Invitation to comparison geometry

Anton Petrunin and Sergio Zamora Barrera

Contents

Part I

Curves

Chapter 1

Curves

Paths. Let $\mathcal{X}$ be a metric space. A continuous map $f\colon [0,1]\to\mathcal{X}$ is called a *path*. If $p=f(0)$ and $q=f(1)$, then we say that f *connects* p *to* q.

If any two points in $\mathcal{X}$ can be connected by a path then $\mathcal{X}$ is called *path connected*. Similarly, a subset $A\subset\mathcal{X}$ is called *path connected* if any two points $p,q\in A$ can be connected by a path that runs in A; equivalently, the subspace A is path connected.

Simple curves.

1.1. Definition. *A path connected subset γ in a metric space is called a* simple curve *if it is* locally *homeomorphic to a real interval; that is, any point $p\in\gamma$ has a neighborhood $U\ni p$ such that the intersection $U\cap\gamma$ is homeomorphic to a real interval.*

It turns out that any curve γ admits a homeomorphism from a real interval or a circle; that is, there is a continuous bijection $G\to\gamma$ with continuous inverse; here (and further) G denotes a circle or real interval. We omit a proof of this statement, but it is not hard.

The homeomorphism $G\to\gamma$ as above is called *parametrization* of γ. The parametrization completely defines the curve. Often will use the same letter for curve and its parametrization, so we can say curve γ has parametrization $\gamma\colon G\to\mathcal{X}$. Note however that any curve admits many different parametrization.

1.2. Exercise. *Find a continuous injective map $\gamma\colon[0,1)\to\mathbb{R}^2$ such that its image is not a simple curve.*

Hint: The image of γ should have a shape of digit 9.

If G is a circle, then the curve $\gamma\colon G\to\mathcal{X}$ is called *closed*. If G is a real interval, then we may say that γ is an *arc*.

Parameterized curves. A *parameterized curve* is defined as a continuous map $\gamma\colon G\to\mathcal{X}$. For a parameterized curve we do not assume that γ is injective; in other words the parameterized curve might have self-intersections.

1.3. Advanced exercise. *Let $\alpha\colon[0,1]\to\mathcal{X}$ be a path from p to q. Assume $p\ne q$. Show that there is a simple path connecting from p to q in $\mathcal{X}$.*

Smooth curves

A curve in the Euclidean space or plane, called *space* or *plane curve* correspondingly.

A space curve can be described by its coordinate functions

$$\gamma(t)=(x(t),y(t),z(t)).$$

Plane curves can be considered as a partial case of space curves with $z(t)\equiv 0$.

If each of the coordinate functions $x(t),y(t),z(t)$ of the space curve γ is a smooth (that is, it has derivatives of all orders everywhere in its domain) then the parameterized curve is called *smooth*.

If the *velocity vector*

$$\gamma'(t)=(x'(t),y'(t),z'(t))$$

does not vanish at all points, then the parameterized curve γ is called *regular*.

A simple space curve is called *smooth and regular* if it admits a smooth and regular parametrization correspondingly. Regular smooth curves are among the main objects in differential geometry; the term *smooth curve* often used for *smooth regular curve*.

1.4. Exercise. *The function*

$$f(t)=\begin{cases}0 & \text{if } t\leqslant 0,\\ \frac{t}{e^{1/t}} & \text{if } t>0.\end{cases}$$

is smooth.[1]

Show that $\gamma(t)=(f(t),f(-t))$ gives a smooth parametrization of the curve S formed by the union of two half-axis in the plane.

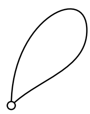
[1] The existence of all derivatives $f^{(n)}(x)$ at $x\ne 0$ is evident and direct calculations show that $f^{(n)}(0)=0$ for any n.

Show that any smooth parametrization of S has vanishing velocity vector at the origin. Conclude that the curve S is not regular and smooth.

1.5. Exercise. *Describe the set of real numbers a such that the plane curve $\gamma_a(t) = (t + a\cdot \sin t, a\cdot \cos t)$, $t \in \mathbb{R}$ is*
(a) regular;
(b) simple.

Loops and periodic parametrization. A closed simple curve can be described as an image of a parameterized curve $\gamma\colon [0, 1] \to \mathcal{X}$ such that $p = \gamma(0) = \gamma(1)$; such curves are called *loops*; the point p in this case is called *base* of the loop.

However, it is more natural to present it as a *periodic* parameterized cure $\gamma\colon \mathbb{R} \to \mathcal{X}$; that is, there is a constant ℓ such that $\gamma(t + \ell) = \gamma(t)$ for any t. For example the unit circle in the plane can be described by $2\cdot\pi$-periodic parametrization $\gamma(t) = (\cos t, \sin t)$.



Any smooth regular closed curve can be described by a smooth regular loop. But in general the closed curve that described by a smooth regular loop might fail to be smooth and regular — it might fail to be smooth at its base; an example shown on the diagram.

Implicitly defined curves

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a smooth function; that is, all its partial derivatives defined in its domain of definition. Consider the set S of solution of equation $f(x, y) = 0$ in the plane.

Assume S is path connected. According to implicit function theorem (A.10), the set S is a smooth regular simple curve if 0 is a *regular value* of f. In this case it means that the gradient ∇f does not vanish at any point $p \in S$. In other words, if $f(p) = 0$, then $\frac{\partial f}{\partial x}(p) \neq 0$ or $\frac{\partial f}{\partial y}(p) \neq 0$.

Similarly, assume f, h is a pair of smooth functions defined in $\mathbb{R}^3$. The system of equations $f(x, y, z) = h(x, y, z) = 0$ defines a regular smooth space curve if the set of solutions is path connected and 0 is a regular value of the map $F\colon (x, y, z) \mapsto (f(x, y, z), h(x, y, z))$. In this case it means that the gradients ∇f and ∇h are linearly independent at any point $p \in S$. In other words, if $f(p) = 0$, then at the Jacobian matrix

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}$$

for the map $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ has rank 2 at p.

The described way to define a curve is called *implicit*; if a curve is defined by its parametrization, we say that it is *explicitly defined.* While implicit function theorem guarantees the existence of regular smooth parametrizations, do not expect it to be in a closed form. When it comes to calculations, usually it is easier to work directly with implicit presentation.

1.6. Exercise. *Consider the set in the plane described by the equation*

$$y^2 = x^3.$$

Is it a simple curve? and if "yes", is it a smooth regular curve?

1.7. Exercise. *Describe the set of real numbers a such that the system of equations*

$$\begin{aligned} x^2 + y^2 + z^2 &= 1 \\ x^2 + a{\cdot}x + y^2 &= 0 \end{aligned}$$

describes a smooth regular curve.

Chapter 2

Length

Recall that a sequence

$$a = t_0 < t_1 < \dots < t_k = b.$$

is called a *partition* of the interval $[a, b]$.

2.1. Definition. *Let $\alpha\colon [a, b] \to \mathcal{X}$ be a curve in a metric space. The* length *of a α is defined as*

$$\begin{aligned}\operatorname{length}\alpha = \sup\{|\alpha(t_0) - \alpha(t_1)| + |\alpha(t_1) - \alpha(t_2)| + \dots \\ \dots + |\alpha(t_{k-1}) - \alpha(t_k)|\},\end{aligned}$$

where the exact upper bound is taken over all partitions

$$a = t_0 < t_1 < \dots < t_k = b.$$

The length of α is a nonnegative real number or infinity; the curve α is called rectifiable *if its length is finite.*

The length of a closed curve is defined as the length of a corresponding loop. If a curve is defined on a open or closed-open interval then its length is defined as the exact upper bound for lengths of all its closed arcs.

If α is a space curve, then the above definition says that it length is the exact upper bound of the lengths of polygonal lines $p_0 \dots p_k$ *inscribed* in the curve, where $p_i = \alpha(t_i)$ for a partition $a = t_0 < t_1 < \dots < t_k = b$. If α is closed then $p_0 = p_k$ and therefore the inscribed polygonal line is also closed.

2.2. Exercise. *Let $\alpha\colon [0, 1] \to \mathbb{R}^3$ be a simple curve. Suppose a parametrized curve $\beta\colon [0, 1] \to \mathbb{R}^3$ has that same image as α; that is*

$\beta([0,1]) = \alpha([0,1])$. *Show that*

$$\operatorname{length}\beta \geqslant \operatorname{length}\alpha.$$

2.3. Exercise. *Assume $\alpha\colon [a,b] \to \mathbb{R}^3$ is a smooth curve. Show that*

(a) $\operatorname{length}\alpha \geqslant \int_a^b |\alpha'(t)|\cdot dt$*,*

(b) $\operatorname{length}\alpha \leqslant \int_a^b |\alpha'(t)|\cdot dt$*.*

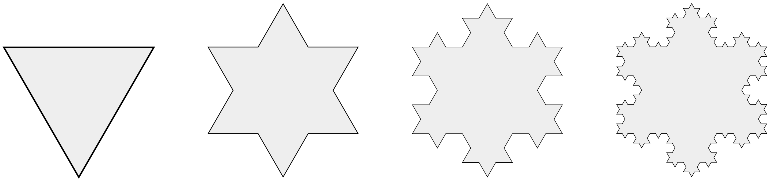
Conclude that

$$\operatorname{length}\alpha = \int\limits_a^b |\alpha'(t)|\cdot dt. \qquad ❶$$

Hints: For (a), apply the fundamental theorem of calculus for each segment in a given partition. For (b) consider a partition such that the velocity vector $\alpha'(t)$ is nearly constant on each of its segments.

Nonrectifiable curves. A classical example of a nonrectifiable curve is the so called *Koch snowflake*; it is a fractal curve that can be constructed the following way:

Start with an equilateral triangle, divide each of its side into three segments of equal length and add an equilateral triangle with base at the middle segment. Repeat this construction recursively to the obtained polygons. Few first iterations of the construction are shown



on the diagram. The Koch snowflake is the boundary of the union of all the polygons.

2.4. Exercise.

(a) Show that Koch snowflake is a closed simple curve; that is, it admits a homeomorphism to a circle.

(b) Show that Koch snowflake is not rectifiable.

Arc length parametrization

We say that a parametrized curve γ has an *arc length parametrization*[1] if for any two values of parameters $t_1 < t_2$, the value $t_2 - t_1$ is the length of $\gamma|_{[t_1,t_2]}$; that is, the closed arc of γ from t_1 to t_2.

Note that a smooth space curve $\gamma(t) = (x(t), y(t), z(t))$ has arc length parametrization if and only if it has unit velocity vector at all times; that is

$$|\gamma'(t)| = \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} = 1;$$

by that reason arc length parametrization of smooth curves with also called *unit-speed curves*. Note that smooth unit-speed curves are automatically regular.

Any rectifiable curve can be parameterized by arc length. For a parametrized smooth curve γ, the arc length parameter s can be written as an integral

$$s(t) = \int\limits_{t_0}^{t} |\gamma'(\tau)|\cdot d\tau.$$

Note that $s(t)$ is a smooth increasing function. Further by fundamental theorem of calculus, $s'(t) = |\gamma'(t)|$. Therefore if γ is regular, then $s'(t) \neq 0$ for any parameter value t. By inverse function theorem (A.9) the inverse function $s^{-1}(t)$ is also smooth. Therefore $\gamma \circ s^{-1}$ — the reparametrization of γ by arclength s — remains smooth and regular.

Most of the time we use s for an arc length parameter of a curve.

2.5. Exercise. *Reparametrize the helix*

$$\gamma_{a,b}(t) = (a\cdot \cos t, a\cdot \sin t, b\cdot t)$$

by arc length.

We will be interested in the properties of curves that are invariant under a reparametrization. Therefore we can always assume that the given smooth regular curve comes with a arc length parametrization. A good property of arc length parametrizations is that it is almost canonical — these parametrizations differ only by a sign and additive constant. On the other hand, often it is impossible to find an arc length parametrization in a closed form which makes it hard to use it calculations; usually it is more convenient to use the original parametrization.

[1] which is also called *natural parametrization*

Convex curves

A simple plane curve is called *convex* if it bounds a convex region.

2.6. Proposition. *Assume a convex closed curve α lies inside the domain bounded by a closed simple plane curve β. Then*

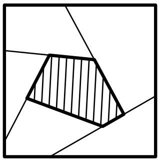
$$\operatorname{length}\alpha \leqslant \operatorname{length}\beta.$$

Note that it is sufficient to show that for any polygon P inscribed in α there is a polygon Q inscribed in β with $\operatorname{perim} P \leqslant \operatorname{perim} Q$, where $\operatorname{perim} P$ denotes the perimeter of P.

Therefore it is sufficient to prove the following lemma.

2.7. Lemma. *Let P and Q be polygons. Assume P is convex and $Q \supset P$. Then*

$$\operatorname{perim} P \leqslant \operatorname{perim} Q.$$



Proof. Note that by the triangle inequality, the inequality

$$\operatorname{perim} P \leqslant \operatorname{perim} Q$$

holds if P can be obtained from Q by cutting it along a chord; that is, a line segment with ends on the boundary of Q that lies in Q.

Note that there is an increasing sequence of polygons

$$P = P_0 \subset P_1 \subset \dots \subset P_n = Q$$

such that P_{i-1} obtained from P_i by cutting along a chord. Therefore

$$\begin{aligned}\operatorname{perim} P = \operatorname{perim} P_0 \leqslant \operatorname{perim} P_1 \leqslant \dots\\ \dots \leqslant \operatorname{perim} P_n = \operatorname{perim} Q\end{aligned}$$

and the lemma follows. □

2.8. Corollary. *Any convex closed plane curve is rectifiable.*

Proof. Any closed curve is bounded; that is, it lies in a sufficiently large square. Indeed the curve can be described as an image of a loop $\alpha\colon [0,1] \to \mathbb{R}^2$, $\alpha(t) = (x(t), y(t))$. The coordinate functions $x(t)$ and $y(t)$ are continous functions defined on $[0,1]$. Therefore the absolute values of both of these functions are bounded by some constant C. That is α lies in the square defined by the inequalities $|x| \leqslant C$ and $|y| \leqslant C$.

By Proposition 2.6, the length of the curve can not exceed the perimeter of the square $8{\cdot}C$, whence the result. $\square$

Recall that convex hull of a set X is the smallest convex set that contains X; in other words convex hull is the intersection of all convex sets containing X.

2.9. Exercise. *Let α be a closed simple plane curve. Denote by K the convex hull of α; let β be the boundary curve of K. Show that*

$$\operatorname{length}\alpha \geqslant \operatorname{length}\beta.$$

Try to show that the statement holds for arbitrary closed plane curve α, assuming that X has nonempty interior.

Crofton formulas*

Consider a plane curve $\alpha\colon [a,b] \to \mathbb{R}^2$. Given a unit vector u, denote by α_u the curve that follows orthogonal projections of α to the line in the direction u; that is

$$\alpha_u(t) = \langle u, \alpha(t)\rangle{\cdot}u.$$

Note that

$$|\alpha'(t)| = |\langle u, \alpha'(t)\rangle|$$

for any t. Note that for any plane vector the magnitude of its average projection is proportional to its magnitude with coefficient; that is,

$$|w| = k{\cdot}\overline{|w_u|},$$

where $\overline{|w_u|}$ denotes the average value of $|w_u|$ for all unit vectors u. (The value k is the average value of $|\cos\varphi|$ for $\varphi \in [0, 2{\cdot}\pi]$; it can be found by integration, but soon we will show another way to find it.)

If the curve α is smooth, then according to Exercise 2.3

$$\begin{aligned}\operatorname{length}\alpha &= \int\limits_a^b |\alpha'(t)|{\cdot}dt = \\ &= \int\limits_a^b k{\cdot}\overline{|\alpha_u'(t)|}{\cdot}dt = \\ &= k{\cdot}\overline{\operatorname{length}\alpha_u}.\end{aligned}$$

This formula and its relatives are called Crofton formulas. To find the coefficient k one can apply it for the unit circle: the left hand side is $2{\cdot}\pi$ — this is the length of unit circle. Note that for any unit vector u, the curve α_u runs back and forth along an interval of length 2. Therefore $\operatorname{length}\alpha_u = 4$ and hence its average value is also 4. It follows that the coefficient k has to satisfy the equation $2{\cdot}\pi = k{\cdot}4$; whence

$$\operatorname{length}\alpha = \tfrac{\pi}{2}{\cdot}\overline{\operatorname{length}\alpha_u}.$$

The Crofton's formula holds for arbitrary rectifiable curves, not necessary smooth; it can be proved using Exercises 2.12.

2.10. Exercise. *Show that any closed plane curve α has length at least $\pi{\cdot}s$, where s is the average of pojections of α to lines. Moreover the equality holds if and only if α is convex.*

Use this statement to give another solution of Exercise 2.9.

2.11. Advanced exercise. *Show that the length of space curve is proportional to the average length of its projections to all lines and to planes. Find the coefficients in each case.*

2.12. Advanced exercises.

(a) Show that the formula ❶ holds for any Lipschitz curve $\alpha\colon [a,b]\to$ $\to\mathbb{R}^3$.

(b) Construct a simple curve $\alpha\colon [a,b]\to\mathbb{R}^3$ such that the velocity vector $\alpha'(t)$ is defined and bounded for almost all $t\in[a,b]$, but the formula ❶ does not hold.

Hint: Use theorems of Rademacher and Lusin (A.13 and A.14).

Semicontinuity of length

Recall that the lower limit of a sequence of real numbers (x_n) is denoted by

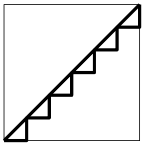
$$\varliminf_{n\to\infty} x_n.$$

It is defined as the lowest partial limit; that is, the lowest possible limit of a subsequence of (x_n). The lower limit is defined for any sequence of real numbers and it lies in the exteded real line $[-\infty,\infty]$

2.13. Theorem. *Length is a lower semi-continuous with respect to pointwise convergence of curves.*

More precisely, assume that a sequence of curves $\alpha_n\colon [a,b]\to\mathcal{X}$ in a metric space $\mathcal{X}$ converges pointwise to a curve $\alpha_\infty\colon [a,b]\to\mathcal{X}$; that is, $\alpha_n(t)\to\alpha_\infty(t)$ for any fixed $t\in[a,b]$ as $n\to\infty$. Then

❷ $$\varliminf_{n\to\infty} \operatorname{length}\alpha_n \geqslant \operatorname{length}\alpha_\infty.$$



Note that the inequality ❷ might be strict. For example the diagonal α_∞ of the unit square can be approximated by a sequence of stairs-like polygonal curves α_n with sides parallel to the sides of the square (α_6 is on the picture). In this case

$$\operatorname{length}\alpha_\infty = \sqrt{2} \quad \text{and} \quad \operatorname{length}\alpha_n = 2$$

for any n.

Proof. Fix a partition $a = t_0 < t_1 < \dots < t_k = b$. Set

$$\Sigma_n := |\alpha_n(t_0) - \alpha_n(t_1)| + \dots + |\alpha_n(t_{k-1}) - \alpha_n(t_k)|.$$
$$\Sigma_\infty := |\alpha_\infty(t_0) - \alpha_\infty(t_1)| + \dots + |\alpha_\infty(t_{k-1}) - \alpha_\infty(t_k)|.$$

Note that $\Sigma_n \to \Sigma_\infty$ as $n \to \infty$ and $\Sigma_n \leqslant \operatorname{length}\alpha_n$ for each n. Hence

❸ $$\varliminf_{n\to\infty} \operatorname{length}\alpha_n \geqslant \Sigma_\infty.$$

If α_∞ is rectifiable, we can assume that

$$\operatorname{length}\alpha_\infty < \Sigma_\infty + \varepsilon.$$

for any given $\varepsilon > 0$. By ❸ it follows that

$$\varliminf_{n\to\infty} \operatorname{length}\alpha_n > \operatorname{length}\alpha_\infty - \varepsilon$$

for any $\varepsilon > 0$; whence ❷ follows.

It remains to consider the case when α_∞ is not rectifiable; that is $\operatorname{length}\alpha_\infty = \infty$. In this case we can choose a partition so that $\Sigma_\infty > L$ for any real number L. By ❸ it follows that

$$\varliminf_{n\to\infty} \operatorname{length}\alpha_n > L$$

for any L; whence

$$\varliminf_{n\to\infty} \operatorname{length}\alpha_n = \infty$$

and ❷ follows. □

Length metric

Let $\mathcal{X}$ be a metric space. Given two points x, y in $\mathcal{X}$, denote by $d(x, y)$ the exact lower bound for lengths of all paths connecting x to y; if there is no such path we assume that $d(x, y) = \infty$.

Note that function d satisfies all the axioms of metric except it might take infinite value. Therefore if any two points in $\mathcal{X}$ can be connected by a rectifiable curve, then d defines a new metric on $\mathcal{X}$; in this case d is called *induced length metric.*

Evidently $d(x,y) \geqslant |x-y|$ for any pair of points $x, y \in \mathcal{X}$. If the equality holds for any pair, then then the metric is called *length metric* and the space is called *length-metric space.*

Most of the time we consider length-metric spaces. In particular the Euclidean space is a length-metric space. A subspaces A of length-metric space $\mathcal{X}$ might be not a lenght-metric space; the induced length distance between points x and y in the subspace A will be denoted as $|x-y|_A$; that is $|x-y|_A$ is the exact lower bound for the length of paths in A.

2.14. Exercise. *Let $A \subset \mathbb{R}^3$ be a closed subset. Show that A is convex if and only if*

$$|x-y|_A = |x-y|_{\mathbb{R}^3}.$$

2.15. Exercise. *Let us denote by $\mathbb{S}^1$ the unit circle in the plane; that is,*

$$\mathbb{S}^1 = \left\{\,(x,y) \in \mathbb{R}^3 \;:\; x^2+y^2=1\,\right\}.$$

Show that

$$|u-v|_{\mathbb{S}^1} = \measuredangle(u,v) := \arccos\langle u,v\rangle$$

for any $u, v \in \mathbb{S}^1$.

Spherical curves

A space curve γ is called *spherical* if it runs in the unit sphere; that is, $|\gamma(t)| = 1$ for any t.

2.16. Exercise. *Let us denote by $\mathbb{S}^2$ the unit sphere in the space; that is,*

$$\mathbb{S}^2 = \left\{\,(x,y,z) \in \mathbb{R}^3 \;:\; x^2+y^2+z^2=1\,\right\}.$$

Show that

$$|u-v|_{\mathbb{S}^2} = \measuredangle(u,v) := \arccos\langle u,v\rangle$$

for any $u, v \in \mathbb{S}^2$.

Hint: Use Exercise 2.15 and the following map $f\colon (r,\theta,\varphi) \mapsto (r,\theta,0)$ in spherical coordinates. Note that f is distance nonexpanding and it maps $\mathbb{R}^3$ to a half-plane and $\mathbb{S}^2$ to one of its meridians.

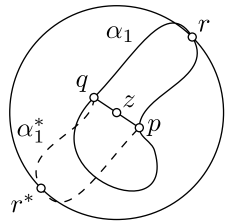
2.17. Hemisphere lemma. *Any closed curve of length $< 2\cdot\pi$ in $\mathbb{S}^2$ lies in an open hemisphere.*

This lemma is a keystone in the proof of Fenchel's theorem given below. The lemma is not as simple as you might think — try to prove it yourself. I learned the following proof from Stephanie Alexander.

Proof. Let α be a closed curve in $\mathbb{S}^2$ of length $2\cdot\ell$.

Assume $\ell < \pi$.

Let us divide α into two arcs α_1 and α_2 of length ℓ, with endpoints p and q. According to Exercise 2.16, $\measuredangle(p,q) \leqslant \ell < \pi$. Denote by z be the midpoint between p and q in $\mathbb{S}^2$; that is z is the midpoint of an equator arc from p to q. We claim that α lies in the open north hemisphere with north pole at z. If not, α intersects the equator in a point, say r. Without loss of generality we may assume that r lies on α_1.



The north hemisphere corresponds to the disc and the south hemisphere to the complement of the disc.

Rotate the arc α_1 by angle π around the line thru z and the center of the sphere. The obtained arc α_1^* together with α_1 forms a closed curve of length $2\cdot\ell$ that passes thru r and its antipodal point r^*. Therefore

$$\tfrac{1}{2}\cdot \operatorname{length}\alpha = \ell \geqslant \measuredangle(r,r^*) = \pi,$$

a contradiction. □

2.18. Exercise. *Describe a simple closed spherical curve that does not pass thru a pair of antipodal points and does not lie in any hemisphere.*

2.19. Exercise. *Suppose that a closed simple spherical curve α divides $\mathbb{S}^2$ into two regions of equal area. Show that*

$$\operatorname{length}\alpha \geqslant 2\cdot\pi.$$

2.20. Exercise. *Consider the following problem, find a flaw in the given solution. Come up with a correct argument.*

Problem. Suppose that a closed plane curve α has length at most 4. Show that α lies in a unit disc.

Wrong solution. Note that it is sufficient to show that diameter of α is at most 2; that is, the distance between any two pairs of points p and q of α an not exceed 2.

The length of α can not be smaller then the closed inscribed polygonal line which goes from p to q and back to p. Therefore

$$2\cdot|p-q| \leqslant \operatorname{length}\alpha \leqslant 4.$$

□

2.21. Advanced exercises. *Given points $v, w \in \mathbb{S}^2$, denote by w_v the closest point to w_u on the equator with pole at v; in other words, if $w^\perp$ is the projection of w to the plane perpendicular to v, then w_v is the unit vector in the direction of $w^\perp$. The vector w_v is defined if $w \neq \pm v$.*

1. *Show that for any spherical curve α we have that*

$$\operatorname{length}\alpha = \overline{\operatorname{length}\alpha_v},$$

where $\overline{\operatorname{length}\alpha_v}$ denotes the average length for all $v \in \mathbb{S}^2$. (This is a spherical analog of Crofton's formula.)

2. *Give another proof of hemisphere lemme using part (1).*

Chapter 3

Space curves

Acceleration of unit-speed curve

Recall that any regular smooth curve can be parametrized by its arc length. The obtained parametrized curve, say γ, remains to be smooth and it has unit speed; that is, $|\gamma'(s)| = 1$ for all s.

The following proposition states that the acceleration vector is perpendicular to the velocity vector if the speed remains constant.

3.1. Proposition. *Assume γ is a smooth unit-speed space curve. Then $\gamma'(s) \perp \gamma''(s)$ for any s.*

The scalar product (also known as dot product) of two vectors v and w will be denoted by $\langle v, w\rangle$. Recall that the derivative of a scalar product satisfies the product rule; that is if $v = v(t)$ and $w = w(t)$ are smooth vector-valued functions of a real parameter t, then

$$\langle v, w\rangle' = \langle v', w\rangle + \langle v, w'\rangle.$$

Proof. The identity $|\gamma'| = 1$ can be rewritten as $\langle \gamma', \gamma'\rangle = 1$. Therefore

$$2{\cdot}\langle \gamma'', \gamma'\rangle = \langle \gamma', \gamma'\rangle' = 0,$$

whence $\gamma'' \perp \gamma'$. $\square$

Curvature

For a unit speed smooth space curve γ the magnitude of its acceleration $|\gamma''(s)|$ is called its *curvature* at the time s. If γ is simple, then we can say that $|\gamma''(s)|$ is the curvature at the point $p = \gamma(s)$ without ambiguity. The curvature is usually denoted by $k(s)$ or $k(s)_\gamma$ and in the latter case it might be also denoted by $k(p)$ or $k(p)_\gamma$.

The curvature measures how fast the curve turns; if you drive along a plane curve, curvature tells how much to turn the steering wheel at the given point (note that it does not depend on your speed). In general, the term *curvature* is used for different types of geometric objects, and it always measures how much it deviates from being *straight*; for curves, it measures how fast it deviates from a straight line.

3.2. Exercise. *Show that any regular smooth spherical curve has curvature at least 1 at each time.*

Hint: Differentiate the identity $\langle\gamma(s),\gamma(s)\rangle = 1$ a couple of times.

Tangent indicatrix

Let γ be a regular smooth space curve. Let us consider another curve

❶ $$\tau(t) = \tfrac{\gamma'(t)}{|\gamma'(t)|}$$

that is called *tangent indicatrix* of γ. Note that $|\tau(t)| = 1$ for any t; that is, τ is a spherical curve.

