The Complex Plane and The Euclidean Plane

Algebraic Approach to Plane Geometry

Michael Taylor

Math. Dept., UNC
E-mail address: met@math.unc.edu

2010 *Mathematics Subject Classification.* 22-01

Key words and phrases. complex numbers, Euclidean geometry, Pythagorean theorem, inner products, isometries, triangles, angle measurements, congruence, similarity, squares, rectangles, polygons, circles, calculus, curves, arclength, exponential and trigonometric functions, Euler formula, real measurement of angles, π, area, Euclidean numbers, linear fractional transformations, inner product spaces, volume, area of spheres

Contents

Introduction 1

Chapter 1. Basic algebraic results on Euclidean geometry 13

§1.1. Metric properties of $\mathbb{C}$ 15

§1.2. Rigid motions on $\mathbb{C}$ 17

§1.3. Lines and line segments 20

§1.4. Rays, wedges, and angles 24

§1.5. Triangles 26

§1.6. Congruence of triangles 31

§1.7. Isosceles triangles and equilateral triangles 37

§1.8. Similarity 39

§1.9. Squares, rectangles, and other quadrilaterals 42

§1.10. Circles 46

Chapter 2. Bringing in calculus 53

§2.1. Outline of calculus 56

§2.2. Curves, arclength, and the real measurement of angles 61

§2.3. Exponential and trigonometric functions, and Euler's formula 67

§2.4. More on Euler's formula, trigonometric functions, and π 73

§2.5. Regular polygons 79

§2.6. Area 84

§2.7. Heron's formula 91

§2.8. Euclidean numbers 94

§2.9. Linear fractional transformations 97

§2.10. Higher dimensions 107

Bibliography 123

Index 125

Introduction

The set $\mathbb{C}$ of complex numbers arose as an extension of the field $\mathbb{R}$ of real numbers in which one could solve the equation $z^2 = -1$. This was obtained by adding an imaginary unit i, satisfying $i^2 = -1$. A complex number $z \in \mathbb{C}$ has the form

$$z = x + iy, \quad x, y \in \mathbb{R}. \tag{0.0.1}$$

We have $\mathbb{R} \hookrightarrow \mathbb{C}$ via $x \mapsto x + i0$. Given z as in (0.0.1), we define $\operatorname{Re} z,\ \operatorname{Im} z \in \mathbb{R}$ by

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y, \quad \text{so} \quad z = \operatorname{Re} z + i \operatorname{Im} z. \tag{0.0.2}$$

It was soon realized that it is convenient to visualize $\mathbb{C}$ as the *complex plane*, and to utilize the insights of Euclidean geometry. Turning this around, one can perceive that the algebraic structure of $\mathbb{C}$ provides an ideal set of tools with which to obtain basic geometric results set out in the books of Euclid. We describe how this works here.

In this introduction we lay out the the basic algebraic structure of $\mathbb{C}$. In subsequent sections we examine the metric properties of $\mathbb{C}$, discuss the rigid motions of $\mathbb{C}$, and derive basic results of Euclidean plane geometry, including the crown jewel of the subject, the Pythagorean theorem.

We start with addition. If z is as in (0.0.1) and also $w = u + iv,\ u, v \in \mathbb{R}$, we define

$$z + w = (x + u) + i(y + v). \tag{0.0.3}$$

Given that we want multiplication in $\mathbb{C}$ to satisfy

$$(x + i0)(u + i0) = xu, \quad (0 + iy)(0 + iv) = -yv, \tag{0.0.4}$$

and expect to obtain the various commutative, associative, and distributive laws, we are led to the following definition of multiplication on $\mathbb{C}$:

$$(x+iy)(u+iv) = (xu-yv) + i(xv+yu). \tag{0.0.5}$$

One readily verifies the standard commutative, associative, and distributive laws, given these definitions of addition and multiplication on $\mathbb{C}$, as a consequence of their counterparts on $\mathbb{R}$.

The set $\mathbb{C}$ also possesses an important map called *complex conjugation*: for $x, y \in \mathbb{R}$,

$$z = x+iy \Longrightarrow \overline{z} = x - iy. \tag{0.0.6}$$

Note that

$$2\,\mathrm{Re}\, z = z + \overline{z}, \quad 2\,\mathrm{Im}\, z = -i(z-\overline{z}). \tag{0.0.7}$$

One readily verifies that, for $z, w \in \mathbb{C}$,

$$\overline{z+w} = \overline{z} + \overline{w}, \quad \overline{zw} = \overline{z}\,\overline{w}. \tag{0.0.8}$$

Another important identity is

$$z\overline{z} = x^2 + y^2. \tag{0.0.9}$$

With this in mind, we define the *absolute value* of z as

$$|z| = (z\overline{z})^{1/2} = \sqrt{x^2+y^2}. \tag{0.0.10}$$

Note that

$$z \neq 0 \Longrightarrow |z| > 0. \tag{0.0.11}$$

Also, by (0.0.8), together with the commutative and associative laws of multiplication,

$$|zw| = |z|\,|w|. \tag{0.0.12}$$

Furthermore,

$$z \neq 0 \Rightarrow z\Big(\frac{1}{|z|^2}\overline{z}\Big) = 1 \Rightarrow \frac{1}{z} = \frac{1}{|z|^2}\overline{z}. \tag{0.0.13}$$

In such a case, we can divide:

$$\frac{w}{z} = w\,\frac{1}{z} = \frac{1}{|z|^2}w\overline{z}. \tag{0.0.14}$$

The algebraic properties established above make $\mathbb{C}$ an object known as a field. The identity (0.0.12) makes it a normed field.

Let us now outline the material covered in the subsequent sections. Chapter 1 consists of Sections 1.1–1.10 and deals with basic results on metric properties of $\mathbb{C}$, including rigid transformations, and such geometric concepts as lines, triangles, congruences, similarity, squares, rectangles, and circles.

In §1.1 we describe metric properties of $\mathbb{C}$. Our study starts with the formula

$$|z+w|^2 = |z|^2 + |w|^2 + 2\langle z, w\rangle, \tag{0.0.15}$$

bringing in the inner product $\langle z, w\rangle = \operatorname{Re} z\overline{w}$ of two elements of $\mathbb{C}$. We say $z \perp w$ provided $\langle z, w\rangle = 0$. We define the distance $d(z, w) = |z - w|$ and establish the triangle inequality.

In §1.2 we define the class $\mathrm{Isom}(\mathbb{C})$ of rigid motions of $\mathbb{C}$ to consist of maps $F : \mathbb{C} \to \mathbb{C}$ that preserve distance. We show that $\mathrm{Isom}(\mathbb{C}) = \mathrm{Isom}^+(\mathbb{C}) \cup \mathrm{Isom}^-(\mathbb{C})$, whose elements have the form

$$F(z) = az + b, \quad a\overline{z} + b, \quad a \in S^1,\ b \in \mathbb{C}, \tag{0.0.16}$$

respectively, where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$.

In §1.3 we discuss lines and line segments. We discuss the notions of orthogonal (perpendicular) lines and parallel lines, and show that if L is a line, $p \in \mathbb{C}$, then through p there is a unique line perpendicular to L and, if $p \notin L$, a unique line parallel to L.

In §1.4 we discuss rays and wedges. A ray is a half-line, $\{p + tu : t \geq 0\}$, and a wedge $W_{p,u,v}$ is a pair of rays, emanating from the same base point p. We define the angle measurement $\Omega(W_{p,u,v})$ as an element $\omega \in S^1$, and give a formula for bisecting an angle. Eventually we want to measure angles in "radians," but that comes later, in §2.2.

In §1.5 we consider triangles, $\mathcal{T} = \triangle(A, B, C)$, with vertices at $A, B, C \in \mathbb{C}$. We define angle measurements at each vertex, e.g., $\Omega_{\mathcal{T}}(A)$, as elements of S^1, and show that

$$\Omega_{\mathcal{T}}(A)\Omega_{\mathcal{T}}(B)\Omega_{\mathcal{T}}(C) = -1. \tag{0.0.17}$$

We also formally state the Pythagorean theorem, that if $\mathcal{T} = \triangle(A, B, C)$ is a right triangle, with right angle at C, then the associated side lengths satisfy

$$|c|^2 = |a|^2 + |b|^2. \tag{0.0.18}$$

This result is essentially contained in (0.0.15).

In §1.6 we study congruence of triangles, which occurs if $\mathcal{T}' = F(\mathcal{T})$ for some $F \in \mathrm{Isom}(\mathbb{C})$. We establish three types of criteria guaranteeing congruence:

(0.0.19)
(a) side-angle-side,
(b) angle-side-angle,
(c) side-side-side.

We also characterize which triples of positive numbers can be sidelengths of a triangle.

In §1.7 we study isosceles triangles, which have two sides of equal length. Say the two sides meet at the vertex C. We show that there exists $F \in \mathrm{Isom}(\mathbb{C})$ such that F preserves C and switches A and B. Hence the angle measurements of A and B are equal,

$$\Omega_{\mathcal{T}}(A) = \Omega_{\mathcal{T}}(B). \tag{0.0.20}$$

Conversely, if (0.0.20) holds, $\mathcal{T}$ is isosceles. An equilateral triangle is one all of whose sidelengths are equal. It is isosceles in three ways, yielding a variety of symmetries.

In §1.8 we study similarity. A map $F : \mathbb{C} \to \mathbb{C}$ is similar provided its application scales distances by some fixed factor. Such a map has a form like (0.0.16), except $a \in \mathbb{C}$ can be any nonzero element, not necessarily satisfying $|a| = 1$. We see that two triangles with identical angle measurements are similar. We display three similar triangles arising from a right triangle, divided by a perpendicular dropped from the right angle to the opposite side, and relate this to the proof of the Pythagorean theorem given in Book VI of Euclid.

Section 1.9 studies squares, rectangles, and more general figures, including parallelograms, as well as other quadrilaterals (4-gons), and further figures with n vertices (n-gons).

Section 1.10 introduces circles, of the form

$$S_R(a) = \{z \in \mathbb{C} : |z - a| = R\}, \tag{0.0.21}$$

given $R > 0,\ a \in \mathbb{C}$. The case $S_1(0)$ is S^1, encountered earlier. We characterize when two circles $S_R(a)$ and $S_r(b)$ intersect. We show that when $v \in S^1$, $v \neq \pm 1$, the triangle $\triangle(v, -1, 1)$ is a right triangle, with right angle at v. We relate this fact to a geometrical construction of the square root of a positive number. We note that

$$\varphi(x) = \frac{x + i}{x - i} \text{ yields } \varphi : \mathbb{R} \to S^1 \setminus \{1\}, \text{ bijective,} \tag{0.0.22}$$

and present an application of this to specifying the rational points on S^1, yielding a description of Pythagorean triples.

Chapter 2, which consists of Sections 2.1–2.10, brings in more sophistication, motivated by the study of circles. The subject needed for an adequate treatment of the geometrical properties of circles, such as arclength and area, is introduced in §2.1 – calculus. Our discussion is necessarily abbreviated; one can find full details in Chapter 4 of [**4**]. We define the derivative and the integral and discuss the fundamental theorem of calculus, which relates these two concepts. We discuss some other key propositions, such as the mean value theorem and the inverse function theorem. Another key tool introduced here is power series, of the form

$$f(x) = \sum_{k=0}^{\infty} a_k x^k. \tag{0.0.23}$$

We show that if (0.0.23) converges for $|x| < R$, then f is differentiable on the interval $(-R, R)$, and its derivative $f'(x)$ is given by term-by-term differentiation of the series (0.0.23).

Section 2.2 uses methods of calculus to treat curves and arclength. A curve in $\mathbb{C}$ is given by a continuous map $\gamma : [a, b] \to \mathbb{C}$. We define its length in terms of a sequence of approximations by unions of line segments, obtaining $\ell(\gamma)$ by passing to the limit (if it exists). We show that if γ is continuously differentiable (i.e., C^1), then

$$\ell(\gamma) = \int_a^b |\gamma'(t)|\, dt. \tag{0.0.24}$$

We discuss the operation of reparametrizing γ, replacing it by $\sigma(t) = \gamma(u(t))$, where $u : [\alpha, \beta] \to [a, b]$ is a C^1 map with C^1 inverse. Using the inverse function theorem, we show that a C^1 curve with nowhere vanishing derivative can be reparametrized by arclength, so $|\sigma'(t)| \equiv 1$. We apply this to the unit circle S^1, obtaining the parametrization

$$\text{cis} : \mathbb{R} \longrightarrow S^1, \quad \text{cis}(t) = \cos t + i \sin t. \tag{0.0.25}$$

We use this to define a real angle measurement. If $\omega \in S^1$, then

$$\Omega(W_{0,1,\omega}) = \omega = \text{cis}\, t, \quad \angle(W_{0,1,\omega}) = t, \tag{0.0.26}$$

for $-\pi < t \le \pi$, where the number π is defined by

$$\pi = \frac{1}{2}\ell(S^1). \tag{0.0.27}$$

We also show that

$$\frac{d}{dt}\,\text{cis}\, t = i\ \text{cis}\, t \tag{0.0.28}$$

and we give an integral formula for π. More will follow in subsequent sections.

In §2.3 we introduce the exponential function. This is given as the power series

$$e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k, \quad z \in \mathbb{C}, \tag{0.0.29}$$

derived to produce a solution to the differential equation

$$\frac{d}{dt} e^{at} = a e^{at}, \quad a \in \mathbb{C}. \tag{0.0.30}$$

An analysis of $\gamma(t) = e^{it}$ shows that it is a unit-speed parametrization of S^1, and we obtain the famous Euler formula

$$e^{it} = \operatorname{cis} t. \tag{0.0.31}$$

Applying (0.0.30) with $a = i$ gives a second proof of (0.0.28). Also, by (0.0.27), π is the smallest positive number satisfying

$$e^{\pi i} = -1. \tag{0.0.32}$$

We also show that

$$e^{a+b} = e^a e^b, \quad a, b \in \mathbb{C}, \tag{0.0.33}$$

hence $e^{i(s+t)} = e^{is}e^{it}$. We show that, if $\mathcal{T} = \triangle(A, B, C)$ is a triangle, then

$$\Omega_{\mathcal{T}}(A) = \alpha = e^{is},\ \Omega_{\mathcal{T}}(B) = \beta = e^{it},\ \Omega_{\mathcal{T}}(C) = \gamma = e^{iu}, \tag{0.0.34}$$

where either $s, t, u \in (0, \pi)$ or $s, t, u \in (-\pi, 0)$, and deduce from (0.0.17) that

$$s + t + u = \pi \quad \text{or} \quad -\pi. \tag{0.0.35}$$

Section 2.4, true to its title, provides more results on Euler's formula, trigonometric functions, and π. Supplementing (0.0.32), we have

$$e^{\pi i/2} = i,\ e^{\pi i/4} = \frac{1+i}{\sqrt{2}},\ e^{\pi i/3} = \frac{1}{2} + \frac{i}{2}\sqrt{3},\ e^{\pi i/6} = \frac{\sqrt{3}}{2} + \frac{i}{2}. \tag{0.0.36}$$

One consequence of the last formula is the identity

$$\frac{\pi}{6} = \int_0^{1/2} \frac{ds}{\sqrt{1-s^2}}. \tag{0.0.37}$$

We show how a power series expansion of $(1 - s^2)^{-1/2}$ can be integrated term by term to yield an infinite series for π that converges fairly rapidly. Other series arise from examining

$$\tan t = \frac{\sin t}{\cos t}, \quad \frac{d}{dt} \tan t = 1 + \tan^2 t, \tag{0.0.38}$$

yielding

$$\tan^{-1} y = \int_0^y \frac{dx}{1+x^2}, \tag{0.0.39}$$

whose integrand can be expanded in a power series and integrated term by term. From (0.0.36) we have

$$\tan \frac{\pi}{6} = \frac{1}{\sqrt{3}}, \tag{0.0.40}$$

and this yields an infinite series for $\pi/6$ alternative to (0.0.37). Other identities, yielding other series, include

$$\frac{\pi}{4} = \tan^{-1}\frac{1}{3} + \tan^{-1}\frac{1}{2}, \quad \frac{\pi}{4} = \tan^{-1}\frac{1}{7} + 2\tan^{-1}\frac{1}{3}. \tag{0.0.41}$$

As another approach to evaluating π, we approach π as a solution to $\sin x = 0$ using the iteration

$$x_{k+1} = x_k + \sin x_k. \tag{0.0.42}$$

As seen in §2.4, starting with $x_0 = 3$ yields x_3, which agrees with π to more than 15 digits.

Section 2.5 studies regular polygons. The regular n-gon $\mathcal{P}_n$ has n vertices lying on S^1, satisfying $z^n - 1 = 0$. The vertices are given by

$$\omega_n^k = e^{2k\pi i/n}, \quad 0 \le k \le n-1. \tag{0.0.43}$$

The identities in (0.0.36) yield formulas for the vertices of $\mathcal{P}_n$ for $3 \le n \le 8$, except for $n = 5$ and $n = 7$. We show that $\mathcal{P}_5$ can be constructed based on the identity

$$\cos \frac{2\pi}{5} = \frac{\sqrt{5} - 1}{4}. \tag{0.0.44}$$

This involves looking at

$$z^4 + z^3 + z^2 + z + 1 = 0, \tag{0.0.45}$$

and setting $w = z + 1/z$, obtaining a quadratic equation for w, solved by the quadratic formula. An attempt to treat $z^7 - 1 = 0$ in a parallel fashion yields a cubic equation for w, whose solution does not have a form amenable to classic construction via compass and straightedge.

Section 2.6 treats the problem of defining and computing the areas of various planar figures. If $S \subset \mathbb{C}$ is a bounded set, we define $\text{Cont}^+(S)$ and $\text{Cont}^-(S)$ in terms of how many tiles in a tiling of $\mathbb{C}$ it takes to contain S, resp., how many are contained in S, and passing to the limit as the sizes of the tiles vanishes. If these quantities are equal, we say S is *contented*, and set

$$\text{Area}\, S = \text{Cont}^{\pm}(S). \tag{0.0.46}$$

We give a criterion for S to be contented, in terms of the nature of the boundary of S. We treat areas of squares, rectangles, and triangles, and use some results on these objects to show that, if S is contented,

$$\text{Area}\, S = \text{Area}(F(S)), \quad \forall\, F \in \text{Isom}(\mathbb{C}). \tag{0.0.47}$$

From there, we consider the disk, $D_1 = \{z \in \mathbb{C} : |z| \le 1\}$. We show that

$$\text{Area}\, D_1 = \pi. \tag{0.0.48}$$

More generally, if Π_t denotes the "pie slice"

$$\Pi_t = \{re^{is} : 0 \le r \le 1,\, 0 \le s \le t\} \tag{0.0.49}$$

given $t \in (0, 2\pi]$, we show that

$$\text{Area}\, \Pi_t = \frac{1}{2} t, \tag{0.0.50}$$

the computation (0.0.48) being the special case $t = 2\pi$. Behind such computations is the result that if

$$\mathcal{O} = \{z = x + iy : a \le x \le b,\ \varphi(x) \le y \le \psi(x)\}, \tag{0.0.51}$$

where $\varphi, \psi : [a, b] \to \mathbb{R}$ are continuous, $\varphi \le \psi$, then

$$\text{Area}\, \mathcal{O} = \int_a^b [\psi(x) - \varphi(x)]\, dx. \tag{0.0.52}$$

Section 2.7 treats Heron's formula for the area of a triangle. It says that if $\mathcal{T}$ is a triangle with side lengths a, b, and c, then

$$(\text{Area}\, \mathcal{T})^2 = s(s-a)(s-b)(s-c), \tag{0.0.53}$$

where

$$s = \frac{1}{2}(a + b + c). \tag{0.0.54}$$

We include this just because the formula (0.0.53) looks intriguing. The proof is a straightforward application of the Pythagorean theorem.

Section 2.8 discusses the set $\mathbb{E} \subset \mathbb{C}$ of "Euclidean numbers." This set contains $\{0, 1\}$ and is closed under the operations of taking intersections of various Euclidean lines and/or Euclidean disks. The set $\mathbb{E}$ is designed to consist of those points in $\mathbb{C}$ that can be constructed via compass and straightedge. We show that $\mathbb{E}$ is a field, i.e.,

$$a, b \in \mathbb{E} \Rightarrow a + b, a - b, ab, \frac{a}{b} \in \mathbb{E}, \tag{0.0.55}$$

the last conclusion holding provided $b \neq 0$, and we show that

$$z \in \mathbb{E},\ w \in \mathbb{C},\ w^2 = z \Longrightarrow w \in \mathbb{E}. \tag{0.0.56}$$

The set $\mathbb{E}$ is also seen to be countable.

Section 2.9 considers an important extension of the family of maps $z \mapsto az + b$, making up $\text{Isom}^+(\mathbb{C})$ and more generally $\text{Sim}^+(\mathbb{C})$. This larger class consists of linear fractional transformations, which have the form

$$L_A(z) = \frac{az + b}{cz + d}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{0.0.57}$$

Here A is an invertible 2×2 matrix, i.e., $A \in G\ell(2, \mathbb{C})$. If $c \neq 0$, this fraction is well defined except at $z = -d/c$. We extend the map to

$$L_A : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}, \quad \widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}, \tag{0.0.58}$$

and note that

$$L_A \circ L_B = L_{AB}. \tag{0.0.59}$$

We note that L_A maps the upper half plane $\mathcal{U}$ onto itself if $A \in S\ell(2,\mathbb{R})$ and it maps the unit disk $\mathcal{D}$ onto itself if $A \in SU(1,1)$, a class of matrices defined in (2.9.18). In fact, we see that $S\ell(2,\mathbb{R})$ acts transitively on $\mathcal{U}$ and $SU(1,1)$ acts transitively on $\mathcal{D}$. The key property connecting linear fractional transformations to classical plane geometry if that each such transformation L maps each circle to either a circle or an extended line (i.e., $\ell \cup \{\infty\}$, where ℓ is a line in $\mathbb{C}$). This is established in Proposition 2.9.2. We pay special attention to L_A for $A \in S\ell(2,\mathbb{R}) \cap SU(1,1)$, bringing in the hyperbolic functions $\sinh t$ and $\cosh t$, and examining analogues of Euler's formula. We bring these results to bear on a classical three circles problem: given disjoint circles C_1, C_2, C_3 in $\mathbb{C}$, no one of which separates the other two, find a fourth circle S that is tangent to each C_j.

Section 2.10 gives a brief taste of the algebraic approach to Euclidean geometry in higher dimensions. We endow $\mathbb{R}^n$ with a natural inner product, leading to a norm, with a triangle inequality established via Cauchy's inequality, as in §1.1. We bring in the more general class of n-dimensional inner product spaces, and their associated norms. We introduce the Gram-Schmidt process, leading to the construction of orthonormal bases for any such n-dimensional inner product space V, and also discuss the construction of an orthogonal projection P_W of V onto a linear subspace W.

We single out $\mathrm{O}(V)$ as the class of linear transformations on V preserving the inner product, and note that a general element $F \in \mathrm{Isom}(V)$ has the form

$$F(v) = Av + y, \quad A \in \mathrm{O}(V), y \in V. \tag{0.0.60}$$

We also bring in the determinant

$$\det : M(n,\mathbb{R}) \longrightarrow \mathbb{R}, \tag{0.0.61}$$

specified in Proposition 2.10.6 by three simple rules, and show that

$$\mathrm{O}(\mathbb{R}^n) = \mathrm{SO}(n) \cup \mathrm{O}^-(n), \tag{0.0.62}$$

where the two subclasses consist of $A \in \mathrm{O}(\mathbb{R}^n)$ such that $\det A = 1$ and $\det A = -1$, respectively. Specializing to $n = 3$, we define the cross product $u \times v$ of vectors $u, v \in \mathbb{R}^3$, and relate $\langle w, u \times v \rangle$ to a 3×3 determinant. This is a convenient tool to show how the cross product behaves under the application of an element of $\mathrm{SO}(3)$.

We then move to a discussion of volume, and define $\mathrm{Cont}^{\pm}(S)$ for a bounded set $S \subset \mathbb{R}^n$, using a tiling construction similar to that used for $n = 2$ in §2.6. If these two quantities are equal, $\mathrm{Vol}\, S$ is their common value, and we say S is contented. We discuss the fact that $\operatorname{Vol} S$ is preserved under maps in $\operatorname{Isom}(\mathbb{R}^n)$, and furthermore, for invertible $A \in M(n, \mathbb{R})$,

$$\operatorname{Vol} A(S) = |\det A| \operatorname{Vol} S. \tag{0.0.63}$$

We describe a further extension, to a change of variable formula for the integral of a function f under a C^1 map on n-dimensional space with a C^1 inverse, referring to [**5**] for a demonstration. We also present formulas for the calculation of the $(n-1)$-dimensional area of an $(n-1)$-dimensional surface $M \subset \mathbb{R}^n$. With these formulas in hand we turn to calculations of

$$V_n = \operatorname{Vol} B^n, \quad A_{n-1} = \operatorname{Area} S^{n-1}, \tag{0.0.64}$$

where B^n is the unit ball in $\mathbb{R}^n$ and S^{n-1} is the unit sphere in $\mathbb{R}^n$. We obtain

$$A_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad V_n = \frac{1}{n} A_{n-1}, \tag{0.0.65}$$

where $\Gamma(z)$ is a special function, known as Euler's gamma function. Examples include

$$A_2 = 4\pi,\ V_3 = \frac{4\pi}{3}, \quad A_3 = 2\pi^2,\ V_4 = \frac{\pi^2}{2}, \tag{0.0.66}$$

the first pair going back to Archimedes.

To end Section 2.10, we look back at our 2D approach, through algebraic operations on $\mathbb{C}$, and note that in n dimensions the product shifted, to a product of a matrix and a vector. We look at another classical algebra, the 4D algebra of quaternions, whose elements have the form

$$\xi = a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R}, \tag{0.0.67}$$

also written $\xi = a + u,\ u = bi + cj + dk \in \mathbb{R}^3$, where i, j, k denotes the standard basis of $\mathbb{R}^3$. We define a product on $\mathbb{H}$ so that

$$ij = k = -ji,\ jk = i = -kj,\ ki = j = -ik, \quad i^2 = j^2 = k^2 = -1. \tag{0.0.68}$$

In particular, if also $v \in \mathbb{R}^3$, then

$$uv = -u \cdot v + u \times v. \tag{0.0.69}$$

This product is not commutative, but it is associative. We have

$$\xi\overline{\xi} = |\xi|^2, \quad \overline{\xi} = a - bi - cj - dk. \tag{0.0.70}$$

Using the associative law, we have

$$|\xi\eta|^2 = |\xi|^2|\eta|^2, \tag{0.0.71}$$

if also $\eta \in \mathbb{H}$. This makes $\mathbb{H}$ a normed division ring. We deduce from (0.0.69) and (0.0.71) a short proof of the purely vectorial identity

$$|u \times v|^2 = |u|^2|v|^2 \sin^2\theta, \tag{0.0.72}$$

where θ is the angle between u and v in the plane in $\mathbb{R}^3$ spanned by these two vectors.

In another vein, we have that

$$\mathrm{Sp}(1) = \{\xi \in \mathbb{H} : |\xi| = 1\}, \tag{0.0.73}$$

which is the unit sphere in $\mathbb{H}$, i.e., S^3, is a multiplicative group, and the map λ defined by

$$\lambda(\xi, \eta)\zeta = \xi\zeta\overline{\eta}, \tag{0.0.74}$$

provides a two-to-one, surjective homomorphism

$$\lambda : \mathrm{Sp}(1) \times \mathrm{Sp}(1) \longrightarrow \mathrm{SO}(4). \tag{0.0.75}$$

Furthermore, we have the following variant of Euler's formula. Assume $u \in \mathbb{R}^3$, $|u| = 1$, and regard $u \in \mathbb{H}$. Then

$$e^{tu} = \cos t + (\sin t)u, \tag{0.0.76}$$

and such elements belong to $\mathrm{Sp}(1)$.

Chapter 1

Basic algebraic results on Euclidean geometry

Building on the results on the algebraic structure on $\mathbb{C}$ presented in the Introduction, we develop various geometrical consequences here. Section 1.1 discusses metric properties of $\mathbb{C}$, equipped with the distance function $d(z, w) = |z - w|$. A key tool is the identity

$$|z + w|^2 = |z|^2 + |w|^2 + 2\langle z, w\rangle, \tag{1.0.1}$$

bringing in the inner product $\langle z, w\rangle = \operatorname{Re} z\overline{w}$. As we will see, this contains the essence of the Pythagorean theorem.

Section 1.2 introduces the class $\operatorname{Isom}(\mathbb{C})$ of maps $F : \mathbb{C} \to \mathbb{C}$ that preserve distance, and gives this class an algebraic characterization.

Section 1.3 introduces the concepts of lines and line segments, perpendicular lines, and parallel lines, and establishes some basic properties.

In §1.4 we discuss rays and wedges, and define the angle measurement of a wedge as an element ω of the unit circle S^1. A related concept, measuring angles in "radians," requires more work, and will appear in §2.2.

In §1.5 we consider triangles, $\mathcal{T} = \triangle(A, B, C)$, with vertices $A, B, C \in \mathbb{C}$. We define angle measurements at each vertex, e.g., $\Omega_{\mathcal{T}}(A) \in S^1$. We also formally state the Pythagorean theorem, and derive it from (1.0.1).

In §1.6 we study congruence of triangles, which occurs if $\mathcal{T}' = F(\mathcal{T})$ for some $F \in \operatorname{Isom}(\mathbb{C})$. We establish various classical criteria for two triangles to be congruent, stated in terms involving side lengths and angle measurements.

