1

Fundamental Principles of Discrete Mechanics

This chapter is dedicated to the foundations of discrete mechanics. The notion of space is defined directly as a set of topological elements: edges and surfaces. These geometric elements exist at every scale and cannot be reduced to a point like in a continuum; as a result, we must abandon the concept of local differentiation, as well as inertial and non-inertial frames of reference. Some of the classical principles of physics can be kept, such as the weak equivalence principle and the principle of relativity, and some new physical principles are encountered for the first time, such as Hodge–Helmholtz decomposition. We also require new axioms and hypotheses: the accumulation of stresses and the duality of mechanical actions of all kinds.

1.1. Definitions of discrete mechanics

1.1.1. *Notion of discrete space–time*

A method of positioning ourselves within space and time is essential if we wish to represent the universe around us, whether the universe of our daily lives, or the wider universe governed by the laws of general relativity. Positioning systems (GPS and Galileo) have become indispensable tools for many human activities such as transportation, well-drilling and so on. But in fact, to move toward a nearby target, we do not need to know our position exactly with respect to some absolute reference; we simply need to know the path to our target. The various theories of mechanics (Newtonian, quantum, continuum, relativistic, etc.) do not contradict each other – quite the opposite – but the connections between them have not yet been definitively established. Each theory of mechanics only describes a part of reality. The concepts, analysis tools and hypotheses of each theory vary. Ultimately, a unified theory of mechanics might not be strictly necessary.

The theory of discrete mechanics presented here assumes that there exists a time, the present, that describes the state of a physical system instantaneously. Although this image of the present exists as such, an observer located within space can only perceive its environment at later moments in time, since waves (light, sound, tidal waves) travel at finite velocity. The present can, therefore, only be perceived by an exterior observer in the form of a mathematical model that provide an instantaneous description of every phenomenon in the physical system.

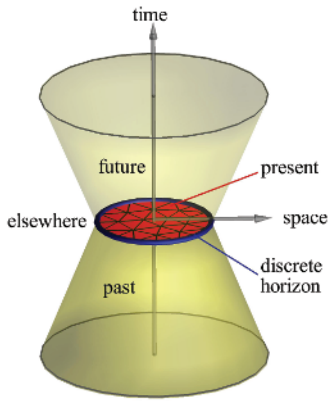

Figure 1.1. *Light cone in space–time. For a color version of this figure, see www.iste.co.uk/caltagirone/mechanics.zip*

Figure 1.1 shows the configuration of space–time in discrete mechanics; this model is borrowed from cosmology, where the present only makes sense for events unfolding at the origin. Any event that can influence or be influenced by an event unfolding at the origin is contained in the two cones whose summits are joined at the origin: the lower cone, which represents the past, and the upper cone, which represents the future. This light cone defines a so-called causal structure. For example, since the distance between the Earth and the Sun is large, we only receive light from the Sun 8 min after it was emitted. Any light signal emitted from the Earth would take just as long to reach the Sun. Events that occur during this period of time cannot be perceived by an observer; these events are said to be located elsewhere. In cosmology, the displacement $\mathcal{A}\mathcal{B}$ of a point in space–time is represented by the four-vector $(c\, dt, d\mathbf{x})$, and the present is restricted to the origin. In both discrete and continuum mechanics, the present unites all elements within a single causal structure; in Figure 1.1, the boundary of this structure is a circle, which we shall call the discrete horizon. Every event in this space is linked by cause and effect, the radius of the circle r_h is independent of time and every event within the circle is known instantaneously. Even if a specific observer located at some point of this space cannot directly perceive every event unfolding in the present instantaneously, the instantaneous field of all problem variables exists and can be represented by a mathematical model. Time is assumed to unfold linearly.

Figure 1.2 shows two spaces. The first has a finite horizon as its boundary, and the second is a sphere without a boundary but which is nonetheless finite; in both cases, all events unfolding on these surfaces are connected by the propagation of various types of waves through space. On the space with a boundary, events will necessarily be influenced by boundary conditions, which will be defined later. On the sphere, interactions will cumulate as the system evolves over time.

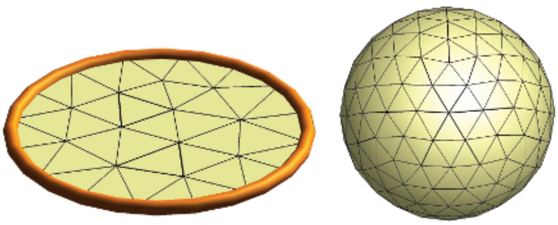

Figure 1.2. *Space with a discrete horizon as a boundary (left) and a sphere without a boundary but which is nonetheless finite (right)*

Thus, the discrete horizon defines a space on which separate phenomena can be described by a mathematical model at the same moment in time. For example, atmospheric models give an image of the present time and can be used to predict the weather over the next few days. Neglecting absorption, sound waves require over 300 h to travel the $40,000$ km of the circumference of the Earth, and light waves require slightly over 0.1 s. Even if an event at one point is not perceived instantaneously from another point, we can still construct an instantaneous image of the atmospheric currents. To make accurate predictions, we need a good representation of the present, which can be achieved by collecting a large amount of precise data. However, the chaotic and to some extent random nature of the turbulent evolution of atmospheric flows limits the prediction range of the model to just a few days.

The approach adopted by discrete mechanics draws heavily from the classical view of mechanics, where every interaction is defined directly. The interactions conventionally described as "actions at a distance", such as variations in gravity due to the Moon and Sun, are predictable and can be taken into account in the mathematical model. However, we cannot represent the cause-and-effect relations of more rapid events, such as the collapse of two black holes producing the gravitational waves predicted by general relativity.

Nevertheless, Newtonian mechanics is an alternative that is compatible with reality. Newtonian mechanics is widely thought to only be valid at velocities far below the speed of light, but in fact Newton's theory has remarkable properties when extended to the propagation of waves. The velocity is not the only relevant distinction between the relativistic and Newtonian theories of mechanics; relativistic kinematics and dynamics are other examples. Ultimately, the objective of the discrete perspective presented here is to investigate whether Newtonian mechanics is capable of describing all types of fluid and solid behavior, as well as the propagation of all types of waves.

1.1.2. *Notion of a discrete medium*

The perspective presented and explored here abandons the hypothesis of a continuum, which defined all of the problem variables, physical properties, etc., at every point. In continuum mechanics, Newton's law of dynamics, also known as Newton's second law, is formulated at a point. To express the spatial variation of the vector quantities on which the theory is based, we are forced to define the concepts of frame of reference and differentiation. Newton himself contributed to the development of infinitesimal calculus, even though he originally represented vectors as bipoints [NEW 90]. Later work expanded this continuous approach, which has various advantages, but also disadvantages that can generate artifacts.

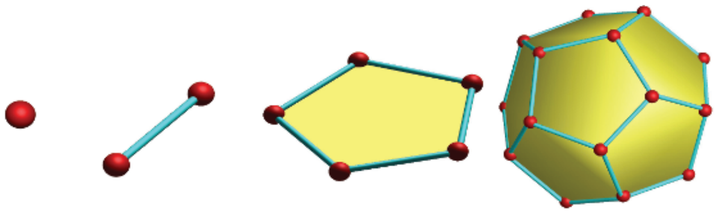

Figure 1.3. *Construction of a discrete medium from points, edges, surfaces and volumes*

The field of discrete mechanics is built upon connected objects such as those shown in Figure 1.3. First, we consider points that are not absolutely positioned within space; we shall work within a local frame of reference that positions objects relatively to one other. Two points and a straight line define an edge, or bipoint. This introduces two important ideas: the distance d between the points and a direction, also defined in relative terms. From multiple edges, we can construct a surface, which is necessarily planar; to represent non-planar surfaces, we can reduce them to triangles, which are planar by definition. Finally, by assembling multiple planar surfaces, we can construct volumes. The concept of dimension (one, two or three) is abandoned. For example, a (2D) plane constructed from three points simply defines two unit vectors: one associated with any given edge and the other normal to the plane. In three-dimensional space, this pair of vectors would necessarily be orthogonal. These topologies are all described as primal topologies; we shall then need to define a dual topology based on localizations such as the barycenter of a face or a volume.

Consider the elementary primal topology shown in Figure 1.4. This example will allow us to introduce notation for later. The three edges Γ with unit vector $\mathbf{t}$ form a planar surface $\mathcal{S}$ with normal vector $\mathbf{n}$; the vectors $\mathbf{t}$ and $\mathbf{n}$ are necessarily orthogonal, $\mathbf{t} \cdot \mathbf{n} = 0$. This condition is strictly necessary to establish the desired properties of the operators of the primal and dual topologies. The orientations of the unit vectors $\mathbf{t}$ and $\mathbf{n}$ are arbitrary but mutually dependent.

