Introduction: Some historical markers

The aim of this introduction is threefold. First, to give a brief account of a paper of Jacques Hadamard (1865–1963) which is at the genesis of the theory of global properties of surfaces of nonpositive Gaussian curvature. Then, to make a survey of several different notions of nonpositive curvature in metric spaces, each of which axiomatizes in its own manner an analogue of the notion of nonpositive Gaussian curvature for surfaces and more generally of nonpositive sectional curvature for Riemannian manifolds. In particular, we mention important works by Menger, Busemann and Alexandrov in this domain. Finally, the aim of the introduction is also to review some of the connections between convexity theory and the theory of nonpositive curvature, which constitute the main theme that we develop in this book.

The work of Hadamard on geodesics

The theory of spaces of nonpositive curvature has a long and interesting history, and it is good to look at its sources. We start with a brief review of the pioneering paper by Hadamard, "Les surfaces à courbures opposées et leurs lignes géodésiques" [87]. This paper, which was written by the end of the nineteenth century (1898), can be considered as the foundational paper for the study of global properties of nonpositively curved spaces.

Before that paper, Hadamard had written a paper, titled "Sur certaines propriétés des trajectoires en dynamique" (1893) [84], for which he had received the Borodin prize of the *Académie des Sciences*.[1] The prize was given for a contest whose topic was "To improve in an important point the theory of geodesic lines". The interest at that point was in *global* properties of geodesics, and it was motivated by Poincaré's recent works on trajectories of differential equations. We note that whereas the 1898 paper [87] by Hadamard concerns surfaces of negative curvature, the 1893 paper [84] essentially concerns geodesics on surfaces of positive curvature. One of the results proved in Hadamard's paper [87] is that on a closed surface of positive curvature, two closed geodesics have necessarily a common point. The members of the jury of the Borodin contest were Picard, Lévy, Appell, Darboux and Poincaré. Poincaré, acting as the referee, wrote a Comptes Rendus Note [179] on the result of the contest

[1]In 1892, Hadamard had already obtained the Grand Prix of the *Académie des Sciences* for a paper on Riemann's zeta function [83]. The subject of the contest that year was to fill a gap in one of Riemann's proofs, and Hadamard's contribution followed from the work he did in his doctoral thesis, *Essai sur l'étude des fonctions données par leur développement de Taylor*, devoted to complex function theory and written under Émile Picard and Jules Tannery. The paper [83] by Hadamard is interesting for the subject of the present book because, as we shall see in Chapter 6, it contains one of the earliest uses of the notion of convex function in analysis.

containing a short summary of Hadamard's work. This report ends with the following words:

> It may seem at first sight that the Memoir which we have just analyzed contains very few results, especially if we consider that one of them is not new.[2] But the committee decided that the author showed a great ingenuity of mind and that he highlighted a great number of new ideas which in all probability will be fruitful one day. It is only the lack of time which prevented him from taking full advantage of them.

In another Comptes Rendus Note which appeared a year later [180], Poincaré, recalling the circumstances of the Borodin contest, writes the following:

> The properties of geodesic lines deserve all the attention of geometers. Indeed, this problem is the simplest among all the problems of dynamics, but we already encounter in it the main difficulties of this kind of questions. Therefore, it is in studying thoroughly the geodesic lines that one can best be familiar with these difficulties. This is one of the reasons for which the Academy decided to hold a competition on that subject.

Hadamard and Poincaré were the first mathematicians who strongly emphasized the importance of topological methods in the study of spaces of nonpositive curvature. In the introduction to his 1898 paper [87], Hadamard writes[3]:

> The only theory which has to be studied profoundly, as a basis to the current work, is the *Analysis situs*, which, as one can expect after reading the work of Poincaré, plays an essential role in everything that will follow.

