words, the change is in notation, not in interpretation, since \( x \) is of the same kind as any number. By contrast, for Euler variables are indeterminate quantities that are all possible numbers of all possible types. They are not “placeholders” as in the modern interpretation. They are different in kind from the values they can take. In different terms, the variable is not the result of a formal generalization, but of a generalization analogous to that from species to genus, which concerns the essence of the concept of quantity.

Hence a variable quantity can be determined in infinitely many ways, since absolutely all numbers can be substituted for it. Nor is the symbol of the variable quantity exhausted until all definite numbers have been assigned to it. Thus a variable quantity encompasses within itself absolutely all numbers, both positive and negative, integers or rational, irrational and complex numbers. Even zero and complex numbers are not excluded from the signification of a variable quantity. [Euler, 1990, §3]
This has several implications for the Eulerian notion of function to which I will now turn.

Euler’s definition of function is:

A function of a variable quantity is an analytic expression composed in any way whatsoever of the variable quantity and numbers or constant quantities.

Hence every analytic expression, in which all component quantities except the variable $z$ are constants, will be a function of that $z$; thus $a + 3z$; $az - 4z^2$; $az + b\sqrt{a^2 - z^2}$; $c^z$; etc. are functions of $z$. [Euler, 1990, §4]

From this definition it follows that a function will also be a variable quantity(§5) for simple reasons:

Since it is permitted to substitute all determined values for the variable quantity, the function takes on innumerable determined values; nor is any determined value excluded from those which the function can take, since the variable quantity includes complex values. [Euler, 1990, §5]

For Euler functions are not rules that relate the elements of one set to the elements of another set, but relations expressed analytically. Implicitly, Euler distinguishes between functional relations and their expressions, although the terms used are different:

The form of a function is changed, either by introducing a different variable or if the same variable is kept, the transformation consists in expressing the same function in a different way. [Euler, 1990, §27]

For instance, if in the function $a^4 - 4a^3z + 6a^2z^2 - 4az^3 + z^4$ we replace $a - z$ with $y$, we get the simpler expression $y^4$. The transformation affects only the way we represent the functional relation between the variable $z$ and the variable $a^4 - 4a^3z + 6a^2z^2 - 4az^3 + z^4$. Therefore, it would seem that what Euler means by function is the actual relation behind the analytic expression. This would be correct, however, were it not for the fact that Euler’s definition of the concept of function implies that the form of the function, i.e. its analytic expression, also plays a part in understanding the concept.

To understand the nature of the Eulerian concept we must answer two questions [Ferraro, 2000]:

55
Q1 A functional relation being given, what are the conditions it must meet so it can be considered a function?

Q2 A string of symbols being given, what are the conditions it must meet so it can be considered a function?

The answer to Q1 is that a functional relation is considered a function if “one was able to associate with it an algorithm consisting of symbols and rules of calculation. No function was given without a special calculus concerning it.”[Ferraro, 2000] It seems that for Euler being expressible in an analytic form is not a sufficient condition for a functional relation to become a function. Also, he requires the knowledge of the rules according to which the value of a function would be calculated for a given value of the variable. If we consider simple functions such as $x^2$ or $x + 2$, then these rules are the rules of arithmetic. A different set of rules applies to functions such as the sine and the cosine, the logarithmic and the exponential. To illustrate this, consider the function $f(x) = a + x + x^4 + \sin x$, in which we suppose the sine function unknown. In order to claim that $f(x)$ is a function, we would have to know how to calculate $f(x)$ for a particular value of $x$. However, without knowing the rules for calculating the sine function, the rules of calculation of $f(x)$ will also be unknown.

Euler classifies functions according to their rules of calculation. Thus, Euler first distinguishes between algebraic and transcendental functions. Algebraic functions relate the variables using basic arithmetical operations, “addition, subtraction, multiplication, division, raising to a power, and extraction of roots”[Euler, 1990], while transcendental functions include exponentials and logarithms. If the variable is affected by a radical sign, then it is irrational, and non-irrational otherwise. Irrational functions are further divided into explicit and implicit functions, where an implicit function is given by an equation of the form $Z^7 = nz^3$ which could not have been solved in Euler’s time. Non-irrational functions too are
further divided into polynomial and rational functions. The final picture is as follows:

1. algebraic
   
   (a) non-irrational
      
      i. polynomial
      
      ii. rational
   
   (b) irrational
      
      i. explicit
      
      ii. implicit

2. transcendental.

To recapitulate, a functional relation becomes a function when (a) it is expressed analytically and (b) its specific rules of calculation have been provided. These rules lie at the basis of Euler’s classification of functions, and the usual way of introducing them is through expansion into infinite series. The *Introductio in analysin infinitorum* contains several examples to choose from, but I will focus on the sine and cosine functions. Euler develops the rules of calculation of the two functions starting from their geometrical interpretation as functional relations between lines in a circle and the properties derived from this. He begins by considering a circle of radius 1 and circumference of $2\pi$ in which the variable quantity is an arc of the circle, labelled $z$. The next step is to enumerate the properties of sine and cosine known from geometry:

Every sine and cosine lies between +1 and -1. Further, we have $\cos z = \sin \frac{\pi}{2} - z$ and $\sin z = \cos \frac{\pi}{2} - z$.

We also have $\sin z^2 + \cos z^2 = 1$… We note further that if $y$ and $z$ are two arcs, then

$\sin y + z = \sin y \cos z + \cos y \sin z$ and

$\cos y + z = \cos y \cos z - \sin y \sin z$. Likewise

$\sin y - z = \sin y \cos z - \cos y \sin z$ and

$\cos y - z = \cos y \cos z + \sin y \sin z$. [Euler, 1990, §128]
The formula $\sin^2 z + \cos^2 z = 1$ can be also written as $(\cos z + i \sin z)(\cos z - i \sin z) = 1$. The introduction of complex terms might seem unjustified; however, “although these factors are complex, still they are quite useful in combining and multiplying arcs.” [Euler, 1990](#132).

The introduction of complex terms allows Euler to obtain a general formula for $\cos nz$ and $\sin nz$ through a process of regular (not mathematical) induction. First, he observes that

$$(\cos z + i \sin z)(\cos y + i \sin y) = \cos (y + z) + \sin (y + z) \quad (4.1)$$

From the formulas for $\cos (y + z)$ and $\sin (y + z)$ quoted above, we get by substitution in (4.1):

$$(\cos z + i \sin z)(\cos y + i \sin y) = \cos (y + z) + \sin (y + z) \quad (4.2)$$

Likewise

$$(\cos z - i \sin z)(\cos y - i \sin y) = \cos (y + z) - \sin (y + z) \quad (4.3)$$

If instead of two variables we had three, then, by following the same steps we get

$$(\cos x \pm i \sin x)(\cos y \pm i \sin y)(\cos z \pm i \sin z) = \cos (x + y + z) \pm \sin (x + y + z) \quad (4.4)$$

For $x = y = z$ equations (4.3) and (4.4) become:

$$(\cos x \pm i \sin x)^2 = \cos 2x \pm i \sin 2x$$

$$(\cos x \pm i \sin x)^3 = \cos 3x \pm i \sin 3x$$
From these two cases Euler concludes that the same relation holds for higher powers, i.e.

\[(\cos x \pm i \sin x)^n = \cos nx \pm i \sin nx \quad (4.5)\]

The lack of rigor introduced by this inductive inference might make a modern mathematician cringe, but although the justification is not perfect, the conclusion is still correct. In order to meet the more stringent requirements of modern mathematics, we would have to consider (4.5) an assumption, rather than a conclusion and try to prove that if (4.5) is true, then \[(\cos x \pm i \sin x)^{n+1} = \cos (n+1)x \pm i \sin (n+1)x\] is also true. That Euler’s generalized formula is correct even by modern standards can easily be proven:

\[
(\cos x + i \sin x)^{n+1} = (\cos x + i \sin x)(\cos x + i \sin x)^n
\]

\[
(\cos x + i \sin x)(\cos nx + i \sin nx) = \cos x \cos nx
\]

\[
- \sin x \sin nx + i(\cos x \sin nx + \sin x \cos nx) = \cos x + nx + i \sin x + nx = \cos (n+1)x + i \sin (n+1)x
\]

The generalized formula allows us to express the sine and cosine in a form to which we can apply the binomial theorem:

\[
\cos nz = \frac{(\cos z + i \sin z)^n + (\cos z - i \sin z)^n}{2} \quad (4.6)
\]

\[
\sin nz = \frac{(\cos z + i \sin z)^n - (\cos z - i \sin z)^n}{2} \quad (4.7)
\]
Which gives

\[
\cos nz = (\cos nz)^n - \frac{n(n-1)}{1\cdot2}(\cos z)^{n-2}(\sin z)^2 \\
+ \frac{n(n-1)(n-2)(n-3)}{1\cdot2\cdot3\cdot4}(\cos z)^{n-4}(\sin z)^4 \\
- \frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{1\cdot2\cdot3\cdot4\cdot5\cdot6}(\cos z)^{n-6}(\sin z)^6 + \ldots
\]  

(4.9)  

(4.10)  

(4.11)  

(4.12)

Euler assumes that for an infinitely large \(n\) and an infinitely small \(z\), \(nz\) will be a finite number, \(v\). Moreover, since \(z\) is an infinitely small arc, \(\sin z = z\) and \(\cos z = 1\).\(^3\) An in-depth analysis of Euler’s view on infinite quantities, both large and small, will be done in the next subsection. For now it suffices to draw attention upon the connection between the development of the rules of calculation of a function and assumptions regarding infinite quantities. The introduction of the infinitely small leads to a simplified\(^4\) form of the series:

\[
\cos v = 1 - \frac{v^2}{1\cdot2} + \frac{v^4}{1\cdot2\cdot3\cdot4} - \frac{v^6}{1\cdot2\cdot3\cdot4\cdot5\cdot6} + \ldots
\]

(4.13)

This formula is the rule of calculation of the cosine. A similar rule can be deduced for the sine.

To recapitulate, we tried to answer Q1, i.e. given a functional relation, what are the conditions that it must meet so it can be considered a function? The answer is that a functional relation becomes a function once it is expressed analytically and its rules of calculation have been given, with the expression “rule of calculation” referring to the procedures of calculating the value of a function for any value of the variable, as well as the method of derivation and

\(^3\)Euler takes these formulas not as approximations, but as actual equalities.

\(^4\)Since for Euler \(\sin z\) is equal to \(z\), the result is not an approximation, but a more manageable analytical expression of the series.
integration of a function. Only when we can calculate the value of the function for a given
value of the variable, deduce its derivative and its integral can we say that the function is
known. One thing to note is that the rules of calculation must refer to variables in Euler’s
sense, i.e. universal quantities, as a consequence of the definition of function. As such, the
rules for calculating the value of a function are abstracting away from any particular inter-
pretation of the variables in question. This answer is supported first, by Euler’s classification
of functions and second, by his presentation of the sine and cosine, which begins by assuming
the geometrical interpretation of the two functional relations and ends by expressing them
in non-geometrical terms. As it can be seen in equation (4.13), the cosine is detached from
its geometrical background and expressed in an algebraic form. If before Euler the values of
the sine and cosine were calculated using geometrical methods which implicitly assumed that
the variable was a geometrical quantity, after Euler the two functions (together with \( \tan x \)
and \( \cot x \)) were generalized. This is in agreement with the note made earlier regarding the
process of abstraction presupposed by the rules of calculation.

The answer to Q2 is similar to Q1: a string of symbols respecting the syntactical rules of
algebra was a function if the rules of calculation were known and expressible in an analytical
form. Knowledge of the functional relation being implicit. To illustrate this, consider the
following expression (F): \( a + x + y + \text{qwerty}(x, y, z) \). (F) is a string of symbols, probably a
function, but since we do not know the rules of calculation of the function \( \text{qwerty}(x, y, z) \),
(F) itself will be no more than a pseudo-mathematical expression. Assuming we know that
\( \text{qwerty}(x, y, z) = x + y + z \), (F) becomes a function and the functional relation whose ana-
lytic expression (F) is also brought to light, for now we know that the relation holds between
\( a + x + y + \text{qwerty}(x, y, z) \) and \( x, y, z \). More precisely, we know the functional relation ex-
hibited by an analytical expression if we know which quantities are interdependent and how
this dependency can be evaluated.
We can get a better picture of Eulerian functions by looking at some of their properties. Since functions are variables, they will have the same properties, i.e. they can be incremented or decremented and they “encompass” within themselves absolutely all numbers of all types. The modern idea of function defined on the set of rational numbers is excluded from the outset by considering functions to be Eulerian variables. More importantly, because functions are variables it follows that they are also continuous, but not in the modern mathematical sense of continuity. For Euler an object was continuous if it was taken as an unbroken object. As a variable a function was necessarily continuous, i.e. an unbroken object, due to the definition of variable as “essence of quantity.” From here it follows that a function has to be defined by a single analytic expression. Functions such as

\[
\begin{align*}
f(x) &= \begin{cases} 
x - 2 & : x \leq 2 \\
3x - 6 & : x > 2
\end{cases}
\end{align*}
\]

(4.14)
despite being continuous in a modern sense are nevertheless discontinuous for Euler, since the two formulas are not equivalent.

Another important property of functions is their capability to be expanded into infinite series. This property was illustrated above, in the presentation of the sine and cosine functions, but in the fourth chapter of *Introductio in analysin infinitorum* Euler shows how any kind of function (algebraic or transcendental) can be expressed as a series:

Thus there is no doubt that any function of \( z \) can be given the form \( Az^\alpha + Bz^\beta + Cz^\gamma + Dz^\delta + \ldots, \)
where the exponents \( \alpha, \beta, \gamma, \delta, \) etc. are any real numbers. [Euler, 1990, §59]
4.1.2 Infinitely Small Quantities

Originally, the main concern of the calculus was finding the value of a geometrical variable for a given curve. Solutions to these problems depended on the nature of the particular curve under scrutiny, so the method used to find the subtangent, for instance, of one curve could not have been applied to find the subtangent of a different curve. In contrast, the calculus aimed to provide a unitary approach to these problems, i.e. to develop algorithms. This is illustrated in the *Introductio in analysin infinitorum*, where Euler emphasizes the role played by the algorithm of expansion into infinite series for the introduction of functions and in the previous subsection we have seen how this was applied to the sine and cosine functions. It is also shown in the preface of the *Institutiones calculi differentialis* where Euler notes the importance of Leibniz’s work on the calculus:

> We are no less indebted to Leibniz insofar as this calculus at that time was viewed as individual tricks, while he put it into the form of a discipline, collected its rules into a system, and gave a crystal-clear explanation. From this there followed great aids in the further development of this calculus, and some of the open questions whose answers were sought were pursued through certain definite principles. [Euler, 2000, Preface]

The search for generalization inevitably led to a departure from the geometrical “tricks” in the form of rephrasing geometrical problems in terms of functions and variables. This paved the way for the introduction of algebraic algorithms to deal with these problems. Not surprisingly, in the Preface of the *Institutiones calculi differentialis* Euler points out that “everything is kept within the bounds of pure analysis, so that in the explanation of the rules of this calculus there is no need for any geometric figures.” [Euler, 2000, Preface] The algorithms introduced in infinitesimal calculus (henceforth CI) in the *Institutiones calculi differentialis* are the same as those for the calculus of finite differences (henceforth CFD):

---

5See Fig.3.1 on p. 41
The analysis of the infinite, which we begin to treat now, is nothing but a special case of the method of differences, explained in the first chapter, wherein the differences are infinitely small, while previously the differences were assumed to be finite. [Euler, 2000, §114]

How does Euler justify this claim? There are three notions we must clarify before attempting to answer:

1. The object of CI
2. The nature of infinitely small quantities
3. The algorithms of CFD.

I shall turn now to a detailed analysis of these points.

Object of the Calculus

In the Institutiones calculi differentialis Euler defines the CI as

\[\ldots \text{a method for determining the ratio of the vanishing increments that any functions take on when the variable, of which they are functions, is given a vanishing increment.} \]

[Euler, 2000, Preface]

The definition suggests that for Euler CI is not a calculus of differences, but rather a calculus of increments. The difference is that in the calculus of differences, we must know the values of the function for two different known values of the variable in order to determine the variation of the function. In the calculus of increments, all we need to know is one value of the variable and the increment. Then the next value of the variable can be easily determined, and so will the corresponding values of the function. The difference between More importantly, Euler’s calculus is not a study of functions, but of ratios of vanishing increments. In Euler’s foundations of calculus, functions play a role insofar as they are variables, i.e. they can be incremented or decremented when the value of their variable is increased. The ratios of vanishing increments are the objects of investigation of the calculus. It is only in Lagrange’s
*Leçons sur le calcul des fonctions* and *Théorie des fonctions analytiques* that “function” becomes the central concept of the differential calculus, by defining the differential of a function as the functional coefficient of a variable in an expansion in power series.⁶

Simply put, for Euler the calculus tries to answer this question: for a given function of variable \( x \), “if the quantity \( x \) is increased or decreased, by how much is the function changed, whether it increases or decreases?” [Euler, 2000, Preface] The *Institutiones calculi differentialis* can be viewed as giving two answers, one in the Preface, another in the fourth chapter. The answer in the Preface considers a determinate increment \( \omega \) of the variable \( x \) and tries to find the corresponding increment of the function \( x^2 \). In modern notation,

\[
\begin{align*}
f(x) &= x^2 \\
f(x + \omega) &= (x + \omega)^2 \\
f(x + \omega) &= x^2 + \omega^2 + 2x\omega \\
f(x + \omega) - f(x) &= 2x\omega + \omega^2
\end{align*}
\]

For an increment \( \omega \) of the variable the function is incremented by \( 2x\omega + \omega^2 \). “Hence, the increase in \( x \) is to the increase of \( x^2 \) as \( \omega \) is to \( 2x + \omega + \omega^2 \), that is, as 1 is to \( 2x + \omega \)” [Euler, 2000, Preface]. In general, the ratio of the increment of any function and the increment of the variable or variables is, in Euler’s view, the foundation of the whole analysis of the infinite. So far, the increment \( \omega \) had been treated as a determinate quantity, be it finite or infinite (small or large). Euler’s next step is to consider \( \omega \) a vanishing increment, i.e. \( \omega \) “goes to zero”. Therefore, the ratio \( \frac{f(x+\omega)-f(x)}{\omega} = 2x + \omega \) will approach \( 2x \). Of course, if \( \omega \) approaches zero both increments vanish so their ratio should approach \( \frac{0}{0} \), but Euler insists that their

⁶See the next chapter on this issue.
ratio is nevertheless a determinate quantity.

The approach in the fourth chapter differs from the one in the Preface in that it includes the idea of “going to zero” within the notion of increment. More precisely, instead of beginning by incrementing the variable with a determinate quantity and calculating the ratio as the increment approaches zero, in the fourth chapter Euler increments the variable by an infinitesimal quantity \( dx \), calculates the corresponding variation of the function \( dy \) and only then calculates the ratio of the two infinitesimals. For [Bos, 1974], this inconsistency shows that Euler’s view on infinitely small quantities does not influence his calculus. In order to analyze this claim we must turn to Euler’s considerations on the nature of the infinitely small.