Section 1.7 considers the family of isosceles triangles and the subfamily of equilateral triangles. We see equivalent characterizations in terms of sidelengths and in terms of angle measurements.

Section 1.8 deals with similarity, arising via maps $F : \mathbb{C} \to \mathbb{C}$ whose application scales distances by some fixed factor. We see that two triangles with identical angle measurements are similar. We display three similar triangles arising from a right triangle, divided by a perpendicular dropped from the right angle to the opposite side, and relate this to a proof of the Pythagorean theorem given in Book VI of Euclid.

Section 1.9 studies squares, rectangles, and more general figures, including parallelograms, as well as other quadrilaterals (4-gons), and further figures with n vertices (n-gons).

Section 1.10 introduces circles, of the form

$$S_R(a) = \{z \in \mathbb{C} : |z - a| = R\}, \tag{1.0.2}$$

and establishes some basic properties. As we discuss in the introduction to Chapter 2, further study of circles demands more sophisticated tools.

1.1. Metric properties of ℂ

A key to the metric structure of $\mathbb{C}$ is provided by the following computation of the absolute value of a sum. Given $z, w \in \mathbb{C}$, we have

$$\begin{aligned} |z+w|^2 &= (z+w)(\overline{z}+\overline{w}) \\ &= |z|^2+|w|^2+z\overline{w}+w\overline{z} \\ &= |z|^2+|w|^2+2\operatorname{Re} z\overline{w}. \end{aligned} \tag{1.1.1}$$

The last term has special importance. We set

$$(z,w)=z\overline{w}, \quad \langle z,w\rangle = \operatorname{Re}(z,w)=\operatorname{Re} z\overline{w}, \tag{1.1.2}$$

the latter quantity called the inner product of z and w. Then (1.1.1) says

$$|z+w|^2=|z|^2+|w|^2+2\langle z,w\rangle. \tag{1.1.3}$$

Note that $(z,w)=\overline{(w,z)}$ and $\langle z,w\rangle = \langle w,z\rangle$. Also, if $z=x+iy,\ w=u+iv,\ x,y,u,v\in\mathbb{R}$, then

$$z\overline{w}=xu+yv+i(yu-xv), \tag{1.1.4}$$

hence

$$\langle z,w\rangle = xu+yv. \tag{1.1.5}$$

Definition. Given $z,w\in\mathbb{C}$,

$$z\perp w \Longleftrightarrow \langle z,w\rangle=0. \tag{1.1.6}$$

When (1.1.6) holds, we say z is orthogonal (or perpendicular) to w. Note that $z\perp w \Leftrightarrow w\perp z$. The following consequence of (1.1.3) merits special attention.

Proposition 1.1.1. *Given $z,w\in\mathbb{C}$,*

$$|z+w|^2=|z|^2+|w|^2 \Longleftrightarrow z\perp w. \tag{1.1.7}$$

Our next result is a special case of the *triangle inequality.* This terminology will be explained a little later.

Proposition 1.1.2. *Given $z,w\in\mathbb{C}$,*

$$|z+w|\le |z|+|w|. \tag{1.1.8}$$

Proof. We compare the formula (1.1.3) for $|z+w|^2$ with the computation

$$(|z|+|w|)^2=|z|^2+|w|^2+2|z|\,|w|. \tag{1.1.9}$$

The fact that (1.1.8) follows from (1.1.3) and (1.1.9) is the content of the following result, a special case of *Cauchy's inequality.* □

Proposition 1.1.3. *Given $z,w\in\mathbb{C}$,*

$$\langle z,w\rangle \le |z|\,|w|. \tag{1.1.10}$$

Proof. Generally, for $\zeta \in \mathbb{C}$, the analogue of (0.0.9) implies $\operatorname{Re}\zeta \le |\zeta|$. Hence

$$\operatorname{Re} z\overline{w} \le |z\overline{w}| = |z|\,|w|, \tag{1.1.11}$$

the last identity by (0.0.12), plus the fact that $|w| = |\overline{w}|$. □

NOTE. Replacing z by $-z$ in (1.1.10) yields $-\langle z, w\rangle \le |z|, |w|$, hence

$$|\langle z, w\rangle| \le |z|\,|w|. \tag{1.1.12}$$

Having these results, we define the *distance* between points in $\mathbb{C}$:

Definition. Given $p, q \in \mathbb{C}$, the distance between p and q is

$$d(p, q) = |p - q|. \tag{1.1.13}$$

Note that $d(p, q) = d(q, p)$. Also

$$d(p + z, q + z) = d(p, q), \quad \forall\, z \in \mathbb{C}. \tag{1.1.14}$$

The following consequence of Proposition 1.1.2 is a more general form of the triangle inequality.

Proposition 1.1.4. *Given $p, q, r \in \mathbb{C}$,*

$$d(p, r) \le d(p, q) + d(q, r). \tag{1.1.15}$$

Proof. Set $z = p-q$, $w = q-r$, so $z+w = p-r$. Then (1.1.14) is equivalent to the statement that

$$|z + w| \le |z| + |w|, \tag{1.1.16}$$

which is the content of (1.1.8). □

Returning to the concept of orthogonality in (1.1.6), we have the following useful result.

Proposition 1.1.5. *Given nonzero $z, w \in \mathbb{C}$,*

$$z \perp w \Longleftrightarrow \frac{z}{w} \in i\mathbb{R}, \tag{1.1.17}$$

where

$$i\mathbb{R} = \{iy : y \in \mathbb{R}\}. \tag{1.1.18}$$

Proof. Indeed,

$$\operatorname{Re} z\overline{w} = \operatorname{Re} |w|^2 \frac{z}{w} = |w|^2 \operatorname{Re} \frac{z}{w}. \tag{1.1.19}$$

□

1.2. Rigid motions on $\mathbb{C}$

We say a map $F : \mathbb{C} \to \mathbb{C}$ is a *rigid motion* (or an *isometry*) provided

$$d(F(z), F(w)) = d(z, w), \quad \forall\, z, w \in \mathbb{C}. \tag{1.2.1}$$

We denote the set of all such maps by $\mathrm{Isom}(\mathbb{C})$. We have

$$F_1, F_2 \in \mathrm{Isom}(\mathbb{C}) \Longrightarrow F_1 \circ F_2 \in \mathrm{Isom}(\mathbb{C}), \tag{1.2.2}$$

where the composition $F_1 \circ F_2$ is defined by

$$F_1 \circ F_2(z) = F_1(F_2(z)). \tag{1.2.3}$$

Here are some examples of isometries:

$$\tau_a(z) = z + a, \quad a \in \mathbb{C}, \tag{1.2.4}$$

$$\rho_\omega(z) = \omega z, \quad \omega \in \mathbb{C},\ |\omega| = 1, \tag{1.2.5}$$

$$C(z) = \overline{z}. \tag{1.2.6}$$

Note that each of these maps is invertible, and the inverses are isometries:

$$\tau_a^{-1} = \tau_{-a}, \quad \rho_\omega^{-1} = \rho_{1/\omega}, \quad C^{-1} = C. \tag{1.2.7}$$

To formulate our first main result, we set

$$\begin{aligned} \mathrm{O}(\mathbb{C}) &= \{F \in \mathrm{Isom}(\mathbb{C}) : F(0) = 0\}, \\ \mathrm{O}_+(\mathbb{C}) &= \{\rho_\omega : \omega \in S^1\}, \\ \mathrm{O}_-(\mathbb{C}) &= \{C \circ \rho_\omega : \omega \in S^1\}, \end{aligned} \tag{1.2.8}$$

where

$$S^1 = \{\omega \in \mathbb{C} : |\omega| = 1\}. \tag{1.2.9}$$

Here is our first result.

Proposition 1.2.1. *We have*

$$\mathrm{O}(\mathbb{C}) = \mathrm{O}_+(\mathbb{C}) \cup \mathrm{O}_-(\mathbb{C}). \tag{1.2.10}$$

As a preliminary, we establish:

Lemma 1.2.2. *If $F \in \mathrm{O}(\mathbb{C}),\ z, w \in \mathbb{C}$,*

$$\langle F(z), F(w)\rangle = \langle z, w\rangle. \tag{1.2.11}$$

Proof. By (1.1.3),

$$\begin{aligned} -2\langle z, w\rangle &= |z - w|^2 - |z|^2 - |w|^2 \\ &= d(z, w)^2 - d(z, 0)^2 - d(w, 0)^2, \end{aligned} \tag{1.2.12}$$

and similarly

$$-2\langle F(z), F(w)\rangle = d(F(z), F(w))^2 - d(F(z), 0)^2 - d(F(w), 0)^2. \tag{1.2.13}$$

Since the right sides of (1.2.12) and (1.2.13) coincide, we have (1.2.11). □

Corollary 1.2.3. *If* $F \in \mathrm{O}(\mathbb{C}),\ z, w \in \mathbb{C}$,

$$z \perp w \Longrightarrow F(z) \perp F(w), \tag{1.2.14}$$

hence

$$F(iz) = \pm i F(z). \tag{1.2.15}$$

Proof. The result (1.2.14) is immediate from (1.2.11). Now, if $z \neq 0$, since $z \perp iz$, $F(z) \perp F(iz)$, so by Proposition 1.1.5,

$$\frac{F(iz)}{F(z)} = ir, \quad r \in \mathbb{R},\ |r| = \frac{|F(iz)|}{|F(z)|} = 1, \tag{1.2.16}$$

so $r = \pm 1$. □

Proof of Proposition 1.2.1. Given $F \in \mathrm{O}(\mathbb{C})$, suppose $F(1) = \omega$. Then $\omega \in S^1$, and

$$G = \rho_\omega^{-1} \circ F \in \mathrm{O}(\mathbb{C}), \quad G(1) = 1. \tag{1.2.17}$$

By Corollary 1.2.3, $G(i) = \pm i$. We claim that

$$\begin{aligned} G(i) = i &\Longrightarrow G(z) \equiv z, \\ G(i) = -i &\Longrightarrow G(z) \equiv \overline{z}. \end{aligned} \tag{1.2.18}$$

This result will yield (1.2.10). To start, suppose

$$G \in \mathrm{O}(\mathbb{C}), \quad G(1) = 1,\ G(i) = i. \tag{1.2.19}$$

Pick $z = x + iy$ and set $G(z) = u + iv,\ x, y, u, v \in \mathbb{R}$. Then

$$\begin{aligned} \langle z, 1 \rangle = x, &\qquad \langle z, i \rangle = y, \\ \langle G(z), 1 \rangle = u, &\quad \langle G(z), i \rangle = v, \end{aligned} \tag{1.2.20}$$

so Lemma 1.2.2 implies $u = x,\ v = y$, and we have the first half of (1.2.10). The proof of the second half is similar.

Corollary 1.2.4. *If* $F \in \mathrm{Isom}(\mathbb{C})$*, then either*

$$F(z) = \omega z + a, \quad \forall\, z \in \mathbb{C}, \tag{1.2.21}$$

for some $a \in \mathbb{C},\ \omega \in S^1$*, or*

$$F(z) = \omega \overline{z} + a, \quad \forall\, z \in \mathbb{C}, \tag{1.2.22}$$

for some such a, ω.

Corollary 1.2.5. *If* $F \in \mathrm{Isom}(\mathbb{C})$*, then* $F : \mathbb{C} \to \mathbb{C}$ *is one-to-one and onto, with inverse* $F^{-1} \in \mathrm{Isom}(\mathbb{C})$.

Remark. Given $F \in \mathrm{Isom}(\mathbb{C})$, we can write

$$F(z) = F_0(z) + a, \quad F_0 \in \mathrm{O}(\mathbb{C}),\ a \in \mathbb{C}, \tag{1.2.23}$$

and then, for $z, w \in \mathbb{C}$,

$$F(z) - F(w) = F_0(z) - F_0(w) = F_0(z - w). \tag{1.2.24}$$

We denote by $\mathrm{Isom}^+(\mathbb{C})$ the set of isometries of the form (1.2.21) and by $\mathrm{Isom}^-(\mathbb{C})$ the set of isometries of the form (1.2.22). Note that if $F^\pm, G^\pm \in \mathrm{Isom}^\pm(\mathbb{C})$, then

$$\begin{aligned} &F^+ \circ G^+, F^- \circ G^- \in \mathrm{Isom}^+(\mathbb{C}), \\ &F^+ \circ G^-, F^- \circ G^+ \in \mathrm{Isom}^-(\mathbb{C}), \end{aligned} \tag{1.2.25}$$

and

$$(F^\pm)^{-1} \in \mathrm{Isom}^\pm(\mathbb{C}). \tag{1.2.26}$$

We now introduce the concept of *congruence*. Given two subsets of $\mathbb{C}$, $S_1, S_2 \subset \mathbb{C}$, we say S_1 and S_2 are congruent provided there exists $F \in \mathrm{Isom}(\mathbb{C})$ such that

$$F(S_1) = S_2, \tag{1.2.27}$$

where $F(S) = \{F(z) : z \in S\}$. If (1.2.27) holds for some $F \in \mathrm{Isom}^+(\mathbb{C})$, we say F is a rotational congruence. If it holds with $F \in \mathrm{Isom}^-(\mathbb{C})$, we say F is a reflection congruence.

1.3. Lines and line segments

A line in $\mathbb{C}$ is a set of points of the form

$$L_{u,v} = \{ut + v : t \in \mathbb{R}\} = \{\ell_{u,v}(t) : t \in \mathbb{R}\}, \tag{1.3.1}$$

given $u, v \in \mathbb{C},\ u \neq 0$. Note that, by scaling $\mathbb{R}$, we have

$$L_{u,v} = L_{ru,v}, \quad \forall\, r \in \mathbb{R} \setminus 0. \tag{1.3.2}$$

In particular,

$$L_{u,v} = L_{\omega,v}, \quad \omega = \frac{u}{|u|} \in S^1, \tag{1.3.3}$$

where

$$S^1 = \{z \in \mathbb{C} : |z| = 1\}. \tag{1.3.4}$$

Note that if L is a line and $F \in \mathrm{Isom}(\mathbb{C})$, then $F(L)$ is also a line. Indeed,

$$\begin{aligned} F(L_{u,v}) = &\ L_{\omega u, \omega v + a}, \quad \text{for } F \text{ in (3.21)}, \\ &\ L_{\omega \overline{u}, \omega \overline{v} + a}, \quad \text{for } F \text{ in (3.22)}. \end{aligned} \tag{1.3.5}$$

In particular,

$$F(\mathbb{R}) = L_{\omega,a}, \quad \text{for } F \text{ in (3.21) or (3.22)}. \tag{1.3.6}$$

Equivalently, for such F,

$$\mathbb{R} = F^{-1}(L_{\omega,a}). \tag{1.3.7}$$

Here is a simple result.

Proposition 1.3.1. *Given two distinct points $p, q \in \mathbb{C}$, there is a line through p and q.*

Proof. Just take the line

$$\ell(t) = \ell_{q-p,p}(t) = p + (q-p)t. \tag{1.3.8}$$

□

There is also a uniqueness result, but first we treat the following basic proposition.

Proposition 1.3.2. *Consider the lines*

$$L_{\omega,a},\ L_{u,v}, \quad \omega, u \in S^1. \tag{1.3.9}$$

If $\omega \neq \pm u$, these lines intersect in exactly one point. If $\omega = \pm u$, these lines either coincide or have empty intersection.

Proof. Take $F \in \text{Isom}(\mathbb{C})$ such that (1.3.6) holds. Then

$$F^{-1}(L_{u,v}) = L_{\alpha,\beta}, \quad |\alpha| = 1, \ \beta \in \mathbb{C}, \tag{1.3.10}$$

and

$$\omega = \pm u \Longleftrightarrow \alpha = \pm 1. \tag{1.3.11}$$

Thus it suffices to consider the lines $\mathbb{R}$ and

$$L_{\alpha,\beta} = \{\alpha t + \beta : t \in \mathbb{R}\}. \tag{1.3.12}$$

Note that, for $t \in \mathbb{R}$,

$$\text{Im}(\alpha t + \beta) = at + b, \quad a = \text{Im}\,\alpha, \ b = \text{Im}\,\beta, \tag{1.3.13}$$

and, given $\alpha \in S^1$,

$$\alpha \neq \pm 1 \Longleftrightarrow a \neq 0. \tag{1.3.14}$$

If $a \neq 0$, then $at + b = 0$ if and only if $t = -b/a$, and the lines $L_{\alpha,\beta}$ and $\mathbb{R}$ have exactly one point of intersection. If $a = 0$, then $at + b \equiv b$, and this vanishes for all t if and only if $b = 0$, and vanishes for no t if and only if $b \neq 0$. This completes the proof. □

Remark. Given the lines in (1.3.9), if $\omega/u \in \mathbb{R}$ and the lines do not coincide, we say they are *parallel*.

Given Proposition 1.3.2, and given a line L through distinct points p and q, if L' is another line through p, and L' does not coincide with L, it follows that L' intersects L only at p. This gives the uniqueness of the line arising in Proposition 1.3.1. Here is another consequence of Proposition 1.3.2.

Proposition 1.3.3. *If L is a line in $\mathbb{C}$ and $p \in \mathbb{C}$, $p \notin L$, there is a unique line L' through p that does not intersect L, and it is parallel to L.*

Proof. We can write $L = L_{\omega,v}$, as in (1.3.3). Then $L' = L_{\omega,p}$ contains p and is parallel to L. If L'' is another line through p, it is expressible as $L'' = L_{\alpha,p}$, for some $\alpha \neq \pm\omega$. By Proposition 1.3.2, it must intersect L. □

Here is a useful related result.

Proposition 1.3.4. *Let L be a line in $\mathbb{C}$, $p \in \mathbb{C}$. Then there is a unique line $L_p^\perp$ through p that intersects L orthogonally.*

Proof. Write $L = L_{\omega,v}$, as in (1.3.3). Then $L_p^\perp$ is given by

$$L_p^\perp = L_{i\omega,p}. \tag{1.3.15}$$

□

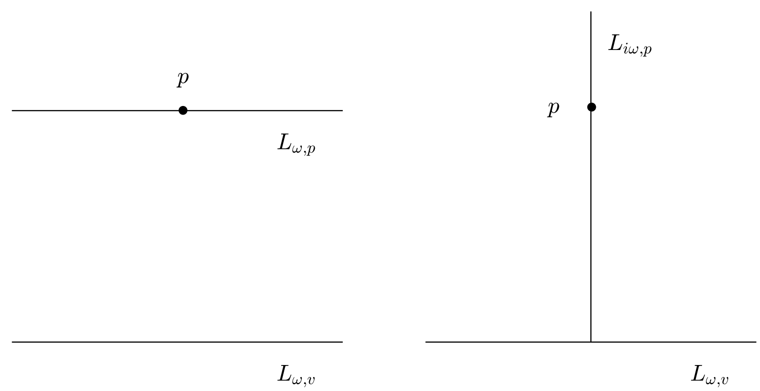

Figure 1.3.1. Drawing a parallel and dropping a perpendicular

Remark. One says we drop a perpendicular from p through L. See Figure 1.3.1 for an illustration of Propositions 1.3.3–1.3.4.

We now define the *reflection* across a line.

Proposition 1.3.5. *Let $L = L_{\omega,a}$ be a line. There is a unique*

$$R_L \in \mathrm{Isom}^-(\mathbb{C}), \text{ such that } R_L(z) = z, \ \forall\, z \in L. \tag{1.3.16}$$

Proof. By (1.3.6), we can take $F \in \mathrm{Isom}^+(\mathbb{C})$ such that $L_{\omega,a} = F(\mathbb{R})$, so it suffices to treat the case $L = \mathbb{R}$. In this case, the map satisfying (1.3.16) is simply

$$C(z) = \overline{z}. \tag{1.3.17}$$

Thus, for L itself, R_L in (1.3.16) is $R_L = F \circ C \circ F^{-1}$. □

Associated to the notion (1.3.1) of a line in $\mathbb{C}$ is the notion of a *line segment*, which has the form

$$L_{u,v,I} = \{ut + v : t \in I\}, \quad I = [a, b], \tag{1.3.18}$$

where $a, b \in I,\ a < b$. Such a line segment has endpoints

$$p = ua + v, \quad q = ub + v. \tag{1.3.19}$$

We say the line segment $L_{u,v,I}$ connects p and q.

We have the following analogue of Proposition 1.3.1

Proposition 1.3.6. *Given two distinct point $p, q \in \mathbb{C}$, there is a unique line segment connecting p and q.*

Proof. For existence, take the line

$$\ell(t) = p + (q - p)t, \quad t \in [0, 1]. \tag{1.3.20}$$

Uniqueness again follows from Proposition 1.3.2. □

1.4. Rays, wedges, and angles

A *ray* in $\mathbb{C}$ is a half-line. Given $p \in \mathbb{C}$, a ray emanating from p has the form

$$R_{p,u} = \{p + tu : t \in [0, \infty)\}, \tag{1.4.1}$$

with $u \in \mathbb{C},\ u \neq 0$. Scaling, we can take $u \in S^1$, defined by (1.3.4).

A *wedge* in $\mathbb{C}$ is an ordered pair of rays issuing from a common point p:

$$W_{p,u,v} = \{R_{p,u}, R_{p,v}\}, \quad u, v \in \mathbb{C} \setminus 0. \tag{1.4.2}$$

Again, we could assume $u, v \in S^1$. We say p is the vertex of $W_{p,u,v}$. We define the *angle measurement* of this wedge as

$$\Omega(W_{p,u,v}) = \Omega(u, v) = \frac{u}{v} \frac{|v|}{|u|} = \frac{u\overline{v}}{|u|\,|v|}. \tag{1.4.3}$$

In other words, $\Omega(W_{p,u,v}) = \omega$, where $\omega \in S^1$ satisfies

$$\frac{u}{|u|} = \omega \frac{v}{|v|}, \tag{1.4.4}$$

or equivalently,

$$\omega v = r u, \quad r \in (0, \infty). \tag{1.4.5}$$

Note that

$$\Omega(u, v) = \overline{\Omega(v, u)}. \tag{1.4.6}$$

It is frequently convenient to bring in the sets

$$S^1_+ = \{z \in S^1 : \operatorname{Im} z > 0\}, \quad S^1_- = \{z \in S^1 : \operatorname{Im} z < 0\}, \tag{1.4.7}$$

so we have a disjoint union

$$S^1 = S^1_+ \cup S^1_- \cup \{1, -1\}. \tag{1.4.8}$$

The following result will allow us to *bisect* an angle.

Proposition 1.4.1. *Given $\omega \in S^1_+$, there is a unique $\alpha \in S^1_+$ such that*

$$\alpha^2 = \omega. \tag{1.4.9}$$

Similarly, given $\omega \in S^1_-$, there is a unique $\alpha \in S^1_-$ satisfying (1.4.9).

Proof. Set $r = |1 + \omega|$, so

$$r^2 = (1 + \omega)(1 + \overline{\omega}) = 2 + \omega + \overline{\omega}. \tag{1.4.10}$$

We claim that

$$\alpha = \frac{1}{r}(1 + \omega) \tag{1.4.11}$$

works. Indeed,

$$\begin{aligned} r^2\alpha^2 &= (1+\omega)^2 = 1 + 2\omega + \omega^2, \\ r^2\omega &= (2+\omega+\overline{\omega})\omega = 2\omega + \omega^2 + 1. \end{aligned} \tag{1.4.12}$$

This covers existence. For uniqueness, if also $\beta \in S^1$,

$$\alpha^2 = \beta^2 = \omega \Rightarrow \Big(\frac{\alpha}{\beta}\Big)^2 = 1 \Rightarrow \frac{\alpha}{\beta} = \pm 1. \tag{1.4.13}$$

□

Remark. More generally, given $\omega \in S^1_+$ (resp., S^1_-) and $n \in \mathbb{N},\ n \geq 2$, there exists $\alpha \in S^1_+$ (resp., S^1_-) such that $\alpha^n = \omega$. For $n \geq 3$, there is a unique such α with maximal real part. This result is more subtle when $n \geq 3$ than for $n = 2$. We will see a neat derivation in §2.5.

Proposition 1.4.1 yields the following result on bisecting a wedge.

Proposition 1.4.2. *Let $W_{p,u,v}$ be a wedge. Assume*

$$|u| = |v|, \quad \Omega(W_{p,u,v}) = \omega \in S^1_+. \tag{1.4.14}$$

Take $\alpha \in S^1_+$ satisfying (1.4.9). *Then*

$$W_{p,u,\alpha v}, \quad W_{p,\alpha v,v} \tag{1.4.15}$$

are wedges, each with angle measurement α. We say the ray

$$R_{p,\alpha v} \tag{1.4.16}$$

bisects the wedge $W_{p,u,v}$. Furthermore,

$$R_{p,\alpha v} = R_{p,u+v}. \tag{1.4.17}$$

Remark. Similar results hold if, in (1.4.14), $\omega \in S^1_-$.

See Figure 1.4.1 for an illustration of angle bisection. We also note the following.

Corollary 1.4.3. *In the setting of Proposition 1.4.2, reflection R_L across the line $L = L_{u+v,p}$ gives $R_L \in \mathrm{Isom}^-(\mathbb{C})$ such that*

$$R_L(u) = v, \quad R_L(v) = u, \quad R_L(z) = z,\ \forall\, z \in L_{u+v,p}. \tag{1.4.18}$$

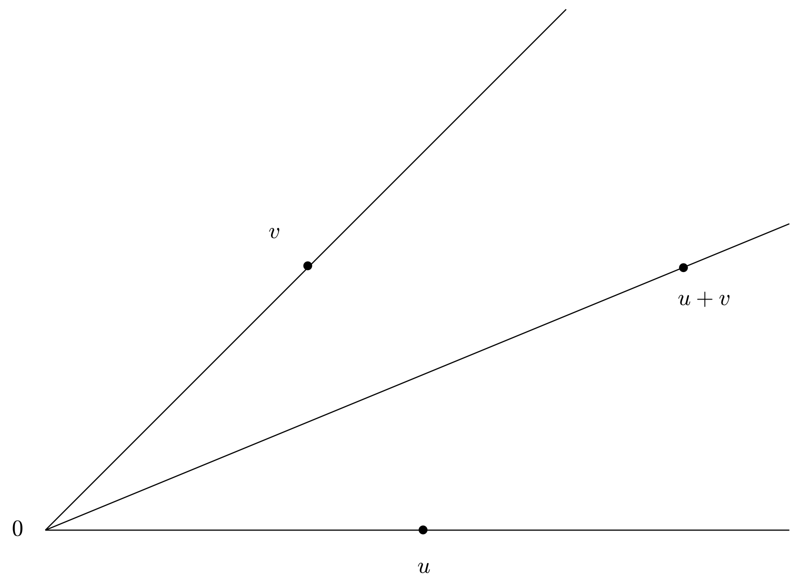

Figure 1.4.1. Bisecting a wedge

1.5. Triangles

Given three points $A, B, C \in \mathbb{C}$, not collinear, we associate a triangle $\mathcal{T} = \triangle(A, B, C)$. The points A, B, and C are the vertices of $\mathcal{T}$. The sides of $\mathcal{T}$ are the line segments, a, b, and c, opposite these vertices. See Figure 1.5.1.

With slight abuse of notation, we also set

$$c = B - A, \quad a = C - B, \quad b = A - C. \tag{1.5.1}$$

Note that

$$a + b + c = 0. \tag{1.5.2}$$

Recall from §1.4 that, for nonzero $z, w \in \mathbb{C}$, we define the angle measurement $\Omega(z, w) \in S^1$ to be the element $\omega \in S^1$ such that $\omega w = rz,\ r > 0$. Equivalently,

$$\Omega(z, w) = \mathcal{U}\Big(\frac{z}{w}\Big), \tag{1.5.3}$$

where $\mathcal{U} : \mathbb{C} \setminus 0 \to S^1$ is given by

$$\mathcal{U}(z) = \frac{z}{|z|}. \tag{1.5.4}$$

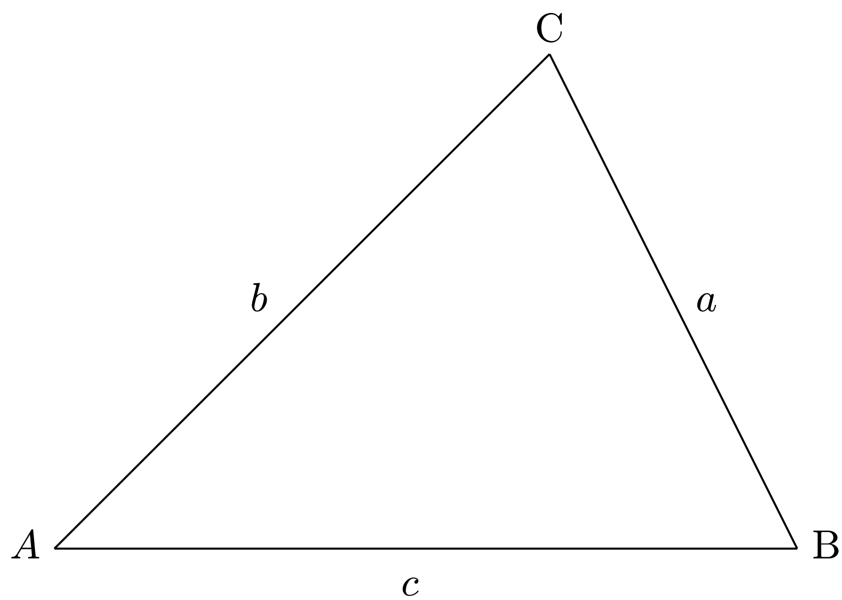

Figure 1.5.1. Triangle

We now define the angle measurements of the three angles of $\mathcal{T}$ as

$$\begin{aligned} \Omega_{\mathcal{T}}(A) &= \Omega(-b, c), \\ \Omega_{\mathcal{T}}(B) &= \Omega(-c, a), \\ \Omega_{\mathcal{T}}(C) &= \Omega(-a, b). \end{aligned} \tag{1.5.5}$$

Equivalently,

$$\Omega_{\mathcal{T}}(A) = \mathcal{U}\Big(-\frac{b}{c}\Big), \quad \Omega_{\mathcal{T}}(B) = \mathcal{U}\Big(-\frac{c}{a}\Big), \quad \Omega_{\mathcal{T}}(C) = \mathcal{U}\Big(-\frac{a}{b}\Big). \tag{1.5.6}$$

This leads immediately to the following result on the angles of a triangle.