In the conclusion of the paper, Hadamard writes:

> As for the method which we have used, we can consider it as an application of two principles set up by M. Poincaré in his study of differential equations.
> First, our conclusions highlight, once again, the fundamental role that the *Analysis situs* plays in these questions. That it is absurd to study integral curves drawn on a fixed domain without taking into account the form of that domain is a truth upon which one may think that it is useless to insist. However, this truth was unsuspected until the works of M. Poincaré.
> Secondly, the importance of periodic solutions which this geometer recognized in his *Traité de Mécanique céleste* is equally manifested in the current question. Here also, they constitute the "only breach through which we can penetrate into a spot which until now was reputed to be inaccessible".[4]

One may add that Poincaré was particularly interested in the applications of the study of geodesics in dynamics and especially in astronomy. In the Comptes Rendus Note [179] in which he makes a summary of the work done by Hadamard in his paper [87] which we shall describe below, Poincaré writes:

[2] Indeed, as the report shows, one of the main results was already obtained by Kneser and by Lyapunov.
[3] In this volume, all the translations from the French are mine.
[4] Hadamard's quote is from Poincaré's *Méthodes nouvelles de la Mécanique céleste*, t. I, p. 82.

> We can consider [after the work done by Hadamard] that the problem is completely solved. The importance of the result can be highlighted by the following considerations:
> When we shall address the problem of the stability of the solar system from a strictly mathematical viewpoint, we shall be confronted with questions which are very similar: The trajectories are equivalent to the geodesics since, like them, they can be defined by equations of the calculus of variations [...].

Let us now look more closely at the content of Hadamard's paper [87].

Hadamard considers a surface S equipped with a Riemannian metric of nonpositive curvature. The surface is smoothly embedded in $\mathbb{R}^3$, the Riemannian metric is induced from that inclusion and the set of points at which the curvature is zero is finite.

Hadamard starts by noting that this surface is necessarily unbounded since if one of the coordinates assumed a maximum or a minimum, then, in a neighborhood of that point, the surface would be situated at one side of its tangent plane and therefore the curvature would be positive. He then proves that such a surface can always be decomposed into the union of a compact region and of a collection of *infinite sheets* (we are translating the term "nappes infinies" used by Hadamard). Each infinite sheet is homeomorphic to a cylinder, and is connected to the compact region along a boundary curve. Hadamard constructs examples of surfaces with nonpositive curvature having an arbitrary number of infinite sheets. He classifies the infinite sheets into two types: the *flared* infinite sheets ("nappes infinies évasées"), which we call "funnels", and the *unflared* infinite sheets. An unflared infinite sheet is characterized by the fact that one can continuously push to infinity a homotopically non-trivial closed curve on such a sheet while keeping its length bounded from above. To each unflared infinite sheet, Hadamard associates an *asymptotic direction*. The simple closed curves that connect the funnels to the compact region can be taken to be closed geodesics that are pairwise disjoint.

Let us now suppose that the fundamental group of S is finitely generated (Hadamard says that S has "finite connectivity"). For such a surface, Hadamard describes a finite collection of homotopically non-trivial and pairwise non-homotopic simple closed curves $C_1, \dots, C_n$ such that any non-trivial homotopy class of closed curves on the surface can be uniquely represented by a finite word written in the letters $C_1, \dots, C_n, C_1^{-1}, \dots, C_n^{-1}$. Hadamard pushes forward this analysis to make it include not only representations of closed curves, but also representations of long segments of open infinite curves. In fact, with these elements, Hadamard initiates the approximation of long segments of bi-infinite geodesic lines by closed geodesics, and therefore he also initiates the theory of symbolic representation of bi-infinite geodesics on the surface, that is, the theory of representation of elements of the geodesic flow associated to S by bi-infinite words in the finite alphabet $\{C_1, \dots, C_n, C_1^{-1}, \dots, C_n^{-1}\}$. We note in passing that the theory of symbolic dynamics for the geodesic flow associated to a nonpositively curved surface was thoroughly developed about twenty years later by Marston Morse, and that Morse, in his paper [156], refers to Hadamard as his source of inspiration. Hadamard calls a closed curve a "contour", and the curves $C_1, \ldots, C_n$ the "elementary contours". There are two sorts of closed curves in the collection $C_1, \ldots, C_n$ which Hadamard considers: the curves that are the boundaries of the infinite sheets, and those that correspond to the genus of the surface. (The latter come in pairs and they correspond to the "holes" of the compact part of the surface S, as Hadamard calls them.) He notices that there is (up to circular permutation) one and only one relation in the representation he obtains that allows us to consider any of these contours as a product of the elementary contours, and therefore, by eliminating one of the contours, he obtains a symbolic representation that is unique up to circular permutation.

In the same paper, Hadamard proves that in each homotopy class of paths on S with fixed endpoints, there is a unique geodesic.[5] He also obtains an analogous result for the free homotopy classes of closed curves on S that are neither homotopic to a point nor to the core curve of a flared infinite sheet.