In the Preface of the *Institutiones calculi differentialis* Euler makes clear that in reality infinitely small quantities are nothing:

> They [vanishing increments] are called *differentials*, and since they are without quantity, they are also said to be *infinitely small*. Hence, by their nature they are to be so interpreted as absolutely nothing, or they are considered to be equal to nothing. [Euler, 2000, Preface]

“Absolutely nothing” here means zero. In the 18th century a number was viewed as “a determination of quantity and was generated from the flow of quantities; in particular, zero was generated from a quantity that became nothing.” [Ferraro, 2004, p. 45] Since numbers are quantities, it follows that 0 could not have been conceived as a number. Actually, in his *Elements of Algebra* Euler excludes 0 from the enumeration of natural and negative numbers.

As can be observed from Newton and Euler’s calculations of the ratios of infinitesimals,
these quantities have a double nature. Consider Euler’s calculation of the ratio of \( f(x) = x^2 \):

\[
\begin{align*}
  f(x) &= x^2 \quad (4.15) \\
  f(x + \omega) &= (x + \omega)^2 \quad (4.16) \\
  f(x + \omega) &= x^2 + 2x\omega + \omega^2 \quad (4.17) \\
  f(x + \omega) - f(x) &= 2x\omega + \omega^2 \quad (4.18) \\
  \frac{f(x + \omega) - f(x)}{\omega} &= 2x + \omega^2 \quad (4.19) \\
  \omega \to 0 \quad (4.20) \\
  \frac{f(x + \omega) - f(x)}{\omega} &= 2x \quad (4.21) \\
  f(x + \omega) - f(x) &= 2x \quad (4.22)
\end{align*}
\]

In the equations from (4.15) to (4.19) \( \omega \) is considered a regular variable quantity that could increment or decrement a given quantity. However, after (4.20), \( \omega \) is completely ignored and Euler maintains that this does not affect the accuracy of the result. In general, if differentials were not zeros from a mathematical perspective, then the calculus would fail to meet the same standards of rigor as geometry. The reason is that no matter how small these quantities are, the errors they introduce eventually add up so that the end result will not be exact:

…if these infinitely small quantities, which are neglected in calculus, are not quite nothing, then necessarily an error must result that will be greater the more these quantities are heaped up. [Euler, 2000, Preface]

If \( \omega \) would have been considered 0 before (4.15) then the difference \( f(x + \omega) - f(x) \) would be 0 and the left hand of (4.19) would be reduced to \( \frac{0}{0} \). If all increments are zeros, what is the point of using different symbols to denote them?

Since all nothings are equal, it seems superfluous to have different signs to designate such a quantity. [Euler, 2000](§84)

67
To solve this conundrum, Euler makes use of the distinction between two means of comparing quantities: arithmetic and geometric. I will talk, following Euler, about arithmetic and geometric ratios, but these expressions do not refer to two different interpretations of $a/b$.

“Arithmetic ratio” and “geometric ratio” should be understood as referring to the comparison of terms of an arithmetic and geometric progression, respectively. Thus, the arithmetic ratio of two quantities $a$ and $b$ is $a - b$, while the geometric ratio is $a/b$. In more precise terms, the geometric ratio “is found by resolving the question, How many times is one of those numbers greater than the other? This is done by dividing one by the other; and the quotient will express the ratio required.” [Euler et al., 1822, §440] For Euler the ratio of two infinitesimal quantities is geometric.

Everyone knows that when zero is multiplied by any number, the product is zero and that $n \cdot 0 = 0$, so that $n : 1 = 0 : 0$. Hence, it is clear that any two zeros can be in a geometric ratio, although from the perspective of arithmetic, the ratio is always of equals. [Euler, 2000]§85 From here it follows that we must use different symbols to denote the zeros, even though they are essentially equal. The reason is that $\frac{0}{0}$ can have any value whatsoever, and there would be no means to distinguish between them (e.g. $\frac{0}{0} = \frac{2}{3}$ and $\frac{0}{0} = \frac{4}{5}$ are indistinguishable, if we use 0 to denote differentials). For instance, $dx = adx = 0$, but the ratio $\frac{adx}{dx} = a$, is a determinate quantity.

Newton’s theory of ultimate ratios, which resembles Euler’s view of geometric ratios, had a geometric justification. Ultimate ratios had a finite value because they were proportional to the ratio of the sides of a triangle. This relation was deduced from the similarity between two triangles. However, Euler wanted to reduce the dependency of the calculus on geometry, so he had to justify the finite values of ratios differently.

The intuition behind Euler’s account of ratios of differentials is that of the terms of two series which increase or decrease differently for the same value of the index. Compare the
series with general terms $\frac{1}{n^2}$ and $\frac{1}{n^3}$:

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1 = \frac{1}{n^2}$</td>
<td>$\frac{1}{1}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{9}$</td>
<td>$\frac{1}{16}$</td>
<td>$\frac{1}{25}$</td>
</tr>
<tr>
<td>$S_2 = \frac{1}{n^3}$</td>
<td>$\frac{1}{1}$</td>
<td>$\frac{1}{8}$</td>
<td>$\frac{1}{27}$</td>
<td>$\frac{1}{64}$</td>
<td>$\frac{1}{125}$</td>
</tr>
</tbody>
</table>

It can be observed that for the same value of $n$, $S_2(n)$ is smaller than $S_1(n)$. The terms of the two series become smaller and smaller and approach 0, but they do so at a different rate.

The ratio of the rates is analogous to the ratio of infinitesimals. This resembles Newton’s method of fluxions. Just as a fluxion is the velocity with which fluents change in time, the ratio of infinitesimals can also be viewed as a velocity only not relative to time, but to the terms of an arithmetical progression, e.g. $n, n+1, n+2, n+3, \ldots$

To sum up, Eulerian infinitesimals are zero, but from a formal point of view they are treated as regular quantities. This is the “before” interpretation of infinitesimals, i.e. the infinitesimals in equations (4.15)-(4.19). However, infinitesimals also have a “modal” aspect, a disposition to “tend” towards 0. This would be the “after” interpretation, as appears in (4.21).

The presentation of the calculus in the Preface of the Introductio in analysin infinitorum only seems at odds with the one in the fourth chapter of the same book. In reality they are emphasizing different aspects of the concept of infinitesimal. The presentation in the Preface focused on the modal aspect, i.e. on an increment approaching nothing, while the fourth chapter emphasizes the formal aspect of infinitesimals, i.e. zero quantities. By treating differentials formally as zeros, Euler confers geometrical rigor to the calculus. The double nature of the infinitesimal lies at the basis of Euler’s extension of the rules of cfd to ci, which will be the focus of the next subsection.
4.1.3 Calculus of Finite Differences (cfd) and Infinitesimal Calculus (ci)

The first chapter of the *Institutiones calculi differentialis* presents the cfd. The method employed here is similar to the one used in the Preface to present the ci. Considering $y$, a function of $x$, how does $y$ change if $x$ is incremented by a finite quantity $\omega$? For each increment of $x$ there corresponds a different value of $y$: The next step is to define the variation of $y$ for each increment of the variable $x$:

\[
\begin{align*}
x & \quad x + \omega & \quad x + 2\omega & \quad x + 3\omega & \quad x + 4\omega & \quad x + 5\omega \\
\Delta y & = y - y' & \Delta y' = y'' - y' & \Delta y'' = y''' - y'' & \ldots
\end{align*}
\]

To each increment of $x$ there is a corresponding variation of $y$:

\[
\begin{align*}
x & \quad x + \omega & \quad x + 2\omega & \quad x + 3\omega & \quad x + 4\omega & \quad x + 5\omega \\
\Delta y & = y & \Delta y' & \Delta y'' & \Delta y''' & \Delta y'''
\end{align*}
\]

From here, we can calculate higher order differentials such as the second ($\Delta \Delta y = \Delta y' - \Delta y$; $\Delta \Delta y' = \Delta \Delta y'' - \Delta \Delta y'$; \ldots), third ($\Delta^3 y$; $\Delta^3 y'$; $\Delta^3 y''$; \ldots) and fourth ($\Delta^4 y$; $\Delta^4 y'$; $\Delta^4 y''$; \ldots) order. Obviously, the values of the increments depend on the nature of $y$. For instance, if $y = x$, all increments of $y$ will be constant ($\Delta y = \omega$; $\Delta y' = \omega$; \ldots) and all increments of order higher than one will be zero ($\Delta^2 y = 0$; $\Delta^3 y''' = 0$; \ldots).

For a function $y = p + q$ where both $p$ and $q$ are functions of $x$, for a finite increment of $x$ the increment of $y$ is $y' - y = (p' + q') - (p + q) = \Delta p + \Delta q$. With equal ease we obtain
the rules of differentiation for basic algebraic functions:

\[ y = p + q \quad \Delta y = \Delta p + \Delta q \]
\[ y = p - q \quad \Delta y = \Delta p - \Delta q \]
\[ y = pq \quad \Delta y = p\Delta q + q\Delta p + \Delta p\Delta q \]
\[ y = p/q \quad \Delta y = \frac{q\Delta p - p\Delta q}{q^2 + q\Delta q} \]

In general, the increment of order \( n \) of a function \( y \) for the increment \( \omega \) of \( x \) is given by the relation:

\[ \Delta^n y = P\omega^n + Q\omega^{n+1} + R\omega^{n+2} + S\omega^{n+3} + \ldots \]  \hspace{1cm} (4.23)

Proving (4.23) would take too long and it would not present anything new. It is worth mentioning nonetheless, since it is the starting point of Euler’s presentation of the differential calculus:

In this chapter, and in all of the analysis of the infinite, the increment \( \omega \) by which we let the variable \( x \) increase will be infinitely small, so that it vanishes; that is, it is equal to 0. Hence it is clear that the increase, or the difference, of the function \( y \) will also be infinitely small. With this hypothesis, each term of the expression

\[ P\omega + Q\omega^2 + R\omega^3 + S\omega^4 + \ldots \]  \hspace{1cm} (4.24)

will vanish when compared with its predecessor, so that only \( P\omega \) will remain. [Euler, 2000, §113]

In this sense the differential calculus is a special case of the calculus of finite differences. From a formal point of view, the change from CFD to CI is expressed by the replacement of \( \Delta \) with \( d \) and in (4.23) by replacing the finite differences with the infinitely small differences. If \( \omega = dx \) then the terms containing higher-order differentials will be 0 relative to \( dx \), just as \( dx \) is 0 relative to any finite quantity. The proof of this is easy. We know that \( dx = 0 \), it follows that a finite quantity cannot be incremented by an infinitesimal, i.e. \( a \pm dx = a \).
Therefore, an infinitesimal “vanishes” in comparison to a finite quantity.

\[ a \pm ndx = a \]
\[ a \pm ndx - a = 0 \]
\[ (a \pm ndx) : a = a : a = 1 \]

In a similar manner we can prove that differentials of higher powers vanish in comparison to the same differential of power 1:

\[ (dx \pm dx^2) : dx = \frac{dx \pm dx^2}{dx} = 1 \pm dx = 1 \]

Because \( dx \pm dx^n = dx \), the first-order differential of \( y \) will be \( Pdx \), that is, \( dy = Pdx \).

For Euler second-order differentials arise from second-order finite differences, not from first-order differentials:

These[second order differentials] arise from second differences...when we let \( \omega \) become the infinitely small \( dx \). If we suppose that the variable \( x \) increases by equal increments, then the second value \( x' \) becomes equal to \( x + dx \), and the following will be \( x'' = x + dx \), \( x''' = x + 3dx \)...Since the first differences \( dx \) are constant, the second differences vanish, and so the second differential of \( x \), that is, \( d^2x \), is equal to 0.

[Euler, 2000, §124]

4.2 Foundations of Mechanics

Euler developed his foundations of mechanics in parallel with those of the calculus, although the latter were published later. [Suisky, 2009] notes that the Mechanica is an application of Euler’s mathematical methods to mechanical problems. The task at hand is to find Euler’s justification for representing mechanical quantities using mathematics. A key part of this is the analogy between Euler’s view on the essence of material bodies and the nature of the infinitely small.
Euler envisages two ways in which one might be able to find the essential properties of material bodies. The first would be to examine all bodies and individuate all their common properties. As such, it is impossible to pursue, since the totality of bodies is potentially infinite. The second way would be to focus on a smaller set of bodies and individuate those properties in virtue of which they are bodies.

For a property which is such that when it is absent in a thing, we would not regard the thing as a body, can justly be regarded as a general property of bodies. [Euler, 1862](14)

This process is purely rational so the end result will be the logically essential properties of bodies, i.e. those properties which, if absent, the body would not be a material body. Furthermore, the essence of material bodies is not a common property of all bodies (the same way shape might be), but a condition of possibility for a thing to be considered a material body. Although Euler does not explicitly mention this, he is assuming the existence of immaterial bodies. This is explainable, considering his religious background, but need not concern us.

Following this method, Euler finds impenetrability to be the essence of material bodies:

_Every material body must occupy in space a particular location, and it is impossible for two bodies to be at the same location at the same time._[Euler, 1862](85)

This is an _a priori_ truth, although Euler does not use this terminology. Should a body lack this property, then it is possible for two bodies to occupy the same location, but in this case there would be no way to distinguish between the two bodies. Moreover, as explained in _An Introduction to Natural Science_, a body must be impenetrable, otherwise it would be inaccessible to our senses, since rays of light would “pass through” the body without being reflected, so they could not be seen. Likewise, if there were penetrable objects, then they could not be sensed by touch, since our hands would pass through them without meeting any resistance.
From impenetrability three other general properties of bodies follow: *extent, mobility* and *persistence*. Moreover, since each property presupposes other properties, the latter will necessarily also be properties of material bodies. Extent necessarily follows from impenetrability since the latter concept is defined by reference to location, which requires that the body should also have the property of extent. If a body does not have extent, then the property of impenetrability cannot be attributed to it, so it cannot be considered a material body.

1. That *extent* is a necessary property of material bodies is a conclusion we can arrive at from experience:

   We are not only convinced by our experience that all material bodies that we know possess extent, but our concept of bodies incorporates extent in such a way, that we can exclude all things without extent from the category of material bodies. [Euler, 1862, §9]

A number of properties follow with necessity from the idea of bodies having extent:

- **Divisibility:** “Everything that has extent is divisible, and divisibility can be continued *ad infinitum*; therefore all material bodies are infinitely divisible.” [Euler, 1862, §11]

- **No atomic components:** “Although, in view of its divisibility, a body must be regarded as a composite thing, it is nevertheless not composed of simple things” [Euler, 1862](§13)

- **Dimensionality:** “Every material body . . . is always extended in the so called three dimensions of length.” [Euler, 1862](§15)

Regarding the first, Euler draws upon an analogy with geometry to prove that extent necessarily entails divisibility. Given a segment of length $l$, any division of the line will result in a number of smaller segments of definite length, $l/2$, $l/5$, $l/100$, etc. This process can be carried *ad infinitum*, since every smaller segment would have to have a certain size and extent, and in virtue of it having extent, it can be further divided. The argument can also be applied to real material bodies, for after repeated divisions if we reach a part that cannot
be divided, then that part has no size. However, it is impossible for this to happen, for there is no reason to believe that a part that is a third or a quarter of the material body has size, and at the same time deny that the millionth part of the body has no size.

Even though a body is infinitely divisible, Euler does not think this entails that bodies are composed of indivisible parts. Of course, they are composed of parts, but there are no “ultimate” parts (atoms). The reason being that atoms must have extent. Which means they can be divided even further. Finally, saying that material bodies have extent means that material bodies have extent in all three dimensions. There are bodies that have extent, but cannot be considered material bodies. Lines are geometrical bodies which have extent, but only in one dimension. Surfaces too have extent but only in two dimensions. Material bodies must have three dimensions, and none more. Euler does not explain why this is the case, but merely postulates that

A body must have threefold extent, in length, in breadth and in depth. There are no further kinds of extent. [Euler, 1862](§15)

2. Mobility is defined as the possibility of a material body to occupy another location. Just as with the property of extent, there are two ways to prove the necessity of mobility. First, if we assume that impenetrability is the essence of material bodies, then a body cannot be anything but mobile. Assuming the contrary would mean that a body is permanently attached to a location. This can only be the case if an external force acts continuously on the body keeping it in the same location; the force must be external “for in the thing itself one can find nothing that would prevent it from being moved from this position.” [Euler, 1862](§38) Therefore, even if an immovable body is conceivable, the cause of this immovability is not inherent in the body itself, so we can consider mobility a necessary property of material bodies. Alternatively, mobility is also proved by observing that there is nothing in the concept of space or the concept of body that would “tie down” a body to a particular location. There is no sufficient
reason to believe that a body would not move should a force act upon it.

3. **Persistence** is the property of bodies to remain in the same state in the absence of any external influences. If the body is in a state of rest, then there is no reason for it to change its position on its own:

> There is however no reason why it should move in one direction rather than in a different one, and in the absence of any sufficient reason for this, we can conclude with certainty that a body, once at rest, will remain in that state, unless an external cause arises that can set it in motion. [Euler, 1862, §26]

A similar argument applies to the case of bodies in motion. If a body moves in a straight line, then it will keep doing so in the absence of any external interactions, for there are no inherent causes of sudden changes in its behavior. On the other hand, should the body’s trajectory be an arc, then this must be caused by the action of an external force. This does not mean that if a body moves in a straight line then no force is acting upon it. A force may affect two parameters of a body’s dynamic: the trajectory and the velocity. A curvilinear motion is necessarily brought about only by the action of a function, but also the velocity of a body can be altered (increased/decreased) only through the action of an external force. Persistence is a direct consequence of mobility. If a body is mobile, then unless it also has the property of persistence, its direction, velocity, and state would change without any external cause.

It will turn out that the applicability of the calculus to mechanics rests upon this view of bodies. This is expressed in Euler’s *Mechanica sive motus scientia exposita* which contains the analytical representation of mechanics using the mathematical tools of the differential calculus. His attitude towards the use of the analytical method is expressed in the Preface:

> But as with all writings composed without analysis, and that mainly falls to be the lot of Mechanics, for the reader to be convinced of the very truth of these propositions offered, an examination of these propositions cannot be followed with sufficient clarity and distinction: thus as the same questions, if changed a little, cannot be resolved from what is given, unless one enquires using analysis, and these same propositions are explained by the analytical method. Thus, I always have the same trouble, when
I might chance to glance through Newton's *Principia* or Hermann's *Phoronomiam*, that comes about in using these, that whenever the solutions of problems seem to be sufficiently well understood by me, that yet by making only a small change, I might not be able to solve the new problem using this method. [Euler, 1736, Preface]

The use of analytical tools for mechanical problems has the same motivation as their use in geometry: generality. The calculus does not apply, like geometrical methods, to particular problems, but to classes of mechanical problems. Obviously, this is possible because the fundamental concepts of the calculus, i.e. variables, constants, functions, have been defined (by Euler) independent of geometry, although not independently of geometrical intuitions.