The line thru $\gamma(s)$ in the direction of $\tau(s)$ is called *tangent line* at s.

We say that smooth regular curve γ_1 at s_1 is *tangent* to a smooth regular curve γ_2 at s_2 if $\gamma_1(s_1) = \gamma_2(s_2)$ and the tangent line of γ_1 at s_1 coinside with the tangent line of γ_2 at s_2; if both of the curves are simple we can also say that they are tangent at the point $p = \gamma_1(s_1) = \gamma_2(s_2)$ without ambiguity.

If γ has a unit speed parametrization, then $\tau(t) = \gamma'(t)$. In this case we have the following expression for curvature: $k(t) = |\tau'(t)| = |\gamma''(t)|$.

In general case we have

❷ $$k(t) = \tfrac{|\tau'(t)|}{|\gamma'(t)|}.$$

Indeed, for an arc length parametrization $s(t)$ we have $s'(t) = |\gamma'(t)|$. Therefore

$$\begin{aligned} k(t) &= \left|\tfrac{d\tau}{ds}\right| = \\ &= \left|\tfrac{d\tau}{dt}\right| / \left|\tfrac{ds}{dt}\right| = \\ &= \tfrac{|\tau'(t)|}{|\gamma'(t)|}. \end{aligned}$$

It follows that indicatrix of a smooth regular curve γ is regular if the curvature of γ does not vanish.

3.3. Exercise. *Use the formulas ❶ and ❷ to show that for any smooth regular space curve γ we have the following expressions for its curvature:*

(a)

$$k(t) = \frac{|\gamma''(t)^{\perp}|}{|\gamma'(t)|^2},$$

where $\gamma''(t)^{\perp}$ denotes the projection of $\gamma''(t)$ to the normal plane of $\gamma'(t)$;

(b)

$$k(t) = \frac{|\gamma''(t) \times \gamma'(t)|}{|\gamma'(t)|^3},$$

where $\times$ denotes the vector product *(also known as* cross product*).*

Hint: Prove and use the following identities:

$$\gamma''(t) - \gamma''(t)^{\perp} = \tfrac{\gamma'(t)}{|\gamma'(t)|} \cdot \langle \gamma''(t), \tfrac{\gamma'(t)}{|\gamma'(t)|} \rangle,$$
$$|\gamma'(t)| = \sqrt{\langle \gamma'(t), \gamma'(t) \rangle}.$$

3.4. Exercise. *Apply the formulas in the previous exercise to show that if f is a smooth real function, then its graph $y = f(x)$ has curvature*

$$k(p) = \frac{|f''(x)|}{(1 + f'(x)^2)^{\frac{3}{2}}}$$

at the point $p = (x, f(x))$.

Total curvature

Let $\gamma \colon \mathbb{I} \to \mathbb{R}^3$ be a regular smooth curve and τ its tangent indicatrix. Recall that without loss of generality we can assume that γ has a unit speed parametrization; in this case $\tau(s) = \gamma'(s)$ and hence

$$k(s) := |\gamma''(s)| = |\tau'(s)|;$$

that is, the curvature of γ at time s is the speed of the tangent indicatrix τ at the same time moment.

The integral

$$\Phi(\gamma) := \int\limits_{\mathbb{I}} k(s) {\cdot} ds$$

is called *total curvature of γ*.

3.5. Exercise. *Find the curvature of the helix*

$$\gamma_{a,b}(t) = (a\cdot \cos t, a\cdot \sin t, b\cdot t),$$

its tangent indicatrix and the total curvature of its arc $t \in [0, 2\cdot\pi]$.

3.6. Observation. *The total curvature of a smooth regular curve is the length of its tangent indicatrix.*

Proof. It is sufficient to prove the observation for a unit-speed space curve $\gamma\colon \mathbb{I} \to \mathbb{R}^3$. Denote by τ its tangent indicatrix. Then

$$\begin{aligned}\Phi(\gamma) &:= \int\limits_{\mathbb{I}} k(s)\cdot ds = \\ &= \int\limits_{\mathbb{I}} |\tau'(s)|\cdot ds = \\ &= \operatorname{length}\tau.\end{aligned}$$

□

3.7. Fenchel's theorem. *The total curvature of any closed regular space curve is at least* $2\cdot\pi$.

Proof. Fix a closed regular space curve γ; we can assume that it is described by a loop $\gamma\colon [a, b] \to \mathbb{R}^3$; in this case $\gamma(a) = \gamma(b)$ and $\gamma'(a) = \gamma'(b)$.

Consider its tangent indicatrix $\tau = \gamma'$. Recall that $|\tau(s)| = 1$ for any s; that is, τ is a closed spherical curve.

Let us show that τ can not lie in a hemisphere. Assume the contrary; without loss of generality we can assume that τ lies in the north hemisphere defined by the inequality $z > 0$ in (x, y, z)-coordinates. It means that $z'(t) > 0$ at any time, where $\gamma(t) = (x(t), y(t), z(t))$. Therefore

$$z(b) - z(a) = \int\limits_a^b z'(s)\cdot ds > 0.$$

In particular, $\gamma(a) \neq \gamma(b)$, a contradiction.

Applying the observation (3.6) and the hemisphere lemma (2.17), we get that

$$\Phi(\gamma) = \operatorname{length}\tau \geqslant 2\cdot\pi.$$

□

3.8. Exercise. *Show that a closed space curve* γ *with curvature at most* 1 *can not be shorter than the unit circle; that is,* $\operatorname{length}\gamma \geqslant 2\cdot\pi$.

3.9. Advanced exercise. *Suppose that γ is a smooth regular space curve that does not pass thru the origin. Consider the spherical curve defined as $\sigma(t)=\frac{\gamma(t)}{|\gamma(t)|}$ for any t. Show that*

$$\operatorname{length}\sigma < \Phi(\gamma)+\pi.$$

Moreover, if γ is closed, then

$$\operatorname{length}\sigma \leqslant \Phi(\gamma).$$

Note that the last inequality gives an alternative proof of Fenchel's theorem. Indeed, without loss of generality we can assume that the origin lies on a chord of γ; in this case the spherical curve σ passes thru a pair of antipodal points in $\mathbb{S}^2$; whence

$$\operatorname{length}\sigma \geqslant 2\cdot\pi.$$

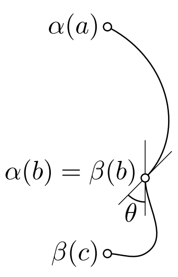
Piecewise smooth curves

Assume $\alpha\colon [a,b]\to\mathbb{R}^3$ and $\beta\colon [b,c]\to\mathbb{R}^3$ are two curves such that $\alpha(b)=\beta(b)$. Then one can combine these two curves into one $\gamma\colon [a,c]\to\mathbb{R}^3$ by the rule

$$\gamma(t)=\begin{cases}\alpha(t) & \text{if} \quad t\leqslant b,\\ \beta(t) & \text{if} \quad t\geqslant b.\end{cases}$$

The obtained curve γ is called the *concatenation* of α and β. (The condition $\alpha(b)=\beta(b)$ ensures that the map $t\mapsto\gamma(t)$ is continuous.)

The same definition of cancatination can be applied if α and/or β are defied on semiopen intervals $(a,b]$ and/or $[b,c)$.



The concatenation can be also defined if the end point of the first curve coincides with the starting point of the second curve; if this is the case, then the time intervals of both curves can be shifted so that they fit together.

If in addition $\beta(c)=\alpha(a)$ then we can do cyclic concatination of these curves; this way we obtain a closed curve.

If $\alpha'(b)$ and $\beta'(b)$ are defined then the angle $\theta=\measuredangle(\alpha'(b),\beta'(b))$ is called *external angle* of γ at time b.

A space curve γ is called *piecewise smooth and regular* if it can be presented as a concatination of finite number of smooth regular curves; if γ is closed, then the concatination is assumed to be cyclic.

If γ is a concatination of smooth regular arcs $\gamma_1, \dots, \gamma_n$, then the total curvature of γ is defined as a sum of the total curvatures of γ_i and the external angles; that is,

$$\Phi(\gamma) = \Phi(\gamma_1) + \dots + \Phi(\gamma_n) + \theta_1 + \dots + \theta_{n-1}$$

where θ_i is the external angle at the joint γ_i and γ_{i+1}; if γ is closed, then

$$\Phi(\gamma) = \Phi(\gamma_1) + \dots + \Phi(\gamma_n) + \theta_1 + \dots + \theta_n,$$

where θ_n is the external angle at the joint γ_n and γ_1.

3.10. Generalized Fenchel's theorem. *Let γ be a closed piecewise smooth regular space curve. Then*

$$\Phi(\gamma) \geqslant 2\cdot\pi.$$

Proof. Suppose γ is a cyclic concatenation of n smooth regular arcs $\gamma_1, \dots, \gamma_n$. Denote by $\theta_1, \dots, \theta_n$ its external angles. We need to show that

$$\Phi(\gamma_1) + \dots + \Phi(\gamma_n) + \theta_1 + \dots + \theta_n \geqslant 2\cdot\pi. \qquad ❸$$

Consider the tangent indicatrix $\tau_1, \dots, \tau_n$ for each arc $\gamma_1, \dots, \gamma_n$; these are smooth spherical arcs.

The same argument as in the proof of Fenchel's theorem, shows that the curves $\tau_1, \dots, \tau_n$ can not lie in an open hemisphere.

Note that the spherical distance from the end point of τ_i to the starting point of τ_{i+1} is equal to the external angle θ_i (we enumerate modulo n, so $\gamma_{n+1} = \gamma_1$). Therefore if we connect the end point of τ_i to the starting point of τ_{i+1} by a short arc of a great circle in the sphere, then we add $\theta_1 + \dots + \theta_n$ to the total length of $\tau_1, \dots, \tau_n$.

Applying the hemisphee lemma (2.17) to the obtained closed curve, we get that

$$\operatorname{length} \tau_1 + \dots + \operatorname{length} \tau_n + \theta_1 + \dots + \theta_n \geqslant 2\cdot\pi.$$

Applying the observation (3.6), we get ❸. □

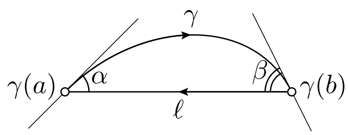
3.11. Chord lemma. *Let ℓ be the chord to a smooth regular arc $\gamma\colon [a, b] \to \mathbb{R}^3$. Assume γ meets ℓ at angles α and β at its ends; that is*

$$\alpha = \measuredangle(w, \gamma'(a)) \quad \textit{and} \quad \beta = \measuredangle(w, \gamma'(b)),$$

where $w = \gamma(b) - \gamma(a)$. Then

$$\Phi(\gamma) \geqslant \alpha + \beta.$$

Proof. Let us parameterize the chord ℓ from $\gamma(b)$ to $\gamma(a)$ and consider the cyclic concatenation $\bar\gamma$ of γ and ℓ. The closed curve $\bar\gamma$ has two external angles $\pi-\alpha$ and $\pi-\beta$. Since curvature of ℓ vanish, we get that



$$\Phi(\bar\gamma)=\Phi(\gamma)+(\pi-\alpha)+(\pi-\beta).$$

According to the generalized Fenechel's theorem (3.10),

$$\Phi(\bar\gamma)\geqslant 2\cdot\pi;$$

hence the result. □

3.12. Exercise. *Show that the estimate in the chord lemma is optimal.*

That is, given two points p,q and two nonzero vectors u,v in $\mathbb{R}^3$, show that there is a smooth regular curve γ that starts at p in the direction of u and ends at q in the direction of v such that $\Phi(\gamma)$ is arbitrarily close to $\measuredangle(w,u)+\measuredangle(w,v)$, where $w=q-p$.

Polygonal lines

Polygonal lines are partial case of piecewise smooth regular curves; each arc in its concatenation is a line segment. Since the curvature of a line segment vanish, the total curvature of polygonal line is the sum of its external angles.

3.13. Exercise. *Let a,b,c,d and x be distinct points in $\mathbb{R}^3$. Show that the total curvature of polygonal line $abcd$ can not exceed the total curvature of $abxcd$; that is,*

$$\Phi(abcd)\le\Phi(abxcd).$$

Use this statement to show that any closed polygonal line has curvature at least $2\cdot\pi$.

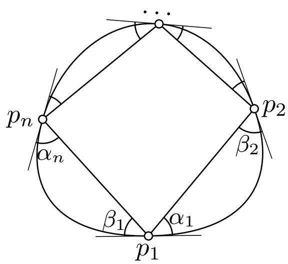
Hint: Use that exterior angle of a triangle equals to the sum of the two remote interior angles; for the second part apply the induction on number of vertexes.

3.14. Proposition. *Assume a polygonal line $\hat\gamma=p_1\dots p_n$ is inscribed in a smooth regular curve γ. Then*

$$\Phi(\gamma)\geqslant\Phi(\hat\gamma).$$

Moreover if γ is closed we can assume that the inscribed polygonal line $\hat\gamma$ is also closed.

Proof. Since the curvature of line segments vanishes, the total curvature of polygonal line is the sum of external angles $\theta_i = \pi - \measuredangle p_{i-1}p_ip_{i+1}$.



Assume $p_i = \gamma(t_i)$. Set

$$w_i = p_{i+1} - p_i, \qquad v_i = \gamma'(t_i),$$
$$\alpha_i = \measuredangle(w_i, v_i), \qquad \beta_i = \measuredangle(w_{i-1}, v_i).$$

In case of closed curve we use indexes modulo n, in particular $p_{n+1} = p_1$.

Note that $\theta_i = \measuredangle(w_{i-1}, w_i)$. Therefore

$$\theta_i \leqslant \alpha_i + \beta_i.$$

By the chord lemma, the total curvature of the arc of γ from p_i to p_{i+1} is at least $\alpha_i + \beta_{i+1}$.

Therefore if γ is a closed curve, we have

$$\begin{aligned}\Phi(\hat\gamma) &= \theta_1 + \dots + \theta_n \leqslant \\ &\leqslant \beta_1 + \alpha_1 + \dots + \beta_n + \alpha_n = \\ &= (\alpha_1 + \beta_2) + \dots + (\alpha_n + \beta_1) \leqslant \\ &\leqslant \Phi(\gamma).\end{aligned}$$

If γ is an arc, the argument is analogous:

$$\begin{aligned}\Phi(\hat\gamma) &= \theta_2 + \dots + \theta_{n-1} \leqslant \\ &\leqslant \beta_2 + \alpha_2 + \dots + \beta_{n-1} + \alpha_{n-1} \leqslant \\ &\leqslant (\alpha_1 + \beta_2) + \dots + (\alpha_{n-1} + \beta_n) \leqslant \\ &\leqslant \Phi(\gamma).\end{aligned}$$

□

3.15. Exercise.

(a) *Draw a smooth regular plane curve γ which has a self-intersection, such that $\Phi(\gamma) < 2\cdot\pi$.*

(b) *Show that if a smooth regular curve $\gamma\colon [a,b] \to \mathbb{R}^3$ has a self-intersection, then $\Phi(\gamma) > \pi$.*

3.16. Proposition. *The equality case in the Fenchel's theorem holds only for convex plane curves; that is, if the total curvature of a smooth regular space curve γ is equal to $2\cdot\pi$, then it is a convex plane curve.*

The proof is an application of Proposition 3.14.

Proof. Consider an inscribed quadraliteral $abcd$ in γ. By the definition of total curvature, we have that

$$\begin{aligned}\Phi(abcd) &= (\pi - \measuredangle dab) + (\pi - \measuredangle abc) + (\pi - \measuredangle bcd) + (\pi - \measuredangle cda) = \\ &= 4{\cdot}\pi - (\measuredangle dab + \measuredangle abc + \measuredangle bcd + \measuredangle cda))\end{aligned}$$

Note that

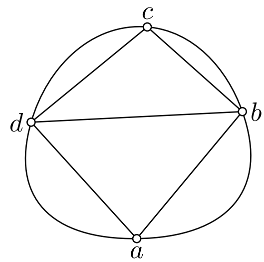
❹ $$\measuredangle abc \leqslant \measuredangle abd + \measuredangle dbc \quad\text{and}\quad \measuredangle cda \leqslant \measuredangle cdb + \measuredangle bda.$$

The sum of angles in any triangle is π. Therefore combining these inequalities, we get that

$$\begin{aligned}&\Phi(abcd) \geqslant 4{\cdot}\pi - (\measuredangle dab + \measuredangle abd + \measuredangle bda) - (\measuredangle bcd + \measuredangle cdb + \measuredangle dbc) = \\ &= 2{\cdot}\pi.\end{aligned}$$

By 3.14,

$$\Phi(abcd) \leqslant \Phi(\gamma) \leqslant 2{\cdot}\pi.$$

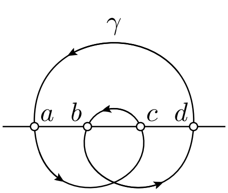


Therefore we have equalities in ❹. It means that the point d lies in the angle abc and the point b lies in the angle cda. That is, $abcd$ is a convex plane quadraliteral.

It follows that any quadraliteral inscribed in γ is convex plane quadraliteral. Therefore all points of γ lie in one plane and the domain bounded by γ is convex; that is, γ is a convex plane curve. □

3.17. Exercise. *Suppose that a closed curve γ crosses a line at four points a, b, c and d. Assume that these points appear on the line in the order a, b, c, d and they appear on the curve γ in the order a, c, b, d. Show that*

$$\Phi(\gamma) \geqslant 4{\cdot}\pi.$$

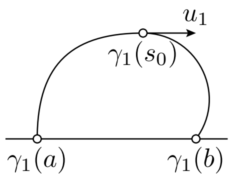


A line crossing a curve at four points as in the exercise is called *alternating quadrisecants.* It turns out that any *nontrivial knot* admits an alternating quadrisecants [1]; it implies the so called Fáry–Milnor theorem — *the total curvature any knot exceeds* $4{\cdot}\pi$.

Bow lemma

3.18. Lemma. *Let $\gamma_1\colon [a,b]\to\mathbb{R}^2$ and $\gamma_2\colon [a,b]\to\mathbb{R}^3$ be two smooth unit-speed curves; denote by $k_1(s)$ and $k_2(s)$ their curvatures at s. Suppose that $k_1(s)\geqslant k_2(s)$ for any s and the curve γ_1 is a simple arc of a convex curve; that is, it runs in the boundary of a covex plane figure. Then the distance between the ends of γ_1 can not exceed the distance between the ends of γ_2; that is,*

$$|\gamma_1(b)-\gamma_1(a)|\leqslant|\gamma_2(b)-\gamma_2(a)|.$$



Proof. Denote by τ_1 and τ_2 the tangent indicatrixes of γ_1 and γ_2 correspondingly.

Let $\gamma_1(s_0)$ be the point on γ_1 that maximize the distance to the line thru $\gamma(a)$ and $\gamma(b)$. Consider two unit vectors

$$u_1=\tau_1(s_0)=\gamma_1'(s_0)\quad\text{and}\quad u_2=\tau_2(s_0)=\gamma_2'(s_0).$$

By construction the vector u_1 is parallel to $\gamma(b)-\gamma(a)$ in particular

$$|\gamma_1(b)-\gamma_1(a)|=\langle u_1,\gamma_1(b)-\gamma_1(a)\rangle$$

Since γ_1 is an arc of convex curve, it indicatrix $\tau(s)$ runs in one direction along the unit circle. Suppose $s\leqslant s_0$, then

$$\begin{aligned}
\measuredangle(\gamma_1'(s),u_1)&=\measuredangle(\tau_1(s),\tau_1(s_0))=\\
&=\operatorname{length}(\tau_1|_{[s,s_0]})=\\
&=\int_s^{s_0}|\tau_1'(t)|\cdot dt=\\
&=\int_s^{s_0}k_1(t)\cdot dt\geqslant\\
&\geqslant\int_s^{s_0}k_2(t)\cdot dt=\\
&=\int_s^{s_0}|\tau_2'(t)|\cdot dt=\\
&=\operatorname{length}(\tau_1|_{[s,s_0]})\geqslant\\
&\geqslant\measuredangle(\tau_2(s),\tau_2(s_0))=\\
&=\measuredangle(\gamma_2'(s),u_2),
\end{aligned}$$

The same argument shows that

$$\measuredangle(\gamma_1'(s), u_1) \geqslant \measuredangle(\gamma_2'(s), u_2)$$

for $s \geqslant s_0$; therefore the inequality holds for any s. Since the vectors $\gamma_1'(s), u_1, \gamma_2'(s), u_2$ are unit, it follows that

$$\langle \gamma_1'(s), u_1 \rangle \leqslant \langle \gamma_2'(s), u_2 \rangle.$$

Integrating the last inequality, we get that

$$\begin{aligned}
|\gamma_1(b) - \gamma_1(a)| &= \langle u_1, \gamma_1(b) - \gamma_1(a) \rangle = \\
&= \int_a^b \langle u_1, \gamma_1'(s) \rangle \cdot ds \leqslant \\
&\leqslant \int_a^b \langle u_2, \gamma_2'(s) \rangle \cdot ds = \\
&= \langle u_2, \gamma_2(b) - \gamma_2(a) \rangle \leqslant \\
&\leqslant |\gamma_2(b) - \gamma_2(a)|.
\end{aligned}$$

Hence the result. □

3.19. Exercise. *Let $\gamma\colon [a, b] \to \mathbb{R}^3$ be a smooth regular curve and $0 < \theta \leqslant \frac{\pi}{2}$. Suppose*

$$\Phi(\gamma) \leqslant 2 \cdot \theta.$$

(a) Show that

$$|\gamma(b) - \gamma(a)| > \cos \theta \cdot \operatorname{length} \gamma.$$

(b) Use part (a) to give another solution of 3.15b.

(c) Show that the inequality in (a) is optimal; that is, given θ there is a smooth regular curve γ such that $\frac{|\gamma(b) - \gamma(a)|}{\operatorname{length} \gamma}$ is arbitrarily close to $\cos \theta$.

Hint: Choose a value $s_0 \in [a, b]$ that splits the total curvature into two equal parts, θ in each. Observe that $\measuredangle(\gamma'(s_0), \gamma'(s)) \leqslant \theta$ for any s. Use this inequality the same way as in the proof of the bow lemma.

3.20. Exercise. *Suppose that two points p and q lie on a unit circle and dividing it in two arcs with lengths $\ell_1 < \ell_2$. Show that if a curve γ runs from p to q and has curvature at most 1, then either*

$$\operatorname{length} \gamma \leqslant \ell_1 \quad \textit{or} \quad \operatorname{length} \gamma \geqslant \ell_2.$$

3.21. Exercise. *Suppose $\gamma\colon [a,b]\to\mathbb{R}^3$ is a smooth regular loop with curvature at most 1. Show that*

$$\operatorname{length}\gamma \geqslant 2\cdot\pi.$$

DNA inequality*

Recall that curvature of a spherical curve is at least 1 (Exercise 3.2). In particular the length of spherical curve can not exceed its total curvature. The following theorem shows that the same inequality holds for *closed* curves in a unit ball.

3.22. Theorem. *Let γ be a smooth regular closed curve that lies in a unit ball. Then*

$$\Phi(\gamma) \geqslant \operatorname{length}\gamma.$$

Proof. Without loss of generality we can assume the curve is described by a loop $\gamma\colon [0,\ell]\to\mathbb{R}^3$ parameterized by its arc length, so $\ell = \operatorname{length}\gamma$. We can also assume that the origin is the center of the ball. It follows that

$$\langle\gamma'(s),\gamma'(s)\rangle = 1, \qquad |\gamma(s)|\leqslant 1$$

and in particular

$$\begin{aligned}\langle\gamma''(s),\gamma(s)\rangle &\geqslant -|\gamma''(s)|\cdot|\gamma(s)| \geqslant\\ &\geqslant -k(s)\end{aligned}$$

for any s, where $k(s) = |\gamma''(s)|$ is the curvature of γ at s.

Since γ is closed, we have that $\gamma'(0) = \gamma'(\ell)$ and $\gamma(0) = \gamma(\ell)$. Therefore

$$\begin{aligned}0 &= \langle\gamma(\ell),\gamma'(\ell)\rangle - \langle\gamma(0),\gamma'(0)\rangle =\\ &= \int_0^\ell \langle\gamma(s),\gamma'(s)\rangle'\cdot ds =\\ &= \int_0^\ell \langle\gamma'(s),\gamma'(s)\rangle\cdot ds + \int_0^\ell \langle\gamma(s),\gamma''(s)\rangle\cdot ds \geqslant\\ &\geqslant \ell - \Phi(\gamma),\end{aligned}$$

whence the result. ☐

This theorem was proved by Don Chakerian [2]; for plane curves it was proved earlier by István Fáry [3]. We present the proof given by Don Chakerian in [4]; few other proofs of this theorem are discussed by Serge Tabachnikov [5]. He also conjectured the following closely related statement:

3.23. Theorem. *Suppose a closed regular smooth curve γ lies in a convex figure with the perimeter $2{\cdot}\pi$. Then*

$$\Phi(\gamma) \geqslant \operatorname{length}\gamma.$$

It was proved by Jeffrey Lagarias and Thomas Richardson [6]; latter a simpler proof was found by Alexander Nazarov and Fedor Petrov [7]. The proof is elementary, but annoyingly difficult; we do not present it here.

Nonsmooth curves*

3.24. Theorem. *For any regular smooth space curve γ we have that*

$$\Phi(\gamma) = \sup\{\Phi(\beta)\},$$

where the least upper bound is taken for all polygonal lines β inscribed in γ; if γ is closed we assume that so is β.

Proof. Note that the inequality

$$\Phi(\gamma) \geqslant \Phi(\beta)$$

follows from 3.14; it remains to show

❺ $$\Phi(\gamma) \leqslant \sup\{\Phi(\beta)\}.$$

Let $\gamma\colon [a,b] \to \mathbb{R}^3$ be a smooth curve. Fix a partition $a = t_0 <$ $< \dots < t_n = b$ and consider the corresponding inscribed polygonal line $\beta = p_0 \dots p_n$. (If γ is closed, then $p_0 = p_n$ and β is closed as well.)

Let $\chi = \xi_1 \dots \xi_n$ be a spherical polygonal line with the vertexes $\xi_i = \frac{p_i - p_{i-1}}{|p_i - p_{i-1}|}$. We can assume that χ has constant speed on each arc and $\chi(t_i) = \xi_i$ for each i. The spherical polygonal line χ will be called tangent indicatrix for β.

Consider a sequence of finer and finer partitions, denote by β_n and χ_n the corresponding inscribed polygonal lines and their tangent indicatrixes. Note that since γ is smooth, the idicatrixes χ_n converge pointwise to τ — the tangent indicatrix of γ. By semi-continuity of the length (2.13), we get that

$$\begin{aligned}\Phi(\gamma) &= \operatorname{length}\tau \leqslant \\ &\leqslant \varliminf_{n\to\infty} \operatorname{length}\chi_n = \\ &= \varliminf_{n\to\infty} \Phi(\beta_n) \leqslant \\ &\leqslant \sup\{\Phi(\beta)\},\end{aligned}$$

where the last supremum is taken over all partitions and their corresponding inscribed polygonal lines β; whence ❺ follows. □

The theorem above can be used to define total curvature for arbitrary curves, not necessary (piecewise) smooth and regular. We say that a parameterized curve is trivial if it is constant; that is, it stays at one point.

3.25. Definition. *The total curvature of a nontrivial parameterized space curve γ is the exact upper bound on the total curvatures of inscribed nondegenerate polygonal lines; if γ is closed then we assume that the inscribed polygonal lines are closed as well.*

3.26. Exercise. *Show that the total curvature is lower semicontinuous with respect to pointwise convergence of curves. That is, if a sequence of curves $\gamma_n\colon [a,b] \to \mathbb{R}^3$ converges pointwise to a nontrivial curve $\gamma_\infty\colon [a,b] \to \mathbb{R}^3$, then*

$$\varliminf_{n\to\infty} \Phi(\gamma_n) \geqslant \Phi(\gamma_\infty).$$

Hint: Modify the proof of semi-continuity of length (Theorem 2.13).