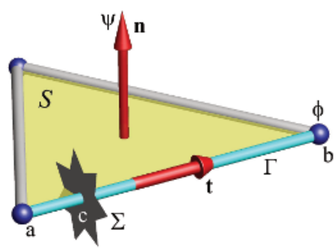

Figure 1.4. *The elementary topology of mechanics in a discrete medium: three straight edges Γ between vertices, defining a planar face $\mathcal{S}$. The normal vector $\mathbf{n}$ of this face and any of the vectors along Γ are orthogonal; in other words, $\mathbf{t} \cdot \mathbf{n} = 0$. The edge Γ can optionally be intercepted by a discontinuity Σ at the point c located between the vertices a and b of Γ. The quantities Φ and $\boldsymbol{\Psi}$ denote the scalar and vector potentials, respectively. For a color version of this figure, see www.iste.co.uk/caltagirone/mechanics.zip*

The edge Γ of length d lies between its two vertices a and b. This is the edge on which the equations of motion will later be defined. Scalar quantities such as the scalar potential ϕ are defined at points, whereas the vector potential is defined orthogonally to the plane $\mathcal{S}$. The plane Σ that intersects Γ at c divides the edge $[a, b]$ into two parts, for example representing two media of different types or with different physical properties. The interface Σ could represent a discontinuity, a shock wave, or a contact discontinuity. The velocity vector is defined on the edge Γ as a scalar constant directed from a to b.

The notion of a continuum is set aside by this approach. The topologies – the edge, surface and volume shown in Figure 1.3 – can never be reduced to a single point. The concept of differentiation is also abandoned. The measure d of the length of the edge Γ can be made arbitrarily small, but everything else scales proportionally, conserving angles. Thus, even though the medium is no longer continuous, we can nonetheless consider scales as small as permitted by the continuity properties of the macroscopic approach. The mean free path of individual molecules is always a lower bound for the scale.

1.2. Properties of discrete operators

The discrete operators are defined straightforwardly from the elementary topology in Figure 1.4. First, the discrete gradient is simply defined as a difference. For example, the gradient of the scalar potential ϕ is $\nabla \phi = (\phi_b - \phi_a)/d$. The gradient vector is not equivalent to the analogous concept in a continuum. Here, the gradient vector is an oriented scalar with direction $\mathbf{t}$. The primal curl $\nabla_p \times \mathbf{V}$ of the vector $\mathbf{V}$ is defined on the unit vector $\mathbf{n}$ as the circulation around the edges of the oriented surface $\mathcal{S}$. The divergence, for example of a vector, can be defined at a given point in terms of the (net) flux of every edge leading to or away from this point. The fourth operator that we shall define is the dual curl $\nabla_d \times \boldsymbol{\psi}$; the components of $\boldsymbol{\psi}$ are orthogonal to each primal surface $\mathcal{S}$.

The gradient and dual curl operators can be used to project the action of various phenomena onto each oriented edge Γ, which also serves as the basis for conservation of momentum and various other vector quantities, including the components of the velocity $\mathbf{V}$.

With this topological structure, some of the operators defined above are exact in the sense that the numerical error is zero. This is the case for the gradient, which is defined as a difference, as well as the primal curl, which is computed from Stokes' theorem as the line integral of the velocity vector around the contour Γ. The two other operators, the divergence and the dual curl, incur numerical errors whose magnitude depends on the quality of the mesh and the construction of the dual space.

Although classical mechanics relies heavily on the divergence theorem to redefine the flux of a surface onto a volume and then at a point, discrete mechanics derives the equations of motion from the fundamental theorem of analysis and corollaries such as Stokes' theorem. These key theorems are briefly recalled below. If $\mathbf{F}(x)$ is uniformly and continuously differentiable on $[a, b]$, then the fundamental theorem of analysis, also known as the fundamental theorem of integral and differential calculus, states that:

$$\left\{ \begin{array}{l} \mathbf{F}'(x) = \mathbf{f}(x) \\ \\ \displaystyle\int_a^b f(t)\, dt = F(b) - F(a) \end{array} \right. \qquad [1.1]$$

Stokes' theorem is a corollary of this result. It allows us to compute the curl of a surface as the line integral around a contour:

$$\int_{\Gamma} \mathbf{V} \cdot \mathbf{t} \, dl = \iint_{\Sigma} \nabla \times \mathbf{V} \cdot \mathbf{n} \, ds. \qquad [1.2]$$

The scope of this theorem is remarkable. We can compute the curl of a surface without explicitly knowing the velocity itself; we simply need to know its components on some closed contour. Armed with this result, the concept of a frame of reference, needed to define the velocity vector at specific points in absolute terms, becomes less essential. Furthermore, since the curl is not defined at points or on curves, this operator can only be defined by taking limits in continuum mechanics.

The divergence theorem, also known as the Green–Ostrogradski theorem, is another corollary of Stokes' theorem that allows us to quantify the source or sink behavior of any quantity defined at a point:

$$\iint_{\Sigma} \mathbf{V} \cdot \mathbf{n} \, ds = \iiint_{\Omega} \nabla \cdot \mathbf{V} \, dv. \qquad [1.3]$$

Since the discrete equations of motion are derived on an edge Γ, we need to know how to interpret the product of two functions on this edge. The generalized mean value theorem for definite integrals (which follows from Rolle's theorem) states that $\exists\, c \in [a, b]$ such that:

$$\int_a^b f(x)\, g(x)\, dx = f(c) \int_a^b g(x)\, dx.$$

This theorem, which only holds in one dimension, will be extremely useful to us. When working with two-phase flows, this theorem can be used to bound the density in terms of its values at the endpoints a and b.

Two key properties from continuum mechanics, $\nabla \times \nabla p = 0$ and $\nabla \cdot \nabla \times \mathbf{V} = 0$, will prove especially important with discrete media. Labeling discrete quantities with the subscript h, we can easily show that the discrete curl of a discrete gradient is zero on the primal topology:

$$\begin{cases} \displaystyle\int_a^b \nabla p \cdot \mathbf{t} \, dl = p_b - p_a \\ \displaystyle\int_\Gamma \nabla p \cdot \mathbf{t} \, dl = 0 \\ \displaystyle\iint_{\mathcal{S}} \nabla \times \left(\nabla p\right) \cdot \mathbf{n} \, ds = 0 \\ \nabla_h \times \left(\nabla_h \, p\right) = 0 \end{cases} \qquad [1.4]$$

Similarly, the discrete divergence of the discrete primal curl computed on the dual volume is zero:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} \Gamma_i = \sum_{i=1}^{n} \iint_s \nabla \times \mathbf{V} \cdot \mathbf{n} \, ds = 0 \\ \displaystyle\iint_{\mathcal{S}} \left(\nabla \times \mathbf{V}\right) \cdot \mathbf{n} \, ds = 0 \\ \displaystyle\iiint_{\mathcal{V}} \nabla \cdot \left(\nabla \times \mathbf{V}\right) ds = 0 \\ \nabla_h \cdot \left(\nabla_h \times \mathbf{V}\right) = 0 \end{cases} \qquad [1.5]$$

Figure 1.5 illustrates why the property $\nabla \times \nabla \phi = 0$ holds on the primal topology, and why $\nabla \cdot \nabla \times \boldsymbol{\psi} = 0$ holds on the dual topology.

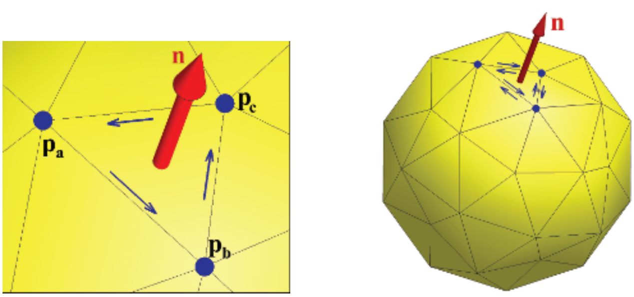

Figure 1.5. *Properties of operators:* $\nabla \times \nabla \phi = 0$ *(left);* $\nabla \cdot \nabla \times \boldsymbol{\psi} = 0$ *(right)*

These properties allow the discrete equations of motion to express the acceleration as the sum of a gradient and a curl, or in other words as a Hodge–Helmholtz decomposition. Hodge–Helmholtz decomposition is usually encountered when solving the Navier–Stokes equations, where it is used to separate the irrotational and solenoidal parts of various vectors. In this way, the velocity correction associated with the irrotational term can be used to construct a divergence-free field [ANG 12, CAL 15c]. More generally, Hodge–Helmholtz decomposition gives two orthogonal terms satisfying certain boundary conditions that are useful in various other fields, such as imaging, fingerprint recognition and so on. Since any vector can be decomposed into these two components, it often makes sense to find expressions for each component of a physical vector *a priori*. We can do this for both the velocity and the acceleration.