In the Preface Euler also mentions the fundamental role played by infinitely small bodies:

Moreover I have striven in the classification of the work to distinguish between bodies which can be moved and those which are fixed, as being either free or not. The division supplied by me has taken into account the innate character of the bodies themselves, as in the first case I shall investigate the motion of infinitely small bodies, or as it were of points, then indeed I shall move on to investigate bodies of finite size and these can be either rigid or flexible, or I may proceed to extended bodies which in turn are entirely free. For as in geometry, in which the measurement of bodies is discussed, it is customary to begin with a discussion of points, thus also the motion of bodies of finite sizes cannot be explained, unless first the motion of points, from which the bodies are considered to be composed, has been carefully examined. [Euler, 1736, Preface]

The existence of infinitely small bodies is a consequence of the property of extent of material bodies. As mentioned above, in virtue of having the property of extent, material bodies also have the property of infinite divisibility. This means that infinitely small bodies are, in theory, possible. The analogy between the concept of infinitely small quantities and that of infinitely small bodies justifies the application of the differential calculus to problems of mechanics. One question we must answer is whether this application is justified by more than this analogy.

The *Mechanica* begins by defining the fundamental concepts of mechanics: motion, rest, place, and velocity. The place occupied by a body **"is a part of the immense or boundless space**
which constitutes the whole world.” [Euler, 1736, §4] Based on this concept, we define absolute motion as the translation of a body from one place to another, rest as the preservation of place, and velocity as the variation of the distance in time. Relative motion and relative rest on the other hand, assume a bounded space in which the translation or lack thereof takes place. Euler seems to focus more on the relative notions than on the absolute ones, since the latter are not easy to manipulate:

Moreover, because of the immense nature of space and of its unbounded nature...we are unable to form a fixed idea of this[absolute motion/rest]. Thus, in place of this immense space within which bodies can move, we are accustomed to defining a finite space and the limits within which bodies can move, from which we can indicate the states of motion and of rest of bodies. [Euler, 1736, §7]

The notion of infinite divisibility is not the only connection between the foundations of mechanics and those of the calculus. The idea of translation of a point of mass from one location to another is analogous to the continuous generation of geometrical curves. Translation has to be continuous, otherwise, if a body were to move from one place to another in discrete “jumps”, then it would have to be annihilated at the starting point and then recreated at the destination, “which cannot be done according to the laws of nature, except by agreeing to a miracle” [Euler, 1736, §13]. [Suisky, 2009] points out that Euler describes translation and motion “in terms of finite and infinitesimal increments of the path” [Suisky, 2009, p. 121]. That this is the case can be seen in Proposition 3 of the Mechanica:

33. In motion with any non-uniformity, the smallest elements of the distance are considered to be traversed by uniform motions [Euler, 1736, §33].

The demonstration of this proposition draws upon an analogy with the geometrical division of curves into infinitely small segments. Just as a curve can be divided into segments, non-uniform motion can be divided into infinitely small increments of uniform velocities: “For either the elements are actually traversed in a uniform motion, or with the change of the
speed by an element of this kind is so small, that the increment or decrement can be ignored without error.”[Euler, 1736, §33]. Mathematically, this means that if the initial velocity is \( c \), the velocity in the second element is \( c + dc \), the third \( c + 2dc + ddc \), i.e. \((c+dc)+ d(c+dc)\), and so on.

However, since in the calculus of infinitely small quantities the increment \( dx \) of a variable was considered zero, one might be tempted to think that representing the increment of velocity in a similar manner would mean that there is no increment at all, since \( dc = 0 \). In consequence, this would be the same as saying that there is no variation of velocity. The reply to this objection is that velocity is the geometrical ratio of an infinitesimal increment of space and of an infinitesimal increment of time so the ratio has a determinate value. Therefore, in the case of non-uniform motion, the increment \( dv \neq 0 \).

Since the cause of the change of state of a body (from motion to rest and vice versa) is not inherent in the body (due to the property of persistence), it follows that a body changes state only as a result of an external action. This cause is called force:

A force is an action on a free body that either leads to the motion of the body at rest or changes the motion of that body.[Euler, 1736, §99]

With the introduction of forces the Mechanica turns towards geometry. Euler proves the principle of composition of forces (PCF) and the principle of the lever (PL) geometrically.

Take for instance, Proposition 14, in the second chapter of Mechanica

For the given effect of an absolute force on a point at rest, to find the effect of the same absolute force on the same point in some kind of motion.

The solution of this problem is not analytical, but geometrical. Euler assumes that the point is initially at \( A \) then assumes that if the point is at rest then in \( dt \) he will be at \( C \) as a result of a force acting upon it. If the force does not act, then the point will reach, in \( dt \), position \( B \). The length of \( AB \) is \( cdt \). The effect of the force will be the segment \( BD = AC \).
Euler’s introduction of functions was a giant leap forward in decoupling calculus from geometry. However, when solving various mechanical problems, he still appeals to geometrical figures and methods. Therefore, we can say that although the function came to play a greater role in his calculus, the intuitions behind it were still geometrical.

The relevance of geometry to Euler’s mechanics goes a bit further than this. Geometry connects the calculus to an idealized version of reality. Consider Euler’s remarks on the nature of material bodies and compare them with the concept of point. Both geometrical points and material bodies are infinitely divisible and neither of them is composed of simpler parts. The two entities differ in kind. An infinitely small quantity is necessarily impenetrable, while the same does not hold for the point.

This explains why Euler uses geometry to represent mechanical problems. Points resemble infinitely small bodies, and on the basis of this relation we can represent the latter through the former. If we wanted to know whether this relation of resemblance is actually accurate, we would have to know Euler’s definition of point. I am assuming that he has in mind the Euclidean definition of point as that which has no parts.

The use of the calculus is also connected to geometry via the concept of infinitely small. Points and infinitesimals also have properties in common. Just as a point has no length, an infinitesimal is nothing, i.e. it is 0. The infinitesimal too has no parts, just as the point
and the material body. This is why analytical expressions can be considered representations of geometrical objects. The reliance on geometry as an intermediary between geometry and calculus was completely removed by Lagrange's work in mechanics and the calculus. This will be the topic of the next chapter.
Chapter 5

Foundational Research in the 18th Century: Lagrange

Euler’s *Mechanica* marks a turn to analysis from geometry, but it does not go all the way in establishing mechanics as a full blown analytical domain. The solutions to the problems presented in the *Mecanica* still depended on geometrical representations. Things are quite different with the *Méchanique Analitique*. In the preface to the second edition Lagrange says

> No figures will be found in this work. The methods that I will present do not require constructions, nor geometrical or mechanical justifications, but only algebraic operations.¹

The differences between the two treatises go further than just their relation with geometry. First, Euler and Lagrange posit different principles and concepts at the foundations of mechanics: Euler begins with a distinction between internal and external principles derived from the essence of material bodies, while Lagrangian mechanics is founded upon the principle of virtual work. A second difference regarding the two mechanics is the different emphasis they

---

¹“On ne trouvera point de Figures dans cet Ouvrage. Les méthodes que j’y expose ne demandent ni constructions, ni raisonnemens géométriques ou mécaniques, mais seulement des opérations algébriques” [Lagrange, 1853, p. III].
place on theoretical concepts, such as force and vis viva. While Euler takes “force” to be a theoretical primitive, Lagrange does not try to establish forces as mechanical concepts, and the same applies for other theoretical concepts as well, e.g. mass, vis viva. On a second level there are the differences between the foundations of the calculus in which Eulerian and Lagrangian mechanics are couched. One such example is Lagrange’s rejection of infinitesimals and his attempt to define derivatives using Taylor’s theorem.

Obviously, there are also similarities between the two views on mechanics and the calculus. Euler and Lagrange share with Jean le Rond d’Alembert and Pierre Louis de Maupertuis the Cartesian “belief in a mathematically structured reality and in the capability of the human mind to condense this reality into a deductive symbolical structure with only a few first, indisputable principles.”[Pulte, 1998, p. 156] Moreover, despite the differences mentioned above, both Euler and Lagrange emphasize the algebraic character of the calculus and its importance in the development of algorithms. Needless to say, they also share the belief that calculus must be separated from geometry and mechanics, although it is being applied to solve classes of problems from these domains.

Lagrange’s work on the foundations of the calculus is divided in two distinct components presented in Lagrange’s letters to Euler from 1755-1756, and in the Théorie des fonctions analytiques and Leçons sur le calcul des fonctions [Fraser, 1985]. A similar, but unrelated, change of mind is present in the foundations of mechanics, as Lagrange strips the principle of least action (\(\text{PLA}\)) of the fundamental role attributed to it in [Lagrange, 1867a], and places the principle of virtual velocities (\(\text{PVV}\)), later referred to as the principle of virtual work (\(\text{PVW}\)), as the basis of mechanics in [Lagrange, 1853].
5.1 Lagrange’s Criticism of His Predecessors

In the first chapter of his *Théorie des fonctions analytiques* Lagrange makes a definitive critique of previous attempts to provide the calculus with sound foundations. It must be noted that his objections are not original, some of them having been presented in Bishop Berkeley’s *The Analyst*, and most of them have been mentioned in previous chapters. Thus, the way in which the “geometers” [Lagrange, 1797] Leibniz, the Bernoullis and l’Hôpital employed the concept of infinitesimal made their calculations inaccurate. In their works, infinitely small quantities were quantities after all, and, as such, they could increment or decrement another quantity. They considered that the results of the calculus could be considered accurate only if the errors introduced by these small variations cancel out. However, Lagrange notes that it would be difficult to demonstrate that the errors are always compensated:

\[\ldots\text{it also seems to me\ldots that the genuine metaphysics of this calculus consists in that the error resulting from this false supposition [infinitesimals] is rectified or compensated by the error generated by the procedures of the calculus.}^{2}\]

Also, in his method of fluxions, Newton tried to avoid the assumption of infinitely small quantities:

\[\text{In order to avoid assuming infinitely small quantities, Newton considered mathematical quantities as generated by movement and he searched for a method to determine directly the velocities, or rather the ratio of the variable velocities with which these quantities are generated\ldots This method, or this calculus, agrees with the differential calculus in content and in what regards the basic operations, and differs only in its metaphysics which seems to be clearer, since everybody thought to have an idea of velocity}^{3}.\]

---

\[^{2}\ldots\text{d’ailleurs il me semble que\ldots la véritable métaphysique de ce calcul consiste en ce que l’erreur résultant de cette fausse supposition est redressée ou compensée par celle qui naît des procédés mêmes du calcul\ldots}^{2}\]

\[^{3}\text{\textquoteleft\textquoteleft Newton, pour éviter la supposition des infiniment petits, a considéré les quantités mathématiques comme engendrées par le mouvement, et il a cherché une méthode pour déterminer directement les vitesses variables avec lesquelles ces quantités sont produites \ldots Cette méthode ou ce calcul s’accorde pour le fond et pour les opérations, avec le calcul différentiel, et n’en diffère que par la métaphysique qui paraît en effet plus claire, parce que tout le monde a ou croit avoir une idée de la vitesse.	extquoteright}\textquoteright}^{3}[^{4}][\text{Lagrange, 1797, §5}]\]
However, for Lagrange the problem with the method of fluxions is that while we might have an *intuitive* idea of velocity, we do not have an idea of *instantaneous* velocity, which is what the fluxion actually is. Moreover, insofar as the fluxion is interpreted as a velocity it is essentially a physical concept and as such cannot be included among the foundational concepts of a mathematical discipline such as the calculus. For this reason, as noted in 3.2.1, Newton introduced the method of ultimate ratios.

Apart from their name, there is no real difference between ultimate ratios and the 18th-century concept of limit as Lagrange and Carnot [Carnot and Browell, 1832, §78] observed. So Lagrange’s objection applies equally to both. To be more precise, the notion of limit as introduced by d’Alembert in 1754 in his *Encyclopédie* and used in the 18th century is

One quantity is the limit of another if the second can approach the first nearer than by a given quantity, so that the difference between them is absolutely inassignable. [Boyer, 1949, p. 247]

However, when applying this concept we are considering “quantities in the state in which they cease, so to speak, to be quantities”,⁴ because they become nothing (i.e. zero). Since both terms of the ratio become simultaneously zero, the ratio will be equal to \( \frac{0}{0} \) which “no longer offers a clear and precise idea to the mind.”⁵

Lagrange’s historical introductions to the calculus in the *Théorie des fonctions analytiques* and to mechanics in *Mécanique Analytique* is an uncommon occurrence in the 18th century. Also peculiar is his changing his views on the foundations of the calculus and mechanics.

### 5.2 Foundations of Lagrangian Calculus

Although the presentation of Eulerian foundations of calculus began with an analysis of the concept of function, a similar strategy would be redundant here, since Lagrange uses a similar

---

⁴ “... de considérer les quantités dans l’état où elles cessent, pour ainsi dire, d’être quantités...”[Lagrange, 1797, §5]

⁵ “... n’offre plus à l’esprit une idée claire et précise...”[Lagrange, 1797, §5]
concept. For him, a function is any expression into which constants and variable quantities enter [Lagrange, 1806]. Although Lagrange did not detail the definition any further, he attributes the same properties to functions as Euler, i.e. they are continuous, differentiable and expandable into infinite series. However, functions had a more important role for Lagrange than they had for Euler. Recall that for Euler geometric ratios of infinitely small increments constitute the object of study of the differential calculus. For Lagrange, on the other hand, it is functions themselves who have this role. The calculus is a study of functions, not of geometric ratios:

But in algebra, we consider functions insofar as they result from the operations of arithmetic, generalized and translated to symbols, whereas in the calculus of functions per se, we consider the functions that result from the algebraical operation of series expansion, when we attribute an indeterminate increment to one or more quantities of the function.6

To get a better picture of Lagrange’s suggestion, we should look at his analysis of the increment of a general function. I will use as a template the second lesson [Lagrange, 1806], although the same procedure is also used in Théorie des fonctions analytiques.

Consider a function $f(x)$ in which we substitute $x$ for $x + i$, where $i$ is an indeterminate quantity. The function will become $f(x + i)$ which expanded into series will be

$$f(x + i) = f(x) + pi + qi^2 + ri^3 + \ldots$$

where $p$, $q$ and $r$ are functions of $x$ “derived from the primitive function $f x$, and independent from the quantity $i$.”7 and whose analytical expression depends on the expression of

---

6“Mais, en algèbre, on ne considère les fonctions qu’autant qu’elles résultent des opérations de l’arithmétique, généralisées et transportées aux letters, au lieu que, dans le calcul des fonctions proprement dit, on considère les fonctions qui résultent de l’opération algébrique du développement en série, lorsqu’on attribue à une ou à plusieurs quantités de la fonctions, des accroissements indéterminés.” [Lagrange, 1806, p. 4]

7“des dérivées de la fonction primitive de $f x$ et indépendantes de la quantité $i$.”[Lagrange, 1806, p. 7]
In order to secure the status of derivative functions as substitutes of differentials, without appealing to infinitesimals or geometrical concepts, Lagrange must show that derivative functions exist for any primitive function whatsoever.

However, as observed in the previous chapter, Eulerian functions focused on functional relations between variables, where the latter were understood to be any quantity whatsoever, including those quantities for which the value of the function cannot be calculated due to the definition of the algebraic operations applied to the functional variables\(^8\) or because the function is expressed as a solution to an equation, or of an integral that cannot be calculated.\(^9\) These cases were accepted as mere peculiarities and were not given too much thought at the time. Similarly, when we say that all functions can be expanded into infinite series, we must be aware that there will be exceptional cases in which the algebraic operation of expanding into series cannot be applied, just as other algebraic operations admit exceptions. Lagrange does not seem to consider this an issue:

We are thus assured that since \(f(x)\) expresses any algebraic function of \(x\), the function \(f(x + i)\) can, generally speaking, be developed into a series of this form:

\[
fx + ip + i^2 q + i^3 r + i^4 s + \&c
\]

in which \(p, q, r, \&c\), will be new functions of \(x\) derived from the primitive function \(fx\).

If the function \(fx\) is not algebraic, we can nevertheless assume that the development of \(f(x + i)\) has in general the same form, considering as particular exceptions those cases in which this development contains powers of \(i\) other than positive integers\(^{10}\)

\(^8\)Consider \(f(x) = \frac{1}{1-x}\) for \(x = 1\). The value of \(f(x)\) at \(x = 1\) cannot be calculated due to the definition of the operation of division.

\(^9\)These transcendental functions, as Euler called them, were discussed in the previous chapter.

\(^{10}\)Nous sommes donc assurés que \(f(x)\) exprimant une fonction quelconque de \(x\) algébrique, la fonction \(f(x + i)\) peut, généralement parlant, se développer en une série de cetter forme,

\[
fx + ip + i^2 q + i^3 r + i^4 s + \&c
\]

dans laquelle, \(p, q, r, \&c\), seront de nouvelles fonctions de \(x\) dérivées de la fonction primitive \(fx\).

Si la fonction \(fx\) n’est pas algébrique, on peut néanmoins supposer que le développement de \(f(x + i)\) soit
Therefore, Lagrange’s aim is to show that in general algebraic functions can be represented as in (5.1) “and he wishes to construct the demonstration so as to provide a plausible account within his algebraic framework for the possibility of exceptional values” [Fraser, 1987].

To this end, he begins by proving that in the series resulting from the expansion of an algebraic function \( f(x + i) \) there are no fractional or negative powers of \( i \) unless we give \( x \) particular values. For instance, if there is a negative power of \( i \) in the formula (5.1), then it is possible to find \( \frac{a}{i^n} \) \((n > 0; n \in \mathbb{N}^*)\) among the terms of the series. However, this implies that for \( i = 0 \), \( \frac{a}{i^n} \) will be infinite; this means that \( f(x + i) = f(x) + ip + i^2q + i^3r + i^4s + &c + \infty \) will be infinite too. For Lagrange, this is possible only for particular values of the variable \( x \) so the validity of his claim is not endangered. Of course, by modern standards, this proof is tentative at best, but recall that following Euler’s lead, 18th-century mathematicians had a different understanding of variables than we do today.

As explained in section 4.1.1, the relation between variables and particular values was similar to the one between genus and species. It follows that when trying to determine whether a function \( f \) has a property \( P \), it does not matter whether there are particular values for which \( f \) does not have \( P \), as long as the formal relation holds between indeterminate values.\(^\text{11}\)

Note that the expansion of \( f(x + i) \) has the form (5.1) only if \( f(x) \) is algebraic. Lagrange explicitly restricts his argument to algebraic functions, admitting that for transcendental

---

\(^\text{11}\)In 4.1.1 I cited Ferraro [Ferraro, 2000] to illustrate this point. I will repeat that citation here: en général de la même forme, en regardant comme des exceptions particulières les cas où ce développement contiendrait d’autres puissances de \( i \) que des puissances positives et entières. [Lagrange, 1806, p. 9] What is legitimate for the variable could not be legitimate for all its occasional values. Consequently, given any property \( P \) of \( x \), there might exist exceptional values at which the property fails. A proof involving the variables \( x, y, \ldots \) was valid and rigorous as long as the variables \( x, y, \ldots \) remained indeterminate; but this was no longer the case if one gave a determinate value to \( x, y, \ldots \) Thus, if one expanded \( f(x) \) into a power series \( \sum_{n=1}^{\infty} a_n x^n \) and made no assumptions concerning the individual values of variables, then the equality \( \sum_{n=1}^{\infty} a_n x^n = f(x) \) was considered globally valid even if there might exist certain occasional values at which the general relation \( \sum_{n=1}^{\infty} a_n x^n = f(x) \) did not furnish a numerical equality: these points were “not significant” [Ferraro, 2000, p. 119].
functions it is not possible to know whether the expansion of \( f(x + i) \), where \( f(x) \) is transcendental, has the same form as (5.1). He does show, however, that particular transcendental functions can be expanded into series such as (5.1), but does not generalize this result\(^\text{12}\).