Proposition 1.5.1. *If $\mathcal{T} = \triangle(A, B, C)$ is a triangle with vertices A, B, C, then*

$$\Omega_{\mathcal{T}}(A)\Omega_{\mathcal{T}}(B)\Omega_{\mathcal{T}}(C) = -1. \tag{1.5.7}$$

Proof. From (1.5.4) we have $\mathcal{U}(z)\mathcal{U}(w) = \mathcal{U}(zw)$. Hence (1.5.6) implies that the left side of (1.5.7) is equal to $\mathcal{U}(-1) = -1$. □

The following is an important complement to Proposition 1.5.1.

Proposition 1.5.2. *In the setting of Proposition 1.5.1, the angle measurements* (1.5.5) *are either all elements of* S^1_+ *or all elements of* S^1_-*, where*

$$S^1_+ = \{\omega \in S^1 : \operatorname{Im}\omega > 0\}, \quad S^1_- = \{\omega \in S^1 : \operatorname{Im}\omega < 0\}. \tag{1.5.8}$$

In light of the formula (1.5.6), this is a consequence of the following result.

Lemma 1.5.3. *If* $a, b, c \in \mathbb{C} \setminus 0$ *satisfy* $a + b + c = 0$ *and all the fractions*

$$\frac{b}{c}, \quad \frac{c}{a}, \quad \frac{a}{b} \tag{1.5.9}$$

belong to $\mathbb{C} \setminus \mathbb{R}$*, then their imaginary parts are either all positive or all negative.*

Proof. It suffices to compare the imaginary parts of

$$1 + \frac{a}{b}, \quad -1 - \frac{b}{a}, \quad \frac{a}{b}. \tag{1.5.10}$$

□

We have that all the angle measurements (1.5.5) belong to S^1_+ if and only if all the ratios in (1.5.9) have negative imaginary part.

Here is another useful formula for angle measurements. By (1.4.3), for nonzero $z, w \in \mathbb{C}$,

$$\langle z, w\rangle = |z|\,|w| \operatorname{Re}\Omega(z, w). \tag{1.5.11}$$

It follows that, in the setting of Proposition 1.5.1,

$$\begin{aligned} \operatorname{Re}\Omega_{\mathcal{T}}(A) &= -\frac{\langle b, c\rangle}{|b|\,|c|}, \\ \operatorname{Re}\Omega_{\mathcal{T}}(B) &= -\frac{\langle c, a\rangle}{|c|\,|a|}, \\ \operatorname{Re}\Omega_{\mathcal{T}}(C) &= -\frac{\langle a, b\rangle}{|a|\,|b|}. \end{aligned} \tag{1.5.12}$$

On the other hand, since $a + b + c = 0$, we have from (1.1.3) that

$$\begin{aligned} 2\langle a, b\rangle &= +|a|^2 + |b|^2 - |c|^2, \\ 2\langle b, c\rangle &= -|a|^2 + |b|^2 + |c|^2, \\ 2\langle c, a\rangle &= +|a|^2 - |b|^2 + |c|^2. \end{aligned} \tag{1.5.13}$$

Here is a classification of the angles of a triangle $\mathcal{T} = \mathcal{T}(A, B, C)$. We say that the angle of $\mathcal{T}$ at A is

$$\begin{aligned} \text{an acute angle} &\Longleftrightarrow \operatorname{Re}\Omega_{\mathcal{T}}(A) > 0, \\ \text{a right angle} &\Longleftrightarrow \operatorname{Re}\Omega_{\mathcal{T}}(A) = 0, \\ \text{an obtuse angle} &\Longleftrightarrow \operatorname{Re}\Omega_{\mathcal{T}}(A) < 0. \end{aligned} \tag{1.5.14}$$

A similar characterization applies to the angles at B and C. Note from (1.5.12) that

$$A \text{ is a right angle} \Longleftrightarrow b \perp c. \tag{1.5.15}$$

Here is a key property of triangles.

Proposition 1.5.4. *Each triangle has at least two acute angles.*

In light of Propositions 1.5.1–1.5.2, this is a consequence of the following result.

Lemma 1.5.5. *If* $\omega_1, \omega_2 \in S^1_+$ *and* $\operatorname{Re}\omega_1, \operatorname{Re}\omega_2 < 0$*, then*

$$\operatorname{Im}\omega_1\omega_2 < 0. \tag{1.5.16}$$

Hence if also $\omega_3 \in S^1$ *and* $\omega_1\omega_2\omega_3 = -1$*, then* $\operatorname{Im}\omega_3 < 0$.

Proof. Suppose that for $j = 1, 2$,

$$\omega_j = -c_j + is_j, \quad c_j, s_j > 0. \tag{1.5.17}$$

Then

$$\omega_2\omega_2 = c_1c_2 - s_1s_2 - i(c_1s_2 + s_1c_2), \tag{1.5.18}$$

and we have (1.5.16). Then if $\omega_3 \in S^1$,

$$\begin{aligned} (\omega_1\omega_2)\omega_3 = -1 &\Rightarrow \omega_3 = -\overline{(\omega_1\omega_2)} \\ &\Rightarrow \operatorname{Im}\omega_3 < 0, \end{aligned} \tag{1.5.19}$$

as asserted. □

If one of the angles of a triangle $\mathcal{T}$ is a right angle, we say $\mathcal{T}$ is a right triangle.

The following result is the Pythagorean theorem. It was essentially established in Proposition 1.1.1, but at that point we had not introduced the terminology for its classical formulation. Here it is.

Proposition 1.5.6. *Assume* $\mathcal{T} = \triangle(A, B, C)$ *is a right triangle, with right angle at* C*. Then*

$$|c|^2 = |a|^2 + |b|^2. \tag{1.5.20}$$

Proof. Parallel to (1.5.15), C is a right angle if and only if $a \perp b$. Then the identity (1.5.20) is immediate from the first identity in (1.5.13). □

1.6. Congruence of triangles

Let $\mathcal{T} = \triangle(A, B, C)$ be a triangle, with associated $a, b, c \in \mathbb{C}$ given by (1.5.1). Suppose $F \in \mathrm{Isom}(\mathbb{C})$, and consider

$$\mathcal{T}' = F(\mathcal{T}), \tag{1.6.1}$$

a triangle with vertices

$$A' = F(A), B' = F(B), C' = F(C), \tag{1.6.2}$$

and associated a', b', c', given by

$$c' = B' - A' = F_0(c), \text{ etc.}, \tag{1.6.3}$$

where, as in (1.2.23), we have

$$F(z) = F_0(z) + p, \quad F_0 \in \mathrm{O}(\mathbb{C}),\ p \in \mathbb{C}. \tag{1.6.4}$$

Since F preserves distances, we have, of course,

$$|a'| = |a|,\ |b'| = |b|,\ |c'| = |c|. \tag{1.6.5}$$

Let us compare angles, e.g.,

$$\begin{aligned} \Omega_{\mathcal{T}'}(A') &= \Omega(-b', c') = \mathcal{U}\Big(-\frac{b'}{c'}\Big), \\ \Omega_{\mathcal{T}}(A) &= \Omega(-b, c) = \mathcal{U}\Big(-\frac{b}{c}\Big). \end{aligned} \tag{1.6.6}$$

There are two cases to consider:

(i) $F \in \mathrm{Isom}^+(\mathbb{C}) \Rightarrow F_0(z) = \alpha z$,
(ii) $F \in \mathrm{Isom}^-(\mathbb{C}) \Rightarrow F_0(z) = \alpha \overline{z}$,
for some $\alpha \in S^1$.

$$\begin{aligned} &\text{Case i.} \qquad \frac{b'}{c'} = \frac{b}{c}, \quad \text{so } \Omega_{\mathcal{T}'}(A') = \Omega_{\mathcal{T}}(A), \\ &\text{Case ii.} \qquad \frac{b'}{c'} = \frac{\overline{b}}{\overline{c}}, \quad \text{so } \Omega_{\mathcal{T}'}(A') = \overline{\Omega_{\mathcal{T}}(A)}. \end{aligned}$$

Similar calculations hold for the other angles. We summarize.

Proposition 1.6.1. *If $F \in \mathrm{Isom}^+(\mathbb{C})$ and $F(\mathcal{T}) = \mathcal{T}'$, then*

$$\Omega_{\mathcal{T}'}(A') = \Omega_{\mathcal{T}}(A), \quad \Omega_{\mathcal{T}'}(B') = \Omega_{\mathcal{T}}(B), \quad \Omega_{\mathcal{T}'}(C') = \Omega_{\mathcal{T}}(C). \tag{1.6.7}$$

If $F \in \mathrm{Isom}^-(\mathbb{C})$, then

$$\Omega_{\mathcal{T}'}(A') = \overline{\Omega_{\mathcal{T}}(A)}, \quad \Omega_{\mathcal{T}'}(B') = \overline{\Omega_{\mathcal{T}}(B)}, \quad \Omega_{\mathcal{T}'}(C') = \overline{\Omega_{\mathcal{T}}(C)}. \tag{1.6.8}$$

We now consider the existence and uniqueness (up to a rigid motion) of a triangle, given partial geometric data. To begin, we take

$$\alpha \in S^1_+, \quad |b|,\ |c| \in \mathbb{R}^+, \tag{1.6.9}$$

and seek a triangle $\mathcal{T} = \triangle(A, B, C)$ such that

$$|B - A| = |c|, \quad |C - A| = |b|, \quad \Omega_{\mathcal{T}}(A) = \alpha. \tag{1.6.10}$$

Since these data are invariant under application of elements of $\text{Isom}^+(\mathbb{C})$, we can translate and rotate and see that it suffices to treat the case

$$A = 0, \quad B = |c|. \tag{1.6.11}$$

Then the conditions (1.6.10) impose the requirement on $C \in \mathbb{C}$ that

$$|C| = |b|, \quad \mathcal{U}\Big(\frac{C}{B}\Big) = \alpha, \tag{1.6.12}$$

hence

$$C = |b|\alpha. \tag{1.6.13}$$

This establishes the following.

Proposition 1.6.2. *Given the data* (1.6.9)*, there exists a triangle* $\mathcal{T} = \triangle(A, B, C)$ *satisfying* (1.6.10)*. If* $\mathcal{T}'$ *is another such triangle, then* $\mathcal{T}' = F(\mathcal{T})$ *for some* $F \in \text{Isom}^+(\mathbb{C})$*.*

REMARKS. The same analysis works if one takes $\alpha \in S^1_-$ instead of $\alpha \in S^1_+$. Also, if $\mathcal{T}'$ satisfies (1.6.10) except that α is replaced by $\overline{\alpha}$, then there exists $F \in \text{Isom}^-(\mathbb{C})$ such that $F(\mathcal{T}') = \mathcal{T}$.

In classical Euclidean geometry, the uniqueness result given above is stated as the use of

$$\text{side-angle-side} \tag{1.6.14}$$

to establish congruence of two triangles.

In the situation described above, we were given one angle of a triangle and (the lengths of) two adjacent sides. We move on to the scenario where we are given (the length of) one side, and (the angle measurements of) two adjacent angles. To be specific, we take

$$\alpha, \beta \in S^1_+, \quad |c| \in \mathbb{R}^+, \tag{1.6.15}$$

and seek a triangle $\mathcal{T} = \triangle(A, B, C)$ such that

$$|B - A| = |c|, \quad \Omega_{\mathcal{T}}(A) = \alpha, \quad \Omega_{\mathcal{T}}(B) = \beta. \tag{1.6.16}$$

In light of Propositions 1.5.1–1.5.2, we must impose the following constraint on α and β:

$$\exists\, \gamma \in S^1_+ \text{ such that } \alpha\beta\gamma = -1. \tag{1.6.17}$$

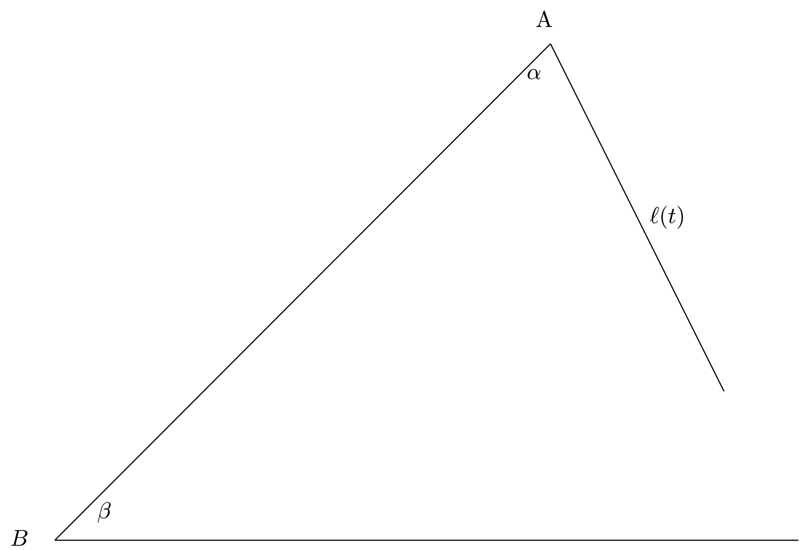

Figure 1.6.1. Illustration of angle-side-angle construction

Given (1.6.15), this holds if and only if

$$\alpha\beta \in S^1_+. \tag{1.6.18}$$

To proceed, this time it is convenient to take a translation and rotation such that

$$B = 0, \quad A = |A|\beta = |c|\beta. \tag{1.6.19}$$

See Figure 1.6.1. We will then find $C \in \mathbb{R}^+$ such that $\triangle(A, B, C)$ satisfies (1.6.16). To find C, we look at the ray,

$$\ell(t) = A + t\overline{\gamma}, \quad t \geq 0. \tag{1.6.20}$$

We have

$$\operatorname{Im} \ell(t) = \operatorname{Im} A - t \operatorname{Im} \gamma. \tag{1.6.21}$$

Since $\operatorname{Im} \gamma > 0$,

$$\operatorname{Im} \ell(t_0) = 0 \text{ for } t_0 = \frac{\operatorname{Im} A}{\operatorname{Im} \gamma} > 0. \tag{1.6.22}$$

This ray intersects $\mathbb{R}^+$ at

$$C = \ell(t_0) = \operatorname{Re}\ell(t_0) = \operatorname{Re} A + (\operatorname{Im} A)\frac{\operatorname{Re}\gamma}{\operatorname{Im}\gamma}. \tag{1.6.23}$$

Now the conditions $\alpha, \beta, \gamma \in S^1_+,\ \alpha\beta\gamma = -1$ imply that $\beta\gamma \in S^1_+$, i.e.,

$$\operatorname{Im}\beta\gamma = \operatorname{Re}\beta\ \operatorname{Im}\gamma + \operatorname{Im}\beta\ \operatorname{Re}\gamma > 0, \tag{1.6.24}$$

so

$$\begin{aligned} C &= |c|\Big(\operatorname{Re}\beta + \operatorname{Im}\beta\frac{\operatorname{Re}\gamma}{\operatorname{Im}\gamma}\Big) \\ &= \frac{|c|}{\operatorname{Im}\gamma}\ \operatorname{Im}(\beta\gamma) > 0. \end{aligned} \tag{1.6.25}$$

We have the following.

Proposition 1.6.3. *Given the data* (1.6.15), *and given the constraint* (1.6.17) *on α and β, there exists a triangle $\mathcal{T} = \triangle(A, B, C)$ satisfying* (1.6.16). *In such a case,*

$$\Omega_{\mathcal{T}}(C) = \gamma. \tag{1.6.26}$$

If $\mathcal{T}'$ is another such triangle, then $\mathcal{T}' = F(\mathcal{T})$ for some $F \in \operatorname{Isom}^+(\mathbb{C})$.

Remark. A similar analysis works if one takes $\alpha, \beta \in S^1_-$ instead of $\alpha.\beta \in S^1_+$. Then the constraint (1.6.15) is modified to say $\gamma \in S^1_-$. Also, if $\mathcal{T}'$ satisfies (1.6.15) except that α, β are replaced by $\overline{\alpha}, \overline{\beta}$, then there exists $F \in \operatorname{Isom}^-(\mathbb{C})$ such that $F(\mathcal{T}') = \mathcal{T}$.

In classical Euclidean geometry, the uniqueness result given above is stated as the use of

$$\text{angle-side-angle} \tag{1.6.27}$$

to establish congruence of the triangles.

We move to what is called the

$$\text{side-side-side} \tag{1.6.28}$$

argument to establish congruence of two triangles. Let $\mathcal{T} = \triangle(A, B, C)$ and $\mathcal{T}' = \triangle(A', B', C')$ be triangles, with associated sides a, b, c and a', b', c', respectively. This time, we assume

$$|a| = |a'|, \quad |b| = |b'|, \quad |c| = |c'|. \tag{1.6.29}$$

Recall from (1.5.12)–(1.5.13) that

$$2\langle b, c\rangle = |b|^2 + |c|^2 - |a|^2, \tag{1.6.30}$$

and hence

$$\operatorname{Re}\Omega_{\mathcal{T}}(A) = -\frac{1}{2|b|\,|c|}(|b|^2 + |c|^2 - |a|^2). \tag{1.6.31}$$

Similarly,

$$\operatorname{Re}\Omega_{\mathcal{T}'}(A') = -\frac{1}{2|b'|\,|c'|}(|b'|^2 + |c'|^2 - |a'|^2). \tag{1.6.32}$$

Hence, if (1.6.29) holds, we have

$$\alpha = \Omega_{\mathcal{T}}(A),\ \alpha' = \Omega_{\mathcal{T}'}(B') \Rightarrow \operatorname{Re}\alpha = \operatorname{Re}\alpha'. \tag{1.6.33}$$

Since $\alpha, \alpha' \in S^1$, it follows that

$$\text{either } \alpha' = \alpha \text{ or } \alpha' = \overline{\alpha}. \tag{1.6.34}$$

We pair this identity with the last two identities of (1.6.29) and see that the side-angle-side argument, establishing Proposition 1.6.2 and subsequent comments, applies. This yields the following.

Proposition 1.6.4. *Given triangles $\mathcal{T} = \triangle(A, B, C)$ and $\mathcal{T}' = \triangle(A', B', C')$ such that* (1.6.29) *holds, there exists $F \in \operatorname{Isom}(\mathbb{C})$ such that $\mathcal{T}' = F(\mathcal{T})$.*

Proposition 1.6.4 is the uniqueness result for triangles, given side lengths $|a|, |b|$, and $|c|$. We now investigate the existence question. In light of the triangle inequality, this requires the constraints

$$|a| \le |b| + |c|, \quad |b| \le |c| + |a|, \quad |c| \le |a| + |b|. \tag{1.6.35}$$

Furthermore, strict inequality must hold for the vertices not to be collinear. In the next result, we use the notation a, b, c in place of $|a|, |b|, |c|$, for convenience.

Proposition 1.6.5. *Let a, b, c be positive numbers, satisfying*

$$a < b + c, \quad b < c + a, \quad c < a + b. \tag{1.6.36}$$

Then there is a triangle $\mathcal{T} = \triangle(A, B, C)$ with such sidelengths.

Proof. We take $A = 0,\ B = c \in \mathbb{R}^+$, and need to find $C \in \mathbb{C}$ of the form

$$C = c + az, \quad z \in S^1 \setminus \mathbb{R}. \tag{1.6.37}$$

The triangle $\triangle(A, B, C)$ will have the desired sidelengths provided

$$|c + az| = b. \tag{1.6.38}$$

Since

$$|c + az|^2 = (c + az)(c + a\overline{z}) = c^2 + a^2 + ac(z + \overline{z}), \tag{1.6.39}$$

the requirement is

$$c^2 + a^2 - b^2 = -ac(z + \overline{z}) = -2acx, \tag{1.6.40}$$

for some $z = x + iy \in S^1,\ x, y \in \mathbb{R},\ y \neq 0$. We can find such z satisfying (1.6.40) if and only if

$$-x = \frac{c^2 + a^2 - b^2}{2ac} \in (-1, 1). \tag{1.6.41}$$

If this holds, then (1.6.40) holds, with

$$y = \pm\sqrt{1 - x^2}. \tag{1.6.42}$$

We claim that (1.6.36) implies (1.6.41).

Indeed, given $a, b, c > 0$, (1.6.36) is equivalent to

$$|c - a| < b < c + a, \tag{1.6.43}$$

and this holds

$$\begin{aligned} &\Leftrightarrow a^2 + c^2 - 2ac < b^2 < a^2 + c^2 + 2ac \\ &\Leftrightarrow a^2 + c^2 - b^2 < 2ac \text{ and} \\ &\quad\ b^2 - a^2 - c^2 < 2ac \\ &\Leftrightarrow \frac{|a^2 + c^2 - b^2|}{2ac} < 1. \end{aligned} \tag{1.6.44}$$

This completes the proof. □

1.7. Isosceles triangles and equilateral triangles

A triangle is said to be isosceles provided two of its sides have equal lengths. For example, take $\mathcal{T} = \triangle(A, B, C)$ and assume

$$|A - C| = |B - C|, \quad \text{i.e., , } |b| = |a|. \tag{1.7.1}$$

In such a case, let $\mathcal{T}' = \triangle(A', B', C')$, where

$$A' = B, \quad B' = A, \quad C' = C. \tag{1.7.2}$$

We see that

$$|a'| = |b'| = |a| = |b|, \tag{1.7.3}$$

and in particular $|a'| = |a|$, $|b'| = |b|$. Also

$$\Omega_{\mathcal{T}'}(C') = \overline{\Omega_{\mathcal{T}}(C)}. \tag{1.7.4}$$

The side-angle-side result, Proposition 1.6.2, or rather the comments following it, apply. Hence we have:

Proposition 1.7.1. *If $\mathcal{T}$ is isosceles, satisfying* (1.7.1)*, and if $\mathcal{T}'$ has vertices given by* (1.7.2)*, then*

$$\exists\, F \in \mathrm{Isom}^{-}(\mathbb{C}) \text{ such that } F(\mathcal{T}) = \mathcal{T}'. \tag{1.7.5}$$

In other words, $F \in \mathrm{Isom}^{-}(\mathbb{C})$ has the property

$$F(A) = B, \quad F(B) = A, \quad F(C) = C. \tag{1.7.6}$$

It follows from Proposition 1.6.1 that

$$\Omega_{\mathcal{T}'}(A') = \overline{\Omega_{\mathcal{T}}(A)}, \quad \Omega_{\mathcal{T}'}(B') = \overline{\Omega_{\mathcal{T}}(B)}, \tag{1.7.7}$$

and since $\Omega_{\mathcal{T}'}(A) = \overline{\Omega_{\mathcal{T}}(A)}$ and $\Omega_{\mathcal{T}'}(B) = \overline{\Omega_{\mathcal{T}}(B)}$, we have:

Corollary 1.7.2. *If $\mathcal{T}$ is isosceles and* (1.7.1) *holds, then*

$$\Omega_{\mathcal{T}}(A) = \Omega_{\mathcal{T}}(B). \tag{1.7.8}$$

We can obtain a converse result as follows. Suppose $\mathcal{T} = \triangle(A, B, C)$ is a triangle, and replace the hypothesis (1.7.1) by the hypothesis that (1.7.8) holds. Again define the triangle $\mathcal{T}' = \triangle(A', B', C')$ by (1.7.2). Since $|c| = |c'|$, we can use the angle-side-angle result, Proposition 1.6.3, or rather the comments following it, to obtain:

Proposition 1.7.3. *If $\mathcal{T} = \triangle(A, B, C)$ and the equi-angularity result* (1.7.8) *holds, then there exists $F \in \mathrm{Isom}^{-}(\mathbb{C})$ such that* (1.7.6) *holds. Hence $\mathcal{T}$ is isosceles.*

If $\mathcal{T} = \triangle(A,B,C)$ is a triangle all of whose sides have the same length, so

$$|A-C| = |B-C| = |A-B|, \text{ i.e., } |a| = |b| = |c|, \tag{1.7.9}$$

we say $\mathcal{T}$ is an equilateral triangle. Such $\mathcal{T}$ is isosceles in three ways, and repeated applications of Proposition 1.7.1 and Corollary 1.7.2 yield the following.

Proposition 1.7.4. *If $\mathcal{T} = \triangle(A,B,C)$ is an equilateral triangle, then there exist $F_j \in \mathrm{Isom}^-(\mathbb{C})$ such that*

$$\begin{aligned} F_1 &: \triangle(A,B,C) \longrightarrow \triangle(B,A,C), \\ F_2 &: \triangle(A,B,C) \longrightarrow \triangle(A,C,B), \\ F_3 &: \triangle(A,B,C) \longrightarrow \triangle(C,B,A). \end{aligned} \tag{1.7.10}$$

Also, there exists $G \in \mathrm{Isom}^+(\mathbb{C})$ such that

$$\begin{aligned} G &: \triangle(A,B,C) \longrightarrow \triangle(B,C,A), \\ G &: \triangle(B,C,A) \longrightarrow \triangle(C,A,B), \\ G &: \triangle(C,A,B) \longrightarrow \triangle(A,B,C). \end{aligned} \tag{1.7.11}$$

Furthermore,

$$\Omega_{\mathcal{T}}(A) = \Omega_{\mathcal{T}}(B) = \Omega_{\mathcal{T}}(C). \tag{1.7.12}$$

As with Proposition 1.7.3, we also have the converse result:

Proposition 1.7.5. *If $\mathcal{T} = \triangle(A,B,C)$ is a triangle and (1.7.12) holds, then $\mathcal{T}$ is an equilateral triangle.*

1.8. Similarity

Given $r \in (0,\infty)$, we have the scaling transformation

$$\delta_r : \mathbb{C} \longrightarrow \mathbb{C}, \quad \delta_r(z) = rz. \tag{1.8.1}$$

We say $G : \mathbb{C} \to \mathbb{C}$ is a similarity transformation if

$$G(z) = F \circ \delta_r(z), \quad F \in \mathrm{Isom}(\mathbb{C}), \tag{1.8.2}$$

and we write $G \in \mathrm{Sim}(\mathbb{C})$. If (1.8.2) holds with $F \in \mathrm{Isom}^{\pm}(\mathbb{C})$, we write $G \in \mathrm{Sim}^{\pm}(\mathbb{C})$. In light of Corollary 1.2.4, $G \in \mathrm{Sim}^{+}(\mathbb{C})$ if and only if

$$G(z) = uz + p, \quad \forall\, z \in \mathbb{C}, \tag{1.8.3}$$

for some $u, p \in \mathbb{C}$ ($u \neq 0$), and $G \in \mathrm{Sim}^{-}(\mathbb{C})$ if and only if

$$G(z) = u\overline{z} + p, \quad \forall\, z \in \mathbb{C}, \tag{1.8.4}$$

for some such u, p.

If $\mathcal{T} = \triangle(A,B,C)$ and $\mathcal{T}' = \triangle(A',B',C')$, and

$$G \in \mathrm{Sim}(\mathbb{C}), \quad G(\mathcal{T}) = \mathcal{T}', \tag{1.8.5}$$

we say $\mathcal{T}$ and $\mathcal{T}'$ are similar, and write $\mathcal{T} \sim \mathcal{T}'$ (or $\mathcal{T}' \sim \mathcal{T}$). Note that if a, b, c are as in (1.5.1) and a', b', c' are similarly defined, then

$$|a'| = r|a|, \quad |b'| = r|b|, \quad |c'| = r|c|, \tag{1.8.6}$$

if G has the form (1.8.2). Thus (1.8.6) holds with $r = |u|$ if G has the form (1.8.3) or (1.8.4).

It is elementary that if $\mathcal{T}' = \delta_r(\mathcal{T})$, then

$$\Omega_{\mathcal{T}'}(A') = \Omega_{\mathcal{T}}(A), \quad \Omega_{\mathcal{T}'}(B') = \Omega_{\mathcal{T}}(B), \quad \Omega_{\mathcal{T}'}(C') = \Omega_{\mathcal{T}}(C). \tag{1.8.7}$$

Hence we can apply Proposition 1.6.1 to conclude:

Proposition 1.8.1. *Suppose* (1.8.5) *holds. If* $G \in \mathrm{Sim}^{+}(\mathbb{C})$*, then* (1.8.7) *holds. If* $G \in \mathrm{Sim}^{-}(\mathbb{C})$*, then*

$$\Omega_{\mathcal{T}'}(A') = \overline{\Omega_{\mathcal{T}}(A)}, \quad \Omega_{\mathcal{T}'}(B') = \overline{\Omega_{\mathcal{T}}(B)}, \quad \Omega_{\mathcal{T}'}(C') = \overline{\Omega_{\mathcal{T}}(C)}. \tag{1.8.8}$$

Here is an important converse.

