Invitation to comparison geometry

Anton Petrunin and Sergio Zamora Barrera

Contents

1 Preliminaries **5**
1.1 Metric spaces 5
1.2 Intrinsic metric 5
1.3 Rademacher's theorem 5
1.4 Inverse function theorem 5
1.5 Picard theorem 5

2 Length **6**
2.1 Length of curve 6
2.2 Semicontinuity of length 8
2.3 Axioms of length 9
2.4 Crofton formula 12
2.5 Applications 14

3 Total curvature **17**
3.1 Smooth regular curves 17
3.2 General definition 18
3.3 Crofton again 21
3.4 DNA inequality 21
3.5 Curves of finite total curvature 23
3.6 Total signed curvature 24

4 Fáry–Milnor theorem **26**
4.1 Tame knots 26
4.2 Fáry–Milnor theorem 27
4.3 Milnor's proof 28
4.4 Fáry's proof 29
4.5 Proof of Alexander and Bishop 31

5 Osculating circlines **34**
5.1 Acceleration of unit-speed curve 34
5.2 Signed curvature 35

5.3 Osculating circline 35
5.4 Spiral theorem 36

6 Supporting circlines **38**
6.1 Definitions 38
6.2 Supporting test 39
6.3 Lens lemma 40
6.4 Convexity and inflection points 41
6.5 Four-vertex theorem 43
6.6 Moon in a puddle 46

7 Surfaces **48**
7.1 Embedded surfaces 48
7.2 Tangent plane 49
7.3 Curvatures 50
7.4 Immersed surfaces 51

8 Bounded principle curvatures **53**
8.1 Lagunov's example 54
8.2 On embedded sphere 56

9 Convex surfaces **57**
9.1 Embedded surfaces 57
9.2 Immersed surfaces 58
9.3 Morse-type proof 59
9.4 Proof via equidistant surfaces 61

10 Geodesics **64**
10.1 Closest point projection 64
10.2 Geodesics 65
10.3 Liberman's lemma 66
10.4 Bound on total curvature 67

11 Parallel transport **71**
11.1 Parallel fields 71
11.2 Parallel transport 71
11.3 Geodesic curvature 72

12 Gauss–Bonnet formula **76**
12.1 Signed area on the sphere 76
12.2 Gauss–Bonnet formula 78
12.3 The remarkable theorem 83
12.4 Simple geodesic 85

13 Local comparison **88**
13.1 First variation formula 88
13.2 Exponential map 90
13.3 Polar coordinates 90
13.4 Local comparison 93

14 Global comparison **94**
14.1 Formulation 94
14.2 Names and history 96
14.3 Local part 96
14.4 Alexandrov's lemma 97
14.5 Reformulations of comparison 98
14.6 Nonnegative curvature 100
14.7 Inheritance lemma 103
14.8 Nonpositive curvature 105

15 Saddle surfaces **107**

A Semisolutions **108**

Bibliography **111**

Chapter 1

Preliminaries

In this chapter we discuss results which will be used further in the sequel. The reader is not expected to know proofs of these statements, but it is better to check that his intuition agrees with each.

1.1 Metric spaces

1.2 Intrinsic metric

1.3 Rademacher's theorem

1.4 Inverse function theorem

1.5 Picard theorem

Chapter 2

Length

The material of this and the following chapters overlaps largely with [2, Chapter 5].

2.1 Length of curve

2.1. Definition. *Consider a* plane curve $\alpha\colon [a,b] \to \mathbb{R}^2$*; a continuous mapping from the real interval* $[a,b]$ *to the Euclidean plane* $\mathbb{R}^2$*.*

If $\alpha(a) = p$ *and* $\alpha(b) = q$*, we say that* α *is a* curve from p to q*.*

A curve $\alpha\colon [a,b] \to \mathbb{R}^2$ *is called* closed *if* $\alpha(a) = \alpha(b)$*.*

A curve α *is called* simple *if it is described by an injective map; that is* $\alpha(t) = \alpha(t')$ *if and only if* $t = t'$*. However, a closed curve* $\alpha\colon [a,b] \to \mathbb{R}^2$ *is called simple if it is injective everywhere except at the ends; that is, if* $\alpha(t) = \alpha(t')$ *for* $t < t'$ *then* $t = a$ *and* $t' = b$*.*

2.2. Advanced exercise. *Let* $\alpha\colon [0,1] \to \mathbb{R}^2$ *from* p *to* q*. Assume* $p \neq q$ *Show that there is a simple curve* $\beta\colon [0,1] \to \mathbb{R}^2$ *from* p *to* q *that runs in the image of* α*; that is for any* $t \in [0,1]$ *there is* $t' \in [0,1]$ *such that* $\beta(t) = \alpha(t')$*.*

Recall that a sequence

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

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

2.3. Definition. *Let* $\alpha\colon [a,b] \to \mathbb{R}^2$ *be a curve. The* length *of* α *is defined as*

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

where the exact upper bound is taken over all partitions

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

Note that $\operatorname{length}\alpha \in [0, \infty]$*; the curve* α *is called* rectifiable *if its length is finite.*

Informally, one could say that the length of a curve is the exact upper bound of the lengths of polygonal lines *inscribed* in the curve.

2.4. Exercise. *Assume* $\alpha\colon [a, b] \to \mathbb{R}^2$ *is a smooth curve, in particular the velocity vector* $\alpha'(t)$ *is defined and depends continuously on* t*. Show that*

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

2.5. Exercise. *Construct a nonrectifiable curve* $\alpha\colon [0, 1] \to \mathbb{R}^2$*.*

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

2.6. Proposition. *Assume a convex figure* A *bounded by a curve* α *lies inside a figure* B *bounded by a 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.

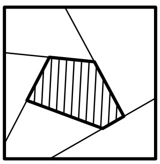
Therefore it is sufficient to prove the following lemma.

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 curve is rectifiable.*

Proof. Any closed curve is bounded; that is, it lies in a sufficiently large square.

By Proposition 2.6, the length of the curve can not exceed the perimeter of the square, hence the result. □

2.2 Semicontinuity of length

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

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

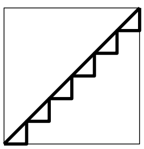
It is defined as the lowest partial limit; that is, the lowest possible limit of a subsequence of (x_n). The lower limit is defined for any sequence of real numbers and it lies in the exteded real line $[-\infty, \infty]$

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

More precisely, assume that a sequence of curves $\alpha_n \colon [a,b] \to \mathbb{R}^2$ converges pointwise to a curve $\alpha_\infty \colon [a,b] \to \mathbb{R}^2$; that is, $\alpha_n(t) \to \alpha_\infty(t)$ for any fixed $t \in [a,b]$ as $n \to \infty$. Then

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

Note that the inequality ❶ 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 $\varepsilon > 0$ and choose a partition $a = t_0 < t_1 < \dots < t_k = b$ such that

$$\operatorname{length}\alpha_\infty < |\alpha_\infty(t_0) - \alpha_\infty(t_1)| + \dots + |\alpha_\infty(t_{k-1}) - \alpha_\infty(t_k)| + \varepsilon.$$

Set

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

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

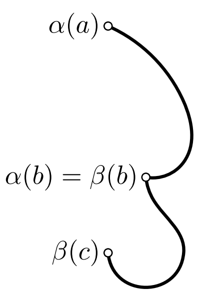
$$\varliminf_{n\to\infty} \operatorname{length}\alpha_n \geqslant \operatorname{length}\alpha_\infty - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we get ❶. □

2.3 Axioms of length

Concatenation. Assume $\alpha\colon [a,b] \to \mathbb{R}^2$ and $\beta\colon [b,c] \to \mathbb{R}^2$ are two curves such that $\alpha(b) = \beta(b)$. Then one can combine these two curves into one $\gamma\colon [a,c] \to \mathbb{R}^2$ by the rule $\gamma(t) = \alpha(t)$ for $t \leqslant b$ and $\gamma(t) = \beta(t)$ for $t \geqslant b$. The obtained curve γ is called the *concatenation* of α and β and is denoted as $\gamma = \alpha * \beta$.

Note that



$$\operatorname{length}(\alpha * \beta) = \operatorname{length}\alpha + \operatorname{length}\beta$$

for any two curves α and β such that the concatenation $\alpha * \beta$ is defined.

Reparametrization. Assume $\alpha\colon [a,b] \to \mathbb{R}^2$ is a curve and $\tau\colon [c,d] \to [a,b]$ is a continuous strictly monotonic onto map. Consider the curve $\alpha'\colon [c,d] \to \mathbb{R}^2$ defined by $\alpha' = \alpha \circ \tau$. The curve α' is called a *reparametrization* of α.

Note that

$$\operatorname{length}\alpha' = \operatorname{length}\alpha$$

whenever α' is a reparametrization of α.

2.10. Proposition. *Let ℓ be a functional that returns a value in $[0,\infty]$ for any curve $\alpha\colon [a,b] \to \mathbb{R}$.*

Assume it satisfies the following properties:

(i) *(Normalization) If* $\alpha\colon [a,b] \to \mathbb{R}^2$ *is a linear curve,*[1] *then*

$$\ell(\alpha) = |\alpha(a) - \alpha(b)|.$$

(ii) *(Additivity) If the concatenation* $\alpha * \beta$ *is defined, then*

$$\ell(\alpha * \beta) = \ell(\alpha) + \ell(\beta).$$

(iii) *(Motion invariance) The functional* ℓ *is invariant with respect to motions of the plane; that is, if* m *is an isometry of the plane, then*

$$\ell(m \circ \alpha) = \ell(\alpha)$$

for any curve α.

(iv) *(Reparametrization invariance) If* α' *is a reparametrization of a curve* α *then*

$$\ell(\alpha') = \ell(\alpha).$$

(In fact linear reparametrizations will be sufficient.)

(v) *(Semi-continuity) If a sequence of curves* $\alpha_n\colon [a,b] \to \mathbb{R}^2$ *converges pointwise to a curve to a curve* $\alpha_\infty\colon [a,b] \to \mathbb{R}^2$, *then*

$$\varliminf_{n\to\infty} \ell(\alpha_n) \geqslant \ell(\alpha_\infty).$$

Then

➊ $$\ell(\alpha) = \operatorname{length}\alpha$$

for any plane curve α.

Proof. Note that from normalization and additivity, the identity

➋ $$\ell(\beta) = \operatorname{length}\beta$$

holds for any polygonal line β that is linear on each edge.

Note that the following two inequalities

➌ $$\ell(\alpha) \leqslant \operatorname{length}\alpha$$

➍ $$\ell(\alpha) \geqslant \operatorname{length}\alpha$$

imply ➊; we will prove them separately.

Fix a curve $\alpha\colon [a,b] \to \mathbb{R}^2$ and a partition $a = t_0 < t_1 < \ldots < {} < t_k = b$. Consider the curve $\beta\colon [a,b] \to \mathbb{R}^2$ defined as the linear

[1] That is $\alpha = w + v\cdot t$ for some vectors w and v.

segment from $\alpha(t_i)$ to $\alpha(t_{i+1})$ on each interval $t \in [t_i, t_j]$. By the definition of length,

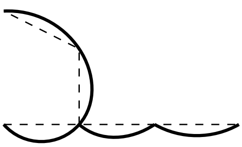
$$\operatorname{length}\beta \leqslant \operatorname{length}\alpha.$$

Since the map $\alpha\colon [a,b] \to \mathbb{R}^2$ is continuous, one can find a sequence of partitions of $[a,b]$ such that the corresponding curves β_n converge to α pointwise. Applying the semi-continuity of ℓ, ❷ and the definition of length, we get that

$$\begin{aligned}\ell(\alpha) \leqslant \varliminf_{n\to\infty} \ell(\beta_n) &= \\ = \varliminf_{n\to\infty} \operatorname{length}\beta_n &\leqslant \\ \leqslant \operatorname{length}\alpha.&\end{aligned}$$

Hence ❸ follows.

Note that a curve $\alpha\colon [a,b] \to \mathbb{R}^2$ with a partition $a = t_0 < t_1 < \ldots < t_k = b$ can be considered as a concatenation



$$\alpha = \alpha_1 * \alpha_2 * \cdots * \alpha_k$$

where α_i is the restriction of α to $[t_{i-1}, t_i]$.

Observe that there is a sequence of motions m_i of the plane so that

$$m_i \circ \alpha(t_i) = m_{i+1} \circ \alpha(t_i)$$

for any i and the points

$$m_1 \circ \alpha(t_0), m_1 \circ \alpha(t_1), \ldots m_k \circ \alpha(t_k)$$

lie in that order on a single line. For the concatenation

$$\gamma = (m_1 \circ \alpha_1) * (m_2 \circ \alpha_2) * \cdots * (m_k \circ \alpha_k)$$

we have

$$\ell(\gamma) = \ell(\alpha).$$

Assume α is rectifiable. In this case we can find a sequence of partitions of $[a,b]$ such that reparametrizations of γ_n converge to a linear segment γ'_∞; denote these reparametrizations by γ'_n. Also, $\operatorname{length}\gamma'_\infty = \operatorname{length}\alpha$; indeed, since γ'_∞ is linear,

$$\begin{aligned}\operatorname{length}\gamma'_\infty &= |\gamma'_\infty(a) - \gamma'_\infty(b)| = \\ &= \lim_{n\to\infty} \Sigma_n = \\ &= \operatorname{length}\alpha.\end{aligned}$$

where Σ_n is the sum in the definition of length for the n-th partition. Hence it is sufficient to choose a sequence of partitions such that $\Sigma_n \to$ $\to \operatorname{length}\alpha$.

Applying additivity, invariance of ℓ with respact to motions and reparametizations, we get that

$$\begin{aligned}\ell(\alpha) &= \lim_{n\to\infty} \ell(\gamma_n) = \\ &= \lim_{n\to\infty} \ell(\gamma'_n) \geqslant \\ &\geqslant \ell(\gamma'_\infty) = \\ &= \operatorname{length}\alpha.\end{aligned}$$

Hence ❹ follows.

If α is not rectifiable, a similar constriction produce an approximation of an arbitrary long line segment. (We need to run zig-zag to reduce the distance $|\gamma'_\infty(a) - \gamma'_\infty(b)|$.) It follows that

$$\ell(\alpha) \geqslant |\gamma'_\infty(a) - \gamma'_\infty(b)|.$$

Since $|\gamma'_\infty(a) - \gamma'_\infty(b)|$ can take arbitrary large value, we get $\ell(\alpha) = \infty$. □

2.11. Exercise. *Construct a functional ℓ that satisfies all the conditions in Proposition 2.10 except the semi-continuity.*

2.4 Crofton formula

Let α be a plane curve and u a unit vector. Denote by α_u the orthogonal projection of α to a line ℓ in the direction of u; that is, $\alpha_u(t) \in \ell$ and $\alpha(t) - \alpha_u(t) \perp \ell$ for any t.

2.12. Crofton formula. *The length of any plane curve α is proportional to the average of the lengths of its projections α_u for all unit vectors u. Moreover for any plane curve α we have*

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

where $\overline{\operatorname{length}\alpha_u}$ *denotes the average value of* $\operatorname{length}\alpha_u$.

Proof. First let us show that the formula

❶ $$\operatorname{length}\alpha = k\cdot\overline{\operatorname{length}\alpha_u},$$

holds for some fixed coefficient k. It will follow once we show that both sides of formula satisfy the length axioms in 2.10.

The normalization can be achieved by adjusting k.

The semi-continuity of the right hand side follows since length α_u is semi-continuous and therefore the average has to be semi-continuous.

It is straightforward to check the remaining properties.

It remains to find k. Let us apply the formula ❶ to the unit circle. The circle has length $2{\cdot}\pi$ and its projection to any line has length 4 — it is a segment of length 2 traveled back and forth. Evidently the average value is also 4, so

$$2{\cdot}\pi = k{\cdot}4,$$

hence $k = \frac{\pi}{2}$. □

Reformulation via number of intersections. Given a unit vector u and a real number ρ, consider the line of vectors w on the plane satisfying the equation

$$\langle u, w\rangle = \rho,$$

where $\langle u, w\rangle$ denotes the scalar product. Any line on the plane admits exactly two such presentations with pairs (u, ρ) and $(-u, -\rho)$. A pair (u, ρ) describes uniquely an *oriented* line — that is a line with a chosen unit normal vector.

Fix a unit vector u_0 and denote by $u(\varphi)$ the result of rotating u_0 counterclockwise by the angle φ. Denote by $\ell(\varphi, \rho)$ the oriented line asociated to the pair $(u_0(\varphi), \rho)$. To describe alllines, we need all pairs $(\varphi, \rho) \in (-\pi, \pi] \times \mathbb{R}$.

For a curve α, set $n_\alpha(\varphi, \rho)$ to be the number of parameter values t such that $\alpha(t)$ lies on the line $\ell(\varphi, \rho)$. The value $n_\alpha(\varphi, \rho)$ is a non-negative integer or ∞. Note that if α is a simple curve, then $n_\alpha(\ell)$ is the number of intersections of α with ℓ.

2.13. Another Crofton formula. *For any curve α,*

$$\operatorname{length}\alpha = \tfrac{1}{4}{\cdot}\iint\limits_{(-\pi,\pi]\times\mathbb{R}} n_\alpha(\rho, \varphi){\cdot}d\rho{\cdot}d\varphi.$$

the integral is to be understood in the sense of Lebesgue.

By definition of average value,

$$\overline{\operatorname{length}\alpha_u} = \frac{1}{2{\cdot}\pi}{\cdot}\int\limits_{-\pi}^{\pi} \operatorname{length}\alpha_{u(\varphi)}{\cdot}d\varphi.$$

Therefore the proof of this reformulation of the Crofton follows from the following observation.

2.14. Observation. *If $u = u(\varphi)$, then*

$$\operatorname{length}\alpha_u = \int\limits_{\mathbb{R}} n_\alpha(\rho,\varphi)\cdot d\rho;$$

The proof is straightforward for those who understand Lebesgue integral.

Variations. The same argument can be used to derive other formulas of the same type. For example.

Recall that a big circle in a sphere is the intersection of the sphere with a plane passing thru its center. For example, the equator as well as the meridians are big circles.

2.15. Spherical Crofton formula. *The length of any curve α in the unit sphere is π times the average number of its crossings with big circles.*

More presciently, given a unit vector u, denote by $n_\alpha(u)$ the number of crossings of α and the equator with pole at u. Then

$$\operatorname{length}\alpha = \pi\cdot\overline{n_\alpha(u)}.$$

Equivalently,

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

where α_u denotes the curve obtianed by closest point projection of α to the equator with pole at u.

2.16. Exercise. *Come up with Crofton formulas for curves in the Euclidean space via projections to lines and to planes. Find the coefficients in those formulas.*

2.5 Applications

Alternative proof of Proposition 2.6. Note that

$$\operatorname{length}\beta_u \geqslant \operatorname{length}\alpha_u$$

for any unit vector u. Indeed α_u runs back and forth along a line segment and β_u has to run at least that much.

It follows that

$$\overline{\operatorname{length}\beta_u} \geqslant \overline{\operatorname{length}\alpha_u}.$$

It remains to apply the Crofton formula. □

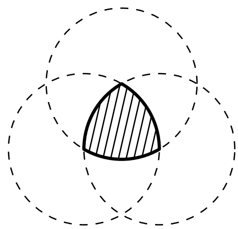
Recall that the diameter of a plane figure F is defined as the least upper bound on the distances between pairs of its points; that is,

$$\operatorname{diam} F = \sup \{\, |x - y| \;:\; x, y \in F \,\}.$$

The equilateral triangle with side 1 gives an example of a convex figure of diameter 1 that cannot be covered by a round disc of diameter 1.

2.17. Exercise. *Assume F is a convex figure of diameter 1 and D is the round disc of diameter 1. Show that*

$$\operatorname{perim} F \leqslant \operatorname{perim} D.$$



A convex figure F has constant width a if the orthogonal projection of F to any line has length a. There are many non-circular shapes of constant width. A nontrivial example is the Reuleaux triangle shown on the picture; it is the intersection of three round disks of the same radius, each having its center on the boundary of the other two. The following exercise is the so called Barbier's theorem.

2.18. Exercise. *Show that figures with constant width a have the same perimeter (which equals $\pi{\cdot}a$ — the perimeter of the round disc of diameter a).*

2.19. Exercise. *Let γ be a closed curve in the unit sphere of length shorter than $2{\cdot}\pi$. Show that γ lies in a hemisphere.*

2.20. Exercise. *Let α be a closed curve of length π. Show that it lies between a pair of parallel lines at distance 1 from each other.*

2.21. Exercise. *A spaceship flies around a nonrotating planet of unit radius and comes back to the original position; it was able to take a picture of every point on the surface of the planet.*

Try to use the Crofton formulas to get a lower bound on the length of its trajectory (does not need to be exact, but should be larger than $2{\cdot}\pi$).

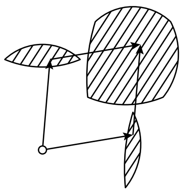
What do you think could be the shortest trajectory?

The Hausdorff distance $d_H(F, G)$ between two closed bounded sets F and G in the plane is defined as the exact lower bound on $\varepsilon > 0$ such that the ε-neightborhood of F contains G and the ε-neightborhood of G contains F.

2.22. Exercise. *Assume F and G are two closed convex figures on the plane such that $d_H(F,G) < \varepsilon$. Show that*

$$|\operatorname{perim} F - \operatorname{perim} G| < 2{\cdot}\pi{\cdot}\varepsilon.$$

Given two sets A and B on the plane, the set C is called their *Minkowski sum* (briefly $C = A + B$) if C is formed by adding each vector in A to each vector in B; that is,



$$C = \{\, a + b \;:\; a \in A,\; b \in B \,\}.$$

Note that if A and B are convex then so is $C = A + B$.

Indeed, A is convex if and only if for any pair of points $a_0, a_1 \in A$ and any $t \in [0,1]$, the point $a_t = (1-t){\cdot}a_0 + t{\cdot}a_1$ belongs to A. Similarly, B is convex if and only if for any pair of points $b_0, b_1 \in B$ and any $t \in [0,1]$, the point $b_t = (1-t){\cdot}b_0 + t{\cdot}b_1$ belongs to A.

Fix a pair of points $c_0, c_1 \in C$; by the definition of Minkowski sum, there are two pairs of points $a_0, a_1 \in A$ and $b_0, b_1 \in B$ such that $c_0 = a_0 + b_0$ and $c_1 = a_1 + b_1$. Then

$$\begin{aligned} c_t &= (1-t){\cdot}c_0 + t{\cdot}c_1 = \\ &= (1-t){\cdot}(a_0 + b_0) + t{\cdot}(a_1 + b_1) = \\ &= [(1-t){\cdot}a_0 + t{\cdot}a_1] + [(1-t){\cdot}b_0 + t{\cdot}b_1] = \\ &= a_t + b_t. \end{aligned}$$

That is, $c_t \in C$ for any $t \in [0,1]$, hence the result.

2.23. Exercise. *Show that*

$$\operatorname{perim}(A + B) = \operatorname{perim} A + \operatorname{perim} B$$

for any pair of convex figures in the plane.

2.24. Exercise. *Use Exercise 2.23 and Lemma 2.7 to give another solution of Exercise 2.22.*

2.25. Exercise. *Let γ be a curve that lies inside a convex figure F in the plane. Assume that*

$$2{\cdot}\operatorname{length}\gamma \geqslant n{\cdot}\operatorname{perim} F$$

for some integer n. Show that there is a line ℓ that intersects γ in at least n distinct points.

Chapter 3

Total curvature

3.1 Smooth regular curves

Here we introduce the so called *total curvature of a curve*. In general the term *curvature* is used for something that measures how much a geometric object deviates from being straight; total curvature is not an exception — as you will see, if the total curvature of a curve is zero, then the curve runs along a straight line.

Let $\alpha\colon [a,b] \to \mathbb{R}^3$ be a *smooth regular* curve — smooth means that the velocity vector $\alpha'(t)$ is defined and is continuous with respect to t, and regular means that $\alpha'(t) \neq 0$ for all t. If the curve α is closed then we assume in addition that $\alpha'(a) = \alpha'(b)$.

Denote by $\tau(t)$ the unit vector in the direction of $\alpha'(t)$; that is, $\tau(t) = \frac{\alpha'(t)}{|\alpha'(t)|}$. Then $\tau\colon [a,b] \to \mathbb{S}^2$ is another curve which is called *tangent indicatrix* of α. The length of τ is called the *total curvature of* α; that is,

$$\operatorname{TotCurv}\alpha := \operatorname{length}\tau.$$

3.1. Exercise. *Show that*

$$\operatorname{TotCurv}\alpha \geqslant 2\cdot\pi$$

for any smooth closed regular curve α.

Moreover, the equality holds if and only if α is a closed convex curve lying in a plane.

The above exercise is the so called Fenchel's theorem.

3.2 General definition

The total curvature of a polygonal line is defined as the sum of its external angles.

More precisely, for a polygonal line $\beta = p_0 \dots p_n$, the external angle at the vertex p_i is defined as $\alpha_i = \pi - \measuredangle p_{i-1}p_ip_{i+1}$. The total curvature of the polygonal line $\beta = p_0 \dots p_n$ is defined as the sum

$$\operatorname{TotCurv}\beta = \alpha_1 + \dots + \alpha_{n-1};$$

it is defined if the polygonal line is *nondegenerate*; that is, $p_{i-1} \neq p_i$ for any i.

If the polygonal line $p_0 \dots p_n$ is closed; that is $p_0 = p_{n+1}$ you add two more angles

$$\alpha_0 + \alpha_1 + \dots + \alpha_{n-1} + \alpha_n,$$

where $\alpha_0 = \pi - \measuredangle p_n p_0 p_1$ and $\alpha_n = \pi - \measuredangle p_{n-1} p_n p_0$.

One can define the tangent indicatrix of a polygonal line β as a spherical polygonal line (each edge is an arc of a big circle in the sphere) whose vertexes are the unit vectors $\xi_1, \dots, \xi_n$ in the directions of $p_1 - p_0$, $p_2 - p_1, \dots, p_n - p_{n-1}$ correspondingly; if the polygonal line is closed then we add one more vertex ξ_0 in the direction of $p_0 - p_n$ and two more edges $\xi_0\xi_1$ and $\xi_n\xi_0$ so the indicatrix of a closed polygonal line is a closed spherical polygonal line.

Note that the total curvature of a polygonal line is the length of its tangent indicatrix.