3.27. Exercise. *Show that Fenchel's theorem holds for any nontrivial closed curve γ; that is,*

$$\Phi(\gamma) \geqslant 2\cdot\pi.$$

3.28. Exercise. *Assume that a curve $\gamma\colon [a,b] \to \mathbb{R}^3$ has finite total curvature. Show that γ is rectifiable.*

Construct a rectifiable curve $\gamma\colon [a,b] \to \mathbb{R}^3$ that has infinite total curvature.

A good survey on curves of finite total curvature is written by John Sullivan [8].

Chapter 4

Torsion

Frenet frame

Let γ be a smooth regular space curve. Without loss of generality, we may assume that γ has arc length parametrization, so the velocity vector $\tau(s) = \gamma'(s)$ is unit.

Assume its curvature does not vanish at some time moment s; in other words, $\gamma''(s) \neq 0$. Then we can define the so called *normal vector* at s as

$$\nu(s) = \frac{\gamma''(s)}{|\gamma''(s)|}.$$

Note that

$$\tau'(s) = \gamma''(s) = k(s)\cdot\nu(s).$$

According to 3.1, $\nu(s) \perp \tau(s)$. Therefore the vector product

$$\beta(s) = \tau(s) \times \nu(s)$$

is a unit vector which makes the triple $\tau(s), \nu(s), \beta(s)$ an oriented orthonormal basis in $\mathbb{R}^3$; in particular, we have that

❶
$$\begin{aligned} &\langle\tau,\tau\rangle = 1, \quad \langle\nu,\nu\rangle = 1, \quad \langle\beta,\beta\rangle = 1, \\ &\langle\tau,\nu\rangle = 0, \quad \langle\nu,\beta\rangle = 0, \quad \langle\beta,\tau\rangle = 0. \end{aligned}$$

The orthonormal basis $\tau(s), \nu(s), \beta(s)$ is called *Frenet frame* at s; the vectors in the frame are called *tangent*, *normal* and *binormal* correspondingly. Note that the frame $\tau(s), \nu(s), \beta(s)$ is defined only if $k(s) \neq 0$.

The plane Π_s thru $\gamma(s)$ spanned by vectors $\tau(s)$ and $\nu(s)$ is called *osculating plane* at s; equivalently it can be defined as a plane thru

$\gamma(s)$ that is perpendicular to the binormal vector $\beta(s)$. This a unique plane that has *second order of contact* with γ at s; that is, $\rho(\ell) = o(\ell^2)$, where $\rho(\ell)$ denotes the distance from $\gamma(s+\ell)$ to Π_s.

Torsion

Let γ be a smooth unit-speed space curve and $\tau(s), \nu(s), \beta(s)$ is its Frenet frame. The value

$$\varkappa(s) = \langle \nu'(s), \beta(s) \rangle$$

is called *torsion* of γ at s.

Note that the torsion $\varkappa(s)$ is defined at each s with nonzero curvature. Indeed, if $k(s) \neq 0$ then Frenet frame $\tau(s), \nu(s), \beta(s)$ is defined at s. Moreover since the function $s \mapsto k(s)$ is continuous, it must be positive in an open interval containing s; therefore Frenet frame is also defined in this interval. Clearly $\tau(s)$, $\nu(s)$ and $\beta(s)$ depend smoothly on s in their domains of definition. Therefore $\nu'(s)$ is defined and so is the torsion $\varkappa(s)$.

The torsion measures how fast the osculating plane rotated when one travels along γ.

4.1. Exercise. *Given real numbers a and b, calculate curvature and torsion of the helix*

$$\gamma_{a,b}(t) = (a \cdot \cos t, a \cdot \sin t, b \cdot t).$$

Conclude that for any $k > 0$ and $\varkappa$ there is a helix with constant curvature k and torsion $\varkappa$.

Frenet formulas

Assume the Frenet frame $\tau(s), \nu(s), \beta(s)$ of curve γ is defined at s. Recall that

$$\tau'(s) = k(s) \cdot \nu(s). \tag{❷}$$

Let us write the remaining derivatives $\nu'(s)$ and $\beta'(s)$ in the frame $\tau(s), \nu(s), \beta(s)$.

First let us show that

$$\nu'(s) = -k(s) \cdot \tau(s) + \varkappa(s) \cdot \beta(s). \tag{❸}$$

Since the frame $\tau(s), \nu(s), \beta(s)$ is orthonormal it is equivalent to the following three identities:

$$\langle \nu', \tau\rangle = -k, \qquad \langle \nu', \nu\rangle = 0, \qquad \langle \nu', \beta\rangle = \kappa,$$

The last identity follows from the definition of torsion. Differentiating $\langle \nu, \nu\rangle = 1$ in ❶, we get that

$$2\cdot\langle \nu', \nu\rangle = 0;$$

whence the second identity follows. Differentiating the identity $\langle \tau, \nu\rangle = {}$ $= 0$ in ❶; we get that

$$\langle \tau', \nu\rangle + \langle \tau, \nu'\rangle = 0.$$

Applying ❷, we get that

$$\begin{aligned}\langle \nu', \tau\rangle &= -\langle \tau', \nu\rangle = \\ &= -k\cdot\langle \nu, \nu\rangle = \\ &= -k.\end{aligned}$$

It proves the first equality $\langle \nu', \tau\rangle = -k$ and whence ❸ follows.

Taking derivatives of the third identity in ❶, we get that $\beta' \perp \beta$. Further taking derivatives of the other identities with β in ❶, we get that

$$\begin{aligned}\langle \beta', \tau\rangle &= -\langle \beta, \tau'\rangle = -k\cdot\langle \beta, \nu\rangle = 0\\ \langle \beta', \nu\rangle &= -\langle \beta, \nu'\rangle = \kappa\end{aligned}$$

Since the frame $\tau(s), \nu(s), \beta(s)$ is orthonormal, it follows that

❹ $$\beta'(s) = -\kappa(s)\cdot\nu(s).$$

The equations ❷, ❸ and ❹ are called Frenet formulas. All three can be written as one matrix identity:

$$\begin{pmatrix}\tau'\\ \nu'\\ \beta'\end{pmatrix} = \begin{pmatrix}0 & k & 0\\ -k & 0 & \kappa\\ 0 & -\kappa & 0\end{pmatrix}\cdot\begin{pmatrix}\tau\\ \nu\\ \beta\end{pmatrix}.$$

4.2. Exercise. *Deduce the formula ❹ from ❷ and ❸ by differentiating the identity $\beta = \tau\times\nu$.*

4.3. Exercise. *Let γ be a regular space curve with nonvanishing curvature. Show that γ lies in a plane if and only if its torsion vanishes.*

Hint: Show and use that the binormal vector is constant.

4.4. Exercise. *Let γ be a smooth regular space curve and τ,ν,β its Frenet frame. Show that*

$$\beta=\frac{\gamma'\times\gamma''}{|\gamma'\times\gamma''|}.$$

Use this formula to show that its torsion is

$$\kappa=\frac{\langle\gamma'\times\gamma'',\gamma'''\rangle}{|\gamma'\times\gamma''|^2}.$$

Curves of constant slope

We say that a smooth regular space curve γ has *constant slope* if its velocity vector makes a constant angle with a fixed direction. The following theorem was proved by Michel Ange Lancret [9] more than two centuries ago.

4.5. Theorem. *Let γ be a smooth regular curve; denote by k and κ its curvature and torsion. Suppose $k(s)>0$ at any s. Then γ has constant slope if and only if the ratio $\frac{\kappa}{k}$ is constant.*

Note that the assumption in the theorem implicitely implies that $k\neq 0$; otherwise $\frac{\kappa}{k}$ is undefined.

The proof is an application of Frenet formulas. The following exercise will guide you thru the proof of the theorem.

4.6. Exercise. *Suppose that γ is a smooth regular space curve with nonvanishing curvature, τ,ν,β is its Frenet frame and k, κ are its curvature and torsion.*

(a) Assume that $\langle w,\tau\rangle$ is constant for a fixed nonzero vector w. Show that

$$\langle w,\nu\rangle=0.$$

Use it to show that

$$\langle w,-k\cdot\tau+\kappa\cdot\beta\rangle=0.$$

Use these two identities to show that $\frac{\kappa}{k}$ is constant; it proves the "only if" part of the theorem.

(b) Assume that $\frac{\kappa}{k}$ is constant, show that the vector $w=\frac{\kappa}{k}\cdot\tau+\beta$ is constant. Conclude that γ has constant slope; it proves the "if" part of the theorem.

Assume γ is a smooth unit speed curve and s_0 is a fixed real number. Then the curve

$$\alpha(s) = \gamma(s) + (s_0 - s)\cdot\gamma'(s)$$

is called *evolvent* of γ. Note that if $\ell(s)$ denotes the tangent line of γ at s, then $\alpha(s) \in \ell(s)$ and $\alpha'(s) \perp \ell$ for any s.

4.7. Exercise. *Show that evolvent of a constant slope curve lies in a plane.*

Hint: Show that $\langle w, \alpha\rangle$ is constant if γ makes constant angle with a fixed vector w and α is the evolvent of γ.

Spherical curves

4.8. Theorem. *A smooth regular space curve γ lies in a unit sphere if and only if the following identity*

$$\left|\tfrac{k'}{\kappa}\right| = k\cdot\sqrt{k^2-1}.$$

holds for its curvature k and torsion κ.

Note that the identity implicitely implies that the torsion κ of the curve is nonzero; otherwise the left hand side would be undefined while right hand side is defined. The proof is another application of Frenet formulas; we present it in a form of guided exercise:

4.9. Exercise. *Suppose γ is a smooth unit-speed space curve. Denote by τ, ν, β is its Frenet frame and k, κ its curvature and torsion.*

Assume that γ is spherical; that is, $|\gamma(s)| = 1$ for any s. Show that

(a) *$\langle\tau, \gamma\rangle = 0$; conclude that $\langle\nu, \gamma\rangle^2 + \langle\beta, \gamma\rangle^2 = 1$.*
(b) *$\langle\nu, \gamma\rangle = -\frac{1}{k}$;*
(c) *$\langle\beta, \gamma\rangle' = \frac{\kappa}{k}$; conclude that if γ is closed, then $\kappa(s) = 0$ for some s.*
(d) *Use (a)–(c) to show that*

$$\left|\tfrac{k'}{\kappa}\right| = k\cdot\sqrt{k^2-1}.$$

It proves the "only if" part of the theorem.

Now assume that γ is a space curve that satisfies the idenity in (d).

(e) *Show that $p = \gamma + \frac{1}{k}\cdot\nu + \frac{k'}{k^2\cdot\kappa}\cdot\beta$ is constant; conclude that γ lies in a unit sphere the centered at p.*

It proves the "if" part of the theorem.

For a unit speed curve γ with nonzero curvature and torsion at s, the sphere Σ_s with the center

$$p(s) = \gamma(s) + \tfrac{1}{k(s)}\cdot\nu(s) + \tfrac{k'(s)}{k^2(s)\cdot\varkappa(s)}\cdot\beta(s)$$

that pass thru $\gamma(s)$ is called *osculating sphere* of γ at s. This a unique sphere that has *third order of contact* with γ at s; that is, $\rho(\ell) = o(\ell^3)$, where $\rho(\ell)$ denotes the distance from $\gamma(s+\ell)$ to Σ_s.

Fundamental theorem of curves

4.10. Theorem. *Let $k(s)$ and $\varkappa(s)$ be two smooth real valued functions defined on a real interval $\mathbb{I}$. Suppose $k(s) > 0$ for any s. Then there is a smooth unit-speed curve $\gamma\colon \mathbb{I}\to\mathbb{R}^3$ with curvature $k(s)$ and torsion $\varkappa(s)$ for every s. Moreover γ is uniquely defined up to a rigid motion of the space.*

Proof. Fix a parameter value s_0, a point $\gamma(s_0)$ and an oriented orthonormal frame $\tau(s_0)$, $\nu(s_0)$, $\beta(s_0)$. Consider the following initial value problem:

$$\begin{cases}\gamma' = \tau,\\ \tau' = k\cdot\nu,\\ \nu' = -k\cdot\tau + \varkappa\cdot\beta,\\ \beta' = -\varkappa\cdot\nu.\end{cases}$$

It has four vector equations, so it can be rewritten as a system of 12 scalar equations. By A.12, it has a solution which is defined in a maximal subinterval $\mathbb{J}$ containing s_0.

Note that

$$\begin{aligned}
\langle\tau,\tau\rangle' &= 2\cdot\langle\tau,\tau'\rangle = 2\cdot k\cdot\langle\tau,\nu\rangle = 0,\\
\langle\nu,\nu\rangle' &= 2\cdot\langle\nu,\nu'\rangle = -2\cdot k\cdot\langle\nu,\tau\rangle + 2\cdot\varkappa\cdot\langle\nu,\beta\rangle = 0,\\
\langle\beta,\beta\rangle' &= 2\cdot\langle\beta,\beta'\rangle = -2\cdot\varkappa\langle\beta,\nu\rangle = 0,\\
\langle\tau,\nu\rangle' &= \langle\tau',\nu\rangle + \langle\tau,\nu'\rangle = k\cdot\langle\nu,\nu\rangle - k\cdot\langle\tau,\tau\rangle + \varkappa\cdot\langle\tau,\beta\rangle = 0,\\
\langle\nu,\beta\rangle' &= \langle\nu',\beta\rangle + \langle\nu,\beta'\rangle = 0,\\
\langle\beta,\tau\rangle' &= \langle\beta',\tau\rangle + \langle\beta,\tau'\rangle = -\varkappa\cdot\langle\nu,\tau\rangle + k\cdot\langle\beta,\nu\rangle = 0.
\end{aligned}$$

That is, the values $\langle\tau,\tau\rangle$, $\langle\nu,\nu\rangle$, $\langle\beta,\beta\rangle$, $\langle\tau,\nu\rangle$, $\langle\tau,\nu\rangle$, $\langle\beta,\tau\rangle$ are constant functions of s. Since we choose $\tau(s_0)$, $\nu(s_0)$, $\beta(s_0)$ to be an oriented orthonormal frame, we have that the $\tau(s)$, $\nu(s)$, $\beta(s)$ is oriented orthonormal for any s.

In particular $|\gamma'(s)| = 1$ for any s.

Assume $\mathbb{J} \neq \mathbb{I}$. Then an end of $\mathbb{J}$, say a, lies in the interior of $\mathbb{I}$. By Theorem A.12, at least one of the values $\gamma(s)$, $\tau(s)$, $\nu(s)$, $\beta(s)$ escapes to infinity as $s \to a$. But this is impossible since the vectors $\tau(s)$, $\nu(s)$, $\beta(s)$ remain unit and $|\gamma'(s)| = |\tau(s)| = 1$ — a contradiction. Whence $\mathbb{J} = \mathbb{I}$.

Assume there are two curves γ_1 and γ_2 with the given curvature and torsion functions. Applying a motion of the space we can assume that the $\gamma_1(s_0) = \gamma_2(s_0)$ and the Frenet frames of the curves coincide at s_0. Then $\gamma_1 = \gamma_2$ by uniqueness of solution of the system (A.12). That is, the curve is unique up to a rigid motion of the space. $\square$

4.11. Exercise. *Assume a curve $\gamma\colon \mathbb{R} \to \mathbb{R}^3$ has constant curvature and torsion. Show that γ is a helix, possibly degenerate to a circle; that is, in a suitable coordinate system we have*

$$\gamma(t) = (a\cdot \cos t, a\cdot \sin t, b\cdot t)$$

for some constants a and b.

Hint: Use the second statement in 4.1.

4.12. Advanced exercise. *Let γ be a smooth regular space curve such that the distance $|\gamma(t) - \gamma(t + \ell)|$ depends only on ℓ. Show that γ is a helix, possibly degenerate to a line or a circle.*

Hint: Note that the function

$$\rho(\ell) = |\gamma(t + \ell) - \gamma(t)|^2 = \langle \gamma(t + \ell) - \gamma(t), \gamma(t + \ell) - \gamma(t)\rangle$$

is smooth and does not depend on t. Express speed, curvature and torsion of γ in terms of derivatives $\rho^{(n)}(0)$ and apply 4.11.

Chapter 5

Plane curves

Signed curvature

Suppose γ is a smooth unit-speed plane curve, so $\tau(s) = \gamma'(s)$ is its unit tangent vector.

Let us rotate $\tau(s)$ by angle $\frac{\pi}{2}$ counterclockwise; denote the obtained vector by $\nu(s)$. The pair $\tau(s), \nu(s)$ is an oriented orthonormal frame in the plane which is analogous to the Frenet frame for space curves; we will keep the name *Frenet frame* for it.

Recall that $\gamma''(s) \perp \gamma'(s)$ (see 3.1). Therefore

❶ $$\tau'(s) = k(s)\cdot\nu(s).$$

for some real number $k(s)$; the value $k(s)$ is called *signed curvature* of γ at s. Note that up to sign it equals to the curvature of γ at s as it defined on page 19; the sign tells which direction γ turns — if it turns left, then it is positive. If we want to emphasise that we work with *nonsigned* curvature of the curve, we call it *absolute curvature* — it is absolute value of signed curvature.

Note that if we reverse the parametrization of γ or change the orientation of the plane, then the signed curvature changes its sign.

Since $\tau(s), \nu(s)$ is an orthonormal frame, we have that

$$\langle\tau,\tau\rangle = 1, \qquad \langle\nu,\nu\rangle = 1, \qquad \langle\tau,\nu\rangle = 0,$$

Differentiating these idenitites we get that

$$\langle\tau',\tau\rangle = 0, \qquad \langle\nu',\nu\rangle = 0, \qquad \langle\tau',\nu\rangle + \langle\tau,\nu'\rangle = 0,$$

By ❶, $\langle\tau',\nu\rangle = k$ and therefore $\langle\tau,\nu'\rangle = -k$. Whence we get

❷ $$\nu'(s) = -k(s)\cdot\tau(s).$$

The equations ❶ and ❷ are Frenet formulas for plane curves. They could be also written in a matrix form:

$$\begin{pmatrix}\tau'\\ \nu'\end{pmatrix} = \begin{pmatrix}0 & k\\ -k & 0\end{pmatrix}\cdot\begin{pmatrix}\tau\\ \nu\end{pmatrix}.$$

5.1. Exercise. *Let $\gamma_0\colon [a,b]\to\mathbb{R}^2$ be a smooth regular curve and τ its tangent indicatrix. Consider another curve $\gamma_1\colon [a,b]\to\mathbb{R}^2$ defined by $\gamma_1(t)=\gamma_0(t)+\tau(t)$. Show that*

$$\operatorname{length}\gamma_0 \leqslant \operatorname{length}\gamma_1.$$

The curves γ_0 and γ_1 in the exercise above describe tracks of idealized bicycle with the distance 1 from rear to front wheel. Thus by the exercise, the front wheel have to have the longer track. For more on geometry of bicycle tracks see [10] and the references there in.

5.2. Theorem. *Let $k(s)$ be a smooth real valued function defined on a real interval $\mathbb{I}$. Then there is a smooth unit-speed curve $\gamma\colon\mathbb{I}\to\mathbb{R}^2$ with signed curvature $k(s)$ at every s. Moreover γ is uniquely defined up to a rigid motion of the plane.*

This is the fundamental theorem of plane curves; it is direct analog of 4.10 and it can be proved along the same lines. We give a slightly simpler proof.

Proof. Fix $s_0\in\mathbb{I}$. Consider the function

$$\theta(s)=\int\limits_{s_0}^{s} k(t)\cdot dt.$$

Note that by the fundamental theorem of calculus, we have $\theta'(s) = k(s)$ for any s.

Set

$$\tau(s)=(\cos[\theta(s)],\sin[\theta(s)])$$

and let $\nu(s)$ be its counterclockwise rotation by angle $\frac{\pi}{2}$; so

$$\nu(s)=(-\sin[\theta(s)],\cos[\theta(s)]).$$

Consider the curve

$$\gamma(s)=\int\limits_{s_0}^{s}\tau(s)\cdot ds.$$

Since $|\gamma'|=|\tau|=1$, the curve γ is unit-speed and τ,ν is its Frenet frame.

Note that

$$\begin{aligned}\gamma''(s) = \tau'(s) = \\ &= (\cos[\theta(s)]', \sin[\theta(s)]') = \\ &= \theta'(s)\cdot(-\sin[\theta(s)], \cos[\theta(s)]) = \\ &= k(s)\cdot\nu(s).\end{aligned}$$

That is $k(s)$ is the signed curvature of γ at s.

We proved the existence; it remains to prove uniqueness. Assume γ_1 and γ_2 are two curves that satisfy the assumptions of the theorem. Applying a rigid motion, we can assume that $\gamma_1(s_0) = \gamma_2(s_0) = 0$ and the Frenet frame of both curves at s_0 is formed by the coordinate frame $(1, 0), (0, 1)$. Let us denote by τ_1, ν_1 and τ_2, ν_2 the Frenet frames of γ_1 and γ_2 correspondingly. The triples γ_i, τ_i, ν_i satisfy the same system system of ordinary differential equations

$$\begin{cases}\gamma_i' = \tau_i,\\ \tau_i' = k\cdot\nu_i,\\ \nu_i' = -k\cdot\tau_i.\end{cases}$$

Motreover, they have the same the initial values at s_0. Therefore $\gamma_1 = \gamma_2$. $\square$

Note that from the proof of theorem we obtain the following corollary:

5.3. Corollary. *Suppose $\gamma\colon [a, b] \to \mathbb{R}^2$ is a smooth unit-speed curve. Then there is a smooth function $\theta\colon [a, b] \to \mathbb{R}$ such that*

$$\gamma'(s) = (\cos[\theta(s)], \sin[\theta(s)]) \quad \textit{and} \quad \theta'(s) = k(s)$$

for any s, where $k(s)$ denotes the signed curvature of γ.

Total signed curvature

Let γ be a smooth unit-speed plane curve. The integral of its signed curvature is called *total signed curvature* and it denoted by $\Psi(\gamma)$; so if θ and γ is as in 5.3, then

❸ $$\Psi(\gamma) = \int_a^b k(s)\cdot ds = \theta(b) - \theta(a).$$

Since $\left|\int k(s)\cdot ds\right| \leqslant \int |k(s)|\cdot ds$, we have that

❹ $$|\Psi(\gamma)| \leqslant \Phi(\gamma)$$

for any smooth regular plane curve γ.

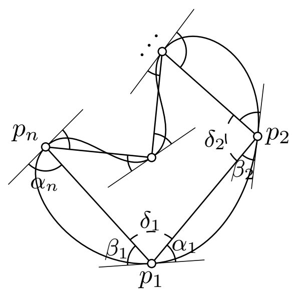
5.4. Proposition. *The total signed curvature of any closed simple smooth plane curve γ is $\pm 2{\cdot}\pi$; it is $+2{\cdot}\pi$ if the region bounded by γ lies on the left from it and $-2{\cdot}\pi$ otherwise.*

This proposition is a differential-geometric analog of the theorem about sum of the internal angles of a polygon (A.19) which we use in the proof. A more conceptual proof was given by Heinz Hopf [11], [12, p. 42].

Proof. Without loss of generality we may assume that γ is oriented in such a way that the region bounded by γ lies on the left from it. We can also assume that it parametrized by arc length.

Consider a closed polygonal line $p_1 \dots p_n$ inscribed in γ. We can assume that the arcs between the vertexes are sufficiently small; in this case the polygonal line is simple and each arc γ_i from p_i to p_{i+1} have small total absolute curvature, say $\Phi(\gamma_i) < \pi$ for each i.

As usual we use indexes modulo n, in particular $p_{n+1} = p_1$. Assume $p_i = \gamma(t_i)$. Set



$$w_i = p_{i+1} - p_i, \qquad v_i = \gamma'(t_i),$$
$$\alpha_i = \measuredangle(v_i, w_i), \qquad \beta_i = \measuredangle(w_{i-1}, v_i),$$

where $\alpha_i, \beta_i \in (-\pi, \pi]$ are oriented angles — α_i is positive if w_i points to the left from v_i.

By ❸, the value

$$\text{❺} \qquad \Psi(\gamma_i) - \alpha_i - \beta_{i+1}$$

is a multiple of $2{\cdot}\pi$. Since $\Phi(\gamma_i) < \pi$, by chord lemma (3.11), we also have that $|\alpha_i| + |\beta_i| < \pi$. By ❹, we have that $|\Psi(\gamma_i)| \leqslant \Phi(\gamma_i)$; therefore the value in ❺ vanishes, or equivalently

$$\Psi(\gamma_i) = \alpha_i + \beta_{i+1}$$

for each i.

Note that $\delta_i = \pi - \alpha_i - \beta_i$ is the internal angle of $p_1 \dots p_n$ at p_i; $\delta_i \in (0, 2{\cdot}\pi)$ for each i. Recall that the sum of internal angles of an n-gon is $(n-2){\cdot}\pi$ (see A.19); that is,

$$\delta_1 + \dots + \delta_n = (n-2){\cdot}\pi.$$

Therefore

$$\begin{aligned}\Psi(\gamma) &= \Psi(\gamma_1) + \dots + \Psi(\gamma_n) = \\ &= (\alpha_1 + \beta_2) + \dots + (\alpha_n + \beta_1) = \\ &= (\beta_1 + \alpha_1) + \dots + (\beta_n + \alpha_n) = \\ &= (\pi - \delta_1) + \dots + (\pi - \delta_n) = \\ &= n{\cdot}\pi - (n-2){\cdot}\pi = \\ &= 2{\cdot}\pi.\end{aligned}$$

□

5.5. Exercise. *Draw a smooth regular closed plane curve with zero total signed curvature.*

5.6. Exercise. *Let $\gamma\colon [a,b] \to \mathbb{R}$ be a smooth regular plane curve with Frenet frame τ, ν. Given a real parameter d, consider the curve $\gamma_d(t) = \gamma(t) + d{\cdot}\nu(t)$; it is called a* parallel curve of γ on the signed distance d.

(a) Show that γ_d is a regular curve if $d{\cdot}k(t) \neq 1$ for any t, where $k(t)$ denotes the signed curvature of γ.

(b) Set $\ell(d) = \operatorname{length}\gamma_d$. Show that

$$\ell(d) = \ell(0) - d{\cdot}\Psi(\gamma)$$

for all d sufficiently close to 0. Show that in general, this formula does not hold for all d.

Osculating circline

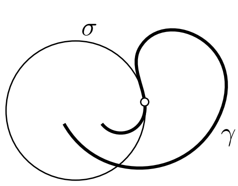
As a direct corollary of Theorem 5.2, we get the following:

5.7. Proposition. *Given a point p, a unit vector τ and a real number k, there is a unique smooth unit-speed curve $\sigma\colon \mathbb{R} \to \mathbb{R}^2$ that starts at p in the direction of τ and has constant signed curvature k.*

Moreover, if $k = 0$, then $\sigma(s) = p + s{\cdot}\tau$ which runs along the line; if $k \neq 0$, then σ runs around the circle of radius $\frac{1}{|k|}$ and center $p + \frac{1}{k}{\cdot}\nu$, where τ, ν is an oriented orthonoral frame.

Further we will use the term *circline* for *a circle or a line.*

5.8. Definition. *Let γ be a smooth unit-speed plane curve; denote by $k(s)$ the signed curvature of γ at s.*



The unit-speed curve σ of constant signed curvature $k(s)$ that starts at $\gamma(s)$ in the direction $\gamma'(s)$ is called the osculating circline *of γ at s.*

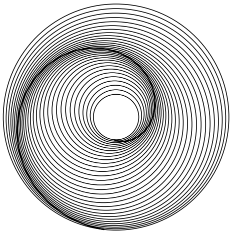
The center and radius of the osculating circle at a given point are called *center of curvature* and *radius of curvature* of the curve at that point.

The osculating circle σ_s can be also defined as the (necessarily unique) circline that has *second order of contact* with γ at s; that is, $\rho(\ell) = o(\ell^2)$, where $\rho(\ell)$ denotes the distance from $\gamma(s+\ell)$ to σ_s.

Spiral lemma

The following lemma was proved by Peter Tait [13] and later rediscovered by Adolf Kneser [14].

5.9. Lemma. *Assume that γ is a smooth regular plane curve with strictly decreasing positive signed curvature. Then the osculating circles of γ are nested; that is, if σ_s denoted the osculating circle of γ at s, then σ_{s_0} lies in the open disc bounded by σ_{s_1} for any $s_0 < s_1$.*



It turns out that osculating circles of the curve γ give a peculiar foliation of an annulus by circles; it has the following property: if a smooth function is constant on each osculating circle it must be constant in the annulus [see 15, Lecture 10]. Also note that the curve γ is tangent to a circle of the foliation at each of its points. However, it does not run along a circle.