Some operators and combinations of the above operators, including the Laplacian $\nabla^2 \phi = \nabla \cdot \nabla \phi$ or $\nabla^2 \boldsymbol{\psi} = \nabla \cdot \nabla \boldsymbol{\psi}$, will not be used; they can lead to artifacts or, for vectors, increase the tensor order of the operators. There is no suitable representation for tensors of orders two or higher in the discrete setting. For example, the gradient of a vector $\nabla \mathbf{V}$, which has a clearly defined meaning in a continuum, cannot be represented in discrete mechanics. One of our key objectives is therefore to determine whether the operators cited above suffice to describe all possible physical behavior.

It might not be immediately obvious whether tensors are strictly necessary to describe all behavior and physical phenomena observed at macroscopic levels. Second-order tensors provide a valid description of the behavior of anisotropic materials, but these tensors do not necessarily reflect the underlying laws of mechanics. Solid mechanics, general relativity, fluid mechanics and other fields frequently include higher order tensors in their laws. For example, in solid mechanics, shear stresses are defined as the gradient of a vector; however, they could alternatively be expressed just in terms of the curl. Over time, the distinction between constitutive laws and the fundamental laws describing the underlying physics has faded. The laws of physics can in fact be phrased equivalently in terms of vectors or tensors. Maxwell's equations are a good example of this; each formulation has specific advantages and disadvantages.

1.3. Invariance under translation and rotation

The equations of motion must satisfy Galileo's principle of relativity; we should not be able to distinguish between the physical phenomena of a system in uniform motion and those of the same system at rest. If an object is held by an observer aboard a train traveling at velocity $\mathbf{V}_t$, it will fall at his or her feet if dropped without acceleration. The velocity of the uniform motion must therefore somehow "cancel out" in the equations of motion. The problem of uniform rotation is similar but trickier than linear motion. In particular, any shear stresses in a uniformly rotating object must be independent of the frame of reference. This is known as the principle of material frame-indifference and was introduced by Truesdell [TRU 65]. Consider a translational motion $\mathbf{V}_t$ and a rotational motion $\mathbf{V}_r$:

$$\begin{cases} \mathbf{V}_t = u_0\,\mathbf{e}_x + v_0\,\mathbf{e}_y + w_0\,\mathbf{e}_z \\ \\ \mathbf{V}_r = \mathbf{\Omega} \times \mathbf{r} \end{cases} \tag{1.6}$$

where (u_0, v_0, w_0) are the constant components of the translation vector and $\mathbf{\Omega}$ represents the uniform rotation.

Let us now apply the discrete operators with the same properties as their differential counterparts to these vectors. In particular, the properties $\nabla \cdot \nabla \times \mathbf{V} = 0$ and $\nabla \times \nabla f = 0$ are satisfied, where $\mathbf{V}$ is a vector and f is a scalar. Applying these operators to these uniform velocities yields:

$$\begin{cases} \nabla \cdot \mathbf{V}_t = 0 \qquad \nabla \times \mathbf{V}_t = 0 \\ \\ \nabla \cdot \mathbf{V}_r = 0 \\ \\ \nabla \times \mathbf{V}_r = 2\,\mathbf{\Omega} \\ \\ \nabla \times \mathbf{\Omega} = 0 \end{cases} \tag{1.7}$$

The case of translational motion is straightforward. The divergence and primal curl of $\mathbf{V}_t$ are identically zero; if a harmonic field is present, it can be eliminated in the same way. The divergence of the rotation field is indeed zero, but the primal curl is merely constant throughout the domain, provided that the latter is connected. The circulation around an arbitrary contour (that does not enclose the origin) reduces to the curl of the rotational velocity, which is constant. By equation [1.7], the dual curl of the primal curl is zero. At least for some of the terms of the equations of motion, we can therefore hope to establish complete invariance between frames of reference. The principle of material frame-indifference cited by Truesdell is automatically satisfied if the shear stresses are expressed in terms of these discrete operators.

The non-linearity of the inertial term in the equations of motion creates extra terms. If we split the velocity field into $\mathbf{V} = \mathbf{V}' + \mathbf{V}_r = \mathbf{V}' + \mathbf{\Omega} \times \mathbf{r}$, the acceleration can be written as:

$$\begin{cases} \gamma = \dfrac{d\mathbf{V}}{dt} = \dfrac{\partial \mathbf{V}}{\partial t} - \mathbf{V} \times \nabla \times \mathbf{V} + \dfrac{1}{2} \nabla \left|\mathbf{V}\right|^2 \\ \\ \gamma = \dfrac{d\mathbf{V}'}{dt} = \dfrac{\partial \mathbf{V}'}{\partial t} - \mathbf{V}' \times \nabla \times \mathbf{V}' + \dfrac{1}{2} \nabla \left|\mathbf{V}'\right|^2 + 2\,\mathbf{\Omega} \times \mathbf{V}' + \mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r} \end{cases} \tag{1.8}$$

where $2\,\mathbf{\Omega} \times \mathbf{V}'$ is the Coriolis acceleration and $\mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r}$ is the centrifugal acceleration. The centrifugal term $\mathbf{\Omega} \times \mathbf{\Omega} \times \mathbf{r}$ derives from a scalar potential and is equal to $-\nabla \left(|\mathbf{\Omega} \times \mathbf{r}|^2 /2 \right)$. For uniform rotational motion, the centrifugal acceleration is locally and instantaneously opposite to the centripetal acceleration.

The derivatives $d\mathbf{V}/dt = d\mathbf{V}'/dt$ are identical in an absolute frame of reference and a moving frame. The Coriolis and centrifugal accelerations only need to be considered when changing between frames of reference. In Newtonian mechanics, interactions between two frames are assumed to be instantaneous. General relativity is required for a more realistic description of actions at a distance. Here, in the discrete setting, we shall work in a local frame, and every interaction is defined by the wave celerity, which is necessarily finite. Any uniform translational celerity $\mathbf{V}_t$ or rotational motion $\mathbf{V}_r$ is imperceptible and cannot be detected by an observer attached to the local frame. This is both a disadvantage and an opportunity; we cannot describe absolute motion, but on the other hand our velocity and position with respect to the universe ultimately do not matter anyway. In our approach, the universe will, therefore, be limited to a horizon that is described as discrete (Figure 1.1). Inertial analysis plays a special role in discrete analysis and is discussed in Chapter 2.

1.4. Weak equivalence principle

Since Galileo's time, various experiments have been performed to investigate the action of gravity on two different types of mass that appear to be accelerated identically, regardless of the internal structure or composition of the object being accelerated – this is the equality between gravitational mass and inertial mass (also known as the Weak Equivalence Principle, or WEP). Einstein would later extend the WEP into a stronger equivalence principle that relates to the fact that the velocity is always bounded by the celerity of light in a vacuum in special relativity.

Today, the weak equivalence principle has been repeatedly verified by an impressive array of experiments. The equivalency is quantified by the Eötvös ratio $\eta = 2\,|\gamma_1 - \gamma_2| \,/\, |\gamma_1 + \gamma_2|$, where γ_1 and γ_2 are the accelerations of the two masses. The acceleration can be measured independently from any frame of reference to extremely high accuracy. The modern view is that the WEP holds exactly, with an Eötvös ratio of $\eta < 10^{-14}$; the article by Will [WIL 09] presents the various experiments that have been conducted over the last century. Other experiments led by France and the United States are also currently underway. It is hoped that they will achieve even greater levels of precision (10^{-15} or 10^{-18}) to confirm (or refute) the exactness of the Weak Equivalence Principle (WEP).

The equivalence principle allows us to rewrite the fundamental law of dynamics established by Newton in his *Principia* [NEW 90] in the following form:

$$\boldsymbol{\gamma} = \mathbf{g}. \qquad [1.9]$$

This equality might seem self-evident, since an isolated observer cannot distinguish the effect of gravity from his own specific acceleration. The equality had been known since Galileo, but Newton formulated his second law as $m\,\boldsymbol{\gamma} = \mathbf{F}$ even though gravity was the most important force at the time. This law would prove to be somewhat problematic in the field of electromagnetism, but would be adopted by the field of dynamics nonetheless.