3.2. Exercise. *Let a, b, c, d and x be distinct points in $\mathbb{R}^3$. Show that*

$$\operatorname{TotCurv} abcd \le \operatorname{TotCurv} abxcd.$$

3.3. Exercise. *Use Exercise 3.2 to prove an analog of Fenchel's theorem (Exercise 3.1) for closed polygonal lines.*

We gave two definitions of total curvature: the first one is given in Section 3.1 via the tangent indicatrix — it works for smooth regular curves; the second is via external angles — it works for polygonal lines. The latter can be used to define total curvature of arbitrary curves.

Let $\alpha\colon [a,b] \to \mathbb{R}^3$ be a curve and $a = t_0 < \dots < t_n = b$ a partition. Set $p_i = \alpha(t_i)$. Then the polygonal line $p_0 \dots p_n$ is said to be inscribed in α.

3.4. Definition. *The total curvature of a nonconstant curve α is the exact upper bound on the total curvatures of inscribed nondegenerate polygonal lines; if α is closed then we assume that the inscribed polygonal lines are closed as well.*

We need to assume that the curve is nonconstant, otherwise it does not admit inscribed polygonal lines that are not trivial.

3.5. Exercise. *Show that the total curvature is lower semi-continuous with respect to pointwise convergence of curves. That is, if a sequence of curves* $\alpha_n\colon [a,b]\to\mathbb{R}^3$ *converges pointwise to a curve* $\alpha_\infty\colon [a,b]\to$ $\to\mathbb{R}^3$*, then*

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

Hint: Modify the proof of semi-continuity of length (Theorem 2.9).

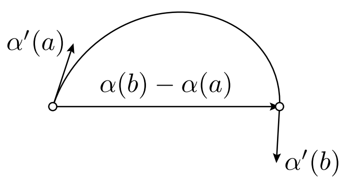
The following definition tells us that the two definitions agree.

3.6. Theorem. *For smooth regular curves the two definitions of total curvature agree; that is, for any regular curve, the length of its tangent indicatrix is equal to the exact upper bound of the total curvatures of inscribed nondegenerate polygonal lines.*

Note that from the theorem and Exercise 3.3, we get a generalization of Fenchel's theorem (Exercise 3.1) — it works for arbtrary closed curves, not necessary smooth and regular.

3.7. Lemma. *Let* $\alpha\colon [a,b]\to\mathbb{R}^3$ *be a smooth regular curve. Consider three unit vectors* λ, μ *and* ν *in the directions of* $\alpha'(a)$, $\alpha(b)-\alpha(a)$ *and* $\alpha'(b)$ *correspondingly. Then*

$$\operatorname{TotCurv}\alpha \geqslant \measuredangle(\lambda,\mu)+\measuredangle(\mu,\nu).$$



Proof. The tangent indicatrix τ runs from λ to ν in the unit sphere $\mathbb{S}^2$.

Note that τ can not be separated from μ by an equator. Indeed the vector

$$\alpha(b)-\alpha(a)=\int\limits_a^b \alpha'(t)\cdot dt$$

points in the same direction as μ. Therefore if the indicatrix $\tau=\frac{\alpha'}{|\alpha'|}$ lies in a hemisphere then μ lies in the same hemisphere.

Fix an equator ℓ in general position. If ℓ intersects the spherical polygonal line $\lambda\mu\nu$ at one point, then ℓ separates λ from ν and therefore it must intersect τ. If ℓ intersects the spherical polygonal line $\lambda\mu\nu$ at two points, then ℓ separates μ from λ and ν and therefore it must intersect τ at least twice — τ must cross ℓ and then come back. It follows that for almost all equators the number of intesections with the spherical polygonal line $\lambda\mu\nu$ can not exceed the number of intersections with τ. By the spherical Crofton formula (2.15), τ is longer than the spherical polygonal line $\lambda\mu\nu$. But the polygonal line $\lambda\mu\nu$ has length $\measuredangle(\lambda,\mu)+\measuredangle(\mu,\nu)$, hence the result. □

Let us sketch an alternative proof of the lemma which is built on Fenchel's theorem.

Alternative proof of the lemma. Note that the curve α can be extended to a smooth regular closed curve $\hat\alpha$ by an arc β that starts from $\alpha(b)$ in the same direction as α. Then turns and joins the segment $[\alpha(b),\alpha(a)]$, runs along the segment until it is close to $\alpha(a)$ turns and smoothly joints α at $\alpha(a)$.

Note that the total curvature of β can be made arbitrarily close to $2\cdot\pi-\measuredangle(\lambda,\mu)-\measuredangle(\mu,\nu)$. Indeed, β needs a bit more than $\pi-\measuredangle(\mu,\nu)$ to turn an join the segment $[\alpha(b),\alpha(a)]$ and bit more than $\pi-\measuredangle(\lambda,\mu)$ to turn an join the segment α.

By Fenchel's theorem,

$$\operatorname{TotCurv}\hat\alpha\geqslant 2\cdot\pi.$$

Evidently

$$\operatorname{TotCurv}\hat\alpha=\operatorname{TotCurv}\alpha+\operatorname{TotCurv}\beta,$$

hence the lemma follows. □

Proof of 3.6. Let $\alpha\colon[a,b]\to\mathbb{R}^3$ be a smooth curve. Fix a partition $a=t_0<\dots<t_n=b$ and consider the corresponding inscribed polygonal line $\beta=w_0\dots w_n$. Let $\chi=\xi_1\dots\xi_n$ be its tangent indicatrix — this is a spherical polygonal line; we assume that $\chi(t_i)=\xi_i$ and it has constant speed on each arc.

Consider a sequence of finer and finer partitions, denote by β_n and χ_n the corresponding inscribed polygonal line and its tangent indicatrix; since α is smooth, the χ_n converge pointwise to τ — the thangent indicatrix of α. By semi-continuity of the length functional, we get

$$\begin{aligned}\operatorname{TotCurv}\alpha&=\operatorname{length}\tau\leqslant\\&\leqslant\varliminf_{n\to\infty}\operatorname{length}\chi_n=\\&=\varliminf_{n\to\infty}\operatorname{TotCurv}\beta_n\leqslant\\&\leqslant\sup\{\operatorname{TotCurv}\beta\},\end{aligned}$$

where the last supremum is taken over all partitions and their corresponding inscribed polygonal lines β.

It remains to prove that

$$\text{❶} \qquad \operatorname{TotCurv}\alpha \geqslant \operatorname{TotCurv}\beta,$$

for any polygonal line β inscribed in α. Let ζ_i be the unit vector in the direction of $\alpha'(t_i)$. Consider the spherical polygonal line $\gamma = \zeta_0\xi_1\zeta_1\xi_2\dots\xi_n\zeta_n$; recall that $\chi = \xi_0\dots\xi_n$. By the triangle inequality,

$$\operatorname{length}\gamma \geqslant \operatorname{length}\chi = \operatorname{TotCurv}\beta.$$

By Lemma 3.7,

$$\operatorname{TotCurv}\alpha \geqslant \operatorname{length}\gamma,$$

hence ❶ follows. □

3.3 Crofton again

Given a curve α in $\mathbb{R}^3$ and a unit vector u, denote by $\alpha_{u^\perp}$ and α_u the projection of α to the plane perpendicular to u and the line parallel to u correspondingly.

To prove the following proposition, apply the spherical Crofton formula to the tangent indicatrix of α.

3.8. Proposition. *Let α be a polygonal line in $\mathbb{R}^3$. Show that*

$$\begin{aligned}\operatorname{TotCurv}\alpha &= \overline{\operatorname{TotCurv}\alpha_{u^\perp}} = \\ &= \overline{\operatorname{TotCurv}\alpha_u}.\end{aligned}$$

Note that since the curve α_u runs back and forth along one line, every time it changes direction contributes π to the total curvature of α_u. Therefore the total curvature of α_u is $n\cdot\pi$, where n is the number of changes of direction. Since n has to be even, $\operatorname{TotCurv}\alpha_u$ may take values $2\cdot\pi$, $4\cdot\pi$, $6\cdot\pi$ and so on.

3.9. Exercise. *Use the proposition and the observation above to give yet another proof of Fenchel's theorem (Exercise 3.1).*

3.4 DNA inequality

3.10. Theorem. *Let α be a closed curve that lies in a unit disc. Then*

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

Note that if $\operatorname{length}\alpha \leqslant 2\cdot\pi$, then Fenchel's theorem gives a better estimate, for longer curves this gives something new.

Proof. Assume α is a polygonal line.

Fix a unit vector u. Note that the curve α_u can run at most length 2 in one direction; therefore the number of turns has to be at least $\frac{1}{2}\cdot\operatorname{length}\alpha$. Since each turn of α_u contributes π to its total curvature, we get

$$\operatorname{TotCurv}\alpha_u \geqslant \tfrac{\pi}{2}\cdot\operatorname{length}\alpha_u.$$

The same inequality holds for the average values of left and right hand sides; that is,

$$\overline{\operatorname{TotCurv}\alpha_u} \geqslant \tfrac{\pi}{2}\cdot\overline{\operatorname{length}\alpha_u}.$$

Applying the Crofton's formula and Proposition 3.8 we get the result.

It remains to reduce the general case to polygonal lines. Given $\varepsilon > 0$, we choose an inscribed polygonal line β such that

$$\operatorname{length}\alpha < \operatorname{length}\beta + \varepsilon.$$

By the definition of total curvature (3.4) and from the first part of the proof

$$\begin{aligned}\operatorname{TotCurv}\alpha \geqslant \operatorname{TotCurv}\beta &\geqslant \\ &\geqslant \operatorname{length}\beta > \\ &> \operatorname{length}\alpha - \varepsilon.\end{aligned}$$

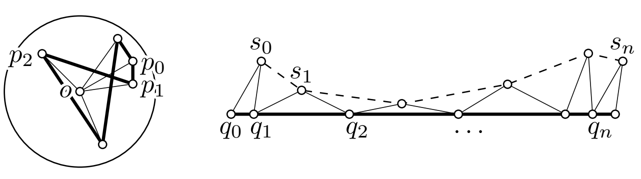
The statement follows since ε was arbitrary. □

Alternative proof. The same argument as above shows that it is sufficient to consider a closed polygonal line $\beta = p_0p_1\dots p_{n-1}$ in the unit disc. We assume that $p_n = p_0$, $p_{n+1} = p_1$ and so on. Denote by α_i the external angle at p_i.

Denote by o the center of the disc. Consider a sequence of triangles

$$\triangle q_0q_1s_0 \cong \triangle p_0p_1o,\ \ \triangle q_1q_2s_1 \cong \triangle p_1p_2o, \dots$$

such that the points $q_0, q_1 \dots$ lie on one line in that order and all the s_i's lie on one side from this line.



Note that

$$|s_n - s_0| = \operatorname{length}\beta.$$

Therefore
$$|s_0 - s_1| + \dots + |s_{n-1} - s_n| \geqslant \operatorname{length}\beta.$$

Note that

$$|q_i - s_{i-1}| = |q_i - s_i| = |p_i - o| \leqslant 1$$

and
$$\measuredangle s_{i-1} q_i s_i \leqslant \alpha_i$$
for each i. Therefore
$$|s_{i-1} - s_i| < \alpha_i$$
for each i.

It follows that

$$\begin{aligned}\operatorname{TotCurv}\beta &= \alpha_1 + \dots + \alpha_n \geqslant \\ &\geqslant |s_0 - s_1| + \dots |s_{n-1} - s_n| \geqslant \\ &\geqslant \operatorname{length}\beta.\end{aligned}$$

Hence the result. □

With minor modifications both proofs given above work in the 3-dimensional case (an in higher dimensions). The following more general result was proved by Jeffrey Lagarias and Thomas Richardson in [3], an other proof is given by Alexander Nazarov and Fedor Petrov in [4].

3.11. Theorem. *Let α be a closed curve that lies in a convex plane figure bounded by a curve γ Then the average curvature of α is not less than the average curvature of γ. Since* $\operatorname{TotCurv}\gamma = 2\cdot\pi$, *it can be written as*
$$\frac{\operatorname{TotCurv}\alpha}{\operatorname{length}\alpha} \geqslant \frac{2\cdot\pi}{\operatorname{length}\gamma}.$$

3.5 Curves of finite total curvature

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

We say that a curve $\alpha\colon [a,b] \to \mathbb{R}^3$ *does not stop* if α is not constant on any subinterval of $[a,b]$.

3.13. Exercise. *Assume that the curve α does not stop and its total curvature is less than π. Show that α is simple; that is, it has no self-intersections.*

3.14. Exercise-definition. *Assume that a curve $\alpha\colon [a,b] \to \mathbb{R}^3$ does not stop and has finite total curvature. Show that the direction of exit and entrance is defined for any point.*

That is for any $t_0 \in [a,b)$ the unit vector

$$v(\varepsilon) = \frac{\alpha(t_0+\varepsilon)-\alpha(t_0)}{|\alpha(t_0+\varepsilon)-\alpha(t_0)|}$$

converges as $\varepsilon \to 0^+$; its limit is called the direction of exit and it will be denoted by $\alpha^+(t_0)$

Analogously, for any $t_0 \in (a,b]$ the unit vector

$$w(\varepsilon) = \frac{\alpha(t_0-\varepsilon)-\alpha(t_0)}{|\alpha(t_0-\varepsilon)-\alpha(t_0)|}$$

converges as $\varepsilon \to 0^+$; its limit is called the direction of entrance and it will be denoted by $\alpha^-(t_0)$.

3.15. Exercise. *Assume that a curve $\alpha\colon [a,b] \to \mathbb{R}^3$ does not stop and has finite total curvature. Show that*

$$\alpha^+(t) = -\alpha^-(t)$$

at all $t \in [a,b]$ except possibly on a countable subset.

3.16. Exercise. *Assume a sequence of curves $\alpha_n\colon [a,b] \to \mathbb{R}^3$ converges uniformly to a curve $\alpha_\infty\colon [a,b] \to \mathbb{R}^3$ and*

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

Show that

$$\operatorname{TotCurv}\alpha_n \to \infty \quad \textit{as} \quad n \to \infty.$$

3.6 Total signed curvature

Let us define the *total signed curvature* of a polygonal line in the plane as the sum of the signed external angles; the external angle has positive sign if the line turns left and negative sign if the line turns right; the signed external angle is undefined if a pair of adjacent edges overlap; that is if at one vertex the polygonal line turns in the exact opposite direction. In particular the total signed curvature is defined for any simple polygonal line in the plane.

3.17. Exercise. *Assume that the total signed curvature of a closed polygonal line in the plane is defined. Show that it is a multiple of* $2{\cdot}\pi$.

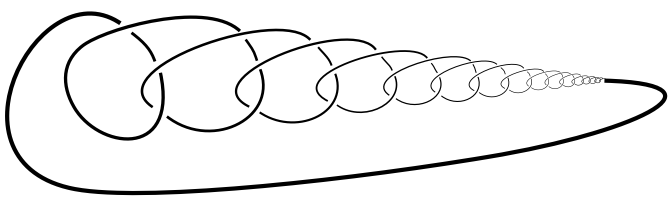
3.18. Exercise. *Show that the total signed curvature of any closed simple polygonal line in the plane is* $\pm 2{\cdot}\pi$*; if the bounded region lies on the left from the curve then it is* $+2{\cdot}\pi$ *and if it lies on the right then it is* $-2{\cdot}\pi$.

Chapter 4

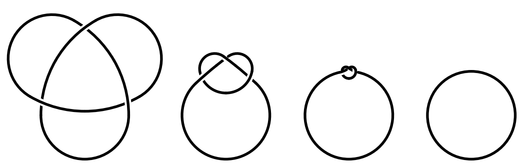
Fáry–Milnor theorem

4.1 Tame knots

It is tricky to make a formal definition that captures the intuitive meaning of *knot*. An attempt to define knots as simple closed curves leads to pathological examples as the one show on the diagram — these are the so called *wild knots*. If one adds that the curve has to



be smooth and regular, then these examples disappear, but it is still tricky to give right definition of *deformation* — the following diagram shows that it can not be defined as a continuous family of closed simple



smooth regular curves. Observe that all curves on the diagram are smooth and regular for all times including the last moment.

We define a *knot* (more precicely *tame knot*) as a simple closed polygonal line in the Euclidean space $\mathbb{R}^3$.

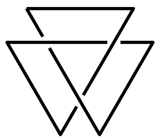
The notation $\triangle abc$ is used for the triangle abc; that is, a polygonal line with three edges and vertexes a, b and c. Let us denote by $\blacktriangle abc$ the convex hull of the points a, b and c; $\blacktriangle abc$ is the solid triangle with the vertexes a, b and c. The points a, b and c are assumed to be distinct, but they might lie on one line; that is, for us a degenerate triangle counts as a legitimate triangle.

We define a *triangular isotopy of a knot* to be the generation of a new knot from the original one by means of the following two operations:

Assume $[pq]$ is an edge of the knot and x is a point such that the solid triangle $\blacktriangle pqx$ has no common points with the knot except for the edge $[pq]$. Then we can replace the edge $[pq]$ in the knot by the two adjacent edges $[px]$ and $[xq]$.

We can also perform the inverse operation. That is, if for two adjacent edges $[px]$ and $[xq]$ of a knot the triangle $\blacktriangle pqx$ has no common points with the knot except for the points on the edges $[px]$ and $[xq]$, then we can replace the two adjacent edges $[px]$ and $[xq]$ by the edge $[pq]$.

Polygons that arise from one another by a finite sequence of triangular isotopies are called *isotopic.*



A knot that is not isotopic to a triangle is called nontrivial.

The trefoil knot shown on the diagram gives a simple example of nontrivial knot. A proof that the trefoil knot is nontrivial can be found in any textbook on knot theory, we do not give it here. The most elementary and visual proof is based on the so called *tricolorability* of knot diagrams.

4.1. Exercise. *Let x and y be two points on the adjacent edges $[p_1p_2]$ and $[p_2p_3]$ of a knot $\beta = p_1p_2p_3\dots p_n$. Assume that the solid triangle $\blacktriangle xp_2y$ intersects β only along $[xp_2]\cup[p_2y]$. Show that the knot $\beta' = p_1xyp_3\dots p_n$ is isotopic to β.*

4.2 Fáry–Milnor theorem

We will give some proofs of the following theorem.

4.2. Theorem. *The total curvature of any nontrival knot is at least $4{\cdot}\pi$.*

The famous Fáry–Milnor theorem states that the inequality is strict; that is, the total curvature of any nontrival knot *exceeds* $4{\cdot}\pi$. It is easy to construct a trefoil knot with total curvature arbitrary close to $4{\cdot}\pi$; therefore this result is optimal.

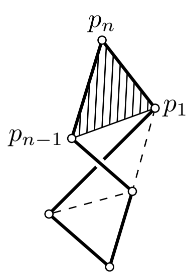
The question was raised by Karol Borsuk [5] and answered independently by István Fáry and John Milnor [6, 7]; later other proofs were found.

4.3 Milnor's proof

In the proof we will use the following fact.

4.3. Proposition. *Assume that a height function $(x, y, z) \to z$ has only one local minimum and one local maximum on a closed simple polygonal line and all the vertexes of the polygonal line are at different height. Then the line is a trivial knot.*

The proof is a simple application of the definition of isotopy, given in the previous section.



Proof. Let $\beta = p_1 \dots p_n$ be the closed simple polygonal line such that the height function $(x, y, z) \to z$ has one local minimum one local maximum. Note that on each of the two arcs of β from the min-vertex to the max-vertex the height function increases monotonically.

Consider the three vertexes with the largest height; they have to include the max-vertex and two more. Note that these three vertexes are consequent in the polygonal line; without loss of generality we can assume that they are p_{n-1}, p_n, p_1.

Note that the solid triangle $\blacktriangle p_{n-1}p_np_1$ does not intersect any edge β except the two adjacent edges $[p_{n-1}p_n] \cup [p_np_1]$. Indeed, if $\blacktriangle p_{n-1}p_np_1$ intersects $[p_1p_2]$, then, since p_2 lies below $\blacktriangle p_{n-1}p_np_1$, $[p_1p_2]$ must intersect $[p_{n-1}p_n]$ which is impossible since β is simple. The same way one can show that $\blacktriangle p_{n-1}p_np_1$ can not intersect $[p_{n-2}p_{n-1}]$. The remaining edges lie below $\blacktriangle p_{n-1}p_np_1$, hence they can not intersect this triangle.

Applying a triangular isotopy, to $\blacktriangle p_{n-1}p_np_1$ we get a closed simple polygonal line $\beta' = p_1 \dots p_{n-1}$ which is isotopic to β.

Since all the vertexes p_i have different height, the assumption of the proposition holds for β'.

Repeating this procedure $n-3$ times we get a triangle. Hence β is a trivial knot. □

Milnor's proof of 4.2. Let α be a simple closed polygonal line. Assume its total curvature is less that $4{\cdot}\pi$. Then by Proposition 3.8,

$$\operatorname{TotCurv}\alpha_u < 4{\cdot}\pi$$

for some unit vector u. Moreover, we can assume that u points in a generic direction; that is, u is not perpendicular to any edge or diagonal of α.

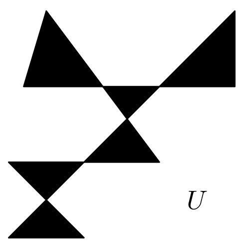
The total curvature of α_u is π times the number of turns of α_u which has to be an even number. It follows that the number of turns of α_u is at most 2; it cannot be less than 2 for a generic direction, therefore it is exactly 2.

That is, if we rotate the space so that u points upward, then the height function has exactly one minimum and one maximum; by Proposition 4.3, α is a trivial knot — hence the result. □

4.4 Fáry's proof

Let us give a sketch of another proof, based on the original idea of István Fáry.

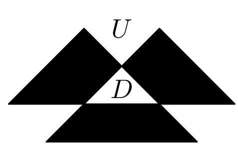
Fáry's proof of 4.2. Consider a projection of the knot to a plane in general position. That is, we assume that the self-intersections of the projection are at most double and the projection of each edge is not degenerate. The obtained closed polygonal line $\beta = p_1p_2\dots p_n$ divides the plane into domains, one of which is unbounded, denote it by U, and the others are bounded.



First note that all domains can be colored in a chessboard order; that is, they can be colored in black and white in such a way that domains with common borderline get different colors. If the unbounded domain is colored in white and every other domain is colored in black then one can untie the knot by flipping these domains one by one.

4.4. Exercise. *Give a formal proof of the last statement; that is, show that if the only undbounded domain is white then β is isotopic to a triangle.*

Therefore among the bounded domains there is a white domain, denote it by D. The domain D cannot adjoin U, since they have the same color. Fix a point o in this domain.



For each i, set

$$\begin{aligned}\varphi_i &= \pi - \measuredangle p_{i-1}p_ip_{i+1},\\ \psi_i &= \measuredangle p_{i-1}op_i,\\ \theta_i &= \measuredangle op_ip_{i+1}.\end{aligned}$$

Here indexes are taken modulo n; in particular, $p_n = p_0$.

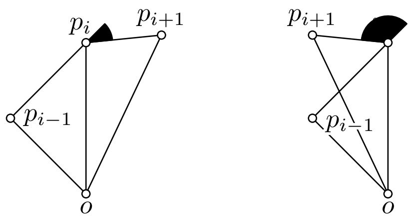
Note that φ_i is the external angle at p_i; therefore

$$\operatorname{TotCurv}\beta = \varphi_1 + \dots + \varphi_n$$

Direct calculations show that

$$\varphi_i \geqslant \psi_i + \theta_{i-1} - \theta_i.$$

In the two pictures below, φ_i is the solid angle and the angles ψ_i, θ_{i-1} and θ_i are just as drawn. We have equality on the first picture and strict inequality on the second picture.



It follows that

$$\varphi_1 + \dots + \varphi_n \geqslant \psi_1 + \dots + \psi_n.$$

The last sum is the total angle at which β is seen from o counted with multiplicity. The boundary of D contributes at least $2\cdot\pi$ to this sum and the boundary of U contributes with other $2\cdot\pi$; since their boundaries do not overlap we get

$$\psi_1 + \dots + \psi_n \geqslant 4\cdot\pi,$$

hence the result.

This is true for the projection of the knot to any plane in general position. The remaining planes contribute nothing to the average value. Therefore by Proposition 3.8, the total curvature of the original knot is at least $4\cdot\pi$. □

4.5. Exercise. *Construct a closed smooth simple curve with total curvature arbitrarily close to* $2\cdot\pi$ *such that its projection to any plane has at least* 10 *self-intersections.*

4.5 Proof of Alexander and Bishop

Here we sketch a proof of the Fáry–Milnor theorem given by of Stephanie Alexander and Richard Bishop in [8].

The proof is elementary, but not simple (elementary does not mean simple, it means only that it does not use much theory). It is based on the following two facts that we are already familiar with:

◇ If a closed polygonal line β' is inscribed in a closed polygonal line β then
$$\operatorname{TotCurv}\beta' \leqslant \operatorname{TotCurv}\beta.$$

◇ The total curvature of a doubly covered bigon is $4{\cdot}\pi$; that is,
$$\operatorname{TotCurv}\beta = 4{\cdot}\pi$$
if $\beta = pqpq$ for two distinct points p and q. Similarly if a quadrilateral is sufficiently close to a doubly covered bigon, then its total curvature is close to $4{\cdot}\pi$.

Proof. Let $\beta = p_1 \dots p_n$ be a closed polygonal line that is not a trivial knot; that is, one can not get a triangle from β by applying a sequence of triangular isotopies defined in the previous section.

We proceed by induction on the number $n \geqslant 3$. In the base case $n = 3$ the polygonal line β is a triangle. Therefore, by definition, β is a trivial knot — nothing to show.

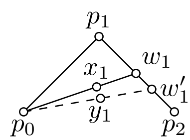
Consider the smallest n for which the statement fails; that is, there is a closed simple polygonal line $\beta = p_1 \dots p_n$ that is not a trivial knot and such that

❶ $$\operatorname{TotCurv}\beta < 4{\cdot}\pi.$$

We use the indexes modulo n; that is, $p_0 = p_n$, $p_1 = p_{n+1}$ and so on. Without loss of generality, we may assume that β is in general position; that is, no four vertexes of β lie on one plane.