Proposition 1.8.2. *Let* $\mathcal{T} = \triangle(A,B,C)$ *and* $\mathcal{T}' = \triangle(A',B',C')$*. If* (1.8.7) *holds, then* $\mathcal{T}' = G(\mathcal{T})$ *for some* $G \in \mathrm{Sim}^{+}(\mathbb{C})$*. If* (1.8.8) *holds, then* $\mathcal{T}' = G(\mathcal{T})$ *for some* $G \in \mathrm{Sim}^{-}(\mathbb{C})$*. In either case,* (1.8.6) *holds for some* $r \in \mathbb{R}^{+}$*.*

Proof. Say $|a'| = s|a|$, $s \in \mathbb{R}^{+}$. Then consider $\mathcal{T}_1 = \delta_s \mathcal{T}$. If (1.8.7) or (1.8.8) holds, it also holds with $\mathcal{T}$ replaced by $\mathcal{T}_1$. But then the angle-side-angle result, Proposition 1.6.3 (and subsequent comments) give $F \in \mathrm{Isom}^{\pm}(\mathbb{C})$ such that $\mathcal{T}' = F(\mathcal{T}_1)$. This establishes the desired result. ☐

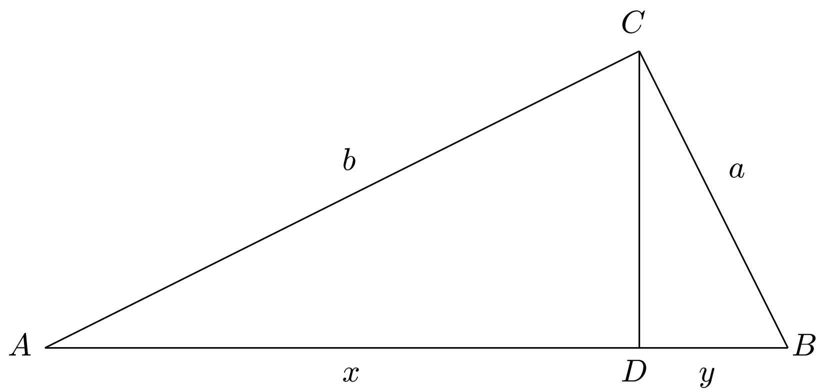

Figure 1.8.1. Three similar triangles

Remark. In light of Proposition 1.5.1, we see that for the three identities in (1.8.7) to hold, it suffices for any two of them to hold. Ditto for (1.8.8).

We now discuss an interesting collection of three similar triangles, depicted in Figure 1.8.1. We have a right triangle $\mathcal{T} = \triangle(A, B, C)$, with right angle at C. We drop a perpendicular from C to the line segment from A to B, intersecting it at D (cf. Proposition 1.3.4). Here, to simplify notation, we let a, b, c stand for $|a|, |b|, |c|$. The point D divides the interval from A to B into two subintervals, of lengths x and y, satisfying

$$x + y = c. \tag{1.8.9}$$

The other triangles are

$$\mathcal{T}_2 = \triangle(A, C, D), \quad \mathcal{T}_3 = \triangle(C, B, D). \tag{1.8.10}$$

Note that $\mathcal{T}$ and $\mathcal{T}_2$ have the vertex A in common, and $\mathcal{T}$ and $\mathcal{T}_3$ have the vertex B in common. We have

$$\Omega_{\mathcal{T}}(A) = \overline{\Omega_{\mathcal{T}_2}(A)}, \quad \Omega_{\mathcal{T}}(B) = \overline{\Omega_{\mathcal{T}_3}(B)}. \tag{1.8.11}$$

All these triangles have right angles, $\mathcal{T}$ at C, $\mathcal{T}_1$ and $\mathcal{T}_2$ at D, Therefore we can invoke Proposition 1.8.2 (keeping in mind its subsequent remark). We have

$$\mathcal{T} \sim \mathcal{T}_2 \sim \mathcal{T}_3. \tag{1.8.12}$$

We conclude from the analogues of (1.8.6) that

$$\mathcal{T} \sim \mathcal{T}_2 \Rightarrow \frac{c}{b} = \frac{b}{x}, \quad \mathcal{T} \sim \mathcal{T}_3 \Rightarrow \frac{c}{a} = \frac{a}{y}. \tag{1.8.13}$$

These identities imply

$$b^2 = cx, \quad a^2 = cy, \tag{1.8.14}$$

and adding these, keeping in mind that $x + y = c$, yields

$$a^2 + b^2 = c^2, \tag{1.8.15}$$

The identity (1.8.15) states the Pythagorean theorem for the right triangle $\mathcal{T}$. Now the Pythagorean theorem has been baked into our infrastructure since Proposition 1.1.1, so we cannot really regard this as a second proof of this theorem. On the other hand, the argument presented above does give the gist of the second proof of the Pythagorean theorem, in Book VI of Euclid. (Euclid's first proof, in Book I, involves area comparisons. For more on this, see the discussion following Proposition 2.6.5.) Anyway, the argument above might provide insight into how people derived the Pythagorean theorem in the old days. Of course, it does not completely show how they thought. Keep in mind that Euclid and his colleagues did not have algebra.

1.9. Squares, rectangles, and other quadrilaterals

Here we look at various classes of four-sided figures, known as quadrilaterals, or 4-gons. First, let us make some general comments about n-gons, for $n \geq 3$. An n-gon $\mathcal{P} = G_n(A_0, A_1, \dots, A_{n-1})$ is specified by n distinct points $A_j \in \mathbb{C},\ 0 \leq j \leq n-1$, which are joined by the sides L_j, line segments from A_j to A_{j+1}, for $0 \leq j \leq n-1$. We make the convention here and below that

$$A_n = A_0, \quad A_{-1} = A_{n-1}. \tag{1.9.1}$$

We impose the requirement that

(1.9.2) these line segments L_j do not intersect except at their endpoints.

We define the angle measure of the angle at A_j in $\mathcal{P}$ as

$$\Omega_{\mathcal{P}}(A_j) = \Omega_{\mathcal{T}_j}(A_j), \quad \mathcal{T}_j = \triangle(A_{j-1}, A_j, A_{j+1}), \tag{1.9.3}$$

using the convention (1.9.1) for $j = 0$ and $j = n-1$. A straightforward extension of Proposition 1.5.1 is that, if $\mathcal{P} = G_n(A_0, \dots, A_{n-1})$,

$$\Omega_{\mathcal{P}}(A_0)\Omega_{\mathcal{P}}(A_1)\cdots\Omega_{\mathcal{P}}(A_{n-1}) = (-1)^n. \tag{1.9.4}$$

Specializing to $n = 4$, we first describe rectangles. A rectangle $R = G_4(A, B, C, D)$ is a 4-gon all of whose angles are right angles. See Figure 1.9.1. To see better what R looks like, we can find $F \in \text{Isom}(\mathbb{C})$ such that

$$F(R) = R' = G_4(A', B', C', D') = G_4(0, a, C', ib), \quad a, b > 0, \tag{1.9.5}$$

whose sides include L_0', from 0 to a, and L_3', from ib to 0. The line segment L_1' connects a to C' and L_2' connects C' to ib. For the angles at $B' = a$ and at $D' = ib$ to be right angles, we require

$$L_1' \perp L_0' \text{ and } L_2' \perp L_3', \tag{1.9.6}$$

i.e.,

$$C' - a \in i\mathbb{R} \text{ and } C' - ib \in \mathbb{R}. \tag{1.9.7}$$

This forces

$$\operatorname{Re} C' = a,\ \operatorname{Im} C' = b, \quad \text{hence } C' = a + ib. \tag{1.9.8}$$

We have specified R'. Note that the calculations (1.9.6)–(1.9.8) only require the angles at A', B', and D' to be right angles. This forces C' to be a right angle. This is consistent with the case $n = 4$ of (1.9.4).

One consequence of (1.9.8) is the following.

Proposition 1.9.1. *Opposite sides of a rectangle $R = G_4(A, B, C, D)$ have equal lengths, i.e.,*

$$|A - B| = |D - C| \text{ and } |A - D| = |B - C|. \tag{1.9.9}$$

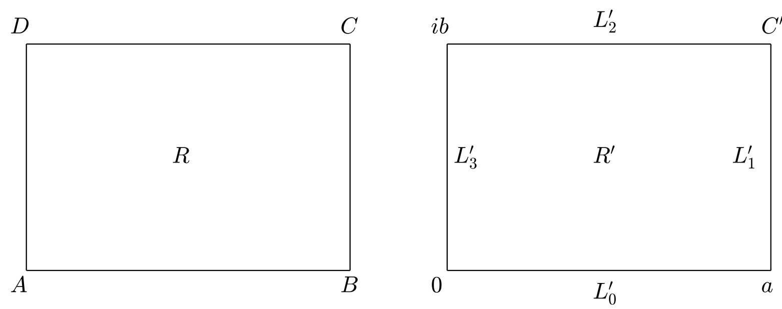

Figure 1.9.1. Rectangles $R = G_4(A, B, C, D)$, and R'

With a and b as in (1.9.5), we say the rectangle R has sidelengths a and b.

A square is a rectangle S whose sidelengths are all equal, say to c. In such a case, calculations parallel to (1.9.5)–(1.9.8) yield a transformation $F \in \mathrm{Isom}(\mathbb{C})$ such that

$$F(S) = S' = G_4(0, c, c + ic, ic). \tag{1.9.10}$$

Translating by $-c(1 + i)/2$ yields a square S'', with vertices

$$-\frac{c}{2}(1 + i), \quad \frac{c}{2}(1 - i), \quad \frac{c}{2}(1 + i), \quad \frac{c}{2}(-1 + i), \tag{1.9.11}$$

which is clearly invariant under $\rho_i(z) = iz$, yielding for the original square $S = G_4(A, B, C, D)$ a rotational symmetry about its center $(A + B + C + D)/4 = p$, i.e., S is invariant under $F_p(z) = i(z - p) + p$.

We next introduce a family of 4-gons more general than the class of rectangles. We say $P = G_4(A, B, C, D)$ is a parallelogram provided its opposite sides are parallel, i.e.,

$$D - A \parallel C - B, \quad B - A \parallel C - D. \tag{1.9.12}$$

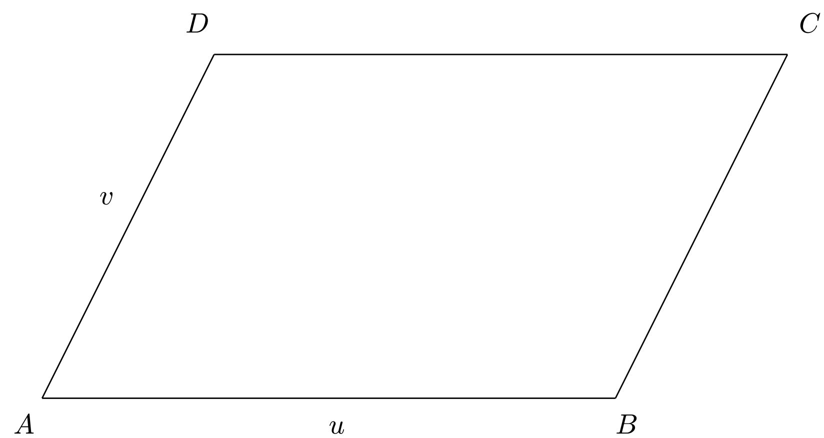

Figure 1.9.2. Parallelogram

See Figure 1.9.2. In such a case, through D there is a unique line L_2, parallel to the line through A and B, and through B there is a unique line L_1, parallel to the line through A and D. Equivalently, if we set

$$u = B - A, \quad v = D - A, \tag{1.9.13}$$

then

$$L_2 = \{D + tu : t \in \mathbb{R}\}, \quad L_1 = \{B + sv : s \in \mathbb{R}\}. \tag{1.9.14}$$

We assume A, B, and D are not collinear, hence

$$\frac{u}{v} \notin \mathbb{R}. \tag{1.9.15}$$

Lemma 1.9.2. *In the setting described above, the lines L_1 and L_2 meet at the point $A + u + v$.*

Proof. Where L_1 and L_2 meet, we have

$$A + v + tu = A + u + sv, \tag{1.9.16}$$

hence

$$(t-1)u = (s-1)v. \tag{1.9.17}$$

Now (1.9.15) implies that $t = 1$ and $s = 1$, as asserted. □

In summary, we have

Proposition 1.9.3. *If* $P = G_4(A, B, C, D)$ *is a parallelogram, with* u, v *given by* (1.9.13)*, then*

$$C = A + u + v. \tag{1.9.18}$$

As a consequence, opposite sides of P *have the same length.*

Corollary 1.9.4. *In the setting of Proposition 1.9.3, the parallelogram* P *is invariant under the map* $F \in \mathrm{Isom}(\mathbb{C})$ *given by*

$$F(z) = -(z-p) + p, \quad p = \frac{1}{4}(A + B + C + D). \tag{1.9.19}$$

Consequently, opposite vertices of P *have the same angular measurements:*

$$\Omega_P(A) = \Omega_P(C), \quad \Omega_P(B) = \Omega_P(D). \tag{1.9.20}$$

Remark. If $|u| = |v|$ in (1.9.13), then all four sides of P have equal length. In such a case, one says P is a rhombus.

There is a result, called the parallelogram law, equating the sum of the squares of the sidelengths of a parallelogram P and the sum of the squares of the lengths of the diagonals of P, i.e., the line segments from A to C and from B to D. As for the latter, note that

$$|A - C| = |u + v|, \quad |B - D| = |u - v|. \tag{1.9.21}$$

To derive the asserted identity, start with

$$\begin{aligned} |u+v|^2 &= |u|^2 + |v|^2 + 2\langle u, v\rangle, \\ |u-v|^2 &= |u|^2 + |v|^2 - 2\langle u, v\rangle, \end{aligned} \tag{1.9.22}$$

and add, to get

$$|u+v|^2 + |u-v|^2 = 2|u|^2 + 2|v|^2. \tag{1.9.23}$$

This is the parallelogram law.

1.10. Circles

A circle in $\mathbb{C}$, centered at a and of radius r, is a set of the form

$$S_r(a) = \{z \in \mathbb{C} : |z - a| = r\}. \tag{1.10.1}$$

The unit circle S^1 we have encountered before is $S^1 = S_1(0)$. Note that

$$z \in \mathbb{C},\ z \neq 0 \Rightarrow \omega = \frac{z}{|z|} \in S^1 \Rightarrow a + r\omega \in S_r(a). \tag{1.10.2}$$

The circle $S_r(a)$ is the boundary of the *disk*

$$D_r(a) = \{z \in \mathbb{C} : |z - a| \leq r\}. \tag{1.10.3}$$

It is of great interest to know when two circles, say $S_R(a)$ and $S_r(b)$, intersect. See Figure 1.10.1 for an illustration. We have the following result.

Proposition 1.10.1. *Given $a \neq b \in \mathbb{C},\ R, r > 0$, the following are equivalent.*

(a) $S_R(a) \cap S_r(b) \neq \emptyset$,

(b) $|R - r| \leq |a - b| \leq R + r$.

If one has strict inequality in (b), the intersection in (a) has two points. If one of the inequalities in (b) is not strict, the intersection in (a) has one point.

Proof. The demonstration is essentially a repeat of the proof of Proposition 1.6.5, with different notation (cf. (1.6.43)). If we bring in the isometry

$$F(z) = \overline{\omega}(z - a), \quad \omega = \frac{b - a}{|b - a|}, \tag{1.10.4}$$

the image circles are $S_R(0)$ and $S_r(d)$, $d = |b - a|$, and, provided $|R - r| \leq d \leq R + r$, they intersect at points $x \pm iy$, given by computations parallel to those done in the proof of Proposition 1.6.5. Then $S_R(a)$ and $S_r(b)$ intersect at $F^{-1}(x \pm iy)$. If one of the inequalities in (b) is not strict, $y = 0$. □

REMARK. Constructions involving using points where two circles (drawn with a compass) intersect play a fundamental role in classical Euclidean geometry, though this aspect is not particularly emphasized here, except in §2.8 on Euclidean numbers.

The Euclideans do not seem to have proved a result like Proposition 1.10.1. As we see from the proof of Proposition 1.6.5, a key ingredient in the proof is the ability to take the square root of a positive number and

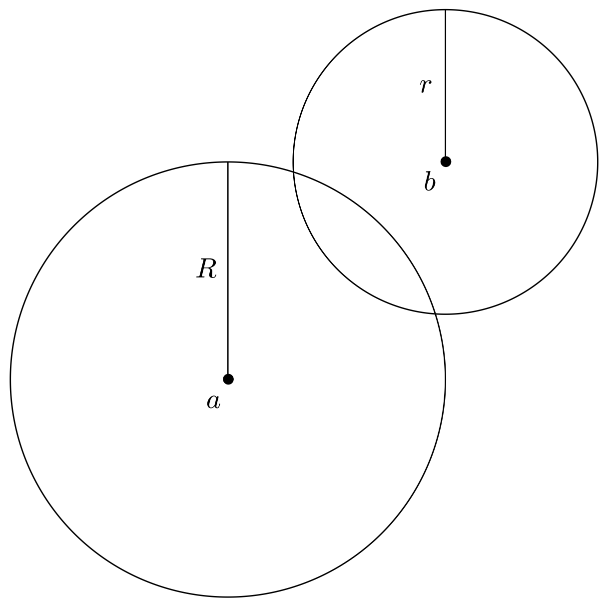

Figure 1.10.1. Intersecting circles

know that it defines a positive real number. Such a foundational issue is treated in Chapter 1 of [**4**], in the development of the real number system.

There is a parallel result for the intersection of a circle $S_R(a)$ and a line L. Given such objects, let $L_a^\perp$ denote the line through a that intersects L orthogonally, given by Proposition 1.3.4. Say L and $L_a^\perp$ intersect at b. Then

$$\begin{aligned} |a-b| < R &\Rightarrow S_R(a) \cap L \text{ has two points,} \\ |a-b| = R &\Rightarrow S_R(a) \cap L \text{ has one point,} \\ |a-b| > R &\Rightarrow S_R(a) \cap L \text{ is empty.} \end{aligned} \tag{1.10.5}$$

We leave the demonstration to the reader.

We next describe an interesting family of right triangles, inscribed in a circle, illustrated in Figure 1.10.2.

Proposition 1.10.2. *Pick* $v \in S^1$, $v \neq \pm 1$. *Then* $\mathcal{T} = \triangle(-1, v, 1)$ *is a right triangle, with right angle at* v.

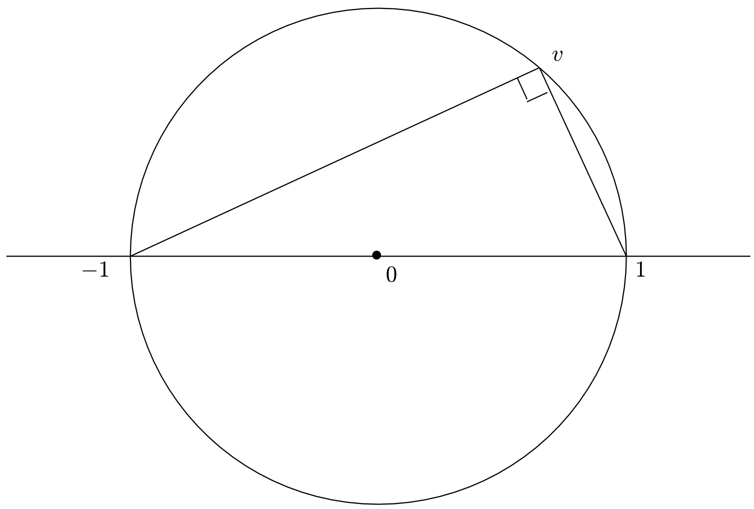

Figure 1.10.2. Right triangle in a circle

Proof. It is equivalent to show that $v+1 \perp v-1$, if $|v|=1$. Indeed,

$$\begin{aligned} \langle v+1, v-1\rangle &= |v|^2 + \langle 1, v\rangle - \langle v, 1\rangle - 1 \\ &= |v|^2 - 1 = 0. \end{aligned} \tag{1.10.6}$$

□

The computation (1.10.6) shows that if $v \in S^1, v \neq 1$, then $(v+1)/(v-1) \in i\mathbb{R}$ (cf. Proposition 1.1.5). Equivalently, if we set

$$\psi(v) = \frac{v+1}{v-1}\, i, \tag{1.10.7}$$

then

$$\psi : S^1 \setminus \{1\} \longrightarrow \mathbb{R}. \tag{1.10.8}$$

For another check of this, note that if $|v|=1,\ v \neq 1$,

$$\frac{v+1}{v-1} = \frac{(v+1)(\overline{v}-1)}{(v-1)(\overline{v}-1)} = -2i\,\frac{\operatorname{Im} v}{|v-1|^2}. \tag{1.10.9}$$

Furthermore, if we set

$$\varphi(x) = \frac{x+i}{x-i}, \tag{1.10.10}$$

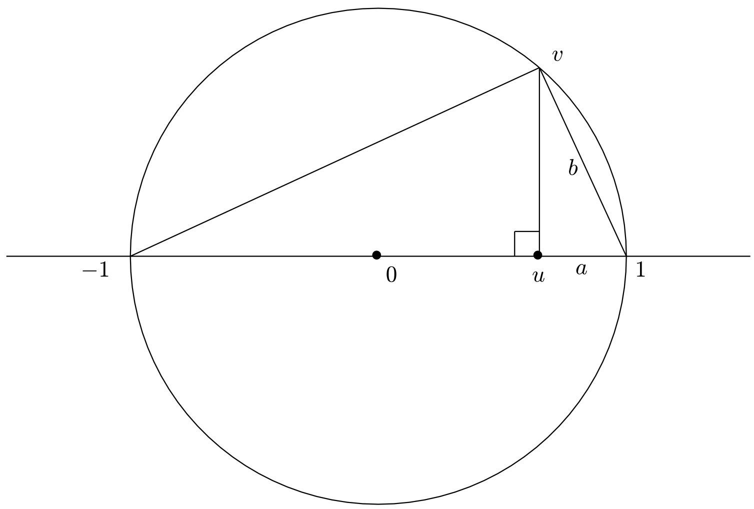

Figure 1.10.3. Geometric construction of $b = \sqrt{2a}$

then

$$\varphi : \mathbb{R} \longrightarrow S^1 \setminus \{1\} \text{ is one-to-one and onto,} \tag{1.10.11}$$

with inverse ψ in (1.10.7)–(1.10.8).

We next describe a geometrical approach to the construction of the square root of a given positive number, illustrated in Figure 1.10.3. To start, assume $a \in (0,1)$, and set $u = 1 - a \in (0,1)$. Let v denote the point in S^1_+ where the line orthogonal to $\mathbb{R}$ through u intersects S^1_+. Explicitly,

$$v = u + iy, \ y > 0, \quad u^2 + y^2 = 1, \tag{1.10.12}$$

so

$$y = \sqrt{1 - u^2}. \tag{1.10.13}$$

Proposition 1.10.3. *In the setting described above,*

$$b = |v - 1| \Longrightarrow b = \sqrt{2a}. \tag{1.10.14}$$

Proof. Since $\triangle(v,u,1)$ is a right triangle with right angle at u, we have

$$\begin{aligned} b^2 &= y^2 + a^2 \\ &= 1 - u^2 + a^2 \\ &= 1 - (1-a)^2 + a^2 \\ &= 2a. \end{aligned} \tag{1.10.15}$$

Second proof. The triangles $\triangle(v,u,1)$ and $\triangle(-1,v,1)$ are both right triangles (thanks to Proposition 1.10.2), and they have a common angle at 1. Hence they are similar. Similarity implies

$$\frac{b}{2} = \frac{a}{b}, \tag{1.10.16}$$

which in turn yields the conclusion in (1.10.14). □

More generally, given $r > 0$, pick $n \in \mathbb{N}$ such that

$$a = \frac{r}{2n^2} \in (0,1). \tag{1.10.17}$$

Then the construction above yields $b = \sqrt{2a}$, giving $b = \sqrt{r}/n$, or

$$\sqrt{r} = nb. \tag{1.10.18}$$

We return to the functions ψ and φ, arising in (1.10.7)–(1.10.11), and derive a result with a number-theoretic flavor. Namely, let $\mathbb{Q} \subset \mathbb{R}$ denote the field of rational numbers, and set

$$\begin{aligned} \mathbb{Q}[i] &= \{x + iy : x, y \in \mathbb{Q}\}, \\ S^1_{\mathbb{Q}} &= S^1 \cap \mathbb{Q}[i] = \{x + iy : x, y \in \mathbb{Q}, x^2 + y^2 = 1\}. \end{aligned} \tag{1.10.19}$$

It follows from the formulas (1.10.7) and (1.10.10), and the result (1.10.11) that

$$\varphi : \mathbb{Q} \longrightarrow S^1_{\mathbb{Q}} \setminus \{1\} \text{ is one-to-one and onto,} \tag{1.10.20}$$

with inverse given by

$$\psi : S^1_{\mathbb{Q}} \setminus \{1\} \longrightarrow \mathbb{Q}. \tag{1.10.21}$$

Note also that

$$x > 0 \Longleftrightarrow \varphi(x) \in S^1_+. \tag{1.10.22}$$

These results apply to the study of Pythagorean triples, which consist of triples (j,k,ℓ) of positive integers satisfying

$$j^2 + k^2 = \ell^2. \tag{1.10.23}$$

We have the following.

Proposition 1.10.4. *For each Pythagorean triple (j, k, ℓ), there is a unique $m/n \in \mathbb{Q}^+$ such that*

$$\varphi\Big(\frac{m}{n}\Big) = \frac{j}{\ell} + \frac{k}{\ell}i. \tag{1.10.24}$$

Conversely, if $m/n \in \mathbb{Q}^+$, then (1.10.24) yields a Pythagorean triple (up to a sign on j).

Since

$$\varphi(x) = \frac{x^2-1}{x^2+1} + \frac{2x}{x^2+1}i, \tag{1.10.25}$$

we have

$$\varphi\Big(\frac{m}{n}\Big) = \frac{m^2-n^2}{m^2+n^2} + \frac{2mn}{m^2+n^2}i, \tag{1.10.26}$$

yielding

$$j = m^2 - n^2, \quad k = 2mn, \quad \ell = m^2 + n^2. \tag{1.10.27}$$

Circles and triangles – inscribed and circumscribed

Let $A, B, C \in \mathbb{C}$ be non-collinear points, hence defining a triangle $\mathcal{T} = \triangle(A, B, C)$. As we will see, there is a unique circle $\mathcal{C}$ through these three points, which is hence said to circumscribe $\mathcal{T}$; alternatively, we say $\mathcal{T}$ is inscribed in $\mathcal{C}$. To find $\mathcal{C}$, we start by looking for all circles that pass through A and B. This family is described as follows.

Proposition 1.10.5. *Let c denote the midpoint of the line segment L_{AB} from A to B, and let $L_c^\perp$ denote the line through c orthogonal to L_{AB}. Pick $p \in L_c^\perp$ and set*

$$r = |p - A| = |p - B|. \tag{1.10.28}$$

Then the circle $S_r(p)$ passes through A and B. All circles containing A and B arise in this fashion.

Proof. Exercise. Use symmetry to show that $|p - A| = |p - B|$. □

To proceed, let b denote the midpoint of the line segment L_{AC} from A to C, and let $L_b^\perp$ denote the line through b orthogonal to L_{AC}. Since L_{AB} and L_{AC} are not parallel, neither are $L_c^\perp$ and $L_b^\perp$. Say

$$p = L_c^\perp \cap L_b^\perp. \tag{1.10.29}$$

Then r in (1.10.28) also satisfies

$$r = |p - A| = |p - C|. \tag{1.10.30}$$

Hence $S_r(p)$ passes through A, B, and C.

Doing the construction above yields a bonus. Namely, we have

$$|p - A| = |p - B| = |p - C| = r, \tag{1.10.31}$$

and the second identity implies

$$p \in L_a^{\perp}, \tag{1.10.32}$$

where a is the midpoint of the line segment L_{BC} and $L_a^{\perp}$ is the line through a orthogonal to L_{BC}.

Proposition 1.10.6. *Given a triangle $\mathcal{T} = \triangle(A, B, C)$, the perpendicular bisectors $L_a^{\perp}, L_b^{\perp}$, and $L_c^{\perp}$ of the sides of $\mathcal{T}$ all meet at a point, namely at the center of the circle that circumscribes $\mathcal{T}$.*

We next show that, given a triangle $\mathcal{T} = \triangle(A, B, C)$, there is a unique circle $\mathcal{C}$ that touches each side of $\mathcal{T}$ at just one point; we say $\mathcal{C}$ is tangent to each side of $\mathcal{T}$, and that $\mathcal{C}$ is inscribed in $\mathcal{T}$.

To start, we look at the problem of finding the circle $\mathcal{C}$, centered at p, tangent to a line L, given $p \notin L$. To produce $\mathcal{C}$, simply drop a perpendicular $L_p^{\perp}$ from p to L, and say $L_p^{\perp}$ intersects L at q. Then

$$\mathcal{C} = S_r(p), \quad r = |p - q|. \tag{1.10.33}$$

To continue, we look for those circles tangent to the two sides L_{AB} and L_{AC} of $\mathcal{T}$. Symmetry considerations imply that the center of such a circle lies on the angle bisector ℓ_A of the angle of $\mathcal{T}$ at A. Similarly, if ℓ_B is the angle bisector of $\mathcal{T}$ at B, and if

$$p = \ell_A \cap \ell_B, \tag{1.10.34}$$

then there is a circle centered at p, tangent to L_{AB} and to L_{AC}, and there is also a circle, centered at p, tangent to L_{AB} and to L_{BC}. Since both circles are tangent to L_{AB}, they must coincide. Thus we have a circle, centered at p, given by (1.10.34), that is tangent to each of the three sides of $\mathcal{T} = \triangle(A, B, C)$. This is the circle inscribed in $\mathcal{T}$.

Such a construction of inscribed circles also yields a bonus. Namely, if this inscribed circle is $S_r(p)$, then we must have

$$p \in \ell_C, \tag{1.10.35}$$

where ℓ_C denotes the angle bisector of $\mathcal{T}$ at C.