Proof. Let $\tau(s), \nu(s)$ be the Frenet frame, $z(s)$ the curvature center and $r(s)$ the radius of curvature of γ at s. By 5.7,

$$z(s) = \gamma(s) + r(s)\cdot\nu(s).$$

Since $k > 0$, we have that $r(s)\cdot k(s) = 1$. Therefore applying Frenet formula ❷, we get that

$$\begin{aligned} z'(s) &= \gamma'(s) + r'(s)\cdot\nu(s) + r(s)\cdot\nu'(s) = \\ &= \tau(s) + r'(s)\cdot\nu(s) - r(s)\cdot k(s)\cdot\tau(s) = \\ &= r'(s)\cdot\nu(s). \end{aligned}$$

Since $k(s)$ is decreasing, $r(s)$ is increasing; therefore $r' \geqslant 0$. It follows that $|z'(s)| = r'(s)$ and $z'(s)$ points in the direction of $\nu(s)$.

Since $\nu'(s) = -k(s)\cdot\tau(s)$, the direction of $z'(s)$ can not have constant direction on a nontrivial interval; that is, the curve $s \mapsto z(s)$ contains no line segments. It follows that

$$\begin{aligned}|z(s_1) - z(s_0)| < \operatorname{length}(z|_{[s_0,s_1]}) =& \\ &= \int\limits_{s_0}^{s_1} |z'(s)|\cdot ds = \\ &= \int\limits_{s_0}^{s_1} r'(s)\cdot ds = \\ &= r(s_1) - r(s_0).\end{aligned}$$

In other words, the distance between the centers of σ_{s_1} and σ_{s_0} is strictly less than the difference between their radiuses. Therefore the osculating circle at s_0 lies inside the osculating circle at s_1 without touching it. □

The curve $s \mapsto z(s)$ is called *evolute* of γ; it traces the centers of curvature of the curve. The evolute of γ can be written as

$$z(t) = \gamma(t) + \tfrac{1}{k(t)}\cdot\nu(t)$$

and in the proof we showed that $(\tfrac{1}{k})'\cdot\nu$ is its the velocity vector.

5.10. Exercise. *Show that the stretched astroid*

$$\alpha(t) = (\tfrac{a^2-b^2}{a}\cdot\cos^3 t, \tfrac{b^2-a^2}{b}\cdot\sin^3 t)$$

is an evolute of the ellipse $\gamma(t) = (a\cdot\cos t, b\cdot\sin t)$.

The following theorem states formally that *if you drive on the plane and turn the steering wheel to the right all the time, then you will not be able to come back to the same place.*

5.11. Theorem. *Assume γ is a smooth regular plane curve with positive and strictly monotonic signed curvature. Then γ is simple.*

Proof of 5.11. Note that $\gamma(s)$ lies on the osculating circle σ_s of γ at s. If $s_1 \neq s_0$, then by lemma σ_{s_0} does not intersect σ_{s_1}. Therefore $\gamma(s_1) \neq \gamma(s_0)$, hence the result. □

The same statement holds without assuming positivity of curvature; the proof requires only minor modifications.

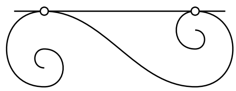
5.12. Exercise. *Show that a 3-dimensional analog of the theorem does not hold. That is, there are self-intersecting smooth regular space curves with strictly monotonic curvature.*

5.13. Exercise. *Assume that γ is a smooth regular plane curve with positive strictly monotonic signed curvature.*

(a) *Show that no line can be tangent to γ at two distinct points.*

(b) *Show that no circle can be tangent to γ at three distinct points.*

Note that part (a) does not hold if we allow the curvature to be negative; an example is shown on the diagram.



Supporting circlines

Suppose γ is a smooth regular plane curve. Recall that a circline σ is tangent to γ at t_0 if $\gamma(t_0) = \sigma(t_1)$ for some t_1 and they share the tangent at these time parameters; that is, the tangent lines of γ at t_0 coincides with the tangent line σ at t_1.

We can (and often will) assume that tangent circline is *cooriented* with the curve; that is, the tangent vectors $\gamma'(t_0)$ and $\sigma'(t_1)$ point in the same direction. If not we can reverse the parametrization of σ. If both curves are given with arc length parametrization, then coorientation means that $\gamma'(t_0) = \sigma'(t_1)$.

If γ is simple we can say that σ is tangent to γ at the point $p = \gamma(t_0)$ without ambiguity.

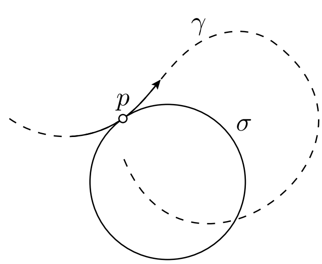
A circline σ supports γ at t_0 if $\gamma(t_0) \in \sigma$ and γ lies on one side of σ. (We assume that t_0 is not an end point of the interval of parameters.) If $p = \sigma(t_0)$ for a single value t_0, then we can also say *σ supports γ at p* without ambiguity.

Note that if σ supports γ at t_0, then σ is tangent to γ at t_0. Indeed, if it is not the case, then $\gamma'(t_0)$ would point inside or outside of σ; therefore γ would cross σ from one side to another. Therefore we can assume that σ is cooriented with γ at t_0. In this case we say that σ supports γ from the left (right) if γ lies on the right (correspondingly left) side from σ.

Note that a circle supports itself on the right and left at the same time at any point.

We say that a circle σ *locally supports* a curve γ at s if it supports its arc $\gamma|_{(s-\varepsilon,s+\varepsilon)}$ for some $\varepsilon > 0$. The same definitions for local support on the left and right are applied.

The circle σ on the diagram locally supports curve γ on the right at p, but does not support it globally — since γ crosses σ at a latter time.



We say that a smooth regular plane curve γ has a *vertex* at s if the signed curvature function is critical at s; that is, if $k'_\gamma(s) = 0$. If γ is simple we could say that the point $p = \gamma(s)$ is a vertex of γ without ambiguity.

5.14. Exercise. *Assume that osculating circle σ_s of a smooth regular plane curve γ supports γ at s. Show that γ has a vertex at s.*

Hint: Apply the spiral lemma (5.9).

Supporting test

The following proposition resembles the second derivative test.

5.15. Proposition. *Assume σ is a circline that locally supports γ at t_0 from the right (correspondingly left). Suppose σ is cooriented to γ at t_0. Then*

$$k(t_0)_\gamma \geqslant k_\sigma \quad (\textit{correspondingly} \quad k(t_0)_\gamma \leqslant k_\sigma).$$

where k_σ is the signed curvature of σ and $k(t_0)_\gamma$ is the signed curvature of γ at t_0.

A partial converse also holds. Namely, suppose a unit-speed circline σ with signed curvature k_σ starts at $\gamma(t_0)$ in the direction $\gamma'(t_0)$. Then σ locally supports γ at t_0 from the right (correspondingly left) if

$$k(t_0)_\gamma > k_\sigma \quad (\textit{correspondingly} \quad k(t_0)_\gamma < k_\sigma).$$

Proof. We prove only the case $k_\sigma > 0$. The 2 remaining cases $k_\sigma = 0$ and $k_\sigma < 0$ can be done essentially the same way.

Since $k_\sigma \neq 0$, the curve σ is a circle. According to Proposition 5.7, σ has radius $r_\sigma = \frac{1}{k_\sigma}$ and it is centered at

$$z = \gamma(t_0) + r{\cdot}\nu(t_0).$$

Consider the function

$$f(t) = |z - \gamma(t)|^2 - \tfrac{1}{k_\sigma^2}.$$

Note that $f(t) \leqslant 0$ (correspondingly $f(t) \geqslant 0$) if and only if $\gamma(t)$ lies on the closed left (correspondingly right) side from σ. It follows that

- if σ locally supports γ at t_0 from the right, then

$$f'(t_0) = 0 \quad \text{and} \quad f''(t_0) \leqslant 0;$$

◇ if σ locally supports γ at t_0 from the left, then

$$f'(t_0) = 0 \quad \text{and} \quad f''(t_0) \geqslant 0;$$

◇ if

$$f'(t_0) = 0 \quad \text{and} \quad f''(t_0) < 0,$$

then σ locally supports γ at t_0 from the right;

◇ if

$$f'(t_0) = 0 \quad \text{and} \quad f''(t_0) > 0,$$

then σ locally supports γ at t_0 from the left;

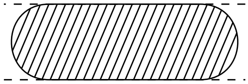
Direct calculations show that

$$\begin{aligned}
f(t_0) &= 0;\\
f'(t_0) &= \langle z-\gamma(t), z-\gamma(t)\rangle'|_{t=t_0} =\\
&= -2\cdot\langle\gamma'(t_0), z-\gamma(t_0)\rangle =\\
&= -2\cdot r\cdot\langle\gamma'(t_0), \nu(t_0)\rangle =\\
&= 0;\\
f''(t_0) &= \langle z-\gamma(t), z-\gamma(t)\rangle''|_{t=t_0} =\\
&= 2\cdot\left(\langle\gamma'(t_0),\gamma'(t_0)\rangle - \langle\gamma''(t_0), z-\gamma(t_0)\rangle\right) =\\
&= 2\cdot\left(\langle\tau(t_0),\tau(t_0)\rangle - r\cdot k(t_0)_\gamma\cdot\langle\nu(t_0),\nu(t_0)\rangle\right) =\\
&= 2\cdot\left(1-\frac{k(t_0)_\gamma}{k_\sigma}\right).
\end{aligned}$$

Hence the result. □

5.16. Exercise. *Assume a closed smooth regular plane curve γ runs between parallel lines on distance 2 from each other. Show that there is a point on γ with absolute curvature at least 1.*

Hint: Note that the curve lies in a figure F as on the diagram. More precisely, F is formed by a rectangle with pair of bases on the lines and two half discs attached to the sides of length 2. Look at the right most position of F that still contains the curve.



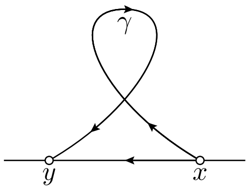
5.17. Exercise. *Assume a closed smooth regular plane curve γ runs inside of a triangle $\triangle$ with inradius 1; that is, the inscribed circle of $\triangle$ has radius 1. Show that there is a point on γ with absolute curvature at least* 1.

Convex curves

Recall that a plane curve is convex if it bounds a convex region.

5.18. Proposition. *Suppose that a closed simple cure γ bounds a figure F. Then F is convex if and only if the signed curvature of γ does not change the sign.*

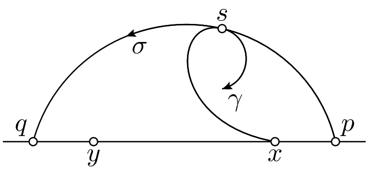
5.19. Lens lemma. *Let γ be a smooth regular simple curve that runs from x to y and distinct from the line segment from x to y. Assume that γ runs on the closed right side (correspondingly left side) of the oriented line xy and only its end points x and y lie on the line. Then γ has a point with positive (correspondingly negative) signed curvature.*



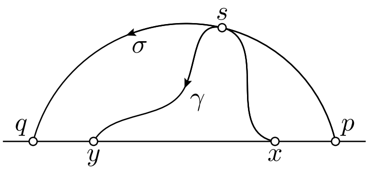
Note that the lemma fails for curves with self-intersections; the curve γ on the diagram always turns right, so it has negative curvature everywhere, but it lies on the right side of the line xy.

Proof. Choose points p and q on ℓ so that the points p, x, y, q appear in the same order on ℓ. We can assume that p and q lie sufficiently far from x and y, so the half-disc with diameter pq contains γ.

Consider the smallest disc segment with chord $[pq]$ that contains γ. Note that its arc σ supports γ at a point $s = \gamma(t_0)$.



Note that the $\gamma'(t_0)$ is tangent to σ at s. Let us parameterise σ from p to q. Then γ and σ are cooriented as s. If not, then the arc of γ from s to y would be trapped in the curvelinear triangle xsp bounded by arcs of σ, γ and the line segment $[px]$. But this is impossible since y does not belong to this triangle.

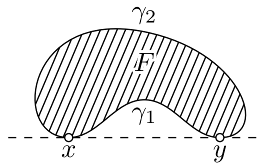


It follows that σ supports γ at t_0 from the right. By 5.15,

$$k(t_0)_\gamma \geqslant k_\sigma,$$

Evidently $k_\sigma > 0$, hence the result. $\square$

Remark. Instead of taking minimal disc segment, one can take a point s on γ that maximize the distance to the line xy. The same argument shows that curvature at s is non-negative, which is slightly weaker than the required positive curvature.



Proof of 5.18. If F is convex, then every tangent line of γ supports γ. If a point moves along γ, the figure F has to stay on one side from its tangent line; that is, we can assume that each tangent line supports γ on one side, say on the right. Applying the supporting test (5.15), we get that $k \geqslant 0$ at each point.

Now assume F is not convex. Then there is a line that supports γ at two points, say x and y that divide γ in two arcs γ_1 and γ_2, both distinct from the line segment xy. Note the one of the arcs is parametrized from x to y and the other from y to x. Applying the lens lemma, we get that the arcs γ_1 and γ_2 contain points with curvatures of opposite signs.

That is, if F is not convex, then curvature of γ changes sign. Equivalently: if curvature of γ does not change sign then F is convex. $\square$

5.20. Exercise. *Suppose γ is a smooth regular simple closed convex plane curve of diameter bigger than 2. Show that γ has a point with absolute curvature less than 1.*

Hint: Note that we can assume that γ bounds a convex figure F, otherwise by 5.18 its curvature changes the sign and therefore it has zero curvature at some point. Choose two points x and y surrounded by γ such that $|x-y|>2$, look at the maximal lens bounded by two arcs with common chord xy that lies in F and apply supporting test (5.15).

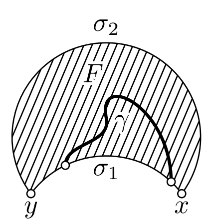
5.21. Exercise. *Suppose γ is a simple smooth regular curve in the plane with positive curvature. Assume γ crosses a line ℓ at the points $p_1, p_2, \dots p_n$ and these points appear on γ in that same order.*

(a) *Show that p_2 can not lie between p_1 and p_3 on ℓ.*

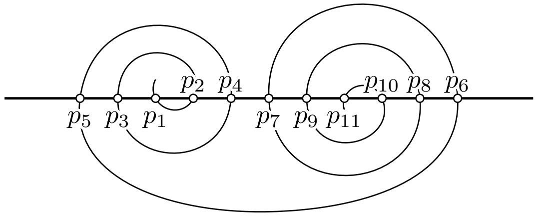
(b) *Show that if p_3 lies between p_1 and p_2 on ℓ then the points appear on ℓ in the following order:*

$$p_1, p_3, \dots, p_4, p_2.$$

(c) *Try to describe all possible orders when p_1 lies between p_2 and p_3 (see the diagram).*



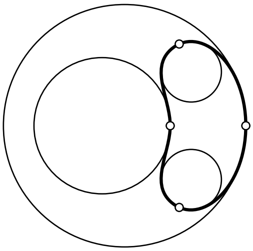
5.22. Exercise. *Let F be a plane figure bounded by two circle arcs σ_1 and σ_2 of signed curvature 1 that run from x to y. Suppose σ_1 is a shorter than σ_2. Assume a simple arc γ runs in F and has the end points on σ_1. Show that the absolute curvature of γ is at least 1 at some parameter value.*



Moon in a puddle

5.23. Theorem. *Assume γ is a simple closed smooth regular plane curve. Then at least two of its osculating circles support γ from the left and at least two from the right.*

The diagram shows for supporting osculating circles, two from inside and two outside the curve for the given curve.

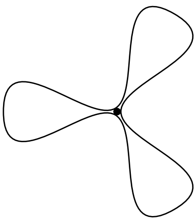


The above theorem is a slight generalization of the following theorem proved by Vladimir Ionin and German Pestov in [16]:

5.24. Theorem. *Assume γ is a simple closed smooth regular plane curve of absolute curvature bounded by 1. Then it surrounds a unit disc.*

This theorem is a direct corollary of 5.23; indeed, since absolute curvature is bounded by 1, every osculating circle has radius at least 1 and by 5.23 two pf these circles are surrounded by γ.

This theorem gives a simple but nontrivial example of the so called *local to global theorems* — based on some local data (in this case the curvature of a curve) we conclude a global property (in this case existence of a large disc surrounded by the curve). For convex curves, this result was known earlier [17, §24].

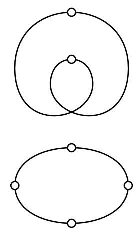


A straightforward approach to the latter theorem would be to start with some disc in the region bounded by the curve and blow it up to maximize its radius. However, as one may see from the diagram it does not always lead to a solution a closed plane curve of absolute curvature bounded by 1 may surround a disc of radius smaller than 1 that can not be enlarged continuously.

Recall that a vertex of a smooth regular curve is defined as a critical point of its signed curvature; in particular, any local minimum (or maximum) of the signed curvature is a vertex.

According to 5.14, if an osculating circle supports the curve at the same point p, then p is a vertex. Therefore 5.23 implies existence of 4 vertexes of γ. That is, we proved the following theorem:

5.25. Four-vertex theorem. *Any smooth regular simple plane curve has at least four vertices.*

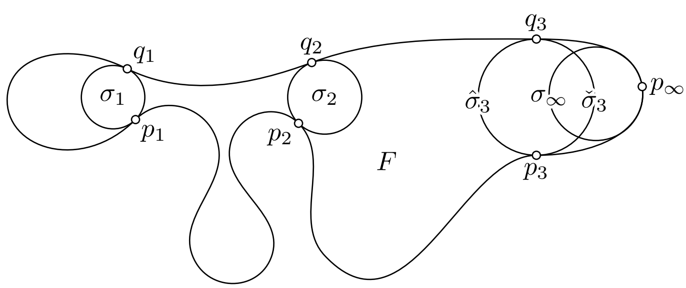


Evidently any closed curve has at least two vertexes — where the minimum and the maximum of the curvature are attained. On the diagram the vertexes are marked; the first curve has one self-intersection and exactly two vertexes; the second curve has exactly four vertexes and no self-intersections.

The four-vertex theorem was first proved by Syamadas Mukhopadhyaya [18] for convex curves. By now it has a large number of different proofs and generalizations. One of my favorite proofs was given by Robert Osserman [19]; the proof of Vladimir Ionin and German Pestov given below is even better.

Proof of 5.23. Denote by F the closed region surrounded by γ; as usual we parametrize γ so that F lies on the left from it.

First let us show that one osculating circle is supporting γ from the left; that is, it lies completely in F — this is the main part of the proof.



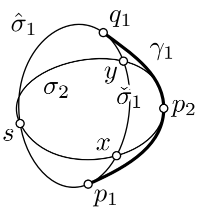
Assume contrary; that is, the osculating circle at each point $p \in \gamma$ does not lie in F. For each point $p \in \gamma$ let us consider the maximal circle σ that lies completely in F and tangent to γ at p; in other words, σ has minimal signed curvature among these circles. Note that σ has to touch γ at another point; otherwise we could increase its radius slightly while keeping the circle in F.

Fix a point p_1 and let σ_1 be the corresponding circle. Denote by γ_1 an arc of γ from p_1 to a first point q_1 on σ_1. Denote by $\hat\sigma_1$ and $\check\sigma_1$

two arcs of σ_1 from p_1 to q_1 such that the cyclic concatenation of $\hat\sigma_1$ and γ_1 surrounds $\check\sigma_1$.

Let p_2 be the midpoint of γ_1 and σ_2 be the corresponding circle.

Note that σ_2 can not intersect $\hat\sigma_1$. Otherwise, if σ_2 intersects $\hat\sigma_1$ at some point s, then σ_2 has two more common points with $\check\sigma_1$ — x and y, one for each arc of σ_2 from p_2 to s. That is σ_1 and σ_2 have common point s, x and y. Therefore $\sigma_1 = \sigma_2$ as two circles with three common points. On the other hand, by construction $p_2 \in \sigma_2$ and $p_2 \notin \sigma_1$ — a contradiction.



Two ovals on the diagram pretend to be circles.

Recall that σ_2 has to touch γ at another point. From above it follows that it can only touch γ_1 and therefore we can choose an arc $\gamma_2 \subset \gamma_1$ that runs from p_2 to a first point q_2 on σ_2. Note that by construction we have that

$$\text{❻} \qquad \operatorname{length}\gamma_2 < \tfrac{1}{2}\cdot \operatorname{length}\gamma_1.$$

Let us repeat this construction recessively. We get an infinite sequence of arcs $\gamma_1 \supset \gamma_2 \supset \dots$. By ❻, we also get that

$$\operatorname{length}\gamma_n \to 0 \quad \text{as} \quad n \to \infty.$$

Therefore the intersection

$$\bigcap_n \gamma_n$$

contains a single point; denote it by p_∞.

Let σ_∞ be the corresponding circle at p_∞; it has to touch γ at another point q_∞. The same argument as above shows that $q_\infty \in \gamma_n$ for any n. It follows that $q_\infty = p_\infty$ — a contradiction.

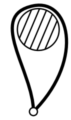
Now suppose that there is only one point $q \in \gamma$ at which osculating circle supports γ from the left. In this case for any point $p \neq q$ on γ the corresponding circle touches γ at another point.

Chose a point $p_1 \neq q$ on γ, take its corresponding circle σ_1 and note that there are two choices for arc γ_1 one of which does not contain q. Repeating the same construction starting from γ_1 we also arrive to a contradiction.

It remains to show that existence of a pair of osculating circlines that support γ form the right. This is done the same way, one only has to change the definition of corresponding circline — given $p \in \gamma$, it has to be the circline of maximal signed curvature that supports γ from the right at p. We leave it as an exercise:

5.26. Exercise. *List the necessary changes in the proof above for the existence of circlines that support γ form the right.* □

Theorem 5.24 admits the following generalization:



5.27. Theorem. *Let γ be a smooth regular simple plane loop. Suppose that absolute curvature of γ does not exceed 1. Then γ surrounds a unit circle.*

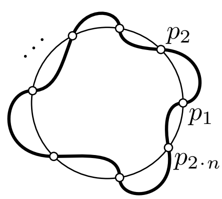
5.28. Exercise. *Describe the modifications in the proof of 5.23 which are necessary to prove 5.27.*

5.29. Exercise. *Assume that a closed smooth regular curve γ lies in a figure F bounded by a closed simple plane curve. Suppose that R is the maximal radius of discs that lies in F. Show that absolute curvature of γ is at least $\frac{1}{R}$ at some parameter value.*

Hint: Note that γ contains a simple loop; apply to it 5.27.

5.30. Advanced exercise. *Suppose γ is a closed simple smooth regular plane curve and σ is a circle. Assume γ crosses σ at the points $p_1, \dots, p_{2\cdot n}$ and these points appear in the same cycle order on γ and on σ. Show that osculating circles at n distinct points of γ lie inside γ and that osculating circles at other n distinct points of γ lie outside of γ. In particular the curve γ has at least $2\cdot n$ vertexes.*

Construct an example of a closed simple smooth regular plane curve γ with only 4 vertexes that crosses a given circle at arbitrary many points.



Hint: Repeat the proof of theorem for each cyclic concatenation of an arc of γ from p_i to p_{i+1} with large arc of the circle.

Recall that the *inverse* of a point x with respect to the unit circle centered at the origin is the point $\hat{x} = \frac{x}{|x|^2}$.

5.31. Advanced exercise. *Suppose γ is a smooth regular plane curve that does not pass thru the origin. Let $\hat{\gamma}$ be the inversion of γ in the unit circle centered at the origin. Show that osculating circline of $\hat{\gamma}$ at s is the inversion of osculating circline of γ at s. Conclude that every vertex of $\hat{\gamma}$ is the inversion of a vertex of γ.*

Hint: Use the definition of osculating circle via order of contact and that inversion maps circles to circlines.

Note that the exercise provides an alternative way to finish the proof of 5.23 — once we proved the existence of two osculating circles that support γ from the left, we can apply to γ inversion with the center surrounded by γ. In this case the curve γ is mapped to a curve $\hat{\gamma}$, the domain inside γ is mapped to the domain outside $\hat{\gamma}$ and the other way around. It follows that if an osculating circle supports the obtained curve $\hat{\gamma}$ on the right then its inversion supports γ from the left and the other way around. That is from the existance of two supporting circles on the left we also get the existance of two supporting circles on the right.

Part II

Surfaces

Chapter 6

Definitions

General definition

Few times we will need the following general definition.

A path connected subset Σ in a metric space is called *surface* (more precisely *embedded surface without boundary*) if any point of $p \in \Sigma$ admits a neighborhood W in Σ which is *homeomorphic* to an open subset in the Euclidean plane; that is, if there is an injective continuous map $U \to W$ from an open set $U \subset \mathbb{R}^2$ such that its invese $W \to U$ is also continuous.

However, as well as in the case of curves we will be mostly interested in smooth surfaces in the Euclidean space describe in the following section.

Smooth surfaces

Recall that a function f of two variables x and y is called *smooth* if all its partial derivatives $\frac{\partial^{m+n}}{\partial x^m \partial y^n} f$ are defined and are continuous in the domain of definition of f.

A path connected set $\Sigma \subset \mathbb{R}^3$ is called a *smooth surface* (or more precisely *smooth regular embedded surface*) if it can be described locally as a graph of a smooth function in an appropriate coordinate system.

More precisely, for any point $p \in \Sigma$ one can choose a coordinate system (x, y, z) and a neighborhood $U \ni p$ such that the intersection $W = U \cap \Sigma$ is formed by a graph $z = f(x, y)$ of a smooth function f defined in an open domain of the (x, y)-plane.

Examples. The simplest example of a surface is the (x, y)-plane

$$\Pi = \left\{ (x, y, z) \in \mathbb{R}^3 \,:\, z = 0 \right\}.$$

The plane Π is a surface since it can be described as the graph of the function $f(x,y)=0$.

All other planes are surfaces as well since one can choose a coordinate system so that it becomes the (x,y)-plane. We can also present a plane as a graph of a linear function $f(x,y)=a{\cdot}x+b{\cdot}y+c$ for some constants a, b and c (assuming the plane is not perpendicular to the (x,y)-plane).

A more interesting example is the unit sphere

$$\mathbb{S}^2=\left\{\,(x,y,z)\in\mathbb{R}^3\;:\;x^2+y^2+z^2=1\,\right\}.$$

This set is not the graph of any function, but $\mathbb{S}^2$ is locally a graph; in fact it can be covered by 6 graphs:

$$\begin{aligned} z=f_\pm(x,y)&=\pm\sqrt{1-x^2-y^2},\\ y=g_\pm(x,z)&=\pm\sqrt{1-x^2-z^2},\\ x=h_\pm(y,z)&=\pm\sqrt{1-y^2-z^2}; \end{aligned}$$

each function $f_\pm, g_\pm, h_\pm$ is defined in an open unit disc. That is, $\mathbb{S}^2$ is a smooth surface.

More conventions. If the surface Σ is formed by a closed set, then it is called *complete*. For example, paraboloids

$$z=x^2+y^2,\qquad z=x^2-y^2$$

or sphere

$$x^2+y^2+z^2=1$$

are complete surfaces, while the open disc in a plane

$$\left\{\,(x,y,z)\in\mathbb{R}^3\;:\;x^2+y^2<1, z=0\,\right\}$$

is a surface which is not complete.

If moreover Σ is a compact set, then it is called *closed surface* (the term is closely related to *closed curve* but has nothing to do with *closed set*).

If a complete surface Σ is noncompact, then it is called *open surface* (again the term *open set* is not relevant).

For example, paraboloids are open surfaces, and sphere is closed.

A closed subset in a surface that is bounded by one or more smooth curves is called *surface with boundary*; in this case the collection of curves is called the *boundary line* of the surface. When we say *surface* we usually mean a surface without boundary; we may use the term *surface with possibly nonempty boundary* if we need to talk about surfaces with and without boundary.