Hadamard then makes the fundamental observation that the distance function from a fixed point in S to a point that moves along a global geodesic in this surface is convex. This observation is at the basis of the definition of nonpositive curvature in the sense of Busemann, which is the main topic of this book.

Next, Hadamard introduces the notion of *asymptotic geodesic rays*, and he constructs such rays as follows. Given a geodesic ray, he considers a sequence of geodesic segments joining a fixed point on the surface to a sequence of points on that ray that tends to infinity. He proves that this sequence of geodesic segments converges to a geodesic ray, which he calls *asymptotic* to the initial ray. He studies the asymptoticity relation and introduces through an analogous construction the notion of (local) geodesics that are asymptotic to a given closed geodesic.

Using an argument that is based on the convexity of the distance function from a point in a funnel to the closed geodesic that connects this funnel to the compact region in S, Hadamard proves that a geodesic that penetrates a funnel cannot get out of it.[6]

Finally, in the case where all the infinite sheets of S are funnels, Hadamard investigates the distribution of geodesics that stay in the compact part of the surface. These geodesics are the flowlines of the geodesic flow associated with the compact nonpositively curved surface with geodesic boundary, obtained from the surface S by deleting the funnels. For a given point x in S, Hadamard considers the set of initial directions of geodesic rays that start at x and stay in that compact surface. He proves that this is a perfect set with empty interior, whereas the set of directions of geodesic rays that tend to infinity is open.

[5]This is Theorem 31 of Part III of [87]. Here, the word geodesic is used, as in Riemannian geometry, in the sense of local geodesic. We warn the reader that this is not the definition that we are adopting in the rest of this book, where a geodesic is a distance-minimizing map between its endpoints (and therefore it is a *global* geodesic). We shall see in this book a generalization of this result of Hadamard to metric spaces satisfying some convexity properties; see Corollary 9.3.3, which follows from a theorem of Alexander–Bishop.

[6]Hadamard already obtained a similar result in [85].

We already mentioned that the setting of the paper [87] by Hadamard is that of a differentiable surface S embedded in $\mathbb{R}^3$, such that at each point of S, the Gaussian curvature is negative except for a finite set of points where it is zero. A theory of *metric spaces with nonpositive curvature*, that is, a theory that does not make any differentiability assumption and whose methods use the distance function alone, without the local coordinates provided by an embedding in Euclidean space or by another Riemannian metric structure, was developed several decades after the paper [87]. It is good to remember, in this respect, that the theory of metric spaces itself was developed long after the theory of spaces equipped with differentiable structures. For instance, Gauss's treatise on the differential geometry of surfaces [74], in which he defines (Gaussian) curvature and proves that this curvature depends only on the intrinsic geometry of the surface and not on its embedding in $\mathbb{R}^3$, was published in 1827,[7] whereas the axioms for metric spaces were set down by Fréchet, some 80 years later.

Of course, the setting of surfaces embedded in $\mathbb{R}^3$ has the advantage of providing visual characteristics for the sign of the curvature. For instance, it is well known that for such a surface, we have the following:

• if the Gaussian curvature at some point is > 0, then the surface, in the neighborhood of that point, is situated on one side of the tangent plane;

• if the Gaussian curvature at some point is < 0, then the surface, at that point, crosses its tangent plane.

We also recall that if the Gaussian curvature at some point is 0, then any one of the above configurations can occur.

Hadamard indicated in the Note [86] how to extend some of the results he proved for surfaces to higher dimensions. The development of these ideas in the general setting of Riemannian manifolds of nonpositive curvature has been carried out by Elie Cartan, in particular in his famous "Leçons sur la géométrie des espaces de Riemann" [50].[8]

In closing this section, we mention a valuable reference on the life and work of Hadamard, [135].

The works of Menger and Wald

A few years after the introduction in 1906 by Fréchet of the axioms for metric spaces, Karl Menger initiated a theory of geodesics in these spaces. A *geodesic* in a metric space is a path whose length is equal to the distance between its endpoints. Menger

[7] We recall that the surfaces considered by Gauss were always embedded in $\mathbb{R}^3$. It is only 30 years after Gauss's paper was written that Riemann introduced the concept of spaces equipped with (Riemannian) metrics with no embedding in $\mathbb{R}^3$ involved in the definition.