The next question is the underlying meaning of the force per unit mass $\mathbf{g}$. Is it exclusive to gravitational forces, or can we view $\mathbf{g}$ as the sum of the forces acting upon any particle of matter? The distinction is irrelevant anyway if gravity is the only force exerted by a body of mass M, viewed as a point, on a particle with (possibly zero) mass m. In this case, the acceleration due to gravity is equal to $-\mathcal{G}\,M/r^2$, derived from the potential $\mathcal{G}\,M/r$, where $\mathcal{G}$ is the universal gravitational constant and r is the distance between the particle and the body of mass M. If the gravitational force exerted by a body reduced to a point mass M is small, we can consider the Taylor expansion at $\mathcal{G}\,M/r\,c^2 << 1$:

$$g_{00} = -1 + 2\,\frac{\mathcal{G}\,M}{r} - 2\left(\frac{\mathcal{G}\,M}{r}\right)^2 + \ldots. \qquad [1.10]$$

Truncating to first order gives the Newtonian potential, and truncating to second order yields a correction term that has order of magnitude 10^{-6} in our solar system. This so-called post-Newtonian theory is able to explain the deviation of photons as they pass near the Sun, which is twice the deviation expected by purely Newtonian mechanics. The deflection of light near the Sun measured by Eddington in 1919 [DYS 20] during an eclipse was the first ever confirmation of the theory of general relativity proposed a few years earlier by Einstein. This theory is phrased in a purely geometric form that isolates the interactions of gravity from the other interactions; it has not been successfully unified with other theories of physics. Special relativity postulates the existence of a maximum velocity, the celerity of light c in a vacuum; this celerity is built into the Lorentz transformation at the origin, and the factor $\gamma = 1/\sqrt{1 - v^2/c^2}$ is used to express relativistic coordinates, velocities, and accelerations, as well as the laws of relativistic kinematics and dynamics.

Some results from fluid dynamics, such as the existence of supersonic flows, for example in de Laval nozzles, show that the material velocity of the medium itself can exceed the local wave velocity (celerity) of sound. Similarly, in media with celerity lower than the speed of light in a vacuum, such as water, charged particles can move faster than the material velocity of the medium; the Vavilov-Cherenkov effect observed in underwater nuclear reactors is one example. These phenomena, where the material velocity of the medium exceeds the local celerity of waves traveling through the medium, do not contradict the absolute upper bound postulated by the theory of relativity, namely the speed of light in a vacuum. Other effects from the field of cosmology are not yet fully understood. One striking example is the cosmic inflation that occurred in the first few moments after the birth of the Universe when its expansion was extremely high. In the 20th century, quantum mechanics discovered wave-particle duality – light for example can be characterized either by its wavelength or as a photon traveling at the speed of light. Discrete mechanics offers an alternative perspective; the local celerity (a scalar) is a characteristic quantity of the problem and the material velocity of the medium (a vector) is an unbounded variable.

1.5. Principle of accumulation of stresses

The evolution of the state taken by a physical system depends on the path followed by the system and the stresses applied to it. The idea of the discrete approach is to describe the evolution of each of the quantities making up the state of the system: velocity, acceleration, pressure, shear stress, flux, temperature and so on. Each variable is only known at time $t + dt$ if it is known at time t (Figure 1.6). We shall label each quantity with an exponent, for example $\mathbf{V}^o$ denotes the velocity at time t^o. Time unfolds uniformly in the positive direction over the full space bounded by the discrete horizon. Our knowledge of the prior state of the system is both useful and inconvenient: we cannot construct a model that gives us the intrinsic state of the system, but we can identify the only possible state from a set of multiple solutions. We need to know the history of the system to predict its future evolution. One description of this is that the medium has continuous memory.

The most illustrative example is the velocity: we can find the velocity by integrating the acceleration, $\mathbf{V} = \gamma \, dt + \mathbf{V}^o$, and we can differentiate the velocity to deduce the acceleration, but these two operations are fundamentally different; when we differentiate, the constant $\mathbf{V}^o$ is permanently lost, and when we integrate, the velocity $\mathbf{V}$ is only known up to a constant, which is why we need to know the velocity at time t^o. The velocity is not bounded by any fixed or predefined value such as the celerity of the medium. The velocity only converges to a limit because the acceleration tends to zero. The acceleration and the velocity are both vectors, but the velocity is a Lagrangian that needs to be aggregated, whereas the acceleration is an absolute quantity. Like the velocity $\mathbf{V}$, the quantities of length, density ρ^o, pressure

p^o, shear stress $\boldsymbol{\omega}^o$, temperature T^o, etc., are updated from the acceleration or the thermal flux $\boldsymbol{\Phi}^o$.

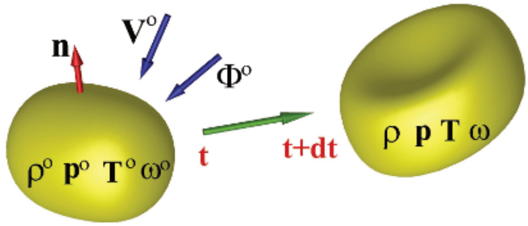

Figure 1.6. *Principle of accumulation of stresses over time from t to $t + dt$. For a color version of this figure, see www.iste.co.uk/caltagirone/mechanics.zip*

The description of motion in Galileo's principle of relativity is phrased in terms of the velocity. Every phenomenon is preserved up to uniform and constant velocity; since Galileo, velocity has stood at the heart of every theory of mechanics, with the acceleration being viewed as just its material derivative. Mass was later associated with the acceleration to introduce the notion of a force, and momentum was developed into a key aspect of contemporary mechanics.

The discrete perspective presented here is of course consistent with this principle of relativity, as well as the principle of weak equivalence, also introduced by Galileo, as an equality between the gravitational mass and the inertial mass. The weak equivalence principle, which only applies to gravitational effects in general relativity, is extended to other phenomena in the discrete approach, rephrasing the laws of dynamics into a law of conservation of acceleration. The current velocity is only known if the velocity $\mathbf{V}^o$ of the mechanical equilibrium at time t^o is known; as a result, it no longer makes sense to consider inertial frames of reference to describe motion.

1.6. Duality-of-action principle

Since we do not wish to consider frames of reference in discrete mechanics, we must replace them by a more local concept that is compatible with the perspective that we have adopted. The acceleration γ is only defined on the edge Γ with unit vector $\mathbf{t}$; we must therefore ensure that all physical effects (viscosity effects, compressibility effects, inertia, gravity, capillary effects, etc.) can also be defined on this edge. Some of these effects, including the last three listed above, will be defined by a source scalar potential ϕ_s that is known at every point but localized on the discrete mechanical stencil in practice (Figure 1.4). The gravitational potential ϕ_g, for example, is defined everywhere in the universe, but others such as the capillary potential are only defined on an interface Σ separating two media. From the physicist's perspective, we do not always fully understand the existence of a potential even if we are able to model how it manifests; for example, we know how to model the gravitational attraction between massive bodies even if we do not properly understand the underlying reasons why they attract each other – perhaps because of the graviton introduced by quantum mechanics. In discrete mechanics, the existence of the gravitational potential ϕ_g is indisputable, much like the existence of various other potentials.

Accordingly, if we assume the existence of an arbitrary potential ϕ_s, what action does it exert upon the motion of a massive or massless body? The hypothesis proposed here is that every potential has both a direct effect and a "dual" effect; this hypothesis, formulated as a postulate, can only be truly accepted once it withstands the test of time, at which point it can be viewed as an axiom. Since every vector can be represented as the sum of the gradient of a scalar potential and the curl of a vector potential, we can construct consistent formulations for each of the various effects captured by the equations of motion. The direct effect will be represented by the component of the gradient that lies in the direction of the unit vector $\mathbf{t}$, and the dual effect will be represented by the dual curl of the corresponding vector potential $\boldsymbol{\psi}_s = (\phi_s \, \mathbf{n})$, where $\mathbf{n}$ is the unit vector of the primal plane. The acceleration induced by this potential can then be expressed as follows:

$$\boldsymbol{\gamma}_s = \nabla \phi_s + \nabla \times (\phi_s \, \mathbf{n}) \, . \qquad [1.11]$$

This description provides a perfectly consistent discrete formulation that presents every effect modeled by the equations of motion as a discrete Hodge–Helmholtz decomposition.