Set $\beta_0 = \beta$. If the solid triangle $\blacktriangle p_0p_1p_2$ intersects β_0 only in the two adjacent edges, then applying the corresponding triangular isotopy, we get a knot β_0' with $n-1$ edges that is inscribed in β_0. Therefore
$$\operatorname{TotCurv}\beta_0 \geqslant \operatorname{TotCurv}\beta_0'.$$
On the other hand, by the induction hypothesis
$$\operatorname{TotCurv}\beta_0' \geqslant 4{\cdot}\pi,$$
which contradicts ❶.

Choose the first point w_1' on the edge $[p_1p_2]$ so that the line segment $[p_0w_1']$ intersects β_0. Denote a point of intersection by y_1.



Choose a point w_1 on $[p_1p_2]$ a bit before w_1' (below we explain how close). Denote by x_1 the point on $[p_0w_1]$ that minimizes the distance to y_1. This way we get a closed polygonal line $\beta_1 = w_1p_2\dots p_n$ with two marked points x_1 and y_1. Denote by m_1 the number of edges in the arc $x_1w_1\dots y_1$ of β_1.

By Exercise 4.1, β_1 is isotopic to β_0; in particular β_1 is a nontrivial knot.

Now let us repeat the procedure for the adjacent edges $[w_1p_2]$ and $[p_2p_3]$ of β_1. If the solid triangle $\blacktriangle w_1p_2p_3$ intersects β_1 only at these two adjacent edges, then we get a contradiction with the induction hypothesis the same way as before. Otherwise we get a new knot $\beta_2 = w_1w_2p_3\dots p_n$ with two marked points x_2 and y_2. Denote by m_2 the number of edges in the broken line $x_2w_2\dots y_2$.

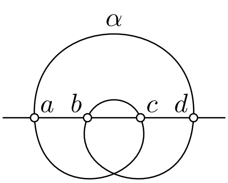
Note that the points x_1, x_2, y_1, y_2 can not appear on β_2 in the same cyclic order; otherwise the broken line $x_1x_2y_1y_2$ can be made to be arbitrary close to a doubly covered bigon which again contradicts ❶.[1]

Therefore we can assume that the arc $x_2w_2\dots y_2$ lies inside the arc $x_1w_1\dots y_1$ in β_2 and therefore $m_1 > m_2$.

Continuing this procedure we get a sequence of polygonal lines $\beta_i = w_1\dots w_ip_{i+1}p_n$ with marked points x_i and y_i such that the number of edges m_i from x_i to y_i decreases as i increases. Clearly $m_i > 1$ for any i and $m_1 < n$. Therefore it requires less than n steps to get a contradiction with the induction hypothesis. □

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

$$\operatorname{TotCurv}\alpha \geqslant 4\cdot\pi.$$



A line crossing a knot at four points as in the exercise is called *alternating quadrisecants.* It turns out that any nontrivial knot admits

[1] More precisely, the choice of w_1 has to be made so that the distance $|x_1 - y_1|$ would be much less that all the distances between y_1 and any point $z \in \beta\cap\blacktriangle p_1p_2p_3$, so we have

$$\measuredangle y_1zx_1 < \tfrac{\varepsilon}{10},$$

where $\varepsilon = 4\cdot\pi - \operatorname{TotCurv}\beta$. In this case, since $y_2 \in \beta\cap\blacktriangle p_1p_2p_3$ and x_2 can be taken arbitrary close to y_2, we have

$$\operatorname{TotCurv} x_1x_2y_1y_2 > 4\cdot\pi - \varepsilon = \operatorname{TotCurv}\beta$$

which can not happen since $x_1x_2y_1y_2$ is inscribed in β.

an alternating quadrisecants [9]; it provides yet another proof of the Fáry–Milnor theorem.

4.7. Advanced exercise. *Show that given any real number* Φ *there is a knot* β *such that any knot isotopic to* β *has total curvature at least* Φ*.*

Hint: Use that there are knots with arbitrary large *bridge number*, see for example [10] and the references therein.

Chapter 5

Osculating circlines

5.1 Acceleration of unit-speed curve

Any regular smooth curve can be parametrized by its length. The obtained curve α has unit speed; that is, $|\alpha'(t)| = 1$ for all t. This is called the *natural parametrization*.

It is straightforward to show any smooth regular curve remains smooth (and surely regular) if equipped with a natural parametrization; here smooth means that all derivatives $\alpha^{(n)}(t)$ are defined for any n and all values of t in the domain of definition of α.

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

5.1. Proposition. *Assume $\alpha\colon [a,b] \to \mathbb{R}^2$ be a smooth unit-speed curve. Then*

$$\alpha'(t) \perp \alpha''(t)$$

for any t.

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. Since $|\alpha'(t)| = 1$, we have

$$\langle \alpha'(t), \alpha'(t)\rangle = 1.$$

Differentiating both sides we get

$$2\cdot\langle\alpha''(t),\alpha'(t)\rangle = 0,$$

hence the result. □

5.2 Signed curvature

Given a vector $v\in\mathbb{R}^2$ denote by $i\cdot v$ the vector obtained from v by the counterclockwise rotation by $\frac{\pi}{2}$. (The "multiplication" by i agrees with the miultiplication by the imaginary unit if one uses complex coordinates on the plane $z = x + i\cdot y$.)

Suppose $\alpha\colon[a,b]\to\mathbb{R}^2$ is a smooth unit-speed curve. Recall that curvature of α at t can be defined as $|\alpha''(t)|$.

The *signed curvature* $\kappa_\alpha(t)$ is uniquely defined by the identity

$$\alpha''(t) = \kappa_\alpha(t)\cdot i\cdot\alpha'(t).$$

Note that by Proposition 5.1 this equation has a solution. Since $|\alpha'(t)| = 1$ we have $|\kappa_\alpha(t)| = |\alpha''(t)|$ for any t.

The signed curvature measures how fast the direction $\tau(t) = \alpha'(t)$ rotates; the signed curvature is positive if τ turns left and negative if τ turns right; if the curve goes straight then its curvature vanishes.

5.3 Osculating circline

It is straightforward to prove the following statement.

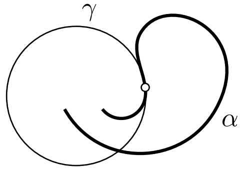
5.2. Proposition. *Given a point p, a unit vector u and a real number κ there is unique smooth unit-speed curve $\gamma\colon\mathbb{R}\to\mathbb{R}^2$ that starts at p in the direction of u and has constant signed curvature κ.*

Moreover, if $\kappa = 0$, then γ runs along the line $\gamma = p + t\cdot u$ and if $\kappa\ne 0$, then γ runs around the circle of radius $\frac{1}{|\kappa|}$ and center $p+\frac{i}{\kappa}\cdot u$.

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

5.3. Definition. *Let α be a smooth unit-speed plane curve; denote by $\kappa_\alpha(t)$ the signed curvature of α at t.*

For $t_0\in[a,b]$, the unit-speed curve γ of constant signed curvature $\kappa_\alpha(t_0)$ that starts at $\alpha(t_0)$ in the direction $\alpha'(t_0)$ is called the osculating circline *of α at t_0.*



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.

5.4 Spiral theorem

The following theorem states 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. This theorem was proved by Peter Tait [see 11] and later rediscovered by Adolf Kneser [see 12].

5.4. Theorem. *Assume α is a smooth regular plane curve with strictly monotonic curvature. Then α is simple.*

The same statement also holds for signed curvature; the proof requires only minor modifications.

5.5. 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.*

The proof of theorem is based on the following lemma.

5.6. Lemma. *Assume that α is a smooth regular plane curve with strictly decreasing positive signed curvature. Then the osculating circles of α are nested; that is, if γ_t denoted the osculating circle of α at t, then γ_{t_0} lies in the open disc bounded by γ_{t_1} for any $t_0 < t_1$.*

The osculating circles of the curve α 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 2, Lecture 10]. Also note that the curve α is tangent to a circle of the foliation at each of its points. However, it does not run along a circle.

Proof. Let $z(t)$ be the curvature center and

$$r(t) = \tfrac{1}{\kappa_\alpha(t)}$$

the radius of curvature of α at t. Note that

$$z(t) = \alpha(t) + r(t){\cdot}i{\cdot}\alpha'(t).$$

Therefore

$$\begin{aligned} z'(t) &= \alpha'(t) + r'(t){\cdot}i{\cdot}\alpha'(t) + r(t){\cdot}i{\cdot}\alpha''(t) = \\ &= \alpha'(t) + r'(t){\cdot}i{\cdot}\alpha'(t) + r(t){\cdot}i{\cdot}\kappa_\alpha(t){\cdot}i{\cdot}\alpha'(t) = \\ &= \alpha'(t) + r'(t){\cdot}i{\cdot}\alpha'(t) - \alpha'(t) = \\ &= r'(t){\cdot}i{\cdot}\alpha'(t). \end{aligned}$$

Since $\kappa_\alpha(t)$ is decreasing, $r(t)$ is increasing; therefore $r' \geqslant 0$. It follows that $|z'(t)| = r'(t)$ and $z'(t) \perp \alpha'(t)$.

Note that the curve $z(t)$ does not have straight arcs; therefore

$$
(*) \qquad \begin{aligned}
|z(t_1) - z(t_0)| &< \int_{t_0}^{t_1} |z'(t)|\cdot dt = \\
&= \int_{t_0}^{t_1} r'(t)\cdot dt = \\
&= r(t_1) - r(t_0).
\end{aligned}
$$

By $(*)$, the osculating circle at t_0 lies inside the osculating circle at t_1 without touching it. $\square$

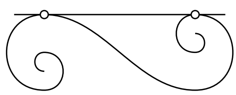
Proof of 5.4. Note that $\alpha(t) \in \gamma_t$ for any t. Applying the lemma we get $\alpha(t_1) \neq \alpha(t_0)$ if $t_1 \neq t_0$. Hence the result. $\square$

The lemma can be used to solve the following exercise.

5.7. Exercise. *Assume that α is a smooth regular plane curve with strictly monotonic curvature.*

(a) *Show that no line can be tangent to α at two distinct points.*
(b) *Show that no circle can be tangent to α at three distinct points.*

Note that part (a) does not hold for smooth regular plane curve with strictly monotonic *signed* curvature; an example is shown on the diagram.



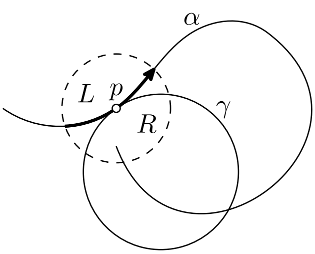
Chapter 6

Supporting circlines

6.1 Definitions

Suppose $\alpha\colon[a,b]\to\mathbb{R}^2$ is a smooth unit-speed plane curve and $t_0\in$ $\in(a,b)$.

A circline γ supports α at t_0 if $\alpha(t_0)\in\gamma$ and γ lies locally on one side of α. If $p=\gamma(t_0)$ for a single value t_0, then we can also say γ *supports α at p* without ambiguity.



More precisely, assume that there is a round neighborhood $U\ni p$ such that for some interval $[a',b']\ni t_0$ the arc $\bar\alpha=\alpha|_{[a',b']}$ has no self-intersection and runs from boundary to boundary of U. In this case α divides U intosets L and R; L lies on the left and R lies on the right from $\bar\alpha$. If $\gamma\cap U$ contains only points of $\bar\alpha$ and R, we say that γ *supports α on the right*; if $\gamma\cap U$ contains only points of $\bar\alpha$ and L, we say that γ *supports α on the left*.

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

Suppose a unit-speed circline γ supports a smooth unit-speed plane curve α at t_0. Without loss of generality we can assume that $\gamma(0)=$ $=\alpha(t_0)$. Then $\gamma'(0)=\pm\alpha'(t_0)$. If not, then the curve α would cross γ transversely and therefore could not stay on the same side for values close to t_0. Therefore reverting the parametrization of γ if necessary we may (and further will) assume that

$$\gamma'(0)=\alpha'(t_0)$$

holds for any supporting circline γ to α at t_0.

6.2 Supporting test

The following proposition resembles the second derivative test.

6.1. Proposition. *Assume γ is a circle that that supports α at t_0 from the rigth (correspondingly left). Then*

$$\kappa(t_0) \geqslant \kappa \quad (\textit{correspondingly} \quad \kappa(t_0) \leqslant \kappa).$$

where κ is the signed curvature of γ and $\kappa(t_0)$ is the signed curvature of α at t_0.

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

$$\kappa(t_0) > \kappa \quad (\textit{correspondingly} \quad \kappa(t_0) < \kappa).$$

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

Since $\kappa \neq 0$, the curve γ is a circle. According to Proposition 5.2, γ has radius $\frac{1}{\kappa}$ and it is centered at

$$z = \alpha(t_0) + \tfrac{i}{\kappa} \cdot \alpha'(t_0).$$

Consider the function

$$f(t) = |z - \alpha(t)|^2 - \tfrac{1}{\kappa^2}.$$

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

◇ if γ supports α at t_0 from the right, then

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

◇ if γ supports α 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 γ supports α at t_0 from the right;

◇ if

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

then γ supports α at t_0 from the left;

Direct calculations show that

$$\begin{aligned}
f(t_0) &= 0;\\
f'(t_0) &= \langle z-\alpha(t), z-\alpha(t)\rangle'|_{t=t_0} =\\
&= -2{\cdot}\langle \alpha'(t_0), z-\alpha(t_0)\rangle =\\
&= -2{\cdot}\langle \alpha'(t_0), \tfrac{i}{\kappa}{\cdot}\alpha'(t_0)\rangle =\\
&= 0;\\
f''(t_0) &= \langle z-\alpha(t), z-\alpha(t)\rangle''|_{t=t_0} =\\
&= 2{\cdot}\left(\langle \alpha'(t_0), \alpha'(t)\rangle - \langle \alpha''(t_0), z-\alpha(t)\rangle\right) =\\
&= 2{\cdot}\left(1-\kappa{\cdot}\frac{1}{\kappa(t_0)}\right)
\end{aligned}$$

Hence the result. □

6.2. Exercise. *Assume α is a closed smooth unit-speed plane curve that runs in a unit disk. Show that there is a point on α with curvature at least* 1.

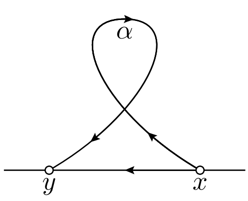
Give two proofs, one based on the DNA inequality 3.10 and another one based on Proposition 6.1.

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

6.3 Lens lemma

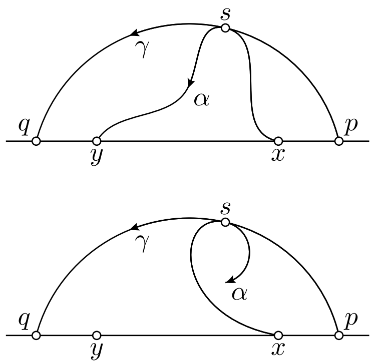
6.4. Lemma. *Let α be a smooth regular simple curve that runs from x to y. Assume that α runs on the right side (correspondingly left side) of the oriented line xy and only its end points x and y lie on the line. Then α has a point with positive (correspondingly negative) curvature.*

Note that the lemma fails for curves with self-intersections; the curve α 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.

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



Note that the $\alpha'(t_0)$ is tangent to γ at s. Moreover $\alpha'(t_0)$ points in the direction of q; that is, if we go along γ in the direction of $\alpha'(t_0)$ then we have to start at p and end at q. If the direction is opposite, then the arc of α from s to y would be trapped in the curvelinear triangle xsp bounded by arcs of γ, α and the line segment $[px]$. But this is impossible since y does not belong to this triangle.

It follows that γ supports α at t_0 from the right. By Proposition 6.1,

$$\kappa(t_0) \geqslant \kappa,$$

where $\kappa(t_0)$ is signed curvature of α at t_0 and κ is the curvature of γ. Evidently $\kappa > 0$, hence the result. □

6.4 Convexity and inflection points

6.5. Exercise. *Assume α is a closed regular simple plane curve with positive signed curvature. Show that α bounds a convex set.*[1]

6.6. Exercise. *Assume α is a closed smooth regular plane curve with positive signed curvature. Show that α is simple if and only if its total curvature is $2{\cdot}\pi$.*

6.7. Exercise. *Assume a smooth regular curve α has curvature at most 1 at any point (that is, $|\kappa_\alpha(t)| \leqslant 1$ for any t). Show that both unit circles tangent to γ at t_0 are supporting.*

Moreover, there is $\varepsilon > 0$ ($\varepsilon = \frac{1}{2}$ will do) such that any arc of α of length $< \varepsilon$ starting at $p = \alpha(t_0)$ cannot enter the unit circle tangent to α at p.

[1] Hint: show that any tangent line to α is supporting.

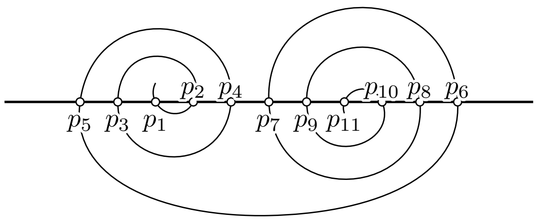
6.8. Exercise. *Suppose α is a simple smooth regular curve in the plane with positive curvature. Assume α crosses a line ℓ at the points $p_1, p_2, \dots p_n$ and these points appear on α in that same order.*

(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 they 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).*

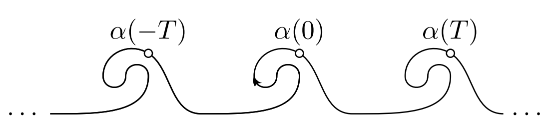


Recall that Conv X denotes the convex hull of the set X; that is, Conv X is the intersection of all convex sets containing X.

6.9. Exercise. *Suppose α is a simple smooth regular curve with positive curvature in the plane. Then the boundary of* Conv α *is formed by an arc of α together with a line segmet connecting the ends of this arc.*

6.10. Exercise. *Suppose α is a simple smooth regular curve in the plane. Show that α lies on one side from one of its tangent lines.*

If the curvature of a curve α vanishes at t_0, then we say that t_0 is inflection value of the parameter, and $p = \alpha(t_0)$ is an inflection point; the later convention might be ambiguous only if α has a self-intersection at p. In other words, t_0 is an inflection value if the osculating circline at t_0 coincides with the tangent line.



6.11. Exercise. *Let $\alpha\colon \mathbb{R} \to \mathbb{R}^2$ be a smooth simple regular plane curve. Assume α is periodic in the following sense: there is $T > 0$ and a vector $v \in \mathbb{R}^2$ such that*

$$\alpha(t + T) = \alpha(t) + v.$$

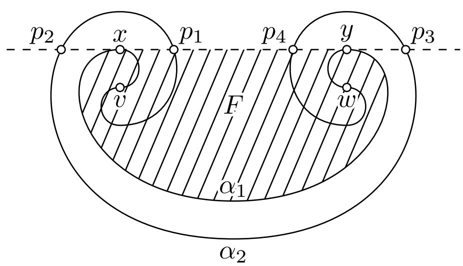
Show that α has at least 2 inflection points in the interval $[0, T]$.

6.12. Theorem. *Let α be a closed simple smooth regular plane curve. Assume α has exactly two inflection points dividing α in two arcs α_1 and α_2. Then*

$$\operatorname{Conv} \alpha_1 \supset \alpha_2 \quad \textit{or} \quad \operatorname{Conv} \alpha_2 \supset \alpha_1.$$

Proof. Let us denote the inflection points by v and w and orient the arcs α_1 and α_2 from v to w; we can then assume that both arcs α_1 and α_2 have positive curvature.

Set $F = \operatorname{Conv} \alpha_1$. By Exercise 6.9, F is bounded by an arc $\bar{\alpha}_1$ of α_1 from x to y and the line segment xy. We may assume that the line xy is horizontal, x lies on the left from y and so α_1 lies below the line.



If $F \not\supset \alpha_2$, then α_2 runs outside of F, so it has to cross the line segment xy. Denote by $p_1, p_2, \ldots$ the points of intersection of α_2 with the line xy as they appear on α_2.

Since α_2 has positive curvature, p_1 lies on left from p_0. If p_2 lies between x and p_1 then the curve α_2 is trapped in the region bounded by the arc xvp_2 of α and the line segment p_2x; therefore it can not reach w — a contradiction. Therefore p_2 lies on the extension of yx behind x.

Further p_3 lies on right from p_2. If p_3 lies between p_1 and x, then α_2 is trapped in the region bounded by the arc xvp_3 of α and the line segment p_3x — a contradiction again. Therefore p_3 lies on the extension of xy behind y and the arc p_2p3 of α surrounds F. Whence

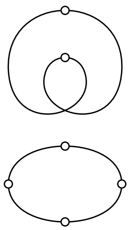
$$\operatorname{Conv} \alpha_2 \supset \alpha_1. \qquad \square$$

6.5 Four-vertex theorem

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

6.13. Exercise. *Assume the osculating circle of a curve α at t_0 supports α at t_0. Show that t_0 is a vertex of α.*

6.14. 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; he second curve has exactly four vertexes and no self-intersections.

The four-vertex theorem was first proved by Syamadas Mukhopadhyaya [13] for convex curves. By now it has a large number of different proofs and generalizations. We will present a proof given by Robert Osserman [14].

Proof. Fix a simple smooth regular closed plane curve α.

Suppose that $2{\cdot}n$ points $p_1, s_1, \dots, p_n, s_n$ appear on a closed curve α in the same cyclic order. Fix a real number κ. Assume that the curvature of α at p_i is at least κ and its curvature at s_i is at most κ. Then each of n arcs p_np_1, $p_1p_2, \dots p_{n-1}p_n$ of α has a point of minimum curvature in its interior. Similarly each of the n arcs s_ns_1, $s_1s_2, \dots s_{n-1}s_n$ of α has a point of maximum curvature.

If one of these local minima coincides with a local maximum, an arc around this point has constant curvature; in this case all these points are vertexes and we have an infinite number of them. If they are all different, then we have at least $2{\cdot}n$ vertexes.

Therefore it is sufficient to show that

❶ *there are at least 4 points p_1, s_1, p_2, s_2 with the described properties for some κ.*

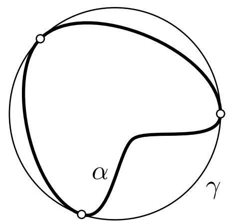
Note that

❷ *α admits a unique* circumscribed circle *γ; that is, a circle of minimal radius that encloses α.*

Denote by r the infimum of radii of circles that enclose α. We can choose a sequence of circles γ_n enclosing α such that their radii $r_n \to r$. Note that all the centers of γ_i lie at a bounded distance from α. Therefore passing to a subsequence we can assume that the centers of γ_n converge to a point o. Note that the circle γ with center o and radius r encloses α; hence the existence of the circumscribed circle follows.

If there are two distinct circumscribed circles, then α lies in the intersection of the discs bounded by these circles. But this intersection is enclosed in a circle of smaller radius — a contradiction. Hence Claim ❷ follows.

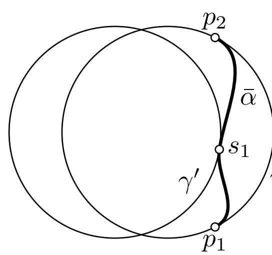
❸ *Assume γ is the circumscribed circle of α. Then γ touches α at least 2 points which divide the γ in arcs no longer than a semicircle.*



If it was not the case, then one could move γ slightly keeping its radius the same so that γ will not touch α at all. But in this case α could be enclosed in a circle of smaller radius — a contradiction.

Let us orient α and γ counterclockwise. Then at the common points the directions of α and γ coincide. Note that these points appear on α and γ in the same order; otherwise α would not be simple.

Denote by κ the signed curvature of γ, since it is oriented counterclockwise, $\kappa = \frac{1}{r} > 0$.



Fix two common points p_1 and p_2 of α and γ. By Proposition 6.1, the curvature of α at p_1 and p_2 is at least κ. Let $\bar{\alpha}$ be the arc of α from p_1 to p_2.

We can assume that the circle is centered at the origin and the points p_1 and p_2 lie on the same vertical line in the right halfplane of the coordinate plane.

Let o be the center of γ. Consider another circle γ' with the same radius r and center o' the leftmost point in the x-axis such that γ' intersects $\bar{\alpha}$.

Denote by s_1 a common point of $\bar{\alpha}$ and γ'; we can assume that s_1 is not an end point of $\bar{\alpha}$. At the point s_1 the directions of $\bar{\alpha}$ and γ' coincide, otherwise α could not be simple — the same argument is used in the proof of Lemma 6.4. Therefore γ' supports α from the left at s_1. By Proposition 6.1, the curvature of α at s_1 is at most κ.

Repeating the same argument for another pair of points p_2, p_3,[2] we prove Claim ❶. Hence the theorem follows. □

6.15. Exercise. *Show that any smooth regular curve of constant width has at least 6 vertexes.*

[2] If p_1 and p_2 divide γ into two semicircles, then we can take $p_3 = p_1$.

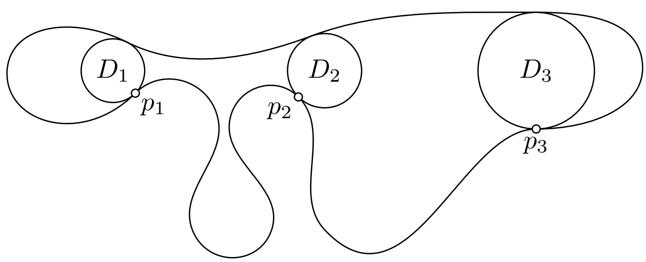
6.6 Moon in a puddle

6.16. Exercise. *Let F be a convex plane figure bounded by a simple closed smooth regular curve α of curvature bounded by 1. Assume α is oriented counterclockwise, so the figure F lies on the left from α. Show that there is a unit circle that is globally supporting α form the left at any given point.*

The following theorem was proved by Vladimir Ionin and German Pestov [15]. It gives a first nontrivial example of the so called "local to global theorems" — based on some local data (in this case the curvature of a curve) we can conclude a global property (in this case existence of a large disc surrounded by the curve). For convex curves, the result was known much earlier [16, §24].

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

Proof. Denote by F the closed region surrounded by α.



Fix $p_1 \in \alpha$. Consider the disc D_1 of maximal radius that is tangent to α at p_1 and lies completely in F.

If the radius of D_1 is at least 1, then the problem is solved. Otherwise note that p_1 is an isolated point of the intersection $D_p \cap \alpha$. Moreover according to Exercise 6.7, there is a fixed value $\varepsilon > 0$ such that any arc of α that starts at p_1 can not end in D_1.