[8] The later editions of Cartan's book [50] contains additional material, in particular the development of the work of Hadamard to higher dimensions.

generalized several results of classical geometry to this new setting and he introduced new methods that did not make any use of local coordinates or of differentials, but only of equalities involving the distance function, and of the triangle inequality. Menger wrote several important papers involving this new notion of geodesic, and he also introduced a notion of "discrete geodesic", which is based on his definition of *betweenness*: a point z in a metric space is said to lie between two distinct points x and y if z is distinct from x and from y and if we have

$$d(x, y) = d(x, z) + d(z, y).$$

In a complete metric space, the existence of a geodesic joining any two points is equivalent to the existence, for any distinct points, of a point that lies between them. We refer the reader to the commented edition of Menger's papers in the volume [151], published on the centenary of Menger's birth. We shall have several opportunities to mention Menger's work in the following chapters, but here, we mention an important notion that he introduced, which we shall not consider further in this book. This notion contains an idea that is at the basis of the various definitions of curvature that make sense in general metric spaces. The idea is to construct "comparison configurations" for sets of points (say of finite cardinality) in a given metric space X. The comparison configurations are built in a model space, which is generally one of the complete simply connected surfaces M_κ of constant Gaussian curvature κ, that is, either the Euclidean plane (of curvature $\kappa = 0$), or a sphere of curvature $\kappa > 0$, or a hyperbolic plane of curvature $\kappa < 0$. The comparison configuration for a given subset $F \subset X$ is a subset $F^\star$ of M_κ, equipped with a map from F to $F^\star$, which is generally taken to be distance-preserving and which is called a "comparison map". Then, one can define notions like curvature at a given point x in X by requiring the *ad hoc* property for the comparison configurations associated to certain classes of subsets contained in a neighborhood of x. Of course, a comparison configuration does not always exist, but in the case where F is of cardinality 3, one can always construct a comparison configuration $F^\star$ of F in any one of the surfaces M_κ with $\kappa \leq 0$. In fact, the triangle inequality in the axioms for a metric space X is equivalent to the fact that one can construct a comparison configuration in the Euclidean plane for any triple of points in X.[9] Let us now consider an example.

Given three pairwise distinct points a, b and c in the Euclidean plane M_0, either they lie on a unique circle (the circumscribed circle), or they lie on a Euclidean straight line. It is useful to consider here such a line as being a circle of radius ∞, in order to avoid taking subcases. Now for any triple of pairwise distinct points in a metric space X, Menger defined its "curvature" as being equal to $1/R$, where R is the radius of a circle in the Euclidean plane which is circumscribed to a comparison configuration associated to that triple. With this definition, the three points in X are aligned (that is, they satisfy a degenerate triangle inequality) if and only if their

[9] We also mention that an early version of the idea of a comparison map is already contained in the very definition of the Gauss map.

curvature is zero. Given an arc (or, say, a one-dimensional object) A contained in X, Menger says that the curvature of A at a point $a \in A$ is equal to κ if for any triple of pairwise distinct points in A that are sufficiently close to a, the curvature of this triple is close to κ.

With this definition, Menger introduced the notion of *curvature* for one-dimensional objects in an arbitrary metric space, and he posed the problem of the definition of curvature for higher-dimensional objects.

Now, we must mention the work of Abraham Wald, a student of Menger, who introduced a notion of two-dimensional curvature in an arbitrary metric space.[10] The definition again uses a limiting process, but now it involves quadruples of points in that space. The problem is that, in general, a quadruple of points in a metric space does not necessarily possess a comparison configuration in the Euclidean plane. However, Wald starts by proving that if the metric space X is a differentiable surface, then, for every point x in X, there exists a real number $\kappa(x)$ satisfying the following property:

($\star$) for each $\epsilon > 0$, there exists a neighborhood $V(x)$ of x such that for every quadruple of points Q in $V(x)$, there is an associated real number $\kappa(Q)$ satisfying $|\kappa(x) - \kappa(Q)| < \epsilon$ such that the quadruple of points Q possesses a comparison configuration in the model surface $M_{\kappa(Q)}$.

Wald then proves that the quantity $\kappa(x)$ is equal to the Gaussian curvature of the surface at the given point x.