Local parametrizations

Let U be an open domain in $\mathbb{R}^2$ and $s\colon U \to \mathbb{R}^3$ be a smooth map. We say that s is regular if its Jacobian has maximal rank; in this case it means that the vectors $\frac{\partial s}{\partial u}$ and $\frac{\partial s}{\partial v}$ are linearly independent at any $(u,v) \in U$ or equivalently $\frac{\partial s}{\partial u} \times \frac{\partial s}{\partial v} \neq 0$, where $\times$ denotes the vector product.

6.1. Proposition. *If $s\colon U \to \mathbb{R}^3$ is a smooth regular embedding of an open connected set $U \subset \mathbb{R}^2$, then its image $\Sigma = s(U)$ is a smooth surface.*

Proof. Let $s(u,v) = (s_1(u,v), s_2(u,v), s_3(u,v))$. Since s is regular the Jacobian matrix

$$\begin{pmatrix} \frac{\partial s_1}{\partial u} & \frac{\partial s_1}{\partial v} \\ \frac{\partial s_2}{\partial u} & \frac{\partial s_2}{\partial v} \\ \frac{\partial s_3}{\partial u} & \frac{\partial s_3}{\partial v} \end{pmatrix}$$

has rank two.

Fix a point $p \in \Sigma$; by shifting the coordinate system we may assume that p is the origin. Permuting the coordinates x, y, z if necessary, we may assume that the matrix

$$\begin{pmatrix} \frac{\partial s_1}{\partial u} & \frac{\partial s_1}{\partial v} \\ \frac{\partial s_2}{\partial u} & \frac{\partial s_2}{\partial v} \end{pmatrix}$$

is invertible. Let $\bar{s}\colon U \to \mathbb{R}^2$ be the projection of s to the (x,y)-coordinate plane; that is, $\bar{s}(u,v) = (s_1(u,v), s_2(u,v))$. Note that the 2×2-matrix above is the Jacobian matrix of $\bar{s}$.

The inverse function theorem implies that there is a smooth regular function h defined on an open set $W \ni 0$ in the (x,y)-plane such that $h(0,0) = (0,0)$ and $\bar{s} \circ h$ is the identity map.

Note that the graph $z = s_3 \circ h(x,y)$ for $(x,y) \in W$ is a subset in Σ. Indeed, if $(u,v) = h(x,y)$, then $x = s_1(u,v)$ and $y = s_2(u,v)$. Therefore the identity $z = s_3 \circ h(x,y)$ can be rewritten as $(x,y,z) = {}$ $= s(u,v)$.

Clearly the graph $z = s_3 \circ h(x,y)$ for $(x,y) \in W$ is open in Σ; that is, Γ a neighborhood of p in Σ that can be described as a graph of a smooth function $f_3 \circ h\colon W \to \mathbb{R}$. Since p is arbitrary, we get that Σ is a surface. $\square$

If s and Σ as in the proposition, then we say that s is a *parametrization* of the surface Σ.

Not all the smooth surfaces can be described by such a parametrization; for example the sphere $\mathbb{S}^2$ cannot. But any smooth surface Σ admits a local parametrization; that is, any point $p \in \Sigma$ admits an open neighborhood $W \subset \Sigma$ with a smooth regular parametrization s. In this case any point in W can be described by two parameters, usually denoted by u and v, which are called *local coordinates* at p. The map s is called a *chart* of Σ.

If W is a graph $z = h(x, y)$ then the map $s\colon (u, v) \mapsto (u, v, h(u, v))$ is a chart. Indeed, s has an inverse $(u, v, h(u, v)) \mapsto (u, v)$ which is continuous; that is, s is an embedding. Further, $\frac{\partial s}{\partial u} = (1, 0, \frac{\partial h}{\partial u})$ and $\frac{\partial s}{\partial v} = (0, 1, \frac{\partial h}{\partial v})$, whence $\frac{\partial s}{\partial u}$ and $\frac{\partial s}{\partial v}$ are linearly independent.

Note that from 6.1, we obtain the following corollary.

6.2. Corollary. *A path connected set $\Sigma \subset \mathbb{R}^3$ is a smooth regular surface if at any point $p \in \Sigma$ it has a local parametrization by a smooth regular map.*

6.3. Exercise. *Consider the following map*

$$s(u, v) = (\tfrac{2\cdot u}{1+u^2+v^2}, \tfrac{2\cdot v}{1+u^2+v^2}, \tfrac{2}{1+u^2+v^2}).$$

Show that s is a chart of the unit sphere centered at $(0, 0, 1)$; describe the image of s.

The map

$$(u, v, 1) \mapsto (\tfrac{2\cdot u}{1+u^2+v^2}, \tfrac{2\cdot v}{1+u^2+v^2}, \tfrac{2}{1+u^2+v^2})$$

is called *stereographic projection.* Note that the point $(u, v, 1)$ and its image $(\tfrac{2\cdot u}{1+u^2+v^2}, \tfrac{2\cdot v}{1+u^2+v^2}, \tfrac{2}{1+u^2+v^2})$ lie on one half-line starting at the origin.

Let $\gamma(t) = (x(t), y(t))$ be a plane curve. Recall that the image of the map

$$(t, \theta) \mapsto (x(t), y(t)\cdot \cos\theta, y(t)\cdot \sin\theta)$$

is called *surface of revolution* of the curve γ around x-axis.

6.4. Exercise. *Assume γ is a closed simple smooth regular plane curve that does not intersect x-axis. Show that surface of revolution of γ around x-axis is a smooth regular surface.*

Golbal parametrizations

A surface can be described by an embedding from a known surface to the space. For example the ellipsoid

$$\Sigma_{a,b,c} = \left\{ (x, y, z) \in \mathbb{R}^3 \ : \ \tfrac{x^2}{a^2} + \tfrac{y^2}{b^2} + \tfrac{z^2}{c^2} = 1 \right\}$$

for some positive numbers a, b, c can be defined as the image of the map $s\colon \mathbb{S}^2 \to \mathbb{R}^3$, defined as the restriction of the map $(x, y, z) \mapsto$ $\mapsto (a{\cdot}x, b{\cdot}y, c{\cdot}z)$ to the unit sphere $\mathbb{S}^2$.

For a surface Σ, a map $s : \Sigma \to \mathbb{R}^3$ is called a *smooth parametrized surface* if for any chart $f\colon U \to \Sigma$ the composition $s \circ f$ is smooth and regular; that is, all partial derivatives $\frac{\partial^{m+n}}{\partial u^m \partial v^n}(s \circ f)$ exist and are continuous in the domain of definition and the following two vectors $\frac{\partial}{\partial u}(s \circ f)$ and $\frac{\partial}{\partial v}(s \circ f)$ are linearly independent.

Evidently the parametric definition includes the embedded surfaces defined previously — as the domain of parameters we can take the surface itself and the identity map as s. But parametrized surfaces are more general, in particular they might have self-intersections.

If Σ is a known surface for example a sphere or a plane, the paramtrized surface $s\colon \Sigma \to \mathbb{R}^3$ might be called by the same name. For example, any embedding $s\colon \mathbb{S}^2 \to \mathbb{R}^3$ might be called topological sphere and if s is smooth and regular, then it might be called smooth sphere. (A smooth regular map $s\colon \mathbb{S}^2 \to \mathbb{R}^3$ which is not necessary an embedding is called *smooth regular immersion*, so we can say that s describes a smooth immersed sphere.) Similarly an embedding $s\colon \mathbb{R}^2 \to \mathbb{R}^3$ might be called topological plane and if s is smooth it might be called smooth plane.

Implicitly defined surfaces

6.5. Proposition. *Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be a smooth function. Suppose that 0 is a regular value of f; that is, $\nabla_p f \neq 0$ if $f(p) = 0$. Then any path connected component Σ of the set of solutions of the equation $f(x, y, z) = 0$ is a surface.*

Proof. Fix $p \in \Sigma$. Since $\nabla_p f \neq 0$ we have

$$\tfrac{\partial f}{\partial x}(p) \neq 0, \quad \tfrac{\partial f}{\partial y}(p) \neq 0, \quad \text{or} \quad \tfrac{\partial f}{\partial z}(p) \neq 0.$$

We may assume $\frac{\partial f}{\partial z}(p) \neq 0$; otherwise permute the coordinates x, y, z.

The implicit function theorem (A.10) implies that a neighborhood of p in Σ is the graph $z = h(x, y)$ of a smooth function h defined on an open domain in $\mathbb{R}^2$. It remains to apply the definition of smooth surface (page 58). $\square$

6.6. Exercise. *Describe the set of real numbers a such that the equation*

$$x^2 + y^2 - z^2 = a$$

describes a smooth regular surface.

Tangent plane

Let $z = f(x, y)$ be a local graph realization of a surface. Assume that a point $p = (x_p, y_p, z_p)$ lies on this graph, so $z_p = f(x_p, y_p)$. The plane spanned by the vectors $(1, 0, (\frac{\partial}{\partial x} f)(x_p, y_p))$ and $(0, 1, (\frac{\partial}{\partial y} f)(x_p, y_p))$ is called the *tangent plane* of Σ at p. The tangent plane to Σ at p is usually denoted by T_p or $\mathrm{T}_p\Sigma$. Vectors in T_p are called *tangent vectors* of Σ at p.

Tangent plane T_p might be considered as a linear subspace of $\mathbb{R}^3$ or as a parallel plane passing thru p. In the latter case it can be interpreted as the best approximation of the surface Σ by a plane at p; it has *first order of contact* with Σ at p; that is, $\rho(q) = o(|p - q|)$, where $q \in \Sigma$ and $\rho(q)$ denotes the distance from q to T_p.

6.7. Proposition. *Let Σ be a smooth surface. A vector w is a tangent vector of Σ at p if and only if there is a curve γ that runs in Σ and has w as a velocity vector at p.*

Note that according to the proposition the tangent plane $\mathrm{T}_p\Sigma$ can be defined as the set of all velocity vectors at p of smooth parameterized curves that run in Σ. In particular the tangent plane to a surface at a given point does not depend on the choice of its local graph representation.

Proof. We can assume that Σ is a graph $z = f(x, y)$; otherwise pass to a local presentation of Σ around p.

"Only if" part. Suppose that $(x(t), y(t))$ denotes the projection of $\gamma(t)$ to the (x, y)-plane. Since γ runs in Σ, we have that

$$\gamma(t) = \big(x(t), y(t), f(x(t), y(t))\big).$$

Therefore

$$\begin{aligned}\gamma' &= (x', y', \tfrac{\partial f}{\partial x}(x, y) \cdot x' + \tfrac{\partial f}{\partial y}(x, y) \cdot y') = \\ &= x' \cdot (1, 0, (\tfrac{\partial}{\partial x} f)(x, y)) + y' \cdot (0, 1, (\tfrac{\partial}{\partial y} f)(x, y)).\end{aligned}$$

That is, $\gamma'(t) \in \mathrm{T}_{\gamma(t)}$ for any t.

"If" part. Without loss of generality we can assume that p is the origin. Fix a tangent vector

$$w = a \cdot (1, 0, (\tfrac{\partial}{\partial x} f)(0, 0)) + b \cdot (0, 1, (\tfrac{\partial}{\partial y} f)(0, 0))$$

and consider the curve $\gamma(t) = (a \cdot t, b \cdot t, f(a \cdot t, b \cdot t))$. By construction γ runs in Σ and the direct calculations show that $\gamma'(0) = w$. $\square$

6.8. Exercise. *Assume $f\colon U\to\mathbb{R}^3$ is a smooth regular chart of a surface Σ and $p=f(u_0,v_0)$. Show that the tangent plane $\mathrm{T}_p\Sigma$ is spanned by vectors $\frac{\partial f}{\partial u}(u_0,v_0)$ and $\frac{\partial f}{\partial v}(u_0,v_0)$.*

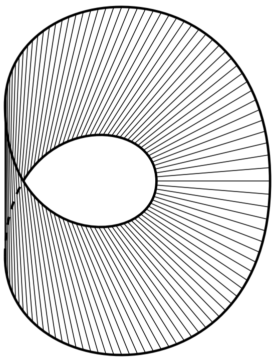
6.9. Exercise. *Let $f:\mathbb{R}^3\to\mathbb{R}$ be a smooth function with a regular value 0 and Σ is a surface described as a connected component of the set of solutions $f(x,y,z)=0$. Show that the tangent plane $\mathrm{T}_p\Sigma$ is perpendicular to the gradient $\nabla_p f$ at any point $p\in\Sigma$.*

6.10. Exercise. *Let Σ be a smooth surface and $p\in\Sigma$. Fix an (x,y,z)-coordinates. Show that a neighborhood of p in Σ is a graph $z=f(x,y)$ of a smooth function f defined on an open subset in (x,y)-plane if and only if the tangent plane T_p is not a* vertical plane*; that is if the projection of T_p to (x,y)-plane does not degenerates to a line.*

Normal vector and orientation

A unit vector that is normal to T_p is usually denoted by ν_p; it is uniquely defined up to sign.

A surface Σ is called *oriented* if it is equipped with a unit normal vector field ν; that is, a continuous map $p\mapsto\nu_p$ such that $\nu_p\perp\mathrm{T}_p$ and $|\nu_p|=1$ for any p. The choice of the field ν is called *orientation* on Σ. A surface Σ is called *orientable* if it can be oriented. Note that each orientable surface admits two orientations ν and $-\nu$.



Möbius strip shown on the diagram gives an example of nonorientable surface — there is no choice of normal vector field that is continuous along the middle of the strip, when you go around it changes the sign.

Note that each surface is locally orientable. In fact each chart $f(u,v)$ admits an orientation

$$\nu=\frac{\frac{\partial f}{\partial u}\times\frac{\partial f}{\partial v}}{\left|\frac{\partial f}{\partial u}\times\frac{\partial f}{\partial v}\right|}.$$

Indeed the vectors $\frac{\partial f}{\partial u}$ and $\frac{\partial f}{\partial v}$ are tangent vectors at p; since they are linearly independent, their vector product does not vanish and it is perpendicular to the tangent plane. Therefore $\nu(u,v)$ is a unit normal vector at $f(u,v)$; evidently $(u,v)\mapsto\nu(u,v)$ is a continuous map.

6.11. Exercise. *Let $h:\mathbb{R}^3\to\mathbb{R}$ be a smooth function with a regular value 0 and Σ is a surface described as a connected component of the set of solutions $h(x,y,z)=0$. Show that Σ is orientable.*

Hint. Show that $\nu=\frac{\nabla h}{|\nabla h|}$ defines a unit normal field on Σ.

In fact any complete smooth surface cuts the space into two connected components. Therefore one could choose an orientation on any complete surface by taking normal vector at each point that points into one of these components. In particular the Möbius strip can not be extended to a complete smooth surface.

Plane sections

6.12. Advanced exercise. *Show that any closed set in the plane can appear as an intersection of this plane with a complete smooth regular surface.*

As a consequence of the exercise above, the section of a smooth regular surface by a plane might look complicated. The following lemma makes possible to perturb the plane so that the section becomes nice.

6.13. Lemma. *Let Σ be a smooth regular surface. Then for any plane Π there is arbitrarily close parallel plane Π' such that each connected component of the intersection $\Sigma\cap\Pi'$ is a smooth regular curve.*

Proof. Assume Π is described by equation $f(x,y,z)=r_0$, where

$$f(x,y,z)=a{\cdot}x+b{\cdot}y+c{\cdot}z.$$

The surface Σ can be covered by a countable set of charts $s_i\colon U_i\to\Sigma$. Note that the composition $f\circ s_i$ is a smooth function. Therefore almost all real numbers r are regular values of each $f\circ s_i$.

Fix such value r sufficiently close to r_0 and consider the plane Π' described by the equation $f(x,y,z)=r$. Note that $\Pi'\parallel\Pi$ and arbitrary close to it. Any point in the intersection $\Sigma\cap\Pi'$ lies in the image of one of the charts. From above it admits a neighborhood which is a regular smooth curve; hence the result. $\square$

Chapter 7

Curvatures

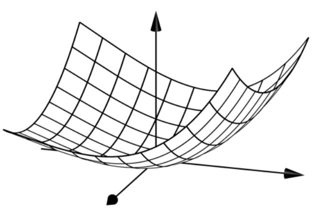
Tangent-normal coordinates

Fix a point p in a smooth surface Σ. Consider a coordinate system (x, y, z) with origin at p such that the (x, y)-plane coincides with T_p. By 6.10 we can present Σ locally around p as a graph of a function f. Note that f satisfies the following additional properties:

$$f(0,0) = 0, \qquad (\tfrac{\partial}{\partial x} f)(0,0) = 0, \qquad (\tfrac{\partial}{\partial y} f)(0,0) = 0.$$

The first equality holds since $p = (0,0,0)$ lies on the graph and the last two equalities mean that the tangent plane at p is horizontal.

Consider the Hessian matrix



❶ $$M_p = \begin{pmatrix} \ell & m \\ m & n \end{pmatrix},$$

where

$$\begin{aligned} \ell &= (\tfrac{\partial^2}{\partial x^2} f)(0,0), \\ m &= (\tfrac{\partial^2}{\partial x \partial y} f)(0,0) = (\tfrac{\partial^2}{\partial y \partial x} f)(0,0), \\ n &= (\tfrac{\partial^2}{\partial y^2} f)(0,0). \end{aligned}$$

The components of the matrix describe the surface at up to the second order at p. In fact the so called *osculating paraboloid*

$$z = \tfrac{1}{2}(\ell \cdot x^2 + 2 \cdot m \cdot x \cdot y + n \cdot y^2)$$

gives the best approximation of the surface at p; it has *second order of contact* with Σ at p.

Given two vectors v, w in the (x, y)-plane, consider the value

$$\mathrm{I\!I}_p(v, w) := (D_w D_v f)(0, 0),$$

where D denotes the directional derivative. The function $(v, w) \mapsto \mapsto \mathrm{I\!I}_p(v, w)$ is called *second fundamental form* at p; it takes two tangent vector v and w at p and spits the real number $\mathrm{I\!I}_p(v, w)$.

The second fundamental form can be written in terms of the Hessian matrix. Indeed if $w = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ and $v = \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)$, then

$$D_w = a \cdot \tfrac{\partial}{\partial x} + b \cdot \tfrac{\partial}{\partial y} \quad \text{and} \quad D_v = c \cdot \tfrac{\partial}{\partial x} + d \cdot \tfrac{\partial}{\partial y}.$$

Therefore

❷
$$\begin{aligned} \mathrm{I\!I}_p(w, v) &:= (D_w D_v f)(0, 0) = \\ &= a \cdot c \cdot \ell + (a \cdot d + b \cdot c) \cdot m + b \cdot d \cdot n = \\ &= \langle M_p \cdot w, v \rangle = \\ &= \langle M_p \cdot v, w \rangle. \end{aligned}$$

Note that from ❷ it follows that $\mathrm{I\!I}_p$ is symmetric; that is,

❸
$$\mathrm{I\!I}_p(v, w) = \mathrm{I\!I}_p(w, v)$$

for any two tangent vectors $v, w \in \mathrm{T}_p$.

Principle curvatures

Note that tangent-normal coordinates give an almost canonical coordinate system in a neighborhood of p; it is unique up to a rotation of the (x, y)-plane and switching the sign of the z-coordinate. Rotating the (x, y)-plane is equivalent to changing its orthonormal basis, which results in the rewriting the Hessian matrix in this new basis.

Since Hessian matrix is symmetric, it is diagonalizable by orthogonal matrices. That is, by rotating (x, y)-plane we can assume that $m = 0$ in ❶. In this case the diagonal components of M_p are called *principle curvatures* of Σ at p; they are uniquely defined up to sign; they are denoted as $k_1(p)$ and $k_2(p)$ or $k_1(p)_\Sigma$ and $k_2(p)_\Sigma$ if we need to emphasize that these are the curvatures of the surface Σ. We will always assume that $k_1 \leqslant k_2$.

The principle curvatures can be also defined as the eigenvalues of M_p; the eigendirections of M_p are called *principle directions* of Σ at p. Note that if $k_1(p) \neq k_2(p)$ then p has exactly two principle directions, which are perpendicular to each other.

Note that if we revert the orientation of Σ, then the principle curvatures at each point switch their signs and indexes.

Normal curvature

Assume we choose the coordinates in the (x,y)-plane so that the Hessian matrix is diagonalized, we can assume that

$$M_p = \begin{pmatrix} k_1(p) & 0 \\ 0 & k_2(p) \end{pmatrix}.$$

According to ❷, the second directional derivative for a vector $w = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ in the (x,y)-plane can be written as

$$(D_w^2 f)(0,0) = a^2 \cdot k_1(p) + b^2 \cdot k_2(p).$$

If w is unit, then the second directional derivative $D_w^2 f(0,0)$ can be intepreted as the signed curvature of the curve formed by the intersection of Σ with the plane thru p spanned by ν_p and w. In this case $\mathrm{I\!I}_p(w,w) = D_w^2 f(0,0)$ is called *normal curvature* in the direction w; it is denoted by $k_w(p)$ or $k_w(p)_\Sigma$.

Since $|w| = 1$, we have $a^2 + b^2 = 1$ which implies the following:

7.1. Observation. *For any point p on an oriented smooth surface Σ, the principle curvatures $k_1(p)$ and $k_2(p)$ are correspondingly minimum and maximum of the normal curvatures at p. Moreover, if θ is the angle between a unit vector $w \in \mathrm{T}_p$ and the first principle direction at p, then*

$$k_w(p) = k_1(p) \cdot \cos^2\theta + k_2(p) \cdot \sin^2\theta.$$

The last identity is the so called *Euler's formula.*

A smooth regular curve on a surface Σ that always runs in the principle directions is called *line of curvature* of Σ.

7.2. Exercise. *Assume that a smooth surface Σ is mirror symmetric with respect to a plane Π. Suppose that Σ and Π intersect along a curve γ. Show that γ is a line of curvature of Σ.*

7.3. Exercise. *Assume V is a body of revolution in $\mathbb{R}^3$ and its boundary is a smooth surface with principle curvatures at most 1 in absolute value. Show that V contains a unit ball.*

Hint: Use 7.2 and 5.24.

More curvatures

Fix an oriented smooth surface Σ and a point $p \in \Sigma$.

The product

$$K(p) = k_1(p) \cdot k_2(p)$$

is called Gauss curvature at p. We may denote it by $K(p)_\Sigma$ if we need to emphasize that this is curvature of Σ. The Gauss curvature can be also interpreted as determinant of the Hessian matrix M_p.

The sum

$$H(p) = \tfrac{1}{2} \cdot (k_1(p) + k_2(p))$$

is called mean curvature at p. We may denote it by $H(p)_\Sigma$ if we need to emphasize that this is curvature of Σ. The mean curvature can be also interpreted as half of the trace of the Hessian matrix M_p.

Note that the Gauss curvature depends only on Σ and p, and not on the choice of the coordinate system. The same is true up to sign for the mean curvature — it changes the sign if we revert the orientation of the surface.

7.4. Exercise. *Show that any surface with positive Gauss curvature is orientable.*

Supporting surfaces

Assume two oriented surfaces Σ_1 and Σ_2 have a common point p. If there is a neighborhood U of p such that $\Sigma_1 \cap U$ lies on one side from Σ_2 in U, then we say that Σ_2 *locally supports* Σ_1 at p.

7.5. Exercise. *Let Σ_1 and Σ_2 be two smooth surfaces. Assume Σ_2* locally supports *Σ_1 at a point p. Show that $\mathrm{T}_p\Sigma_1 = \mathrm{T}_p\Sigma_2$; that is, the tangent planes of Σ_1 and Σ_2 at p coincide.*

By the exercise, we can assume that Σ_1 and Σ_2 are cooriented at p; that is, they have common unit normal vector at p. If not we can revert the orientation of one of the surfaces.

If Σ_1 and Σ_2 are cooriented at p, then we can say that Σ_1 locally supports Σ_2 from *inside* or from *outside*, assuming that the normal vector points *inside* the domain bounded by surface Σ_2 in U.

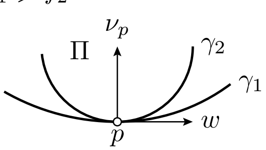
More precisely, we can use for Σ_1 and Σ_2 one tangent-normal coordinate system at p, assuming that the axis z points in the direction of the unit normal vector ν_p to both surfaces. This way we write Σ_1 and Σ_2 locally as graphs: $z = f_1(x, y)$ and $z = f_2(x, y)$ correspondingly. Then Σ_1 locally supports Σ_2 from inside (from outside) if $f_1(x, y) \geqslant f_2(x, y)$ (correspondingly $f_1(x, y) \leqslant f_2(x, y)$) for (x, y) in a sufficiently small neighborhood of the origin.

Note that Σ_1 locally supports Σ_2 from inside at the point p is equivalent to Σ_2 locally supports Σ_1 from outside. Further if we revert the orientation of both surfaces then supporting from inside becomes supporting from outside and the other way around.

7.6. Proposition. *Let Σ_1 and Σ_2 be oriented surfaces. Assume Σ_1 locally supports Σ_2 from inside at the point p (equivalently Σ_2 locally supports Σ_1 from outside). Then $k_1(p)_{\Sigma_1} \geqslant k_1(p)_{\Sigma_2}$ and $k_2(p)_{\Sigma_1} \geqslant$ $\geqslant k_2(p)_{\Sigma_2}$.*

7.7. Exercise. *Give an example of two surfaces Σ_1 and Σ_2 that have common point p with common unit normal vector ν_p such that $k_1(p)_{\Sigma_1} > k_1(p)_{\Sigma_2}$ and $k_2(p)_{\Sigma_1} > k_2(p)_{\Sigma_2}$, but Σ_1 does not support Σ_2 locally at p.*

Proof. We can assume that Σ_1 and Σ_2 are graphs $z = f_1(x, y)$ and $z = f_2(x, y)$ in a common tangent-normal coordinates at p, so we have $f_1 \geqslant f_2$.



Fix a unit vector $w \in \mathrm{T}_p\Sigma_1 = \mathrm{T}_p\Sigma_2$. Consider the plane Π passing thru p and spanned by the normal vector ν_p and w. Let γ_1 and γ_2 be the curves of intersection of Σ_1 and Σ_2 with Π.

Let us orient Π so that ν_p points to the left from w. Further, let us parametrize both curves so that they are running in the direction of w at p and therefore cooriented. In this case the curve γ_1 supports the curve γ_2 from the right.

Therefore we have the following inequality for the normal curvatures of Σ_1 and Σ_2 at p in the direction of w:

❹ $$k_w(p)_{\Sigma_1} \geqslant k_w(p)_{\Sigma_2}.$$

According to 7.1,

$$k_1(p)_{\Sigma_i} = \min\{\, k_w(p)_{\Sigma_i} \;:\; w \in \mathrm{T}_p, |w| = 1 \,\}$$

for $i = 1, 2$. Choose w so that $k_1(p)_{\Sigma_1} = k_w(p)_{\Sigma_1}$. Then by ❹, we have that

$$\begin{aligned} k_1(p)_{\Sigma_1} &= k_w(p)_{\Sigma_1} \geqslant \\ &\geqslant k_w(p)_{\Sigma_2} \geqslant \\ &\geqslant \min\{\, k_w(p)_{\Sigma_2} \,\} = \\ &= k_1(p)_{\Sigma_2}; \end{aligned}$$

that is, $k_1(p)_{\Sigma_1} \geqslant k_1(p)_{\Sigma_2}$.

Similarly, by 7.1, we have that

$$k_2(p)_{\Sigma_i} = \max\{\, k_w(p)_{\Sigma_i} \,\}.$$

Let us fix w so that $k_2(p)_{\Sigma_2} = k_w(p)_{\Sigma_2}$. Then

$$\begin{aligned}
k_2(p)_{\Sigma_2} &= k_w(p)_{\Sigma_2} \leqslant \\
&\leqslant k_w(p)_{\Sigma_1} \leqslant \\
&\leqslant \max\{\, k_w(p)_{\Sigma_1} \,\} = \\
&= k_2(p)_{\Sigma_1};
\end{aligned}$$

that is, $k_2(p)_{\Sigma_1} \geqslant k_2(p)_{\Sigma_2}$. □

7.8. Corollary. *Let Σ_1 and Σ_2 be oriented surfaces. Assume Σ_1 locally supports Σ_2 from inside at the point p. Then*

(a) $H(p)_{\Sigma_1} \geqslant H(p)_{\Sigma_2}$;
(b) If $k_1(p)_{\Sigma_2} \geqslant 0$, then $K(p)_{\Sigma_1} \geqslant K(p)_{\Sigma_2}$.