After showing that \emph{algebraic} functions can generally be expressed as (5.1) Lagrange tries to show that although the forms of the derivative functions depend on the primitive, they are, nevertheless, governed by a “general law”\cite{Lagrange1806}. The law is obtained by a series of substitutions and expansions into series. First, if we replace \( x \) with \( x + o \) in \( f(x + i) \) we obtain \( f(x + i + o) \). Second, replacing \( i \) with \( i + o \) in the same function, we also obtain \( f(x + i + o) \). From the subsequent equality we will obtain the relation between derivative functions:

\[
f(x + i + o) = f(x) + p(i + o) + q(i + o)^2 + r(i + o)^3 + \ldots	ag{5.2}\]

By expanding \((i + o)\) and retaining only the first terms we obtain

\[
f(x + i + o) = f(x) + pi + qi^2 + ri^3 + si^4 + \cdots + op + 2ioq + 3i^2or + 4i^3os + \ldots	ag{5.3}\]

Replacing \( x \) with \( x + o \) in \( f(x) \) gives

\[
f(x + o) = f(x) + op + o^2q + o^3r + \ldots	ag{5.4}\]

Since \( p, q, r, \ldots \) are also functions of \( x \), it follows that replacing \( x \) with \( x + o \) will give

\[
p(x + o) = p(x) + op' + o^2p'' + \ldots
\]

\[
q(x + o) = q(x) + oq' + o^2q'' + \ldots
\]

\(^{12}\)Thus, exponentials and logarithms are discussed in the fourth Lécon and sine and cosine in the fifth.
Then, \( f(x + i + o) \) will be

\[
f(x + i + o) = f(x) + pi + qi^2 + ri^3 + \cdots + i(p + p' o + p'' o^2 + \cdots) + i^2(q(x) + oq' + o^2q'' + \cdots) + \cdots
\]  

(5.5)

Now, (5.3) and (5.5) are equal to \( f(x + i + o) \). If the two expressions are equal, then the coefficients of the products of the same powers of \( i \) and of \( o \) should be equal both in (5.3) and in (5.5). This means that \( 2q = p' \), \( 3r = q' \), etc., which means that

\[
q = \frac{1}{2}p' \quad r = \frac{1}{3}q' \quad s = \frac{1}{4}r' \quad \ldots
\]

(5.6)

At this point Lagrange introduces the well-known notation of derivatives: \( f'(x) = p(x) \), \( f'' = p'(x) \) and \( q = \frac{1}{2} f'' \), \ldots. Replacing in (5.6) we obtain the well-known formula

\[
f(x + i) = f(x) + pi + \frac{i^2}{2} p' + \frac{i^3}{3} q' + \cdots
\]  

(5.7)

This is the general law expressing the relation between the primitive and the derivative functions.

To illustrate this procedure, consider the third Leçons, in which Lagrange applies the law to find the derivative of any power function \( f(x) = x^m \). For \( x + i \) we have \( f(x + i) = (x + i)^m = x^m(1 + \frac{i}{x})^m = x^m(1 + \omega)^m \). The result of expanding \( (1 + \omega)^m \) is \( 1 + \omega F(m) + \ldots \). For another exponent \( n \), the result of expanding is \( 1 + \omega F(n) + \ldots \). Multiplying the two
functions we get

$$(1 + \omega)^{m+n} = 1 + \omega(F(m) + F(n)) + \omega^2 F(m)F(n) + \ldots \quad (5.8)$$

But at the same time the expansion of $(1 + \omega)^{m+n}$ is $1 + \omega(F(m + n)) + \ldots$, which means that $F(m + n) = F(m) + F(n)$:

\[
\begin{align*}
F(m + n) &= F(m) + F(n) \\
F(m + i + n - i) &= F(m + i) + F(n - i) \\
F(m + i) &= F(m) + iF'(m) + \frac{i^2}{2} F''(m) + \ldots \\
F(n - i) &= F(n) - iF'(n) + \frac{i^2}{2} F''(n) + \ldots \\
F(m + n) &= F(m) + F(n) + i(F'(m) - F'(n)) + \frac{i^2}{2} (F''(m) + F''(n)) + \ldots \\
&= F(m) + F(n)
\end{align*}
\]

Therefore,

$$i(F'(m) - F'(n)) + \frac{i^2}{2} (F''(m) + F''(n)) + \ldots = 0 \quad (5.9)$$

which holds only if $F'(m) - F'(n) = 0$, $F''(m) + F''(n) = 0$, etc. This can only be the case if $F'(m) = F'(n) = a$ and $F''(m) = F''(n) = 0$. Given the relation (5.7), it follows that higher-order derivatives will also be 0. Returning to the expansion of $F(m + i)$,

\[
\begin{align*}
F(m + i) &= F(m) + F'(m) = F(m) + ai \\
F(m + (-m)) &= F(0) = F(m) - ma \\
F(m) &= F(0) + ma = b + ma
\end{align*}
\]

But $(1 + \omega)^m = 1 + \omega F(m) + \ldots$, that is $(1 + \omega)^m = 1 + \omega(a m + b) + \ldots$. We can find the values of $a$ and $b$ for special values of $m$. Thus, for $m = 0$: $1 = 1 + \omega b + \ldots$. Since all terms
after 1 contain $\omega$ but no $\omega$ appears on the left hand of the equality, it follows that necessarily $b = 0$. If $m = 1$ then $1 + \omega = 1 + a\omega + \ldots$, which means that $a = 1$. We conclude that $(1+\omega)^m = 1 + m\omega + \ldots$. Since $x^m(1+i/x)^m = x^m(1+\omega)^m = (x+i)^m$, it follows that $(x+i)^m = x^m(1 + m\frac{i}{x} + \ldots) = x^m + mx^{m-1} + \ldots$. However, $(x+i)^m = (x^m)' + i(x^m)' + \frac{i^2}{2}(x^m)'' + \ldots$, therefore, the derivative of the function $f(x) = x^m$ is $f'(x) = mx^{m-1}$.

After finding this result, Lagrange gives "a demonstration both simple and general, and perhaps the only rigorous one, given as yet to the binomial formula for a variable exponent."\(^{13}\)

The demonstration is simple. If $f'(x) = mx^{m-1}$, then taking $f'$ as primitive and deriving again $f'' = m(m-1)x^{m-2}$. Replacing the derivatives in (5.7) gives the general case:

\[(x + i)^m = x^m + mx^{m-1}i + \frac{m(m-1)}{2}x^{m-2}i^2 + \frac{m(m-1)(m-2)}{2 \cdot 3}i^3 + \ldots \quad (5.10)\]

For $x = a$ and $i = b$ we obtain Newton’s binomial:

\[(a + b)^m = a^m + ma^{m-1}b + \frac{m(m-1)}{2}a^{m-2}b^2 + \frac{m(m-1)(m-2)}{2 \cdot 3}a^{m-3}b^3 + \ldots \quad (5.11)\]

Although Lagrange views the proof of the binomial theorem a consequence of (5.7) it might be tempting to view it as a verification of the validity of (5.7). This was a common means of validating the mathematical procedures during the 18th century. The solutions to geometrical problems obtained through the analytical methods of the calculus were compared with the results obtained through purely geometrical methods and from this kind of successes the validity of the calculus, regardless of its foundations. However, under this interpretation Lagrange’s argument is open to one of Berkeley’s objections:

\(^{13}\)... une démonstration aussi simple que générale, et peut être la seule rigoureuse qu’on ait encore donnée de la formule du binôme pour un exposant quelconque. [Lagrange, 1806, p. 17]
And yet it should seem that, whatever errors are admitted in the Premises, proportional errors ought to be apprehended in the Conclusion, be they finite or infinitesimal: and that therefore the ακριβεία of Geometry requires nothing should be neglected or rejected. In answer to this you will perhaps say, that the Conclusions are accurately true, and that therefore the Principles and Methods from whence they are derived must be so too. But this inverted way of demonstrating your Principles by your Conclusions, as it would be peculiar to you Gentlemen, so it is contrary to the Rules of Logic. The truth of the Conclusion will not prove either the Form or the Matter of a Syllogism to be true. [Berkeley, 2009, §19]

Lagrange’s deduction of the general law of derivation shows that for him the relation between the primitive and its derivative is a formal algebraic relation. This way, Lagrange avoids the complications raised by infinitesimals and the appeal to geometry. Moreover, this method of derivation is completely different from the modern, limit-based approach.

However, Lagrange’s law is too separated from geometry [Grabiner, 2005, p. 36]. He defines the derivative of a function as the first coefficient of $i$ in the expansion of the function in Taylor series, but does not explain how this concept applies to solving problems of geometry (such as finding the rate of change of a function). Furthermore, Lagrange did not prove that the series obtained from the expansion of a function is unique, i.e. that there are no other functions whose corresponding series are identical to the series of the function under investigation. Cauchy rightly pointed out that the functions $e^{x^2}$ and $e^{-x^2} + e^{-1/x^2}$ have the same Taylor series around $x = 0$, i.e. $1 - \frac{x^2}{1} + \frac{x^3}{1 \cdot 2} - \frac{x^4}{1 \cdot 2 \cdot 3} + \ldots$ [Cauchy, 1974, p. 277].

It might be thought that Lagrange’s method resembles Euler’s foundations of the calculus in that it uses finite increments. However, in the eighteenth Leçon Lagrange explicitly distinguishes his own foundations from those relying on finite increments. For Lagrange, founding the differential calculus upon the calculus of finite differences can only be done through a process of passing to the limit and, as noted previously, he already rejected the validity of such an approach. Regarding the difference between the calculus of finite differences and Lagrange’s own calculus of derivatives, is that the former represents the terms of the progression
by functions of different quantities, while the latter represents the same terms by different functions of the same quantity. As seen in the previous chapter, for an increment $\omega$ of $x$ there corresponds a second value ($y^I$) of a function $y$. In a similar manner we calculate the rest of the series of values of $y(y^{II}, y^{III}, \ldots)$. From here we compute the differences $\Delta y = y^I - y$. The difference is between the different values of the same function. By contrast, in Lagrange’s calculus the derivative function, which corresponds to $\frac{\Delta y}{\Delta x}$ for $\Delta x \rightarrow 0$, is a function in its own right, obtained from the primitive function or from a derivative function of a lower order following “fixed and uniform rules” [Lagrange, 1806, p. 292]:

The equations in finite differences are nothing more but a series of similar equations between different unknown quantities, by which we can always successively determine each of these quantities.

But the uniform law which governs these equations allow us to consider these unknown quantities as forming a regular series which may have a general term, and the expressions of this term gives the general resolution of all equations.14

The second problem with founding CI upon the CFD is that in moving from the latter to the former we are inevitably assuming the problematic concept of infinitely small quantities. Although the problems raised by this concept are provisionally solved by expressing the ratio of infinitesimals as the limit of a ratio of finite differences, in doing so we are introducing the problematic concept of limit. For Lagrange, “passing from the finite to the infinite always demands a kind of leap, more or less forced, which breaks the law of continuity and changes the form of the functions.”15

In the *Théorie des fonctions analytiques* Lagrange extends the method of derivation to functions of multiple variables. The procedure is simple. For a function $f(x, y)$ of two

---

14"Les équations aux différences finies ne sont autre chose qu’une suite d’équations semblables entre différentes inconnues, par lesquelles on peut toujours déterminer successivement chacune de ces inconnues.

Mais la loi uniforme qui règne entre ces équations fait qu’on peut regarder leurs inconnues comme formant une suite régulière et susceptible d’un terme général, et l’expression de ce terme donne alors la résolution générale de toutes les équations." [Lagrange, 1806, p. 292]

15"le passage du fini à l’infini exige toujours une espèce de saut, plus ou moins forcé, qui rompt la loi de continuité et change la forme des fonctions." [Lagrange, 1806, p. 293]
variables, first we replace $x$ with $x + i$ then expand the function. After this, we replace $y$ with $y + o$ in the expanded function and perform the appropriate algebraic operations.

\[
\begin{align*}
    f(x + i, y) &= f(x, y) + if'(x, y) + \frac{i^2}{2} f''(x, y) + \frac{i^3}{2 \cdot 3} f'''(x, y) + \ldots \\
    f(x + i, y + o) &= f(x, y + o) + if'(x, y + o) + \frac{i^2}{2} f''(x, y + o) + \frac{i^3}{2 \cdot 3} f'''(x, y + o) + \ldots \\
    f(x, y + o) &= f(x, y) + of'_x(x, y) + \frac{o^2}{2} f''_x(x, y) + \frac{o^3}{2 \cdot 3} f'''_x(x, y) + \ldots \\
    f'_x(x, y + o) &= f'_x(x, y) + of''_x(x, y) + \frac{o^2}{2} f'''_x(x, y) + \ldots \\
    f''_x(x, y + o) &= f''_x(x, y) + of'''_x(x, y) + \frac{o^2}{2} f''''_x(x, y) + \ldots \\
    \ldots \end{align*}
\]

In the above deduction of the expansion of a function with two variables, $f'$, $f''$, $f'''$, etc. are the derivative functions relative to the variable $x$, while $f$, $f_m$, $f_m$, etc. are the derivatives relative to $y$. Likewise, $f''_x$ is the first derivative relative to $x$ of the second derivative relative to $y$, and so on. The general term of the expansion is

\[
\frac{i^n o^n}{(1 \cdot 2 \cdot 3 \cdot 4 \ldots m)(1 \cdot 2 \cdot 3 \cdot 4 \ldots n)} f_m^n(x, y)
\]  

This small digression into functions of several variables is required if we are to understand probably the most important discovery of 18th-century calculus: the calculus of variations.
5.3 Euler and Lagrange on the Calculus of Variations

Simply put, the central problem of the calculus of variations is to find the function \( y = y(x) \) from the members of a class of functions such that a given definite integral has an extremum. Under this heading fall a series of problems addressed in the 18th century:

1. Brachistochrone problems
2. Problems of geodesics
3. Isoperimetric problems

The first brachistochrone problem appears in the *Acta Eruditorum* of June 1696 [Kline, 1990] as a challenge to mathematicians raised by John Bernoulli. “The problem is to determine the path down which a particle will slide from one given point to another not directly below in the shortest time.”[Kline, 1990, p. 574] If \( v_1 \) is the initial velocity at point \( P_1 \) and friction is ignored, then the brachistochrone problem is to minimize the integral \( I \)

\[
I = \frac{1}{\sqrt{sg}} \int_{x_1}^{x_2} \sqrt{1 + \left[ y'(x) \right]^2} \, dx
\]

representing the time of descent, \( g \) is the gravitational acceleration and \( \alpha = y_1 - \frac{v_1^2}{2g} \). The solution requires finding the function \( y(x) \) for which \( I \) is a minimum. The problem was solved independently by Newton, Leibniz, L’Hôpital and John Bernoulli.

Problems of geodesics require finding the shortest path between two points on a surface. In the simple case of a planar surface the integral is

\[
I = \int_{x_1}^{x_2} \sqrt{1 + [y'(x)]^2} \, dx
\]
Obviously, the solution will be a straight line. However, 18th-century mathematicians were more interested in the shortest paths on curved surfaces, having in mind the practical applications to navigation. This research was impeded by the lack of knowledge of the curvature of the earth. *Isoperimetric problems* form a subclass of the problems of geodesics, which require finding the functions \( x = x(t) \) and \( y = y(t) \), where \( t_1 \leq t \leq t_2 \), such that

\[
L = \int_{t_1}^{t_2} \sqrt{(x')^2 + (y')^2} dt
\]

is a given constant and the area integral:

\[
J = \int_{t_1}^{t_2} (xy' - x'y) dt
\]

Although extremum problems also appear in the works of Galileo and Newton, the calculus of variations was systematized in collaboration, by Euler and Lagrange. Euler was first introduced to the brachistochrone problem in 1728 when John Bernoulli proposed “the problem of obtaining geodesics on surfaces by using the property that the osculating planes of geodesics cut the surface at right angles” [Kline, 1990, p. 577]. He found a solution the same year and generalized it in 1734 to include resisting media and minimize quantities other than time.

After this, he generalized the results and in 1744 he published the *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes* in which he presented a collection of methods for solving variational problems. As reported in [Fraser, 1985], these methods had a strong geometrical character and were too complex to be applied effectively. A complete analysis of Euler’s *Methodus* would lead us away from our present purposes. Instead, following [Fraser, 1985], I will focus on the one problem presented in the *Methodus* that influenced
Lagrange’s own work on calculus of variations.

5.3.1 Euler

The problem is presented in the third proposition of the Methodus:

If $Z$ is a function of $x$, $y$, and $p$ is determined, so that $dZ = Mdx + Ndy + Pdp$, discover, among all curves with the same abscissa, the one that makes $\int Zdx$ a maximum or a minimum.\(^{16}\)

Let us illustrate this graphically. In 5.1 the relation($Z$) between the abscissa $x$ and the ordinate $y$ is represented by the curve $az$. The segments $MN$ and $NO$ are considered infinitely small. Correspondingly, $m$, $n$ and $o$ are the points on the curve $az$ corresponding to $M$, $N$ and $O$ respectively. Let $AM = x$, $AN = x'$, $AO = x''$, $Mm = y'$, $Nn = y''$ and $Oo = y'''$.

\(^{16}\) “Si $Z$ fuerit functio ipsarum $x$, $y$, & $p$ determinata, ita ut sit $dZ = Mdx + Ndy + Pdp$; invenire, inter omnes curvas eidem abscisse respondentes, eam in qua sit $\int Zdx$ maximum vel minimum.”[Euler, 1743, p. 42]
relations:

\[ p = \frac{y' - y}{dx}, \quad p' = \frac{y'' - y'}{dx} \]

Keep in mind that since \( AN \) is infinitesimal and \( AM = y' - y \) it follows that the difference \( y' - y \) will also be infinitely small. The same notation convention applies to the function \( Z \), i.e. \( Z \) is the value of the function at \( x, y, p \), \( Z' \) the value at \( x', y', y' \) and \( Z'' \) the value at \( x'', y'', p'' \). On the interval \( AZ \) the integral \( \int Z \, dx \) will be \( \int (Z + Z' + Z'' + \ldots) \, dx \), since the abscissa \( x \) takes on all values from \( A \) to \( Z \), and to each infinitely small increment of \( x \) there corresponds a variation of \( y \) and consequently of \( p \) and \( Z \).