Now, without going into too much detail in each case, we can already define three scalar potentials, ϕ_g for gravity, ϕ_i for inertia [CAL 17] and ϕ_c for capillary effects:

$$\begin{cases} \phi_g = \dfrac{\mathcal{G} \, M}{r} \\[2ex] \phi_i = \dfrac{1}{2} \, |\mathbf{V}|^2 \\[2ex] \phi_c = \sigma \, \kappa \end{cases} \qquad [1.12]$$

where $\mathcal{G}$ is the universal gravitational constant, M is the mass of the body, $|\mathbf{V}|^2/2$ is the kinetic energy, σ is the surface tension per unit mass and κ is the curvature. The accelerations $\boldsymbol{\gamma}_g$, $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_c$ may then be defined as follows on the edge Γ:

$$\left\{ \begin{aligned} \boldsymbol{\gamma}_g &= \nabla \left(\frac{\mathcal{G}\, M}{r} \right) + \nabla \times \left(\frac{\mathcal{G}\, M}{r}\, \mathbf{n} \right) \\ \boldsymbol{\gamma}_i &= \nabla \left(\frac{1}{2} \left| \mathbf{V} \right|^2 \right) + \nabla \times \left(\frac{1}{2} \left| \mathbf{V} \right|^2 \mathbf{n} \right) \\ \boldsymbol{\gamma}_c &= \nabla \left(\sigma\, \kappa \right) + \nabla \times \left(\sigma\, \kappa\, \mathbf{n} \right) \end{aligned} \right. \qquad [1.13]$$

Recall that the acceleration $\boldsymbol{\gamma}$ of a body is the only physical vector that exists in an absolute sense. It is equal to the sum of the viscosity and compressibility effects plus the accelerations in equation [1.13].

The gradient operator accounts for any unidirectional effects, such as the action of the Earth's gravity on a body located at its surface; the dual curl operator represents any (possibly polarizable) transverse effects, such as gravity waves. In general, the direct and dual effects might be zero, additive or might cancel each other out. Each of these cases will be examined in more detail in subsequent chapters.

1.7. Physical characteristics of a medium

The system of equations of the discrete formulation does not involve any state equations or any equations describing the rheological behavior of the model; the physical characteristics of the model must simply be known. Their localizations (point, edge, face, volume) on the topology from Figure 1.4 vary from operator to operator. For instance, the density on an edge is denoted ρ_v, whereas the density of the point a is the usual scalar ρ. The standard notion of differentiation (on a continuum) encounters various obstacles in discrete settings, where the discontinuities prevent us from simply differentiating functions as usual. Introducing equations at the discontinuities does not entirely resolve the problem; jump equations are simply conditions that must be satisfied, without truly offering a solution that reflects the continuous variables nearby. The continuity or conservation of mass equation illustrates the typical difficulties that we encounter if we attempt to define a point density ρ and an edge density ρ_v simultaneously.

In continuum mechanics, the two quantities ρ and ρ_v are conflated and cannot be distinguished. This makes it easy to compute the partial derivative in both conservative and non-conservative problems:

$$\frac{\partial \rho}{\partial t} = -\rho \, \nabla \cdot \mathbf{V} - \mathbf{V} \cdot \nabla \rho = -\nabla \cdot (\rho \, \mathbf{V}) \, . \qquad [1.14]$$

The above equation is closely related to the equations of motion from fluid mechanics. The density is viewed as a variable in its own right. In discrete mechanics, the density term ρ_v appears in the equations of motion, but the variable that appears in the equation of conservation of mass is the scalar ρ:

$$\frac{d\rho}{dt} = -\rho \, \nabla \cdot \mathbf{V}. \qquad [1.15]$$

Let us examine the differences between equations [1.14] and [1.15]. The first equation does not make sense in a discrete setting; the density is evaluated at a point, the velocity component $\mathbf{V}$ is evaluated on an edge, and $\mathbf{V} \cdot \nabla \rho$ is then assigned to a point. The gradient of the density, a discontinuous function, is not defined. The conservative form $\nabla \cdot (\rho \, \mathbf{V})$ does not make sense either – the density ρ should be evaluated on edge, but the divergence assigns the product to a point. However, equation [1.15], which gives the Lagrangian formulation of the conservation of mass, does still make sense, since $\rho \, \nabla \cdot \mathbf{V}$ is calculated at the same point as ρ and $\nabla \cdot \mathbf{V}$, and the result is assigned to a point.

Thus, the equations of motion and the equations of conservation of mass must be distinguished in discrete mechanics. The continuity equations are only used to express the evolution of the density in the Lagrangian formulation, whereas the conservation of the acceleration is autonomous. For the equations of motion to be autonomous, the pressure must clearly be described as a direct function of the velocity itself. There are no other constitutive laws that might introduce inconsistencies elsewhere in the problem; the physical properties are simply assumed to be known at each moment in time.

The evolution of the density and the velocity over time, as well as that of any other accumulation-based quantities, is deduced by advection at velocity $\mathbf{V}$. Since we cannot quantify this with $\mathbf{V} \cdot \nabla \rho$, we need to find some other way to apply a translation to each quantity on the edge Γ for a period dt and compute the density ρ_v.

For a two-phase flow with separate phases, ρ_v can be computed by the mean value theorem on the edge Γ, which has length d and endpoints a and b. The velocity $\mathbf{V}$ is assumed to be constant on this edge:

$$\frac{1}{d} \int_a^b \rho \, \mathbf{V} \, dx = \frac{\rho_v}{d} \int_a^b \mathbf{V} \, dx = \rho_v \, \mathbf{V}. \qquad [1.16]$$

Suppose that the interface is at c. If $x < c$, then $\rho = \rho_a$, and if $x > c$, then $\rho = \rho_b$. We find that:

$$\rho_v = \frac{1}{d} \left(\int_a^c \rho_a \, dx + \int_c^b \rho_b \, dx \right). \qquad [1.17]$$

After defining a partition function $\psi \in [0, 1]$, we have that:

$$\rho_v = \psi \, \rho_a + (1 - \psi) \, \rho_b, \qquad [1.18]$$

where $\psi = [ac]/[ab]$.

Any thermodynamic coefficients can simply be defined directly in terms of the mass, volume, heat, etc. We shall define the following quantities classically:

– the isochoric expansion coefficient α;

– the isobaric expansion coefficient β;

– the isothermal compressibility coefficient χ_T;

– the specific heat at constant pressure and constant volume, c_p and c_v;

– the latent heat from a phase change at constant temperature L.

The definitions of these quantities are recalled in the following:

$$\left\{ \begin{array}{lll} \alpha = \dfrac{1}{p} \left(\dfrac{\partial p}{\partial T} \right)_\rho & \beta = -\dfrac{1}{\rho} \left(\dfrac{\partial \rho}{\partial T} \right)_p & \chi_T = \dfrac{1}{\rho} \left(\dfrac{\partial \rho}{\partial p} \right)_T \\ c_p = \left(\dfrac{\partial h}{\partial T} \right)_p & c_v = \left(\dfrac{\partial e}{\partial T} \right)_\rho & L = \dfrac{T}{\Delta \rho} \dfrac{dp}{dT} \end{array} \right.$$

We know that:

$$\left(\frac{\partial \rho}{\partial T} \right)_p \left(\frac{\partial T}{\partial p} \right)_\rho \left(\frac{\partial p}{\partial \rho} \right)_T = -1,$$

and so these coefficients satisfy the relation:

$$\frac{\beta}{\alpha\, p\, \chi_T} = 1.$$

The notion of tensor is needed here if we wanted to express the variation of a quantity at a point in terms of the direction of observation. However, if the direction of observation Δ is fixed, the various quantities, mechanical properties, stresses, etc., can be simply described as scalars or oriented vectors on Γ. Some materials have tensor properties, such as the thermal conductivity, the permeability, and so on, as well as certain mechanical properties. Like the components of the heat flux, displacement, or velocity, any tensor quantities are defined on the edge Γ by projection.

For example, consider the diffusion of the heat flux $\mathbf{\Phi}$ in an anisotropic material. The matrix $[\mathbf{\Lambda}]$ representing the thermal conductivity tensor is diagonalizable. The eigenvectors define the principal directions of the tensor $\mathbf{\Lambda}$; the matrix is diagonal, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$, with respect to its basis of eigenvectors. The flux on the edge $\mathbf{t}$ can be formally defined as $\mathbf{\Phi} \cdot \mathbf{t} = -k\,(T_R - T_P)/L$, where k is the scalar representing the conductivity on the edge and $(T_R - T_P)/L$ is the discrete gradient. We can now identify the heat flux along $\mathbf{t}$ and compute the value of k:

$$k = \boldsymbol{\lambda} \cdot \mathbf{t}, \qquad [1.19]$$

where $\boldsymbol{\lambda}$ is written with respect to the basis of eigenvectors. This property is constant over the whole edge. Note that any quantity defined on an oriented edge with unit vector $\mathbf{t}$ can equivalently be written either as a scalar that is constant over the edge or as an oriented vector.