Proposition 1.10.7. *Given a triangle $\mathcal{T} = \triangle(A, B, C)$, the angle bisectors of $\mathcal{T}$ at its three vertices all meet at a point, namely at the center of the circle inscribed in $\mathcal{T}$.*

Chapter 2

Bringing in calculus

Once circles enter the picture, a proper treatment of Euclidean geometry requires substantially more sophistication. The thing about circles, as opposed to lines, triangles, and other polygons, is that they are, well, *curved*. This makes the computation of various geometric quantities associated to circles more subtle. For example, the length of a line segment is simply the distance between its endpoints. The total length of the perimeter of a polygon is the sum of the lengths of the line segments from one vertex to the next. How about the length of a circle? One can ask similar questions about areas.

These problems were tackled by Archimedes, who approached the computations as limits of a sequence of approximations obtained by replacing a circle by a polygon with many sides. This is the sort of brilliant insight that can lead to a revolution in mathematics. However, Archimedes was too far ahead of his time, and the revolution had to wait for centuries. The revolution was calculus.

Section 2.1 gives a short presentation of basic calculus. We define the derivative and the integral, and discuss the fundamental theorem of calculus, connecting these two concepts. We also treat such key results as the mean value theorem, the inverse function theorem, and the change of variable formula for the integral, as well as functions given by power series and their derivatives.

Section 2.2 uses methods of calculus to treat curves and arclength. We derive the arclength formula

$$\ell(\gamma) = \int_a^b |\gamma'(t)|\, dt, \tag{2.0.1}$$

for an appropriate class of curves $\gamma : [a, b] \to \mathbb{C}$. We discuss parametrizing a curve by arclength. When this is applied to the unit circle, one obtains the parametrization

$$\text{cis} : \mathbb{R} \longrightarrow \mathbb{C}, \quad \text{cis}(t) = \cos t + i \sin t, \tag{2.0.2}$$

bringing in the classical trigonometric functions cos and sin. We also define the real measurement of angles, in radians.

In §2.3 we introduce the exponential function e^z, defined by a power series, valid for all $z \in \mathbb{C}$, so as to solve a basic differential equation, $(d/dt)e^{at} = ae^{at}$, given $a \in \mathbb{C}$. We show that, for $t \in \mathbb{R}$,

$$e^{it} = \text{cis}(t), \tag{2.0.3}$$

which is the famous Euler formula. This leads to a characterization of $\pi = \ell(S^1)/2$ as the smallest positive number such that

$$e^{\pi i} = -1. \tag{2.0.4}$$

We show that, for a triangle, the angle sum is π.

Section 2.4 pursues further results on Euler's formula, trigonometric functions, and π, including numerical evaluation of π.

Section 2.5 studies regular polygons, specifying the vertices of a regular n-gon inscribed in a circle using the function e^{it}. We also derive an algebraic formula for the vertices of a regular pentagon.

Section 2.6 introduces the concept of area of a set $S \in \mathbb{C}$. We define this in terms of taking tilings of $\mathbb{C}$, counting how many tiles are contained in S, and how many tiles it takes to contain S, and passing to the limit as the size of the tiles vanishes. We show that Area S is invariant under the application of $F \in \text{Isom}(\mathbb{C})$. The analysis brings in a popular approach to a proof of the Pythagorean theorem and turns it on its head. We obtain an integral formula for the area of a set of points in $\mathbb{C}$ bounded by the graphs of two continuous functions, and show that if D_1 is the unit disk, then

$$\text{Area}\, D_1 = \pi. \tag{2.0.5}$$

Section 2.7 derives a formula for the area of a triangle, strictly in terms of its sidelengths, known as Heron's formula.

Section 2.8 discusses the set $\mathbb{E} \subset \mathbb{C}$ of "Euclidean numbers," designed to consist of those points in $\mathbb{C}$ that can be constructed via compass and straightedge. It is seen that $\mathbb{E}$ is a field, satisfying

$$z \in \mathbb{E},\ w \in \mathbb{C},\ w^2 = z \Longrightarrow w \in \mathbb{E}. \tag{2.0.6}$$

Also, $\mathbb{E}$ is countable.

In §2.9, we study the class of linear fractional transformations. We apply this to an approach to the three circles problem: given three disjoint circles

$C_j \subset \mathbb{C}$, no one of which separates the other two, find a fourth circle S tangent to each C_j.

Finally, §2.10 gives a brief taste of the algebraic approach to Euclidean geometry in higher dimensions. Topics include the notion of volume of sets in $\mathbb{R}^n$, and areas of $(n-1)$-dimensional surfaces in $\mathbb{R}^n$, with applications to the area of spheres S^{n-1} and volumes of balls in $\mathbb{R}^n$. We contrast the algebraic approaches taken in dimension 2 and in dimension $n > 2$, and see a synthesis in case $n = 4$, involving the algebra of quaternions.

2.1. Outline of calculus

Calculus has two facets, differential calculus and integral calculus, tied together by the fundamental theorem of calculus. Differential calculus studies the rate of change of a function $f(x)$, given in terms of the derivative,

$$f'(x) = \lim_{h\to 0} \frac{1}{h}(f(x+h) - f(x)). \tag{2.1.1}$$

We say f is differentiable at x if this limit exists.

Integral calculus studies the accumulated quantity over an interval $[a, b]$ associated to a function g, given in terms of the integral,

$$\int_a^b g(x)\,dx = \lim_{N\to\infty} \sum_{k=1}^{N} g(x_k)(x_k - x_{k-1}), \tag{2.1.2}$$

where $x_k = a + (k/N)(b-a)$, given that $g : [a, b] \to \mathbb{C}$ is continuous. To say g is continuous at x is to say that

$$x_k \to x \Longrightarrow g(x_k) \to g(x). \tag{2.1.3}$$

Here we describe some basic results on these concepts of calculus. For details, one can see Chapter 4 of [**4**].

First, a quick application of the definition (2.1.1) shows that, for

$$p_1(x) = x, \quad p_2(x) = x^2, \tag{2.1.4}$$

we have

$$p_1'(x) = 1, \quad p_2'(x) = 2x. \tag{2.1.5}$$

Next, one verifies from the definition that if f and g are differentiable,

$$\frac{d}{dx}(fg)(x) = f'(x)g(x) + f(x)g'(x). \tag{2.1.6}$$

Applying this to (2.1.4)–(2.1.5) yields inductively that

$$\frac{d}{dx}x^n = nx^{n-1}, \tag{2.1.7}$$

for $n \in \mathbb{N}$. In addition, a direct check of (2.1.1) yields

$$\frac{d}{dx}x^{-1} = -x^{-2}, \quad \text{for } x \neq 0, \tag{2.1.8}$$

and an argument using (2.1.6) then extends (2.1.7) to $n \in \mathbb{Z}$ (assuming $x \neq 0$ if $n < 0$). Another important general result is the *chain rule*:

$$h(x) = f(g(x)) \Rightarrow h'(x) = f'(g(x))g'(x). \tag{2.1.9}$$

One major theoretical result of differential calculus is the following, known as the mean value theorem.

Proposition 2.1.1. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Assume f is differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{2.1.10}$$

The mean value theorem plays a major role in establishing the following result, known as the inverse function theorem.

Proposition 2.1.2. *Let $f : [a, b] \to \mathbb{R}$ be continuous, and differentiable on (a, b). Say $f(a) = \alpha,\ f(b) = \beta$. Assume that for $x \in (a, b)$, $f'(x) > c > 0$. Then $f : [a, b] \to [\alpha, \beta]$ is one-to-one and onto, with inverse*

$$\varphi : [\alpha, \beta] \longrightarrow [a, b], \tag{2.1.11}$$

and φ is differentiable at each point of (α, β).

Combining Proposition 1.1.2 with the chain rule gives

$$\varphi'(f(x))f'(x) = 1, \quad \forall\, x \in (a, b). \tag{2.1.12}$$

One consequence is that we have

$$p_r : (0, \infty) \longrightarrow (0, \infty), \tag{2.1.13}$$

with $r = 1/n$, inverting $p_n(x) = x^n$. We write

$$p_r(x) = x^r, \tag{2.1.14}$$

with $r = 1/n$. Use of (2.1.12) gives

$$\frac{d}{dx} x^r = r x^{r-1}, \quad x > 0, \tag{2.1.15}$$

for $r = 1/n$. Then we can define (2.1.14) for $r = m/n \in \mathbb{Q}$, and again obtain (2.1.15). Further extension to $r \in \mathbb{R}$ is conveniently accomplished using the exponential function, which will arise in §2.3.

We move to integral calculus, starting with the fundamental theorem of calculus. This has two parts.

Proposition 2.1.3. *Let $f : [a, b] \to \mathbb{C}$ be continuous, and set*

$$g(x) = \int_a^x f(y)\, dy. \tag{2.1.16}$$

Then g is differentiable for $x \in (a, b)$, and

$$g'(x) = f(x). \tag{2.1.17}$$

The proof consists of writing

$$\frac{1}{h}(g(x + h) - g(x)) = \frac{1}{h} \int_x^{x+h} f(y)\, dy, \tag{2.1.18}$$

and taking $h \to 0$.

Proposition 2.1.4. *Let $g : [a,b] \to \mathbb{C}$ and assume g and g' are continuous on $[a,b]$. Then*

$$\int_a^b g'(x)\,dx = g(b) - g(a). \tag{2.1.19}$$

To see this, set

$$G(x) = \int_a^x g'(y)\,dy, \tag{2.1.20}$$

and use Proposition 2.1.3 to show that $G'(x) = g'(x)$. Hence $F(x) = G(x) - g(x)$ satisfies $F'(x) \equiv 0$. The mean value theorem (applied separately to $\operatorname{Re} F$ and $\operatorname{Im} F$) implies F is constant, and one gets (2.1.19).

Another important tool is the following change of variable formula.

Proposition 2.1.5. *Take f as in Proposition 2.1.2, and assume f' is continuous. Then, if $u : [\alpha,\beta] \to \mathbb{C}$ is continuous and $y \in [a,b]$,*

$$\int_a^y u(f(x))f'(x)\,dx = \int_\alpha^{f(y)} u(x)\,dx. \tag{2.1.21}$$

To see this, denote the two sides of (2.1.21) by $L(y)$ and $R(y)$. By Proposition 2.1.3,

$$L'(y) = u(f(y))f'(y) = R'(y), \tag{2.1.22}$$

the second identity also using the chain rule. Hence $L(y) - R(y)$ is constant, but $L(a) = R(a) = 0$, so we have (2.1.21).

The integral plays an essential role in the study of lengths of curves (including circles) and the areas of planar regions bounded by curves (including circles). We go into these matters in Sections 2.2 and 2.6.

One further tool from calculus we will need involves power series, of the form

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \tag{2.1.23}$$

with $a_k, z \in \mathbb{C}$. Now for some classes of coefficients a_k (e.g., $a_k = k!$) this might not converge for any nonzero z. Suppose, however, that it does converge for $z = z_1 \neq 0$. Then the terms in the series are bounded, so one has

$$|a_k R^k| \le A < \infty, \quad \forall\, k \in \mathbb{Z}^+, \text{ with } R = |z_1|. \tag{2.1.24}$$

Hence, given $\alpha \in (0,1)$,

$$\begin{aligned} |z| \le \alpha A \Rightarrow |a_k z^k| &\le \alpha^k A \\ \Rightarrow \sum_k |a_n z^k| &\le A \sum_k \alpha^k = \frac{A}{1-\alpha}, \end{aligned} \tag{2.1.25}$$

and we deduce that the power series in (2.1.23) converges for all $z \in D_R(0)$, uniformly on $D_{\alpha R}(0)$ for each $\alpha < 1$, to a limit f that is continuous on $D_R(0)$. The following result will be essential in the treatment of the exponential function.

Proposition 2.1.6. *Assume the power series* (2.1.23) *converges for* $|z| < R$. *Then* $f(x)$ *is differentiable for* $x \in (-R, R)$, *and* $f'(x)$ *is given by*

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}. \tag{2.1.26}$$

For a demonstration, it is convenient first to show that the power series given in (2.1.26) converges for $|x| < R$, defining

$$g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}, \tag{2.1.27}$$

continuous on $(-R, R)$. It is relatively straightforward to show that one can integrate a convergent power series term by term, obtaining

$$\begin{aligned} \int_0^x g(t)\,dt &= \sum_{k=1}^{\infty} k a_k \int_0^x t^{k-1}\,dt \\ &= \sum_{k=1}^{\infty} a_k x^k = f(x) - f(0), \end{aligned} \tag{2.1.28}$$

and then Proposition 2.1.3 yields (2.1.26).

Remark. One can define the notion of complex differentiability and extend Proposition 2.1.6 to a result on $f'(z)$, for $z \in D_R(0)$. Such matters are treated in Chapter 1 of [**6**], but they will not play a role here.

Once we have Proposition 2.1.6, we have inductively that, for $n \in \mathbb{N}$,

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1) a_k x^{k-n}, \tag{2.1.29}$$

and in particular

$$f^{(n)}(0) = n! a_n, \tag{2.1.30}$$

hence

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k. \tag{2.1.31}$$

For example, take

$$f_r(x) = (1-x)^{-r}, \quad |x| < 1. \tag{2.1.32}$$

Then inductive use of (2.1.15) and the chain rule gives

$$f^{(k)}(x) = r(r+1)\cdots(r+k-1)(1-x)^{-r-k}, \tag{2.1.33}$$

so, granted that this function has a power series of the form (2.1.23), we have

$$(1-x)^{-r} = \sum_{k=0}^{\infty} a_k x^k, \quad a_k = \frac{r(r-1)\cdots(r+k-1)}{k!}, \tag{2.1.34}$$

for $|x| < 1$. One can use the ratio test to show that the right side converges for $|x| < 1$, defining a continuous function g_r on $(-1,1)$. To show that $f_r = g_r$ takes an additional argument, which one can find in §4.3 of [**4**]. This power series, with $r = 1/2$, will appear in calculations for arclength of a circle.

2.2. Curves, arclength, and the real measurement of angles

A curve in $\mathbb{C}$ is given by a continuous function from an interval $I \subset \mathbb{R}$ to $\mathbb{C}$,

$$\gamma : I \longrightarrow \mathbb{C}, \quad \gamma(t) = x(t) + iy(t). \tag{2.2.1}$$

Typically we work with differentiable functions; $\gamma'(t)$ is the velocity of γ, at "time" t, and its "speed" is the magnitude of $\gamma'(t)$:

$$|\gamma'(t)| = \sqrt{x'(t)^2 + y'(t)^2}. \tag{2.2.2}$$

One also calls the image of I under the map γ a curve in $\mathbb{C}$. If $u : J \to I$ is continuous, one-to-one, and onto, the map

$$\sigma : J \longrightarrow \mathbb{C}, \quad \sigma(t) = \gamma(u(t)) \tag{2.2.3}$$

has the same image as γ. We say σ is a reparametrization of γ. We typically require that u be C^1, with C^1 inverse. If also γ is C^1, the chain rule gives

$$\sigma'(t) = u'(t)\gamma'(u(t)). \tag{2.2.4}$$

Assume $I = [a, b]$ and γ is C^1. We want to define the length of this curve. We follow §4.4 of [**4**], which treats curves in n-dimensional Euclidean space $\mathbb{R}^n$.

To start, we take a partition $\mathcal{P}$ of $[a, b]$, given by

$$a = t_0 < t_1 < \cdots < t_N = b, \tag{2.2.5}$$

and set

$$\ell_{\mathcal{P}}(\gamma) = \sum_{j=1}^{N} |\gamma(t_j) - \gamma(t_{j-1})|. \tag{2.2.6}$$

See Figure 2.2.1. We convert this into something that looks like a Riemann sum for $\int_a^b |\gamma'(t)|\,dt$. We have

$$\begin{aligned} \gamma(t_j) - \gamma(t_{j-1}) &= \int_{t_{j-1}}^{t_j} \gamma'(t)\,dt \\ &= \int_{t_{j-1}}^{t_j} [\gamma'(t_j) + \gamma'(t) - \gamma'(t_j)]\,dt \\ &= (t_j - t_{j-1})\gamma'(t_j) + \int_{t_{j-1}}^{t_j} [\gamma'(t) - \gamma'(t_j)]\,dt. \end{aligned} \tag{2.2.7}$$

We deduce that

$$|\gamma(t_j) - \gamma(t_{j-1})| = (t_j - t_{j-1})|\gamma'(t_j)| + r_j, \tag{2.2.8}$$

with

$$|r_j| \le \int_{t_{j-1}}^{t_j} |\gamma'(t) - \gamma'(t_j)|\,dt, \tag{2.2.9}$$

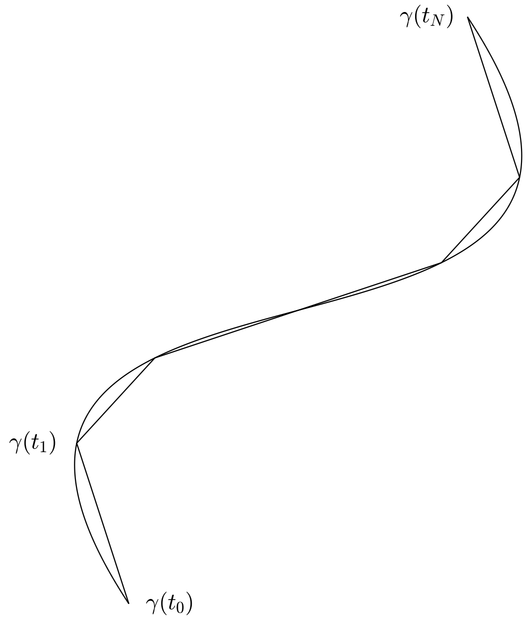

Figure 2.2.1. Approximating $\ell(\gamma)$ by $\ell_{\mathcal{P}}(\gamma)$

If γ' is continuous on $[a, b]$, so is $|\gamma'|$, and in fact both are uniformly continuous on $[a, b]$. Hence

$$s, t \in [a, b],\ |s - t| \le h \Longrightarrow |\gamma'(t) - \gamma'(s)| \le \omega(h), \tag{2.2.10}$$

where $\omega(h) \to 0$ as $h \to 0$. Summing (2.2.8) over j, we get

$$\ell_{\mathcal{P}}(\gamma) = \sum_{j=1}^{N} |\gamma'(t_j)|(t_j - t_{j-1}) + R_{\mathcal{P}}, \tag{2.2.11}$$

with

$$|R_{\mathcal{P}}| \le (b - a)\omega(h)), \quad \text{if each } t_j - t_{j-1} \le h. \tag{2.2.12}$$

Since the sum on the right side of (2.2.11) is a Riemann sum, we have:

Proposition 2.2.1. *Assume $\gamma : [a, b] \to \mathbb{C}$ is a C^1 curve. Then*

$$\ell_{\mathcal{P}}(\gamma) \longrightarrow \int_a^b |\gamma'(t)|\, dt, \tag{2.2.13}$$

as the maximal lengths of the intervals in the partition tend to 0.

We call the limit the length of the curve γ, and write

$$\ell(\gamma) = \int_a^b |\gamma'(t)|\, dt. \tag{2.2.14}$$

Note that if $u : [\alpha, \beta] \to [a, b]$ is a C^1 map with C^1 inverse, and $\sigma = \gamma \circ u$, as in (2.2.3), we have from (2.2.4) that $|\sigma'(t)| = |u'(t)| \cdot |\gamma'(t)|$, and the change of variable formula for integrals gives

$$\int_\alpha^\beta |\sigma'(t)|\, dt = \int_a^b |\gamma'(t)|\, dt, \tag{2.2.15}$$

verifying the geometrically natural result that

$$\ell(\sigma) = \ell(\gamma). \tag{2.2.16}$$

Given such a C^1 curve γ, it is natural to consider the length function

$$\ell_\gamma(t) = \int_a^t |\gamma'(s)|\, ds, \quad \ell'_\gamma(t) = |\gamma'(t)|. \tag{2.2.17}$$

If we assume that γ' is nowhere vanishing on $[a, b]$, the inverse function theorem, Proposition 2.1.2, implies that $\ell_\gamma : [a, b] \to [0, \ell(\gamma)]$ has a C^1 inverse

$$u : [0, \ell(\gamma)] \longrightarrow [a, b], \tag{2.2.18}$$

and then $\sigma = \gamma \circ u : [0, \ell(\gamma)] \to \mathbb{C}$ satisfies

$$\begin{aligned} \sigma'(t) &= u'(t)\gamma'(u(t)) \\ &= \frac{1}{\ell'_\gamma(s)} \gamma'(u(t)), \quad \text{for } t = \ell_\gamma(s),\ s = u(t), \end{aligned} \tag{2.2.19}$$

since the chain rule applied to $u(\ell_\gamma(t)) = t$ yields $u'(\ell_\gamma(t))\ell'_\gamma(t) = 1$. Also, by (2.2.17), $\ell'_\gamma(s) = |\gamma'(s)| = |\gamma'(u(t))|$, so

$$|\sigma'(t)| \equiv 1. \tag{2.2.20}$$

Then σ is a reparametrization of γ, and σ has unit speed. We say σ is a reparametrization by arc length.

We apply these considerations to the circle

$$S^1 = \{z \in \mathbb{C} : |z| = 1\}. \tag{2.2.21}$$

We can parametrize S^1 away from $z = \pm 1$ by

$$\gamma_+(t) = t + i\sqrt{1 - t^2}, \quad \gamma_-(t) = t - i\sqrt{1 - t^2}, \tag{2.2.22}$$

on the intersection of S^1 with $\{z : \operatorname{Im} z > 0\}$ and $\{z : \operatorname{Im} z < 0\}$, respectively. Here $\gamma_\pm : (-1, 1) \to \mathbb{C}$, and both maps are smooth. In fact, we can take

$\gamma_\pm : [-1,1] \to \mathbb{C}$, but these functions are not differentiable at ± 1. We can also parametrize S^1 away from $z = \pm i$, by

$$\gamma_\ell(t) = -\sqrt{1-t^2} + it, \quad \gamma_r(t) = \sqrt{1-t^2} + it, \tag{2.2.23}$$

again with $t \in (-1,1)$. A calculation gives

$$|\gamma_+'(t)|^2 = 1 + \frac{t^2}{1-t^2} = \frac{1}{1-t^2}. \tag{2.2.24}$$

Hence, if $\ell(t)$ is the length of the image $\gamma_+([0,t])$, we have

$$\ell(t) = \int_0^t \frac{1}{\sqrt{1-s^2}}\, ds, \quad \text{for } 0 < t < 1. \tag{2.2.25}$$

The same formula holds with γ_+ replaced by γ_-, γ_ℓ, or γ_r.

These formulas imply that S^1 can be parametrized by arc length. We take $C : \mathbb{R} \to \mathbb{C}$ to be such a parametrization, satisfying

$$C(0) = 1, \quad C'(0) = i. \tag{2.2.26}$$

We are led to the standard trigonometrical functions $\cos t$ and $\sin t$, real-valued functions of t defined by

$$C(t) = \operatorname{cis} t = \cos t + i \sin t. \tag{2.2.27}$$

Also, it is standard to identify the arclength t with the *real angle measurement* (in radians) of the angle made by the wedge $W_{0,1,\omega}$:

$$\Omega(W_{0,1,\omega}) = \omega = \operatorname{cis} t, \quad \angle(W_{0,1,\omega}) = t. \tag{2.2.28}$$

See Figure 2.2.2. Here $\Omega(W_{0,1,\omega})$ is the angle measurement defined in (1.4.3).

We turn to an evaluation of the derivative $C'(t)$. Applying d/dt to the identity

$$\langle C(t), C(t) \rangle = 1 \tag{2.2.29}$$

and using the product formula gives

$$\langle C'(t), C(t) \rangle = 0, \tag{2.2.30}$$

hence $C'(t) \perp C(t)$, or equivalently $C'(t) = \pm i C(t)$ (since $C'(t)$ and $C(t)$ both have unit length). Since $C'(0) = i$, there is only one possibility:

$$C'(t) = iC(t), \quad \text{i.e.,} \quad \frac{d}{dt} \operatorname{cis} t = i \operatorname{cis} t. \tag{2.2.31}$$

In light of (2.2.27), this gives the classical identities

$$\frac{d}{dt} \cos t = -\sin t, \quad \frac{d}{dt} \sin t = \cos t. \tag{2.2.32}$$

In the next section we will provide another path to a direct unit speed parametrization of S^1, connecting trigonometric functions to the exponential function, in the classical Euler formula.

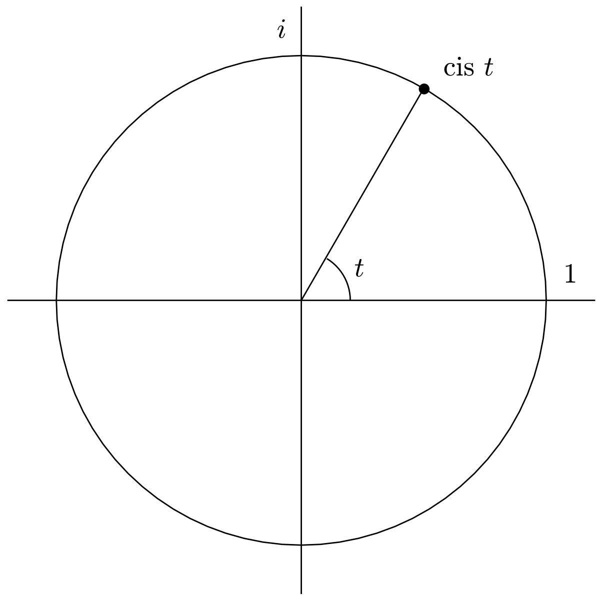

Figure 2.2.2. The circle $C(t) = \text{cis } t$

The famous number π is defined as follows:

$$\pi = \ell(S^1_+). \tag{2.2.33}$$

Equivalently,

$$\frac{\pi}{2} = \ell(S^1_{++}), \tag{2.2.34}$$

where

$$S^1_+ = \{z \in S^1 : \operatorname{Im} z > 0\}, \quad S^1_{++} = \{z \in S^1_+ : \operatorname{Re} z > 0\}. \tag{2.2.35}$$

Going further, the bisector of $W_{0,1,i}$ intersects S^1_{++} at

$$\omega = \frac{1+i}{\sqrt{2}}, \tag{2.2.36}$$

and we have, by (2.2.25),

$$\frac{\pi}{4} = \int_0^{1/\sqrt{2}} \frac{ds}{\sqrt{1-s^2}}. \tag{2.2.37}$$

Behind the results (2.2.34)–(2.2.37) is the following observation.

Proposition 2.2.2. *If* $\gamma : [a, b] \to \mathbb{C}$ *is a smooth curve and* $F \in \mathrm{Isom}(\mathbb{C})$, *then*

$$\ell(\gamma) = \ell(F(\gamma)). \tag{2.2.38}$$

Further results bearing on formulas for π and its numerical evaluation are given in §2.4.

To end this section, we mention the following *law of cosines*. Given nonzero $u, v \in \mathbb{C}$, we have the angle measurement defined in (1.4.3),

$$\Omega(u, v) = \frac{u}{v} \frac{|v|}{|u|} = \frac{u\overline{v}}{|u|\,|v|} \in S^1. \tag{2.2.39}$$

As in (2.2.28), write

$$\Omega(u, v) = \operatorname{cis} t, \quad -\pi < t \le \pi. \tag{2.2.40}$$

Then (cf. (1.5.11)),

$$\langle u, v \rangle = \operatorname{Re} u\overline{v} = |u|\,|v| \operatorname{Re} \Omega(u, v), \tag{2.2.41}$$

which yields the law of cosines:

$$\langle u, v \rangle = |u|\,|v| \cos t. \tag{2.2.42}$$

2.3. Exponential and trigonometric functions, and Euler's formula

Here we present the major insight of Euler relating the exponential and trigonometric functions. We start with a definition of the exponential function, both for real and complex arguments.

We construct the exponential function to solve the differential equation

$$\frac{dx}{dt} = x, \quad x(0) = 1. \tag{2.3.1}$$

We seek a solution as a power series

$$x(t) = \sum_{k=0}^{\infty} a_k t^k. \tag{2.3.2}$$

In such a case, if this series converges for $|t| < R$, then, as stated in §2.1,

$$\begin{aligned} x'(t) &= \sum_{k=1}^{\infty} k a_k t^{k-1} \\ &= \sum_{\ell=0}^{\infty} (\ell+1) a_{\ell+1} t^{\ell}, \end{aligned} \tag{2.3.3}$$

so for (2.3.1) to hold we need

$$a_0 = 1, \quad a_{k+1} = \frac{a_k}{k+1}, \tag{2.3.4}$$

i.e., $a_k = 1/k!$, where $k! = k(k-1)\cdots 2 \cdot 1$. Thus (2.3.1) ia solved by

$$x(t) = e^t = \sum_{k=0}^{\infty} \frac{1}{k!} t^k, \quad t \in \mathbb{R}. \tag{2.3.5}$$

This defines the exponential function e^t, for $t \in \mathbb{R}$.