Euler’s next step is to move the point \( n \) to \( v \), thereby incrementing the ordinate \( y' \).\(^{17}\) If \( \int Z \, dx \) is a maximum or a minimum, then incrementing the ordinate \( y' \) will not bring about any change of the integral, otherwise \( \int Z \, dx \) would not be an extremum. Euler then observes that only the values \( Z \) and \( Z' \) are affected by the variation of \( y' \). Euler calculates the corresponding variations of \( Z \) and \( Z' \):

\[
\begin{align*}
    dZ &= M \, dx + N \, dy + P \, dp \\
    dZ' &= M' \, dx + N' \, dy' + P' \, dp'
\end{align*}
\]

However, we already know that \( p = \frac{y' - y}{dx} \) and \( p' = \frac{y'' - y'}{dx} \). It follows that \( p = \frac{nv}{dx} \) and \( p' = -\frac{nv}{dx} \).

The relations in (5.13) become:

\[
\begin{align*}
    dZ &= M \, dx + N \, dy + P \frac{nv}{dx} \\
    dZ' &= M' \, dx + N' \, dy' - P' \frac{nv}{dx}
\end{align*}
\]

\(^{17}\)“Totius igitur quantitatis \( \int Z \, dx \) valor differentialis ex translatione puncti \( n \) in \( v \) habebitur, si singulorum illorum terminorum, qui quidem hac translatione afficiuntur, valores differentiales quaerantur & in unam sumam addantur.”[Euler, 1743, p. 43]
However, if \( y' \) is increased by \( nv \), \( dx = 0 \) and \( dy = 0 \), since these infinitely small quantities are not affected. It follows that

\[
\begin{align*}
    dZ &= P_{\frac{nv}{dx}} \\
    dZ' &= N' \cdot nv - P'_{\frac{nv}{dx}}
\end{align*}
\]  

(5.15)

Since when \( y \) is increased by \( nv \) the change of \( \int Zdx \) is zero (otherwise the integral would not be an extremum) and the corresponding variation of \( Z \) is \( dZ + dZ' \), it follows that

\[
\begin{align*}
    dZ + dZ' &= 0 \\
    dZ + dZ' &= P_{\frac{nv}{dx}} + (N' \cdot nv - P'_{\frac{nv}{dx}}) \\
    (dZ + dZ')dx &= Pnv + dxN' \cdot nv - P'nv \\
    (dZ + dZ')dx &= nv \cdot (P + N'dx - P') = 0 \\
    P' - P &= dP \\
    (dZ + dZ')dx &= nv \cdot (N'dx - dP) = 0 \\
    N' - \frac{dP}{dx} &= 0
\end{align*}
\]

In the last equation Euler replaces \( N' \) with \( N \) and obtains his formulation of the Euler-Lagrange equations for the calculus of variations.

\[
    N - \frac{dP}{dx} = 0
\]  

(5.16)

(5.14) is obtained from (5.15) by taking into consideration the values of those infinitesimals which were not affected by the original variation, i.e. \( dx = 0 \), \( dy = 0 \). There seems to be a
contradiction here with Euler’s general view of infinitesimals. In chapter 4 we saw that for Euler infinitely small quantities were ultimately nothing, yet now he considers that some of them could be different from 0, i.e. \( dp = \frac{n v}{dx}, dy' = -\frac{n v}{dx}, dy' = n v. \)

These values of infinitesimals in this case should be understood from a modal perspective, i.e. the values of \( dp, dp', dy' \) are the values the displacements would have if the point \( n \) would move to \( v \). To put matters into perspective, consider the overall problem of finding the function \( Z \) such that \( \int Z dx \) is an extremum. If \( n \) would be moved to \( v \) then \( Z \) would change by \( dZ + dZ' \). This variation, however, has to be null, and from here we deduce (5.16).

The modal interpretation of infinitesimals will be more visible in section 5.4 where instead of considering counterfactual variations of geometrical curves, we will focus on the application of infinitesimals to mechanics. However, before tackling this problem, I will present Lagrange’s calculus of variations since this is the version used in his *Méchanique Analitique*.

### 5.3.2 Lagrange

The problem with Euler’s calculus of variation was that it was overtly complex and difficult to apply to specific problems. In a series of letters to Euler, Lagrange systematized and introduced simpler methods for the calculus of variations. However, after the publication of the *Méchanique Analitique* Lagrange’s approach to the calculus of variations changed, as a result of his turn towards a more algebraic differential calculus.

In a letter to Euler sent on August 12th 1755 Lagrange presents his reply to Euler’s *Methodus Inveniendi* and introduces the \( \delta \) operator to the infinitesimal calculus. The definition of the operator is operational. Thus, \( x \) is constant relative to \( \delta \), i.e. in \( \delta x = 0 \), \( \delta y \) represents the differential of \( y \) but is applicable only to problems of maxima and minima to distinguish it from the differential \( dy \) appearing in these problems. \( \delta F y \) represents the variation of the
function $F$ corresponding to an increment of $\delta y$ of $y$. Lagrange postulates that $d\delta F y = \delta dF y$

\[ d^m \delta F y = \delta d^m F y \quad (5.17) \]

The reason I call (5.17) a postulate is that Lagrange justifies it by invoking one of Euler’s earlier works [Euler, 1740]. However, in [Euler, 1740] Euler tries to find a method for finding the equations of curves of the same kind:

Here I call curves of the same kinds such curves that are only different from each other because of the length of a certain constant line, which by assuming some or other values determines these curves. [Euler, 1740, §1]

Euler calls the constant line “modulus”, but modern mathematics calls it “parameter”. This parameter determines an infinite number of curves:

Thus a parameter is a constant invariable line, by which each one of an infinitude of curves is determined; moreover it has different values and thus it is a variable, if it is to refer to different curves. Thus if in the equation $y^2 = ax$, $a$ is taken for the parameter, from the variability of an innumerable parabolas can arise placed on the same axis and having a common vertical axis. [Euler, 1740, §1]

Thus, the treatise does not seem to have too much to do with Lagrange’s postulate. The only connection seems to be Euler’s differentiation of a function of two variables (in modern parlance, $\frac{\partial^2 z}{\partial a \partial x} = \frac{\partial^2 z}{\partial x \partial a}$). However, Lagrange does not present any arguments supporting any logical connection between Euler’s method of derivation and the relation (5.17).
After postulating (5.17), Lagrange lists several results of the integral calculus:

\[ \int z \, du = zu - \int udz \]  \hspace{1cm} (5.18)

\[ \int zd^2u = zdu - udz + \int ud^2z \]  \hspace{1cm} (5.19)

\[ \int zd^3 = zd^2 - dzu + ud^2 - \int ud^3z \]  \hspace{1cm} (5.20)

\[ \int u \int z = \int u \times \int z - \int z \int u \]  \hspace{1cm} (5.21)

\[ \int u \int z = \int Vz \]  \hspace{1cm} (5.22)

\[ \int u \int z = \int \delta Z \]  \hspace{1cm} (5.23)

where \( \int u = H \) and \( V = H - \int u \). Moreover, \( \int u \) is the definite integral of \( u \) from an unspecified value \( x \) to \( x = a \), \( \int z \int u \) is the integral of \( z \int u \), i.e. first we calculate the value of \( \int u \) then the integral of \( z \int u \), and \( \int u \times \int z \) is the product of the two integrals.

After introducing these theoretical considerations Lagrange attacks the first problem discussed above in connection to Euler’s Methodus Inveniendi, i.e. to find the relation between \( x \) and \( y \) that makes the function \( Z \) an extremum. The relation between \( Z \) and the variables \( x, y \), etc. is

\[ \delta Z = N \delta y + P \delta dy + Q \delta d^2y + R \delta d^3y + \ldots \]  \hspace{1cm} (5.24)

With the same ease with which he postulated (5.17) Lagrange introduces the relation

\[ \delta \int Z = \int \delta Z \]  \hspace{1cm} (5.25)
Integrating (5.24) gives:

\[ \int \delta Z = \delta \int Z = \int N \delta y + \int P \delta dy + \int Q \delta d^2 y + \int R \delta d^3 y + \ldots \]

By applying (5.17) to the previous equation thus switching the symbols \( \delta \) and \( d \) Lagrange obtains

\[ \delta \int Z = \int N \delta y + \int P d \delta y + \int Q d^2 \delta y + \int R d^3 \delta y + \ldots \]

After applying (5.18) and cancelling out some of the terms, we obtain

\[ \delta \int Z = \int (N - dP + d^2 Q - \ldots) \delta y \] \hspace{1cm} (5.26)

Since the integral has to be an extremum, then its variation \( \delta \int Z \) must be zero. So,

\[ \int (N - dP + d^2 Q - \ldots) \delta y = 0 \]

\[ N - dP + d^2 Q - \cdots = 0 \]

If \( Z \) is a function only of \( x, dx, y \) and \( dy \) we obtain Lagrange’s formulation of the Euler-Lagrange equations for the calculus of variations:

\[ N - dP = 0 \] \hspace{1cm} (5.27)

As it can be seen from the above deduction of (5.27), at this point of his career, Lagrange assumes the existence of infinitesimals, presumably interpreted in an Eulerian sense. At the same time, the deduction illustrates the emergence of Lagrange’s formalist approach to the
calculus. The operator $\delta$ is introduced as a helpful instrument, not as a symbol that stands for something. We observed this formalist tendency in the previous section, but in that case we were dealing with the foundations of the calculus in Lagrange’s late work.

It is also important to note that this letter is Lagrange’s first attempt to address the problems of the calculus. In his subsequent correspondence with Euler, Lagrange changed his view slightly, as a result of his attempts to introduce the $\delta$ operator to solve mechanical problems. In a letter to Euler written on October 5th 1756, Lagrange tried to apply what he had learned during his mechanical researches to the problem of the brachistochrone, a problem to which he returned later on in 1760 in his *Essai d’une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfinies* [Lagrange, 1867b]. More important than all these works, is the publication of the *Mécanique Analitique*, in which Lagrange applies the variational calculus as he developed it up to that point. The next sections will analyze the *Mécanique Analitique* in greater depth, for now suffice to say that the foundations of calculus presented in the previous chapter are a later development in Lagrange’s thought.

Nevertheless, these “new foundations” of the differential calculus, presented in the *Théorie des fonctions analytiques* and *Leçons sur le calcul des fonctions* lie also at the basis of the calculus of variations. For instance, in the twenty-second lesson Langrange derives the Euler-Lagrange equation, only this time he does it assuming the basis of the calculus defined by means of Taylor power series:

The method of variations, founded upon the use and the combination of the characteristics $d$ and $\delta$ which corresponded to different differentiations, leaves nothing to be desired; but this method having at its basis, like the differential calculus, the assumption of infinitesimals, it is necessary to present it from another point of view to connect it to the calculus of functions ...

18 La méthode des variations, fondée sur l’emploi et la combinaison des caractéristiques $d$ et $\delta$ qui répondent à des différenciations différentes, ne laissait rien à désirer; mais cette méthode ayant, comme le calcul différentiel, la supposition des infiniment petits pour base, il était nécessaire de la présenter sous un autre
The details of Lagrange’s deduction of the Euler-Lagrange equation from the “new foundations” are not particularly relevant for our present purposes, since the method used in the *Méchanique Analitique* is the $\delta$-algorithm.

### 5.4 The Project of Mechanics in Lagrange’s *Méchanique Analitique*

The *Méchanique Analitique* marks a change in Lagrange’s approach to the foundations of mechanics. More precisely, he strips the Principle of Least Action (PLA) from its fundamental role and constructs mechanics in a deductive manner starting with the Principle of Virtual Velocities (PVV). Lagrange constructs statics from the PVV and then uses d’Alembert’s principle (PDA) to logically connect statics with dynamics. Apart from this axiomatic structure, the *Méchanique Analitique* is also characterized by its indifference towards theoretical concepts, particularly the concept of force. The first part of the *Méchanique Analitique* begins with the definition of force and statics:

> Statics is the science of the equilibrium of forces. In general, by *force* or *power* we understand the cause, whatever that might be, which impresses or tends to impress a movement to the body to which the force is applied; and it is also by the quantity of movement impressed that the force must be estimated by.\(^{19}\)

Equilibrium is attained by the mutual “destruction” of forces and the aim of the statics is to provide the laws governing this process of mutual cancelling out. The laws themselves are derived from three principles:

1. the principle of the lever (PL)
2. the principle of the composition of forces (PCF)

---

\(^{19}\)La Statique est la science de l’équilibre des forces. On entend, en général, par *force* ou *puissance* la cause, quelle qu’elle soit, qui imprime ou tend à imprimer du mouvement au corps auquel on la suppose appliquée; et c’est aussi par la quantité du mouvement imprimé, ou prêt à imprimer, que la force ou puissance doit s’estimer. [Lagrange, 1853, §1]
3. the principle of virtual velocities (PVV)

In connection with the PL Lagrange mentions the work of Archimedes, whom he credits as the author of PL, Stevin and Galileo, for reducing Archimedes’ proof, and Huyghens. In Lagrange’s formulation, the PL consists in that “if a straight lever is loaded with two different weights placed arbitrarily on both sides of a fulcrum, at distances proportional to the weights, the lever will be in equilibrium and the fulcrum will be loaded with the sum of the two weights.” Archimedes saw the principle as an axiom of mechanics or as a principle justifiable through experience. Lagrange rejects the latter interpretation, focusing instead on the self-evident, axiomatic character of a restricted PL: “The equilibrium of a straight and horizontal lever, whose extremities are loaded with equal weights and whose fulcrum is at the middle of the lever, is a self-evident truth, because there is no reason why one of the weights would move but not the other, since the bodies are placed at equal distances from the fulcrum.”

The PCF is the parallelogram rule applied to forces: if two forces move a body uniformly following two different directions, they are equivalent to a single force moving the body uniformly following the diagonal of the parallelogram constructed by taking the directions of the two forces as adjacent sides. Lagrange credits Varignon and Stevin rather than Galileo with discovering the principle, but for good reason. Galileo discovered the principle of the composition of motions to determine the curve described by a projectile under the action of gravity. Varignon and Slevin on the other hand explicitly used the composition of forces.

On the question of justifying PCF, Lagrange turns to d’Alembert’s proof, given in the first
volume of the *Opuscules*. d’Alembert’s proof is based on two principles: (i) two equal forces acting on an object are equivalent to a single force dividing the angle between the first two in two equal angles and (ii) if the forces are multiplied by the same number, then the resultant is proportional to the factor of multiplication. Lagrange considers (ii) to be self-evident in regarding forces as quantities. As to (i), Lagrange explains it by considering the *movement* impressed to a body by two forces. Since the two forces necessarily make the body move on a *unique* trajectory, then the movement can be attributed to a single force. Moreover, the direction of this force has to be the bisector of the angle between the two forces, because, as Lagrange argues, there is no sufficient reason to believe otherwise.

Finally, the most important principle is PVV. By *virtual velocity* Lagrange understands the velocity “that a body in equilibrium is disposed to receive, in case the equilibrium would be broken, that is, the velocity that the body really gains in the first instant of its movement; and the principle of virtual velocities claims that the powers are in equilibrium when they are in inverse ratio of their virtual velocities, estimated following the directions of these forces.”22 Lagrange’s taste for the historical development of mechanics is present here too. He attributes the PVV to Guido Ubaldi who was lead to it after studying the lever and the pulleys. Galileo too reached it after studying the inclined plane and considered it a fundamental principle of mechanics. Lagrange also mentions Descartes and Wallis as positing principles similar to the PVV at the basis of their theories of statics. An important concept that Lagrange retains, and will use later on, is Galileo’s concept of *momentum*:

> By *momentum* of a weight or a force applied to a machine, Galileo understands the action, the energy, the effort, the *impetus* of this force to move the machine such that there is an equilibrium between two forces only if their momenta for moving the machine in opposite directions are equal; and it must be seen

---

22 “... celle qu’on corps en équilibre vienne à être rompu, c’est-à-dire la vitesse que ce corps prendrait réellement dans le premier instant de son mouvement; et le principe dont il s’agit consiste en ce que des puissances sont en équilibre quand elles sont en raison inverse de leurs vitesses virtuelles, estimées suivant les directions de ces puissances.” [Lagrange, 1853, §15]
that the momentum is always proportional with the force multiplied by the virtual velocity, depending upon the way in which the force acts.\textsuperscript{23}

Nevertheless, Lagrange does not consider the previous formulation of the PVV general enough to have a foundational character. Moreover, another problem with the above formulation of PVV is that it cannot be expressed mathematically. Compare it with Lagrange’s definition:

\begin{quote}
If a system of however many bodies or points we desire, each pulled by arbitrary forces, is in a state of equilibrium, and if we impress a small movement to the system, by virtue of which each point travels an infinitely small distance which will express his virtual velocity, then the sum of the products of the forces with the distances their initial points travel on the direction of the force, will always be zero, by considering as positive the small distances travelled in the sense of the forces, and as negative those distances travelled in the opposite sense.\textsuperscript{24}
\end{quote}

Lagrange himself observes that in this formulation, the principle can be translated into a general formula that can be applied to solve all problems of statics. In the second section of the first part of the *Méchanique Analitique* Lagrange constructs the formula of the PVV starting from its natural language formulation. He begins by considering a system of two forces in equilibrium. Let $P$ and $Q$ be two forces oriented along the straight lines $p$ and $q$. The infinitely small quantities $dp$ and $dq$ are the displacements along $p$ and $q$ brought about by the two forces. The virtual velocities of $P$ and $Q$ are proportional with the infinitely small displacements $dp$ $dq$ and Lagrange considers that they can be, for sake of simplicity, taken to be equal to these displacements.