Now let X be a metric space that is "Menger convex", that is, a metric space X in which for every pair of distinct points x and y, there exists a point z that lies between them. Suppose furthermore that the Wald two-dimensional curvature exists at each point of X. In other words, suppose that one can associate to each point x in X a real number $\kappa(x)$ satisfying property ($\star$) that is stated above. Under these assumptions, Wald shows that the space X has the structure of a differentiable surface embedded in $\mathbb{R}^3$ that induces the same *length structure* as that of the original metric on X. In other words, the lengths of an arbitrary curve in X, on the one hand measured using the original metric and on the other hand measured using this differentiable surface structure, coincide. Furthermore, at each point a in X, the quantity $\kappa(a)$ is equal to the Gaussian curvature induced by the differentiable embedding of the surface in $\mathbb{R}^3$. With this result, Wald solved a problem that had been posed by Menger in [148], which asked for a metric characterization of Gauss surfaces among Menger convex metric spaces. We refer the reader to the Comptes Rendus Note [208] by Wald, presented by Elie Cartan.

We note in passing that for extra-mathematical reasons, Wald stopped working on this subject soon after he published the solution to Menger's problem, and his research interests switched to statistics and econometry. The story is interesting and it is told by Menger in [152]. It seems that there was no direct continuation to Wald's work.

[10] Wald gave this 2-dimensional curvature the name "surface curvature", and Menger refers to it as "Wald curvature".

Now, after the notion of curvature in metric spaces, we pass to the notion of nonpositive curvature.

The works of Busemann and Alexandrov

For the development of the theory of nonpositively curved metric spaces, we shall consider works that have been carried out in two different directions: the works of H. Busemann and the works of A. D. Alexandrov and his collaborators. Both Busemann and Alexandrov dedicated their life to geometry. Busemann obtained his doctorate in Göttingen in 1931 under Richard Courant, and Alexandrov in 1935 (and his second doctorate in 1937) in Leningrad, under Boris Delone (Delaunay) and Vladimir Fock. Busemann and Alexandrov developed after that two distinct approaches to metrics of nonpositive curvature[11] and they gave rise to two different schools with no real interaction between them (although the two men met in 1954 and later on).[12] The ramifications of these two theories continue to grow today, especially since the rekindling of interest that was given to metric nonpositive curvature by M. Gromov in geometric group theory in the 1970s.

Let us briefly describe the basic underlying ideas of these works. First we need to recall a few definitions. Consider a metric space X in which each point x possesses a neighborhood U such that any two points in U can be joined by a geodesic path in U. A metric space X with such a property is said to be a *locally geodesic* space. We say that such a neighborhood U is a *geodesically convex* neighborhood of x. A *geodesic segment* $[a,b]$ in X is, by definition, the image of a geodesic path in X joining a and b.[13] A *triangle* in U is the union of three geodesic segments $[a,b]$, $[a,c]$ and $[b,c]$ contained in U. The segments $[a,b]$, $[a,c]$ and $[b,c]$ are called the *sides* of this triangle.

We start by presenting Busemann's definition of nonpositive curvature, which has the advantage of being the simplest one to describe, and on which we shall focus in this book.

We say that the space X has *nonpositive curvature in the sense of Busemann* if

[11] Alexandrov's work, unlike the one of Busemann, contains substantial results on manifolds with "curvature bounded from one side", and not only from above.

[12] Busemann received the Lobachevsky prize in 1985, "for his innovative book *The Geometry of Geodesics*" which he had written 30 years before, and which we shall quote extensively in the present book. The Lobachevsky prize is certainly the most prestigious reward in geometry that was given the Soviet Union (and, before that, in Russia; the first recipients of this prize were Sophus Lie in 1897, Wilhelm Killing in 1900 and David Hilbert in 1903). The prize was awarded to Alexandrov in 1951. We also note that in the AMS (1991, 2000 and 2010) subject classification, there is a section devoted to Alexandrov geometry (53C45: Global surface theory à la A. D. Alexandrov) and another section devoted to Busemann geometry (53C70: Direct methods: G-spaces of Busemann, etc.). This shows the importance of the two geometries, but it also shows that they are considered there as separate subjects.