Consider an arc α_1 of α that runs along α from p_1 to the next point in D_1. Denote by F_1 the region that contains D_1 and whose boundary is formed by α_1 and part of the boundary of D_1. From above length $\alpha_1 > \varepsilon$.

Let p_2 be the midpoint of α_1. Let D_2 be the disc of maximal radius that is tangent to α_1 at p_2 and lies completely in F_1. The disc D_2 touches the boundary ∂F_1 at other points, dividing it in at least two arcs.

Note that D_2 can not touch the boundary of D_1, otherwise it would lie inside D_1, which is impossible. Therefore at least one of these arcs, say α_2, do not contain the common boundary of F_1 and D_1. Note that

$$\operatorname{length}\alpha_2 < \tfrac{1}{2}\cdot\operatorname{length}\alpha_1.$$

Again, if the radius of D_2 is at least 1, then the theorem is proved. By Exercise 6.7, it happens if $\operatorname{length}\alpha_2 < \varepsilon$. If the radius is less than 1, denote by F_2 the region that contains D_2 and is bounded by α_2 and a part of the boundary of D_2. Clearly, we can repeat this construction as many times as needed.

Since the length of the arc gets at least twice as small on each step, after several steps the obtained disc D_n will lie completely in F_{n-1} and therefore in F. $\square$

A straightforward modification of the above proof gives the following.

6.18. Theorem. *Suppose α is a closed simple smooth regular plane curve. Denote by F and G the two closed domains bounded by α, say F is bounded and G is unbounded. Then α has at least 2 osculating circlines that lie in F and 2 osculating circlines that lie in G.*

Note that Theorem 6.17 as well as the Four-vertex theorem follow from Theorem 6.18; the first implication is evident and the second follows from Exercise 6.13.

Chapter 7

Surfaces

7.1 Embedded 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 subset $\Sigma \subset \mathbb{R}^3$ 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, any point $p \in \Sigma$ admits a neighborhood U such that in some coordinate system (x, y, z), the intersection $W = U \cap \Sigma$ can be written as a graph $z = f(x, y)$ of a smooth function f defined in an open domain of the (x, y)-plane.

Once we get a local representation of the surface by a graph, we can change it using the Proposition 7.1 below.

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 if 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 $\mathbb{S}^2$ 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. Therefore the unit sphere 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

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

is a surface which is complete.

If a complete surface Σ is compact and connected, then it is called *closed surface* (the term *closed set* is not directly relevant).

If a complete surface Σ is noncompact and connected, 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.

7.2 Tangent plane

Let $z = f(x,y)$ be a local graph realization of a surface. Assume $p = (x_p, y_p, z_p)$ lies on this graph, so $z_p = f(x_p, y_p)$. The plane passing thru p and spanned by the vectors $(\frac{\partial}{\partial x} f)(x_p, y_p)$ and $(\frac{\partial}{\partial y} f)(x_p, y_p)$ is called the *tangent plane* of Σ at p. It can be interpreted as the best approximation of the surface Σ by a plane at p.

The tangent plane to Σ at p is usually denoted by T_p or $\mathrm{T}_p\Sigma$.

It is straightforward to check that the tangent plane does not depend of the local presentation of Σ by a graph.

On local graph representations. The following proposition guarantees the existence of a local graph representation near a given point.

7.1. Proposition. *Assume the tangent of a smooth surface Σ at point p is not perpendicular to the (x, y)-plane. Then a neighborhood of p in Σ can be presented as a graph of smooth function $z = f(x, y)$ defined on an open set of the (x, y)-plane.*

A reader familiar with the inverse function theorem, can consider this proposition as an exercise.

Special coordinate system. 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.

According to Proposition 7.1, 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.

This gives 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.

7.3 Curvatures

Hessian. Fix a point p on a smooth surface Σ and the associated special coordinate system.

Consider the Hessian matrix

$$M_p = \begin{pmatrix} (\frac{\partial^2}{\partial x^2} f)(0,0) & (\frac{\partial^2}{\partial x \partial y} f)(0,0) \\ (\frac{\partial^2}{\partial y \partial x} f)(0,0) & (\frac{\partial^2}{\partial y^2} f)(0,0) \end{pmatrix}.$$

This is a symmetric matrix, therefore by an appropriate rotation of the (x, y)-plane, we can make it diagonal; that is, we can assume that $(\frac{\partial^2}{\partial x \partial y} f)(0,0) = 0$. Then the diagonal elements are called *principle curvatures* of Σ at p; they are uniqueley defined up to sign; They are denoted as $k_1(p)$ and $k_2(p)$. The principle curvatures can be also defined as the eigenvalues of M_p.

The determinant of M_p is $k_1(p){\cdot}k_2(p)$; it is called the *Gauss curvature* of Σ at p. The trace of M_p is $k_1(p)+k_2(p)$; it is called the *mean curvature* of Σ at p.

Form the discussion above, we get that the Gauss curvature depends only on Σ and p, and not on the choice of the coordinate system. Up to sign, the same obsevation is true for the principle curvatures and the mean curvature.

7.2. Exercise. *Assume Σ is a closed surface of with principle curvatures at most 1 and let F be its orthogonal projection to a plane. Show that no circle of curvature larger than 1 can support F from the left.*

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

7.4. Exercise. *Assume a closed surface Σ bounds a convex body. Show that Σ is a sphere with nonnegative Gauss curvature.*

7.4 Immersed surfaces

Parametrizations. A surface can be described by a map 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 $(x,y,z)\mapsto(a{\cdot}x,b{\cdot}y,c{\cdot}z)$ to the unit sphere $\mathbb{S}^2$.

For a surface Σ, a map $s:\Sigma\to\mathbb{R}^3$ is called a *parametrized surface* if it is smooth and regular in the sense defined below. Σ will be called the *domain of parameters* and s the *parametrization*.

Assume Σ is written locally as a graph $z=f(x,y)$ in some coordinate system. Let $\tilde f$ be defined on the same domain as f as $\tilde f(x,y)=(x,y,f(x,y))$.

The map $s\colon\Sigma\to\mathbb{R}^3$ is smooth if for the composition $s\circ\tilde f$, all partial derivatives $\frac{\partial^{m+n}}{\partial x^m\partial y^n}(s\circ\tilde f)$ exist and are continuous in the domain of definition of f.

The map $s\colon\Sigma\to\mathbb{R}^3$ is regular if the vectors $\frac{\partial}{\partial x}(s\circ\tilde f)$ and $\frac{\partial}{\partial y}(s\circ\tilde f)$ are linearly independent at each point of the domain of f.

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.

Immersed surfaces. The parametric definition allows the surfaces to have self-intersections, hence it is more general. The surfaces with possible self-intersections are called *immersed*.

In the described example $\mathbb{S}^2$ is the *domain of parameters* of the surface. We can say that the surface $\Sigma_{a,b,c}$ is a *sphere* since it has the sphere as the domain of parameters.

We may use other domains of parameters, the torus or the sphere with two handles or for surfaces with boundary, the disc, the annulus, the Möbius band and so on. The sphere with n handles is also called the *surface of genus n*.

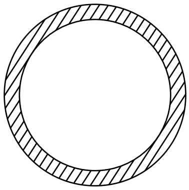
The set of parameters can be more complicated, for example the projective plane — a sphere where opposite points are identified; such set of parameters can not be realized as an embedded surface in $\mathbb{R}^3$, but it can be embedded in a higher dimensional Euclidean space. Another example is the Klein bottle — the nonoriented brother of the torus; it also can not be embedded in the Euclidean space, but it can be immersed with a self-intersection along a closed smooth curve.

Chapter 8

Bounded principle curvatures

Note that there sets in $\mathbb{R}^3$ bounded by a closed surface Σ with principle curvatures at most 1 by absolute value that do not contain a ball of radius 1.

For example the region between two large spheres with similar radii and centers close to each other. This region can be made arbitrary thin and the curvature of the boundary can be made arbitrary close to zero.



The same example works in the plane — a pair of circles with arbitrary small curvature can bound an arbitrary thin region.

8.1. Advanced exercise. *Suppose a set $V \subset \mathbb{R}^3$ is bounded by a closed surface Σ with principle curvatures bounded in absolute value by 1. Assume that V does not contain a ball of radius $\frac{1}{100}$. Show that Σ has two components of the same topological type; that is, both can be written in parametric form with the same parameter domain.*

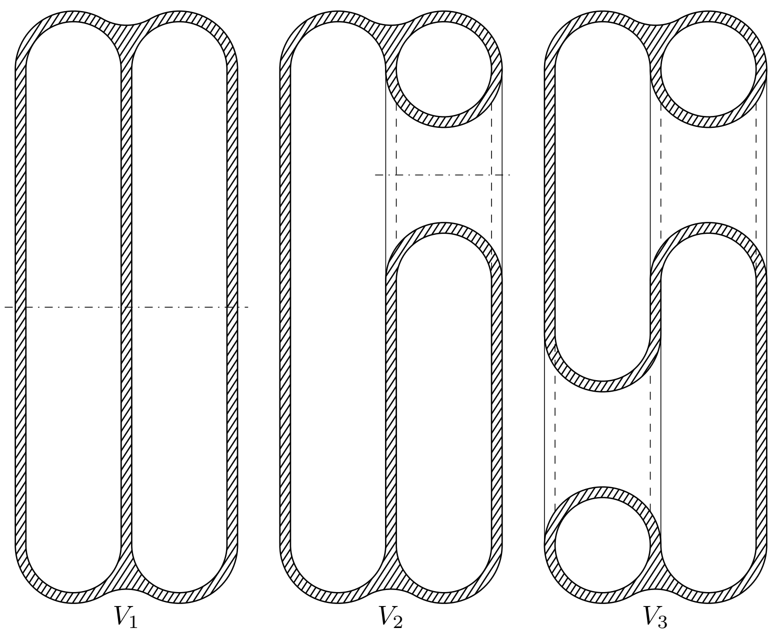
The same example would work for curves if we allow the boundary of the plane figure to be not connected. The following question might look like a right 3-dimensional analog of the moon in a puddle problem (6.17).

8.1 Lagunov's example

8.2. Question. *Assume a set $V \subset \mathbb{R}^3$ is bounded by a closed connected surface Σ with principle curvatures bounded in absolute value by 1. Is it true that V contains a ball of radius 1?*

It turns out that the answer is "no". The following example was constructed by Vladimir Lagunov [17].

Construction. Let us start with a body of revolution V_1 with cross section shown on the diagram. The boundary curve of the cross section consists of 6 vertical line segments smoothly jointed into 3 closed simple curves. The boundary of V_1 has 3 components, each of which is a sphere.



A simple computation shows that if the curvature of all curves is at most 1 then the boundary surface of V_1 has priniple curvatures at most 1 in absolute value.

At most of the places V_1 can be made arbitrary thin, the only thick place is where all tree spheres come together; it could be arranged that the radius of the maximal ball is just a little bit above

$$r_2 = \tfrac{2}{\sqrt{3}} - 1 < \tfrac{1}{6}.$$

This is the radius of the smallest circle tangent to three unit circles that are tangent to each other.

It remains to modify V_1 to make its boundary connected without alloing larger balls inside.

Note that each sphere in the boundary contains two flat discs; they come into pairs close lying to each other. Let us drill thru two of such pairs and reconnect the holes by another body of revolution whose axis is shifted but stays parallel to the axis of V_1. Denote the obtained body by V_2; its cross section of the obtained body is shown on the diagram.

Then repeat the operation for the other two pairs. Denote the obtained body by V_3; the cross section of the obtained body is shown on the diagram.

It is easy to see that the boundary of V_3 is connected and assuming that the holes are large its boundary can be made so that its principle curvatures are still at most 1. □

8.3. Claim. *The surface of V_3 has genus 2.*

Proof. Note that the boundary of V_1 consists of three spheres.

When we drill a hole, we make one hole in two spheres and two holes in one shpere. We reconnect two spheres by a tube and obtain one sphere. Connecting the two holes of the other sphere by a tube we get a torus.

At the second operation we make a torus from the remaining sphere and connect it to the other torus by a tube. This way we get a sphere with two handles; that is, it has genus 2. □

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

8.5. Exercise. *Assume V is a convex body in $\mathbb{R}^3$ bounded by a surface with principle curvatures at most 1. Show that V contains a unit ball.*[1]

8.6. Exercise. *Modify Lagunov's construction to make the boundary surface a sphere with 4 handles.*[2]

8.7. Advanced exercise. *Show that the bound in the Lagunov's example is optimal. That is, if a body $V \subset \mathbb{R}^3$ is bounded by a connected surface Σ with principle curvatures at most 1, then V contains a ball of radius r_2.*

[1] Hint: Consider a maximal ball in V and apply Exercise 7.2 for a right choice of projection.

[2] Hint: Drill an extra hole or combine two examples together.

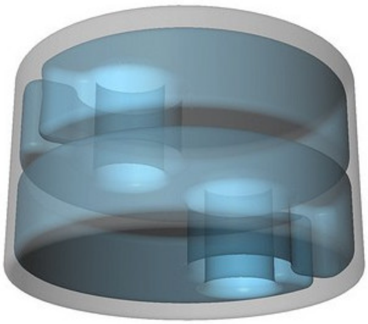
8.2 On embedded sphere

8.8. Advanced exercise. *Note that the body V in the example of Lagunov is constructed by thikening a surface having a singular curve where the surface self-intersects at an angle of* $120°$*. Show that this way one can not obtain a body bounded by a sphere.*

In fact one can show that if a body $V \subset \mathbb{R}^3$ is bounded by a sphere Σ with principle curvatures at most 1, then V contains a ball of radius $r_3 = \sqrt{\frac{3}{2}} - 1 > \frac{1}{5}$, which is the radius of the smallest sphere tangent to three unit spheres that are tangent to each other. Moreover, this bound is optimal.

An example of such a body can be obtained by thickening the so called Bing's house. It is certainly a surface whose singularities are formed by three curves meeting at two points; four ends at each point. The remaining surface of Bing's house is smooth and has bounded principle curvatures; we can assume that they are bounded by an arbitrary small number.

At the singular curves the three pieces of surface have to cross at angles $\frac{2}{3}\cdot\pi$ and at the sigular points 6 pieces of surface should come together forming 6 tringles with vertex in the center of a regular tetrahedron and the bases at its 6 edges. Thickening of a sufficiently large Bing's house of that type produces the optimal bound r_3 on the maximal ball that it contains.



The thickening of Bing's house shown on the picture can not give the optimal bound, but still it can produce an example of an embedded sphere that does not surround a ball of radius 1.

This picture is very similar to the Lagunov's example described above — it can be obtained by filling the rings in the section of V_3 by thickened discs.

This picture was taken from a post of Ken Baker [18]; this post has many other beautiful pictures that help to visualize Bing's house.

Chapter 9

Convex surfaces

9.1 Embedded surfaces

A set X in the Euclidean space 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. Clearly any open convex set is stricly convex; the closed cube (as well as any convex polyhedron) gives an example of a convex set which is not strictly convex.

9.1. Exercise. *Let Σ be a surface with positive Gauss curvature. Show that for any point $p \in \Sigma$ and all sufficiently small $\varepsilon > 0$, the surface Σ divides the ball $B(p, \varepsilon)$ in two regions, one of which is strictly convex.*

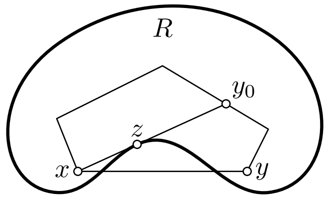
The following theorem gives a global version of the above exercise.

9.2. Theorem. *Assume Σ is a closed or open smooth connected surface with positive Gauss curvature. Then Σ bounds a convex region R. Moreover, if Σ is closed then it is a sphere; that is, Σ admits a smooth regular parametrization by $\mathbb{S}^2$.*

Proof. By Exercise 9.1, one of the regions, say R, bounded by Σ is locally strictly convex; that is, at any point p of R, the intersection of R with a sufficiently small ball centered at p is strictly convex.

Since Σ is connected, so it R. Moreover any two points x and y in the interior of R can be connected by a polygonal line β in the interior of R.

Arguing by contradiction, assume the line segment $[xy]$ does not lie in the in the interior of R. Let y_0 be the first point on β so that the line segment $[xy_0]$ intersects Σ; assume $[xy_0]$ touches Σ at a point z.



By Exercise 9.1, $R \cap B(z, \varepsilon)$ is strictly convex for all sufficiently small $\varepsilon > 0$. On the other hand z lies between two points both in the line segment $[xy_0]$ and $R \cap B(z, \varepsilon)$ — a contradiction.

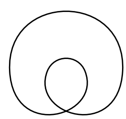
It remains to parameterize Σ by $\mathbb{S}^2$. Fix a point p in the interior of R. By strict convexity of R, for any point $x \in \mathbb{S}^2$ there is unique point $x' \in \Sigma$ that lies on the halfline px; moreover, the map $h \colon x \mapsto x'$ describes a bijection $\mathbb{S}^2 \to \Sigma$.

Applying inverse function theorem in local coordinates of $\mathbb{S}^2$ and Σ, we get that the map h is smooth and regular. Hence the result. $\square$

9.2 Immersed surfaces

It seems that first formulation and proof of the following theorem was given by James Stoker [19] who attributed it to Jacques Hadamard, who proved a closely relevant statement in [20, item 23].

9.3. Theorem. *Any closed connected immersed surface with positive Gauss curvature is embedded.*



In other words such surface can not have self-intersections. Note that an analogous statement does not hold in the plane; on the diagram you can see a closed curve with a self-intersection and positive curvature at all points. Exercise 6.6 gives a condition that guarantees simplicity of a locally convex plane curve; it will be used in the following proof.

Before going into the proof, note that theorems 9.3 and 9.2 imply the following:

9.4. Corollary. *Any closed connected immersed surface with positive Gauss curvature is an embedded sphere that bounds a convex region.*

In the following sections we will give one complete proof and sketch an alternative proof.

The first proof uses a *Morse-type argument* for the height function; that is, we study how the part of the surface that lies below a plane changes when we move the plane upward. Little more careful analysis of this changes would imply the corollary above directly, without using Theorem 9.2.

The sketch use equidistants surfaces and the Gauss map. We will not proove a topological statement relying on intuition.

In the proof we abuse notation slightly; we say a *point of the immersed surface* instead of a *point in the parameter domain of the immersed surface.* So each point of self-intersection is considered as two or more "distinct" points of the surface.

9.3 Morse-type proof

Let Σ be a closed surface with positive Gauss curvature, possibly with self-intersections.

Fix a horizontal plane Π_h defined by the equation $z = h$ in an (x, y, z)-coordinate system. Note that the intersection $W_h = \Sigma \cap \Pi_h$ is formed by a finite collection of closed curves and isolated points. (These curves and isolated points might intersect in the Euclidean space, but they are disjoint in the domain of parameters of Σ.)

Indeed, if $\mathrm{T}_p = \Pi_h$, then, since the principle curvatures are positive, p is a local minimum or local maximum of the height function. In both cases, p is an isolated point of W_h in Σ. If the tangent plane T_p is not Π_h, then it is not perpendicular to (x, z)-plane or (y, z)-plane. Therefore by Proposition 7.1, the surface can be written locally as a graph $x = f(y, z)$ or $y = f(x, z)$; in both cases p lies on the curve $x = f(y, h)$ or correspondingly $y = f(x, h)$.

Summarizing, the closed set $W_h \subset \Sigma$ locally looks like a curve or an isolated point. Since Σ is compact, so is W. Therefore W is a finite disjoint collection of isolated points and closed simple curves in Σ.

Assume α_{h_0} is a closed curve in W_{h_0}. Note that its neighborhood is swept by curves α_h in W_h for $h \approx h_0$. Indeed a neighbohood of α_{h_0} in Σ can be covered by a finite number of graphs of the type $x = f(y, z)$ (or $y = f(x, z)$) and the curves α_h can be described locally as curves of the form $t \mapsto (f(t, h), t, h)$ (or correspondingly $t \mapsto (t, f(t, h), h)$) for $h \approx h_0$.

As α_h is the intersection of a locally convex surface with a plane, the curvature of α_h has fixed sign; so if we choose an orientation for the curves properly, we can assume that they all have positive curvature.

The family α_h depends smoothly on h and the same holds for its tangent indicatrix. Therefore the total signed curvature K_h of α_h depends continuously on h. If $K_h = 2{\cdot}\pi$ for some h, then $K_h = 2{\cdot}\pi$ for every h. It follows since, the function $h \mapsto K_h$ is continuous and its value is a multiple of $2{\cdot}\pi$. In this case, by Exercise 6.6, all curves α_h are simple and each one bounds a convex region in the plane Π_h.

Summarizing, if one of the curves in the constructed family α_h is simple, then each curve in the family is simple and each α_h bounds a convex region in the plane Π_h.

Choose a point $p \in \Sigma$ that minimizes the height function z. Without loss of generality we may assume that p is the origin and therefore the surface lies in the upper half-space.

Fix $h > 0$. The intersection of the set $z \leqslant h$ with the surface may contain several connected components; one of them contains p, denote this component by Σ_h.[1]

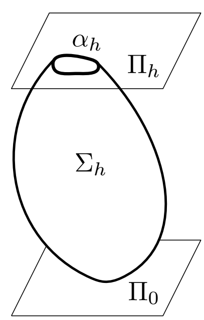
From above, Σ_h is a surface with possibly nonempty boundary. Indeed it might be bounded only by few closed curves in W_h; any isolated point of W_h either lies in Σ_h together with its neighborhood or does not lie in Σ_h.

Note that for small values of h, the surface Σ_h is an embedded disc. Indeed, if $z = f(x, y)$ is a graph represetation of Σ around p, then Σ_h is a graph of f over

$$\Delta = \left\{ (x, y) \in \mathbb{R}^2 \ : \ f(x, y) \leqslant h \right\}.$$

Since the Gauss curvature is positive, the function f is convex and therefore Δ is convex and bounded by a smooth curve; any such set can be parameterized by a disc.

Let $H > 0$ be the maximal value such that Σ_h has no self-intersections for any $h < H$. For a sequence $h_n \to H^-$, choose a point q_n on the boundary of Σ_h and pass to a partial limit q of q_n in Σ; that is, q is a limit of a subsequence of (q_n).



If the tangent plane at q is *not* horizontal, then there is a closed curve α_H in Σ that passes thru q and lies on the plane $z = H$. From the above discussion, the curve α_H, as well as all α_h with $h \approx H$ are closed embedded convex curves. Hence Σ_h has no self-intersections for some $h > H$ — a contradiction.

If the tangent plane at q is horizontal, then the surface Σ_H has no boundary. Since Σ is connected, $\Sigma_H = \Sigma$. Since Σ_h has no self-intersections for $h < H$, we get that Σ is an embedded surface. □

9.5. Exercise. *Show that any open immersed surface with positive Gauss curvature is embedded.*[2]

[1] These components might intersect in the space, but they are disjoint in the domain of parameters. Note also that from the corollary, it follows that there is only one component Σ_h, but we can not use it before the theorem is proved.

[2] Hint: Modify the proof of the theorem.

9.4 Proof via equidistant surfaces

Recall that a surface Σ is called *orientable* if one can choose at each point p of the surface a unit normal vector $\nu(p)$ in such a way that the function $p \mapsto \nu_p$ is continuous in every chart of Σ. For immersed surfaces we should say that ν is a continuous function defined on the parameter domain of the surface. The map ν is called a *Gauss map* of the surface.

9.6. Claim. *Assume Σ is a closed immersed surface with positive Gauss curvature, then it is orientable.*

Proof. Indeed we can choose the unit normal vector $\nu(p)$ in such a way that both principle curvatures are positive. Iin this case the surface lies locally on the side of tangent plane T_p which is opposite from $\nu(p)$.

Evidently this choice is continuous. □

The unit normal described in the proof of the claim will be called *outer normal.*

9.7. Lemma. *Assume Σ is a closed connected immersed surface with positive Gauss curvature. Then the Gauss map $\nu\colon \Sigma \to \mathbb{S}^2$ has a smooth regular inverse; in particular, Σ is a sphere.*

This lemma follows from two facts: (1) if Gauss curvature does not vanish then the Gauss map is regular, in particular this map has a local inverse at each point and (2) the sphere $\mathbb{S}^2$ is *simply connected*; that is, $\mathbb{S}^2$ is connected any closed curve in $\mathbb{S}^2$ can be deformed continuously into a trivial curve that stays at one point. The proof is standard in topology, we hope that the statement is intuitively obvious. The reader might be able to reinvent the theory by trying to prove that if the map $\varphi\colon \mathbb{S}^2 \to \mathbb{S}^2$ is smooth and regular then it has an inverse.

Equidistant surfaces. Assume $\nu\colon \Sigma \to \mathbb{S}^2$ is a Gauss map of a smooth surface Σ. Fix a real number R and consider the map $h_R\colon \Sigma \to$ $\to \mathbb{R}^3$ defined by $h_R\colon p \mapsto p + R{\cdot}\nu(p)$. The map h_R describe the so called *equidistant surface*; it is smooth by definition, but in general it does not have to be regular.

9.8. Lemma. *Suppose $\nu\colon \Sigma \to \mathbb{S}^2$ is a Gauss map of a surface Σ. Assume the corresponding principle curvatures are nonnegative at all points. Then the equidistant surface Σ_R is regular and its principle curvatures are positive and strictly smaller than $\frac{1}{R}$.*

Proof. To prove regularity, let us use the special representation of Σ as a graph $z = f(x, y)$ with the x and y axis in the principle directions of Σ at p.[3]

Due to the choice of directions of x and y axis, for the Gauss map $g(x,y)$, we have

$$\begin{aligned}\tfrac{\partial}{\partial x}g(0,0)&=(k_1,0,0),\\ \tfrac{\partial}{\partial y}g(0,0)&=(0,k_2,0).\end{aligned}$$

Then $h_R=h_0+R{\cdot}g$; therefore

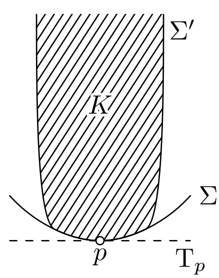
$$\begin{aligned}\tfrac{\partial}{\partial x}h_R&=(1+R{\cdot}k_1,0,0),\\ \tfrac{\partial}{\partial y}h_R&=(0,1+R{\cdot}k_2,0)\end{aligned}$$

which are linearly independent if $R\geqslant 0$ and $k_1,k_2\geqslant 0$.

If Σ bounds a convex closed set K. Then Σ_R bounds K_R — the closed R-neighborhood of K; that is, K_R is the set of all points at distance at most R from K.