More generally, we can define

$$e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k, \quad z \in \mathbb{C}. \tag{2.3.6}$$

The ratio test then shows that the series (2.3.6) is absolutely convergent for all $z \in \mathbb{C}$, and uniformly convergent for $|z| \le R$, for each $R < \infty$. Results given in §2.1 also imply that

$$e^{at} = \sum_{k=0}^{\infty} \frac{a^k}{k!} t^k \tag{2.3.7}$$

solves

$$\frac{d}{dt} e^{at} = a e^{at}, \tag{2.3.8}$$

and this works for each $a \in \mathbb{C}$,

We claim that e^{at} is the unique solution to

$$\frac{dy}{dt} = ay, \quad y(0) = 1. \tag{2.3.9}$$

To see this, compute the derivative of $e^{-at}y(t)$:

$$\frac{d}{dt}(e^{-at}y(t)) = -ae^{-at}y(t) + e^{-at}ay(t) = 0, \tag{2.3.10}$$

where we use the product rule, (2.3.8) (with a replaced by $-a$) and (2.3.9). Thus $e^{-at}y(t)$ is independent of t. Evaluating at $t = 0$ gives

$$e^{-at}y(t) = 1, \quad \forall\, t \in \mathbb{R}, \tag{2.3.11}$$

whenever $y(t)$ solves (2.3.9). Since e^{at} solves (2.3.9), we have $e^{-at}e^{at} = 1$, hence

$$e^{-at} = \frac{1}{e^{at}}. \quad \forall\, t \in \mathbb{R},\ a \in \mathbb{C}. \tag{2.3.12}$$

Thus multiplying both sides of (2.3.11) by e^{at} gives the desired uniqueness:

$$y(t) = e^{at}, \quad \forall\, t \in \mathbb{R}. \tag{2.3.13}$$

We can draw further useful conclusions from applying d/dt to products of exponential functions. In fact, let $a, b \in \mathbb{C}$. Then

$$\begin{aligned} &\frac{d}{dt}\Big(e^{-at}e^{-bt}e^{(a+b)t}\Big) \\ &= -ae^{-at}e^{-bt}e^{(a+b)t} - be^{-at}e^{-bt}e^{(a+b)t} + (a+b)e^{-at}e^{-bt}e^{(a+b)t} \\ &= 0, \end{aligned} \tag{2.3.14}$$

so again we are differentiating a function that is independent of t. Evaluation at $t = 0$ gives

$$e^{-at}e^{-bt}e^{(a+b)t} = 1, \quad \forall\, t \in \mathbb{R}. \tag{2.3.15}$$

Again using (2.3.12), we get

$$e^{(a+b)t} = e^{at}e^{bt}, \quad \forall\, t \in \mathbb{R}, a, b \in \mathbb{C}, \tag{2.3.16}$$

or, setting $t = 1$,

$$e^{a+b} = e^a e^b, \quad \forall\, a, b \in \mathbb{C}. \tag{2.3.17}$$

We next record some properties of $\exp(t) = e^t$ for real t. The power series (2.3.5) clearly gives $e^t > 0$ for $t \geq 0$. Since $e^{-1} = 1/e^t$, we see that $e^t > 0$ for all $t \in \mathbb{R}$. Since $de^t/dt = e^t > 0$, the function is monotone increasing in t. Note that, for $t > 0$,

$$e^t = 1 + t + \frac{t^2}{2} + \cdots > 1 + t \nearrow +\infty, \tag{2.3.18}$$

as $t \to \infty$. Hence

$$\lim_{t\to+\infty} e^t = +\infty. \tag{2.3.19}$$

Since $e^{-t} = 1/e^t$,

$$\lim_{t\to-\infty} e^t = 0. \tag{2.3.20}$$

As a consequence,

$$\exp : \mathbb{R} \longrightarrow (0,\infty) \tag{2.3.21}$$

is one-to-one and onto, with positive derivative, so there is a smooth inverse:

$$L : (0,\infty) \longrightarrow \mathbb{R}. \tag{2.3.22}$$

We call this inverse the natural logarithm:

$$\log x = L(x). \tag{2.3.23}$$

Applying d/dt to

$$L(e^t) = t \tag{2.3.24}$$

gives $L'(e^t)e^t = 1$, hence $L'(e^t) = 1/e^t$, so

$$\frac{d}{dx} \log x = \frac{1}{x}. \tag{2.3.25}$$

Since $\log 1 = 0$, we get

$$\log x = \int_1^x \frac{dy}{y}. \tag{2.3.26}$$

We move on to a study of e^z for purely imaginary z, i.e., of

$$\gamma(t) = e^{it}, \quad t \in \mathbb{R}. \tag{2.3.27}$$

This traces out a curve in $\mathbb{C}$, and we want to understand which curve it is. First we calculate $|e^{it}|$. Recall that $|z|^2 = z\overline{z}$. The basic results (0.0.8) on complex conjugates of sums and products yield

$$\overline{e^z} = \sum_{k=0}^{\infty} \frac{\overline{z}^k}{k!} = e^{\overline{z}}, \tag{2.3.28}$$

for $z \in \mathbb{C}$. In particular,

$$t \in \mathbb{R} \Longrightarrow |e^{it}|^2 = e^{it}e^{-it} = 1. \tag{2.3.29}$$

Hence $\gamma(t) = e^{it}$ lies on the unit circle centered at the origin in $\mathbb{C}$. Also

$$\gamma'(t) = ie^{it} \Longrightarrow |\gamma'(t)| \equiv 1, \tag{2.3.30}$$

so $\gamma(t)$ moves at unit speed on the unit circle. Also $\gamma(0) = 1,\ \gamma'(0) = i$.

It follows that $\gamma(t) = e^{it}$ is the parametrization of the unit circle by arc length introduced in (2.2.27). This establishes the following result, known as Euler's formula.

Proposition 2.3.1. *For* $t \in \mathbb{R}$,

$$e^{it} = \operatorname{cis} t = \cos t + i \sin t. \tag{2.3.31}$$

Given Proposition 2.3.1, the identity

$$\frac{d}{dt} = ie^{it}, \tag{2.3.32}$$

hence gives a second proof of (2.2.31), i.e.,

$$\frac{d}{dt} \operatorname{cis} t = i \operatorname{cis} t, \tag{2.3.33}$$

or equivalently (2.2.32), i.e.,

$$\frac{d}{dt} \cos t = -\sin t, \quad \frac{d}{dt} \sin t = \cos t. \tag{2.3.34}$$

Given Euler's formula, we can use (2.3.17) to derive formulas for sin and cos of the sum of two angles. Indeed, comparing

$$e^{i(s+t)} = \cos(s+t) + i \sin(s+t) \tag{2.3.35}$$

with

$$e^{is}e^{it} = (\cos s + i \sin s)(\cos t + i \sin t) \tag{2.3.36}$$

gives

$$\begin{aligned} \cos(s+t) &= (\cos s)(\cos t) - (\sin s)(\sin t), \\ \sin(s+t) &= (\sin s)(\cos t) + (\cos s)(\sin t). \end{aligned} \tag{2.3.37}$$

Another compact presentation of these sum formulas is simply

$$\operatorname{cis}(s+t) = (\operatorname{cis} s)(\operatorname{cis} t). \tag{2.3.38}$$

Remark. An alternative approach to Euler's formula (2.3.31) is to use the formula for the derivative of $\operatorname{cis} t$, (2.3.33), relying on the proof in (2.2.31), together with the initial condition $\operatorname{cis}(0) = 1$. Then the uniqueness result (2.3.9)–(2.3.13) implies that $\operatorname{cis} t = e^{it}$.

Recall from (2.2.33) that the number π is defined by the identity

$$\pi = \ell(S^1_+). \tag{2.3.39}$$

Equivalently, π is the smallest positive number such that

$$\operatorname{cis} \pi = -1, \tag{2.3.40}$$

or alternatively, by (2.3.31), the smallest positive number such that

$$e^{\pi i} = -1. \tag{2.3.41}$$

Note also that

$$e^{-\pi i} = -1, \quad e^{\pm 2\pi i} = 1. \tag{2.3.42}$$

In fact, we have the following.

Proposition 2.3.2. *The maps*

$$\text{cis} : (0, \pi) \to S^1_+, \quad \text{cis} : (0, 2\pi) \to S^1 \setminus \{1\} \tag{2.3.43}$$

are one-to-one and onto. Similarly,

$$\text{cis}^-(t) = \overline{\text{cis}\, t} = \text{cis}(-t) \tag{2.3.44}$$

defines

$$\text{cis}^- : (0, \pi) \to S^1_-, \quad \text{cis}^- : (0, 2\pi) \to S^1 \setminus \{1\}, \tag{2.3.45}$$

and these maps are one-to-one and onto.

Another important role played by π is in the angle sum formula for triangles.

Proposition 2.3.3. *Let $\mathcal{T} = \triangle(A, B, C)$ be a triangle. Then*

$$\Omega_{\mathcal{T}}(A) = \alpha = e^{is}, \ \Omega_{\mathcal{T}}(B) = \beta = e^{it}, \ \Omega_{\mathcal{T}}(C) = \gamma = e^{iu}, \tag{2.3.46}$$

where either

$$\alpha, \beta, \gamma \in S^1_+, \quad s, t, u \in (0, \pi), \tag{2.3.47}$$

or

$$\alpha, \beta, \gamma \in S^1_-, \quad s, t, u \in (-\pi, 0). \tag{2.3.48}$$

Then

$$s + t + u = \pi \quad \textit{or} \quad -\pi, \tag{2.3.49}$$

in these two respective cases.

Proof. The cases (2.3.47) and (2.3.48) hold by Propositions 1.5.2 and 2.3.2. Proposition 1.5.1 says

$$\alpha\beta\gamma = -1, \tag{2.3.50}$$

hence

$$e^{i(s+t+u)} = -1. \tag{2.3.51}$$

Consider case (2.3.47). Since $\alpha\beta = -\gamma^{-1}$, we also have

$$\alpha\beta \in S^1_+, \tag{2.3.52}$$

hence $s+t \in (0, \pi)$. Thus $s+t+u < 2\pi$, and then (2.3.51) forces $s+t+u = \pi$. The analysis of the case (2.3.48) is similar. □

REMARK. We could skip the use of (2.3.52). In fact, given $\theta \in \mathbb{R}$,

$$e^{i\theta} = -1 \Longleftrightarrow \theta = \pi + 2k\pi, \quad \text{for some} \quad k \in \mathbb{Z}, \tag{2.3.53}$$

and since $s, t, u \in (0, \pi) \Rightarrow s + t + u \in (0, 3\pi)$, (2.3.51) forces this sum to be π.

2.4. More on Euler's formula, trigonometric functions, and π

We continue examining some consequences of the interaction of Euler's formula and π, initiated in (2.3.39)–(2.3.42), which gave

$$e^{\pm \pi i} = -1, \quad e^{\pm 2\pi i} = 1, \quad e^{\pm \pi i/2} = \pm i. \tag{2.4.1}$$

For one, as seen in §1.4,

$$\omega \in S^1_{\pm},\ \alpha = \frac{\omega + 1}{|\omega + 1|} \Rightarrow \alpha \in S^1_{\pm} \text{ and } \alpha^2 = \omega. \tag{2.4.2}$$

Consequently,

$$e^{\pi i/4} = \frac{1+i}{\sqrt{2}}, \quad e^{-\pi i/4} = \frac{1-i}{\sqrt{2}}. \tag{2.4.3}$$

We next examine

$$e^{\pi i/3} = z, \quad z^3 = -1, \quad z^3 + 1 = 0. \tag{2.4.4}$$

We have

$$z^3 + 1 = (z+1)(z^2 - z + 1), \tag{2.4.5}$$

so $e^{\pi i/3} = z$ is a solution to

$$z^2 - z + 1 = 0. \tag{2.4.6}$$

The roots of this equation are given by the quadratic formula:

$$z = \frac{1}{2} \pm \frac{i}{2}\sqrt{3}. \tag{2.4.7}$$

Since $e^{\pi i/3} \in S^1_+$, we have

$$e^{\pi i/3} = \frac{1}{2} + \frac{i}{2}\sqrt{3}. \tag{2.4.8}$$

It follows that

$$\begin{aligned} e^{2\pi i/3} &= \Big(\frac{1}{2} + \frac{i}{2}\sqrt{3}\Big)^2 \\ &= -\frac{1}{2} + \frac{i}{2}\sqrt{3}, \end{aligned} \tag{2.4.9}$$

and

$$\begin{aligned} e^{\pi i/6} &= e^{\pi i/2 - \pi i/3} = e^{\pi i/2} e^{-\pi i/3} \\ &= i\Big(\frac{1}{2} - \frac{i}{2}\sqrt{3}\Big) \\ &= \frac{\sqrt{3}}{2} + \frac{i}{2}. \end{aligned} \tag{2.4.10}$$

The formula (2.4.8) in concert with (2.2.25) leads to the following variant of (2.2.37).

$$\frac{\pi}{6} = \int_0^{1/2} \frac{ds}{\sqrt{1-s^2}}. \tag{2.4.11}$$

This is the value at $t = 1/2$ of the integral (2.2.25) for $\ell(t)$. We can evaluate $\ell(t)$ as a power series in t, in the following way. As noted in §12,

$$(1-x)^{-1/2} = \sum_{k=0}^{\infty} \frac{a_k}{k!} x^k, \quad \text{for } |x| < 1, \tag{2.4.12}$$

where

$$a_0 = 1,\ a_1 = \frac{1}{2},\ a_j = \Bigl(\frac{1}{2}\Bigr)\Bigl(\frac{3}{2}\Bigr)\cdots\Bigl(k-\frac{1}{2}\Bigr). \tag{2.4.13}$$

It follows that

$$(1-s^2)^{-1/2} = \sum_{k=0}^{\infty} \frac{a_k}{k!} s^{2k}, \quad |s| < 1, \tag{2.4.14}$$

with uniform convergence on $[-a, a]$ for each $a \in (0,1)$. Then we can integrate (2.4.14) term-by-term to get

$$\ell(t) = \sum_{k=0}^{\infty} \frac{a_k}{k!} \frac{t^{2k+1}}{2k+1}, \quad 0 \le t < 1. \tag{2.4.15}$$

In particular,

$$\frac{\pi}{6} = \sum_{k=0}^{\infty} \frac{a_k}{k!(2k+1)}\, 2^{-(2k+1)}. \tag{2.4.16}$$

As seen in [**4**], (4.5.48)–(4.5.50), summing (2.4.16) over $0 \le k \le 20$ yields π to within an error $< 10^{-13}$. One gets

$$\pi \approx 3.141592653589\cdots. \tag{2.4.17}$$

Here is another take on the integral

$$\ell(t) = \int_0^t \frac{ds}{\sqrt{1-s^2}}, \quad 0 \le t < 1. \tag{2.4.18}$$

We make a change of variable

$$s = \sin\theta. \tag{2.4.19}$$

Note that, since

$$\sin\theta = \operatorname{Im} e^{i\theta}, \quad \frac{d}{d\theta}\sin\theta = \cos\theta = \operatorname{Re} e^{i\theta}, \tag{2.4.20}$$

the function $\sin\theta$ has positive derivative for $-\pi/2 < \theta < \pi/2$, hence is monotone increasing on $[-\pi/2, \pi/2]$, with inverse

$$\sin^{-1} : [-1,1] \longrightarrow \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big], \tag{2.4.21}$$

smooth on $(-1,1)$. Hence the change of variable (2.4.19) leads to $ds = \cos\theta\, d\theta$, and then to

$$\ell(t) = \int_0^{\varphi} \frac{\cos\theta\, d\theta}{\cos\theta} = \varphi, \quad \sin\varphi = t, \tag{2.4.22}$$

i.e.,

$$\ell(t) = \sin^{-1} t, \quad 0 \le t < 1. \tag{2.4.23}$$

In particular,

$$\ell\Big(\frac{1}{2}\Big) = \sin^{-1}\frac{1}{2} = \frac{\pi}{6}, \tag{2.4.24}$$

by (2.4.10).

We can obtain other formulas for π by bringing in the trigonometric function

$$\tan t = \frac{\sin t}{\cos t}, \quad |t| < \frac{\pi}{2}. \tag{2.4.25}$$

The quotient rule for derivatives gives

$$\begin{aligned} \frac{d}{dt}\tan t &= \frac{\cos^2 t + \sin^2 t}{\cos^2 t} \\ &= 1 + \tan^2 t. \end{aligned} \tag{2.4.26}$$

Thus, to evaluate the integral

$$I(y) = \int_0^y \frac{dx}{1+x^2}, \tag{2.4.27}$$

we can use the substitution $x = \tan t$, with $dx = (1 + \tan^2 t)\, dt$, to get

$$I(y) = \int_0^u dt = u, \quad \tan u = y, \tag{2.4.28}$$

i.e.,

$$I(y) = \tan^{-1} y. \tag{2.4.29}$$

On the other hand, for $|y| < 1$, the integral (2.4.27) has a convenient power series, derived as follows. We have the geometrical series

$$\frac{1}{1-r} = \sum_{k=0}^{\infty} r^k, \quad |r| < 1, \tag{2.4.30}$$

and setting $r = -x^2$ yields

$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k}, \quad |x| < 1. \tag{2.4.31}$$

Hence integrating term-by-term gives

$$\tan^{-1} y = \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{2k+1}, \quad |y| < 1. \tag{2.4.32}$$

To get formulas for π, note that computations in (2.4.3)–(2.4.10) give

$$\tan\frac{\pi}{4} = 1, \quad \tan\frac{\pi}{3} = \sqrt{3}, \quad \tan\frac{\pi}{6} = \frac{1}{\sqrt{3}}. \tag{2.4.33}$$

Only the last of these identities is of use for (2.4.32). It gives

$$\frac{\pi}{6} = \tan^{-1}\frac{1}{\sqrt{3}} = \frac{1}{\sqrt{3}} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \Bigl(\frac{1}{3}\Bigr)^k. \tag{2.4.34}$$

This has a simpler form than (2.4.16) (except for the irrational factor $1/\sqrt{3}$). On the other hand, convergence is somewhat slower. To get accuracy as in (2.4.17), one needs to sum over $0 \le k \le 30$.

We can do better by expressing π as a finite linear combination of terms $\tan^{-1} x_j$ for certain fairly small numbers x_j. The key to this is the following formula for $\tan(a+b)$. Using (2.3.37), we have

$$\begin{aligned} \tan(a+b) &= \frac{\sin(a+b)}{\cos(a+b)} = \frac{\sin a \, \cos b + \cos a \, \sin b}{\cos a \, \cos b - \sin a \, \sin b} \\ &= \frac{\tan a + \tan b}{1 - \tan a \, \tan b}. \end{aligned} \tag{2.4.35}$$

Since $\tan \pi/4 = 1$, we have, for $a, b, a+b \in (-\pi/2, \pi/2)$,

$$\frac{\pi}{4} = a + b \Longleftarrow \frac{\tan a + \tan b}{1 - \tan a \, \tan b} = 1. \tag{2.4.36}$$

Taking $a = \tan^{-1} x, \ b = \tan^{-1} y$ gives

$$\begin{aligned} \frac{\pi}{4} = \tan^{-1} x + \tan^{-1} y &\Longleftarrow x + y = 1 - xy \\ &\Longleftarrow x = \frac{1-y}{1+y}. \end{aligned} \tag{2.4.37}$$

If we set $y = 1/2$, we get $x = 1/3$, so

$$\frac{\pi}{4} = \tan^{-1}\frac{1}{3} + \tan^{-1}\frac{1}{2}. \tag{2.4.38}$$

The power series (2.4.32) for $\tan^{-1}(1/3)$ and $\tan^{-1}(1/2)$ both converge faster than (2.4.34), but that for $\tan^{-1}(1/2)$ converges at essentially the same rate as (2.4.16). We might optimize by taking $x = y$ in (2.4.37), but that yields

$x = y = \sqrt{2} - 1$, and we do not want to plug this irrational number into (2.4.32). Taking a cue from $\sqrt{2} - 1 \approx 0.414$, we set $y = 2/5$, which yields $x = 3/7$, so

$$\frac{\pi}{4} = \tan^{-1}\frac{2}{5} + \tan^{-1}\frac{3}{7}. \tag{2.4.39}$$

Both resulting power series converge a little faster than (2.4.16).

To do better, one can bring in a formula for $\tan(a + 2b)$. Computations somewhat parallel to (2.4.35)–(2.4.37) give rise to the following analogues of (2.4.38)–(2.4.39):

$$\begin{aligned} \frac{\pi}{4} &= \tan^{-1}\frac{1}{7} + 2\tan^{-1}\frac{1}{3}, \\ \frac{\pi}{4} &= \tan^{-1}\frac{7}{23} + \tan^{-1}\frac{1}{4}. \end{aligned} \tag{2.4.40}$$

Details can be found in [**4**], Appendix A.5. All the power series here converge significantly faster than the series in (2.4.16).

Going further, one can bring in a formula for $\tan(a+4b)$ and run through similar computations, producing a number of identities, including the following, known as Machin's formula:

$$\frac{\pi}{4} = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}. \tag{2.4.41}$$

We now discuss another approach to approximating π. Namely, we try to solve

$$\sin x = 0, \tag{2.4.42}$$

by an iterative procedure starting from an initial guess that is known to be a rough approximation to π. The iteration is the following: given an approximation x_k, we set

$$x_{k+1} = x_k + \sin x_k. \tag{2.4.43}$$

This approximation is predicated on the fact that a rapid calculation of $\sin x_k$ is feasible. To analyze how much x_{k+1} is an improvement over x_k, set

$$x_k = \pi + \delta_k. \tag{2.4.44}$$

We only know x_k, not π and not δ_k, but note that

$$\sin x_k = \sin(\pi + \delta_k) = -\sin\delta_k, \tag{2.4.45}$$

so

$$x_{k+1} = \pi + \delta_k - \sin\delta_k. \tag{2.4.46}$$

Now the power series (2.3.7), applied to $e^{it} = \cos t + i\sin t$, implies

$$\sin\delta_k = \delta_k - \frac{1}{3!}\delta_k^3 + O(\delta_k^5), \tag{2.4.47}$$

so we get

$$x_{k+1} = \pi + \delta_{k+1}, \quad \delta_{k+1} = \delta_k - \sin\delta_k = O(\delta_k^3). \tag{2.4.48}$$

This implies very rapid convergence of the sequence (x_k) to its limit (which is π), provided the initial guess is sufficiently close.

To implement this iteration, we start with

$$x_0 = 3, \tag{2.4.49}$$

and implement (2.4.43), obtaining (to 15 digits),

$$\begin{aligned} x_1 &= 3.14112000805987, \\ x_2 &= 3.14159265357220, \\ x_3 &= 3.14159265358979. \end{aligned} \tag{2.4.50}$$

The error $\pi - x_2$ is $< 2 \cdot 10^{-11}$, and all the printed digits of x_3 are accurate. Indeed, complementing (2.4.48), we can deduce from the fact that $\sin\delta_k$ has alternating series that

$$|\delta_k| \le 1 \Longrightarrow |\delta_{k+1}| \le \frac{1}{6}|\delta_k^3|. \tag{2.4.51}$$

Consequently

$$\begin{aligned} |\delta_2| < 2 \cdot 10^{-11} &\Longrightarrow |\delta_3| < \frac{4}{3} \cdot 10^{-33} \\ &\Longrightarrow |\delta_4| < 4 \cdot 10^{-100}. \end{aligned} \tag{2.4.52}$$

Regarding the ability to compute e^{ix} rapidly and accurately for $x \approx 3$, note that the power series (2.3.7) for e^{it} converges very fast, due the the $k!$ in the denominator, especially when $|t| < 1/2$. In the current situation, we can bring in

$$e^{ix} = \alpha^8, \quad \alpha = e^{ix/8}, \quad \alpha^8 = (((\alpha)^2)^2)^2. \tag{2.4.53}$$

Such devices are routinely used in the software for a computer program that has a command to compute $\sin x$ and $\cos x$.

Here is another wrinkle, once you have done the first iteration. If you have just computed x_k and already have $e^{ix_{k-1}}$, then write

$$e^{ix_k} = e^{ix_{k-1}} e^{i\gamma_k}, \quad \gamma_k = x_k - x_{k-1}, \tag{2.4.54}$$

and use the rapidly convergent power series for $e^{i\gamma_k}$.

2.5. Regular polygons

The regular polygon $\mathcal{P}_n$ is specified as having n vertices, in S^1, one vertex being 1, and as being invariant under the rotation ρ_n, given by

$$\rho_n(z) = \omega_n z, \quad \omega_n = e^{2\pi i/n}. \tag{2.5.1}$$

Then the complete set of vertices of $\mathcal{P}_n$ is

$$\omega_n^k = e^{2k\pi i/n}, \quad 0 \le k \le n-1. \tag{2.5.2}$$

These quantities are precisely the roots of the polynomial

$$\mathcal{P}_n(z) = z^n - 1. \tag{2.5.3}$$

Once one has the point ω_n, the remaining vertices can be constructed with a compass, using the observation that

$$|\omega_n^{k+1} - \omega_n^k| = |\omega_n - 1|. \tag{2.5.4}$$

One can start with the points $\omega_n^0 = 1$, $\omega_n^1 = \omega_n$, and inductively produce ω_n^{k+1} from ω_n^k as a point where the circle $S_{r_n}(\omega_n^k)$ intersects S^1, where $r_n = |\omega_n - 1|$.

Some of the basic vertices ω_n have been computed in §2.4. As seen there, we have

$$\begin{aligned} \omega_3 &= e^{2\pi i/3} = -\frac{1}{2} + \frac{i}{2}\sqrt{3}, \\ \omega_4 &= e^{\pi i/2} = i, \\ \omega_6 &= e^{\pi i/3} = \frac{1}{2} + \frac{i}{2}\sqrt{3}, \\ \omega_8 &= e^{\pi i/4} = \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}. \end{aligned} \tag{2.5.5}$$

In the first three cases, one can find $x_n = \operatorname{Re}\omega_n$, an element of $\{0, \pm 1/2\}$, and read off $y_n = \operatorname{Im}\omega_n$ as the point where the line through x_n orthogonal to $\mathbb{R}$ intersects S^1_+. The number ω_8 can be obtained from ω_4 by angle bisection. Alternatively, we have $1/\sqrt{2} = \sqrt{2}/2$, and $\sqrt{2}$ can be constructed as the length of the hypothenuse of a right triangle with base and height 1.

The regular hexagon ($n = 6$) is illustrated in Figure 2.5.1. In this case we have

$$\omega_6 - 1 = \omega_3, \quad \text{hence } |\omega_6 - 1| = 1, \tag{2.5.6}$$

and ω_6 can be constructed as the intersection of S^1_+ with $S_1(1)$. Note incidentally that

$$\omega_3 + 1 = \omega_6, \quad \text{hence } |\omega_3 + 1| = 1, \tag{2.5.7}$$

so ω_3 can be constructed as the intersection of S^1_+ with $S_1(-1)$.

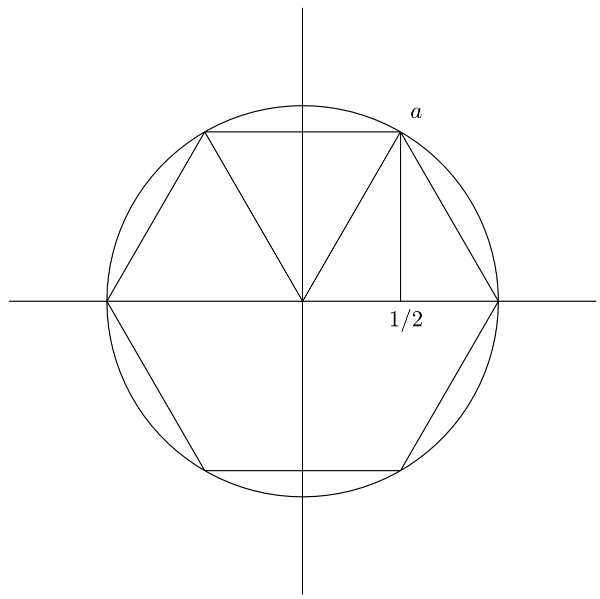

Figure 2.5.1. Regular hexagon, $a = e^{\pi i/3}$

There are two gaps in the sequence of n-gons to which the calculations in (2.5.5) apply, at $n = 5$ and $n = 7$. We discuss each of these in turn.

We first treat ω_5, which is a root of $p_5(z) = z^5 - 1$, leading to the construction of a regular pentagon. To start, write

$$z^5 - 1 = (z-1)(z^4 + z^3 + z^2 + z + 1). \tag{2.5.8}$$

Our task is to find solutions to

$$z^4 + z^3 + z^2 + z + 1 = 0. \tag{2.5.9}$$

It is useful to divide by z^2, obtaining

$$z^2 + z + 1 + \frac{1}{z} + \frac{1}{z^2} = 0. \tag{2.5.10}$$

Now a clever trick enters. Take

$$w = z + \frac{1}{z}, \quad \text{so} \quad w^2 = z^2 + 2 + \frac{1}{z^2}. \tag{2.5.11}$$

We substitute this into (2.5.10), obtaining

$$w^2 + w - 1 = 0. \tag{2.5.12}$$

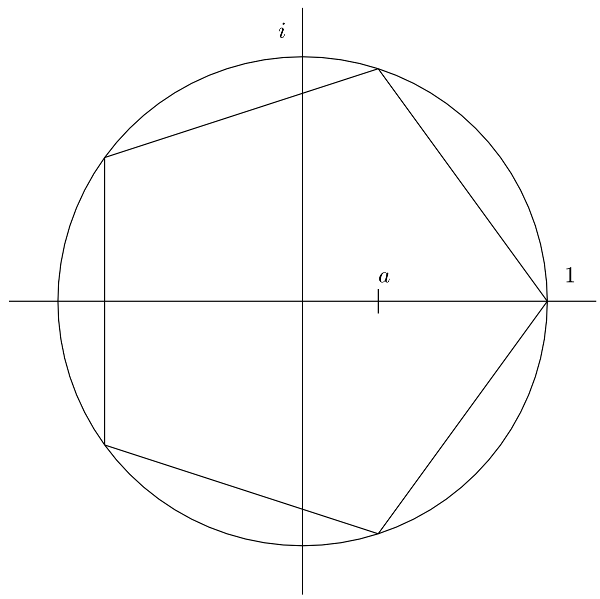

Figure 2.5.2. Regular pentagon, $a = (\sqrt{5} - 1)/4$

The quadratic formula presents the roots as

$$w = -\frac{1}{2} \pm \frac{1}{2}\sqrt{5}. \tag{2.5.13}$$

Note that since $|z| = 1$, $w = z + \overline{z} = 2\,\mathrm{Re}\, z$. We deduce that

$$\mathrm{Re}\,\omega_5 = \cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}. \tag{2.5.14}$$

Going further, we can write $w = z + 1/z$ as

$$z^2 - wz + 1 = 0, \tag{2.5.15}$$

and solve for z. We leave this to the reader. Note that ω_5 can be obtained from (2.5.14) as the point where the line through $\mathrm{Re}\,\omega_5$ orthogonal to $\mathbb{R}$ intersects S^1_+. See Figure 2.5.2.