\textsuperscript{23}“Galilée entend par moment d’un poids ou d’une puissance appliquée à une machine, l’effort, l’action, l’énergie, l’impetus de cette puissance pour mouvoir la machine, de manière qu’il y ait équilibre entre deux puissances, lorsque leurs moments pour mouvoir la machine en sens contraires sont égaux; et il fait voir que le moment est toujours proportionnel à la puissance multipliée par la vitesse virtuelle, dépendante de la manière dont la puissance agit.”[Lagrange, 1853, §16]

\textsuperscript{24}“Si un système quelconque de tant de corps ou points que l’on veut, tirés chacun par des puissances quelconques, est en équilibre, et qu’on donne à ce système un petit mouvement quelconque, en vertu duquel chaque point parcourt un espace infiniment petit qui exprimera sa vitesse virtuelle, la somme des puissances multipliées chacune par l’espace que le point où elle est appliquée parcourt suivant la direction de cette même puissance, sera toujours égale à zéro, en regardant comme positifs les petits espaces parcourus dans le sens des puissances, et comme négatifs les espaces parcourus dans un sens opposé.” [Lagrange, 1853, §17]
If the two forces are in equilibrium, then they must be oriented in opposite directions. Moreover, according to \( \text{pl} \) if the two forces are in equilibrium, then \( P \) and \( Q \) are in an inverse ratio with the displacements \( dp \) and \( dq \), i.e. \( \frac{P}{Q} = -\frac{dq}{dp} \), where the minus sign comes by taking into account their different orientations. Therefore,

\[
Pdp + Qdq = 0 \quad (5.28)
\]

The product \( Pdp \) is called the momentum of \( P \). After reaching the condition of equilibrium in this simple case, Lagrange considers three forces, \( P, Q, R \) on directions \( p, q, r \), causing displacements \( dp, dq \) and \( dr \), and that the three forces are in equilibrium. Lagrange applies the pcf to break down \( Q \) into \( Q' \) and \( Q'' \), where \( Q' \) and \( P \) are opposing each other, just as \( Q'' \) and \( R \) do. Applying (5.28) to the two pairs of functions Lagrange obtains the relations:

\[
Q'dq + Q'' = Qdq
\]

\[
Pdp + Q'dq = 0 \quad (5.29)
\]

\[
Q''dq + Rdr = 0 \quad (5.30)
\]

By adding (5.29) to (5.30) we obtain the pvv for the case of a system acted upon by three forces:

\[
Pdp + Qdq + Rd = 0 \quad (5.31)
\]

Following the same reasoning, the pvv can be proven for four, five, or more forces. First we break down the forces such that we can form pairs of opposing component forces, then we apply \( \text{pl} \) to obtain the relations between the forces of each pair. Finally, we add up these
formulas to obtain the final condition of equilibrium:

\[ Pdp + Qdq + Rdr + \cdots = 0 \] (5.32)

This is the “general formula of Statics for the equilibrium of any system of forces.”\textsuperscript{25}

The considerations above should not be viewed as an inductive justification from experience of the PVV, but as an illustration of (i) formalizing a principle originally expressed in natural language and (ii) the connection between the three principles, PVV, PL and PCF. Regarding (i), we must observe that in switching from the non-mathematical to the mathematical expression of PVV Lagrange assumes a particular view of infinitely small quantities. In Lagrange’s argument, if \( dp, dq, dr, \ldots \) are considered to represent \textit{actual} displacements, the system could not be considered in equilibrium, unless these displacements are viewed as zeros. Otherwise, it would mean that the forces are actually moving the system. However, if the displacements are zeros, then (5.32) would be trivial, i.e. \( 0 + 0 + 0 + \cdots = 0 \). It is also worth noticing that in this context the differentials differ from the ones examined in the previous chapter. Here \( dp, dq, dr, \) etc. are interrelated, such that when one of them varies then the magnitude of the other differentials will depend on the variation of the first. This dependency can be expressed in a more accurate form as the proportion between the differentials and the corresponding forces, i.e. \( \frac{P}{Q} = -\frac{dq}{dp} \). In Lagrange’s formulation of PVV, i.e. (5.32), this interdependency is implicit, but we should not lose track of the order of ideas. The PVV follows from \( \frac{P}{Q} = -\frac{dq}{dp} \), not the other way around. This shows that a mathematical formula can express more than its mathematical equivalents, i.e. formulas obtained from the initial one following mathematical rules.

\textsuperscript{25}“... la formule g\`{e}n\`{e}rale de la Statique pour l’\'{e}quilibre d’un syst\'{e}me quelconque de puissances.” [Lagrange, 1853, Part I, Section II §2]
represent the infinitely small displacement of a point if the system were to receive an infinitely small impulse. This allows Lagrange to apply the formal rules of the infinitesimal calculus and at the same time avoid the problems that arise by treating the differentials as zeros. More importantly, Lagrange’s mathematization of PVV shows the dependence between the foundations of the calculus and those of mechanics. The dependence is not logical; mechanics is not deduced from the calculus. Instead, the foundations of the calculus are conditions of expressibility of the foundations of mechanics. In the absence of infinitesimals, the mathematical formulation of what Lagrange considers an axiom of mechanics (PVV) would not have been possible.

Lagrange realized that the PVV was not sufficiently self-evident to be an axiom. For this reason he tried to provide an alternative justification based on what Lagrange calls the principle of pulleys (PP). Lagrange represents a system of forces \((P, Q, R, \ldots)\) acting on points of mass by a number of block and pulley systems sharing the same rope, each pulley having a mass attached. Since the points of mass and forces are all parts of the system, Lagrange can represent their interconnectedness by considering the pulleys as sharing the same rope. One end of the rope is fixed, while the other is tied to a unit weight (call it \(G\)). Moreover, the rope is looped in each block and pulley, the number of loops corresponding to the magnitude of the force it represents, i.e. if the force is 5 then there will be five loops; if the force is \(P\), then there will be \(P\) loops.

Each block and pulley represents a system composed of a force and a point of mass. The overall system of points of mass and forces is represented by all the block and pulleys connected by the same rope. This means that if one block and pulley system is in equilibrium, then, in virtue of being connected, the next block and pulley is also in equilibrium, and so on. The equilibrium of the whole system can be observed by looking at the state of the

---

26I am indebted to Hepburn [2007] for clarifying this principle.
weight $G$. Now, if we consider that one block and pulley system moves the infinitely small distance $dp$, then, the other block and pulleys will also move by $dq$, $dr$, etc. Because of this, the weight $G$ will descend by

$$Pdp + Qdq + Rdr + Sds + \ldots$$  \hspace{1cm} (5.33)

Therefore, since $G$ always tends to descend, the system will be in equilibrium only if there is no infinitely small displacement of $G$. Lagrange’s condition of equilibrium will be

$$Pdp + Qdq + Rdr + Sds + \cdots = 0$$  \hspace{1cm} (5.34)

For Lagrange the negative value of the sum of products also indicates a state of equilibrium, since it would be impossible for the weight $G$ to ascend by itself, i.e. without the intervention of an external force.

To recapitulate, Lagrange tried to construct an axiomatic mechanics that would preserve the standards of rigor of ancient geometry. To this end, in the first two sections of the *Mécanique Analytique* he tried to show the self-evident character of PVV, PL and PCF. While
the latter two principles might be considered self-evident, we cannot say the same about PVV. Lagrange tried to justify PVV and for this reason he introduced the PP, both as a proof and as an example of the mathematization of PVV.

However, Lagrange’s PP raises a number of issues. First, it might be objected that justifying the PVV by analyzing the behavior of a mechanical system is a kind of circular reasoning. To sum up this objection, remember that for Lagrange the PVV is the theoretical foundation of our understanding of the equilibrium of a mechanical system, but in order to justify the PVV Lagrange presupposes that the equilibrium of a mechanical system is already understood. This objection would be valid only if Lagrange were to present the PP as a justification of PVV. This, however, is not the case. The PP is an illustration of PVV. More precisely, the role of PP is to make the PVV more intuitive by rephrasing it in terms that are closer to experience.

There are two arguments supporting this interpretation and that would make the role of the PP more clear. First, in PP the concept of force is not mentioned and the characteristics of forces are replaced by physical objects or their attributes. For instance, the magnitude of the force is represented by the number of loops of rope around each block and pulley system, the direction of the force is replaced by the direction of the rope around the block and pulley (see fig. 5.2), and so on. This suggests that the PP should not be viewed as a principle per se, but as an illustration aimed at making PVV more intelligible, playing an expositional role similar to the one played by graphs and diagrams in science in general. Second, Lagrange repeatedly mentions the foundational role of PVV, not of PP. However, if PVV would be deduced from PP, then Lagrange should have attributed the foundational role to the latter, not the former.

Although the objection of circularity may be avoided, the PP still has some problems. I will present here just two, both presented in Carl Jacobi’s Vorlesungen über analytische
One such problem is that it considers only the vertical movement of the weight $G$. Lagrange assumes that gravity is the only force acting upon $G$ and from here he inferred that $G$ necessarily descends. However, the pendulum is a clear illustration of the fact that the direction of a force acting on a body and the direction of that body’s movement do not necessarily coincide. Moreover, Lagrange does not seem to distinguish between stable, unstable and median equilibrium. To use Helmut Pulte’s example in [Pulte, 1998, p. 168], suppose the weight $G$ is suspended using a rigid rod and that it is placed perpendicularly above the point of suspension. In this case, $G$ would be in a state of unstable equilibrium and any infinitesimal displacement would change the state of the body.

The formula (5.32) is not general enough, since it is applied to a single body. To reach a general formula, Lagrange applies the method of multipliers. In the following paragraphs I will summarize this method, as it is one of Lagrange’s numerous contributions to analytical mechanics.

Let $L = 0, M = 0, N = 0, \ldots$ be the mathematical equations representing the constraints acting on the system, where $L, M, N, \ldots$ are functions of $x, y, z, x', y', z'$. Through derivation, we get

$$\frac{dL}{dx} = 0 \quad \frac{dM}{dy} = 0 \quad \frac{dN}{dz} = 0$$

which express the relations between the differentials of the same variables, insofar as these differentials are linear. The equations of (5.35) are used to eliminate the same number of differentials from (5.32). The coefficients of the remaining differentials of (5.32) will have to be 0 in order to have equilibrium. Obviously, the same result can be obtained if we simply multiply each equation in (5.35) with an unknown coefficient such that when all these
equations are added to (5.32) the result will still be 0:

\[
Pdp + Qdq + Rdr + \cdots + \lambda dL + \mu dM + \nu dN + \cdots = 0 \quad (5.36)
\]

This is the general equation of motion. The particular equations, i.e. the equations of motion for a particular coordinate, will be obtained by dividing (5.36) by the corresponding infinitesimal:

\[
\begin{align*}
P \frac{dp}{dx} + Q \frac{dq}{dx} + R \frac{dr}{dx} + \cdots + \lambda \frac{dL}{dx} + \mu \frac{dM}{dx} + \nu \frac{dN}{dx} + \cdots &= 0 \\
P \frac{dp}{dy} + Q \frac{dq}{dy} + R \frac{dr}{dy} + \cdots + \lambda \frac{dL}{dy} + \mu \frac{dM}{dy} + \nu \frac{dN}{dy} + \cdots &= 0 \\
P \frac{dp}{dz} + Q \frac{dq}{dz} + R \frac{dr}{dz} + \cdots + \lambda \frac{dL}{dz} + \mu \frac{dM}{dz} + \nu \frac{dN}{dz} + \cdots &= 0 \quad (5.37)
\end{align*}
\]

Lagrange regards the terms \(\lambda dL, \mu dM, \nu dN\) \ldots as moments of general forces.

However, these forces are not \(\lambda, \mu, \ldots\) as it might be thought. Lagrange rewrites \(dL\) as

\[
dL = dL' + dL'' + dL''' + \ldots,
\]

where \(dL'\) has only terms containing \(dx', dy', dz'\), \(dL''\) only terms containing \(dx'', dy'', dz''\) and so on. It follows that we can write

\[
\lambda dL = \lambda dL' + \lambda dL'' + \lambda dL''' + \ldots \quad (5.38)
\]
Each term of (5.38) can be re-written:

\[
\lambda dL' = \lambda \sqrt{\left(\frac{dL'}{dx'}\right)^2 + \left(\frac{dL'}{dy'}\right)^2 + \left(\frac{dL'}{dz'}\right)^2} \times \frac{dL'}{\sqrt{\left(\frac{dL}{dx}\right)^2 + \left(\frac{dL}{dy}\right)^2 + \left(\frac{dL}{dz}\right)^2}}
\]

\[
\lambda dL'' = \lambda \sqrt{\left(\frac{dL''}{dx''}\right)^2 + \left(\frac{dL''}{dy''}\right)^2 + \left(\frac{dL''}{dz''}\right)^2} \times \frac{dL''}{\sqrt{\left(\frac{dL'}{dx'}\right)^2 + \left(\frac{dL'}{dy'}\right)^2 + \left(\frac{dL'}{dz'}\right)^2}}
\]

\[
\lambda dL''' = \lambda \sqrt{\left(\frac{dL'''}{dx'''}\right)^2 + \left(\frac{dL'''}{dy'''}\right)^2 + \left(\frac{dL'''}{dz'''}\right)^2} \times \frac{dL'''}{\sqrt{\left(\frac{dL''}{dx''}\right)^2 + \left(\frac{dL''}{dy''}\right)^2 + \left(\frac{dL''}{dz''}\right)^2}}
\]

\[
\lambda dL'''' = \lambda \sqrt{\left(\frac{dL''''}{dx'''}\right)^2 + \left(\frac{dL''''}{dy'''}\right)^2 + \left(\frac{dL''''}{dz'''}\right)^2} \times \frac{dL''''}{\sqrt{\left(\frac{dL''}{dx''}\right)^2 + \left(\frac{dL''}{dy''}\right)^2 + \left(\frac{dL''}{dz''}\right)^2}}
\]

\[
\lambda dL'''' = \lambda \sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2} \times \frac{dL''''}{\sqrt{\left(\frac{dL''}{dx'''}\right)^2 + \left(\frac{dL''}{dy'''}\right)^2 + \left(\frac{dL''}{dz'''}\right)^2}}
\]

\[
\lambda dL''''' = \lambda \sqrt{\left(\frac{dL'''''}{dx'''''}\right)^2 + \left(\frac{dL'''''}{dy'''''}\right)^2 + \left(\frac{dL'''''}{dz'''''}\right)^2} \times \frac{dL'''''}{\sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2}}
\]

\[
\lambda dL'''''' = \lambda \sqrt{\left(\frac{dL'''''}{dx'''''}\right)^2 + \left(\frac{dL'''''}{dy'''''}\right)^2 + \left(\frac{dL'''''}{dz'''''}\right)^2} \times \frac{dL'''''}{\sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2}}
\]

\[
\lambda dL'''''' = \lambda \sqrt{\left(\frac{dL'''''}{dx'''''}\right)^2 + \left(\frac{dL'''''}{dy'''''}\right)^2 + \left(\frac{dL'''''}{dz'''''}\right)^2} \times \frac{dL'''''}{\sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2}}
\]

\[
\lambda dL'''''' = \lambda \sqrt{\left(\frac{dL'''''}{dx'''''}\right)^2 + \left(\frac{dL'''''}{dy'''''}\right)^2 + \left(\frac{dL'''''}{dz'''''}\right)^2} \times \frac{dL'''''}{\sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2}}
\]

\[
\lambda dL'''''' = \lambda \sqrt{\left(\frac{dL'''''}{dx'''''}\right)^2 + \left(\frac{dL'''''}{dy'''''}\right)^2 + \left(\frac{dL'''''}{dz'''''}\right)^2} \times \frac{dL'''''}{\sqrt{\left(\frac{dL''''}{dx''''}\right)^2 + \left(\frac{dL''''}{dy''''}\right)^2 + \left(\frac{dL''''}{dz''''}\right)^2}}
\]

In turn, each one of these equations represents the moment of a force applied to a body and perpendicular to a curved surface. For instance, in the first equation, \(\lambda \sqrt{\left(\frac{dL'}{dx'}\right)^2 + \left(\frac{dL'}{dy'}\right)^2 + \left(\frac{dL'}{dz'}\right)^2} \times \frac{dL'}{\sqrt{\left(\frac{dL}{dx}\right)^2 + \left(\frac{dL}{dy}\right)^2 + \left(\frac{dL}{dz}\right)^2}}\) is the force acting on the body with coordinates \((x', y', z')\) and perpendicular to the curved surface given by the equation \(dL'\). This result can be generalized to \(\lambda dL\) which can be regarded as the “effect of different forces expressed by \(\lambda \sqrt{\left(\frac{dL}{dx}\right)^2 + \left(\frac{dL}{dy}\right)^2 + \left(\frac{dL}{dz}\right)^2} \times \frac{dL}{\sqrt{\left(\frac{dL}{dx'}\right)^2 + \left(\frac{dL}{dy'}\right)^2 + \left(\frac{dL}{dz'}\right)^2}}\) and applied to bodies located at coordinates \((x', y', z'), (x'', y'', z''), \ldots\) following directions perpendicular to different curved surfaces represented by the equation \(dL = 0\), by varying first \(x', y', z'\), then \(x'', y'', z''\), and so on the others.” [Lagrange, 1853, §8, Part II, Section IV]

Applying these considerations to the general equation (5.36) and to the particular cases of (5.37), these equations come to represent a different system, i.e. a system which includes, apart from the bodies acted upon by the forces \(P, Q, R, \ldots\), a second set of bodies with coordinates \((x', y', z'), (x'', y'', z''), \ldots\) acted upon by the forces \(\lambda \sqrt{\left(\frac{dL}{dx'}\right)^2 + \left(\frac{dL}{dy'}\right)^2 + \left(\frac{dL}{dz'}\right)^2} \times \frac{dL}{\sqrt{\left(\frac{dL}{dx'}\right)^2 + \left(\frac{dL}{dy'}\right)^2 + \left(\frac{dL}{dz'}\right)^2}}\), \(\lambda \sqrt{\left(\frac{dL}{dx'''}\right)^2 + \left(\frac{dL}{dy'''}\right)^2 + \left(\frac{dL}{dz'''}\right)^2} \times \frac{dL}{\sqrt{\left(\frac{dL}{dx'''}\right)^2 + \left(\frac{dL}{dy'''}\right)^2 + \left(\frac{dL}{dz'''}\right)^2}}\), etc.

Lagrange interprets this result in a peculiar way, which intertwines mathematical concepts
with theoretical ones. He regards the product $\lambda dL$ as the moment of a force with magnitude $\lambda$ which tends to vary the function $L$ separately towards the coordinates $(x', y', z')$, $(x'', y'', z'')$, \ldots This is at odds with Lagrange’s emphasis in *Théorie des fonctions analytiques* and *Leçons sur le calcul des fonctions* on the need to eliminate non-mathematical concepts from the calculus. Here he seems to turn a blind eye to the intrusion of the mechanical concept of force into calculus. His only justification is an analogy with the effect of the forces $P$, $Q$, $R$, etc. which tend to change the lines $p$, $q$, $r$, etc. by either increasing or decreasing them. In a three dimensional coordinate system these lines are represented as functions of three variables $(p(x, y, z), q(x, y, z), r(x, y, z))$, which allows the interpretation of the concept of force as tending to change a function. However, it must be kept in mind that $p(x, y, z)$, $q(x, y, z)$, $r(x, y, z)$, being the equations of straight lines, cannot contain any variable with a power greater than 1, while there is no such constraint on $L$, $M$, $N$, etc.

$\lambda$, $\mu$, \ldots were introduced as purely *mathematical* entities, yet Lagrange freely interprets them as forces. In particular, these forces can be viewed as emerging from the internal interactions arising between parts of the system. Therefore, the products $\lambda dL$, $\mu dM$, $\nu dN$ \ldots are the moments of these forces which try to diminish the values of the functions $L$, $M$, $N$. In this double interpretation of $\lambda dL$, $\mu dM$, $\nu dN$ \ldots lies the innovation of Lagrange’s method of multipliers. The reason is that it allows the deduction of the conditions of equilibrium even for systems in which not all forces are known, insofar as their action can be represented as an equation. Moreover, this interpretation does not affect the conditions of equilibrium of the system. Even if $\lambda dL$, $\mu dM$, $\nu dN$ \ldots are considered simple equations, the relation between the forces of the system, represented by (5.32), will still hold.