Since Σ is smooth it is supported at each point p from inside by a small ball $B_\varepsilon(o)$. Then the ball $B_{R+\varepsilon}(o)$ lies in K_R and touches its boundary at the point corresponding to p. Hence the principle curvatures at p are at least $\frac{1}{R+\varepsilon}$.

In the general case, a local chart of Σ can be modified so that it has a piece of the original surface around p and bounds a convex set. Here is one way to do this:



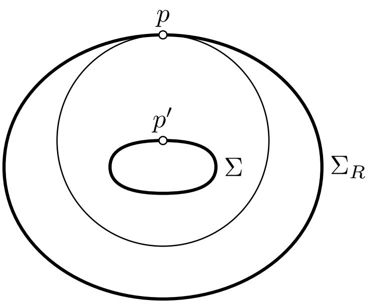
Choose a smooth function $\varphi(x)$ that is convex increasing and such that for sufficiently small $\varepsilon>0$ we have $\varphi(x)=x$ if $x<\varepsilon$ and $\varphi(x)\to\infty$ as $x\to 2{\cdot}\varepsilon$. (Such functions do exist; moreover an explicit formula can be written, but we leave it without a proof.)

Assume $z=f(x,y)$ is a special representation of Σ around p by some convex function f. Direct computations show that $h=\varphi\circ f(x,y)$ is still convex. The surface Σ' described as the graph $z=h(x,y)$ bounds a convex closed set K and the part of Σ' described by the parameters $\{\,(x,y)\;:\;f(x,y)<\varepsilon\,\}$ coincide with a neighborhood of p in Σ. Hence the general case follows. $\square$

Proof assembling. Let $s\colon\mathbb{S}^2\to\mathbb{R}^3$ be the parametization of Σ provided by Lemma 9.7. Then the equidistant surface Σ_R can be parametrized

[3]If we assume that $\nu(p)$ points in the direction of the z-axis, then Σ is given in parametric form $h_0\colon(x,y)\mapsto(x,y,f(x,y))$, where $f=-\frac{k_1}{2}{\cdot}x^2-\frac{k_2}{2}{\cdot}y^2+o(x^2+y^2)$.

by $s_R(u)=s(u)+R{\cdot}u$ for $u\in\mathbb{S}^2$. Rescaling s_R by a factor of $\frac{1}{r}$ we get the map $u\mapsto\frac{1}{R}{\cdot}s(u)+u$ which converges smoothly to the indentity map on the sphere $\mathbb{S}^2$. Therefore Σ_R is embedded for sufficiently large R.



Applying Theorem 9.2, we get that Σ_R bounds a convex set.

By Lemma 9.8, the principle curvatures of Σ_R are smaller than $\frac{1}{R}$. Therefore the same idea as in Exercise 8.5 shows that any point p of Σ_R can be supported by a ball of radius R from inside. Note that the center p' of such ball has to lie on Σ; indeed it lies at distance R in the normal direction. In other words, the map $s_0(u)=s_R(u)-R{\cdot}u$ is injective, or equivalently Σ has no self-intersection. $\square$

Chapter 10

Geodesics

The following exercise might look like a hard 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?

10.1 Closest point projection

10.2. Lemma. *Let K be a closed convex set in $\mathbb{R}^3$. Then for every point $p \in \mathbb{R}^3$ there is unique point $\bar{p} \in K$ that minimizes the distance $|p - x|$ among all points $x \in K$.*

Moreover the map $p \mapsto \bar{p}$ is short; that is,

❶ $$|p - q| \geqslant |\bar{p} - \bar{q}|$$

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 K. Note that if $p \in K$, then $\bar{p} = p$.

Proof. Fix a point p and set

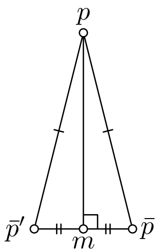
$$\ell = \inf_{x \in K} \{|p - x|\}.$$

Choose a sequence $x_n \in K$ 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 K is closed $\bar{p} \in K$. By construction

$$|p - \bar{p}| = \ell = \lim_{n\to\infty} |p - x_n|.$$

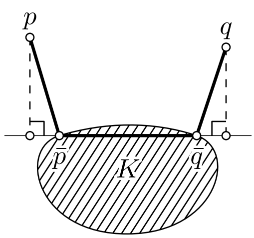
Hence the existence follows.



Assume there are two distinct points $\bar{p}, \bar{p}' \in K$ that minimize the distance to p. Since K is convex, their midpoint $m = \frac{1}{2}\cdot(\bar{p} + \bar{p}')$ lies in K. 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.

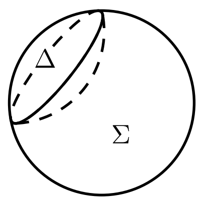
We can assume that $\bar{p} \neq \bar{q}$, otherwise there is nothing to prove. Note that if $p \neq \bar{p}$ (that is, if $p \notin K$), then $\angle p\bar{p}\bar{q}$ is right or obtuse. Otherwise there would be a point x on the line segment $[\bar{q}, \bar{p}]$ that is closer to p than $\bar{p}$. Since K is convex, the line segment $[\bar{q}, \bar{p}]$ and therefore x lie in K. Hence $\bar{p}$ is not closest to p — a contradiction.



The same way we can show that if $q \neq \bar{q}$, 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

$$|p - q| \geqslant |\bar{p} - \bar{q}|. \qquad \square$$

10.2 Geodesics



Let Σ be a surface. Assume γ is a constant speed curve in Σ that connects two points $p, q \in \Sigma$ and minimizes the length among all such curves. Then γ is called a *minimizing geodesic* from p to q.

10.3. Exercise. *Suppose Σ is a smooth closed surface that bounds a convex body K in $\mathbb{R}^3$ and Π is a plane that cuts from Σ a disk Δ. Assume that the reflection of Δ with respect to Π lies inside of Σ. Show that Δ is* convex *with respect to the intrinsic metric of Σ;*

that is, if both ends of a minimizing geodesic in Σ lie in Δ, then the entire geodesic lies in Δ.

Recall that the *diameter* of a set K in the Euclidean space is defined as the exact uppper bound on the distances between pairs of points in K. Let us define the *intrinsic diameter* of a closed surface Σ as the exact upper bound on the lengths of minimizing geodesics in the surface.

10.4. Exercise. *Assume that a closed smooth surface Σ bounds a convex body K of diameter D.*

- *Show that the intrinsic diameter of Σ cannot exceed $\pi\cdot D$.*
- *Show that the area of Σ can not exceed the area of the sphere of radius D.*

10.3 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 [21].

10.5. 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.

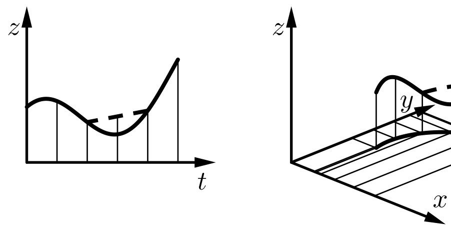
Proof. Denote the graph by Σ. Choose a partition such that $\gamma|_{[t_{i-1},t_{i+1}]}$ is minimizing. If the function z is convex on each interval $[t_{i-1},t_{i+1}]$, then it is convex on the entire interval. Therefore it is sufficient to prove the case when $\gamma\colon [a,b] \to$ is a minimizing 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

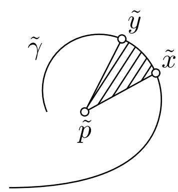
$$K = \{\,(x,y,z)\in\mathbb{R}^3\,:\,(x,y)\in\Delta,\ z\leqslant f(x,y)\,\};$$

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. □

10.6. 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.

10.4 Bound on total curvature

10.7. 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 [22], later David Berg [23] 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

$$W = \{\,(x,y,z)\in\mathbb{R}^3\,:\,z\geqslant f(x,y)\,\}$$

Recall that by definition, geoedesics have constant speed. In the proof we will use the following claim; we do not prove it, but make a physical interpretation which might help to believe it.

10.8. Claim. *Let Σ be a smooth regular surface. A curve γ in Σ is a geodesic if and only if it is a smooth constant-speed curve such that for any t the second derivative $\gamma''(t)$ is normal to the tangent plane $\mathrm{T}_{\gamma(t)}\Sigma$.*

Moreover given an initial point $p \in \Sigma$ and a tangent vector $v \in \mathrm{T}_p\Sigma$ there is unique geodesic γ defined on a maximal open interval containing zero such that $\gamma(0) = p$ and $\gamma'(0) = v$.

This claim provides connection between intrinsic geometry of the surface and its extrinsic geometry which will be important latter. 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.

This connection will play an important role in the proof of the remarkable theorem (see 12.12).

For instance geodesic γ is an object of intrinsic geometry of the surface Σ, while its second derivative γ'' is not. 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 tension 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 $\mathrm{T}_{\gamma(t)}\Sigma$.[1]

The third statement can be also understood using physical intuition — $\gamma(t)$ is the trajectory of a particle that slides on Σ without friction and with initial velocity v. Formally, existence and uniqueness follows from Picard's theorem (the fundamental theorem of ordinary differential equations). □

10.9. Exercise. *Show that two minimizing geodesics in a smooth regular surface can not have more than one point of intersections.*

10.10. Exercise. *Assume that a smooth regular surface Σ is mirror symmetric with respect to a plane Π. Show that no minimizing geodesic in Σ can cross Π more than once.*

10.11. 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. □

Proof of 10.7. Let $\gamma(t) = (x(t), y(t), z(t))$ be a unit speed geodesic on Σ. According to Liberman's lemma and Lemma 10.11, $z(t)$ 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

$$\begin{aligned}\int_a^b |z''(t)| &= z'(a) - z'(b) \leqslant \\ &\leqslant 2\cdot\tfrac{\ell}{\sqrt{1+\ell^2}}.\end{aligned}$$

[1] In fact $\gamma''(t) + \nu\cdot\langle s(\gamma'(t)), \gamma'(t)\rangle = 0$, where s is the shape operator of Σ at $\gamma(t)$ or equivalently, $\gamma''(t) + \nu\cdot \mathrm{I\!I}(\gamma'(t)), \gamma'(t)) = 0$, where $\mathrm{I\!I}$ is the second fundamental form of Σ at $\gamma(t)$.

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,

$$|\gamma''(t)| \leqslant \sqrt{1+\ell^2}|z''(t)|.$$

□

10.12. 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.*[2]

In fact there is a fixed bound on the total curvature of any minimizing geodesic on any closed convex surface [24].

[2] Hint: Use Exercise 10.4.

Chapter 11

Parallel transport

11.1 Parallel fields

Let Σ be a smooth regular surface in the Euclidean space and $\alpha\colon [a,b]\to$ $\to\Sigma$ be a smooth curve. A smooth vector-valued function $t\mapsto v(t)$ is called a *tangent field* on α if the vector $v(t)$ lies in the tangent plane $\mathrm{T}_{\alpha(t)}\Sigma$ for each t.

A tangent field $v(t)$ on α is called *parallel* if $v'(t)\perp\mathrm{T}_{\alpha(t)}$ for any t.

In general the family of tangent planes $\mathrm{T}_{\alpha(t)}\Sigma$ is not parallel. Therefore one can not expect to have a truly parallel family $v(t)$ with $v'\equiv 0$. The condition $v'(t)\perp\mathrm{T}_{\alpha(t)}$ means that this family is as parallel as possible, it rotates together with the tangent plane, but does not rotate inside the plane.

Note that according to Claim 10.8, for any geodesic γ, the velocity field $v(t)=\gamma'(t)$ is parallel along γ.

11.1. Exercise. *Let Σ be a smooth regular surface in the Euclidean space, $\alpha\colon [a,b]\to\Sigma$ a smooth curve and $v(t)$, $w(t)$ parallel vector fields along α.*

(a) Show that $|v(t)|$ is constant.

(b) Show that the angle $\theta(t)$ between $v(t)$ and $w(t)$ is constant.

11.2 Parallel transport

Assume $p=\gamma(a)$ and $q=\gamma(b)$. Given a tangent vector $v\in\mathrm{T}_p$ there is unique parallel field $v(t)$ along α such that $v(a)=v$. The latter follows from Picard's theorem; the uniqueness also follows from Exercise 11.1.

The vector $v(b)\in\mathrm{T}_q$ is called the *parallel transport* of v along α and denoted as $\iota_\alpha(v)$.

As it follow from the Exercise 11.1, parallel transport $\iota_\alpha\colon \mathrm{T}_p \to \mathrm{T}_q$ is an an isometry; it depends on the choice of α — for another curve β connecting p to q in Σ, the parallel transport $\iota_\beta\colon \mathrm{T}_p \to \mathrm{T}_q$ might be different.

To interpret the parallel transport physically, think of walking along α and carrying a perfectly balanced bike wheel in such a way that you touch only its axis keeping it normal to Σ. It should be physically evident that if the wheel is non-spinning at the starting point p, then it will not be spinning after stopping at q.[1] The map that sends the initial position of the wheel to the final position is the parallel transport ι_α.

This physical interpretation was suggested by Mark Levi [25]; it will be used further.

On a more formal level, one can choose a partition $a = t_0 < \ldots < t_n = b$ of $[a,b]$ and consider the sequence of orthogonal projections $\varphi_i\colon \mathrm{T}_{\alpha(t_{i-1})} \to \mathrm{T}_{\alpha(t_i)}$. For a fine partition, the composition

$$\varphi_n \circ \cdots \circ \varphi_1\colon T_p \to \mathrm{T}_q$$

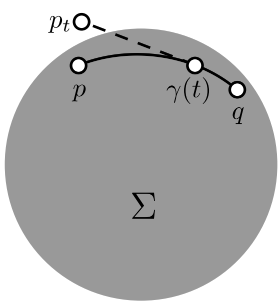
gives an approximation of ι_α.[2]

11.2. Advanced exercise. *Let Σ be a smooth closed strictly convex surface in $\mathbb{R}^3$ and $\gamma\colon [0,\ell] \to \Sigma$ be a unit-speed minimizing geodesic. Set $p = \gamma(0)$, $q = \gamma(\ell)$ and*

$$p_t = \gamma(t) - t\cdot\gamma'(t),$$

where $\gamma'(t)$ denotes the velocity vector of γ at t.

Show that for any $t \in (0,\ell)$, one cannot *see q from p_t; that is, the line segment $[p_tq]$ intersects Σ at a point distinct from q.*[3]



11.3 Geodesic curvature

Suppose Σ is a smooth regular surface. Assume Σ is oriented; in this case terms “left” and “right” can be used the same sense as in the plane.

[1] Indeed, by pushing axis one can not produce torque to spin the wheel.

[2] Each φ_i does not increase the magnitude of a vector and neither the composition. It is straightforward to see that if if the partition is sufficiently fine, then it almost preserves the composition almost preserves magnitudes.

[3] Hint: Show that the concatenation of the line segment $[p_t\gamma(t)]$ and the arc $\gamma|_{[t,\ell]}$ is a minimizing geodesic in the closed set W outside of Σ.

Broken geodesics. For a concatenation of two geodesics in Σ, let us define signed external angle at their common point; the sign is positive if it turns left and negative if it turns right.

A concatenation of minimizing geodesics in Σ will be called *broken geodesic.*

The sum of the signed external angles for a broken geodesic γ in Σ will be called *total geodesic curvature* of γ; it will be denoted as Θ_γ or $\Theta_{\gamma,\Sigma}$ if we need to emphasize that γ is a curve in Σ.

Note that if we change the orientation of the curve, then the total geodesic curvature changes sign.

Smooth regular curves. The total geodesic curvature can be also defined for a smooth unit-speed curve $\gamma\colon [a,b]\to\Sigma$.

Let $\nu\colon \Sigma\to\mathbb{S}^2$ be the Gauss map that defines the orientation on Σ. Then for any t the vectors $\nu(t)=\nu(\gamma(t))$ and the velocity vector $\tau(t)=\gamma'(t)$ are unit vectors that are normal to each other. Denote by $\mu(t)$ the unit vector that is normal to both $\nu(t)$ and $\tau(t)$ that points to the left from γ. Note that the triple $\tau(t),\mu(t),\nu(t)$ is an orthogonal basis for any t.

Since γ is unit-speed, the acceleration $\gamma''(t)$ is perpendicular to $\tau(t)$; therefore at any parameter value t, we have

$$\gamma''(t)=k_g(t)\cdot\mu(t)-k_n(t)\cdot\nu(t),$$

for some real numbers $k_n(t)$ and $k_g(t)$. The numbers $k_n(t)$ and $k_g(t)$ are called *normal* and *geodesic curvature* of γ at t correspondingly.

Note that the geodesic curvature vanishes if γ is a geodesic. It measures how much a given curve diverges from being a geodesic; it is positive if γ turns left and negative if γ turns right.

The total geodesic curvature of γ can be defined as the integral of its geodesic curvature

$$\Theta_\gamma=\int\limits_a^b k_g(t)\cdot dt.$$

If γ is a regular curve then one has to parameterize it by arc length and then apply the definition above.

The given two definitions for regular curve and broken geodesic agree in the following sense. If β_n is a sequence of inscribed broken geodesics in γ for finer and finer partitions, then

$$\Theta_{\beta_n}\to\Theta_\gamma,$$

where the left and right hand sides are defined using the first and the second definitions above. The proof is straightforward, it can be done the same way as the case in the plane.

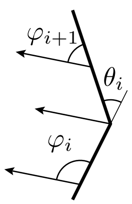
11.3. Exercise. *Let γ be a smooth unit-speed curve in a smooth regular surface Σ with a Gauss map ν. Show that*

$$k_n(t) = \langle\gamma'(t), \nu'(t)\rangle.$$

Piecewise smooth curves. One could also combine both definitions to define total geodesic curvature for piecewise smooth curve γ in Σ; that is a concatenation of smooth regular curves. We need to add the total geodesic curvature of all the edges of γ and the signed external angle at each vertex.

11.4. Proposition. *Assume γ is a closed broken geodesic in a smooth oriented surface Σ that starts and ends at the point p. Then the parallel transport $\iota_\gamma\colon T_p \to \mathrm{T}_p$ is a rotation of the the plane T_p clockwise by angle Θ_γ.*

Moreover, the same statement holds for smooth closed curves and piecewise smooth curves.



Proof. Assume γ is a cyclic concatenation of geodesics $\gamma_1, \dots, \gamma_n$. Fix a tangent vector v at p and extend it to a parallel vector field along γ. Since $w_i(t) = \gamma_i'(t)$ is parallel along γ_i, the angle φ_i between v and w_i stays constant on each γ_i.

If θ_i denotes the external angle at this vertex of switch from γ_i to γ_{i+1}, we have that

$$\varphi_{i+1} = \varphi_i - \theta_i \pmod{2\cdot\pi}.$$

Therefore after going around we get that

$$\varphi_{n+1} - \varphi_1 = -\theta_1 - \dots - \theta_n = -\Theta_\gamma.$$

Hence the the first statement follows.

For the smooth unit-speed curve $\gamma\colon [a, b] \to \Sigma$, the proof is analogous. If $\varphi(t)$ denotes the angle between $v(t)$ and $w(t) = \gamma'(t)$, then

$$\varphi'(t) + k_g(t) \equiv 0$$

Whence the angle of rotation

$$\begin{aligned}\varphi(b) - \varphi(a) &= \int_a^b \varphi'(t)\cdot dt = \\ &= -\int_a^b k_g\cdot dt = \\ &= -\Theta_\gamma\end{aligned}$$

The case of piecewise regular smooth curve is a straightforward combination of the above two cases. □

Chapter 12

Gauss–Bonnet formula

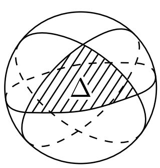
12.1 Signed area on the sphere

12.1. Lemma. *Let Δ be a spherical triangle; that is, Δ is the intersection of three closed half-spheres in the unit sphere $\mathbb{S}^2$. Then*

❶ $$\operatorname{area}\Delta = \alpha + \beta + \gamma - \pi,$$

where α, β and γ are the angles of Δ.

The value $\alpha + \beta + \gamma - \pi$ is called *excess* of the triangle Δ.



Proof. Recall that

❷ $$\operatorname{area}\mathbb{S}^2 = 4{\cdot}\pi.$$

Note that the area of a spherical slice S_α between two meridians meeting at angle α is proportional to α. Since for S_π is a half-sphere, from ❷, we get $\operatorname{area} S_\pi = 2{\cdot}\pi$. Therefore the coefficient is 2; that is,

❸ $$\operatorname{area} S_\alpha = 2{\cdot}\alpha.$$

Extending the sides of Δ we get 6 slices: two S_α, two S_β and two S_γ which cover most of the sphere once, but the triangle Δ and its centrally symmetric copy Δ' are covered 3 times. It follows that

$$2{\cdot}\operatorname{area} S_\alpha + 2{\cdot}\operatorname{area} S_\beta + 2{\cdot}\operatorname{area} S_\gamma = \operatorname{area}\mathbb{S}^2 + 4{\cdot}\operatorname{area}\Delta.$$

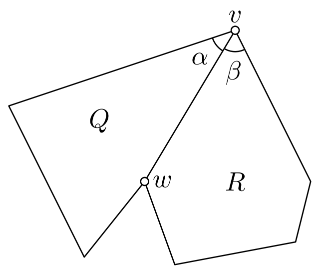
Substituting ❷ and ❸ and simplifying, we get ❶. □

If the contour $\partial\Delta$ of a spherical triangle with angles α, β and γ is oriented such that the triangle lies on the left, then its external angles are $\pi-\alpha$, $\pi-\beta$ and $\pi-\gamma$. Therefore the total geodesic curvature of $\partial\Delta$ is $\Theta_{\partial\Delta}=3\cdot\pi-\alpha-\beta-\gamma$. The identity ❶ can be rewritten as

❹ $$\Theta_{\partial\Delta}+\operatorname{area}\Delta=2\cdot\pi.$$

The formula ❹ holds for an arbitrary spherical polygon bounded by a simple broken geodesic; that is, intersection of finitely many closed half-spheres. The latter can be proved by triangulating the poygon, applying the formula for each triangle in the triangulation and summing up the results.

If a spherical polygon is P divided in two polygons Q and R by a diagonal vw then



$$\Theta_{\partial P}+2\cdot\pi=\Theta_{\partial Q}+\Theta_{\partial R}.$$

Indeed, for the internal angles Q and R at v are α and β, then their external angles are $\pi-\alpha$ and $\pi-\beta$ respectfully. The internal angle of P in this case is $\alpha+\beta$ and its external angle is $\pi-\alpha-\beta$ Clearly we have that

$$(\pi-\alpha)+(\pi-\beta)=(\pi-\alpha-\beta)+\pi;$$

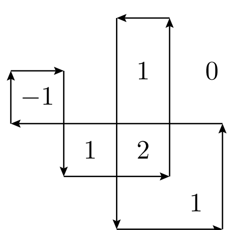
that is, the sum of external angles of Q and R at v is π plus the external angle of P at v. The same holds for the external angles at w and the rest of the external angles of P appear once on Q or R. Therefore if the formula ❹ holds for Q and R, then it holds for P.

12.2. Exercise. *Assume γ is a simple broken geodesic on $\mathbb{S}^2$ that divides its area into two equal parts. Show that $\Theta_\gamma=0$.*

Signed area. The formula ❹ holds modulo $2\cdot\pi$ for any closed broken geodesic, if one use *signed area* surrounded by curve instead of usual area; that is, we count area of the regions taking into account how many times the curve goes around the region.

Namely, we have to choose a *south pole* and state that its region has zero multiplicity. When you cross the curve the mulitplicity changes by ±1; we add 1 if the curve crosses your path from left to right and we subtract 1 otherwise. The signed area surrounded by a closed curve is the sum of area of all the regions counted with multiplicities.

Here is an example of a broken line with multiplicities assuming that the big region has the south pole inside.



This signed-area formula can be proved in a similar way: Apply the formula for each triangle with vertex at the north pole and base at each edge of the broken geodesic. Sum the resulting identities taking each with a sign: plus if the triangle lies on the left from the edge and minus if the triangle lies on the right from edge.

Choosing a different pole will change all the coefficients by the same number. So the resulting formula holds only modulo the area of $\mathbb{S}^2$, which is $4\cdot\pi$ — this will not destroy identity modulo $2\cdot\pi$.

Furthermore, by approximation, the signed-area formula holds for any reasonable curve, say piecewise smooth regular curves on the sphere. Summarizing, we hope the discussion above convinced the reader that the following statement holds.

A domain Δ in a surface is called a *disc* (or more precisely *topological disc*) if it is bounded by a closed simple curve and can be parameterized by a unit plane disc

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

That is there is a continuous bijection $\mathbb{D} \to \Delta$.

12.3. Proposition. *For any closed piecewise smooth regular curve α on the sphere, we have that*

$$\Theta_\alpha + \operatorname{area}\alpha = 0 \pmod{2\cdot\pi},$$

where area α *denotes the signed area surrounded by α and Θ_α the total geodesic curvature of α.*

Moreover, if α is a simple curve that bounds a disc Δ on the left from it, then we have

$$\Theta_\alpha + \operatorname{area}\Delta = 2\cdot\pi.$$

12.2 Gauss–Bonnet formula

12.4. Theorem. *Let Δ be a disc in a smooth oriented surface Σ bounded by a simple piecewise smooth and regular curve $\partial\Delta$ that is oriented in such a way that Δ lies on its left. Then*

❶
$$\Theta_{\partial\Delta} + \iint\limits_{\Delta} G = 2\cdot\pi,$$

where G denotes the Gauss curvature of Σ.

For geodesic triangles this theorem was proved by Carl Friedrich Gauss [26]; Pierre Bonnet and Jacques Binet independently generalized the statement for arbitrary curves. The modern formulation described below was given by Wilhelm Blaschke.

Remarks; (1). For a general compact domain Δ (not necessary a disc) we have that

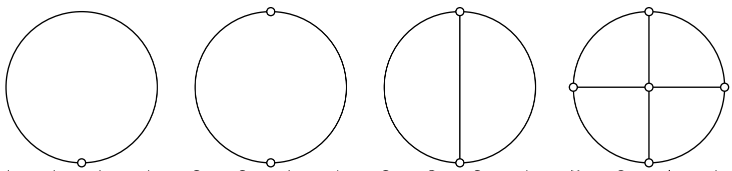
$$\Theta_{\partial\Delta} + \iint\limits_{\Delta} G = 2\cdot\pi\cdot\chi(\Delta), \qquad ❷$$

where $\chi(\Delta)$ is the so called *Euler's characteristic* of Δ. The Euler's characteristic is *topological invariant*, in particular preserved in a continuous deformation.