Proof. By (7.6), we get that $k_1(p)_{\Sigma_1} \geqslant k_1(p)_{\Sigma_2}$ and $k_2(p)_{\Sigma_2} \geqslant k_2(p)_{\Sigma_2}$. Therefore part (a) follows since

$$H(p)_{\Sigma_i} = \tfrac{1}{2}\cdot(k_1(p)_{\Sigma_i} + k_2(p)_{\Sigma_i})).$$

(b). Since $k_2(p)_{\Sigma_i} \geqslant k_1(p)_{\Sigma_i}$ and $k_1(p)_{\Sigma_2} \geqslant 0$, we get that all the principle curvatures $k_1(p)_{\Sigma_1}$, $k_1(p)_{\Sigma_1}$ and $k_2(p)_{\Sigma_1}$ and $k_2(p)_{\Sigma_2}$ are nonnegative. Whence

$$\begin{aligned}
K(p)_{\Sigma_1} &= k_1(p)_{\Sigma_1}\cdot k_2(p)_{\Sigma_1} \geqslant \\
&\geqslant k_1(p)_{\Sigma_2}\cdot k_2(p)_{\Sigma_2} = \\
&= K(p)_{\Sigma_2}.
\end{aligned}$$

□

7.9. Exercise. *Show that any closed surface has a point with positive Gauss curvature.*

Hint. Consider the minimal sphere that encloses the surface.

7.10. Exercise. *Assume that a closed surface Σ surrounds a unit disc. Show that Gauss curvature of Σ is at most 1 at some point.*

Try to prove the same assuming that Σ surrounds a unit circle only.

Hint. Look for a supporting spherical dome with the unit circle as the boundary.

Curve in a surface

Recall that the second fundamental form $\mathrm{I\!I}_p$ is defined on page 67.

7.11. Proposition. *Suppose γ is a smooth curve in a smooth oriented surface Σ with a unit normal field ν. Then, the following identity holds for any time parameter t:*

$$\langle\gamma''(t),\nu_{\gamma(t)}\rangle=\mathrm{I\!I}_{\gamma(t)}(\gamma'(t),\gamma'(t)).$$

Proof. Fix a parameteer value t_0; set $p=\gamma(t_0)$, $v=\gamma'(t_0)$ and $a=$ $=\gamma''(t_0)$; so we need to show that

❺ $$\langle a,\nu_p\rangle=\mathrm{I\!I}_p(v,v).$$

Let $z=f(x,y)$ be the local representaion of Σ in the tangent-normal coordinates at p; we assume that ν points in the derection of ν_p.

Without loss of generality may assume that γ runs in the graph $z=f(x,y)$; so

$$\gamma(t)=\left(x(t),y(t),f(x(t),y(t))\right).$$

Then

$$\begin{aligned}\gamma'&=(x',y',\tfrac{\partial f}{\partial x}\cdot x'+\tfrac{\partial f}{\partial y}\cdot y');\\ \gamma''&=(x'',y'',\tfrac{\partial^2 f}{\partial x^2}\cdot(x')^2+2\cdot\tfrac{\partial^2 f}{\partial x\partial y}\cdot x'\cdot y'+\tfrac{\partial^2 f}{\partial y^2}\cdot(y')^2+\tfrac{\partial f}{\partial x}\cdot x''+\tfrac{\partial f}{\partial y}\cdot y'').\end{aligned}$$

Recall that $p=\gamma(t_0)=(0,0,0)$ and

$$f(0,0)=0,\qquad \tfrac{\partial f}{\partial x}(0,0)=0,\qquad \tfrac{\partial f}{\partial y}(0,0)=0.$$

Therefore

$$\begin{aligned}v&=(x',y',0)\,(t_0);\\ a&=\left(x'',y'',\tfrac{\partial^2 f}{\partial x^2}\cdot(x')^2+2\cdot\tfrac{\partial^2 f}{\partial x\partial y}\cdot x'\cdot y'+\tfrac{\partial^2 f}{\partial y^2}\cdot(y')^2\right)(t_0).\end{aligned}$$

Note that

$$\mathrm{I\!I}_p(v,v)=\left(\tfrac{\partial^2 f}{\partial x^2}\cdot(x')^2+2\cdot\tfrac{\partial^2 f}{\partial x\partial y}\cdot x'\cdot y'+\tfrac{\partial^2 f}{\partial y^2}\cdot(y')^2\right)(t_0);$$

that is, the z-coordinate of the acelaration a equals to $\mathrm{I\!I}_p(v,v)$ which is equivalent to ❺. □

7.12. Corollary. *Let γ be a regular smooth curve that runs in a smooth surface Σ. Suppose $p=\gamma(t_0)$ and $w=\gamma'(t_0)$. Denote by α the* *angle between the unit normal to Σ at p and the unit normal vector in the Frenet frame of γ. Then the following identity holds for the curvature $k(t_0)_\gamma$ of γ at p and the normal curvature $k_w(p)$ of Σ at p in the direction of w:*

$$k(t_0)_\gamma \cdot \cos\alpha = k_w(p).$$

Proof. Denote by ν_Σ the unit normal vector to Σ at p and by ν_γ the unit normal vector in the Frenet frame of γ. Note that $\cos\alpha = \langle \nu_\Sigma, \nu_\gamma\rangle$.

Applying 7.11, we get that

$$\begin{aligned} k_w(p) &= \mathrm{I\!I}_p(w,w) = \\ &= \langle \gamma'', \nu_\Sigma \rangle = \\ &= k(t_0)_\gamma \cdot \langle \nu_\gamma, \nu_\Sigma\rangle = \\ &= k(t_0)_\gamma \cdot \cos\alpha. \end{aligned}$$

□

The corollary above, as well as the statement in the following exercise are proved by Jean Baptiste Meusnier [20].

7.13. Exercise. *Let Σ be a smooth surface, $p \in \Sigma$ and $w \in \mathrm{T}_p\Sigma$ is a unit vector. Assume that $k_w(p) \neq 0$; that is the normal curvature of Σ at p in the direction of w does not vanish.*

Show that the osculating circles at p of smooth regular curves in Σ that run in the direction w sweep out a sphere.

7.14. Exercise. *Let $\gamma(t) = (x(t), y(t))$ be a smooth unit-speed simple plane curve in the upper half-plane. Suppose that Σ is the surface of revolution of γ with respect to the x-axis.*

Express the principle curvatures of Σ at $(x(t), y(t), 0)$ in terms of $y(t)$, $y'(t)$ and $y''(t)$. Conclude that $-\frac{y''(t)}{y(t)}$ is the Gauss curvature of Σ at $(x(t), y(t), 0)$.

Hint: Use 7.2 and 7.12.

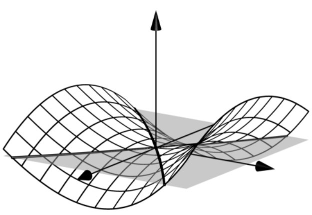
7.15. Exercise. *Assume that a regular smooth curve γ lies in a surface of positive Gauss curvature. Show that curvature of γ does not vanish at any value.*

Chapter 8

Saddle surfaces

Definitions

A surface is called *saddle* if its Gauss curvature at each point is nonpositive; in other words principle curvatures at each point have opposite signs or one of them is zero.



If the Gauss curvature is negative at each point, then the surface is called *strictly saddle*; equivalently it means that the principle curvatures have opposite signs at each point. Note that in this case tangent plane is does not support the surface even locally — moving along the surface in the principle directions at a given point, one gets above and below the tangent plane at this point.

8.1. Exercise. *Let $f\colon \mathbb{R}\to\mathbb{R}$ be a smooth positive function. Show that the surface of revolution of the graph $y=f(x)$ around the x-axis is saddle if and only if f is convex; that is if $f''(x)\geqslant 0$ for any x.*

Hint: Use 7.14.

A surface Σ is called *ruled* if for every point $p\in\Sigma$ there is a line segment $\ell_p\subset\Sigma_p$ thru p that is infinite or has its endpoint(s) on the boundary line of Σ.

8.2. Exercise. *Show that any ruled surface Σ is saddle.*

Hint: Prove and use that each point $p\in\Sigma$ has a direction with vanishing normal curvature.

8.3. Exercise. *Assume Σ is complete saddle surface. Show that for any point $p \in \Sigma$ there is a curve $\gamma\colon [0,\infty) \to \Sigma$ that starts at p such that $t \mapsto |\gamma(t)|$ is an increasing function and $|\gamma(t)| \to \infty$ as $t \to \infty$.*

A tangent direction on a smooth surface with vanishing normal curvature is called *asymptotic*. A smooth regular curve that always run in an asymptotic direction is called *asymptotic line*.

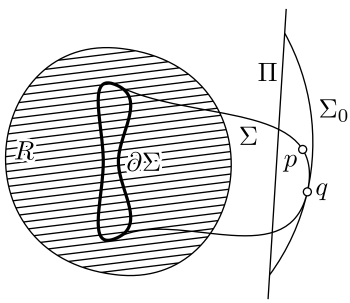
8.4. Advanced exercise. *Let $\Sigma \subset \mathbb{R}^3$ be the graph $z = f(x,y)$ of a smooth function f and γ be a closed smooth asymptotic line in Σ. Assume Σ is strictly saddle in a neighborhood of γ. Show that the projection of γ to the (x,y)-plane cannot be star-shaped.*

Hats

Note that a closed surface cannot be saddle. Indeed consider a smallest sphere that contains a closed surface Σ inside; it supports Σ at some point p and at this point the principle curvature must have the same sign. The following more general statement is proved using the same idea.

8.5. Lemma. *Assume Σ is a compact saddle surface and its boundary line lies in a convex closed region R. Then $\Sigma \subset R$.*

Proof. Assume contrary; that is, there is point $p \in \Sigma$ that does not lie in R. Let Π be a plane that separates p from R; it exists by A.18. Denote by Σ' the part of Σ that lies with p on the same side from Π.



Since Σ is compact, it is surrounded by a sphere S; let σ be the circle of intersection of S and Π. Consider the smallest spherical dome Σ_0 with boundary σ that includes Σ'.

Note that Σ_0 supports Σ at some point q. Without loss of generality we may assume that Σ_0 and Σ are cooriented at q and Σ_0 has positive principle curvatures. In this case Σ_0 supports Δ from outside. By 7.8, $K(q)_\Sigma \geqslant K(q)_{\Sigma_0} > 0$ — a contradiction. $\square$

Note that if we assume that Σ is strictly saddle, then we could arrive to a contradiction by taking a point q on the same side with p and on the maximal distance from Π.

8.6. Exercise. *Let Δ be a smooth regular saddle disc and $p \in \Delta$. Assume that the boundary line $\partial\Delta$ lies in the unit sphere centered at p. Show that* $\operatorname{length}(\partial\Delta) \geqslant 2{\cdot}\pi$.

Hint: Use the lemma above and the hemisphere lemma (2.17).

8.7. Exercise. *Show that an open saddle surface can not lie inside of an infinite circular cone.*

A disc Δ in a surface Σ is called a *hat* of Σ if its boundary line $\partial\Delta$ lies in a plane Π and the remaining points of Δ lie on one side of Π.

8.8. Proposition. *A smooth surface Σ is saddle if and only if it has no hats.*

Note that a saddle surface can contain a closed plane curve. For example the hyperboloid $x^2 + y^2 - z^2 = 1$ contains the unit circle in the (x, y)-plane centered at the origin. However, a plane curve can not bound a disc (as well any compact set) in a saddle surface.

Proof. Since plane is a convex set, the "only if" part follows from 8.5; it remains to prove the "if" part.

Assume Σ is not saddle; that is, it has a point p with strictly positive Gauss curvature; or equivalently, the principle curvatures $k_1(p)$ and $k_2(p)$ have the same sign.

Let $z = f(x, y)$ be a graph representation of Σ in the tangent-normal coordinates at p. Without loss of generality we may assume that both principle curvatures are positive, or equivalently the

$$D_w^2 f(0,0) = \mathrm{I\!I}_p(w, w) > 0$$

for any unit tangent vector $w \in \mathrm{T}_p\Sigma$ (which is the (x, y)-plane).

Since the set of unit vectors is compact, we have that

$$D_w^2 f(0,0) > \varepsilon$$

for some fixed $\varepsilon > 0$ and any unit tangent vector $w \in \mathrm{T}_p\Sigma$. By continuity of the function $(x, y, w) \mapsto D_w^2 f(x, y)$, we have that $D_w^2 f(x, y) > 0$ for (x, y) in a neighborhood of the origin. That is, f is a strictly convex function in a neighborhood of the origin in the (x, y)-plane. In particular the set

$$\Delta_\varepsilon = \left\{\, (x, y, f(x, y)) \in \mathbb{R}^3 \;:\; f(x, y)) \leqslant \varepsilon \,\right\}$$

is a disc for sufficiently small $\varepsilon > 0$ (see 8.10). Note that its boundary line lies on the plane $z = \varepsilon$ and whole disc lies below it; that is, Δ_ε is a hat. $\square$

Note that we proved the following lemma; it will be useful latter.

8.9. Lemma. *Let $z = f(x, y)$ be the local description of a smooth surface Σ in a tangent-normal coordinates at some point $p \in \Sigma$. Assume both principle curvatures of Σ are positive at p. Then the function f is strictly convex in a neighborhood of the origin.*

In the proof above we assumed that the statement in following exercise is evident.

8.10. Exercise. *Let Δ_ε be as in the proof above. Show that Δ_ε is a smooth disc; that is, Δ_ε is the image of regular embedding $\mathbb{D} \to \mathbb{R}^3$, where $\mathbb{D} = \left\{ (x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leqslant 1 \right\}$.*

Hint: Observe that it is sufficient to construct a smooth parametrization of Δ_ε by a closed hemisphere. To do this repeat the argument in 9.3 with the center at a point surrounded by the boundary line of Δ_ε in its plane.

8.11. Exercise. *Let $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation; that is, $T(x, y, z) = (x, y, z) \cdot A$ for a invertible 3×3-matrix A. Show that for any saddle surface Σ the image $T(\Sigma)$ is also a saddle surface.*

Saddle graphs

The following theorem was proved by Sergei Bernstein [21].

8.12. Theorem. *Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be a smooth function. Assume its graph $z = f(x, y)$ is a strictly saddle surface in $\mathbb{R}^3$. Then f is not bounded; that is, there is no constant C such that $|f(x, y)| \leqslant C$ for any $(x, y) \in \mathbb{R}^2$.*

Before going into the proof let us discuss some examples.

Note that the theorem states that a saddle graph can not lie between parallel horizontal planes; applying 8.11 we get that saddle graph can not lie between parallel planes, not necessarily horizontal. The following exercise shows that the theorem does not hold for saddle surface which are not graphs.

8.13. Exercise. *Construct a complete strictly saddle surface that lies between parallel planes.*

Hint: Look for an example among the surfaces of revolution and use 7.14.

The following exercise shows that there are saddle graphs with functions bounded on one side; that is, both (upper and lower) bounds are needed in the proof of Bernshtein's theorem.

8.14. Exercise. *Show that there are positive functions with strictly saddle graphs. In fact the graph $z = \exp(x - y^2)$ is strictly saddle.*

Hint: Look at two section of the graph by planes parallel to (x, y)-plane and to (x, z)-plane and apply Meusnier's theorem (7.12).

Note that according to 8.5, there are no complete saddle surfaces in a parallelepiped that boundary line lies on one of its faces. The following lemma gives an analogous statement for a parallelepiped with an infinite side.

8.15. Lemma. *There is no complete strictly saddle smooth surface with the boundary line in a plane that lies on bounded distance from a line.*

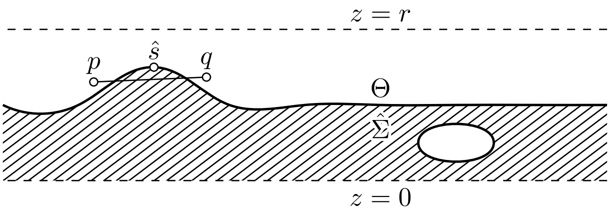
Proof. Note that in a suitable coordinate system, the statement can be reformulated the following way: *There is no complete strictly saddle smooth surface with the boundary line in the (x, y)-plane that lies in a region of the following form:*

$$R = \left\{ (x, y, z) \in \mathbb{R}^3 \ : \ 0 \leqslant z \leqslant r, 0 \leqslant y \leqslant r \right\}.$$

Further we will prove this statement.

Assume contrary, let Σ be such a surface. Consider the projection $\hat{\Sigma}$ of Σ to the (x, z)-plane. It lies in the upper half-plane and below the line $z = r$.

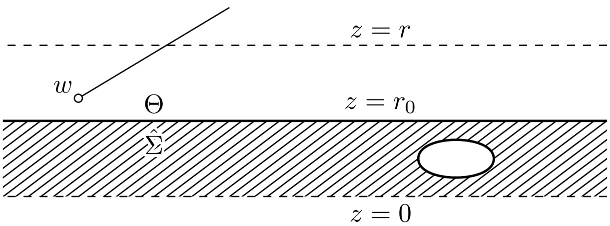
Consider the open upper half-plane $H = \left\{ (x, z) \in \mathbb{R}^2 \ : \ z > 0 \right\}$. Let Θ be the connected component of the complement $H \backslash \hat{\Sigma}$ that contains all the points above the line $z = r$.



Note that Θ is convex. If not then a line segment $[pq]$ for some $p, q \in \Theta$ cuts from $\hat{\Sigma}$ a compact piece. Consider the plane Π thru $[pq]$ that is perpendicular to the (x, z)-plane. Note that π cuts from Σ a compact region Δ. By general position argument 6.13, we can assume that Δ is a compact surface with boundary line in Π and the remaining part of Δ lies on one side from Π. Since the plane Π is a convex, this statement contradicts 8.5.

Summarizing, Θ is an open convex set of H that contains all points above $z = r$. By convexity, together with any point w, the set Θ

contains all points on the half-lines that point up from it. Whence it contain all points with z-coordinate larger than the z-coordinate of w. Since Θ is open it can be described by inequality $z > r_0$. It follows that

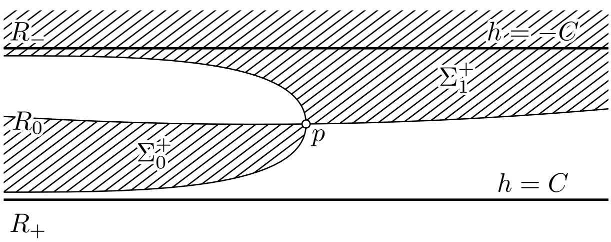


the plane $z = r_0$ supports Σ at some point (in fact at many points). By 7.6, the latter is impossible — a contradiction. □

Proof of 8.12. Denote by Σ the graph $z = f(x, y)$. Assume contrary; that is, Σ lies between two planes $z = \pm C$.

Note f can not be constant. It follows that the tangent plane T_p at some point $p \in \Sigma$ is not horizontal.

Denote by Σ^+ the part of Σ that lies above T_p. Note that it has at least two connected components which are approaching p from both sides in the principle direction with positive principle curvature. Indeed if there would be a curve that runs in Σ^+ and approaches p from both sides then it would cut a disc from Σ with boundary line above T_p and some points below it; the later contradicts 8.5.



The surface Σ seeing from above.

Summarizing, Σ^+ has at least two connected components, denote them by Σ_0^+ and Σ_1^+. Let $z = h(x, y) = a{\cdot}x + b{\cdot}y + c$ be the equation of T_p. Note that Σ^+ contains all points in the region

$$R_- = \{\, (x, y, f(x, y)) \in \Sigma \,:\, h(x, y) < C \,\}$$

which is a connected set and no points in

$$R_+ = \{\, (x, y, f(x, y)) \in \Sigma \,:\, h(x, y) > C \,\}$$

Whence one of the connected components, say Σ_0^+, lies in

$$R_0 = \{\,(x,y,f(x,y)) \in \Sigma \,:\, |h(x,y)| \leqslant C\,\}\,.$$

This set lies on a bounded distance from the line of intersection of T_p with the (x,y)-plane.

Moving the plane T_p little up, we can cut from Σ_0^+ is a complete surface with boundary line lying in this plane (see 6.13). The obtained surface is still on a bounded distance to a line which is impossible by 8.15. □

The following exercise gives a condition that guarantees that a saddle surface is a graph; it can be used in combination with Bernshtein's theorem.

8.16. Advanced exercise. *Let Σ be a smooth saddle disk in $\mathbb{R}^3$. Assume that the orthogonal projection to the (x,y)-plane maps the boundary line of Σ injectively to a convex closed curve. Show that the orthogonal projection to (x,y)-plane is injective on Σ.*

In particular, Σ is the graph $z = f(x,y)$ of a function f defined on a convex figure in the (x,y)-plane.

Remarks

Note that Bernstein's theorem and the lemma in its proof do not hold for saddle surfaces; counterexamples can be found among infinite cylinders over a smooth regular curves. In fact it can be shown that these are only counterexamples; the proof is based on the same idea, but more technical.

By 8.8, saddle surfaces can be defined as smooth surfaces without hats. This definition can be used for arbitrary surfaces not necessarily smooth. Some results, for example Bernshtein's characterization of saddle graphs can be extended to generalized saddle surface, but this class of surfaces is far from being understood. Some nontrivial properties were proved by in Samuil Shefel [22] see also [23, Capter 4].

Chapter 9

Positive Gauss curvature

Convexity

9.1. Exercise. *Suppose that an oriented surface Σ bounds a convex region R.*

(a) Show that Gauss curvature of Σ is nonnegative at each point.

(b) Show that for any point $p \in \Sigma$ and q in the interior of R we have that

$$\langle \nu_p, q - p \rangle > 0,$$

where ν_p is the unit normal vector at p that points in R.

Hint: Show and use that any tangent plane T_p supports Σ at p.

Recall that a region R in the Euclidean space is called *strictly convex* if for any two points $x, y \in R$, any point z between x and y lies in the interior of R.

Clearly any open convex set is strictly convex; the cube (as well as any convex polyhedron) gives an example of a convex set which is not strictly convex. It is easy to see that a convex region is strictly convex if and only if its boundary does not contain a line segment.

The following theorem gives a global description of surfaces with positive Gauss curvature.

9.2. Theorem. *Assume Σ is a complete smooth surface with positive Gauss curvature. Then Σ bounds a strictly convex region.*

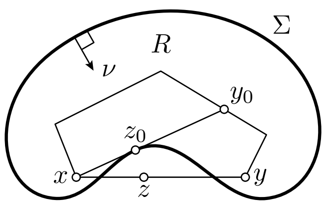
Note that in the proof we have to use that surface is a connected set; otherwise a pair of disjoint spheres which bound two disjoint balls would give a counterexample.

Proof. Since the Gauss curvature is positive, we can choose unit normal field ν on Σ so that the both principle curvatures are positive at any point. Let R be the region bounded by Σ that lies on the side of ν; that is ν points inside of R at any point of Σ.

Fix $p \in \Sigma$; let $z = f(x, y)$ be a local description of Σ in the tangent-normal coordinates at p. By 8.9, f is strictly convex in a neighborhood of the origin. In particular the intersection of a small ball centered at p with the epigraph $z \geqslant f(x, y)$ is strictly convex. In other words, R is *locally strictly convex*; that is, for any point $p \in R$, the intersection of R with a small ball centered at p is strictly convex.

Since Σ is connected, so is R; moreover any two points in the interior of R can be connected by a polygonal line in the interior of R.

Assume the interior of R is not convex; that is, there are points $x, y \in R$ and a point z between x and y that does not lie in the interior of R. Consider a polygonal line β from x to y in the interior of R. Let y_0 be the first point on β such that the chord $[x, y_0]$ touches Σ at some point, say z_0.



Since R is locally strictly convex, $R \cap B(z_0, \varepsilon)$ is strictly convex for all sufficiently small $\varepsilon > 0$. On the other hand z_0 lies between two points in the intersection $[x, y_0] \cap B(z_0, \varepsilon)$. Since $[x, y_0] \subset R$, we arrived to a contradiction.

Therefore the interior of R is a convex sets. Note that the region R is the closure of its interior, therefore R is convex as well.

Since R is locally strictly convex, its boundary Σ contains no line segments. Therefore R is strictly convex. $\square$

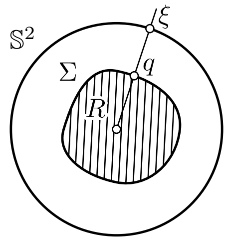
In fact only minor modifications of the proof above imply that any connected locally convex region is convex.

Closed surfaces

9.3. Lemma. *Assume Σ is a closed smooth surface with positive Gauss curvature. Then Σ is a smooth sphere; that is, Σ admits a smooth regular parametrization by $\mathbb{S}^2$.*

Proof. Without loss of generality we can assume that the origin lies in the interior of the convex region R bounded by Σ.

By convexity of R, any half-line starting at the origin intersects Σ at a single point; that is there is a positive function $\rho\colon \mathbb{S}^2 \to \mathbb{R}$ such that Σ is formed by points $q = \rho(\xi)\cdot\xi$ for $\xi \in \mathbb{S}^2$.



Let us show that ρ is a smooth function. Fix a point $p = (x_p, y_p, z_p)$ on Σ. Consider a local implicit description of Σ at p as a solution of equation $h(x, y, z) = 0$ with nonvanishing gradient; so $h(p) = 0$. Note that for any point q in a neighborhood of p, we have that

$$h(q) = 0 \iff q \in \Sigma \iff q = \rho(\xi)\cdot\xi$$

for some $\xi \in \mathbb{S}^2$. In other words h defines implicitly ρ as in the implicit function theorem.

Recall that $\nabla_p h \perp \mathrm{T}_p$. Since the origin lies in the interior of R, it can not lie on T_p; that is $\langle\nabla_p f, p\rangle \neq 0$; or equivalently $D_\eta h(p) \neq 0$, where $\eta = \frac{p}{|p|}$ is the unit vector in the direction of p and D denotes the directional derivative.

Fix a (u, v) chart s on $\mathbb{S}^2$ in a neighborhood of η; note that the map

$$S\colon (u, v, w) \mapsto w\cdot s(u, v)$$

is smooth and regular for $w > 0$; that is, the vectors

$$\tfrac{\partial S}{\partial u}, \quad \tfrac{\partial S}{\partial v}, \quad \tfrac{\partial S}{\partial w} = w$$

are linearly independent. Note that the function $h \circ S$ is smooth and $\frac{\partial h\circ S}{\partial \rho}(p) = D_\eta h(p) \neq 0$. Applying implicit function theorem, we get that ρ is smooth in a neighborhood of η; since η is arbitrary, ρ is smooth on whole $\mathbb{S}^2$. $\square$

If one only needs to show that Σ is a topological sphere, then one only needs to show that ρ is continuous. The latter is a consequence of another classical result it topology — the so called *closed graph theorem.*

Open surfaces

9.4. Lemma. *Suppose Σ is an open surface in with positive Gauss curvature. Then there there is a coordinate system such that Σ is a graph $z = f(x, y)$ of a convex function f defined on a convex open region of (x, y)-plane.*

Proof. The surface Σ is a boundary of an unbounded closed convex region R.

We can assume that the origin lies on Σ. Consider a sequence of points $x_n \in \Sigma$ such that $|x_n| \to \infty$ as $n \to \infty$. Set $u_n = \frac{x_n}{|x_n|}$; this is the unit vector in the direction from x_n. Since the unit sphere is compact, we can pass to a subsequence of (x_n) such that u_n converges to a unit vector u.

Note that for any $q \in \Sigma$, the directions $v_n = \frac{x_n - q}{|x_n - q|}$ converge to u as well. Indeed, the sequences of vectors $\frac{x_n - q}{|x_n - q|}$, $\frac{x_n - q}{|x_n|}$ and $\frac{x_n}{|x_n|}$ have the same limit since $\frac{|q|}{|x_n|} \to 0$ and $\frac{|x_n|}{|x_n - q|} \to 1$ as $n \to \infty$; the latter follows since

$$\begin{aligned}|\tfrac{|x_n|}{|x_n-q|} - 1| &= |\tfrac{|x_n|-|x_n-q|}{|x_n-q|} - 1| \leqslant \\ &\leqslant \tfrac{|p-q|}{|x_n-q|}.\end{aligned}$$

Morover, the half-line from q in the direction of u lies in R. Indeed any point on the half-line is a limit of points on the line segments $[q, x_n]$; since R is closed, all of these poins lie in R.