[13] From now on, we use the term "geodesic" in the sense of "global geodesic", that is, a path whose length is equal to the distance between its endpoints.

every point x in X possesses a geodesically convex neighborhood U such that for any geodesic triangle with sides $[a,b]$, $[a,c]$ and $[b,c]$ contained in U, we have[14]

$$\mathrm{dist}(m, m') \leq (1/2)\mathrm{dist}(b, c),$$

where m and m' are respectively the midpoints of $[a,b]$ and $[a,c]$. This property can be stated in terms of a convexity property of the distance function, defined on the product of any two geodesic segments $[a,b] \times [a,c]$ in X equipped with their natural (barycentric) coordinates. We shall develop this point of view in later chapters.

A nonpositively curved space in the sense of Busemann is sometimes referred to as a "locally convex metric space", or "local Busemann metric space". The terminology "nonpositively curved space", in this sense, is due to Busemann.[15]

A complete Riemannian manifold of nonpositive sectional curvature is an example of a metric space of nonpositive curvature in the sense of Busemann.

Busemann also introduced more general notions of nonpositive curvature in metric spaces. We can mention here the notion of a *space where the distance function is peakless* and the notion of a *space where the capsules are convex*. For these two notions, we refer the reader to Chapter 6 and 9 respectively, and the Notes at the end of these chapters which indicate the usefulness of these notions.

Busemann opened up a wide research area in metric geometry. He formulated a large amount of open problems in this field with applications in various domains such as Lie groups, Finsler geometry, Hilbert geometry, Minkowski geometry, dynamics, asymptotic geometry and the calculus of variations. Together with his students and collaborators, he worked extensively on several of these problems. Some of them are solved today, but others remain wide open and are presently the subject of intense research. The present book is written with the aim of introducing the reader to some of the tools needed to tackle some of these problems.

Busemann died in 1994 after he spent the last years of his life working mostly as a painter. In his last paper, written in collaboration with Phadke and published in 1993, Busemann recalls his beginning in metric geometry. He writes the following ([46] p. 181):

[14] It is interesting to note that an analogue of this property for the case of the sphere was already known in Greek antiquity, namely, in the work of Menelaus on spherical geometry. Indeed, Proposition 27 of Book I of Menelaus' *Spherics* reads as follows: "Let us produce in a triangle ABC an arc DE of great circle that cuts AB and BC into two halves. I say that DE is greater than half of AC". Menelaus' *Spherics* is a masterpiece of Greek geometry (Alexandrian school, 2nd century A.D.). It is a complete treatise in non-Euclidean geometry which is rather unknown to the mathematics community, except for a few historians. The work has been conceived as the non-Euclidean counterpart of Euclid's *Elements*. It survived only in Arabic texts. A German translation exists [126], based on a manuscript of Ibn 'Irāq (10th century), and a new critical English edition based on several other Arabic manuscripts is in preparation [183].

[15] Busemann has some additional hypotheses on the metric spaces that he considers. One such hypothesis is the uniqueness of the prolongation of geodesics. There are interesting examples of spaces that do not satisfy this hypothesis and for which the results that we are interested in are valid, and for that reason we have tried to avoid this hypothesis in this book. We recall the precise definition of the spaces considered by Busemann (which he calls G-spaces) in the Notes on Chapter 2 below.

> Busemann has read the beginning of Minkowski's *Geometrie der Zahlen* in 1926 which convinced him of the importance of non-Riemannian metrics. At the same time he heard a course on point set topology and learned Fréchet's concept of metric spaces. The older generation ridiculed the idea of using these spaces as a way to obtain results of higher differential geometry. But it turned out that a few simple axioms on distance suffice to obtain many non-trivial results of Riemannian geometry and, in addition, many which are quite inaccessible to the classical methods.

In the concluding remarks of that paper, Busemann writes:

> The acceptance of our theory by others was slow in coming, but A. D. Alexandrov and A. M. Gleason were among the first who evinced interest and appreciated the results.

Let us now say a few words on the point of view of A. D. Alexandrov.

Before giving the definition of nonpositive curvature in the sense of Alexandrov, we must recall the notion of angle which two geodesic segments (or more generally two paths) in a metric space, that start at a common point, make at that point.

The notion of angle in a metric space has been introduced by Alexandrov as a generalization of the notion of angle in a surface. The first papers by Alexandrov deal with the intrinsic geometry of surfaces, and the notion of angle is certainly the most important notion in that theory, after the notion of distance and that of length of a path.[16] In fact, Alexandrov introduced several notions of angle, and these notions coincide provided some reasonable hypotheses on the ambient metric spaces are satisfied. For our needs, it suffices to consider the notion of *upper angle* that two geodesic segments with a common initial point make at that point.