If a surface Σ (possibly with boundary) can be divided into f discs by drawing e edges connecting v vertexes, then

$$\chi(\Sigma) = v - e + f.$$

For example the disc $\mathbb{D}$ has Euler's characteristic 1; it can be divided



$1-1+1=1$; $2-2+1=1$; $2-3+2=1$; $5-8+4=1$.

into discs many ways, but each time we have $v-e=f=1$. The latter agrees with ❶ and ❷. It is useful to know that $\chi(\mathbb{S}^2)=2$; $\chi(\mathbb{T}^2)=0$ where $\mathbb{T}^2$ denotes torus; $\chi(S_g)=2-2\cdot g$, where S_g is a surface of genus g; that is, sphere with g handles.

(2). Note that if Σ is a plane then a geodesic in Σ are formed by line segments. In this case the statement of theorem follows from Exercise 3.18.

(3). If Σ is the unit sphere then $G\equiv 1$ and therefore formula ❶ can be rewritten as

$$\Theta_{\partial\Delta} + \operatorname{area}\Delta = 2\cdot\pi,$$

which follows from Proposition 12.3.

We will give an informal proof of 12.4 based on the bike wheel interpretation described above. We suppose that it is intuitively clear that moving the axis of the wheel without changing its direction does not change the direction of the wheel's spikes.

More precisely, assume we keep the axis of a non-spinning bike wheel and perform the following two experiments:

(i) We move it around and bring the axis back to the original position. As a result the wheel might rotate by some angle; let us measure this angle.

(ii) We move the direction of the axis the same way as before without moving the center of the wheel. After that we measure the angle of rotation.

Then the resulting angle in these two experiments is the same.

Consider a surface Σ with a Gauss map $\nu\colon \Sigma \to \mathbb{S}^2$. Note that for any point p on Σ, the tangent plane $\mathrm{T}_p\Sigma$ is parallel to the tangent plane $\mathrm{T}_{\nu(p)}\mathbb{S}^2$; so we can identify these tangent spaces. From the experiments above, we get the following:

12.5. Lemma. *Suppose α is a piecewise smooth regular curve in a smooth regular surface Σ which has a Gauss map $\nu\colon \Sigma \to \mathbb{S}^2$. Then the parallel transport along α in Σ coincides with the parallel transport along the curve $\beta = \nu \circ \alpha$ in $\mathbb{S}^2$.*

12.6. Exercise. *Let Σ be a smooth closed surface with positive Gauss curvature. Given a line ℓ denote by ω_ℓ the closed curve formed by points with tangent planes parallel to ℓ.*[1] *Show that parallel transport around ω_ℓ is the identity map.*

Now we are ready to prove the theorem.

Proof of 12.4. Let α be the boundary $\partial\Delta$ parameterized in such a way that Δ lies on the left from it. Assume p is the point where α starts and ends.

Set $\beta = \nu \circ \gamma$ and $q = \nu(p)$, so the spherical curve β starts and ends at q.

By Lemma 12.5 the parallel transport along α in Σ coincides with the parallel transport along the curve β in $\mathbb{S}^2$. By Proposition 11.4, it follows that

$$\Theta_{\alpha,\Sigma} = \Theta_{\beta,\mathbb{S}^2} \pmod{2\cdot\pi}.$$

By Proposition 12.3,

$$\Theta_{\beta,\mathbb{S}^2} + \operatorname{area}\beta = 0 \pmod{2\cdot\pi}.$$

Therefore

$$\Theta_{\alpha,\Sigma} + \operatorname{area}\beta = 0 \pmod{2\cdot\pi}.$$

[1] Equivalently the normal vector at any point of ω_ℓ is perpendicular to ℓ. If the light falls on Σ from one side parallel to ℓ, then ω_ℓ divides the bright and dark sides of Σ.

Recall that the shape operator $s_p\colon T_p\Sigma \to \mathrm{T}_{\nu(p)}\mathbb{S}^2 = T_p\Sigma$ is the Jacobian of the Gauss map $\nu\colon \Sigma \to \mathbb{S}^2$ at the point p. In appropriately chosen coordinates in T_p, the shape operator can be presented by a diagonal matrix $\left(\begin{smallmatrix} k_1 & 0 \\ 0 & k_2 \end{smallmatrix}\right)$, where k_1 and k_2 are the principle curvatures at p. Therefore, the determinant of s_p is the Gauss curvature at p.

If Σ is a closed surface with positive Gauss curvature, then the Gauss map $\nu\colon \Sigma \to \mathbb{S}^2$ is a smooth bijection. Therefore

$$\iint\limits_{\Delta} G = \operatorname{area}[\nu(\Delta)].$$

In the general case we have to count the area $\nu(\Delta)$ taking orientation and multiplicity of the Gauss map into account. In this case

$$\iint\limits_{\Delta} G = \operatorname{area}\beta,$$

where $\operatorname{area}\beta$ is the signed area surrounded by β; it is defined above. Therefore

❸ $$\Theta_{\alpha,\Sigma} + \iint\limits_{\Delta} G = 0 \pmod{2{\cdot}\pi}.$$

If Δ is a disc in the plane then Gauss curvature vanishes and by Exercise 3.18, we have

$$\Theta_{\partial\Delta} + \iint\limits_{\Delta} G = 2{\cdot}\pi.$$

Assunme that Σ_t is a smooth one parameter family of surfaces with a one parameter family of discs $\Delta_t \subset \Sigma_t$ and α_t is the boundary $\partial\Delta_t$ parameterized in such a way that Δ_t lies on the left from it. The value

$$f(t) = \Theta_{\alpha_t} + \iint\limits_{\Delta} G$$

is continuous in t and by ❸ it has to be constant.

If Σ_0 is a plane, then

$$\Theta_{\partial\Delta_0} + \iint\limits_{\Delta_0} G = 2{\cdot}\pi.$$

Intuitively it is clear that any disc can be obtained as a resultof continuous deformation of plane disc. Therefore

$$\Theta_{\partial\Delta_1} + \iint\limits_{\Delta_1} G = 2{\cdot}\pi$$

for arbitrary disc Δ_1; whence ❶ follows. □

12.7. Exercise. *Assume γ is a closed simple curve with constant geodesic curvature 1 in a smooth convex closed surface Σ. Show that*

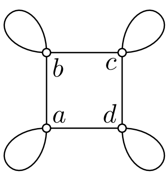
$$\operatorname{length}\gamma \leqslant 2\cdot\pi;$$

that is, the length of γ can not exceed the length of the unit circle in the plane.

12.8. Exercise. *Let γ be a closed simple geodesic on a smooth convex closed surface Σ. Assume $\nu\colon\Sigma\to\mathbb{S}^2$ is a Gauss map. Show that the curve $\alpha=\nu\circ\gamma$ divides the sphere into regions of equal area.*

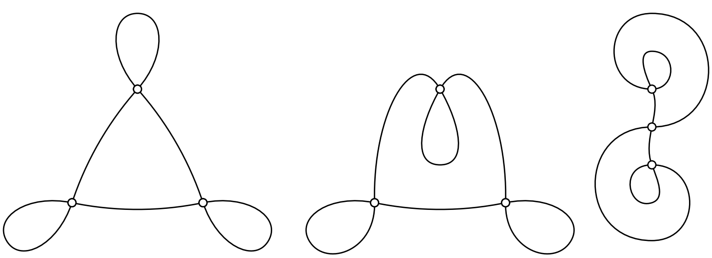
Conclude that

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



12.9. Exercise. *Let Σ be a smooth closed surface with a closed geodesic γ. Assume γ has exactly 4 self-intersection at the points a, b, c and d that appear on γ in the order a,a,b,b,c,c,d,d. Show that Σ can not have positive Gauss curvature.*[2]

12.10. Advanced exercise. *Let Σ be a smooth regular sphere with positive Gauss curvature and $p\in\Sigma$. Suppose γ be a closed geodesic that does not pass thru p. Assume $\Sigma\backslash\{p\}$ parametrized by the plane. Can it happen that in this parametrization, γ looks like one of the curves on the diagram? Say*



as much as possible about possible/impossible diagrams of that type.

[2]Hint: estimate integral of Gauss curvature bounded by a simple geodesic loop.

12.3 The remarkable theorem

Let Σ_1 and Σ_2 be two smooth regular surfaces in the Euclidean space. A map $f\colon \Sigma_1 \to \Sigma_2$ is called length-preserving if for any curve γ_1 in Σ_1 the curve $\gamma_2 = f\circ\gamma_1$ in Σ_2 has the same length. If in addition f is smooth and bijective then it is called *intrinsic isometry.*

A simple example of intrinsic isometry can obtained by warping a plane into a cylinder. The following exercise produce slightly more interesting example.

12.11. Exercise. *Suppose $\gamma(t) = (x(t), y(t))$ is a smooth unit-curve in the plane such that $y(t) = a\cdot\cos t$. Let Σ_γ be the surface of revolution of γ around the x-axis. Show that a small open domain in Σ_γ admits a smooth length-pereserving map to the unit sphere.*

Conclude that any round disc Δ in $\mathbb{S}^2$ of intrinsic radius smaller than $\frac{\pi}{2}$ admits a smooth length preserving deformation; that is there is one parameter family of surfaces with boundary Δ_t, such that $\Delta_0 = \Delta$ and Δ_t is not congruent to Δ_0 for any $t \neq 0$.[3]

12.12. Theorem. *Suppose $f\colon \Sigma_1 \to \Sigma_2$ is an intrinsic isometry between two smooth regular surfaces in the Euclidean space; $p_1 \in \Sigma_1$ and $p_2 = f(p_1) \in \Sigma_1$. Then*

$$G(p_1)_{\Sigma_1} = G(p_2)_{\Sigma_2};$$

that is, the Gauss curvature of Σ_1 at p_1 is the same as the Gauss curvature of Σ_2 at p_2.

This theorem was proved by Carl Friedrich Gauss [26] who called it *Remarkable theorem* (Theorema Egregium). The theorem is indeed remarkable because the Gauss curvature is defined as a product of principle curvatures which might be different at these points; however, according to the theorem, their product can not change.

In fact Gauss curvature of the surface at the given point can be found *intrinsically*, by measuring the lengths of curves in the surface. For example, Gauss curvature $G(p)$ in the following formula for the circumference $c(r)$ of a geodesic circle centered at p in a surface:

$$c(r) = 2\cdot\pi\cdot r - \tfrac{\pi}{3}\cdot G(p)\cdot r^3 + o(r^3).$$

Note that the theorem implies there is no smooth length-preserving map that sends an open region in the unit sphere to the plane.[4] It

[3] In fact any disc in $\mathbb{S}^2$ of intrinsic radius smaller than π admits a smooth length preserving deformation.

[4] There are plenty of non-smooth length-preserving maps from the sphere to the plane; see [27] and the references there in.

follows since the Gauss curvature of the plane is zero and the unit sphere has Gauss curvature 1. In other words, there is no map of a region on Earth without distortion.

Proof. Set $g_1 = G(p_1)_{\Sigma_1}$ and $g_2 = G(p_2)_{\Sigma_2}$; we need to show that

❶ $$g_1 = g_2.$$

Suppose Δ_1 is a small geodesic triangle in Σ_1 that contains p_1. Set $\Delta_2 = f(\Delta_1)$. We may assume that the Gauss curvature is almost constant in Δ_1 and Δ_2; that is, given $\varepsilon > 0$, we can assume that

❷ $$\begin{aligned} |G(x_1)_{\Sigma_1} - g_1| &< \varepsilon, \\ |G(x_2)_{\Sigma_2} - g_2| &< \varepsilon \end{aligned}$$

for any $x_1 \in \Delta_1$ and $x_2 \in \Delta_2$.

Since f is length-preserving the triangles Δ_2 is geodesic and

❸ $$\operatorname{area}\Delta_1 = \operatorname{area}\Delta_2.$$

Moreover, triangles Δ_1 and Δ_2 have the same corresponding angles; denote them by α, β and γ.

By Gauss–Bonnet formula, we get that

❹ $$\iint\limits_{\Delta_1} G_{\Sigma_1} = \alpha + \beta + \gamma - \pi = \iint\limits_{\Delta_2} G_{\Sigma_2}.$$

By ❷,

$$\left| g_1 - \frac{1}{\operatorname{area}\Delta_1} \cdot \iint\limits_{\Delta_1} G_{\Sigma_1} \right| < \varepsilon,$$

$$\left| g_2 - \frac{1}{\operatorname{area}\Delta_2} \cdot \iint\limits_{\Delta_2} G_{\Sigma_2} \right| < \varepsilon.$$

By ❸ and ❹,

$$\frac{1}{\operatorname{area}\Delta_1} \cdot \iint\limits_{\Delta_1} G_{\Sigma_1} = \frac{1}{\operatorname{area}\Delta_2} \cdot \iint\limits_{\Delta_2} G_{\Sigma_2},$$

therefore

$$|g_1 - g_2| < 2\cdot\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, ❶ follows. □

12.4 Simple geodesic

The following theorem provides an interesting application of Gauss–Bonnet formula; it is proved by Stephan Cohn-Vossen [Satz 9 in 28].

12.13. Theroem. *Any open smooth regular surface with positive Gauss curvature has a simple two-sided infinite geodesic.*

12.14. Lemma. *Suppose Σ is an open surface in with positive Gauss curvature in the Euclidean space. Then there is a convex function f defined on a convex open region of (x,y)-plane such that Σ can be presented as a graph $z = f(x,y)$ in some (x,y,z)-coordinate system of the Euclidean space.*

Moreover

❶ $$\iint\limits_{\Sigma} G \leqslant 2{\cdot}\pi.$$

Proof. The surface Σ is a boundary of an unbounded closed convex set K.

Fix $p \in \Sigma$ and consider a sequence of points x_n such that $|x_n - p| \to \to \infty$ as $n \to \infty$. Set $u_n = \frac{x_n - p}{|x_n - p|}$; the unit vector in the direction from p to 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. The half-line from q in the direction of u lies in K. Indeed any point on the half-line is a limit of points on the line segments $[q, x_n]$; since K is closed, all of these poins lie in K.

Let us choose the z-axis in the direction of u. Note that line segments can not lie in Σ, otherwise its Gauss curvature would vanish. It follows that any vertical line can intersect Σ at most at one point. That is, Σ is a graph of a function $z = f(x,y)$. Since K is convex, the function f is convex and it is defined in a region Ω which is convex. The domain Ω is the projection of Σ to the (x,y)-plane. This projection is injective and by the inverse function theorem, it maps open sets in Σ to open sets in the plane; hence Ω is open.

It follows that the outer normal vectors to Σ at any point, points to the south hemisphere $\mathbb{S}^2_- = \left\{\,(x,y,z) \in \mathbb{S}^2 \;:\; z < 0\,\right\}$. Therefore the area of the spherical image of Σ is at most $\operatorname{area} \mathbb{S}^2_- = 2{\cdot}\pi$. The area of this image is the integral of the Gauss curvature along Σ. That is,

$$\begin{aligned}\iint\limits_{\Sigma} G &= \operatorname{area}[\nu(\Sigma)] \leqslant \\ &\leqslant \operatorname{area}\mathbb{S}^2_- = \\ &= 2\cdot\pi,\end{aligned}$$

where $\nu(p)$ denotes the outer unit normal vector at p. Hence ❶ follows. □

Proof of 12.13. Let Σ be an open surface in with positive Gauss curvature and γ a two-sided infinite geodesic in Σ. The following is the key statement in the proof.

12.15. Claim. *The geodesic γ contains at most one simple loop.*

Assume γ has a simple loop ℓ. By Lemma 12.14, Σ is parameterized by a open convex region Ω in the plane; therefore ℓ bounds a disc in Σ; denote it by Δ. If φ is the angle at the base of the loop, then by Gauss–Bonnet,

$$\iint\limits_{\Delta} G = \pi + \varphi.$$

By Lemma 12.14, $\varphi < \pi$; that is γ has no concave simple loops

Assume γ has two simple loops, say ℓ_1 and ℓ_2 that bound discs Δ_1 and Δ_2. Then the disks Δ_1 and Δ_2 have to overlap, otherwise the curvature of Σ would exceed $2\cdot\pi$.

We may assume that $\Delta_1 \not\subset \Delta_2$; the loop ℓ_2 appears after ℓ_1 on γ and there are no other simple loops between them. In this case, after going around ℓ_1 and before closing ℓ_2, the curve γ must enter Δ_1 creating a concave loop. The latter contradicts the above observation.

If a geodesic γ has a self-intersection, then it contains a simple loop. From above, there is only one such loop; it cuts a disk from Σ and goes around it either clockwise or counterclockwise. This way we divide all the self-intersecting geodesics into two sets which we will call *clockwise* and *counterclockwise*.

Note that the geodesic $t \mapsto \gamma(t)$ is clockwise if and only if the same geodesic traveled backwards $t \mapsto \gamma(-t)$ is counterclockwise. By shooting unit-speed geodesics in all directions at a given point $p = \gamma(0)$, we get a one parameter family of geodesics γ_s for $s \in [0,\pi]$ connecting the geodesic $t \mapsto \gamma(t)$ with the $t \mapsto \gamma(-t)$; that is, $\gamma_0(t) = \gamma(t)$ and $\gamma_\pi(t) = \gamma(-t)$. It follows that there are geodesics which aren't clockwise nor counterclockwise. Those geodesics have no self-intersections. □

12.16. Exercise. *Suppose that $f\colon \mathbb{R}^2 \to \mathbb{R}$ is a $\sqrt{3}$-Lipshitz smooth convex function. Show that any geodesic in the surface defined by the graph $z = f(x, y)$ has no self-intersections.*

Chapter 13

Local comparison

13.1 First variation formula

13.1. Proposition. *Assume $(s,t)\mapsto w(s,t)$ be a local parametrization of an oriented smooth regular surface Σ such that $\tfrac{\partial}{\partial s}w\perp\tfrac{\partial}{\partial t}w$, $|\tfrac{\partial}{\partial s}w|=1$ and the vector $\tfrac{\partial}{\partial s}w$ points to the right from $\tfrac{\partial}{\partial t}w$ at any parameter value (s,t).*

Fix a closed real interval $[a,b]$ and consider a one parameter family of curves $\sigma_s\colon[a,b]\to\Sigma$ defined as the coordinate lines $\sigma_s(t)=w(s,t)$. Set $\ell(s)=\operatorname{length}\sigma_s$. Then

$$\ell'(s)=\Theta_{\sigma_s}$$

for any s.

The proof is done by direct calculations.

Proof. Since $\tfrac{\partial}{\partial s}w\perp\tfrac{\partial}{\partial t}w$, we have that

$$\langle\tfrac{\partial}{\partial s}w,\tfrac{\partial}{\partial t}w\rangle=0$$

and therefore

$$\langle\tfrac{\partial^2}{\partial s\partial t}w,\tfrac{\partial}{\partial t}w\rangle+\langle\tfrac{\partial}{\partial s}w,\tfrac{\partial^2}{\partial t^2}w\rangle=\tfrac{\partial}{\partial t}\langle\tfrac{\partial}{\partial s}w,\tfrac{\partial}{\partial t}w\rangle=0.$$

Note that $|\gamma_s'(t)| = |\frac{\partial}{\partial t}w(s,t)|$ and therefore

$$\begin{aligned}\tfrac{\partial}{\partial s}|\gamma_s'(t)| &= \tfrac{\partial}{\partial s}\sqrt{\langle \tfrac{\partial}{\partial t}w(s,t), \tfrac{\partial}{\partial t}w(s,t)\rangle} = \\ &= \frac{\langle \tfrac{\partial^2}{\partial s\partial t}w(s,t), \tfrac{\partial}{\partial t}w(s,t)\rangle}{\sqrt{\langle \tfrac{\partial}{\partial t}w(s,t), \tfrac{\partial}{\partial t}w(s,t)\rangle}} = \\ &= -\frac{\langle \tfrac{\partial}{\partial s}w, \tfrac{\partial^2}{\partial t^2}w\rangle}{|\gamma_s'(t)|} = \\ &= -\frac{\langle \tfrac{\partial}{\partial s}w, \gamma_s''(t)\rangle}{|\gamma_s'(t)|}.\end{aligned}$$

The values $\ell(s)$ do not change if we reparametrize γ_s, so we can assume that for a fixed value s the curve σ_s is unit-speed. Since $|\frac{\partial}{\partial s}w| = 1$ and $\frac{\partial}{\partial s}w$ points to the right from $\frac{\partial}{\partial t}w = \gamma_s'(t)$, the last expression equals to $k_g(s,t)$, where $k_g(s,t)$ denotes the geodesic curvature of σ_s at t. Therefore, for this particular s we have

$$\begin{aligned}\ell'(s) &= \int_a^b \tfrac{\partial}{\partial s}|\gamma_s'(t)|\cdot dt = \\ &= \int_a^b k_g(s,t)\cdot dt = \\ &= \Theta_{\sigma_s}.\end{aligned}$$

Since the left hand side and the right hands side of this formula do not depend on the parametrization of σ_s, this formula holds for all s.[1] □

The parametrization of a surface satisfying the conditions in the proposition are called *semigeodesic coordinates.* The following exercise explains the reason for this name.

13.2. Exercise. *Assume $(s,t) \mapsto w(s,t)$ be a local parametrization of an oriented smooth regular surface Σ as in the proposition above. Show that for any fixed t the curve $\gamma_t(s) = w(s,t)$ is a geodesic.*[2]

[1] One may avid passing the a unit speed parametrization by using the following formula for geodesic curvature which holds for any regular parametrization:

$$k_g(t,s) = \langle \nu(\sigma_s(t)), [\sigma_s'(t), \sigma_s''(t)]\rangle / |\sigma_s'(t)|^3;$$

it will save your thinking by making calculations longer.

[2] Hint: note that in order to show that $\gamma_t''(s) \perp \mathrm{T}_{\gamma_t(s)}$, it is sufficient to show that $\langle \frac{\partial^2}{\partial s^2}w, \frac{\partial}{\partial t}w\rangle = 0$.

13.2 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 $q = \exp_p v$. The map $\exp_p \colon \mathrm{T}_p \to \Sigma$ is defined in a neighborhood of zero. We assume that it is intuitively obvious that the map $\exp_p$ is smooth; formally it follows since the solution of the initial value problem for the equation $\gamma_v''(t) \perp \mathrm{T}_{\gamma_v(t)}$ which describes the geodesic γ_v smoothly depend on the initial data v. Note that the Jacobian of $\exp_p$ at zero is the identity matrix. Therefore frm the inverse function theorem we get the following statement:

13.3. Proposition. *Let Σ be smooth regular 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.

Note that if there are two minimizing geodesics between two points x and y in a surface, then there are two distinct vectors $v, v' \in \mathrm{T}_x$ such that $y = \exp_x v = \exp_x v'$. Therefore by the above proposition we get the following:

13.4. Corollary. *Let Σ be a smooth regular surface. Then for any point $p \in \Sigma$ there is $\varepsilon > 0$ such that any two points x and y in the ε-neightbohood of p in Σ can be connected by a unique minimizing geodesic $[xy]_\Sigma$.*

13.3 Polar coordinates

Proposition 13.3 implies existence of polar coordinates in a neighborhood of any point in p in Σ. That is, any point x in Σ sufficiently close to p can be uniquely described by the distance $|x - p|_\Sigma$ and the direction from p to x.

Assume (θ, r) are the described polar coordinates at p. Namely, assume $\tilde{w}(\theta, r)$ denotes the tangent vector at p with polar coordinates (θ, r) and $w(\theta, r) = \exp_p[\tilde{w}(\theta, r)]$. By the definition of exponential map, for a fixed θ, the curve $\gamma_\theta(t) = w(\theta, t)$ is a unit-speed geodesic that starts at p; in particular $|\frac{\partial}{\partial r} w| = |\gamma_\theta'(r)| = 1$ and $\gamma_\theta''(r) \perp \mathrm{T}_{\gamma_\theta(r)}$.

The curve $\sigma_r(t) = w(t,r)$ is a parametrization of the circle of radius r and center at p in Σ; that is, if $q = \sigma_r(t)$, then $|q-p|_\Sigma = r$. If the latter is not the case, then a minimizing geodesic $[pq]_\Sigma$ would be shorter than r and therefore q would not be described uniquely in the polar coordinates.

Note that $\frac{\partial}{\partial r}w \perp \frac{\partial}{\partial \theta}w$ if $r > 0$; otherwise for small $\varepsilon > 0$ the intrinsic distance from p to $w(\theta \pm \varepsilon, r)$ would be shorter than r, which contradicts the previous statement.

13.5. Proposition. *Let $w(\theta,r)$ and $\tilde w(\theta,r)$ be the polar coordinates of a surface Σ at p and its tangent plane T_p at zero, so $w(\theta,r) = \exp_p[\tilde w(\theta,r)]$. Given a real interval $[a,b]$ consider the one parameter families of circular arcs $\sigma_r\colon [a,b] \to \Sigma$ and $\tilde\sigma_r\colon [a,b] \to \mathrm{T}_p$ $\sigma_r(t) = w(t,r)$ and $\tilde\sigma_r(t) = \tilde w(t,r)$. Set $\ell(r) = \operatorname{length}\sigma_r$ and $\tilde\ell(r) = \operatorname{length}\tilde\sigma_r$.*[3]

(i) If the Gauss curvature of Σ is nonnegative, then

$$\ell(r) \leqslant \tilde\ell(r)$$

for all small $r > 0$.

(ii) If the Gauss curvature of Σ is nonpositive, then

$$\ell(r) \geqslant \tilde\ell(r)$$

for all small $r > 0$.

Taking a limit as $b \to a$, we obtain the following corollary.

13.6. Corollary. *Let $w(\theta,r)$ and $\tilde w(\theta,r)$ be the polar coordinates of a surface Σ at p and its tangent plane T_p at zero, so $w(\theta,r) = \exp_p[\tilde w(\theta,r)]$.*

(i) If the Gauss curvature of Σ is nonnegative, then

$$|\tfrac{\partial}{\partial\theta}w| \leqslant |\tfrac{\partial}{\partial\theta}\tilde w|$$

for all small $r > 0$.

(ii) If the Gauss curvature of Σ is nonpositive, then

$$|\tfrac{\partial}{\partial\theta}w| \geqslant |\tfrac{\partial}{\partial\theta}\tilde w|$$

for all small $r > 0$.