Let us see where such an approach leads when applied to ω_7, which is a root of $p_7(z) = z^7 - 1$, hence a solution to

$$z^6 + z^5 + z^4 + z^3 + z^2 + z + 1 = 0. \tag{2.5.16}$$

Dividing by z^3 gives

$$z^3 + z^2 + z + 1 + z^{-1} + z^{-2} + z^{-3} = 0, \tag{2.5.17}$$

and again we take w as in (2.5.11). We complement (2.5.11) with

$$w^3 = z^3 + 3z + 3z^{-1} + z^{-3}, \tag{2.5.18}$$

and substitute into (2.5.17), obtaining

$$w^3 + w^2 - 2w - 1 = 0. \tag{2.5.19}$$

As opposed to (2.5.12), this is a cubic equation for w. Now there is a formula for the roots of such an equation (cf. [**6**], Appendix A.6), but of course it involves taking cube roots.

As seen in §1.10, square roots can be constructed via compass and straightedge. But it has been shown that in general cube roots cannot be so constructed, and in particular that ω_7 cannot be.

Numbers that can be constructed via compass and straightedge constitute a class that we consider further in §2.8, where the set of such numbers is denoted $\mathbb{E}$. The numbers $\omega_3, \omega_4, \omega_5$, and ω_6 belong to $\mathbb{E}$, but ω_7 does not. In fact, it has been shown that the regular n-gon $\mathcal{P}_n$ is constructible via compass and straightedge if and only if n has the form $n = 2^k$, or more generally,

$$n = 2^k \times \text{ a product of distinct Fermat primes,} \tag{2.5.20}$$

where a Fermat prime is a prime p of the form

$$p = 2^{2^k} + 1. \tag{2.5.21}$$

Examples include

$$2^1 + 1 = 2, \quad 2^2 + 1 = 5, \quad 2^4 + 1 = 17. \tag{2.5.22}$$

Such results are typically demonstrated using an area of algebra known as Galois theory. See [**1**].

A related problem that cannot be solved via compass and straightedge is that of trisecting an angle, that is, solving for $\alpha \in S^1$

$$\alpha^3 = \omega, \quad \omega \in S^1. \tag{2.5.23}$$

Of course, we can solve this using the complex exponential function. Given $\omega \in S^1$, we can write

$$\omega = e^{i\theta}, \quad -\pi < \theta \le \pi, \tag{2.5.24}$$

and then take

$$\alpha = e^{i\theta/3}. \tag{2.5.25}$$

More generally, for each $n \in \mathbb{N}$,

$$\alpha = e^{i\theta/n} \ \text{ solves } \ \alpha^n = \omega. \tag{2.5.26}$$

The general solution to $\alpha^n = \omega$ is then

$$e^{i\theta/n + 2\pi i k/n}, \quad 0 \le k \le n-1. \tag{2.5.27}$$

Here is an iterative procedure to solve

$$z^n = w, \quad w = 1 + \beta, \tag{2.5.28}$$

given $\beta \in \mathbb{C}$ sufficiently small. It does not require $|w|$ to be 1, and avoids going through (2.5.24). Start with the initial approximation

$$\zeta_0 = 1 + \frac{\beta}{n}. \tag{2.5.29}$$

Then, by the binomial formula,

$$\zeta_0^n = 1 + \beta + R\Big(\frac{\beta}{n}\Big), \quad |R(u)| \le C_n|u|^2. \tag{2.5.30}$$

Setting $z = \zeta_0 z_1^{-1}$, we reduce our task to solving

$$z_1^n = 1 + \beta_1, \quad \beta_1 = w^{-1}R\Big(\frac{\beta}{n}\Big) = O(|\beta|^2). \tag{2.5.31}$$

Iterating this procedure produces a sequence ζ_k,

$$\zeta_k = \zeta_{k-1}\Big(1 + \frac{\beta_k}{n}\Big)^{\sigma(k)}, \quad \sigma(k) = (-1)^k, \tag{2.5.32}$$

which converges rapidly to $w^{1/n}$, if $|\beta|$ is sufficiently small.

2.6. Area

Here we present material on the area of various sets $S \subset \mathbb{C}$. To start, given $a > 0$, we say S is a standard square of sidelength a provided it has the form

$$S = \mathrm{Sq}_{z_0}(a) = \{x + iy : x_0 \le x \le x_0 + a,\ y_0 \le y \le y_0 + a\}, \tag{2.6.1}$$

for some $z_0 = x_0 + iy_0 \in \mathbb{C}$. If S is such a set, we say

$$\operatorname{Area} S = a^2. \tag{2.6.2}$$

To proceed, given $\varepsilon > 0$, we tile the plane $\mathbb{C}$ with squares of sidelength ε, given by

$$S_{jk\varepsilon} = \mathrm{Sq}_{(j+ki)\varepsilon}(\varepsilon), \quad j, k \in \mathbb{Z}. \tag{2.6.3}$$

Given a bounded $S \subset \mathbb{C}$, we say

$$\mathrm{Cont}^+(S) \le M\varepsilon^2 \tag{2.6.4}$$

if S is contained in a union of M squares of the form (2.6.3), and

$$\mathrm{Cont}^-(S) \ge N\varepsilon^2 \tag{2.6.5}$$

if S contains a union of N squares of the form (2.6.3).

We then set $\mathrm{Cont}^+(S)$ to be the infimum of the quantities arising in (2.6.4) and we set $\mathrm{Cont}^-(S)$ to be the supremum of the quantities arising in (2.6.5), as ε runs over $(0, 1]$. Clearly

$$\mathrm{Cont}^-(S) \le \mathrm{Cont}^+(S). \tag{2.6.6}$$

If these two quantities coincide, we say S is *contented*, and write

$$\operatorname{Area} S = \mathrm{Cont}^+(S) = \mathrm{Cont}^-(S). \tag{2.6.7}$$

One verifies that if S is a standard square with sidelength a, then the definition via (2.6.4)–(2.6.7) agrees with (2.6.2). Another easy check is that if R is a standard rectangle of sidelengths a and b, of the form

$$R = R_{z_0}(a, b) = \{x + iy : x_0 \le x \le x_0 + a, y_0 \le y \le y_0 + b\}, \tag{2.6.8}$$

then

$$\operatorname{Area} R = ab. \tag{2.6.9}$$

Here are some useful properties of the concepts defined above. Given a set S, let $\overline{S}$ denote the closure of S, consisting of all points $p \in \mathbb{C}$ such that for some $q_k \in S$, $q_k \to p$, and let $\overset{\circ}{S}$ denote the interior of S, consisting of points $p \in S$ such that for some $\varepsilon > 0$,

$$z \in \mathbb{C},\ |z - p| < \varepsilon \Longrightarrow z \in S. \tag{2.6.10}$$

Then

$$\mathrm{Cont}^+(S) = \mathrm{Cont}^+(\overline{S}), \quad \mathrm{Cont}^-(S) = \mathrm{Cont}^-(\overset{\circ}{S}). \tag{2.6.11}$$

Also, if R is a rectangle, as in (2.6.8), and

$$S \subset R, \quad T = R \setminus S, \tag{2.6.12}$$

then

$$\mathrm{Cont}^+(S) + \mathrm{Cont}^-(T) = \mathrm{Area}\, R. \tag{2.6.13}$$

Similarly $\mathrm{Cont}^-(S) + \mathrm{Cont}^+(T) = \mathrm{Area}\, R$, so

$$S \text{ contented} \Rightarrow T \text{ contented and } \ \mathrm{Area}\, S + \mathrm{Area}\, T = \mathrm{Area}\, R. \tag{2.6.14}$$

Another useful observation is that, if $S \subset \mathbb{C}$ is a bounded set,

$$\mathrm{Cont}^-(S) + \mathrm{Cont}^+(bS) \geq \mathrm{Cont}^+(S), \tag{2.6.15}$$

where $bS = \overline{S} \setminus \overset{\circ}{S}$. In particular,

$$\mathrm{Cont}^+(bS) = 0 \Longrightarrow S \text{ is contented.} \tag{2.6.16}$$

And another:

Proposition 2.6.1. *Given $S, T \subset \mathbb{C}$ bounded and contented,*

$$\mathrm{Cont}^+(S \cap T) = 0 \Longrightarrow \mathrm{Area}(S \cup T) = \mathrm{Area}\, S + \mathrm{Area}\, T. \tag{2.6.17}$$

We say a bounded set $S \subset \mathbb{C}$ is a nil set if $\mathrm{Cont}^+(S) = 0$. The union of any two nil sets is a nil set. This follows from the fact that, if S and $T \subset \mathbb{C}$ are bounded,

$$\mathrm{Cont}^+(S \cup T) \leq \mathrm{Cont}^+(S) + \mathrm{Cont}^+(T). \tag{2.6.18}$$

Here is a preliminary result on invariance of area under isometries.

Proposition 2.6.2. *Let $F(z) = az + b$ or $a\overline{z} + b$, and let $S \subset \mathbb{C}$ be a bounded set. Then*

$$\mathrm{Cont}^\pm(F(S)) = \mathrm{Cont}^\pm(S), \tag{2.6.19}$$

provided

$$a \in \{\pm 1, \pm i\}. \tag{2.6.20}$$

Proof. In case $b = 0$, such F preserves the tiling described in (2.6.3). It remains to treat $F(z) = z + b$, which we leave to the reader. □

We move to another class of figures for which we can compute areas. We call a triangle $\mathcal{T} = \triangle(A, B, C)$ a standard right triangle if it has a right angle at C and the sides meeting at C are parallel to the real and imaginary axes, i.e., $A - C \in i\mathbb{R}$, $B - C \in \mathbb{R}$.

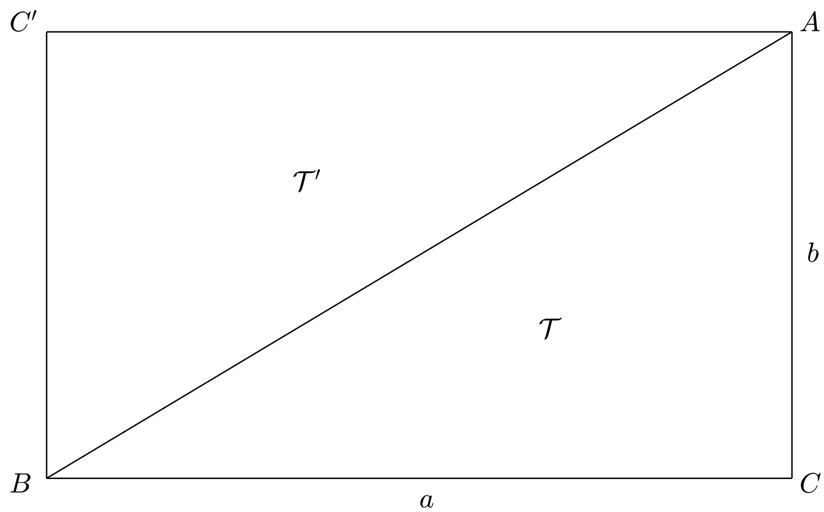

Figure 2.6.1. Finding the area of a standard right triangle

Proposition 2.6.3. *Let $\mathcal{T} = \triangle(A, B, C)$ be a standard right triangle, with sidelengths*

$$a = |B - C|, \quad b = |A - C|. \tag{2.6.21}$$

Then

$$\operatorname{Area} \mathcal{T} = \frac{1}{2} ab. \tag{2.6.22}$$

Proof. The triangle $\mathcal{T}$ has a twin, $\mathcal{T}' = \triangle(A, B, C')$, where

$$C' = C + (A - C) + (B - C). \tag{2.6.23}$$

See Figure 2.6.1. These two triangles are related by an isometry $F : \mathcal{T} \to \mathcal{T}'$ of the form considered in Proposition 2.6.2, with $a = -1$. Hence $\operatorname{Area} \mathcal{T}' = \operatorname{Area} \mathcal{T}$, and, by Proposition 2.6.1, $\operatorname{Area} \mathcal{T} + \operatorname{Area} \mathcal{T}'$ equals the area of a rectangle R with sidelengths a and b. This gives (2.6.22). $\square$

As a preliminary to extending Proposition 2.6.2, we have the following.

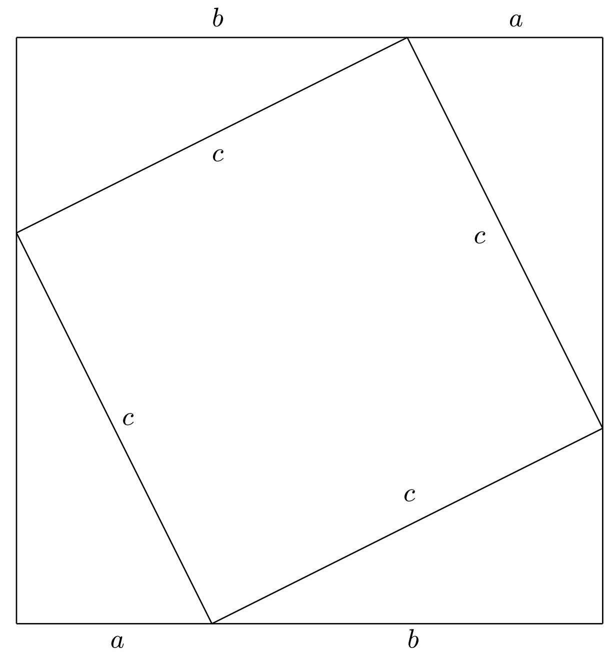

Figure 2.6.2. Area of an arbitrary square

Lemma 2.6.4. *If S_c is an arbitrary square of sidelength c (not necessarily standard), then*

$$\text{Area}\, S_c = c^2. \tag{2.6.24}$$

Proof. Translating S_c, we can assume S_c is obtained by applying a rotation to the standard square $S'_c = \{x + iy : |x|, |y| \le c/2\}$. If S_c itself is not standard, consider the horizontal lines through the top and bottom vertices of S_c, and the vertical lines through the left and right vertices of S_c. They form a rectangle R, in which S_c is inscribed. See Figure 2.6.2. This figure is invariant under the rotation $z \mapsto iz$, so R is in fact a standard square of sidelength $a + b$. R is divided into five figures; one is S_c, and the other four are standard right triangles, of sidelengths a and b. Therefore we have

$$\text{Area}\, R = \text{Area}\, S_c + 4 \cdot \frac{ab}{2}, \tag{2.6.25}$$

hence

$$(a + b)^2 = \text{Area}\, S_c + 2ab, \tag{2.6.26}$$

hence

$$a^2 + b^2 = \operatorname{Area} S_c. \tag{2.6.27}$$

Now the Pythagorean theorem gives $a^2 + b^2 = c^2$, so we have (2.6.24). $\square$

Now for our extension of Proposition 2.6.2.

Proposition 2.6.5. *If $S \subset \mathbb{C}$ is bounded, then*

$$\operatorname{Cont}^{\pm}(F(S)) = \operatorname{Cont}^{\pm}(S), \quad \forall F \in \operatorname{Isom}(\mathbb{C}). \tag{2.6.28}$$

Proof. It suffices to establish this when $F(z) = az$, $|a| = 1$, i.e., F is a rotation. In such a case, if we have a tiling $\{S_{jk\varepsilon} : j, k \in \mathbb{Z}\}$ of the form (2.6.3) and S is contained in the union of M of the squares in this tiling, and S contains N of the squares in this tiling, so (2.6.3) and (2.6.4) hold, it follows that $F(S)$ is contained in the union of M squares of the rotated tiling and contains N squares of this rotated tiling. Thanks to Lemma 2.6.4 (and also Proposition 2.6.1), we also have

$$\operatorname{Cont}^{+}(F(S)) \le M\varepsilon^2, \quad \operatorname{Cont}^{-}(F(S)) \ge N\varepsilon^2. \tag{2.6.29}$$

In the limit we get

$$\operatorname{Cont}^{+}(F(S)) \le \operatorname{Cont}^{+}(S), \quad \operatorname{Cont}^{-}(F(S)) \ge \operatorname{Cont}^{-}(S), \tag{2.6.30}$$

for all bounded S and $F \in \operatorname{Isom}(\mathbb{C})$. Replacing S by $F(S)$ and F by F^{-1} gives the converse inequalities, and we have (2.6.28). $\square$

Remark. Figure 2.6.2 plays a role in some popular proofs of the Pythagorean theorem, in which the flow of the argument is reversed, and one works under the hypothesis that (2.6.24) holds.

As a related matter, we mention that the proof of the Pythagorean theorem in Book I of Euclid uses as a key ingredient the proposition that congruent triangles have equal area. For us, the order of arguments is different. The Pythagorean theorem comes first.

We now bring calculus into the study of area, with the following key result.

Proposition 2.6.6. *Let $\varphi, \psi : [a, b] \to \mathbb{R}$ be continuous, and assume $\varphi \le \psi$. Set*

$$\mathcal{O} = \{z = x + iy : a \le x \le b, \varphi(x) \le y \le \psi(x)\}. \tag{2.6.31}$$

Then $\mathcal{O}$ is contented, and

$$\operatorname{Area} \mathcal{O} = \int_a^b [\psi(x) - \varphi(x)]\, dx. \tag{2.6.32}$$

Proof. Translating, we can assume $a = 0$. For $n \in \mathbb{N}$, tile $\mathbb{C}$ as in (2.6.3) with tiles of edgelength $\varepsilon = (b-a)/n$. Then considering collections of tiles that are contained in $\mathcal{O}$ but, with the addition of a minimum number of extra tiles, contain $\mathcal{O}$, we obtain two-sided bounds of the form

$$N\varepsilon^2 \le \text{Cont}^-(\mathcal{O}) \le \text{Cont}^+(\mathcal{O}) \le M\varepsilon^2, \tag{2.6.33}$$

where $N\varepsilon^2$ and $M\varepsilon^2$ are, for ε sufficiently small, close to the Riemann sums of the form (2.1.2), which approximate

$$\int_a^b g(x)\, dx, \quad g(x) = \psi(x) - \varphi(x). \tag{2.6.34}$$

Passing to the limit gives (2.6.32) □

We can present the unit disk

$$D_1 = \{z \in \mathbb{C} : |z| \le 1\} \tag{2.6.35}$$

in the form (2.6.31), with $a = -1$, $b = 1$,

$$\varphi(x) = -\sqrt{1-x^2}, \quad \psi(x) = \sqrt{1-x^2}, \tag{2.6.36}$$

to obtain

$$\text{Area}\, D_1 = 2\int_{-1}^{1} \sqrt{1-x^2}\, dx. \tag{2.6.37}$$

To evaluate this integral, we can make a change of variable, $x = \sin t$, $-\pi/2 \le t \le \pi/2$, giving $dx = \cos t\, dt$, hence, by Proposition 2.1.5,

$$\text{Area}\, D_1 = 2\int_{-\pi/2}^{\pi/2} \cos^2 t\, dt. \tag{2.6.38}$$

Using

$$\cos 2t = \cos^2 t - \sin^2 t = 2\cos^2 t - 1, \tag{2.6.39}$$

which follows from (2.3.37), we get

$$\text{Area}\, D_1 = \int_{-\pi/2}^{\pi/2} (\cos 2t + 1)\, dt. \tag{2.6.40}$$

Furthermore, via (2.2.32),

$$\int_{-\pi/2}^{\pi/2} \cos 2t\, dt = \frac{\sin 2t}{2}\Big|_{-\pi/2}^{\pi/2} = 0. \tag{2.6.41}$$

Hence our conclusion:

$$\text{Area}\, D_1 = \pi. \tag{2.6.42}$$

We aim to extend this result to a computation of the area of a "pie slice,"

$$\Pi_t = \{re^{is} : 0 \le r \le 1,\ 0 \le s \le t\}, \tag{2.6.43}$$

given $t \in [0, 2\pi]$. Here is the result.

Proposition 2.6.7. *For* Π_t *given by* (2.6.43), $0 \le t \le 2\pi$,

$$\operatorname{Area} \Pi_t = \frac{t}{2}. \tag{2.6.44}$$

Proof. Set

$$A(t) = \operatorname{Area} \Pi_t. \tag{2.6.45}$$

We have

$$A(s+t) = A(s) + A(t), \tag{2.6.46}$$

for $s, t \ge 0,\ s + t \le 2\pi$, as a consequence of Proposition 2.6.5. Hence it suffices to prove (2.6.44) for $t \in (0, \pi/2)$. Following Figure 2.6.3, consider the region

$$\mathcal{O}_\xi = \{x + iy : 0 \le x \le \xi,\ 0 \le y \le \sqrt{1 - x^2}\}, \quad \xi = \sin t. \tag{2.6.47}$$

The ray from 0 to $e^{it} \in S^1$ divides $\mathcal{O}_\xi$ into two parts, one a "pie slice," obtained from Π_t via a rotation, which, by Proposition 2.6.5 has the same area as Π_t. The other is a standard right triangle, with sidelengths $\sin t$ and $\cos t$. By Proposition 2.6.6,

$$\operatorname{Area} \mathcal{O}_\xi = B(\xi) = \int_0^\xi \sqrt{1 - x^2}\, dx, \quad B'(\xi) = \sqrt{1 - \xi^2}. \tag{2.6.48}$$

The division of $\mathcal{O}_\xi$ described above implies

$$B(\sin t) = A(t) + \frac{1}{2} \cos t \sin t. \tag{2.6.49}$$

Applying d/dt gives

$$\begin{aligned} A'(t) &= B'(\sin t) \cos t - \frac{1}{2}(\cos^2 t - \sin^2 t) \\ &= \frac{1}{2} \cos^2 t + \frac{1}{2} \sin^2 t \\ &= \frac{1}{2}. \end{aligned} \tag{2.6.50}$$

Since $A(0) = 0$, we have (2.6.44) □

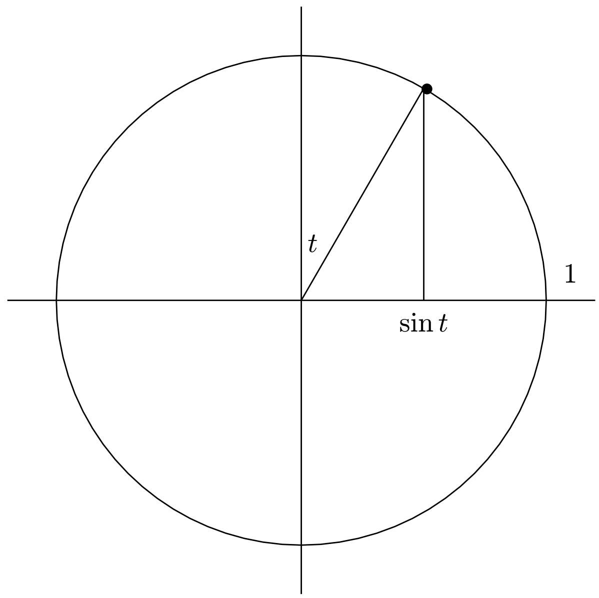

Figure 2.6.3. Area under a circle

2.7. Heron's formula

Let $\mathcal{T} \subset \mathbb{C}$ be a triangle, with vertices A, B, C, having opposing sides of length a, b, c, respectively. Set

$$s = \frac{1}{2}(a + b + c). \tag{2.7.1}$$

Heron's formula for the area of $\mathcal{T}$ is the following

$$(\operatorname{Area} \mathcal{T})^2 = s(s - a)(s - b)(s - c). \tag{2.7.2}$$

To set up a proof, recall from Proposition 1.5.4 that at most one angle of $\mathcal{T}$ is obtuse. We can assume A and B are acute. Drop a perpendicular from C to the line segment from A to B, dividing this segment into two pieces, of length x and y, so

$$x + y = c. \tag{2.7.3}$$

See Figure 2.7.1. Let h denote the length of this perpendicular line segment. Hence

$$\operatorname{Area} \mathcal{T} = \frac{1}{2}ch. \tag{2.7.4}$$

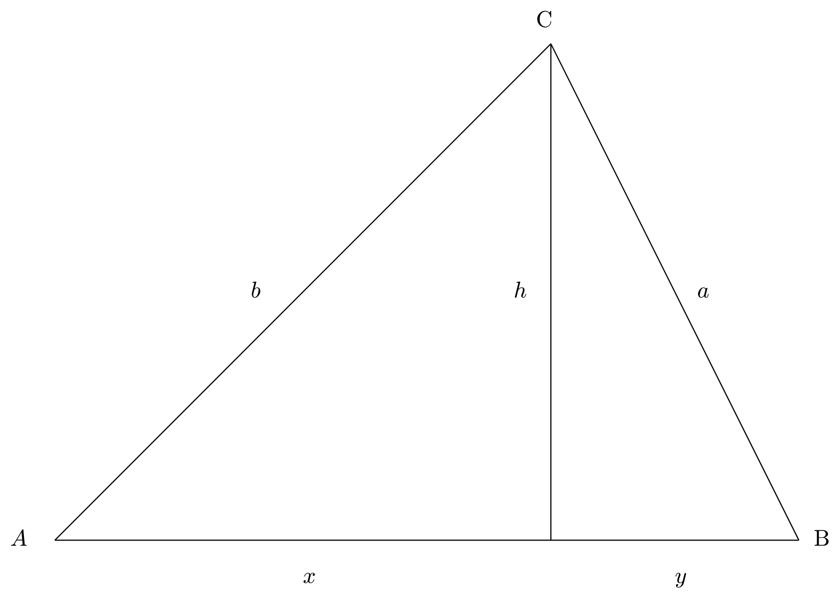

Figure 2.7.1. Setup for computing Area $\mathcal{T}$

Meanwhile, the Pythagorean theorem yields

$$h^2 + x^2 = b^2, \quad h^2 + y^2 = a^2. \tag{2.7.5}$$

We aim to deduce (2.7.2) from (2.7.3)–(2.7.5).

Subtracting in (2.7.5) gives $x^2 - y^2 = b^2 - a^2$, and since

$$x^2 - y^2 = (x+y)(x-y) = c(x-y),$$

we have

$$x - y = \frac{b^2 - a^2}{c}. \tag{2.7.6}$$

Adding (2.7.3) and (2.7.6) gives

$$x = \frac{c^2 + b^2 - a^2}{2c}, \tag{2.7.7}$$

and hence from (2.7.5) we get

$$h^2 = b^2 - x^2 = b^2 - \frac{(c^2 + b^2 - a^2)^2}{4c^2}. \tag{2.7.8}$$

Therefore,

$$\begin{aligned} 4c^2h^2 &= (2bc)^2 - (c^2 + b^2 - a^2)^2 \\ &= (2bc + c^2 + b^2 - a^2)(2bc - c^2 - b^2 + a^2) \\ &= [(b+c)^2 - a^2] \cdot [a^2 - (b-c)^2] \\ &= (a+b+c)(-a+b+c)(a+b-c)(a-b+c) \\ &= 16s(s-a)(s-b)(s-c). \end{aligned} \tag{2.7.9}$$

By (2.7.4),

$$(\text{Area}\, \mathcal{T})^2 = \frac{c^2h^2}{4}, \tag{2.7.10}$$

and we have the desired identity (2.7.2).

2.8. Euclidean numbers

Here we describe collections of numbers (in $\mathbb{C}$), lines, and circles that we call Euclidean, denoted $\mathbb{E}$, $\mathcal{E}_L$, and $\mathcal{E}_C$, respectively, which can be constructed using compass and straightedge. We start with some rules for obtaining objects in these classes.

(A) $0, 1 \in \mathbb{E}$,

(B) If $a, b \in \mathbb{E},\ a \neq b$, the line ℓ_{ab} through a and b belongs to $\mathcal{E}_L$. In particular,

$$\mathbb{R} \in \mathcal{E}_L.$$

(C) If $a, b \in \mathbb{E},\ a \neq b$, the circle $C_{a,b}$ centered at a, passing through b, belongs to $\mathcal{E}_C$.

(D) If $a \in \mathbb{E},\ r > 0,\ r \in \mathbb{E}$, then $S_r(a)$, the circle with center a, radius r, belongs to $\mathcal{E}_C$. In particular,

$$S^1 \in \mathcal{E}_C.$$

Rules (B)–(D) state how to produce elements of $\mathcal{E}_L$ and $\mathcal{E}_C$, given elements of $\mathbb{E}$. We also have the following rules for generating elements of $\mathbb{E}$, given elements of $\mathcal{E}_L$ and $\mathcal{E}_C$.

(a) If $\ell, \ell' \in \mathcal{E}_L$, then $\ell \cap \ell' \subset \mathbb{E}$.

(b) If $C, C' \in \mathcal{E}_C$, then $C \cap C' \subset \mathbb{E}$.

(c) If $\ell \in \mathcal{E}_L,\ C \in \mathcal{E}_C$, then $\ell \cap C \subset \mathbb{E}$.

We define $\mathbb{E}$, $\mathcal{E}_L$, and $\mathcal{E}_C$ to be the smallest collections of numbers, lines, and circles satisfying (A)–(D) and (a)–(c). We now describe other elements of $\mathbb{E}$, $\mathcal{E}_L$, and $\mathcal{E}_C$, applying these rules. We leave the demonstrations of these results as exercises, with some hints.