Lagrange’s foundations of statics are not as rigorous as he might have desired. Nevertheless, the project of the *Méchanique Analytique* is to construct an axiomatic mechanics expressed using the mathematical tools of analysis. Although the principles proposed by
Lagrange are not as self-evident as their status of axioms requires, this does not mean that Lagrange’s mechanical project is a failure.

Lagrange defines dynamics as “the science of accelerating or decelerating forces and the different movements they must produce.”\textsuperscript{28} The accelerating force is, in modern terms, the acceleration. The product of mass and accelerating force, is what Lagrange calls the \textit{elementary force} or the \textit{nascent force}. However, Lagrange gives two interpretations of the product $ma$. On one hand, the product can be viewed as the \textit{pressure} the body may exert in virtue of its having that velocity. Under this interpretation, the body is viewed as an \textit{active} element, able to affect other bodies. According to the second interpretation, the product is the \textit{motor force} required to get accelerate the body from rest to the same velocity. Here, the body is a \textit{passive} element, i.e. it is the subject of a force, not a cause in itself.

The basis of Lagrange’s dynamics is d’Alembert’s principle:

\begin{quote}
If we impress a movement to several bodies which are forced to change because of their mutual action, it is clear that we can regard these movements as being composed of those that the bodies really have, and other movements which are destroyed; from here it follows that the latter must be such that the bodies animated only by these forces will be in equilibrium.\textsuperscript{29}
\end{quote}

The principle, originally presented in d’Alembert’s \textit{Traité de dynamique} (1743), was used to solve the problem of the centre of oscillation of a double pendulum. According to Lagrange, the movement of such a system would be “somewhere in between” the movement of the two bobs taken as simple pendulums. At this point of his exposition, Lagrange is using a very vague and ambiguous language. For instance, he envisages a certain “compensation and repartition”\textsuperscript{[Lagrange, 1853, p. 218]} of the movements of the two bobs, in the sense that the

\textsuperscript{28} “...la science des forces accélératrices ou retardatrices, et des mouvements variés qu’elle doivent produire.”\textsuperscript{[Lagrange, 1853, p. 207]}

\textsuperscript{29} “Si l’on imprime à plusieurs corps des mouvements qu’ils soient forcés de changer à cause de leur action mutuelle, il est clair qu’on peut regarder ces mouvements comme composés de ceux que les corps prendront réellement, et d’autres mouvements qui sont détruits; d’où il suit que ces derniers doivent être tels, que les corps animés de ces seuls mouvements se fassent équilibre. [Lagrange, 1853, §10 p.23]
upper bob will tend to increase the frequency of the oscillations, while the lower one will tend to decrease it. Therefore, concludes Lagrange, there has to be a point, between the two bobs, where if a third bob were to be placed, it would not be neither accelerated, nor decelerated by the other two.

The first step towards finding a solution to the problem of the centre of oscillation had been done, according to Lagrange, by Jacques Bernoulli. Bernoulli considered the simple case of two masses situated on a rigid rod and observed that the upper mass moves slower than it would move if it were not connected to the lower mass and that the lower mass moves faster than it would move if it were on its own. Bernoulli concludes that there is a “communication” of movement and since the two bobs are connected by a rigid rod thus forming a second class lever, this communication will obey the laws of the equilibrium applied to the forces acting on the two masses. For Lagrange, Bernoulli was right to try to relate the problem of the centre of oscillation to statics, but he was misled in using in his calculations velocities defined on finite rather than infinitesimal time intervals.

D'Alembert carried Bernoulli’s idea further and by introducing PDA he managed to “reduce all the laws of the movement of bodies to those of their equilibrium, and [to bring] back Dynamics to Statics.” However, the PDA cannot be used to solve problems of dynamics on its own, but it enables the application of the principles of statics to dynamics, but only for internal forces, i.e. forces that appear between the elements of a system but do not affect the movement of the system as a whole.

The fundament of dynamics is a principle Lagrange develops in the second section of the Dynamics. He begins by considering a system of bodies upon which act different forces. From this system, pick out a body \( m \) regarded as a point of mass, with coordinates \( x, y, z \) at time

---

30 “...réduit toutes les lois du mouvement des corps à celles de leur équilibre, et ramène ainsi la Dynamique à la Statique.” [Lagrange, 1853, p. 223]
\( t \), considered relative to three axes perpendicular on each other. The instantaneous velocities along the three axes are *represented* as \( \frac{dx}{dt} \), \( \frac{dy}{dt} \), \( \frac{dz}{dt} \). Lagrange adds that due to the action of forces these velocities will be incremented in an infinitely small amount of time \((dt)\) by \( d \cdot \frac{dx}{dt} \), \( d \cdot \frac{dy}{dt} \), \( d \cdot \frac{dz}{dt} \). This gives the *accelerating forces* \( \frac{d^2x}{dt^2} \), \( \frac{d^2y}{dt^2} \), \( \frac{d^2z}{dt^2} \).

At this point of his exposition Lagrange makes an important remark:

We can regard these increments \([d \cdot \frac{dx}{dt}, d \cdot \frac{dy}{dt}, d \cdot \frac{dz}{dt}]\) as new velocities given to each body and, by dividing them by \( dt \) we will have the measure of the instantaneous accelerating forces used to produce them; because, for any variable that might be the action of a force, we can always, in *virtue of the nature of the differential calculus*, regard it as constant during an infinitely small amount of time, and the velocity generated by this force is then proportional to the force multiplied by the time.\(^{31}\)

In this paragraph, Lagrange is talking about the measure of *accelerating forces*. An important point for our present purposes is the relation between the *nature of the calculus* and the *measure of accelerating force*. On the basis of this relation we can further calculate the *motor force*. This is done by considering \( dt \) constant relative to the increments of the velocities along the three axes, which means that the accelerating forces will be \( \frac{d^2x}{dt^2} \), \( \frac{d^2y}{dt^2} \), \( \frac{d^2z}{dt^2} \). By multiplying these values with the mass \( m \) we obtain the expressions of the forces acting upon the body \( m \) in the interval \( dt \).

Therefore, we can speak about the relation between the *nature of the calculus* and the *measure of motor force*. Essentially, this relation is of *constraint of representation*, for Lagrange infers from the nature of the calculus that the action of a force (i.e. the increment of the velocity) can be considered constant over an infinitely small time interval and from here he concludes that for a constant time interval the accelerating forces are represented as \( d \cdot \frac{dx}{dt} \), \( d \cdot \frac{dy}{dt} \), \( d \cdot \frac{dz}{dt} \).

\(^{31}\)"On peut regarder ces accroissements comme de nouvelles vitesses imprimées à chaque corps, et, en les divisant par \( dt \), on aura la mesure des forces accélératrices employées immédiatement à les produire; car, quelque variable que puisse être l'action d'une force, on peut toujours, par la nature du calcul différentiel, la regarder comme constante pendant un temps infiniment petit, et la vitesse engendrée par cett force est alors proportionelle à la force multipliée par le temps.[Lagrange, 1853, p. 232, My emphasis]"
More precisely, Lagrange justifies the representation of accelerating forces, and in consequence of motor forces, by referring to the nature of the differential calculus. Whether we understand the differential calculus as a special case of calculus of finite differences as Euler, or as the study of fluents and fluxions as Newton, it is nevertheless true that a function can be considered constant if the increment of its variable is sufficiently small. An illustration of this essential aspect of differential calculus is given in Figure 5.3. As \( D \) approaches \( C \), \( d \) will approach \( c \).

![Figure 5.3: Reduction of a curve to infinitely small polygons. Source: [Newton et al., 1729, p. 45]](image)

Eventually, the length of \( CD \) will be sufficiently small such that \( dc \) will be constant. Applied to the case of the increments of velocity, this translates as

\[
d \cdot \frac{dx}{dt} = \alpha
\]

where \( \alpha \) is an arbitrary constant. However, the measure of a force is the velocity it is able to produce, which means that if we regard \( d \cdot \frac{dx}{dt} \) as a velocity, then it will be equal to the force that produced it multiplied by \( dt \), the time it takes to reach the velocity \( d \cdot \frac{dx}{dt} \). Thus, this essential character of the differential calculus makes possible, at least in Lagrange’s view, the mathematical representation of the accelerating force.

Lagrange now generalizes this result to all bodies of the system. That is, the force acting
on each body will be the result of the composition of three forces, each parallel to one of the three axes. More importantly, the sum of the momenta of all these forces must equal the sum of the momenta of all forces acting upon the entire system. Consider that the forces \( mP, mQ, mR, \ldots \) act upon the entire system, where \( P, Q, R, \ldots \) are accelerating forces. If we apply forces equal in magnitude to \( mP, mQ, mR, \ldots \), but in opposite directions, then the system will be in equilibrium. To make matters easier, Lagrange introduces the following notation convention: \( d \) denotes the differentials relative to time, \( \delta \) represents the virtual velocities. Thus, the moment of the force \( \frac{d^2x}{dt^2}m \) will be \( m\frac{d^2x}{dt^2}\delta x \). The sum of all moments will be:

\[
S\left(\frac{d^2x}{dt^2}\delta x + \frac{d^2y}{dt^2}\delta y + \frac{d^2z}{dt^2}\delta z\right)m + S(P\delta p + Q\delta q + R\delta r + \ldots)m = 0 \tag{5.41}
\]

where the symbol \( S \) stands for the indefinite integral applied to all points of mass.

In the fourth section of the second part of *Méchanique Analitique*, in §1, Lagrange explains that applying (5.41) to particular cases amounts to reducing the number of variables by using the equations describing the particular system. Here, just as in the method of multipliers, Lagrange considers that each system is necessarily describable by a set of equations. Moreover, at least in the case of the method of multipliers, these equations can be interpreted as expressing relations between forces. This is not different from the function I attributed to the second layer of models\(^{32}\), the theoretical description. The equations that expressed the relations between the forces acting on the pendulum’s bob have the same function as the equations \( L = 0, M = 0, \) etc. of the method of multipliers, and as the equations used to reduce the number of variables in (5.41).

For Lagrange, the first part of (5.41), \( S\left(\frac{d^2x}{dt^2}\delta x + \frac{d^2y}{dt^2}\delta y + \frac{d^2z}{dt^2}\delta z\right)m \), represents the effect of the forces of inertia. More precisely, is the result of the internal interactions between the com-

\(^{32}\)see section 2.5
ponents of the system. In the case of the double pendulum this expression would represent the
effects the two bobs have on each other. The second part of (5.41), $S(P\delta p+Q\delta q+R\delta r+\ldots)m$
represents the effects of the active forces, i.e. forces that move the entire system. In the case
of the double pendulum, this expression would represent the forces in virtue of which the rod
$l_1$ connected to the fulcrum oscillates.

Lagrange’s inclination towards generality is illustrated by his attempt to find a single
equation which ignores the particularities of the system to which (5.41) is applied and allows
the deduction of all variable quantities of the system. To find this equation, Lagrange con-
siders a function $\Phi$ of variables $x$, $y$, $z$, ..., $dx$, $dy$, $dz$, ..., $d^2x$, $d^2y$, $d^2z$, .... A change of
the coordinate system, from $xyz$... to $\xi\psi\phi$... will transform $\Phi$ into a function of $\xi$, $\psi$, $\phi$,
..., $d\xi$, $d\psi$, $d\phi$, ..., $d^2\xi$, $d^2\psi$, $d^2\phi$, ...

Next, he differentiates both functions ($\Phi(x, y, z, \ldots, dx, dy, dz, \ldots, d^2x, d^2y, d^2z, \ldots)$, and
$\Phi(\xi, \psi, \phi, \ldots, d\xi, d\psi, d\phi, \ldots, d^2\xi, d^2\psi, d^2\phi, \ldots)$) relative to $\delta$

\[
\begin{align*}
\Phi(x, y, z, \ldots, dx, dy, dz, \ldots, d^2x, d^2y, d^2z, \ldots) &= \Phi(\xi, \psi, \phi, \ldots, d\xi, d\psi, d\phi, \ldots, d^2\xi, d^2\psi, d^2\phi, \ldots) \\
\delta\Phi &= \frac{\delta\Phi}{\delta x} \delta x + \frac{\delta\Phi}{\delta dx} \delta dx + \frac{\delta\Phi}{\delta d^2x} \delta d^2x + \ldots \\
&+ \frac{\delta\Phi}{\delta y} \delta y + \frac{\delta\Phi}{\delta dy} \delta dy + \frac{\delta\Phi}{\delta d^2y} \delta d^2y + \ldots \\
&\quad \ldots \\
&= \frac{\delta\Phi}{\delta \xi} \delta \xi + \frac{\delta\Phi}{\delta \psi} \delta \psi + \frac{\delta\Phi}{\delta \phi} \delta \phi + \ldots \\
&+ \frac{\delta\Phi}{\delta d\xi} \delta d\xi + \frac{\delta\Phi}{\delta d\psi} \delta d\psi + \frac{\delta\Phi}{\delta d\phi} \delta d\phi + \ldots \\
&+ \frac{\delta\Phi}{\delta d^2\xi} \delta d^2\xi + \frac{\delta\Phi}{\delta d^2\psi} \delta d^2\psi + \frac{\delta\Phi}{\delta d^2\phi} \delta d^2\phi + \ldots \\
\end{align*}
\]
Given the relation $d^{m}Fy = \delta d^{m}Fy(5.17)$, Lagrange replaces $\delta d$ and $\delta d^{2}$ with $d\delta$ and $d^{2}\delta$, respectively, which allows him to integrate (5.42). The result has the form

\[
\int (A\delta x + B\delta y + C\delta z + \ldots) + Z = \int (A'\delta \xi + B'\delta \psi + C'\delta \phi + \ldots) + Z' \tag{5.43}
\]

where the coefficients $A, B, C, A', B', C', \ldots$ are calculated from (5.42):

\[
\begin{align*}
A &= \frac{\delta \Phi}{\delta x} - d\frac{\delta \Phi}{\delta x} + d^{2}\frac{\delta \Phi}{\delta x^{2}} - \ldots \\
B &= \frac{\delta \Phi}{\delta y} - d\frac{\delta \Phi}{\delta y} + d^{2}\frac{\delta \Phi}{\delta y^{2}} - \ldots \\
C &= \frac{\delta \Phi}{\delta z} - d\frac{\delta \Phi}{\delta z} + d^{2}\frac{\delta \Phi}{\delta z^{2}} - \ldots \\
A' &= \frac{\delta \Phi}{\delta \xi} - d\frac{\delta \Phi}{\delta \xi} + d^{2}\frac{\delta \Phi}{\delta \xi^{2}} + \ldots \\
B' &= \frac{\delta \Phi}{\delta \psi} - d\frac{\delta \Phi}{\delta \psi} + d^{2}\frac{\delta \Phi}{\delta \psi^{2}} \\
C' &= \frac{\delta \Phi}{\delta \phi} - d\frac{\delta \Phi}{\delta \phi} + d^{2}\frac{\delta \Phi}{\delta \phi^{2}} \tag{5.44}
\end{align*}
\]

Differentiating (5.43) Lagrange obtains:

\[
A\delta x + B\delta y + C\delta z + \ldots - A'\delta \xi - B'\delta \psi - C'\delta \phi = dZ' - dZ \tag{5.45}
\]

125
(5.45) should hold for any variation \( \delta \), since Lagrange no particular value of \( \delta \) was assumed in deducing (5.45). The right side of the equality is an exact differential of \( Z \) and \( Z' \), while the left side is independent of \( d \). The sides cannot be equal unless they are both zero. Therefore,

\[
A \delta x + B \delta y + C \delta z + \cdots = A' \delta \xi + B' \delta \psi + C' \delta \phi + \cdots  
\]

\[
dZ = dZ'  
\]

In order to obtain \( S(d^2x \delta x + d^2y \delta y + d^2z \delta z)m \), Lagrange considers \( \Phi = \frac{1}{2}(dx^2 + dy^2 + d^2z) \) for which

\[
\frac{\delta \Phi}{\delta x} = 0 \quad \frac{\delta \Phi}{\delta dx} = dx \quad \frac{\delta \Phi}{\delta^2 x} = 0  
\]

\[
A = -d^2x \quad B = -d^2y \quad C = -d^2z  
\]

From this it follows that \( \Phi \) contains only first order differentials. Therefore, the equations of \( A', B', C', \ldots \) will also contain only first order differences:

\[
A' = \frac{\delta \Phi}{\delta \xi} - d \frac{\delta \Phi}{\delta d\xi}  
\]

\[
B' = \frac{\delta \Phi}{\delta \psi} - d \frac{\delta \Phi}{\delta d\psi}  
\]

\[
C' = \frac{\delta \Phi}{\delta \phi} - d \frac{\delta \Phi}{\delta d\phi}  
\]

Replacing these values in (5.46) we obtain:

\[
\Phi = -d^2x \delta x - d^2y \delta y - d^2z \delta z  
\]

\[
= \left( \frac{\delta \Phi}{\delta \xi} - d \frac{\delta \Phi}{\delta d\xi} \right) \delta \xi + \left( \frac{\delta \Phi}{\delta \psi} - d \frac{\delta \Phi}{\delta d\psi} \right) \delta \psi  
\]

\[
+ \left( \frac{\delta \Phi}{\delta \phi} - d \frac{\delta \Phi}{\delta d\phi} \right) \delta \phi + \cdots  
\]
The first term of (5.41), \( S\left(\frac{d^2x}{dt^2}\delta x + \frac{d^2y}{dt^2}\delta y + \frac{d^2z}{dt^2}\delta z\right)m \), is \( S\Phi dt^2 m \), i.e. \( S\left(\frac{d^2x+d^2y+d^2z}{2dt^2}\right) m \). Since Lagrange is trying to generalize (5.41), he must express it in a form independent of the initial choice of the system of coordinates. Therefore, he must show that both terms of (5.41) can be expressed in terms of \( \xi, \psi, \phi, \ldots \)

Lagrange defines the function \( T \) as the expression of \( S\left(\frac{d^2x}{dt^2}\delta x + \frac{d^2y}{dt^2}\delta y + \frac{d^2z}{dt^2}\delta z\right)m \) in terms of \( \xi, \psi, \phi, \ldots \) He then finds it obvious that

\[
S\left(\frac{d^2x}{dt^2}\delta x + \frac{d^2y}{dt^2}\delta y + \frac{d^2z}{dt^2}\delta z\right)m = \left(\frac{\delta T}{\delta \delta \xi} - \frac{\delta T}{\delta \xi}\right) \delta \xi + \left(\frac{\delta T}{\delta \delta \psi} - \frac{\delta T}{\delta \psi}\right) \delta \psi + \ldots
\]

(5.51)

Turning to the second term of (5.41), expressing it in terms of \( \xi, \psi, \phi, \ldots \) is easier, since it amounts to translating the equations of the lines \( p, q, r, \ldots \) and the forces in a different coordinate system (\( \xi \psi \phi \ldots \)). Lagrange’s argument goes as follows:

\[
d\Pi = Pdp + Qdq + Rdr + \ldots
\]

(5.52)

\[
\delta \Pi = P\delta p + Q\delta q + R\delta r + \ldots
\]

(5.53)

\[
S \left( P\delta p + Q\delta q + R\delta r + \ldots \right)m = S\delta \Pi m = \delta S\Pi m
\]

(5.54)

(5.55)

(5.54) is obtained from (5.53) by multiplying the latter then summing up the result for the masses of all the bodies in the system. Lagrange defines the function \( V = S\Pi m \). This entails that

\[
\delta S\Pi m = \delta V = \frac{dV}{d\xi} \delta \xi + \frac{dV}{d\psi} \delta \psi + \frac{dV}{d\phi} \delta \phi + \ldots
\]

(5.56)
Adding up (5.51) and (5.56) Lagrange obtains the general formula of dynamics

$$\Xi \delta \xi + \Psi \delta \psi + \Phi \delta \phi + \cdots = 0$$

(5.57)

where

$$\Xi = d \cdot \frac{\delta T}{\delta \xi} - \frac{\delta T}{\delta \xi} + \frac{\delta V}{\delta \xi}$$

$$\Psi = d \cdot \frac{\delta T}{\delta \psi} - \frac{\delta T}{\delta \psi} + \frac{\delta V}{\delta \psi}$$

(5.58)

$$\Phi = d \cdot \frac{\delta T}{\delta \phi} - \frac{\delta T}{\delta \phi} + \frac{\delta V}{\delta \phi}$$

It might be tempting to think of $T$ and $V$ as kinetic and potential energy. However, this would be a misunderstanding of Lagrange’s position. For him, both $T$ and $V$ were nothing else but functions in which the variables represented the coordinates of the bodies. Using (5.57) has two advantages. First, there is no need to individuate the forces acting on every body of the system. We observed this in section 2.3 where the same procedure was used to find the equations of motion of the double pendulum.