Let us choose the z-axis in the direction of u. Note that there is no point $p \in \Sigma$ with vertical tangent plane. Otherwise Σ would contain a vertical half-line starting from p and therefore Gauss curvature of Σ would vanish. It follows that any vertical line can intersect Σ at most at one point. That is, Σ is a graph $z = f(x, y)$ of a smooth convex function f.

Let Ω be the domain of definition of f. Note that Ω is the projection of Σ to (x, y)-plane which is the same as the projection of R. Since R is convex, so is Ω.

Since no tangent plane is vertical, the projection from Σ to the (x, y)-plane is regular. By inverse function theorem, the set Ω is open.

□

9.5. Exercise. *Show that any open surface Σ with positive Gauss curvature is a topological plane; that is, there is an embedding $\mathbb{R}^2 \to \mathbb{R}^3$ with image Σ.*

Try to show that Σ is a smooth plane; that is, the embedding f can be made smooth and regular.

9.6. Exercise. *Show that any open smooth surface Σ with positive Gauss curvature lies inside of an infinite circular cone.*

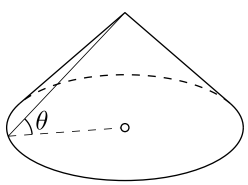
9.7. Exercise. *Let Σ be an open smooth surface with positive Gauss curvature. Show that Σ has a point with arbitrary small Gauss curvature.*

Hint: Observe that the Gauss curvature of the surface of revolution of the graph $y = a \cdot \sin(x/b)$ for $x \in (0, \frac{\pi}{b})$ can not exceed $\frac{1}{b^2}$ (Use 3.4 and 7.12). Try to support the surface Σ from inside by a surface of revolution of the described type with large b.

Chapter 10

Geodesics

We start to study the intrinsic geometry of surfaces. The following exercise should help you to be in the right mood for this; it might look like a tedious problem in calculus, but actually it is an easy problem in geometry.



10.1. Exercise. *There is a mountain of frictionless ice with the shape of a perfect cone with a circular base. A cowboy is at the bottom and he wants to climb the mountain. So, he throws up his lasso which slips neatly over the top of the cone, he pulls it tight and starts to climb. If the angle of inclination θ is large, there is no problem; the lasso grips tight and up he goes. On the other hand if the angle of inclination θ is small, the lasso slips off as soon as the cowboy pulls on it.*

What is the critical angle θ_0 at which the cowboy can no longer climb the ice-mountain?

Hint: Cut the lateral surface of the mountain by a line from the cowboy to the top, unfold it on the plane and try to figure out what is the image of the strained lasso.

Shortest paths

Let p and q be two points on a surface Σ. Recall that $|p-q|_\Sigma$ denotes the induced length distance from p to q; that is, the exact lower bound on lengths of paths in Σ from p to q.

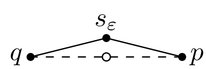
Note that if Σ is smooth, then any two points in Σ can be joined by a piecewise smooth path. Since any such path is rectifiable, the value $|p-q|_\Sigma$ is finite for any pair of points $p,q\in\Sigma$.

A path γ from p to q in Σ that minimize the length is called a *shortest path* from p to q.

The image of a shortest path between p and q in Σ is usually denoted by $[p,q]_\Sigma$. In general there might be no shortest path between two given points on the surface and it might many of them; this is shown in the following two examples. However if we write $[p,q]_\Sigma$, then we assume that a shortest path exists and we made a choice of one of them.

Nonuniqueness. There plenty of shortest paths between the poles on the sphere — each meridian is a shortest path.

Nonexistence. Let us give an example of a surface with a pair of points without shortest path. Consider the surface Σ which is the (x,y)-plane with removed origin with two points $p=(1,0,0)$ and $q=(-1,0,0)$.



Note that $|p-q|_\Sigma=2$. Indeed, given $\varepsilon>$ >0, consider the point $s_\varepsilon=(0,\varepsilon,0)$. Note that the polygonal path $ps_\varepsilon q$ lies in Σ and its length $2\cdot\sqrt{1+\varepsilon^2}$ approaches 2 as $\varepsilon\to0$. It follows that $|p-q|_\Sigma\leqslant2$. On the other hand $|p-q|_\Sigma\geqslant|p-q|_{\mathbb{R}^3}=2$; that is, $|p-q|_\Sigma=2$.

Therefore a shortest path from p to q (if it exists) must have length 2. By triangle inequality any curve of length 2 from p to q must run along the line segment $[p,q]$; in particular it must pass thru the origin. Since the origin does not lie in Σ, there is no shortest from p to q in Σ

10.2. Proposition. *Any two points in a complete smooth surface can be joined by a shortest path.*

Proof. Fix a complete smooth surface Σ with two points p and q. Set $\ell=|p-q|_\Sigma$.

By the definition of induced length metric, there is a sequence of paths γ_n from p to q in Σ such that

$$\operatorname{length}\gamma_n\to\ell\quad\text{as}\quad n\to\infty.$$

Without loss of generality, we may assume that $\operatorname{length}\gamma_n<\ell+1$ for any n and each γ_n is parameterized proportional to its arclength. In particular each path $\gamma_n\colon[0,1]\to\Sigma$ is $(\ell+1)$-Lipschitz; that is,

$$|\gamma(t_0)-\gamma(t_1)|\leqslant(\ell+1)\cdot|t_0-t_1|$$

for any $t_0, t_1 \in [0,1]$. Further the image of γ_n lies in the closed ball $\bar{B}[p, \ell+1]$ for any n. It follows that γ_n is a uniformly continuous sequence of curves with image in a compact set. Whence we can pass to a a converging subsequence of γ_n; denote by $\gamma_\infty\colon [0,1] \to \mathbb{R}^3$ its limit. As a limit of uniformly continuous sequence, γ_∞ is continuous; that is, γ_∞ is a path. Evidently γ_∞ runs from p to q. Since Σ is a closed set, γ_∞ lies in Σ. Finally, by 2.13,

$$\gamma_\infty \leqslant \ell;$$

that is, γ_∞ is a shortest path from p to q. $\square$

Closest point projection

10.3. Lemma. *Let R be a closed convex set in $\mathbb{R}^3$. Then for every point $p \in \mathbb{R}^3$ there is unique point $\bar{p} \in R$ that minimizes the distance $|p-x|$ among all points $x \in R$.*

Moreover the map $p \mapsto \bar{p}$ is short; that is,

$$❶ \qquad |p-q| \geqslant |\bar{p}-\bar{q}|$$

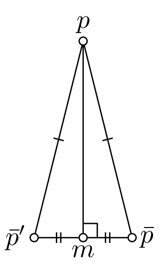
for any pair of points $p, q \in \mathbb{R}^3$.

The map $p \mapsto \bar{p}$ is called the *closest point projection*; it maps the Euclidean space to R. Note that if $p \in R$, then $\bar{p} = p$.

Proof. Fix a point p and set

$$\ell = \inf_{x \in R}\{|p-x|\}.$$

Choose a sequence $x_n \in R$ such that $|p-x_n| \to \ell$ as $n \to \infty$.



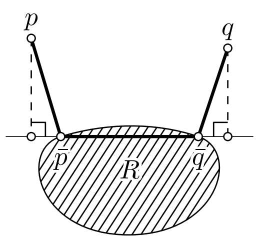
Without loss of generality, we can assume that all the points x_n lie in a ball or radius $\ell+1$ centered at p. Therefore we can pass to a partial limit $\bar{p}$ of x_n; that is, $\bar{p}$ is a limit of a subsequence of x_n. Since R is closed $\bar{p} \in R$. By construction

$$|p-\bar{p}| = \ell = \lim_{n\to\infty} |p-x_n|.$$

Hence the existence follows.

Assume there are two distinct points $\bar{p}, \bar{p}' \in R$ that minimize the distance to p. Since R is convex, their midpoint $m = \frac{1}{2}\cdot(\bar{p}+\bar{p}')$ lies in R. Note that $|p-\bar{p}| = |p-\bar{p}'| = \ell$; that is $\triangle p\bar{p}\bar{p}'$ is isosceles and therefore $\triangle p\bar{p}m$ is right with the right angle at m. Since a leg of a right triangle is shorter than its hypotenuse, we have $|p-m| < \ell$ — a contradiction.

It remains to prove inequality ❶.



We can assume that $\bar p \neq \bar q$, otherwise there is nothing to prove. Note that if $p \neq \bar p$ (that is, if $p \notin R$), then $\angle p\bar p\bar q$ is right or obtuse. Otherwise there would be a point x on the line segment $[\bar q, \bar p]$ that is closer to p than $\bar p$. Since R is convex, the line segment $[\bar q, \bar p]$ and therefore x lie in R. Hence $\bar p$ is not closest to p — a contradiction.

The same way we can show that if $q \neq \neq \bar q$, then $\angle q\bar q\bar p$ is right or obtuse. In all cases it implies that the orthogonal projection of the line segment $[p, q]$ to the line $\bar p\bar q$ contains the line segment $[\bar p, \bar q]$. In particular

$$|p - q| \geqslant |\bar p - \bar q|. \qquad \square$$

10.4. Corollary. *Assume a surface Σ bounds a closed convex region R and $p, q \in \Sigma$. Denote by W the outer closed region of Σ; in other words W is the union of Σ and the complement of R. Then for any curve γ in W that runs from p to q we have*

$$\operatorname{length}\gamma \geqslant |p - q|_\Sigma.$$

Moreover if $\gamma \not\subset \Sigma$, then the inequality is strict.

Proof. The first part of the corollary follows from the lemma and the definition of length. Indeed consider the closest point projection $\bar\gamma$ of γ. Note that $\bar\gamma$ lies in Σ and connects p to q therefore

$$\operatorname{length}\bar\gamma \geqslant |p - q|_\Sigma.$$

Indeed, consider an inscribed polygonal line $p_0 \dots p_n$ in γ. Denote by $\bar p_i$ the closest point projection of p_i to R. Note that the polygonal line $\bar p_0 \dots \bar p_n$ is inscribed in $\bar\gamma$; moreover any inscribed polygonal line in $\bar\gamma$ can appear this way. By 10.3 $|p_i - p_{i-1}| \geqslant |\bar p_i - \bar p_{i-1}|$ for any i. Therefore

$$\operatorname{length} p_0 \dots p_n \geqslant \operatorname{length} \bar p_0 \dots \bar p_n.$$

Taking least upper bound of each side of the inequality for all inscribed polygonal lines $p_0 \dots p_n$ in γ, we get

$$\operatorname{length}\gamma \geqslant \operatorname{length}\bar\gamma.$$

Whence the first statement follows.

To prove the second statement, note that if $s = \gamma(t_1) \notin \Sigma$, then $s \notin R$. Hence there is a plane Π that cuts s from Σ. The curve γ must intersect at least at two points: one point before t_1 and one after; let $a = \gamma(t_0)$ and $b = \gamma(t_2)$ be these points. Note that the arc of γ from a to b is strictly longer that $|a - b|$; indeed on the way γ visits s that is not on the plane Π and therefore not on the lie segment $[a, b]$.

Remove from γ the arc from a to b and glue in the line segment $[a, b]$; denote the obtained curve by γ_1. From above,

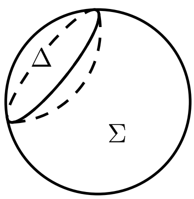
$$\operatorname{length}\gamma > \operatorname{length}\gamma_1$$

Note that γ_1 is runs in W. Therefore by the first part of corollary, we have

$$\operatorname{length}\gamma_1 \geqslant |p - q|_\Sigma.$$

Whence the second statement follows. □

10.5. Exercise. *Suppose Σ is a complete smooth surface that bounds a convex region R in $\mathbb{R}^3$ and Π is a plane that cuts a hat Δ from Σ. Assume that the reflection of the interior of Δ with respect to Π lies in the interior of R. Show that Δ is* convex *with respect to the intrinsic metric of Σ; that is, if both ends of a shortest path in Σ lie in Δ, then the entire geodesic lies in Δ.*



Let us define the *intrinsic diameter* of a closed surface Σ as the exact upper bound on the lengths of shortest paths in the surface.

10.6. Exercise. *Assume that a closed smooth surface Σ with positive Gauss curvature lies in a unit ball. Show that the intrinsic diameter of Σ cannot exceed π.*

Hint: Use 10.3.

Geodesics

A smooth curve γ on a smooth surface Σ is called *geodesic* if its acceleration $\gamma''(t)$ is perpendicular to the tangent plane $\mathrm{T}_{\gamma(t)}$ for each t.

Geodesics can be understood as the trajectories of a particle that slides on Σ without friction. In this case the force that keeps the particle on Σ must be perpendicular to Σ. By the second Newton's laws of motion, we get that the acceleration γ'' is perpendicular to $\mathrm{T}_{\gamma(t)}$.

10.7. Exercise. *Assume that a smooth surface Σ is mirror symmetric with respect to a plane Π. Suppose that Σ and Π intersect along a curve γ. Show that γ is a geodesic of Σ.*

10.8. Lemma. *Any geodesic γ has constant speed; that is $|\gamma'(t)|$ is constant.*

Proof. Since $\gamma'(t)$ is a tangent vector at $\gamma(t)$, we have that $\gamma''(t) \perp \perp \gamma'(t)$, or equivalently $\langle\gamma'', \gamma'\rangle = 0$ for any t. Whence

$$\begin{aligned}\langle\gamma', \gamma'\rangle' &= 2\cdot\langle\gamma'', \gamma'\rangle = \\ &= 0.\end{aligned}$$

That is $|\gamma'(t)|^2 = \langle\gamma'(t), \gamma'(t)\rangle$ is constant. ☐

10.9. Proposition. *Given a tangent vector v to a smooth surface Σ at a point p There is a unique geodesic $\gamma\colon \mathbb{I} \to \Sigma$ defined on a maximal open interval $\mathbb{I} \ni 0$ that starts at p with velocity vector v; that is, such that $\gamma(0) = p$ and $\gamma'(0) = v$.*

Moreover

(a) the map $(p, v, t) \mapsto \gamma(t)$ is smooth in its domain of definition.

(b) if Σ is complete, then $\mathbb{I} = \mathbb{R}$; that is, the maximal interval is whole real line.

Sketch of proof. The first part of the proposition and part (a) follows from existence and uniqueness of a solution of initial value problem (A.12). One only needs to rewrite the condition $\gamma''(t) \perp \mathrm{T}_{\gamma(t)}$ as a differential equation $\gamma''(t) = \mathrm{I\!I}_{\gamma(t)}(\gamma'(t), \gamma'(t))$.

The part (b) follows from 10.8. Indeed by A.12, if the maximal interval is not whole real line, then the curve γ must escape to infinity. But the later is impossible since γ runs with constant speed. ☐

Exponential map

Let Σ be smooth regular surface and $p \in \Sigma$. Given a tangent vector $v \in \mathrm{T}_p$ consider a geodesic γ_v in Σ that runs from p with the initial velocity v; that is, $\gamma(0) = p$ and $\gamma'(0) = v$.

The point $q = \gamma_v(1)$ is called *exponential map* of v, or briefly $q = \exp_p v$. By 10.9, the map $\exp_p\colon \mathrm{T}_p \to \Sigma$ is smooth and defined in a neighborhood of zero in T_p and if Σ is compete, it is defined on the whole space T_p.

Note that the Jacobian of $\exp_p$ at zero is the identity matrix. Therefore by the inverse function theorem (A.9), we get the following statement:

10.10. Proposition. *Let Σ be smooth surface and $p \in \Sigma$. Then the exponential map $\exp_p\colon \mathrm{T}_p \to \Sigma$ is a smooth regular parametrization* *of a neighborhood of p in Σ by a neighborhood of* 0 *in the tangent plane* T_p.

Moreover for any $p \in \Sigma$ there is $\varepsilon > 0$ such that for any $x \in \Sigma$ such that $|x - p|_\Sigma < \varepsilon$ the map $\exp_x \colon \mathrm{T}_x \to \Sigma$ is a smooth regular parametrization of the ε-neighborhood of x in Σ by the ε-neighborhood of zero in the tangent plane T_x.

Shortest paths and geodesics

10.11. Claim. *Let Σ be a smooth regular surface. Then any shortest path γ in Σ parameterized proportional to its length is a geodesic in Σ. In particular γ is a smooth curve.*

A partial converse to the first statement also holds: a sufficiently short arc of any geodesic is a shortest path. More precisely, given a smooth surface Σ there is a positive function ρ on Σ such that if a geodesic γ starts at $p \in \Sigma$ and has length at most $\rho(p)$ then it is a shortest path.

A geodesic might not form a shortest path, but if this is the case, then it is called *minimizing geodesic*. Note that according to the claim, any shortest path is a reparametrization of a minimizing geodesic.

This claim provides connection between intrinsic geometry of the surface and its extrinsic geometry. This connection will be important latter; in particular it will play the key role in the proof of the so called remarkable theorem.

Intrinsic means that it can be expressed in terms of measuring things inside the surface, for example length of curves or angles between the curves that lie in the surface. Extrinsic means that we have to use ambient space in order to measure it.

For instance shortest path γ is an object of intrinsic geometry of the surface Σ, while definition of geodesic is not intrinsic — it requires the second derivative γ'' which needs the ambient space. Note that there is a smooth bijection between the cylinder $z = x^2$ and the plane $z = 0$ that preserves the lengths of all curves; in other words cylinder can be *unfolded* on the plane. Such a bijection sends geodesics in the cylinder to geodesics on the plane and the other way around; however a geodesic on the cylinder might have nonvanishing second derivative while geodesics on the plane are straight lines with vanishing second derivative.

Informal sketch. The smoothness should be intuitively obvious; at least the curve should be twice differentiable otherwise it can be shortened.

Let us give an informal physical explanation why $\gamma''(t) \perp \mathrm{T}_{\gamma(t)}\Sigma$. One may think about the geodesic γ as of stable position of a stretched elastic thread that is forced to lie on a frictionless surface. Since it is frictionless, the force density $N(t)$ that keeps the geodesic γ in the surface must be therefore proportional to the normal vector to the surface at $\gamma(t)$. The tension in the thread has to be the same at all points (otherwise the thread would move back or forth and it would not be stable). The sum of tensions at the ends of small arc is roughly proportional to the angle between the tangent lines at the ends of the arc. Passing to the limit as the length of the arc goes to zero, we get that the density of this force $F(t)$ is proportional to $\gamma''(t)$. According to the second Newton's law, we have $F(t) + N(t) = 0$; which implies that $\gamma''(t)$ is perpendicular to $\mathrm{T}_{\gamma(t)}\Sigma$.

Fix a point $p \in \Sigma$. Let $\varepsilon > 0$ be as in 10.10. Assume a geodesic γ of length less than ε from p to q does not minimize the length between its endpoints. Then there is a shortest path from p to q, which becomes a geodesic if parameterized by its arc length. That is there are two geodesics from p to q of length smaller than ε. In other words there are two vectors $v, w \in \mathrm{T}_p$ such that $|v| < \varepsilon$, $|w| < \varepsilon$ and $q = \exp_p v = \exp_p w$. But according to 10.10, the exponential map is injective in ε-neighborhood of zero — a contradiction. $\square$

10.12. Exercise. *Show that two shortest paths in a smooth regular surface can not have more than one point of intersections.*

Show by example that a geodesics can cross each other arbitrary number of times.

10.13. Exercise. *Assume that a smooth regular surface Σ is mirror symmetric with respect to a plane Π. Show that no shortest path in Σ can cross Π more than once.*

Liberman's lemma

A curve of constant speed $\gamma\colon [a, b] \to \Sigma$ is called a *geodesic* if for some partition $a = t_0 < t_1 < \dots < t_n = b$, each arc $\gamma|_{[t_{i-1}, t_{i+1}]}$ is a minimizing geodesic.

The following lemma was proved by Joseph Liberman [24].

10.14. Liberman's lemma. *Assume γ is a geodesic on the graph $z = f(x, y)$ of a concave function f defined on an open subset of the plane. Consider a reparametrization $(x(t), y(t), z(t))$ of γ such that the curve $t \mapsto (x(t), y(t))$ is a unit-speed curve. Then $z(t)$ is a concave function.*

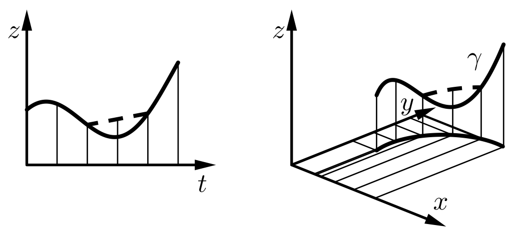
If we draw a line parallel to the z-axis thru each point of γ, we get a surface which can be developed on the plane — that is, it can be parametrized by a strip in the plane between parallel lines so that the length of all curves in the strip are preserved after the mapping. If we assume that the strip is oriented vertically on the plane then the curve becomes a graph of a function and the theorem states that this function is concave.

Proof. Denote the graph by Σ. Choose a partition such that $\gamma|_{[t_{i-1},t_{i+1}]}$ is minimizing. If the function z is convex on each interval $[t_{i-1},t_{i+1}]$, then it is convex on the entire interval. Therefore it is sufficient to prove the case when $\gamma\colon [a,b]\to$ is a minimizing geodesic.

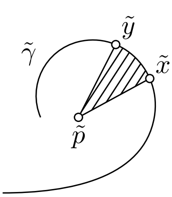
Further, passing to a finer partition, we can assume that the projection of γ to the (x,y)-plane lies completely in a closed disc Δ in the domain of definition of f; moreover the distance from the projection of γ to the boundary of the disc is much larger than the length of γ. In this case the curve lies in the boundary of a closed convex set

$$K=\left\{\,(x,y,z)\in\mathbb{R}^3\;:\;(x,y)\in\Delta,\;z\leqslant f(x,y)\,\right\};$$

so we can apply the lemma on closest point projection.



If the function z is not concave, then there is another function $\check z\geqslant z$ with shorter graph such that $\check z(a)=z(a)$, $\check z(b)=z(b)$. Consider the curve $\check\gamma(t)=(x(t),y(t),\check z(t))$; $\check\gamma$ lies higher than γ and therefore can be on the boundary or outside of K. The closest point projection of $\check\gamma$ to K gives a curve connecting the endpoints of γ, by construction it runs in Σ, and by the lemma on closest point projection it is shorter than γ — a contradiction. $\square$



10.15. Exercise. *Assume γ is a minimizing geodesic on a smooth closed convex surface Σ and p in the interior of a convex set bounded by Σ. Consider the cone C with the tip at p and with ruling the half-lines passing thru the points of γ. Let us develop C on the plane; the curve γ*

becomes a plane curve $\tilde\gamma$ and the tip of the cone is mapped to a point $\tilde p$.

Show that $\tilde\gamma$ is convex toward to $\tilde p$; that is for any sufficiently small arc $\tilde x\tilde y$ of $\tilde\gamma$, the curvelinear triangle $\tilde p\tilde x\tilde y$ is convex.

Bound on total curvature

10.16. Theorem. *Assume Σ is a graph $z = f(x, y)$ of a convex ℓ-Lipschitz function f defined on an open set in the (x, y)-plane. Then the total curvature of any geodesic in Σ is at most $2\cdot\ell$.*

The above theorem was proved by Vladimir Usov [25], later David Berg [26] pointed out that the same proof works for geodesics in closed epigraphs of ℓ-Lipschitz functions which are not necessary concave; that is, sets of the type

$$W = \left\{\, (x, y, z) \in \mathbb{R}^3 \;:\; z \geqslant f(x, y) \,\right\}$$

10.17. Lemma. *Assume $f\colon [a, b] \to \mathbb{R}$ is a smooth function. Let $(x(t), y(t))$ be a unit-speed parametrization of the graph $y = f(x)$. Then f is concave if and only if the function $t \mapsto y(t)$ is concave.*

Proof. We can assume that the function $t \mapsto x(t)$ is increasing.

Note that

$$y'(t) = \frac{f'(x(t))}{\sqrt{1 + (f'(x(t)))^2}}.$$

It follows that $y'(t)$ is nonincreasing if and only if $f'(x)$ is nonincreasing, hence the result. $\square$

Proof of 10.16. Let $\gamma(t) = (x(t), y(t), z(t))$ be a unit speed geodesic on Σ. According to Liberman's lemma and Lemma 10.17, $z(t)$ is concave.

Since the slope of f is at most ℓ, we have

$$|z'(t)| \leqslant \tfrac{\ell}{\sqrt{1+\ell^2}}.$$

If γ is defined on the interval $[a, b]$, then

$$\begin{aligned}\int\limits_a^b |z''(t)| &= z'(a) - z'(b) \leqslant \\ &\leqslant 2\cdot\tfrac{\ell}{\sqrt{1+\ell^2}}.\end{aligned}$$

Further, note that z'' is the projection of γ'' to the z-axis. Since f is ℓ-Lipschitz, the tangent plane $T_{\gamma(t)}\Sigma$ cannot have slope greater than ℓ for any t. Because γ'' is perpendicular to that plane,

$$|\gamma''(t)| \leqslant \sqrt{1+\ell^2}\cdot|z''(t)|.$$

Therefore

$$\begin{aligned}\Phi(\gamma) &= \int\limits_a^b |\gamma''(t)|\cdot dt \leqslant \\ &\leqslant \sqrt{1+\ell^2}\cdot \int\limits_a^b |z''(t)|\cdot dt \leqslant \\ &\leqslant 2\cdot\ell.\end{aligned}$$

□

10.18. Exercise. *Assume f is a convex function $\tfrac{3}{2}$-Lipschitz function defined on the (x,y)-plane. Show that any geodesic γ on the graph $z=f(x,y)$ is simple; that is, it has no self-intersections.*

Construct a convex function 2-Lipschitz function defined on the (x,y)-plane with a nonsimple geodesic γ on its graph $z=f(x,y)$.

Hint: Use the theorem and 3.15. The suggested argument does not give the optimal bound the Lipschitz constant that guarantee that γ is simple, but later we will give the exact bound.

10.19. Advanced exercise. *Suppose a smooth surface Σ bounds a convex set K in the Euclidean space. Assume K contains a unit ball and has diameter D. Find an upper bound of the total curvatures of minimizing geodesics in Σ.*

Hint: Use Exercise 10.6.

In fact there is a fixed bound on the total curvature of any minimizing geodesic on any closed convex surface [27].

Appendix A

Review

Here we state and discuss results from different branches of mathematics which were used further in the book. The reader is not expected to know proofs of these statements, but it is better to check that his intuition agrees with each.

A.1 Metric spaces

Metric is a function that returns a real value $\operatorname{dist}(x, y)$ for any pair x, y in a given nonempty set $\mathcal{X}$ and satisfies the following axioms for any triple x, y, z:

(a) Positiveness:
$$\operatorname{dist}(x, y) \geqslant 0.$$

(b) $x = y$ if and only if
$$\operatorname{dist}(x, y) = 0.$$

(c) Symmetry:
$$\operatorname{dist}(x, y) = \operatorname{dist}(y, x).$$

(d) Triangle inequality:
$$\operatorname{dist}(x, z) \leqslant \operatorname{dist}(x, y) + \operatorname{dist}(y, z).$$

A set with a metric is called *metric space* and the elements of the set are called *points*.

Shortcut for distance. Usually we consider only one metric on a set, therefore we can denote the metric space and its underlying set by the same letter, say $\mathcal{X}$. In this case we also use a shortcut notations

$|x - y|$ or $|x - y|_{\mathcal{X}}$ for the *distance* $\operatorname{dist}(x, y)$ *from* x *to* y *in* $\mathcal{X}$. For example, the triangle inequality can be written as

$$|x - z|_{\mathcal{X}} \leqslant |x - y|_{\mathcal{X}} + |y - z|_{\mathcal{X}}.$$

Examples. Euclidean space and plane as well as real line will be the most important examples of metric spaces for us. In these examples the introduced notation $|x-y|$ for the distance from x to y has perfect sense as a norm of the vector $x - y$. However, in general metric space the expression $x-y$ has no sense, but anyway we use expression $|x-y|$ for the distance.