Proof. From the above discussion, the polar coordinates $w(\theta,r)$ are semigeodesic; that is $w(\theta,r)$ satisfies the conditions in the first variation formula (13.1). In particular if $\ell(r) = \operatorname{length}\sigma_r$, then

$$\ell'(r) = \Theta_{\sigma_r}$$

[3] Note that angular measure of $\tilde\sigma_r$ is $b-a$; therefore $\tilde\ell(r) = r\cdot(b-a)$.

for any $r > 0$.

By Gauss–Bonnet formula, the last identity can be rewritten as

❶ $$\ell'(r) = 2{\cdot}(b - a) - \iint\limits_{\Delta_r} G,$$

where Δ_r is the sector in Σ in the polar coordinates at p

$$\{\, w(t,s) \;:\; a \leqslant t \leqslant b,\; 0 \leqslant s \leqslant r \,\};$$

which is bounded by two geodesics from p with angle $b - a$ and a circular arc that meets these geodesics at right angle.

Since the plane has vanishing Gauss curvature, we have

❷ $$\tilde\ell'(r) = 2{\cdot}(b - a),$$

which agrees with the formula for the length of the arc $\tilde\ell(r) = 2{\cdot}\pi{\cdot}r$.

If the Gauss curvature of Σ is nonnegative, the equations ❶ and ❷ imply that

$$\ell'(r) \leqslant \tilde\ell'(r)$$

for any small r.

If the Gauss curvature of Σ is nonnegative, the same equations imply that

$$\ell'(r) \geqslant \tilde\ell'(r)$$

for any small r.

Since $\ell(0) = \tilde\ell(0)$, integrating the inequalities proves both statements. □

The following exercise provides a stronger statement. It almost follow from the proof above, but one has to make an extra observation.

13.7. Exercise. *Assume Σ is a smooth regular surface and $p \in \Sigma$, denote by $\ell(r)$ the circumference of the circle with the center at p and radius r in Σ and let $\tilde\ell(r) = 2{\cdot}\pi{\cdot}r$ the circumference of the plane circle of radius r.*

(i) Show that if Gauss curvature of Σ is nonnegative, then the function $r \mapsto \ell(r)$ is concave for small $r > 0$. Conclude that the function $r \mapsto \frac{\ell(r)}{\tilde\ell(r)}$ is nonincreasing for small $r > 0$.

(ii) Show that if Gauss curvature of Σ is nonpositive, then the function $r \mapsto \ell(r)$ is convex for small $r > 0$. Conclude that the function $r \mapsto \frac{\ell(r)}{\tilde\ell(r)}$ is nondecresing for small $r > 0$.

13.4 Local comparison

The following proposition is a special case of the so a comparison theorem, proved by Harry Rauch [29].

13.8. Theorem. *Let Σ be a smooth regular surface and $p \in \Sigma$. Assume $\tilde\gamma\colon[a,b]$ is a curve the tangent plane $\mathrm{T}_p\Sigma$ that runs in a sufficiently small neighborhood of the origin; consider the curve*

$$\gamma = \exp_p \circ\gamma$$

in Σ.

(i) If Gauss curvature of Σ is nonnegative, then

$$\operatorname{length}\gamma \leqslant \operatorname{length}\tilde\gamma$$

(ii) If Gauss curvature of Σ is nonpositive, then

$$\operatorname{length}\gamma \geqslant \operatorname{length}\tilde\gamma.$$

The proof is a direct application of Corollary 13.6.

Proof. Let us denote $\tilde w(\theta,r)$ and $w(\theta,r)$ the polar coordinates of T_p and Σ at p Recall that

$$\begin{aligned}
&\tfrac{\partial\tilde w}{\partial\theta}\perp\tfrac{\partial\tilde w}{\partial r}; && |\tfrac{\partial\tilde w}{\partial r}|=1;\\
&\tfrac{\partial w}{\partial\theta}\perp\tfrac{\partial w}{\partial r}; && |\tfrac{\partial w}{\partial r}|=1;
\end{aligned}$$

By Corollary 13.6, we also have

$$|\tfrac{\partial\tilde w}{\partial\theta}|\geqslant|\tfrac{\partial\tilde w}{\partial\theta}|; \qquad |\tfrac{\partial\tilde w}{\partial\theta}|\leqslant|\tfrac{\partial\tilde w}{\partial\theta}|;$$

if Gauss curvature is nonnegative or nonpositive correspondingly.

It is sufficient to show that

❶ $$|\gamma'(t)|\leqslant|\tilde\gamma'(t)| \quad \text{or, correspondingly} \quad |\gamma'(t)|\geqslant|\tilde\gamma'(t)|$$

for any t.

Note that both curves $\gamma(t)$ and $\tilde\gamma(t)$ described the same way in the polar coordinates; denote these coordinates by $(\theta(t),r(t))$. Then

$$\begin{aligned}
\gamma'(t)|^2&=|\tfrac{\partial w}{\partial\theta}\cdot\theta'(t)+\tfrac{\partial w}{\partial r}\cdot r'(t)|^2=\\
&=|\tfrac{\partial w}{\partial\theta}|^2\cdot|\theta'(t)|^2+|r'(t)|^2
\end{aligned}$$

The same way

$$\tilde\gamma'(t)|^2=|\tfrac{\partial\tilde w}{\partial\theta}|^2\cdot|\theta'(t)|^2+|r'(t)|^2;$$

hence ❶ follows. □

AFTER THIS LINE READ AT YOUR OWN RISK!!!

Chapter 14

Global comparison

14.1 Formulation

A minimizing geodesic between points x and y in a surface Σ will be denoted as $[xy]$ or $[xy]_\Sigma$; the latter notation is used if we need to emphasise that the geodesic lies in Σ. If we write $[xy]$, then we assume that a minimizing geodesic exists and we made a choice of one of them.

In general minimizing geodesic might be not unique for example any meridian in the sphere is a minimizing geodesic between its poles. If Σ is complete, then a minimizing geodesic always exists.

A *geodesic triangle* in a surface Σ is a triple of points $x, y, z \in \Sigma$ with choice of minimizing geodesics $[xy]$, $[yz]$ and $[zx]$. The points x, y, z are called *vertexes* of the geodesic triangle, the minimizing geodesics $[xy]$, $[yz]$ and $[zx]$ are called its sides; the triangle itself is denoted by $[xyz]$.

The length of one (and therefore any) minimizing geodesic $[xy]_\Sigma$ will be denoted by $|x-y|_\Sigma$; it is called *intrinsic distance* from x to y in Σ. If defined, then $|x-y|_\Sigma$ is the exact lower bound on the lengths of curves from x to y in Σ.

A triangle $[\tilde{x}\tilde{y}\tilde{z}]$ in the plane $\mathbb{R}^2$ is called *model triangle* of the triangle $[xyz]$, briefly $[\tilde{x}\tilde{y}\tilde{z}] = \tilde{\triangle}xyz$, if its corresponding sides are equal; that is,

$$|\tilde{x}-\tilde{y}|_{\mathbb{R}^2} = |x-y|_\Sigma, \quad |\tilde{y}-\tilde{z}|_{\mathbb{R}^2} = |y-z|_\Sigma, \quad |\tilde{z}-\tilde{x}|_{\mathbb{R}^2} = |z-x|_\Sigma.$$

A pair of minimizing geodesics $[xy]$ and $[xz]$ starting from one point x is called *hinge* and denoted as $[x\,{}^y_z]$. The angle between these geodesics at x is denoted by $\measuredangle[x\,{}^y_z]$. The corresponding angle $\measuredangle[\tilde{x}\,{}^{\tilde{y}}_{\tilde{z}}]$ in the model triangle $[\tilde{x}\tilde{y}\tilde{z}] = \tilde{\triangle}xyz$ is denoted by $\tilde{\measuredangle}(x\,{}^y_z)$.

A surface Σ is called *simply connected* if any closed simple curve in Σ bounds a disc. Equivalently any closed curve in Σ can be continuously deformed into a trivial curve (trivial means that it stays at one point). A plane or sphere are examples of simply connected surfaces, while torus or cylinder are not simply connected.

14.1. Comparison theorem. *Let Σ be a complete smooth regular surface with a geodesic triangle $[xyz]$.*

(i) If Σ has nonnegative Gauss curvature, then

$$\measuredangle[x\,{}^y_z] \geqslant \tilde\measuredangle(x\,{}^y_z).$$

(ii) If Σ is simply connected and has nonpositive Gauss curvature, then

$$\measuredangle[x\,{}^y_z] \leqslant \tilde\measuredangle(x\,{}^y_z).$$

Let us make few remarks on the formulation.

The angle $\measuredangle[x\,{}^y_z]$ is a number in the interval $[0,\pi]$. If the triangle $[xyz]$ bounds a disc Δ and θ is the external angle at x which used in Gauss–Bonnet formula, then $\measuredangle[x\,{}^y_z] = |\pi - \theta|$. The corresponding internal angle might be $\measuredangle[x\,{}^y_z]$ or $2{\cdot}\pi - \measuredangle[x\,{}^y_z]$ depending on which side lies the disc Δ.

- ◇ Since the angles of any plane triangle sum up to π, the part (i) of the theorem implies that angles of any triangle in a surface with nonnegative Gauss curvature have sum at least π.
- ◇ The triangle may not bound a disc[1], but if it does, then by Gauss–Bonnet formula the sum of its *internal* angles is at least π.

These two statements are closely related, but they are not the same. Note that if α is the angle in the comparison theorem, then the internal angle might be α or $2{\cdot}\pi-\alpha$; while Gauss–Bonnet formula gives a lower bound on the sum of internal angles it does not forbid that each of these angles is close to $2{\cdot}\pi$. However the latter is impossible by the comparison theorem.

First note that without condition that Σ is simply connected, the statement *(ii)* does not hold. For example the equator $z=0$ of the hyperboloid (which is not simply connected)

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

forms a triangle with all angles π, which contradict the comparison in *(ii)*.

14.2. Exercise. *Let Σ be a complete smooth regular simply connected surface with nonpositive Gauss curvature. Show that any two points in Σ are connected by a unique geodesic.*

[1]For example equator on the cyclinder is formed by a geodesic triangle that does not bound a disc.

14.2 Names and history

Part *(i)* of this theorem is called *Toponogov comparison theorem*; it is was proved by Paolo Pizzetti [30] and latter independently by Alexandr Alexandrov [31]; generalizations were obtained by Victor Toponogov [32], Mikhael Gromov, Yuri Burago and Grigory Perelman [33].

Part *(ii)* is called *Cartan–Hadamard theorem*; it was proved by Hans von Mangoldt [34] and generalized by Elie Cartan [35], Jacques Hadamard [20], Herbert Busemann [36], Willi Rinow in [37], Mikhael Gromov [38, p. 119], Stephanie Alexander and Richard Bishop in [39].

14.3 Local part

First we prove the following local version of comparison theorem and then use it to prove the global version.

14.3. Theorem. *The comparison theorem (14.1) holds in a small neighborhood of any point.*

That is, if Σ be a complete smooth regular surface, then any point $p \in \Sigma$ admits a neighborhood $U \ni p$ such that

(i) If Σ has nonnegative Gauss curvature, then for any geodesic triangle $[xyz]$ in U we have

$$\measuredangle[x\,{}^y_z] \geqslant \tilde{\measuredangle}(x\,{}^y_z).$$

(ii) If Σ has nonpositive Gauss curvature, then for any geodesic triangle $[xyz]$ in U we have

$$\measuredangle[x\,{}^y_z] \leqslant \tilde{\measuredangle}(x\,{}^y_z).$$

Note that we can assume that U is simply connected therefore this condition is not necessary to include in part *(ii)*.

Proof. Assume $y = \exp_x v$ and $z = \exp_x w$ for two small vectors $v, w \in \mathrm{T}_x$. Note that

$$\begin{aligned}
\measuredangle[x\,{}^v_w]_{\mathrm{T}_x} &= \measuredangle[x\,{}^y_z]_\Sigma,\\
|x - v|_{\mathrm{T}_x} &= |x - y|_\Sigma,\\
|x - w|_{\mathrm{T}_x} &= |x - z|_\Sigma.
\end{aligned}$$

If the Gauss curvature is nonnegative, consider the line segment $\tilde\gamma$ joining v to w in the tangent plane T_x and set $\gamma = \exp_x \circ\tilde\gamma$. By Rauch comparison theorem (13.8), we have

$$\operatorname{length}\gamma \leqslant \operatorname{length}\tilde\gamma.$$

Since $|v-w|_{\mathrm{T}_x} = \operatorname{length}\tilde\gamma$ and $|y-z|_\Sigma \leqslant \operatorname{length}\gamma$, we get

$$|v-w|_{\mathrm{T}_x} \geqslant |y-z|_\Sigma.$$

Therefore

$$\tilde\measuredangle(x\,{}^z_y) \geqslant \measuredangle[x\,{}^z_y].$$

If the Gauss curvature is nonpositive, consider a minimizing geodesic γ joining y to z in Σ and let $\tilde\gamma$ be the corresponding curve joining v to w in T_x; that is, $\gamma = \exp_x\circ\tilde\gamma$. By Rauch comparison theorem (13.8), we have

$$\operatorname{length}\gamma \geqslant \operatorname{length}\tilde\gamma.$$

Since $|v-w|_{\mathrm{T}_x} \leqslant \operatorname{length}\tilde\gamma$ and $|y-z|_\Sigma = \operatorname{length}\gamma$, we get

$$|v-w|_{\mathrm{T}_x} \geqslant |y-z|_\Sigma.$$

Therefore

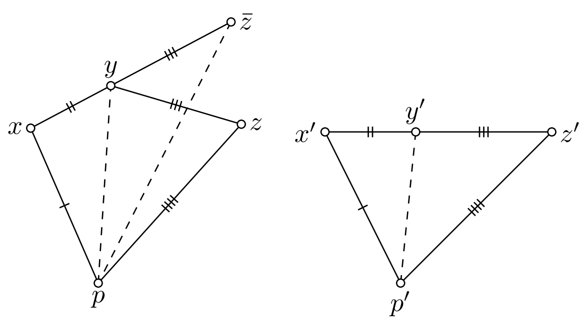
$$\tilde\measuredangle(x\,{}^z_y) \geqslant \measuredangle[x\,{}^z_y].$$ □

14.4 Alexandrov's lemma

It this section we prove the following lemma in the plane geometry.

14.4. Lemma. *Assume $[pxyz]$ and $[p'x'y'z']$ be two quadraliterals in the plane with equal corresponding sides. Assume that the sides $[x'y']$ and $[y'z']$ extend each other; that is, y' lies on the line segment $[x'z']$. Then the following expressions have the same signs:*

(i) $|p-y| - |p'-y'|$;
(ii) $\measuredangle[x\,{}^p_y] - \measuredangle[x'\,{}^{p'}_{y'}]$;
(iii) $\pi - \measuredangle[y\,{}^p_x] - \measuredangle[y\,{}^p_z]$;



Proof. In the proof we use the following *monotonicity property*: if two sides adjacent to an angle in a plane triangle are fixed, then the angle is increases if the opposite side increase.

Take a point $\bar z$ on the extension of $[xy]$ beyond y so that $|y-\bar z|=|y-z|$ (and therefore $|x-\bar z|=|x'-z'|$).

From monotonicity, the following expressions have the same sign:

(i) $|p-y|-|p'-y'|$;

(ii) $\measuredangle[x\,{}^{y}_{p}]-\measuredangle[x'\,{}^{y'}_{p'}]=\measuredangle[x\,{}^{\bar z}_{p}]-\measuredangle[x'\,{}^{z'}_{p'}]$;

(iii) $|p-\bar z|-|p'-z'|$;

(iv) $\measuredangle[y\,{}^{\bar z}_{p}]-\measuredangle[y'\,{}^{z'}_{p'}]$;

The statement follows since

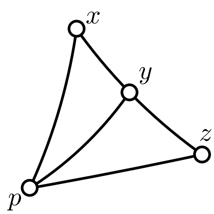
$$\measuredangle[y'\,{}^{z'}_{p'}]+\measuredangle[y'\,{}^{x'}_{p'}]=\pi$$

and

$$\measuredangle[y\,{}^{\bar z}_{p}]+\measuredangle[y\,{}^{x}_{p}]=\pi. \qquad \square$$

Further we will use the following reformulation of this lemma that is using language of comparison triangles and angles.

14.5. Reformulation. *Assume $[pxz]$ be a triangle in a surface Σ and the point y lies on the side $[xz]$. Consider its model triangle $[\tilde p\tilde x\tilde z]=\tilde\triangle pxz$ and let $\tilde y$ be the corresponding point on the side $[\tilde x\tilde z]$. Then the following expressions have the same signs:*



(i) $|p-y|_\Sigma-|\tilde p-\tilde y|_{\mathbb{R}^2}$;

(ii) $\tilde\measuredangle(x\,{}^{p}_{y})-\tilde\measuredangle(x\,{}^{p}_{z})$;

(iii) $\pi-\tilde\measuredangle(y\,{}^{p}_{x})-\tilde\measuredangle(y\,{}^{p}_{z})$;

14.5 Reformulations of comparison

In this section we formulate conditions equivalent to the conclusion of the comparison theorem (14.1).

A triangle $[xyz]$ in a surface is called *fat* (or correspondingly *thin*) if for any two points p and q on the sides of the triangle and the corresponding points $\tilde p$ and $\tilde q$ on the sides of its model triangle $[\tilde x\tilde y\tilde z]=\tilde\triangle xyz$ we have $|p-q|\geqslant|\tilde p-\tilde q|$ (or correspondingly $|p-q|\leqslant|\tilde p-\tilde q|$).

14.6. Proposition. *Let Σ be a complete smooth regular surface. Then the following three conditions are equivalent:*

(i^+) For any geodesic triangle $[xyz]$ in Σ we have

$$\measuredangle[x\,{}^{y}_{z}]\geqslant\tilde\measuredangle(x\,{}^{y}_{z}).$$

(ii^+) For any geodesic triangle $[pxz]$ in Σ and y on the side $[xz]$ we have

$$\tilde\measuredangle(x\,{}^{p}_{y})\geqslant\tilde\measuredangle(x\,{}^{p}_{z}).$$

(iii$^+$) Any geodesic triangle in Σ is fat.

Similarly, following three conditions are equivalent:

(i$^-$) For any geodesic triangle $[xyz]$ in Σ we have

$$\measuredangle[x\,{}^y_z] \leqslant \tilde\measuredangle(x\,{}^y_z).$$

(ii$^-$) For any geodesic triangle $[pxz]$ in Σ and y on the side $[xz]$ we have

$$\tilde\measuredangle(x\,{}^p_y) \leqslant \tilde\measuredangle(x\,{}^p_z).$$

(iii$^-$) Any geodesic triangle in Σ is thin.

Proof. We will prove the implications *(i$^+$)⇒(ii$^+$)⇒(iii$^+$)⇒(i$^+$).* The implications *(i$^-$)⇒(ii$^-$)⇒(iii$^-$)⇒(i$^-$)* can be done the same way.

(i$^+$)⇒(ii$^+$). Note that $\measuredangle[y\,{}^p_x] + \measuredangle[y\,{}^p_z] = \pi$. By *(i$^+$)*,

$$\tilde\measuredangle(y\,{}^p_x) + \tilde\measuredangle(y\,{}^p_z) \leqslant \pi.$$

It reamains to apply Alexandrov's lemma (14.5).

(ii$^+$)⇒(iii$^+$). Applying *(i$^+$)* twice, first for $y \in [xz]$ and then for $w \in [px]$, we get that

$$\tilde\measuredangle(x\,{}^w_y) \geqslant \tilde\measuredangle(x\,{}^p_y) \geqslant \tilde\measuredangle(x\,{}^p_z)$$

and therefore

$$|w - y|_\Sigma \geqslant |\tilde w - \tilde y|_{\mathbb{R}^2},$$

where $\tilde w$ and $\tilde y$ are the points corresponding to w and y points on the sides of the model triangle. Hence the implication follows.

(iii$^+$)⇒(i$^+$). Since the triangle is fat, we have

$$\tilde\measuredangle(x\,{}^w_y) \geqslant \tilde\measuredangle(x\,{}^p_z)$$

for any $w \in\,]xp]$ and $y \in\,]xz]$. Note that $\tilde\measuredangle(x\,{}^w_y) \to \measuredangle[x\,{}^p_z]$ as $w, y \to x$, whence the implication follows. □

In the following exercises you can apply the globalization theorem.

14.7. Exercise. *Let Σ be a complete smooth regular surface with nonnegative Gauss curvature. Show that for any four distinct points the following inequality holds:*

$$\tilde\measuredangle(p\,{}^x_y) + \tilde\measuredangle(p\,{}^y_z) + \tilde\measuredangle(p\,{}^z_x) \leqslant 2{\cdot}\pi.$$

14.8. Exercise. *Let Σ be a complete smooth regular surface and γ be a unit-speed geodesic in Σ and $p \in \Sigma$.*

Consider the function

$$h(t) = |p - \gamma(t)|^2_\Sigma - t^2.$$

(a) *Show that if the Gauss curvature of* Σ *is nonnegative then* h *is a concave function.*
(b) *Show that if* Σ *is simply connected and the Gauss curvature of* Σ *is nonpositive then* h *is a convex function.*

14.9. Exercise. *Let* $\tilde{x}_1 \dots \tilde{x}_n$ *be a convex plane polygon and* $x_1 \dots x_n$ *be a broken geodesic in a complete simply connected surface* Σ *with nonpositive curvature. Assume that* $|x_i - x_{i-1}|_\Sigma = |\tilde{x}_i - \tilde{x}_{i-1}|_{\mathbb{R}^2}$ *and* $\measuredangle[x_{i-1}\,{}^{x_i}_{x_{i+1}}] \geqslant \measuredangle[x_{i-1}\,{}^{x_i}_{x_{i+1}}]$ *for each* i*. Show that*

$$|x_1 - x_n|_\Sigma \geqslant |\tilde{x}_1 - \tilde{x}_n|_{\mathbb{R}^2}.$$

For $\Sigma = \mathbb{R}^2$, the exercise above is the so called *arm lemma*; you can use it without proof.

14.10. Exercise. *Let* x' *and* y' *be the midpoints of minimizing geodesics* $[px]$ *and* $[py]$ *in a complete smooth regular surface* Σ*.*
(a) *Show that if the Gauss curvature of* Σ *is nonnegative, then*

$$2{\cdot}|x' - y'|_\Sigma \geqslant |x - y|_\Sigma.$$

(b) *Show that if* Σ *is simply connected and has nonpositive Gauss curvature, then*

$$2{\cdot}|x' - y'|_\Sigma \leqslant |x - y|_\Sigma.$$

14.6 Nonnegative curvature

In this section we will prove part *(i)* of the comparison theorem (14.1) assuming that Σ is compact; the general case requre only minor modifications.

Since Σ is compact, from the local theorem (14.3), we get that there is $\varepsilon > 0$ such that the inequality

$$\measuredangle[x\,{}^y_z] \geqslant \tilde{\measuredangle}(x\,{}^y_z).$$

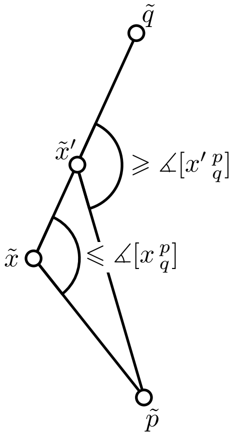
holds for any hinge $[x\,{}^y_z]$ such that $|x - y| + |x - z| < \varepsilon$. The following lemma states that in this case the same holds for any hinge $[x\,{}^y_z]$ such that $|x - y| + |x - z| < \frac{3}{2}{\cdot}\varepsilon$. Applying the lemma few times we will get that the comparison holds for arbitrary hinge, which will prove part *(i)*.

14.11. Key lemma. *Let* Σ *be a complete smooth regular surface. Assume that the comparison*

$$\measuredangle[x\,{}^y_z] \geqslant \tilde{\measuredangle}(x\,{}^y_z) \qquad ❶$$

holds for any hinge $[x\,^y_z]$ with $|x-y|+|x-z|<\tfrac{2}{3}\cdot\ell$. Then the comparison ❶ *holds for any hinge $[x\,^y_z]$ with $|x-y|+|x-z|<\ell$.*

Proof. Given a hinge $[x\,^p_q]$ consider a triangle in the plane with angle $\measuredangle[x\,^p_q]$ and two adjacent sides $|x-p|$ and $|x-q|$. Let us denote by $\tilde\Upsilon[x\,^p_q]$ the third side of this triangle; let us call it *model side* of the hinge.



Note that the inequalities

$$\measuredangle[x\,^p_q]\geqslant\tilde\measuredangle(x\,^p_q)\quad\text{and}\quad\tilde\Upsilon[x\,^p_q]\geqslant|p-q|$$

are equivalent. So it is sufficient to prove that

❷ $$\tilde\Upsilon[x\,^p_q]\geqslant|p-q|.$$

for any hinge $[x\,^p_q]$ with $|x-p|+|x-q|<\ell$.

Given a hinge $[x\,^p_q]$ such that

$$\tfrac{2}{3}\cdot\ell\leqslant|p-x|+|x-q|<\ell,$$

let us construct a new smaller hinge $[x'\,^p_q]$; that is,

❸ $$|p-x|+|x-q|\geqslant|p-x'|+|x'-q|$$

and such that

❹ $$\tilde\Upsilon[x\,^p_q]\geqslant\tilde\Upsilon[x'\,^p_q].$$

Assume $|x-q|\geqslant|x-p|$, otherwise switch the roles of p and q in the following construction. Take $x'\in[xq]$ such that

❺ $$|p-x|+3\cdot|x-x'|=\tfrac{2}{3}\cdot\ell$$

Choose a geodesic $[x'p]$ and consider the hinge $[x'\,^p_q]$ fromed by $[x'p]$ and $[x'q]\subset[xq]$. Then ❸ follows since the length of $[x'p]$ can not exceed the total length of $[x'x]$ and $[x'p]$.