These constitutive laws, state laws, etc., are only required to describe the behavior of a medium, whether fluid or solid, in terms of specific scalar or vector variables, such as the temperature, the pressure, the mechanical stress and so on. If the properties of the medium vary as a function of the direction (the anisotropic case), it can be useful to summarize the behavior of the medium with a symmetric tensor, which allows us to easily compute the stress in any direction. However, this tensor is always defined by a basis, typically an orthonormal basis, and therein lies the problem. The extremely general nature of this approach introduces difficulties that will need to be addressed later, for example using the principle of material frame-indifference, which allows the tensor properties to be introduced into the conservation equations.

The Cauchy stress tensor defined at a point has six independent coefficients, which are expressed in terms of the velocity or the displacement. This results in 81 coefficients for the elasticity tensor in the stress–strain relationship defined by

Hooke's law. In an isotropic medium, the number of coefficients is reduced to just two, the Lamé coefficients. Whether these Lamé coefficients are representative depends on whether the material is fluid or solid.

This robust connection between the constitutive equations and the conservation equations can, however, be broken without undermining the representativeness of the model constructed from the fundamental equations of motion. Any anisotropy or inhomogeneity in the medium does not directly affect these fundamental equations. The physical properties – the viscosity of the primal surface and the compressibility coefficient defined at the endpoints of each edge – are functions of the scalar variables of pressure, temperature and density, or vector variables such as the vector potential. Anisotropy is handled analogously to inhomogeneity, viscosity is defined on planes, and compressibility is defined at points; these properties can vary over space and of course over time.

1.8. Composition of velocities and accelerations

Consider a particle following a rectilinear trajectory along some axis Δ. At time t^o, the velocity of this particle is given by $\mathbf{V}^o$ and can be viewed as a scalar V^o defined on the oriented axis Δ; the acceleration is assumed to be constant and equal to γ. After a period dt has elapsed, the velocity of the particle is $\mathbf{V} = \mathbf{V}^o + \gamma \, dt$, provided that dt is sufficiently small – it can be deduced by integrating with respect to time if not. In discrete mechanics, this period dt is chosen to be as small as necessary. Suppose now that an additional acceleration $\boldsymbol{\delta}$ is applied to the particle, say $\boldsymbol{\delta} = \gamma$: the acceleration of the particle is now equal to 2γ and its velocity at time $t + dt$ is $\mathbf{V} = \mathbf{V}^o + 2\gamma dt$. Thus, when the acceleration is doubled, the velocity is modified by an amount that depends on dt. Although the acceleration and the velocity are both vectors, they perform completely different roles in mechanics.

This distinction is essential if we hope to understand the laws of physics, especially when summing vectors. In mathematics, the sum of two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is simply defined as the vector $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$. This still holds in Newtonian mechanics, but in relativistic mechanics a Lorentz transformation is required, meaning that the equation cited above no longer holds. Acceleration and velocity are independent concepts. The sum of two accelerations is always *a priori* equal to $\gamma_1 + \gamma_2$ in a single frame of reference but the same is not true for two velocities. For example, consider a single particle: it can be acted upon by two accelerations simultaneously, but it can only have one single instantaneous velocity. We need to distinguish between the quantities applied to the particle, the external accelerations and the internal quantities of the particle, its own acceleration and velocity.

Some of the thought experiments performed in the early 20th Century explored the idea of two trains passing each other by. Newtonian mechanics needed to be questioned in light of the newly proposed postulate that the velocity is necessarily bounded by the speed of light in a vacuum. This postulate is now widely accepted but is not an axiom. The trains considered by these thought experiments moved at constant speed, so with zero acceleration, and the fact that they are passing by one another is not especially significant if they do not interact, for example in a two-way tunnel. To conduct thought experiments with two particles, the causality principle must hold; in other words, the particles must interact and share the same discrete space–time. Therefore, we shall assume axiomatically that the accelerations sum additively. By contrast, the velocities are additive, free from any constraints and shall be viewed as secondary variables. This illustrates the benefit of formulating the equations of motion in terms of the acceleration rather than the velocity or the displacement. Any constraints on the velocity necessarily arise from the fact that the acceleration tends to zero; if so, we need to know the velocity at previous times to know the velocity at the present time.

Consider the case of steady motion where the particle has zero acceleration, $\gamma = 0$. By Newton's first law, there is no force acting on the particle, or more precisely the sum of all forces acting upon the particle is zero; we can rephrase this in terms of the acceleration: "Every body remains at rest or in uniform rectilinear motion unless acted upon by some acceleration which forces it to change its state". More precisely, it is the sum of the accelerations that must be zero: for example, in the case of a rigidly rotating flow with constant velocity, the non-zero centripetal and centrifugal accelerations compensate one another exactly as two equal gradients of scalar potentials. The velocity $\mathbf{V} = \mathbf{\Omega} \times \mathbf{r}$ is constant in the absence of external action. This solution can be superimposed with any other solution without modifying the latter, provided that there are no strong interactions between the two. As well as constant translational motion (Galilean frame of reference), this rotational motion needs to be eliminated from the field of solutions of the equations of motion; in fact, this can only be done up to the gradient of a scalar potential. Two theoretical difficulties need to be addressed: the mathematical behavior at infinity, as well as the physical notion of time, which is necessary for this type of flow to arise. Indeed, to generate this flow, we must take the transverse speed to be infinite or wait for an infinite period of time.

Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two vectors viewed as mathematical objects, and let $\mathbf{V}_1$ and $\mathbf{V}_2$ be two material velocities. The mathematical and mechanical composition rules of these vectors are not the same:

– Galilean transformation:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2; \quad [1.20]$$

– relativistic transformation:

$$\mathbf{V} = \frac{\mathbf{V}_1 + \mathbf{V}_2}{1 + \dfrac{\mathbf{V}_1\, \mathbf{V}_2}{c^2}}. \qquad [1.21]$$

The second relation is more general.

However, there is an issue – we are comparing a material velocity against a celerity (wave velocity). It does not seem justifiable to conflate the velocity and the celerity in a term of the form $\mathbf{V}/c$, since at a fundamental level these quantities relate to completely different concepts, despite having the same units. The celerity is an intrinsic property of the medium under fixed conditions (solid, gas, vacuum); it determines the velocity of the wavefront. The material velocity, on the other hand, governs the advection of a physical quantity, whether massive or massless. The two concepts are quite simply distinct, yet the Lorentz transformation unites them into a dimensionless ratio that is assumed to be universally valid.

By contrast, the acceleration of a particle is always equal to the sum of the accelerations applied to this particle: $\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2$. The composition of accelerations simply reduces to the conservation of acceleration, i.e. Newton's second law. For example, the local acceleration of an object subject to gravitational attraction from both the Earth and the Moon is equal to the vector sum of the accelerations induced by each body.

The relativistic composition rule for the acceleration $\overline{\boldsymbol{\gamma}}$ involves the Lorentz factor γ, whenever it is specified in terms of the velocity:

$$\overline{\boldsymbol{\gamma}} = \boldsymbol{\gamma} \left(\frac{\sqrt{1 - \frac{V^2}{c^2}}}{1 - \frac{v\, V}{c^2}} \right)^3 . \qquad [1.22]$$

This expression gives a description of the acceleration and may differ from one inertial frame to another. With a local frame, the law of dynamics $F = m\boldsymbol{\gamma}$ is identical in both Newtonian and relativistic mechanics, provided that m is the mass that is undergoing the motion.

The velocity $\mathbf{V}$ is in fact a Lagrangian that is updated by the acceleration $\mathbf{V} = \mathbf{V}^o + \boldsymbol{\gamma}\, dt$. Like the pressure, the density, the temperature or the shear-rotation stress, the velocity is computed from its previous value at time t^o. Adopting the perspective that the velocity is the accumulation of the acceleration offers one important advantage; it eliminates unphysical solutions such as those which arise from superimposing velocities exceeding the speed of the light; the velocity $\mathbf{V}^o$ is a mechanical equilibrium and $\mathbf{V}$ is another equilibrium state, since the acceleration vanishes as the velocity approaches its limit value, the celerity c. Moreover, the velocity does not appear directly in the expression of the acceleration but only indirectly via operators such as the divergence and the curl that eliminate any uniform motion.