Let X be a locally geodesic metric space, let U be a geodesically convex open subset of X and let us consider a geodesic triangle Δ in U, with sides $[a,b]$, $[a,c]$ and $[b,c]$. Let us define the *upper angle* α of the triangle Δ at the vertex a. For every point x on the segment $[a,b]$ and for every point y on the segment $[a,c]$, we consider a triangle $\Delta_{x,y}$ in X with sides $[a,x]$, $[a,y]$ and $[x,y]$, where $[a,x]$ and $[a,y]$ are subsegments of $[a,b]$ and $[a,c]$ respectively and where $[x,y]$ is a geodesic segment joining the points x and y. Let $\Delta^\star_{x,y}$ be an associated comparison triangle in the Euclidean plane and let $\alpha_{x,y}$ be the angle of $\Delta^\star_{x,y}$ at the vertex that corresponds to the vertex a of $\Delta_{x,y}$. Then, the *upper angle* of the triangle Δ at the vertex a is defined as

$$\alpha = \limsup_{x,y\to a} \alpha_{x,y}.$$

The *angular excess* of the triangle Δ is then defined as

$$\delta(\Delta) = \alpha + \beta + \gamma - \pi,$$

where α, β and γ are the upper angles of Δ at the vertices a, b and c.

[16] One of the basic papers of Alexandrov on that theory is [4]. We also refer the reader to [5] and [6].

Finally, the space X is said to be *nonpositively curved* in the sense of Alexandrov if every point in X possesses a geodesically convex neighborhood U such that the angular excess of any triangle in U is ≤ 0.

Let us note that in the case where the space X is a differentiable surface, the angular excess $\delta(\Delta)$ is the classical *total curvature*, that is, the integral of the Gaussian curvature over the region $R \subset X$ which is bounded by the triangle Δ. Indeed, the Gauss-Bonnet theorem applied to a disk D with piecewise geodesic boundary embedded in a differentiable surface $S \subset \mathbb{R}^3$ says that

$$\tau + \omega = 2\pi\chi,$$

where $\omega = \int_D dS$ is the total curvature of the disk D, that is, the integral of the Gaussian curvature with respect to the area element of that disk. In the special case considered, χ is the Euler characteristic of the disk, which is equal to 1, and τ is the sum of the *rotations* of the boundary of the disk at the vertices. At a vertex whose angle is α, the rotation is equal to $\pi - \alpha$. This gives

$$\omega = 2\pi - (\pi - \alpha) - (\pi - \beta) - (\pi - \gamma) = \delta(\Delta).$$

Thus, the notion of nonpositive curvature in the sense of Alexandrov generalizes the classical notion of nonpositive curvature for differentiable surfaces.

Complete Riemannian manifolds with nonpositive sectional curvature are also examples of nonpositively curved metric spaces in the sense of Alexandrov.

It should be noted that a metric space which is nonpositively curved in the sense of Alexandrov is also nonpositively curved in the sense of Busemann, but that the converse is not true. For instance, any finite-dimensional normed vector space whose unit ball is strictly convex is nonpositively curved in the sense of Busemann, but if the norm of such a space is not associated to an inner product, then this space is not nonpositively curved in the sense of Alexandrov. Alexandrov mentions this example in [5], p. 197.

In positive curvature the angular excess is positive, and one should also note in this respect that the fact that the angular excess of a spherical triangle is positive was known since Greek antiquity; it is contained in Proposition 12 of Book I of Menelaus' *Spherics*, see [126] and [183].

One can note that the techniques that are used by Alexandrov in all his works rely heavily on the notion of angle in a metric space, whereas the techniques of Busemann seldom use this notion.

Finally, it is fair to mention a predecessor to Alexandrov in his work on angle comparison in triangles, namely, Paolo Pizzetti, who wrote five papers on the subject in 1907, [176], [177]; see the report [161] by Pambuccian and Zamfirescu, and the other references given there.

Convexity

To end this introduction, we would like to make a few comments on convexity theory in relation with nonpositive curvature, but before that, we mention a particular link between the study of convex polyhedra and that of the differential geometry of surfaces.