If we say *plane* or *space* we mean *Eucledean* plane or space. However the plane (as well as the space) admits many other metrics, for example the so called Manhattan metric from the following exercise.

A.1. Exercise. *Consider the function*

$$\operatorname{dist}(p, q) = |x_p - x_q| + |y_p - y_q|,$$

where $p = (x_p, y_p)$ *and* $q = (x_q, y_q)$ *are points in the coordinate plane* $\mathbb{R}^2$. *Show that* dist *is a metric on* $\mathbb{R}^2$.

Let us mention another example: the *discrete space* — arbitrary nonempty set $\mathcal{X}$ with the metric defined as $|x - y|_{\mathcal{X}} = 0$ if $x = y$ and $|x - y|_{\mathcal{X}} = 1$ otherwise.

Subspaces. Any subset of a metric space is also a metric space, by restricting the original metric to the subset; the obtained metric space is called a *subspace*. In particular, all subsets of Euclidean space are metric spaces.

Balls. Given a point p in a metric space $\mathcal{X}$ and a real number $R \geqslant 0$, the set of points x on the distance less then R (or at most R) from p is called open (or correspondingly closed) ball of radius R with center at p. The *open ball* is denoted as $B(p, R)$ or $B(p, R)_{\mathcal{X}}$; the second notation is used if we need to emphasize that the ball lies in the metric space $\mathcal{X}$. Formally speaking

$$B(p, R) = B(p, R)_{\mathcal{X}} = \{\, x \in \mathcal{X} \,:\, |x - p|_{\mathcal{X}} < R \,\}.$$

Analogously, the *closed ball* is denoted as $\bar{B}[p, R]$ or $\bar{B}[p, R]_{\mathcal{X}}$ and

$$\bar{B}[p, R] = \bar{B}[p, R]_{\mathcal{X}} = \{\, x \in \mathcal{X} \,:\, |x - p|_{\mathcal{X}} \leqslant R \,\}.$$

A.2. Exercise. *Let* $\mathcal{X}$ *be a metric space.*

(a) Show that if $\bar{B}[p, 2] \subset \bar{B}[q, 1]$ *for some points* $p, q \in \mathcal{X}$, *then* $\bar{B}[p, 2] = \bar{B}[q, 1]$.

(b) Construct a metric space $\mathcal{X}$ *with two points* p *and* q *such that* $B(p, \tfrac{3}{2}) \subset B(q, 1)$ *and the inclusions is strict.*

Calculus

In this section we will extend standard notions from calculus to the metric spaces.

A.3. Definition. *Let $\mathcal{X}$ be a metric space. A sequence of points $x_1, x_2, \dots$ in $\mathcal{X}$ is called* convergent *if there is $x_\infty \in \mathcal{X}$ such that $|x_\infty - x_n| \to 0$ as $n \to \infty$. That is, for every $\varepsilon > 0$, there is a natural number N such that for all $n \geqslant N$, we have*

$$|x_\infty - x_n| < \varepsilon.$$

In this case we say that the sequence (x_n) converges *to x_∞, or x_∞ is the* limit *of the sequence (x_n). Notationally, we write $x_n \to x_\infty$ as $n \to \infty$ or $x_\infty = \lim_{n\to\infty} x_n$.*

A.4. Definition. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces. A map $f\colon \mathcal{X} \to \mathcal{Y}$ is called* continuous *if for any convergent sequence $x_n \to x_\infty$ in $\mathcal{X}$, we have $f(x_n) \to f(x_\infty)$ in $\mathcal{Y}$.*

Equivalently, $f\colon \mathcal{X} \to \mathcal{Y}$ is continuous if for any $x \in \mathcal{X}$ and any $\varepsilon > 0$, there is $\delta > 0$ such that

$$|x - x'|_{\mathcal{X}} < \delta \quad \textit{implies} \quad |f(x) - f(x')|_{\mathcal{Y}} < \varepsilon.$$

A.5. Exercise. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces $f\colon \mathcal{X} \to \mathcal{Y}$ is* distance non-expanding map*; that is,*

$$|f(x) - f(x')|_{\mathcal{Y}} \leqslant |x - x'|_{\mathcal{X}}$$

for any $x, x' \in \mathcal{X}$. Show that f is continuous.

A.6. Definition. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces. A continous bijection $f\colon \mathcal{X} \to \mathcal{Y}$ is called a* homeomorphism *if its inverse $f^{-1}\colon \mathcal{Y} \to \mathcal{X}$ is also continuous.*

If there exists a homeomorphism $f\colon \mathcal{X} \to \mathcal{Y}$, we say that $\mathcal{X}$ is homeomorphic *to $\mathcal{Y}$, or $\mathcal{X}$ and $\mathcal{Y}$ are* homeomorphic.

If a metric space $\mathcal{X}$ is homeomorphic to a known space, for example plane, sphere, disc, circle and so on, we may also say that $\mathcal{X}$ is a *topological* plane, sphere, disc, circle and so on.

A.7. Definition. *A subset A of a metric space $\mathcal{X}$ is called* closed *if whenever a sequence (x_n) of points from A converges in $\mathcal{X}$, we have that $\lim_{n\to\infty} x_n \in A$.*

A set $\Omega \subset \mathcal{X}$ is called open *if for any $z \in \Omega$, there is $\varepsilon > 0$ such that $B(z, \varepsilon) \subset \Omega$.*

An open set Ω that contains a given point p is called *neighborhood of* p.

A.8. Exercise. *Let Q be a subset of a metric space $\mathcal{X}$. Show that A is closed if and only if its complement $\Omega = \mathcal{X}\backslash Q$ is open.*

A.2 Multivariable calculus

A map $\boldsymbol{f}\colon \mathbb{R}^n \to \mathbb{R}^k$ can be thought as array of functions

$$f_1, \dots, f_k\colon \mathbb{R}^n \to \mathbb{R}.$$

The map $\boldsymbol{f}$ is called *smooth* if each function f_i is smooth; that is, all partial derivatives of f_i are defined in the domain of definition of $\boldsymbol{f}$.

Inverse function theorem gives a sufficient condition for a smooth function to be invertible in a neighborhood of a given point p in its domain. The condition is formulated in terms of partial derivative of f_i at p.

Implicit function theorem is a close relative to inverse function theorem; in fact it can be obtained as its corollary. It is used for instance when we need to pass from parametric to implicit description of curves and surface.

Both theorems reduce the existence of a map satisfying certain equation to a question in linear algebra. We use these two theorems only for $n \leqslant 3$.

These two theorems are discussed in any course of multivariable calculus, the classical book of Walter Rudin [28] is one of my favorites.

A.9. Inverse function theorem. *Let $\boldsymbol{f} = (f_1, \dots, f_n)\colon \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map. Assume that the Jacobian matrix*

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

is invertible at some point p in the domain of definition of $\boldsymbol{f}$. Then there is a smooth function $\boldsymbol{h}\colon \mathbb{R}^m \to \mathbb{R}^n$ defined is a neighborhood Ω_q of $q = \boldsymbol{f}(p)$ that is local inverse of $\boldsymbol{f}$ at *p; that is, there are neighborhoods $\Omega_p \ni p$ such that $\boldsymbol{f}$ defines a bijection $\Omega_p \to \Omega_q$ and $\boldsymbol{f}(x) = y$ if and only if $x = \boldsymbol{h}(y)$ for any $x \in \Omega_p$ and any $y \in \Omega_q$.*

A.10. Implicit function theorem. *Let $\boldsymbol{f} = (f_1, \dots, f_n)\colon \mathbb{R}^{n+m} \to \mathbb{R}^n$ be a smooth map, $m, n \geqslant 1$. Let us consider $\mathbb{R}^{n+m}$ as a product* *space* $\mathbb{R}^n \times \mathbb{R}^m$ *with coodinates* $x_1, \ldots, x_n, y_1, \ldots, y_m$. *Consider the following matrix*

$$M = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

formed by first n columns of the Jacobian matrix. Assume M *is invertible at some point* p *in the domain of definition of* $\boldsymbol{f}$ *and* $\boldsymbol{f}(p) = 0$. *Then there is a neighborhood* $\Omega_p \ni p$ *and smooth function* $\boldsymbol{h}\colon \mathbb{R}^m \to \mathbb{R}^n$ *defined is a neighborhood* $\Omega_0 \ni 0$ *that for any* $(x_1, \ldots, x_n, y_1, \ldots y_m) \in \Omega_p$ *the equality*

$$\boldsymbol{f}(x_1, \ldots, x_n, y_1, \ldots y_m) = 0$$

holds if and only if

$$(x_1, \ldots x_n) = \boldsymbol{h}(y_1, \ldots y_m).$$

If the assumption in the theorem holds for any point p such that $\boldsymbol{f}(p) = 0$, then we say that 0 is a regular value of $\boldsymbol{f}$. The following lemma states that most of the values of smooth map are regular; in particular generic smooth function satisfies the assumption of the theorem.

A.11. Sard's lemma. *Almost all values of a smooth map* $f\colon U \to \to \mathbb{R}^m$ *defined on an open set* $U \subset \mathbb{R}^n$ *are regular.*

The words *almost all* means that with exception o set of zero Lebesgue measure. In particular if one chooses a random value equidistributed in arbitrary small ball $B \subset \mathbb{R}^m$ then it is a regular value of f with probability 1.

A.3 Initial value problem

The following theorem guarantees existence and uniqueness of a solution of an initial value problem for a system of ordinary differential equations

$$\begin{cases} x_1'(t) &= f_1(x_1, \ldots, x_n, t), \\ & \cdots \\ x_n'(t) &= f_n(x_1, \ldots, x_n, t), \end{cases}$$

where each $x_i = x_i(t)$ is a real valued function defined on a real interval $\mathbb{I}$ and each f_i is a smooth function defined on $\mathbb{R}^n \times \mathbb{I}$.

The array functions $(f_1, \dots, f_n)$ can be considered as one vector-valued function $\boldsymbol{f}\colon \mathbb{R}^n \times \mathbb{I} \to \mathbb{R}^n$ and the array $(x_1, \dots, x_n)$ can be considered as a vector $\boldsymbol{x} \in \mathbb{R}^n$. Therefore the system can be rewritten as one vector equation

$$\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}, t).$$

A.12. Theorem. *Suppose $\mathbb{I}$ is a real interval and $\boldsymbol{f}\colon \mathbb{R}^n \times \mathbb{I} \to \mathbb{R}^n$ is a smooth function. Then for any initial data $\boldsymbol{x}(t_0) = \boldsymbol{u}$ the differential equation*

$$\boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}, t)$$

has a unique solution $\boldsymbol{x}(t)$ defined at a maximal subinterval $\mathbb{J}$ of $\mathbb{I}$ that contains t_0. Moreover

(a) if $\mathbb{J} \ne \mathbb{I}$, that is, if an end a of $\mathbb{J}$ lies in the interior of $\mathbb{I}$, then $\boldsymbol{x}(t)$ diverges for $t \to a$;

(b) the function $(\boldsymbol{u}, t_0, t) \mapsto \boldsymbol{x}(t)$ is smooth.

A.4 Real analysis

Recall that a function f is called Lipschitz if there is a constant L such that

$$|f(x) - f(y)| \leqslant L \cdot |x - y|$$

for values x and y in the domain of definition of f. This definition works for maps between metric spaces, but we will use it for real-to-real functions only.

A.13. Rademacher's theorem. *Let $f\colon [a, b] \to \mathbb{R}$ be a Lipschitz function. then derivative $f'(x)$ is defined for alomst all $x \in [a, b]$. Moreover the derivative f' is a bounded measurable function defined almost everywhere in $[a, b]$ and it satisfies the fundamental theorem of calculus; that is, the following identity*

$$f(b) - f(a) = \int_a^b f'(x) \cdot dx,$$

holds if the integral understood in the sense of Lebesgue.

It is often helps to work with measurable functions; it makes possible to extend many statements about continuous function to measurable functions.

A.14. Lusin's theorem. *Let $\varphi\colon [a, b] \to \mathbb{R}$ be a measurable function. Then for any $\varepsilon > 0$, there is a continuous function $\psi_\varepsilon\colon [a, b] \to \mathbb{R}$ that coincides with φ outside of a set of measure at most ε. Moreover, φ is bounded above and/or below by some constants then we can assume that so is ψ_ε.*

A.5 Topology

We sometimes characterization of homeomorphism.

A.15. Theorem. *A continuous bijection f between compact metric spaces has continuous inverse; that is f is a homeomorphism.*

The first part of the following theorem is proved by Camille Jordan, the second part is due to Arthur Schoenflies.

A.16. Theorem. *The complement of any closed simple plane curve γ has exactly two connected components.*

Moreover the there is a homeomorphism $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ that maps the unit circle to γ. In particular γ bounds a topological disc.

This theorem is known for simple formulation and quite hard proof. By now many proofs of this theorem are known. For the first statement, a very short proof based on somewhat developed technique is given by Patrick Doyle [29], among elementary proofs, one of my favorites is the proof given by Aleksei Filippov [30].

We use the following smooth analog of this theorem.

A.17. Theorem. *The complement of any closed simple smooth regular plane curve γ has exactly two connected components.*

Moreover the there is a diffeomorphism $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ that maps the unit circle to γ.

The proof of this statement is much simpler. An amusing proof of can be found in [31].

A.6 Convexity

A set X in the Euclidean space is called *convex* if for any two points $x, y \in X$, any point z between x and y lies in X. It is called *strictly convex* if for any two points $x, y \in X$, any point z between x and y lies in the interior of X.

From the definition, it is easy to see that intersection of arbitrary family of convex sets is convex. The intersection of all convex sets containing X is called *convex hull* of X; it is the minimal convex set containing the given set X.

A.18. Lemma. *Let $K \subset \mathbb{R}^3$ be a closed convex subset. Then for any point $p \notin K$ there is a plane Π that separates K from p; that is, K and p lie on the opposite open half-spaces of Π.*

A function of two variables $(x, y) \mapsto f(x, y)$ is called convex if its epigraph $z \geqslant f(x, y)$ is a convex set. This is equivalent to the so called

Jensen's inequality

$$f\,(t{\cdot}x_1 + (1-t){\cdot}x_2) \le t{\cdot}f(x_1) + (1-t){\cdot}f(x_2)$$

for $t \in [0,1]$. If f is smooth, then the condition is equivalent to the following inequality for second directional derivative:

$$D_w^2 f \geqslant 0$$

for any vector $w \neq 0$ in the (x,y)-plane.

A.7 Elementary geometry

A.19. Theorem. *The sum of sum of all the internal angles of a simple n-gon is $(n-2){\cdot}\pi$.*

Proof. The proof is by induction on n. For $n = 3$ it says that sum of internal angles of a triangle is π, which is assumed to be known.

First let us show that for any $n \geqslant 4$, any n-gon has a diagonal that lies inside of it. Assume this is holds true for all polygons with at most $n-1$ vertex.

Fix an n-gon P, $n \geqslant 4$. Applying rotation if necessary, we can assume that all its vertexes have different x-coordinates. Let v be a vertex of P that minimizes the x-coordinate; denote by u and w its adjacent vertexes. Let us choose the diagonal uw if it lies in P. Otherwise the triangle $\triangle uvw$ contains another vertex of P. Choose a vertex s in the interior of $\triangle uvw$ that maximizes the distance to line uw. Note that the diagonal vs lyes in P; if it is not the case then vs crosses another side pq of P, one of the vertexes p or q has larger distance to the line and it lies in the interior of $\triangle uvw$ — a contradiction.

Note that the diagonal divides P into two polygons, say Q and R, with smaller number of sides in each, say k and m correspondingly. Note that

❶ $$k + m = n + 2;$$

indeed each side of P appears once as a side of P or Q plus the diagonal appears twice — once as a side in Q and once as a side of R. Note that the sum of angles of P is the sum of angles of Q and R, which by the induction hypothesis are $(k-2){\cdot}\pi$ and $(m-2){\cdot}\pi$ correspondingly. It remains to note that ❶ implies

$$(k-2){\cdot}\pi + (m-2){\cdot}\pi = (n-2){\cdot}\pi.$$

□

The following theorem says that triangle inequality holds for angles between half-lines from a fixed point. In particular it implies that a unit sphere with angle metric is a metric space.

A.20. Theorem. *The inequality*

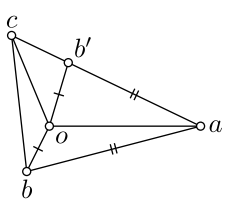
$$\measuredangle aob + \measuredangle boc \geqslant \measuredangle aoc$$

holds for any three half-lines oa, ob and oc in the Euclidean space.

The following lemma says that angle of a triangle monotonically depends on the opposit side, assuming the we keep the remaining two sides fixed. It is a simple statement in elementary geometry; in particular it follows directly from the cosine rule.

A.21. Lemma. *Let x, y, z, x', y' and z' be 6 points such that $|x-y| = |x'-y'| > 0$ and $|y-z| = |y'-z'| > 0$. Then*

$$\measuredangle xyz \geqslant \measuredangle x'y'z' \quad \textit{if and only if} \quad |x-z| \geqslant |x'-y'|.$$



Proof of A.20. We can assume that $\measuredangle aob < \measuredangle aoc$; otherwise the statement is evident. In this case there is a half-line ob' in the angle aoc such that

$$\measuredangle aob = \measuredangle aob',$$

so in particular we have that

$$\measuredangle aob' + \measuredangle b'oc = \measuredangle aoc.$$

Without loss of generality we can assume that $|o-b| = |o-b'|$ and b' lies on a line segment ac, so

$$|a-b'| + |b'-c| = |a-c|.$$

Then by triangle inequality

❷
$$\begin{aligned} |a-b| + |b-c| \geqslant |a-c| &= \\ &= |a-b'| + |b'-c|. \end{aligned}$$

Note that in the triangles aob and aob' the side ao is shared, $\measuredangle aob = \measuredangle aob'$ and $|o-b| = |o-b'|$. By side-angle-side congruence condition, we have that $\triangle aob \cong \triangle aob'$; in particular $|a-b'| = |a-b|$. Therefore from ❷ we have that

$$|b-c| \geqslant |b'-c|.$$

Applying the angle monotonicity (A.21) we get that

$$\measuredangle boc \geqslant \measuredangle b'oc.$$

Whence

$$\begin{aligned}\measuredangle aob + \measuredangle boc \geqslant \measuredangle aob' + \measuredangle b'oc =\\ = \measuredangle aoc.\end{aligned}$$

□

Appendix B

Homework assignments

HWA-01. Exercises: A.2, 1.4, 1.5, 1.7, 2.14.

HWA-02. Exercises: 2.4(b), 2.5, 2.9, 2.15, 2.20.

HWA-03. Exercises: 2.16, 2.18, 2.19, 3.3a, 3.5.

HWA-04. Exercises: 3.2, 3.4, 3.15, 3.17 + 1.12.17 in the Toponogov's book.

HWA-05. Exercises: 3.8, 3.12, 3.13, 4.6, 5.5.

HWA-06. Exercises: 3.20, 5.12, 5.13, 5.14, 8.13.

MIDTERM. 5 problems, 20 points each:

- ⋄ 2 theorems from the following list; in the brackets I give the corresponding statement in the Toponogov's book, if it exists.
  2.7, 2.13, 2.17 (Prob. 1.10.4),
  3.1, 3.7 (Thm. 1.10.1), 3.10, 3.11 (Lem. 1.11.2), 3.14, 3.16, 3.18 (Prob. 1.10.6), 4.10 (Thm. 1.9.1),
  5.4 (Prob. 1.7.7), 5.9, 5.15 (Prob. 1.7.1), 5.19, ~~5.23~~.
- ⋄ 2 problems from HWA's;
- ⋄ 1 new problem.

HWA-07. Exercises: 3.19, 4.9(a+b), 5.17, 5.21(a+b), 5.26.

HWA-08. Exercises: 4.9(finish), 5.28, 6.3, 6.4, 6.6.

HWA-09. Exercises: 6.8, 6.9, 7.2, 7.4, 7.9.

HWA-10. Exercises: 7.10, 7.13, 7.14, 8.11, 8.14.

HWA-11. Exercises: 7.3, 8.6, 8.7, 9.1, 9.6 + try to think about this 10.1.

HWA-12. Exercises: 9.7, 10.5, 10.6, 10.7, 10.12.

Appendix C

Semisolutions

Exercise 5.14

Using the spiral lemma. If γ does not have a vertex at s then $k'(s) \neq 0$ and therefore the curvature of a small arc around s is monotonic. By spiral lemma the osculating circles at this arc are nested. In particular the curve γ crosses the osculating circle σ_s at s; that is, σ_s is not a local support of γ at s.

We proved that *if γ does not have a vertex at s then the osculating circle σ_s is not supporting at s.* The latter is equivalent to the requred statement: *if the osculating circle σ_s supports γ at s, then γ has a vertex at s.*

By direct calculations. Assume the osculating circline σ_s is a circle. Then its center is $p = \gamma(s) + \frac{1}{k(s)} \cdot \nu(s)$. Since σ_s is supporting γ at s, we have that the function

$$f(t) = \langle p - \gamma(t), p - \gamma(t) \rangle$$

has a minimum or maximum at s.

Note that

$$\begin{aligned}
f'(t) &= 2 \cdot \langle p - \gamma(t), -\tau(t) \rangle; \\
f''(s) &= 2 \cdot \langle -\tau(t), -\tau(t) \rangle - 2 \cdot \langle p - \gamma(t), k(t) \cdot \nu(t) \rangle = \\
&= 2 - 2 \cdot \langle p - \gamma(t), k(t) \cdot \nu(t) \rangle; \\
f'''(s) &= -2 \cdot \langle -\tau(t), k(t) \cdot \nu(t) \rangle - \\
&\quad - 2 \cdot \langle p - \gamma(t), k'(t) \cdot \nu(t) \rangle + 2 \cdot \langle p - \gamma(t), -k^2(t) \cdot \tau(t) \rangle = \\
&= -2 \cdot \langle p - \gamma(t), k'(t) \cdot \nu(t) \rangle + 2 \cdot \langle p - \gamma(t), -k^2(t) \cdot \tau(t) \rangle.
\end{aligned}$$

Therefore

$$\begin{aligned} f'(s) &= -2\cdot\langle \tfrac{1}{k(s)}\cdot\nu(s), \tau(s)\rangle = \\ &= 0. \\ f''(s) &= 2 - 2\cdot\tfrac{k(s)}{k(s)} = \\ &= 0. \\ f'''(s) &= -2\cdot\tfrac{k'(s)}{k(s)}. \end{aligned}$$

Therefore if f has a local minimum or maximum at s, then $f'''(s) = 0$ and therefore $k'(s) = 0$.

Bibliography

[1] E. Denne. *Alternating quadrisecants of knots*. Thesis (Ph.D.)–University of Illinois at Urbana-Champaign. ProQuest LLC, Ann Arbor, MI, 2004, p. 119.

[2] G. D. Chakerian. "An inequality for closed space curves". *Pacific J. Math.* 12 (1962), pp. 53–57.

[3] István Fáry. "Sur certaines inégalites géométriques". *Acta Sci. Math. Szeged* 12.Leopoldo Fejér et Frederico Riesz LXX annos natis dedicatus, Pars A (1950), pp. 117–124.

[4] G. D. Chakerian. "An inequality for closed space curves". *Pacific J. Math.* 12 (1962), pp. 53–57.

[5] Serge Tabachnikov. "The tale of a geometric inequality". *MASS selecta.* Amer. Math. Soc., Providence, RI, 2003, pp. 257–262.

[6] J. Lagarias and T. Richardson. "Convexity and the average curvature of plane curves". *Geom. Dedicata* 67.1 (1997), pp. 1–30.

[7] A. I. Nazarov and F. V. Petrov. "On a conjecture of S. L. Tabachnikov". *Algebra i Analiz* 19.1 (2007), pp. 177–193.

[8] John M. Sullivan. "Curves of finite total curvature". *Discrete differential geometry.* Vol. 38. Oberwolfach Semin. 2008, pp. 137–161.

[9] Michel Ange Lancret. "Mémoire sur les courbes à double courbure". *Mémoires présentés à l'Institut des Sciences, Lettres et Arts, par divers savants, et lus dans ses assemblées. Sciences mathématiques et physiques.* 1 (1802), pp. 416–454.

[10] R. Foote, M. Levi, and S. Tabachnikov. "Tractrices, bicycle tire tracks, hatchet planimeters, and a 100-year-old conjecture". *Amer. Math. Monthly* 120.3 (2013), pp. 199–216.

[11] Heinz Hopf. "Über die Drehung der Tangenten und Sehnen ebener Kurven". *Compositio Math.* 2 (1935), pp. 50–62.

[12] Heinz Hopf. *Differential geometry in the large.* Second. Vol. 1000. Lecture Notes in Mathematics. Notes taken by Peter Lax and John W. Gray, With a preface by S. S. Chern, With a preface by K. Voss. Springer-Verlag, Berlin, 1989, pp. viii+184.

[13] P. G. Tait. "Note on the circles of curvature of a plane curve." *Proc. Edinb. Math. Soc.* 14 (1896), p. 26.

[14] A. Kneser. „Bemerkungen über die Anzahl der Extreme der Krümmung auf geschlossenen Kurven und über verwandte Fragen in einer nichteuklidischen Geometrie." *Heinrich Weber Festschrift.* 1912.

[15] D. Fuchs and S. Tabachnikov. *Mathematical omnibus.* Thirty lectures on classic mathematics. American Mathematical Society, Providence, RI, 2007.

[16] G. Pestov and V. Ionin. "On the largest possible circle imbedded in a given closed curve". *Dokl. Akad. Nauk SSSR* 127 (1959), pp. 1170–1172.

[17] W. Blaschke. *Kreis und Kugel.* Verlag von Veit & Comp., Leipzig, 1916.

[18] S. Mukhopadhyaya. "New methods in the geometry of a plane arc". *Bull Calcutta Math Soc* 1 (1909), pp. 31–37.

[19] R. Osserman. "The four-or-more vertex theorem". *Amer. Math. Monthly* 92.5 (1985), pp. 332–337.

[20] J. B. Meusnier. "Mémoire sur la courbure des surfaces". *Mem des savan etrangers* 10.1776 (1785), pp. 477–510.

[21] S. N. Bernstein. "Sur un théoréme de géométrie et son application aux équations aux dérivées partielles du type elliptique". *Comm. de la Soc. Math de Kharkov (2eme ser.), 15 (1915–17), 38–45.* 26 (1927). See also: Über ein geometrisches Theorem und seine Anwendung auf die partiellen Differential gleichungen vom elliptischen Typus, Math. Zeit, 26, pp. 551–558.

[22] ().

[23] S. Alexander, V. Kapovitch, and A. Petrunin. *An invitation to Alexandrov geometry: CAT (0) spaces.*

[24] J. Liberman. "Geodesic lines on convex surfaces". *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 32 (1941), pp. 310–313.

[25] V. V. Usov. "The length of the spherical image of a geodesic on a convex surface". *Sibirsk. Mat. Ž.* 17.1 (1976), pp. 233–236.

[26] I. D. Berg. "An estimate on the total curvature of a geodesic in Euclidean 3-space-with-boundary." *Geom. Dedicata* 13 (1982), pp. 1–6.

[27] N. Lebedeva and A. Petrunin. "On the total curvature of minimizing geodesics on convex surfaces". *Algebra i Analiz* 29.1 (2017). Reprinted in St. Petersburg Math. J. **2**9 (2018), no. 1, 139–153, pp. 189–208.

[28] Walter Rudin. *Principles of mathematical analysis.* 1976.

[29] P. H. Doyle. "Plane separation". *Proc. Cambridge Philos. Soc.* 64 (1968), p. 291.

[30] A. F. Filippov. "An elementary proof of Jordan's theorem". *Uspehi Matem. Nauk (N.S.)* 5.5(39) (1950), pp. 173–176.

[31] Gregory R. Chambers and Yevgeny Liokumovich. "Converting homotopies to isotopies and dividing homotopies in half in an effective way". *Geom. Funct. Anal.* 24.4 (2014), pp. 1080–1100.