1.9. Discrete curvature

The concept of curvature is essential in mechanics; it is encountered with various phenomena such as inertia, gravity, capillarity and so on. The notions of principal curvature, mean curvature, Gaussian curvature, etc., were originally introduced in planar geometry in an intuitive form before being formalized into a rigorous mathematical framework for more general spaces. Curvature is defined by the Riemann curvature tensor; this tensor specifies the acceleration at which two neighboring geodesics split from one another on a curved space. In general relativity, gravity is viewed as a manifestation of the curvature of space–time.

The work performed by Gauss allowed a formula to be established for the curvature K of a surface. In Riemann coordinates, this is equivalent to finding an expression for the completely covariant Riemann tensor R_{xyxy}. In two dimensions, this tensor can be stated as follows:

$$R_{xyxy} = -\frac{1}{2}\left(\frac{\partial^2 g_{xx}}{\partial y^2} + \frac{\partial^2 g_{yy}}{\partial x^2}\right) = K, \qquad [1.23]$$

where g_{xx} and g_{yy} denote the coefficients of the metric in Riemann coordinates. Most work on the curvature of space has relied on metrics defined for specific frames of reference, both inertial and non-inertial; this allows the curvature to be expressed in terms of tensors of orders greater than two in a manner consistent with the concepts of a continuum and differentiation at a point. The fundamental principles of analysis, differential geometry and vector analysis that are conventionally used to introduce the notion of curvature are not compatible with the discrete approach, nor is performing an *a posteriori* discretization of the results of the continuous approach a good solution. Alternative approaches in the field of discrete differential geometry have considered meshed topologies, but they are mostly intended for computations and numerical simulations, without attempting to fully detach themselves from the equations of a continuum.

Discrete mechanics [CAL 15a] revisits the equations of mechanics from a different perspective: we assume the existence of objects, edges and oriented surfaces that cannot be reduced to a single point by scaling. The usual notions of frame of reference, differentiation, etc., are no longer suitable and must therefore be abandoned. Like the other quantities used to establish the equations of motion, the curvature requires a discrete definition valid for any primal and dual topologies.

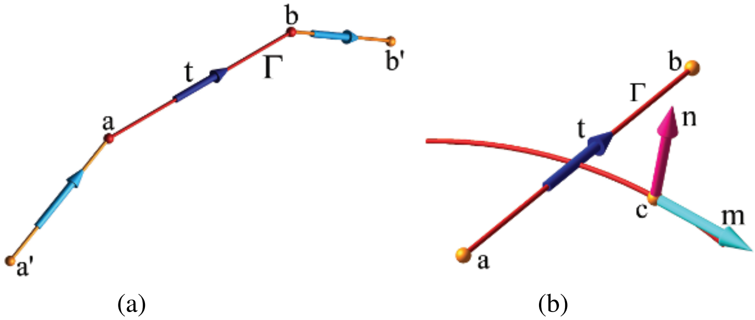

Figure 1.7. *Longitudinal and transverse curvatures, calculated as the divergence of the vectors* $\mathbf{t}$ *and* $\mathbf{m}$ *in curvilinear coordinates. For a color version of this figure, see www.iste.co.uk/caltagirone/mechanics.zip*

Consider the elementary stencils shown in Figure 1.7; Figure 1.7a shows a discrete curvilinear contour of edges. Each edge Γ has unit vector $\mathbf{t}$ and is rectilinear between its endpoints a and b. Figure 1.7b shows the curvilinear edge orthogonal to Γ with unit vector $\mathbf{m}$; at the point c, the vector $\mathbf{n}$ satisfies $\mathbf{n} = \mathbf{t} \times \mathbf{m}$.

The discrete curvature of the edge Γ has two components: the longitudinal curvature κ_l and the transverse curvature κ_t. The first, defined at the point a or b, represents the variation in the direction of the vector $\mathbf{t}$ in discrete curvilinear coordinates, whereas the second, defined at the point c, defines the variation of the curvature along the axis with unit vector $\mathbf{m}$. As in continuous differential geometry, the mean curvature is defined as the sum of the principal curvatures $\kappa = \kappa_l + \kappa_t$. Like the other quantities of the discrete approach, κ_t can be viewed either as a vector or a scalar defined on the unit vector $\mathbf{n}$. The curvature κ_l is a scalar defined at the points of the primal topology.

If the primal topology is known, we can compute each unit vector $\mathbf{t}$ from the local coordinates of the points. The principal curvatures are expressed in terms of discrete operators on the primal topology as follows:

$$\begin{cases} \kappa_l = \nabla \cdot \mathbf{t} \\ \kappa_t \, \mathbf{n} = \nabla \times \mathbf{t} \end{cases} \qquad [1.24]$$

These are formal definitions that introduce metrics that must be computed from the points of the primal topology where the unit vectors $\mathbf{t}$ and $\mathbf{n}$ are defined.

Whenever the curvature is encountered in a physical phenomenon, our objective is to find the acceleration γ_c of the phenomenon; for example, in the context of an interfacial effect, the curvature is used to find the capillary acceleration γ_c; this quantity may be obtained by projecting the action of variations in the surface curvature onto Γ using the gradient and dual curl operators. The capillary acceleration is therefore given by:

$$\gamma_c = \nabla \left(\sigma_l \, \nabla \cdot \mathbf{t} \right) + \nabla \times \left(\sigma_t \, \nabla \times \mathbf{t} \right), \qquad [1.25]$$

where σ_l and σ_t are the surface tension per unit mass of the two media separated by the interface Σ. Even with constant curvature, the interface may experience motion due to the Marangoni effect. Note that these effects are fully defined on the interface Σ and the overpressure is simply the consequence of their action on a closed surface such as a drop or a bubble.

In fact, the formulation of equation [1.25] is perfectly generic and can be extended to gravity, inertial effects and so on. Figure 1.8 shows how the acceleration γ_c is defined on the edge Γ; the two actions represented by the divergence of $\mathbf{t}$ and the curl of $\mathbf{t}$ are both projected onto the principal edge Γ.

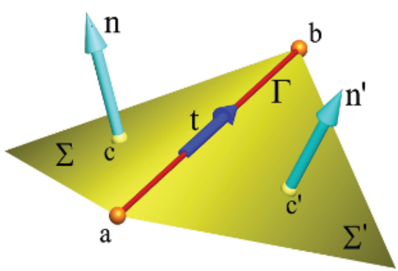

Figure 1.8. *Longitudinal and transverse curvatures, computed as the divergence of the vector* $\mathbf{t}$ *and the primal curl of* $\mathbf{t}$*, respectively, on the contour of* Σ*. For a color version of this figure, see www.iste.co.uk/caltagirone/mechanics.zip*

The velocity induced by each effect on the edge Γ does not need to be known directly; the velocity $\mathbf{V}$ also depends on all other effects and the value $\mathbf{V}^o$ of the mechanical equilibrium at time t^o and can be updated accordingly, $\mathbf{V} = \mathbf{V}^o + \gamma dt$. If we compute the curvature of a circle of radius R represented by equidistant points connected by edges using the expression from equation [1.24], we recover the exact solution $\kappa = 1/R$ regardless of the number of markers in the primal topology. In three dimensions, we recover the notion of mean curvature $\kappa = \kappa_l + \kappa_t$ from differential geometry. As we saw elsewhere, these two effects can double up; for example, the mean curvature of a sphere is twice that of a cylinder. They can also cancel, for example in surfaces with zero curvature such as catenoids. The minimal surface of a soap film pressing against two circles of radius R takes the form of a catenary rotated around an axis – also known as a catenoid. Assuming atmospheric pressure on either side of the film, this surface has zero mean curvature.

Equation [1.25] is stated in a generic form; any vector may be written as the sum of a gradient and a dual curl, defining any internal actions in terms of the divergence and primal curl operators. This decomposition is the cornerstone of discrete mechanics and in particular the equations of motion.