However, in that case, the theoretical description of the system employed the concept of energy. An implicit assumption in using the Lagrangian in 2.3 was that the energy of the system is conserved. For Lagrange, on the other hand, the conservation of energy is a consequence of the fundamental principle of dynamics ((5.57)). The second advantage is that we can choose the coordinate system that best suits our purposes. As long as we can express the relation between the system $xyz$ and the system of our choice as a function, we can always obtain equations of the form (5.58), in which the only variables are the coordinates of the new system.

This chapter provides an overview of Lagrange’s work on the calculus and mechanics. Concerning the calculus, we can say in Lagrange’s works we find traces of instrumentalism. This
is justified by his changing view of the foundations of calculus throughout his career. In the
* Mécanique Analitique* he uses infinitesimals without questioning their accuracy. However,
in *Théorie des fonctions analytiques* and *Leçons sur le calcul des fonctions* he approaches
foundational problems with the intention of providing a solid basis for the construction of the
methods of the calculus. This does not mean that he excluded infinitesimals altogether, but
quite the opposite. Lagrange justified the use of infinitesimals by showing that their results
are accurate, despite the conceptual issues they raise.

There are a couple of things to remember about Lagrange’s work. First, he attempts to
construct an axiomatic mechanics, based on the principle of the lever (PL), the principle of
composition of forces (PCF) and the principle of virtual velocities (PVV). Since the latter is too
counter-intuitive to be considered a self-evident truth, Lagrange tries to justify it by applying
the principle of the pulleys (PP), which is more an illustration of the concept of equilibrium
rather than a principle. The relation between statics and dynamics is d’Alembert’s principle
(PDA), which does not play a deductive role, i.e. as a premise which taken together with the
principles of statics allows the deduction of the principles of dynamics. Instead, the PDA is a
methodological principle which justifies the elimination of internal forces, i.e. forces arising
from the interaction between parts of the system.

Second, Lagrange is the first mathematician who treated the foundations of mechanics as
a subject of inquiry in its own right, rather than a didactic requirement. In his early works,
he used infinitesimals without taking into consideration their lack of precision. However,
in *Théorie des fonctions analytiques* and *Leçons sur le calcul des fonctions* he constructs the
differential calculus as a calculus of functions rather than ratios of infinitesimals, and proves
that the results of the calculus of differentials are sound whether it uses infinitesimals or
derivatives.
Chapter 6

Conclusions

In the previous pages I suggested a theory of scientific models as representations, according to which a model has three components: an idealized system, a theoretical description of this system and the corresponding mathematical formalisms. The idealized system is defined by a series of sentences expressed in natural language which identify the components of the system and the relations between them. The system is idealized because there need be no perfect real-world correspondent to it, e.g. there are no frictionless pendulums.

The second component is the theoretical description, a set of mathematical equations which express the relations between the theoretical entities of the system, e.g. force, kinetic or potential energy, wave amplitudes, and so on. Finally, the mathematical component is the set of equations and their solutions which represent the evolution in time of the relevant properties of the idealized system.

I supported this suggestion by analyzing the algorithms used for deducing the equations of motion of different kinds of pendulums, and tried to answer the following question: How do the three components interact? Here I introduced the term “bridge principle”, loosely defined in 2.5 as “any assertion, whether justified or simply postulated, which establishes a
correspondence between the elements of one layer and those of another”\textsuperscript{1}. Obviously, this definition is too vague to be included within a theory of models. This is why I dedicated chapters 3-5 to an in-depth analysis of the separation between the three layers (mathematical, theoretical, idealized description), which largely took place in the 18th century. This analysis aimed at finding within the history of 18th-century science a set of principles that could be considered bridge principles. I admit that the suggested theory of models needs a stronger defense than the one given in chapter 2. However, since the problem of bridging the gaps between layers is the most difficult challenge my proposed theory of models has to face, solving it takes precedence over a more thorough development of the theory.

In chapter 3 I presented the status of 18th century science focusing on its conceptual entanglement. During this period, different scientific traditions, which oftenly overlapped, competed against each other, usually on the topic of precedence of their foundational principles. More importantly, although the domains of study were delineated (e.g. mechanics was the study of rest and motion, geometry the study of curves, etc.), at the theoretical layer different sciences used concepts unrelated to their field of study. A case in point is Newton’s theory of fluxions, where he used the non-mathematical concept of velocity to define the rate of variation of a quantity. This problem was ignored up until Lagrange’s \textit{Théorie des fonctions analytiques} and \textit{Leçons sur le calcul des fonctions}, which focused on the delineation of sciences at a theoretical level.

During this period, mechanics, although a mathematized science, still relied on geometry to perform its calculations and predictions. This meant that geometrical proofs could become cumbersome for more complex mechanical problems, which involved several bodies, and that these proofs could not be extended to solve similar problems. The introduction of functions in Euler’s \textit{Introductio in analysin infinitorum} changed this. Instead of drawing curves and

\textsuperscript{1}see p. 28
representing forces as lines, solutions to mechanical problems now implied the arithmetical
manipulation of symbols.

However, explaining how geometry can be used to represent empirical phenomena is easi-
er than explaining how functions, differentials or integrals perform the same task. Using
geometry, trajectories are represented by curves, forces by lines, with the magnitude of the
force proportional to the line’s length, etc. Then we determine the relations between these
geometrical objects (e.g. proportionality). This approach is more intuitive than the manip-
ulation of symbols required by analytical solutions to mechanical problems.

Therefore, in order to understand how and why functions entered in use in science, I
looked at Euler’s work on the topic. As it turns out, the 18th-century concept of function,
as well as those of variable and constant, are very different from the modern ones. There
are two differences between the older and the more recent concepts of function. The first
is conceptual. Euler regarded functions as analytical expressions of real relations between
quantities, while modern mathematics regards functions as rules of correspondence between
sets. The second difference, derived from the first, is that the modern conception does not
admit exceptional values of its variables, therefore it is more accurate. Eulerian variables
were “essences” of quantities which meant that they were at the same time complex, ra-
tional, irrational numbers, regardless of the fact that the function might not be defined for
particular values of its variables.

The history of the concept of functional relation, briefly discussed in 4.1, shows that the
concept of analytical function is the offspring of geometry. To make things clearer, I used the
A functional relation can be expressed both as a geometrical curve, (or surface, body, if there
are several variables) and as an analytical expression. This distinction is presupposed by the
graphical representation of functions in modern mathematics. For instance, the second power
of a quantity can be expressed both as \( f(x) = x^2 \) and as a parabola.

Therefore, the representational power of functions can be considered a “trait” inherited from geometrical representations of functional relations between quantities. Geometry expressed relations between quantities as curves in a system of coordinates. This is how velocity and acceleration were represented. However, the coordinates of a point are numbers which means that we can treat coordinates as analytic variables. In turn, this implies that the function represents the curve in analytical form. Since geometry was used to represent mechanical quantities (forces, velocities, accelerations, etc.), the introduction of functions to represent geometrical curves shifted this burden from geometry to functions and related concepts. The historical sequence can be reduced to: (1) empirical phenomenon \( \rightarrow \) (2) idealization \( \rightarrow \) (3) theoretical description \( \rightarrow \) (4) geometrical representation \( \rightarrow \) (5) functional representation, where “\( \rightarrow \)” stands for “is represented by”. Eventually, the fourth step turned out to be redundant.

In section 3.2.1 I presented Newton’s view on the relation between mechanics and geometry to support my argument concerning the representational role of geometry, i.e. to support the subsequence (1)-(4). In support of the subsequence (4)-(5) I presented in chapters 3 and 4 a study of the development of the calculus as a means to solve the geometrical problems raised by applying geometry to mechanical problems. Finally, in section 5.1, I showed how Lagrange’s criticism of his predecessors cut the final ties that bound the calculus to geometry.

This holds not only for the concept of function, but also for that of infinitesimal. Infinitely small quantities, or differentials, also had their origins in geometry, as illustrated by Newton’s method of fluxions and ultimate ratios (see section 3.2.1). Euler tried to avoid the problems raised in [Berkeley, 2009] and introduced geometric ratios. Although at first sight geometric ratios are geometric only in name, the underlying intuition is geometric. Euler re-interpreted Newton’s generation of curves by the movement of a point \textit{in time} as generation of a series
in terms of infinitely small finite increments. Lagrange’s *Théorie des fonctions analytiques* and *Leçons sur le calcul des fonctions* cut this tie too.

The connection between geometry and calculus mentioned above is most visible in the second volume of *Mechanica*, where Euler applies the mechanical concepts developed in the first volume together with his view on infinitesimals to some mechanical problems. His solutions begin with a graphical representation of the trajectory and of the forces acting upon an infinitely small body. Here, the forces are represented by lines, and the bodies by points on the curves representing their trajectory. The next step of his solutions is to give a description of the system from a geometrical perspective, i.e. as relations between curves. Infinitesimals are introduced only in the last part to derive the final result. The main characteristic of Euler’s solutions is that they are not applications of the calculus to a problem of mechanics, but to a problem of geometry. It might be thought that in the second volume of *Mechanica* geometry plays the role of a redundant middleman. This is not the case and in the next paragraphs I will explain why.

At the basis of Euler’s mechanics lie the concepts of rest and motion. The reason they are given this foundational role lies in Euler’s conception of the nature of bodies (see section 4.2). The essential property of bodies is impenetrability, from which it follows that bodies necessarily have the properties of mobility, persistence and extent. Motion and rest are foundational concepts of mechanics because the disposition to be moved by a force and to maintain a state of rest are essential properties of bodies (mobility and persistence). Moreover, in virtue of having the property of extent, bodies are also infinitely divisible (see 4.2 for the argument).

Can we now say what fulfills the role of bridge principle between the mathematical and the theoretical in Euler’s work? My answer yes, but we have to settle for something less rigorous than the term “principle” might lead us to expect. Euler’s conception of bodies as infinitely divisible justifies him to consider bodies analogous to points on a curve. This is
illustrated by the solutions given in the second volume of the Mechanica, and is justified by the classical definition of points as “things” without parts. As argued in 4.2, for Euler bodies too are not composed of finite parts or infinitely small parts. So a first bridge principle would be: *An infinitely small body can be considered a geometrical point.* This would connect the theoretical layer to the idealized system. We must observe three things. First, that the principle is justified by the fact that bodies and points share some essential properties. Second, despite this similarity, the infinitely small bodies and points differ in kind. The difference is given by the fact that bodies, unlike points, are impenetrable. Thirdly, the principle is postulated. Since infinitesimal bodies and points differ in kind, they cannot be substituted for one another. In order to represent, geometrically, an infinitely small body as a point, an intentional act must occur, through which a cognitive agent decides to use points to represent bodies. This decision is justified by the similarity between the two kinds of entities.

The geometric point is also similar to infinitely small (numerical) quantities. Just as a point has no length, an infinitesimal is nothing (see 4.1.2), i.e. it is 0. Also, the infinitesimal does not have any parts, although it too can be divided *ad infinitum*. This means that we can formulate a second bridge principle: *A geometrical point can be considered an infinitely small (numerical) quantity.* The three observations made above apply here too. The principle is introduced by an intentional act of a cognitive agent, justified by the resemblance between the two kinds of entities.

Do we find something similar in Lagrange? We saw that in *Méchanique Analitique* he tried to construct an axiomatic mechanics. The main axiom of this system is the PVV which he considers the foundation of statics. Dynamics is connected with statics via PDA. However, this connection is not logical. Dynamics is not *deduced* from statics and the PDA is not an axiom of dynamics in the same way that PVV is an axiom of statics. PDA is a methodological principle which justifies the use of the laws and principles of statics. It cannot have the same
axiomatic status as PVV because it concerns only the internal interactions between the bodies composing a system, and not the movement of the entire system.

However, as Jacobi argued [Pulte, 1998], Lagrange’s mechanics is not a proper axiomatic system for two reasons. First, its fundamental axiom (PVV) is not self-evident. Second, PDA does not allow the deduction of dynamics from statics, as it should do if mechanics were an axiomatic system. Nevertheless, I think that although these principles could not serve as axioms, they play the role of bridge principles that connect the three layers of models. Thus, the principle of the lever (PL) connects the first two components of a model, i.e. the idealized system and the theoretical description.

As observed in chapter 4 and repeated above, for Euler the cause of a body’s change of state is necessarily external. A body cannot go from rest to motion or vice versa on its own, but only if a force acts upon it. If a body is at rest, then either there are no forces acting upon it, or the forces that do act cancel each other out. Since for Lagrange the aim of statics is to provide the laws by which the forces of a system cancel each other out, only the second case can be considered interesting for statics.

However, the equilibrium of a single body is explained by PCF. If $P, Q, R$ act upon a body simultaneously, then by adding $P$ and $Q$ and then adding the result to $R$, the result should be null. The principle of the lever applies to systems of several interacting bodies, i.e. systems in which the forces acting upon one composing body can be expressed as functions of the forces acting upon the other components. It would be impossible to apply PCF in this case, since it requires that the forces have the same origin. It follows that the simplest system is one containing only two bodies.

However, two bodies can be connected in several ways, some of which might be more complex than the lever, while others might not be of much. For instance, if the two bodies, having different masses, are tied to each end of a rope which passes through a pulley,
then there would be no way of equilibrating the system. If their masses are equal, then the
system will be in equilibrium, even though one of the bodies is higher than the other. By
contrast, the lever allows us to “determine” the condition of equilibrium when charged with
two weights of different masses \( \frac{P}{Q} = -\frac{q}{p} \). The mathematical expression of the condition of
equilibrium of a lever is then applied to the three-body systems, four-body systems, and so
on, after reducing the number of forces by applying the PCF.

Therefore, we can consider the PL a paradigmatic example of describing an idealized sys-
tem in theoretical terms. How does it connect the first two layers of a model? A given
idealized system can be divided into groups of two interacting bodies. Each pair of bodies is
described in theoretical terms, following the example set by PL. In the case of statics at least,
the theoretical description of the pair has the form \( Pdp + Qdq = 0 \), as shown in Lagrange’s
justification of PVV. Here I am not presenting an algorithm for constructing the theoretical
description of an idealized system, but the way in which the theoretical description of any
idealized system is justified. In this sense, the PL, together with PCF, are bridge principles
connecting the first two layers of a model, the theoretical description and the idealized sys-
tem.

What connects the mathematical layer to the theoretical description? It depends. In
statics, the PVV seems to have this function, while in dynamics, the PVV in conjunction with
PDA, which makes possible the application of PVV. In Lagrange’s case, it is much harder
to identify a bridge principle between the mathematical and the theoretical layer, mainly
because he tries to completely eliminate forces from mechanics. This makes the identification
of bridge principles quite difficult.

Nevertheless, by looking at his demonstration of the general principle of dynamics, we
can see that an argument similar to the one made concerning the other bridge principle, can
also be made here. Lagrange begins by considering the simplest case, that of a simple body
acted upon by forces which he then generalizes, by adding all the momenta of all the forces act on all masses, i.e. \(S\left(\frac{d^2\delta x}{dt^2} + \frac{d^2\delta y}{dt^2} + \frac{d^2\delta z}{dt^2}\right)m\). This is guaranteed by PDA.

In the previous chapter I did note that the *Méchanique Analitique* is completely anti-metaphysical, refusing to engage in any debates on the nature of mass, force, body, etc. and this is one point in which Lagrange and Euler differ greatly. Euler tried to build mechanics on a solid basis of *a priori* principles, and to a certain extent he succeeded. Lagrange, on the other hand, tries to construct an axiomatic system beginning with a few intuitive principles (PP, PL), which can be grasped from experience\(^2\).

The conclusion is that the bridge principles connecting the parts of a model depend on the nature of the mathematical concepts involved, as illustrated by the part played by infinitesimals both in Euler and in Lagrange’s mechanics. Also, these principles depend on the overall theory in which we are working. Although both Euler and Lagrange can be said to have been working within the confines of Newtonian mechanics, their views on the nature of mechanics and the calculus shaped the way in which the two domains combine.

For these reasons it would be difficult, perhaps impossible, to provide a definition of bridge principles which would cover any possible case. I would suggest that the three-layered structure of models be viewed more as a framework for the historical analysis of models, rather than a theory of models. The identification of the different layers of models and the ways in which these layers have been connected could expand our understanding of the interactions between the development of mathematics and that of scientific theories.

\(^2\)What are lever and pulleys if not the most common items on an 18th century dock?
Bibliography


Euler, L. (1736). Mechanica sive motus scientia analytice exposita. In *Opera Omnia*, volume II. [77, 78, 79]

Euler, L. (1740). De infinitis curvis eiusdem generis seu methodus inveniendi aequationes pro infinitis curvis eiusdem generis. *Commentarii academicae scientiarum Petropolitanae*, 7. [102]

Euler, L. (1862). Anleitung zur naturlehre. In *Opera Omnia*, volume III. [73, 74, 75, 76]


Newton, I. and Stewart, J. (1745). *Sir Isaac Newton’s Two treatises: Of the quadrature of curves, and Analysis by equations of an infinite number of terms, explained: containing the treatises themselves, translated into English, with a large commentary; in which the demonstrations are supplied where wanting, the doctrine illustrated, and the whole accommodated to the capacities of beginners, for whom it is chiefly designed*. Printed by James Bettenham, at the expense of the Society for the encouragement of learning; and sold by John Nourse and John Whiston. [48, 49]