In his historical report [68], Werner Fenchel traces back the origin of Alexandrov's work on the intrinsic geometry of convex surfaces to some early work on convex polyhedra. He first recalls Cauchy's rigidity result of 1812 stating that if two combinatorially equivalent convex polyhedra in $\mathbb{R}^3$ have congruent corresponding faces, then the polyhedra are themselves congruent.[17] Fenchel then reports that Cauchy, in a note he made for the French Academy of Sciences, wrongly announced that this result on polyhedra immediately implies that there is no closed convex surface that admits isometric deformations, a result that had already been claimed, also with a false proof, by Newton, around the year 1770.[18] Still, the problem was posed, and the relation between the rigidity of convex polyhedra and the rigidity of convex surfaces was clear: by replacing the condition on the isometry between the faces by a local (infinitesimal) condition, one is naturally led to the problem of finding local conditions on two closed convex surfaces under which these surfaces are isometric. The list of mathematicians who worked on this problem includes the names of J. H. Jellett, H. Liebmann, H. Weyl, S. Cohn-Vossen, A. D. Alexandrov, A. V. Pogorelov and others. We refer the reader to Fenchel's paper for this fascinating story.

There are also relations between convexity and nonnegative curvature. We mention as an example and without further comment the following result of Alexandrov and Pogorelov obtained in 1948: if X is a length metric space homeomorphic to the 2-sphere, then X has nonnegative curvature if and only if X is isometric to a convex surface S in $\mathbb{R}^3$, and in this case the surface S is unique up to rigid motions of $\mathbb{R}^3$. The result proved in particular a conjecture made by Hermann Weyl.

Although there are some parallel between the two theories, the general theory of nonnegative curvature has a different flavor than that of nonpositive curvature, and we do not consider it in this book.

[17] Cauchy, in his paper *Des polygones et les polyèdres*, cf. [52], traces back this work on polyhedra to Euclid, namely, he says that the rigidity statement is contained in Definitions 9 and 10 of Book XI of the *Elements*. The definitions say respectively that "similar solid figures are those contained by similar planes equal in multitude" and "Equal and similar solid figures are those contained by similar planes equal in multitude and in magnitude". Legendre, in a report on Cauchy's work which he wrote a year after the work has been presented to the Academy (the report is published in the *Correspondance sur l'École Polytechnique*, Tome II, p. 361–367, 1811), writes the following: "[...] I encouraged the author to continue his researches on polyhedra, with the goal of proving an interesting theorem which is assumed by Definitions 9 and 10 of the XIth book of Euclid and which is not yet proved. This theorem of which I talked much in the Notes to my *Geometry* and to which I added the necessary restrictions so that it would not be the object of the objection made by Robert Simpson in his edition of Euclid's *Elements*, can be stated as follows: *Two convex polyhedra are equal whenever they are contained by the same number of polygons which are mutually equal and placed together in the same manner*".

[18] Cauchy's arguments were corrected later on by H. Lebesgue among others.

We turn back to the relation between convexity theory and the theory of spaces of nonpositive curvature. First of all, as we have already said, it had already been realized by Hadamard that the convexity of the distance function in a nonpositively curved space is responsible for many of the global properties of that space. We also mentioned that this idea has been extensively explored by Busemann, who defined nonpositive curvature precisely by a convexity property of the distance function, and who showed, using this new definition, that most of the important properties of a nonpositively curved Riemannian manifolds are valid in a setting which is much wider than that of Riemannian geometry. Secondly, many of the basic results in convexity theory have the flavor of nonpositive curvature, and we mention as an important class of examples the "local-implies-global" properties, such as the fact that a locally convex function is globally convex, or the fact that a local geodesic in a Busemann space (*i.e.* in a simply connected metric space that is nonpositively curved in the sense of Busemann) is a global geodesic, and there are many others.

In the chapters that follow, we describe these facts in some detail.

We already noted that the field of metric geometry, since its first developments by Busemann, Alexandrov and others, has always had interactions with several fields of mathematics including Lie groups, Finsler geometry, the calculus of variations, Hilbert geometry, Minkowski geometry and other non-Euclidean geometries. More recently, new interactions appeared between metric geometry and other fields, like Teichmüller theory, the theory of lattices in semisimple Lie groups, geometric group theory, Tits buildings and geometric topology. We barely touch upon these topics in this book, but we mention them here as a motivation for the reader.