In practice, the nature of the problem of computing the curvatures κ_l and κ_t changes according to whether the interface is defined analytically by markers, by a phase function, by an indicator function, etc. We will revisit the purely numerical aspects of this problem in Chapter 5. As a demonstration, let us compute the curvature of a zero-curvature surface – the catenoid from classical differential geometry. In a parametric form, the catenoid may be expressed as $x = \cosh(k\,u)\ \cos v/k$, $y = \cosh(k\,u)\ \sin v/k$, $z = u$, where k is a real constant. Writing $r(u, v)$ for a point on the surface, the normal vector $\mathbf{n}$ at this point is given by:

$$\mathbf{n} = \frac{\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}}{|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}|} = \begin{cases} -\dfrac{\cos v}{\cosh(k\,u)} \\ -\dfrac{\sin v}{\cosh(k\,u)} \\ \dfrac{\sin(k\,u)}{\cosh(k\,u)} \end{cases} \quad [1.26]$$

After computing the metrics $[g_{ij}]$ on the surface, the principal curvatures are as follows:

$$\begin{cases} \kappa_1 = -\dfrac{k}{\cosh^2(k\,u)} \\ \kappa_2 = \dfrac{k}{\cosh^2(k\,u)} \end{cases} \quad [1.27]$$

and we find that the mean curvature $\kappa = \kappa_1 + \kappa_2$ is locally zero, as expected.

On a discrete surface for which a parameterization is not available, the computation is performed directly from equation [1.24]. The computation of the longitudinal curvature from the expression of κ_l is relatively straightforward; similarly, the transverse curvature κ_t can be computed as the variation of the vector $\mathbf{m}$ along the curvilinear edge. In the general case, for arbitrary primal topologies, the mean curvature can be computed using existing techniques from discrete differential geometry or based on the ideas presented here. Figure 1.9 shows two surface geometries, a sphere and a catenoid, constructed from a relatively regular triangulation.

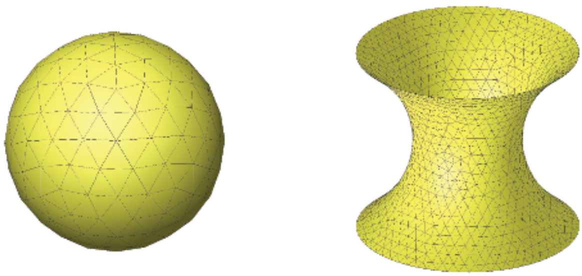

Figure 1.9. *Surface topologies of a sphere and a catenoid based on a regular triangulation*

With this example topology, we recover the theoretical (continuous) value of the mean curvature of the sphere to a high degree of accuracy, even though the edges Γ are oriented arbitrarily. In the case of the catenoid, we recover the value of zero predicted by differential geometry from its parametric function.

It is then easy to compute the unit vectors $\mathbf{t}$ from the coordinates of the vertices of each triangle and deduce the curvatures κ_l and κ_t. The duality hypothesis allows us to write the acceleration on the edge Γ as the sum of the gradient of a scalar potential and the dual curl of a vector potential. Techniques based on elaborate concepts from differential geometry and numerical methods of evaluating the curvature are not all so accurate nor easy to implement. Here, we only require the values of the curvatures κ_l and κ_t at the points a and c; depending on the methodological context, the best way to evaluate them is as follows:

$$\boldsymbol{\gamma}_c = \nabla \left(\sigma_l \, \kappa_l \right) + \nabla \times \left(\sigma_t \, \kappa_t \, \mathbf{n} \right). \qquad [1.28]$$

Defining a single non-localized mean curvature is not equivalent to defining the two longitudinal and transverse curvatures from equation [1.28] at the points and barycenters of the cells of the primal topology. The formulation presented here aligns with the approach adopted by [VIN 04], who evaluate the mixed partial derivatives to compute the Hessian of the principal curvatures, even though these derivatives are not considered by classical methodologies. The mixed derivatives are evaluated at the center of each face of an MAC (marker-and-cell)-type mesh. The use of a single curvature, defined on the points of the primal mesh, leads to a loss of information and accuracy.

Returning to the standard continuum approach, consider a point M on a curved surface S defined implicitly by the equation $f(x, y, z) = 0$. The principal radii of curvature of S are computed on two orthogonal planes that are perpendicular to the tangent plane at the point M. If $\mathbf{t}_1 = (a_1, b_1, c_1)$ and $\mathbf{t}_2 = (a_2, b_2, c_2)$ are two unit vectors of the tangent plane, and $\mathbf{n} = \nabla f$ is the normal vector of the interface at the point M, the radius of curvature of the surface in the plane $(\mathbf{n}, \mathbf{t}_1)$ is given by the ratio $R_1 = \|\mathbf{n}\|/Hes(\mathbf{n}, \mathbf{t}_1)$, where $Hes(\mathbf{n}, \mathbf{t}_1) = \mathbf{t}_1^t H \mathbf{t}_1$ is the Hessian of the surface at the point M and H is the Hessian matrix of second-order partial derivatives of f:

$$R_1 = \frac{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}}{a_1^2 \frac{\partial^2 f}{\partial x^2} + b_1^2 \frac{\partial^2 f}{\partial y^2} + c_1^2 \frac{\partial^2 f}{\partial z^2} + 2a_1 b_1 \frac{\partial^2 f}{\partial x \partial y} + 2a_1 c_1 \frac{\partial^2 f}{\partial x \partial z} + 2b_1 c_1 \frac{\partial^2 f}{\partial y \partial z}}. \qquad [1.29]$$

The second tangent vector $\mathbf{t}_2$ can then easily be determined from its vector products with $\mathbf{n}$ and $\mathbf{t}_1$, since it satisfies $\mathbf{t}_1 \cdot \mathbf{t}_2 = 0$ and $\mathbf{n} \cdot \mathbf{t}_2 = 0$. The second radius of curvature R_2 is obtained similarly by repeating this approach in the second plane $(\mathbf{n}, \mathbf{t}_2)$. The mean curvature of the surface at the point M is given by $\kappa = 1/R_1 + 1/R_2$.

Let $\mathbf{f}$ be the vector of derivatives $\partial f/\partial x_j - \partial f/\partial x_k$ for each of the three components i. After observing that $\nabla f \cdot \nabla \times \mathbf{f} = 0$, we can define one of the tangent vectors $\mathbf{t}$ as the unit vector of $\mathbf{f}$ by setting $\mathbf{t} = \mathbf{f}/\|\mathbf{f}\|$.

The notion of discrete curvature is consistent with these results from differential geometry, as is the fact that we need to define different quantities at both the points and the barycenters of each cell of the primal topology. The mixed second derivative terms in equation [1.29] are taken into account by the discrete approach in the expression of the transverse curvature. An important result from differential analysis shows that the curvature of the surface is independent of the choice of basis $(\mathbf{t}_1, \mathbf{t}_2)$; this is another cornerstone of discrete mechanics, where the unit vector $\mathbf{t}$ of the edge Γ is defined with respect to a local frame of reference.

1.10. Axioms of discrete mechanics

In summary, the above-described principles – WEP and the causality principle – are assumed without reservation; these principles are consistent with every physical experiment that has ever been performed to date. Furthermore, Hodge–Helmholtz decomposition allows us to decompose any physical vector into an irrotational component and a solenoidal component. We define scalar and vector potentials for the acceleration in order to accumulate the compressive and shear stresses. This is known as the principle of accumulation of stresses.

The other axioms of discrete mechanics may be stated as follows:

– the fundamental principle of dynamics takes the form of a composition rule for accelerations; the acceleration $\boldsymbol{\gamma}$ of a particle is equal to the sum of the accelerations $\mathbf{g}$ acting upon this particle, $\boldsymbol{\gamma} = \mathbf{g}$;

– the acceleration is an absolute quantity in a single frame of reference;

– the velocity is defined as the accumulation of the acceleration $\mathbf{V} = \mathbf{V}^o + \boldsymbol{\gamma}\, dt$;

– like the velocity, the quantities of pressure, stress, density and temperatures are defined as accumulators;

– the propagation of relaxed longitudinal waves is directed and has velocity equal to the celerity;

– the effects of longitudinal and transverse propagation are disjoint;

– absolute time is assumed to exist and unfolds in the positive direction;

– there is a discrete horizon within which all points are connected by causality;

– any vector may be written as the sum of a solenoidal component and an irrotational component (Hodge–Helmholtz);

– for physical phenomena derived from a potential, there always exists both a direct effect and a dual effect; these effects can either cumulate or cancel.

Discrete mechanics moves away from various ideas that served as milestones in the history of the evolution of classical mechanics. The following concepts are abandoned:

– the notion of a continuum;

– the hypothesis of a local equilibrium;

– the necessity of using tensors;

– the composition rule of velocities viewed as vectors;

– the restriction of the velocity to the speed of light in a vacuum.

We shall set aside quantities such as force, momentum and energy for now while deriving the equations of motion; like other quantities discussed above, these secondary quantities will be derived later by integrating the primary variables. The hypothesis that the material velocity is bounded by the speed of light is not rejected as such by the field of discrete mechanics; it is simply not required as part of the theoretical framework.