Further, note that $|p-x|+|x-x'|, |p-x'|+|x'-x|<\tfrac{2}{3}\cdot\ell$. In particular,

❻ $$\measuredangle[x\,^p_{x'}]\geqslant\tilde\measuredangle(x\,^p_{x'})\ \text{ and }\ \measuredangle[x'\,^p_x]\geqslant\tilde\measuredangle(x'\,^p_x).$$

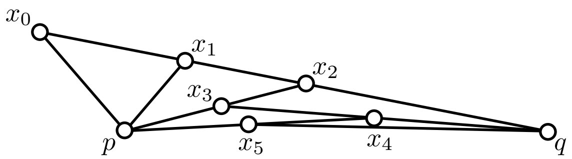
Consider the model triangle $[\tilde x\tilde x'\tilde p]=\tilde\triangle xx'p$. Take $\tilde q$ on the extension of $[\tilde x\tilde x']$ beyond x' such that $|\tilde x-\tilde q|=|x-q|$ (and therefore $|\tilde x'-\tilde q|=|x'-q|$). From ❻,

$$\measuredangle[x\,^p_q]=\measuredangle[x\,^p_{x'}]\geqslant\tilde\measuredangle(x\,^p_{x'})\quad\Rightarrow\quad\tilde\Upsilon[x\,^q_p]\geqslant|\tilde p-\tilde q|.$$

Since $\measuredangle[x'\,{}^p_x] + \measuredangle[x'\,{}^p_q] = \pi$, ❻ implies

$$\pi - \tilde\measuredangle(x'\,{}^p_x) \geqslant \pi - \measuredangle[x'\,{}^p_x] \geqslant \measuredangle[x'\,{}^p_q].$$

Therefore $|\tilde p - \tilde q| \geqslant \tilde\Upsilon[x'\,{}^q_p]$ and ❹ follows.



Set $x_0 = x$. Let us apply inductively the above construction to get a sequence of hinges $[x_n\,{}^p_q]$ with $x_{n+1} = x'_n$. By ❹ and triangle inequality, both sequences

$$s_n = \tilde\Upsilon[x_n\,{}^p_q] \quad \text{and} \quad r_n = |p - x_n| + |x_n - q|$$

are nonincreasing.

The sequence might terminate at some n only if $r_n < \frac{2}{3}\cdot\ell$. In this case, by the assumptions of the lemma,

$$s_n = \tilde\Upsilon[x_n\,{}^p_q] \geqslant |p - q|.$$

Since sequence s_n is nonincreasing;

$$s_0 = \tilde\Upsilon[x\,{}^p_q] \geqslant |p - q|,$$

whence inequality ❷ follows.

If the sequence does not terminate, then $r_n \geqslant \frac{2}{3}\cdot\ell$ for all n. Since (r_n) is nonincreasing, $r_n \to r \geqslant |p - q|_\Sigma$ as $n \to \infty$.

Let us show that $\measuredangle[x_n\,{}^p_q] \to \pi$ as $n \to \infty$.

Indeed assume $\measuredangle[x_n\,{}^p_q] \leqslant \pi - \varepsilon$ for some $\varepsilon > 0$. Without loss of generality we can assume that $x_{n+1} \in [x_n q]$; otherwise switch p and q further. Note that $|x_n - x_{n+1}|, |p - x_n| > \frac{\ell}{100}$. Therefore by comparison

$$|p - x_{n+1}| < \tilde\Upsilon[x_n\,{}^{p}_{x_{n+1}}] < |p - x_n| + |x_n - x_{n+1}| - \delta$$

for some fixed $\delta = \delta(\varepsilon) > 0$. Therefore $r_n - r_{n+1} > \delta$. The latter can not hold for large n, otherwise the sequence r_n would not converge.

It follows that for any $\varepsilon > 0$ we have that $\measuredangle[x_n\,{}^p_q] > \pi - \varepsilon$ for all large n; that is $\measuredangle[x_n\,{}^p_q] \to \pi$ as $n \to \infty$.

Since $\measuredangle[x_n\,{}^p_q] \to \pi$, we have $s_n - r_n \to 0$ as $n \to \infty$; that is $s_n \to r$. Since the sequence (s_n) is nonincreasing and $r \geqslant |p - q|$, we get

$$s_n \geqslant |p - q|$$

for any n. In particular

$$\tilde{\Upsilon}[s\,{}^{p}_{q}] = s_0 \geqslant |p-q|,$$

so we obtain ❷. $\square$

14.12. Exercise. *Assume a disc Δ lies in complete smooth regular surface Σ with nonnegative Gauss curvature and bounded by a closed broken geodesic $x_1\dots x_n$ with positive exterior angles, that is when you travel along the boundary, you always turn to the side where Δ is.*

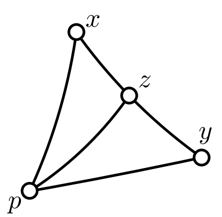
Show that there is a convex plane polygon $\tilde{x}_1\dots\tilde{x}_n$ which sides are equal to the corresponding sides of $x_1\dots x_n$ and with internal angles at not bigger than in the corresponding angles of $x_1\dots x_n$.

14.7 Inheritance lemma

The following lemma will play key role in the proof of part *(ii)* of the comparison theorem (14.1).

14.13. Inheritance Lemma. *Assume that a triangle $[pxy]$ in a surface Σ* decomposes *into two triangles $[pxz]$ and $[pyz]$; that is, $[pxz]$ and $[pyz]$ have common side $[pz]$, and the sides $[xz]$ and $[zy]$ together form the side $[xy]$ of $[pxy]$.*

If both triangles $[pxz]$ and $[pyz]$ are thin, then so is $[pxy]$.



We shall need the following lemma in plane geometry.

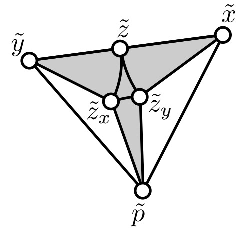
14.14. Lemma. *Let $\blacktriangle\tilde{p}\tilde{x}\tilde{y}$ be a solid plane triangle; that is, $\blacktriangle\tilde{p}\tilde{x}\tilde{y} = \mathrm{Conv}\{\tilde{p},\tilde{x},\tilde{y}\}$. Given $\tilde{z}\in[\tilde{x}\tilde{y}]$, consider points $\dot{p},\dot{x},\dot{z},\dot{y}$ in the plane such that*

$$\begin{aligned}
&|\dot{p}-\dot{x}| = |\tilde{p}-\tilde{x}|, && |\dot{p}-\dot{y}| = |\tilde{p}-\tilde{y}|, && |\dot{p}-\dot{z}| \leqslant |\tilde{p}-\tilde{z}|,\\
&|\dot{x}-\dot{z}| = |\tilde{x}-\tilde{z}|, && |\dot{y}-\dot{z}| = |\tilde{y}-\tilde{z}|,
\end{aligned}$$

where points $\dot{x}$ and $\dot{y}$ lie on either side of $[\dot{p}\dot{z}]$. Then there is a short map

$$F\colon \blacktriangle\tilde{p}\tilde{x}\tilde{y} \to \blacktriangle\dot{p}\dot{x}\dot{z} \cup \blacktriangle\dot{p}\dot{y}\dot{z}$$

that maps $\tilde{p}$, $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ to $\dot{p}$, $\dot{x}$, $\dot{y}$ and $\dot{z}$ respectively.



Proof. Note that

$$\begin{aligned}|\dot x-\dot y|&\leqslant|\dot x-\dot z|+|\dot z-\dot y|=\\&=|\tilde x-\tilde z|+|\tilde z-\tilde y|\\&=|\tilde x-\tilde y|.\end{aligned}$$

Applying monotonicity propery, we get that

$$\measuredangle[\dot p\,{}^{\dot x}_{\dot y}]\leqslant\measuredangle[\tilde p\,{}^{\tilde x}_{\tilde y}].$$

It follows that there are nonoverlapping triangles $[\tilde p\tilde x\tilde z_y]\cong[\dot p\dot x\dot z]$ and $[\tilde p\tilde y\tilde z_x]\cong[\dot p\dot y\dot z]$ inside triangle $[\tilde p\tilde x\tilde y]$.

Connect points in each pair $(\tilde z,\tilde z_x)$, $(\tilde z_x,\tilde z_y)$ and $(\tilde z_y,\tilde z)$ with arcs of circles centered at $\tilde y$, $\tilde p$, and $\tilde x$ respectively. Define F as follows.

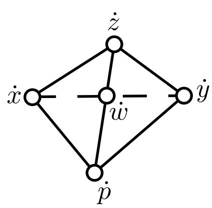
- ◇ Map $\blacktriangle\tilde p\tilde x\tilde z_y$ isometrically onto $\blacktriangle\dot p\dot x\dot y$; similarly map $\blacktriangle\tilde p\tilde y\tilde z_x$ onto $\blacktriangle\dot p\dot y\dot z$.
- ◇ If a point w lies in one of the three circular sectors, say at distance r from center of the circle, let $F(w)$ be the point on the corresponding segment $[\dot p\dot z]$, $[\dot x\dot z]$ or $[\dot y\dot z]$ whose distance from the lefthand endpoint of the segment is r.
- ◇ Finally, if w lies in the remaining curvilinear triangle $\tilde z\tilde z_x\tilde z_y$, set $F(w)=\dot z$.

By construction, F satisfies the remaining conditions of the lemma. □

Proof of Inheritance lemma 14.13. Construct model triangles $[\dot p\dot x\dot z]=$ $=\tilde\triangle(pxz)$ and $[\dot p\dot y\dot z]=\tilde\triangle(pyz)$ so that $\dot x$ and $\dot y$ lie on opposite sides of $[\dot p\dot z]$.

Suppose

$$\tilde\measuredangle(z\,{}^{p}_{x})+\tilde\measuredangle(z\,{}^{p}_{y})<\pi.$$



Then for some point $\dot w\in[\dot p\dot z]$, we have

$$|\dot x-\dot w|+|\dot w-\dot y|<|\dot x-\dot z|+|\dot z-\dot y|=|x-y|.$$

Let $w\in[pz]$ correspond to $\dot w$; that is, $|z-w|=|\dot z-\dot w|$. Since $[pxz]$ and $[pyz]$ are thin, we have

$$|x-w|+|w-y|<|x-y|,$$

contradicting the triangle inequality.

Thus

$$\tilde\measuredangle(z\,{}^{p}_{x})+\tilde\measuredangle(z\,{}^{p}_{y})\geqslant\pi.$$

By Alexandrov's lemma (14.5), this is equivalent to

❶ $$\tilde\measuredangle(x\,{}^p_z) \leqslant \tilde\measuredangle(x\,{}^p_y).$$

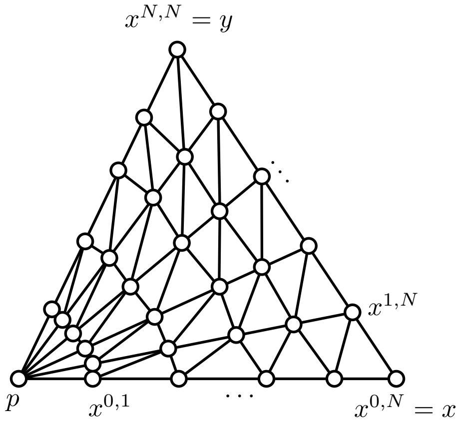
Let $[\tilde p\tilde x\tilde y] = \tilde\triangle(pxy)$ and $\tilde z \in [\tilde x\tilde y]$ correspond to z; that is, $|x - z| = |\tilde x - \tilde z|$. Inequality ❶ is equivalent to $|p - z| \leqslant |\tilde p - \tilde z|$. Hence Lemma 14.14 applies; let $F\colon \blacktriangle\tilde p\tilde x\tilde y \to \blacktriangle\dot p\dot x\dot z \cup \blacktriangle\dot p\dot y\dot z$ be the provided short map.

Fix v, w on the sides of $[pxy]$; let $\tilde v, \tilde w$ be the corresponding points on the sides of the model triangle $[\tilde p\tilde x\tilde y] = \tilde\triangle pxy$ and $\dot v, \dot w$ be the corresponding points on the sides of the model triangles $[\dot p\dot x\dot z] = \tilde\triangle pxz$ and $[\dot p\dot y\dot z] = \tilde\triangle pyz$. Denote by ℓ the length of shortest curve from $\dot v$ to $\dot w$ in $\blacktriangle\dot p\dot x\dot z \cup \blacktriangle\dot p\dot y\dot z$. Since F is short, $|\tilde v - \tilde w|_{\mathbb{R}^2} \geqslant \ell$. Since both triangles $[pxz]$ and $[pyz]$ are thin, $\ell \geqslant |v - w|_\Sigma$.

It follows that $|\tilde v - \tilde w|_{\mathbb{R}^2} \geqslant |v - w|_\Sigma$ for any v and w; that is, the triangle $[pxy]$ is thin. □

14.8 Nonpositive curvature

Assume Σ is a complete smooth regular surface with nonpositive curvature. As it follow from Exercise 14.2 any two points x and y in Σ are joined by unique geodesic $[xy]$.



Note that the geodesic $[xy]$ depends continuously on its endpoints x and y. That is, if $\gamma_{[xy]}\colon [0,1] \to \Sigma$ is the constant speed parametrization of $[xy]$ from x to y, then the map $(x, y, t) \mapsto \gamma_{[xy]}(t)$ is continuous in three arguments. Indeed, assume contrary, that is $x_n \to x$, $y_n \to y$ and $t_n \to t$ as $n \to \infty$ and $\gamma_{[x_ny_n]}(t_n)$ does not converge to $\gamma_{[xy]}(t)$. Then we can pass to a subsequence such that $\gamma_{[x_ny_n]}(t_n)$ converges to a point distinct from $w \neq \gamma_{[xy]}(t)$. Note that $w \notin [xy]$. Therefore there will be two distinct geodesics from x to y; one is $[xy]$ and the other is the limit of $[x_ny_n]$ which passes thru w.

Proof of part (ii) of the comparison theorem (14.1). Fix a triangle $[pxy]$; by Proposition 14.6, it is sufficient to show that the triangle $[pxy]$ is thin.

Fix large integer N and divide $[xy]$ by points $x = x^{0,N}, \dots, x^{N,N} = y$ into N equal parts. Further divide each geodesic $[p\,x^{i,N}]$ into N equal parts by points $p = x^{i,0}, \dots, x^{i,N}$. Since the geodesic depends continuously on its end points, we can assume that each triangle $[x^{i,j}\,x^{i,j+1}\,x^{i+1,j+1}]$ and $[x^{i,j}\,x^{i+1,j}\,x^{i+1,j+1}]$ is small; in particular, by local comparison (14.3), each of these triangles is thin.

Now we show that the thin property propagates to $[pxy]$ by repeated application of the inheritance lemma (14.13):

- ◇ First, for fixed i, sequentially applying the lemma shows that the triangles $[x\,x^{i,1}\,x^{i+1,2}]$, $[x\,x^{i,2}\,x^{i+1,2}]$, $[x\,x^{i,2}\,x^{i+1,3}]$, and so on are thin.

In particular, for each i, the long triangle $[x\,x^{i,N}\,x^{i+1,N}]$ is thin.

- ◇ Applying the lemma again shows that the triangles $[x\,x^{0,N}\,x^{2,N}]$, $[x\,x^{0,N}\,x^{3,N}]$, and so on are thin.

In particular, $[pxy] = [p\,x^{0,N}\,x^{N,N}]$ is thin. □

14.15. Exercise. *Assume γ_1 and γ_2 be two geodesics in complete smooth regular simply connected surface Σ with nonpositive Gauss curvature. Show that the function*

$$h(t) = |\gamma_1(t) - \gamma_2(t)|_\Sigma$$

is convex.

Chapter 15

Saddle surfaces

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.

Note that a closed surface can not 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 exercise can be solved using the same idea.

15.1. Exercise. *Show that a smooth surface Σ is saddle if and only if it has no hats; that is no disc Δ in Σ that boundary lies in a plane and the remaining points of the Δ lie on one side of the plane.*

A surface Σ is called *ruled* if thru every point of Σ there is a straight line that lies on Σ.

15.2. Exercise. *Show that any ruled surface is saddle.*

Appendix A

Semisolutions

Exercise ??. First let us show that Dido's problem follows from the isoperimetric inequality.

Assume F is a figure bounded by a straight line and a curve of length ℓ whose endpoints belong to that line. Let F' be the reflection of F in the line. Note that the union $G = F \cup F'$ is a figure bounded by a closed curve of length $2{\cdot}\ell$.

Applying the isoperimetric inequality, we get that the area of G can not exceed the area of round disc with the same circumference $2{\cdot}\ell$ and the equality holds only if the figure is congruent to the disc. Since F and F' are congruent, Dido's problem follows.

Now let us show that the isoperimetric inequality follows from the Dido's problem.

Assume G is a convex figure bounded by a closed curve of length $2{\cdot}\ell$. Cut G by a line that splits the perimeter in two equal parts — ℓ each. Denote by F and F' the two parts. Applying the Dido's problem for each part, we get that that are of each does not exceed the area of half-disc bounded by a half-circle. The two half-disc could be arranged into a round disc of circumference ℓ, hence the isoperimetric inequality follows.

Exercise 2.16. Let $\alpha\colon [a,b] \to \mathbb{R}^3$ be a curve. Given a unit vector u, denote by α_u the projection of α on a line in the direction of u; denote by $\alpha_{u^\perp}$ the of α on a plane perpendicular to u.

Two formulas

$$\operatorname{length}\alpha = k{\cdot}\overline{\operatorname{length}\alpha_u}$$

and

$$\operatorname{length}\alpha = k'{\cdot}\overline{\operatorname{length}\alpha_{u^\perp}}$$

can be proved the same way as the Crofton's formula in the plane.

It remains to find the coefficients k and k'. It is sufficient to calculate the average projection of unit segment to a line and to a plane. We need to find two integrals

$$k = \fint_{\mathbb{S}^2} |x|\cdot d\,\mathrm{area}$$

and

$$k' = \fint_{\mathbb{S}^2} \sqrt{1-x^2}\cdot d\,\mathrm{area},$$

where $\mathbb{S}^2 = \{\, (x,y,z)\in\mathbb{R}^3 \,:\, x^2+y^2+z^2=1 \,\}$ is the unit sphere in the Euclidean space and $\fint$ denotes the average value — since the area of unit sphere is $4\cdot\pi$, we have

$$\fint_{\mathbb{S}^2} f(x,y,z)\cdot d\,\mathrm{area} = \tfrac{1}{4\cdot\pi}\cdot\int_{\mathbb{S}^2} f(x,y,z)\cdot d\,\mathrm{area}$$

Note that in the cylindrical coordinates

$$(x, \varphi = \arctan\tfrac{y}{z}, \rho = \sqrt{y^2+z^2}),$$

we have $d\,\mathrm{area} = dx\cdot d\varphi$. Therefore

$$k = \fint_{[-1,1]} |x|\cdot dx = \tfrac{1}{2}$$

and

$$k' = \fint_{[-1,1]} \sqrt{1-x^2}\cdot dx = \tfrac{\pi}{4}.$$

Comment. Note that $\frac{k'}{k} = \frac{\pi}{2}$ is the coefficient in the 2-dimensional Crofton formula. This is not a coincidence — think about it.

Exercise 3.1. Assume contrary, that is there is a closed smooth regular curve α such that $\operatorname{TotCurv}\alpha < 2\cdot\pi$.

The tangent indicatrix τ of α is a curve in a sphere; by the definition of total curvature, the length of τ is the total curvature of α; in particular

$$\operatorname{length}\tau < 2\cdot\pi.$$

By Exercise 2.19, τ lies in an open hemisphere. If u is the center of the hemisphere, then

$$\langle u,\tau(t)\rangle > 0 \quad \text{and therefore} \quad \langle u,\alpha'(t)\rangle > 0$$

for any t. Therefore the function $t \mapsto \langle u,\alpha(t)\rangle$ is strictly increasing. In particular, if α is defined on the time interval $[a,b]$, then

$$\langle u,\alpha(a)\rangle < \langle u,\alpha(a)\rangle.$$

But α is closed; that is $\alpha(a)=\alpha(b)$ — a contradiction.

Now let us prove the equality case. First note that it is sufficient to show that τ runs around an equator.

Assume τ is not an equator, from above we know that τ can not lie in an open hemisphere. Note that we can shorten τ by a small chord. The obtained curve τ' is shorter than $2{\cdot}\pi$ and therefore lies in an open hemisphere. Applying this construction for shorter and shorter chord and passing to the limit we get that τ lies in closed hemisphere. Denote its center by u as before, then

$$\langle u,\tau(t)\rangle \geqslant 0 \quad \text{and therefore} \quad \langle u,\alpha'(t)\rangle \geqslant 0$$

for any t. Since α is closed we have that $\langle u,\alpha(t)\rangle$ is constant; that is, runs in a plane perpendicular to u and τ lies in an equator perpendicular to u.

So τ is a curve that runs along equator, has length $2{\cdot}\pi$ and does not lie in a open hemisphere. Since τ is not an equator, it have to run along half-equator back and forth. In this case τ lies in an other closed hemisphere and has some points in its interior. The latter contradicts closeness of α the same way as above.

Bibliography

[1] G. Lawlor. "A new area-maximization proof for the circle". *The Mathematical Intelligencer* 20.1 (1998), pp. 29–31.

[2] D. Fuchs and S. Tabachnikov. *Mathematical omnibus*. Thirty lectures on classic mathematics. American Mathematical Society, Providence, RI, 2007.

[3] J. Lagarias and T. Richardson. "Convexity and the average curvature of plane curves". *Geom. Dedicata* 67.1 (1997), pp. 1–30.

[4] A. I. Nazarov and F. V. Petrov. "On a conjecture of S. L. Tabachnikov". *Algebra i Analiz* 19.1 (2007), pp. 177–193.

[5] K. Borsuk. "Sur la courbure totale des courbes fermées". *Ann. Soc. Polon. Math.* 20 (1947), 251–265 (1948).

[6] I. Fáry. "Sur la courbure totale d'une courbe gauche faisant un nœud". *Bull. Soc. Math. France* 77 (1949), pp. 128–138.

[7] J. Milnor. "On the total curvature of knots". *Ann. of Math. (2)* 52 (1950), pp. 248–257.

[8] S. Alexander and R. Bishop. "The Fary-Milnor theorem in Hadamard manifolds". *Proc. Amer. Math. Soc.* 126.11 (1998), pp. 3427–3436.

[9] E. Denne. *Alternating quadrisecants of knots*. Thesis (Ph.D.)–University of Illinois at Urbana-Champaign. ProQuest LLC, Ann Arbor, MI, 2004, p. 119.

[10] J. Schultens. "Additivity of bridge numbers of knots". *Math. Proc. Cambridge Philos. Soc.* 135.3 (2003), pp. 539–544.

[11] P. G. Tait. "Note on the circles of curvature of a plane curve." *Proc. Edinb. Math. Soc.* 14 (1896), p. 26.

[12] 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.

[13] S. Mukhopadhyaya. "New methods in the geometry of a plane arc". *Bull Calcutta Math Soc* 1 (1909), pp. 31–37.

[14] R. Osserman. "The four-or-more vertex theorem". *Amer. Math. Monthly* 92.5 (1985), pp. 332–337.

[15] 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.

[16] W. Blaschke. *Kreis und Kugel*. Verlag von Veit & Comp., Leipzig, 1916.

[17] V. N. Lagunov. "On the largest sphere contained in a closed surface". *Sibirsk. Mat. Ž.* 1 (1960), pp. 205–232.

[18] K. Baker. *Bing's House*. Mar. 25, 2010. URL: `https://sketchesoftopology.wordpress.com/2010/03/25/bings-house/`.

[19] J. J. Stoker. "Über die Gestalt der positiv gekrümmten offenen Flächen im dreidimensionalen Raume". *Compositio Mathematica* 3 (1936), pp. 55–88.

[20] J. Hadamard. “Sur certaines propriétés des trajectoires en dynamique”. *J. math. pures appl.* 3 (1897), pp. 331–387.

[21] J. Liberman. “Geodesic lines on convex surfaces”. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 32 (1941), pp. 310–313.

[22] 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.

[23] 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.

[24] 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.

[25] M. Levi. “A “bicycle wheel” proof of the Gauss–Bonnet theorem”. *Exposition. Math.* 12.2 (1994), pp. 145–164.

[26] K. F. Gauss. “Disquisitiones generales circa superficies curvas”. *Commentationes Societatis Regiae Scientiarum Gottingensis recentiores* 6 (classis mathematicae) (1828). [Translated to English in *General investigations of curved surfaces* 1902.], pp. 99–146.

[27] A. Petrunin and A. Yashinski. “Piecewise isometric mappings”. *Algebra i Analiz* 27.1 (2015), pp. 218–247.

[28] S. Cohn-Vossen. „Totalkrümmung und geodätische Linien auf einfachzusammenhängenden offenen vollständigen Flächenstücken“. . . 1(43).2 (1936), pp. 139–164.

[29] H. E. Rauch. “A contribution to differential geometry in the large”. *Ann. of Math. (2)* 54 (1951), pp. 38–55.

[30] P. Pizzetti. “Paragone fra due triangoli a lati uguali”. *Atti della Reale Accademia dei Lincei, Rendiconti (5) Classe di Scienze Fisiche, Matematiche e Naturali* 16 (1) (1907), pp. 6–11.

[31] A. D. Alexandrov. *Vnutrennyaya Geometriya Vypuklyh Poverhnosteı.* OGIZ, Moscow-Leningrad,] 1948, p. 387.

[32] V. A. Toponogov. “On convexity of Riemannian spaces of positive curvature”. *Dokl. Akad. Nauk SSSR (N.S.)* 115 (1957), pp. 674–676.

[33] Yu. Burago, M. Gromov, and G. Perelman. “A. D. Aleksandrov spaces with curvatures bounded below”. *Uspekhi Mat. Nauk* 47.2(284) (1992), pp. 3–51, 222.

[34] H. von Mangoldt. “Ueber diejenigen Punkte auf positiv gekrümmten Flächen, welche die Eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein”. *J. Reine Angew. Math.* 91 (1881), pp. 23–53.

[35] É. Cartan. *Leçons sur la Géométrie des Espaces de Riemann.* 1928.

[36] H. Busemann. “Spaces with non-positive curvature”. *Acta Math.* 80 (1948), pp. 259–310.

[37] W. Rinow. *Die innere Geometrie der metrischen Räume.* Die Grundlehren der mathematischen Wissenschaften, Bd. 105. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1961, pp. xv+520.

[38] M. Gromov. “Hyperbolic groups”. *Essays in group theory.* Vol. 8. Math. Sci. Res. Inst. Publ. Springer, New York, 1987, pp. 75–263.

[39] S. Alexander and R. Bishop. “The Hadamard–Cartan theorem in locally convex metric spaces”. *Enseign. Math. (2)* 36.3-4 (1990), pp. 309–320.