1. $\mathbb{Z} \subset \mathbb{E}$. Start with (A) and (B), and use (D) and (c) repeatedly.

2. If $\ell \in \mathcal{E}_L$ and $p \in \mathbb{E}$, then $\ell_p^{\perp}$, the line through p, perpendicular to ℓ, lies in $\mathcal{E}_L$.

Start with a circle $S_r(p)$, with $r \in \mathbb{N}$ large enough that $S_r(p) \cap \ell$ consists of two points, say a, b. Then p is the midpoint of a and b. Take circles $S_R(a)$

and $S_R(b)$, with $R \in \mathbb{N},\ R > |a-b|$. Then $\ell_p^\perp$ is the line through the two points in $S_R(a) \cap S_R(b)$.
Consequences:

$$i\mathbb{R} \in \mathcal{E}_L, \quad i \in \mathbb{E}, \quad ki \in \mathbb{E},\ \forall\, k \in \mathbb{Z}.$$

3. If $\ell \in \mathcal{E}_L,\ q \in \mathbb{E},\ q \notin \ell$, and ℓ'_q is the line through q parallel to ℓ, then $\ell'_q \in \mathcal{E}_L$. Indeed, take $\ell'_q = (\ell_q^\perp)_q^\perp$.

4. $a \in \mathbb{E} \Rightarrow \operatorname{Re} a,\ \operatorname{Im} a,\ |a| \in \mathbb{E}$. Drop perpendiculars from a to $\mathbb{R}$ and to $i\mathbb{R}$. Also, C_{0a} intersects $\mathbb{R}$ at $\pm|a|$.

5. If $x, y \in \mathbb{R} \cap \mathbb{E}$, then $-x, x+y, |y| \in \mathbb{R} \cap \mathbb{E}$. Furthermore, $x+iy \in \mathbb{E}$.

For the latter part, we have $\ell_x^\perp \in \mathcal{E}_L$, and $x+iy \in \ell_x^\perp \cap S_{|y|}(x)$.

6. If $a, b \in \mathbb{E}$, then $a+b \in \mathbb{E}$. Also $ia \in \mathbb{E}$. Use Exercises 4–5. Consequence:

$$\mathbb{Z}[i] \subset \mathbb{E}.$$

7. Assume $a, b \in \mathbb{E},\ a, b > 0$. Then

$$ab,\ \frac{1}{a} \in \mathbb{E}.$$

The line through 0 and $1+ib$ intersects $\ell_a^\perp$ at $a+iab = z$, and $ab = \operatorname{Re} z$. The line through 0 and $a+i$ intersects $\ell_1^\perp$ at $1+i(1/a) = w$, and $1/a = \operatorname{Im} w$. Corollary:

$$\mathbb{Q}[i] \subset \mathbb{E}.$$

8. More generally,

$$a, b \in \mathbb{E} \Longrightarrow ab,\ \frac{a}{b} \in \mathbb{E},$$

the latter holding provided $b \neq 0$.

For ab, express the product in terms of $\operatorname{Re} a, b$ and $\operatorname{Im} a, b$. Next, write $a/b = a\bar{b}/|b|^2$ and apply Exercise 7 to get $1/|b|^2 \in \mathbb{E}$.
Remark. Exercises 6 and 8 imply $\mathbb{E}$ is a field.

We next consider square roots.

9. Given $\omega, \alpha \in S^1,\ \alpha^2 = \omega,\ \omega \in \mathbb{E} \Longrightarrow \alpha \in \mathbb{E}$.

Indeed, if $\omega \neq -1$, we have

$$\alpha = \pm\frac{\omega+1}{|\omega+1|}.$$

10. If $x \in \mathbb{E},\ x > 0$, then $\sqrt{x} \in \mathbb{E}$.

For the geometrical construction yielding this result, see §1.10.

11. Given $z \in \mathbb{E}$, if $w \in \mathbb{C}$ and $w^2 = z$, then $w \in \mathbb{E}$.

Indeed, if $z \neq 0$, set

$$z = r\omega, \quad r = |z|, \ \omega = \frac{z}{|z|}.$$

Then $r, \omega \in \mathbb{E}$ and

$$w = \pm\sqrt{r}\alpha,$$

with α as in Exercise 9.

12. Let $p(z) = a_2 z^2 + a_1 z + a_0$, $a_j \in \mathbb{E}$, $a_2 \neq 0$. Then its roots, i.e., the solutions to $p(z) = 0$, belong to $\mathbb{E}$.

Use the quadratic formula.

13. $\mathbb{E}$ is the smallest subfield of $\mathbb{C}$ that is closed under taking square roots.

Results stated above give half of this. For the other half, one needs that elements of $\mathbb{E}$ produced in (b)–(c) arise as solutions to quadratic equations.

14. $\mathbb{E}$ is countable. Each application of (B)–(D) produces one element of $\mathcal{E}_L$ or one element of $\mathcal{E}_C$. Each application of (a)–(c) produces one or two elements of $\mathbb{E}$.

15. We say an element of $\mathrm{Isom}(\mathbb{C})$, resp. $\mathrm{Isom}^{\pm}(\mathbb{C})$ belongs to $\mathrm{Isom}(\mathbb{E})$, resp. $\mathrm{Isom}^{\pm}(\mathbb{E})$, provided $F : \mathbb{E} \to \mathbb{E}$. Then one has $F \in \mathrm{Isom}^{+}(\mathbb{E})$ if and only if

$$F(z) = az + p, \quad a \in S^1 \cap \mathbb{E}, \ p \in \mathbb{E},$$

and $F \in \mathrm{Isom}^{-}(\mathbb{E})$ if and only if

$$F(z) = a\overline{z} + p, \quad a \in S^1 \cap \mathbb{E}, \ p \in \mathbb{E}.$$

2.9. Linear fractional transformations

Here we consider an important extension of the family of maps $z \mapsto az + b$, making up $\mathrm{Isom}^+(\mathbb{C})$ and more generally $\mathrm{Sim}^+(\mathbb{C})$. This larger class consists of linear fractional transformations. They have the form

$$L_A(z) = \frac{az+b}{cz+d}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.9.1}$$

Here A is an invertible 2×2 matrix; we say $A \in G\ell(2,\mathbb{C})$. We do not want the denominator in this fraction to be $\equiv 0$, and we do not want the numerator to be a constant multiple of the denominator. This leads to the requirement that $\det A \neq 0$. If $c \neq 0$, L_A is well defined on $\mathbb{C} \setminus \{-d/c\}$. We extend L_A to

$$L_A : \widehat{\mathbb{C}} \longrightarrow \widehat{\mathbb{C}}, \quad \widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}, \tag{2.9.2}$$

by setting

$$L_A(-d/c) = \infty, \quad \text{if } \ c \neq 0. \tag{2.9.3}$$

and

$$\begin{aligned} L_A(\infty) = \frac{a}{c} \ &\text{ if } \ c \neq 0, \\ \infty \ &\text{ if } \ c = 0. \end{aligned} \tag{2.9.4}$$

If also $B \in G\ell(2,\mathbb{C})$, a calculation gives

$$L_A \circ L_B = L_{AB}. \tag{2.9.5}$$

In particular L_A is bijective on $\widehat{\mathbb{C}}$, with inverse $L_{A^{-1}}$.

The set $G\ell(2,\mathbb{C})$ is a group, i.e., a nonempty set of matrices G with the property that

$$A, B \in G \Longrightarrow AB, \ A^{-1} \in G. \tag{2.9.6}$$

The result (2.9.5) says the map $A \mapsto L_A$ is a group homomorphism.

Note that $L_{sA} = L_A$ for each nonzero $s \in \mathbb{C}$. In particular, $L_A = L_{A_1}$ for some A_1 of determinant 1; we say $A_1 \in S\ell(2,\mathbb{C})$. Given $A_j \in S\ell(2,\mathbb{C})$, $L_{A_1} = L_{A_2}$ if and only if $A_1 = \pm A_2$.

Note that if a, b, c, d are all real, then L_A in (2.9.1) preserves $\mathbb{R} \cup \{\infty\}$. In this case we have $A \in G\ell(2,\mathbb{R})$. We still have $L_{sA} = L_A$ for all nonzero s, but we need $s \in \mathbb{R}$ to get $sA \in G\ell(2,\mathbb{R})$. We can write $L_A = L_{A_1}$ for $A_1 \in S\ell(2,\mathbb{R})$ if $A \in G\ell(2,\mathbb{R})$ and $\det A > 0$. We can also verify that

$$A \in S\ell(2,\mathbb{R}) \Longrightarrow L_A : \mathcal{U} \to \mathcal{U}, \tag{2.9.7}$$

where

$$\mathcal{U} = \{z : \operatorname{Im} z > 0\} \tag{2.9.8}$$

is the upper half-plane. In more detail, for $a, b, c, d \in \mathbb{R},\ z = x + iy$,

$$\frac{az+b}{cz+d} = \frac{(az+b)(c\overline{z}+d)}{(cz+d)(c\overline{z}+d)} = \frac{R}{P} + iy\frac{ad-bc}{P}, \tag{2.9.9}$$

with

$$R = ac|z|^2 + bd + (ad+bc)x \in \mathbb{R}, \quad P = |cz+d|^2 > 0, \text{ if } y \neq 0, \tag{2.9.10}$$

which gives (2.9.7).

We now single out for attention the following linear fractional transformation:

$$\varphi(z) = \frac{z-i}{z+i}, \quad \varphi(z) = L_{A_0}(z),\ A_0 = \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}. \tag{2.9.11}$$

Note that

$$\varphi(x+iy) = \frac{x+i(y-1)}{x+i(y+1)} \Longrightarrow |\varphi(x+iy)|^2 = \frac{x^2+(y-1)^2}{x^2+(y+1)^2}. \tag{2.9.12}$$

In particular, $|\varphi(x+iy)| < 1$ if and only if $y > 0$. We have

$$\varphi : \mathcal{U} \to \mathcal{D}, \quad \varphi : \mathbb{R} \cup \{\infty\} \to S^1 = \partial\mathcal{D}, \tag{2.9.13}$$

where

$$\mathcal{D} = \{z : |z| < 1\} \tag{2.9.14}$$

is the unit disk. The bijectivity of φ on $\widehat{\mathbb{C}}$ implies that φ is bijective in (2.9.13).

Conjugating the $S\ell(2,\mathbb{R})$ action on $\mathcal{U}$ by φ yields the mappings

$$M_A = L_{A_0 A A_0^{-1}} : \mathcal{D} \longrightarrow \mathcal{D}. \tag{2.9.15}$$

In detail, if A is as in (2.9.1), with a, b, c, d real, and if A_0 is as in (2.9.11),

$$\begin{aligned} A_0 A A_0^{-1} &= \frac{1}{2i}\begin{pmatrix} (a+d)i - b + c & (a-d)i + b + c \\ (a-d)i - b - c & (a+d)i + b - c \end{pmatrix} \\ &= \begin{pmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix}. \end{aligned} \tag{2.9.16}$$

Note that

$$|\alpha|^2 - |\beta|^2 = \det A = ad - bc. \tag{2.9.17}$$

It follows that

$$\begin{aligned} A_0 S\ell(2,\mathbb{R}) A_0^{-1} &= \left\{ \begin{pmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix} \in G\ell(2,\mathbb{C}) : |\alpha|^2 - |\beta|^2 = 1 \right\} \\ &= SU(1,1), \end{aligned} \tag{2.9.18}$$

the latter identity defining the group $SU(1,1)$. Hence we have linear fractional transformations

$$L_B : \mathcal{D} \to \mathcal{D}, \quad L_B(z) = \frac{\alpha z + \beta}{\overline{\beta} z + \overline{\alpha}}, \tag{2.9.19}$$
$$B = \begin{pmatrix} \alpha & \beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix} \in SU(1,1), \qquad |\alpha|^2 - |\beta|^2 = 1.$$

Note that for such B as in (2.9.19),

$$L_B(e^{i\theta}) = \frac{\alpha e^{i\theta} + \beta}{\overline{\beta} e^{i\theta} + \overline{\alpha}} = e^{i\theta} \frac{\alpha + \beta e^{-i\theta}}{\overline{\alpha} + \overline{\beta} e^{i\theta}}, \tag{2.9.20}$$

and in the last fraction the numerator is the complex conjugate of the denominator. This directly implies the result $L_B : S^1 \to S^1$ for such B.

We have the following important transitivity properties.

Proposition 2.9.1. *Given the groups $S\ell(2,\mathbb{R})$ and $SU(1,1)$ defined above,*

$$S\ell(2,\mathbb{R}) \text{ acts transitively on } \mathcal{U}, \text{ via } (2.9.1), \tag{2.9.21}$$

and

$$SU(1,1) \text{ acts transitively on } \mathcal{D}, \text{ via } (2.9.19). \tag{2.9.22}$$

Proof. To demonstrate (2.9.21), take $p = a + ib \in \mathcal{U}$ $(a \in \mathbb{R},\ b > 0)$. Then $L_p(z) = bz + a = (b^{1/2}z + b^{-1/2}a)/b^{-1/2}$ maps i to p. Given another $q \in \mathcal{U}$, we see that $L_p L_q^{-1}$ maps q to p, so (2.9.21) holds. The conjugation (2.9.18) implies that (2.9.21) and (2.9.22) are equivalent.

We can also demonstrate (2.9.22) directly. Given $p \in \mathcal{D}$, we can pick $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 - |\beta|^2 = 1$ and $\beta/\overline{\alpha} = p$. Then $L_B(0) = p$, so (2.9.22) holds. ☐

We list some building blocks for the group of linear fractional transformations, namely (with $a \neq 0$)

$$\delta_a(z) = az, \quad \tau_b(z) = z + b, \quad \iota(z) = \frac{1}{z}. \tag{2.9.23}$$

We call these respectively (complex) dilations, translations, and inversion about the unit circle $S^1 = \{z : |z| = 1\}$. These have the form (2.9.1) with A given respectively by

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.9.24}$$

We can produce the inversion ι_D about the boundary of a disk $D = D_r(p)$ as

$$\iota_D = \tau_p \circ \delta_r \circ \iota \circ \delta_{1/r} \circ \tau_{-p}. \tag{2.9.25}$$

The reader can work out the explicit form of this linear fractional transformation. Note that ι_D leaves ∂D invariant and interchanges p and ∞.

Recall that the linear fractional transformation φ in (2.9.11) was seen to map $\mathbb{R} \cup \{\infty\}$ to S^1. Similarly its inverse, given by $-i\psi(z)$ with

$$\psi(z) = \frac{z+1}{z-1}, \tag{2.9.26}$$

maps S^1 to $\mathbb{R} \cup \{\infty\}$; equivalently ψ maps S^1 to $i\mathbb{R} \cup \{\infty\}$. To see this directly, write

$$\psi(e^{i\theta}) = \frac{e^{i\theta}+1}{e^{i\theta}-1} = \frac{-i}{\tan \theta/2}. \tag{2.9.27}$$

These are special cases of an important general property of linear fractional transformations. To state it, let us say that en extended line is a set $\ell \cup \{\infty\}$, where ℓ is a line in $\mathbb{C}$.

Proposition 2.9.2. *If L is a linear fractional transformation, then L maps each circle to a circle or an extended line, and L maps each extended line to a circle or an extended line.*

To begin the proof, suppose $D \subset \mathbb{C}$ is a disk. We investigate where L maps ∂D.

Claim 1. If L has a pole at $p \in \partial D$ (i.e., if $L(p) = \infty$), then L maps ∂D to an extended line.

Proof. Making use of the transformations (2.9.23), we have $L(\partial D) = L'(S^1)$ for some linear fractional transformation L', so we need to check only the case $D = \{z : |z| < 1\}$, with L having a pole on S^1, and indeed we can take the pole to be at $z = 1$. Thus we look at

$$\begin{aligned} L(e^{i\theta}) &= \frac{ae^{i\theta}+b}{e^{i\theta}-1} \\ &= -\frac{a+b}{2}\,\frac{1}{\tan\theta/2} + \frac{a-b}{2}, \end{aligned} \tag{2.9.28}$$

whose image is clearly an extended line.

Claim 2. If L has no pole on ∂D, then L maps ∂D to a circle.

Proof. One possibility is that L has no pole in $\mathbb{C}$. Then $c = 0$ in (2.9.1). This case is elementary.

Next, suppose L has a pole at $p \in D$. Composing (on the right) with various linear fractional transformations, we can reduce to the case $D = \{z : |z| < 1\}$, and making further compositions (via Proposition 2.9.1), we need only deal with the case $p = 0$. So we are looking at

$$L(z) = \frac{az+b}{z}, \quad L(e^{i\theta}) = a + be^{-i\theta}. \tag{2.9.29}$$

Clearly the image $L(S^1)$ is a circle.

If L has a pole at $p \in \mathbb{C} \setminus \overline{D}$, we can use an inversion about ∂D to reduce the study to that done in the previous paragraph. This finishes Claim 2.

To finish the proof of Proposition 2.9.2, there are two more claims to establish.

Claim 3. If $\ell \subset \mathbb{C}$ is a line and L has a pole on ℓ, or if L has no pole in $\mathbb{C}$, then L maps $\ell \cup \{\infty\}$ to an extended line.

Claim 4. If $\ell \subset \mathbb{C}$ is a line and L has a pole in $\mathbb{C} \setminus \ell$, then L maps $\ell \cup \{\infty\}$ to a circle.

We leave Claims 3–4 as exercises for the reader.

We now take a look at the group $G_{\mathcal{HU}}$ of linear fractional transformations that preserve both the upper half plane $\mathcal{U}$ and the unit disk $\mathcal{D}$. These have the form

$$L_A, \quad A \in S\ell(2,\mathbb{R}) \cap SU(1,1), \tag{2.9.30}$$

hence

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad a, b \in \mathbb{R},\ a^2 - b^2 = 1. \tag{2.9.31}$$

Since $L_A = L_{-A}$, we can take $a > 0$. Note that

$$A^{-1} = \begin{pmatrix} a & -b \\ -b & a \end{pmatrix}. \tag{2.9.32}$$

It is convenient to parametrize the group $G_{\mathcal{HU}}$ by setting

$$a = \cosh t, \quad b = \sinh t, \quad t \in \mathbb{R}, \tag{2.9.33}$$

where we bring in the hyperbolic functions

$$\cosh t = \frac{1}{2}(e^t + e^{-t}), \quad \sinh t = \frac{1}{2}(e^t - e^{-t}). \tag{2.9.34}$$

A computation gives $\cosh^2 t - \sinh^2 t = 1$. In such a case, we can write A in (2.9.31) as

$$H(t) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} = (\cosh t)I + (\sinh t)K, \tag{2.9.35}$$

at least for $a > 0$, where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.9.36}$$

As usual, I denotes the identity matrix.

The casual resemblance of the right side of (2.9.35) to the right side of Euler's formula (2.3.31) is worth pursuing. Let's start with a derivative formula for $H(t)$. By (2.3.8), we have the derivative formulas

$$\frac{d}{dt}\cosh t = \sinh t, \quad \frac{d}{dt}\sinh t = \cosh t. \tag{2.9.37}$$

Applying this to (2.9.35) gives

$$H'(t) = (\sinh t)I + (\cosh t)K = KH(t), \tag{2.9.38}$$

using the fact that $K^2 = I$. This invites comparison with (2.3.32), or more generally (2.3.8). Indeed, we can form the matrix exponential

$$e^{tK} = \sum_{n=0}^{\infty} \frac{t^n}{n!} K^n, \tag{2.9.39}$$

and a computation parallel to (2.3.3) yields

$$\frac{d}{dt} e^{tK} = K e^{tK}. \tag{2.9.40}$$

Now we compare e^{tK} with $H(t)$ via a computation parallel to (2.3.10):

$$\frac{d}{dt}\Big(e^{-tK} H(t)\Big) = -K e^{-tK} H(t) + e^{-tK} K H(t) = 0, \tag{2.9.41}$$

hence

$$e^{-tK} H(t) = I, \quad \forall\, t \in \mathbb{R}. \tag{2.9.42}$$

A similar computation gives

$$e^{-tK} e^{tK} = I, \quad \forall\, t \in \mathbb{R}, \quad \text{hence} \quad e^{-tK} = \big(e^{tK}\big)^{-1}, \tag{2.9.43}$$

so

$$e^{tK} = H(t). \tag{2.9.44}$$

Also, somewhat parallel to (2.3.14)–(2.3.16), we have, for $s, t \in \mathbb{R}$,

$$\frac{d}{dt} e^{(s+t)K} e^{-tK} = 0, \tag{2.9.45}$$

hence $e^{(s+t)K}e^{-tK} = e^{sK}$, so

$$e^{(s+t)K} = e^{sK}e^{tK}, \quad \forall\, s,t \in \mathbb{R}, \tag{2.9.46}$$

or, equivalently,

$$H(s+t) = H(s)H(t), \quad \forall\, s,t \in \mathbb{R}. \tag{2.9.47}$$

Returning to the action of $H(t)$ as a linear fractional transformation, we have

$$\Phi_t(z) = L_{H(t)}(z) = \frac{(\cosh t)z + \sinh t}{(\sinh t)z + \cosh t}, \tag{2.9.48}$$

and (2.9.47), together with (2.9.5), yields

$$\Phi_{s+t} = \Phi_s \circ \Phi_t, \quad s,t \in \mathbb{R}. \tag{2.9.49}$$

Note that

$$\Phi_t(0) = \frac{\sinh t}{\cosh t} = \tanh t, \tag{2.9.50}$$

where the last identity defines $\tanh t$. We have

$$\tanh : \mathbb{R} \longrightarrow (-1,1), \quad \tanh t \to \pm 1 \text{ as } t \to \pm\infty. \tag{2.9.51}$$

Also, by (2.9.34) and basic identities discussed in §2.1,

$$\frac{d}{dt}\tanh t = \frac{\cosh^2 t - \sinh^2 t}{\cosh^2 t} = 1 - \tanh^2 t = \frac{1}{\cosh^2 t}. \tag{2.9.52}$$

In particular the derivative is > 0 for all t, so the inverse function theorem (Proposition 2.1.2) implies the map (2.9.51) is bijective. Hence each $x \in (-1,1)$ can be uniquely written as $x = \tanh t$ for $t \in \mathbb{R}$. Combining (2.9.52) with Proposition 2.1.5 gives

$$\int_0^x \frac{dy}{1-y^2} = \tanh^{-1} x, \quad |x| < 1, \tag{2.9.53}$$

via the substitution $y = \tanh t$.

Note that applying (2.9.49) to $z = 0$ gives

$$\Phi_s(\tanh t) = \tanh(s+t), \quad s,t \in \mathbb{R}. \tag{2.9.54}$$

As an alternative derivation, note that (2.9.48) gives

$$\Phi_s(z) = \frac{z + \tanh s}{(\tanh s)z + 1}, \tag{2.9.55}$$

and then (2.9.54) is equivalent to the identity

$$\tanh(s+t) = \frac{\tanh s + \tanh t}{1 + (\tanh s)(\tanh t)}, \tag{2.9.56}$$

which can also be established by calculations parallel to (2.4.35).

Here is one significant consequences of these calculations regarding Φ_t.

Proposition 2.9.3. *Let $S \subset \mathcal{D}$ be a circle centered at a point $p \in (-1,1)$. Then we can find $t \in \mathbb{R}$ such that $\Phi_t(S) \subset \mathcal{D}$ is a circle centered at* 0.

Proof. Say S intersects $(-1,1)$ at $\tanh s_1$ and $\tanh s_2$, $s_1 < s_2$. Then $\Phi_t(S) \subset \mathcal{D}$ is a circle, centered at a point in $(-1,1)$, that intersects $(-1,1)$ at $\tanh(s_1+t)$ and $\tanh(s_2+t)$. Now

$$s_2 + t = -(s_1+t) \Longleftrightarrow t = -\frac{s_1+s_2}{2}. \tag{2.9.57}$$

Thus, if t satisfies (2.9.57), it follows that $\Phi_t(S)$ intersects $(-1,1)$ at

$$\tanh\frac{s_1-s_2}{2} \quad \text{and at} \quad \tanh\frac{s_2-s_1}{2}. \tag{2.9.58}$$

Symmetry considerations imply $\Phi_t(S)$ is centered at 0. □

Remark. Using (2.9.56), we have

$$\tanh(s_1-s_2) = \frac{\tanh s_1 - \tanh s_2}{1-(\tanh s_1)(\tanh s_2)}, \tag{2.9.59}$$

and

$$\begin{aligned} x &= \tanh\frac{s}{2}, \quad y = \tanh s \\ &\Longrightarrow y = \frac{2x}{1+x^2} \\ &\Longrightarrow x = \frac{1-\sqrt{1-y^2}}{y}. \end{aligned} \tag{2.9.60}$$

We can readily extend Proposition 2.9.3 to a broader setting.

Proposition 2.9.4. *If $S \subset \mathcal{D}$ is a circle, we can find $A \in SU(1,1)$ such that*

$$L_A(S) \subset \mathcal{D} \quad \textit{is centered at } 0. \tag{2.9.61}$$

Proof. Rotate $\mathcal{D}$ so that S is centered at a point in $(-1,1)$. □

We can use Proposition 2.9.4 to tackle the following classical problem.

Three circles problem. Let C_1, C_2, C_3 be three disjoint circles in $\mathbb{C}$. Assume no one circle separates the other two. In this situation, find a fourth circle S that is tangent to each C_j.

Solution. Using a translation and dilation, we can assume

$$C_1 = S^1 = \{z : |z| = 1\}. \tag{2.9.62}$$

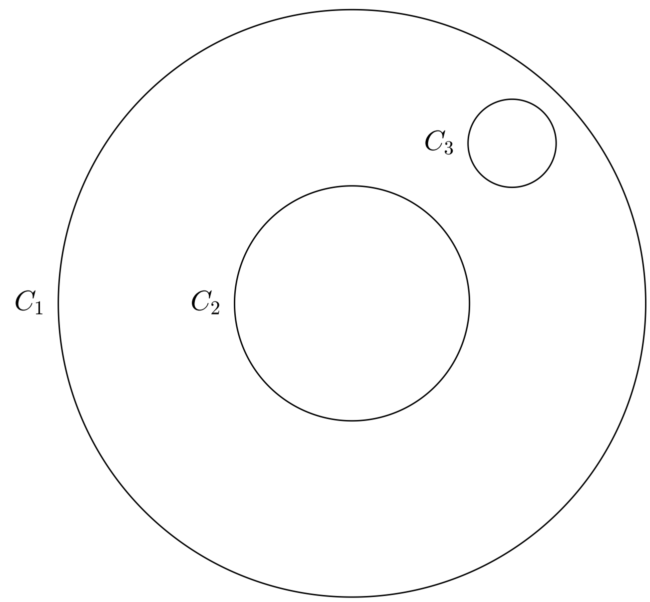

Figure 2.9.1. Setup for solution to three circles problem

Then either $C_2, C_3 \subset \mathcal{D}$ or $C_2, C_3 \subset \mathbb{C} \setminus \overline{\mathcal{D}}$. In the latter case, use the inversion $z \mapsto 1/z$ to arrange that

$$C_2,\ C_3 \subset \mathcal{D}. \tag{2.9.63}$$

Now, using Proposition 2.9.4, apply a linear fractional transformation, preserving $\mathcal{D}$, to arrange that C_2 is centered at 0, so

$$C_2 = \{z : |z| = a\}, \quad 0 < a < 1. \tag{2.9.64}$$

See Figure 2.9.1. Let

$$d = \frac{1-a}{2}, \quad \Gamma = \Big\{z : |z| = a + d = \frac{1+a}{2}\Big\}. \tag{2.9.65}$$

Say C_3 has center $p \in \mathcal{D}$ and radius r, i.e.,

$$C_3 = \{z : |z - p| = r\}, \tag{2.9.66}$$

and set

$$\Sigma = \{z : |z - p| = r + d\}. \tag{2.9.67}$$

Then Σ intersects Γ at two points. Pick one, call it q. Note that q has distance d from both C_1 and C_2. Furthermore, it has distance $r + d$ from p, hence distance d from C_3. Hence the circle

$$S = \{z : |z - q| = d\} \tag{2.9.68}$$

is tangent to C_1, C_2, and C_3.

Hyperbolic and spherical geometry

One can put distance functions $d_{\mathcal{U}}$ and $d_{\mathcal{D}}$ on the upper half-plane $\mathcal{U}$ and the disk $\mathcal{D}$ that have the properties

$$\begin{aligned} \mathrm{Isom}^+(\mathcal{U}) &= \{L_A : A \in S\ell(2, \mathbb{R})\}, \\ \mathrm{Isom}^+(\mathcal{D}) &= \{L_A : A \in \mathrm{SU}(1,1)\}. \end{aligned} \tag{2.9.69}$$

Such spaces are called the Poincaré upper half-plane and the Poincaré disc. These are non-Euclidean geometries, known as hyperbolic geometries. The map $\varphi : \mathcal{U} \to \mathcal{D}$ given in (2.9.11) is an isometry between the two spaces. One can see how this is done in §5.12 of [**6**], and also see applications to deep results in complex analysis, including the Riemann mapping theorem and Picard's theorems.

In another direction, one can use stereographic projection to produce a bijective map of $\widehat{\mathbb{C}}$ onto S^2 (the unit sphere in $\mathbb{R}^3$, discussed in the next section). This yields maps

$$\widetilde{L}_A : S^2 \longrightarrow S^2, \quad A \in S\ell(2, \mathbb{C}), \tag{2.9.70}$$

satisfying $\widetilde{L}_{AB} = \widetilde{L}_A \circ \widetilde{L}_B$, and preserving angles. One has

$$\mathrm{Isom}^+(S^2) = \{\widetilde{L}_A : A \in \mathrm{SU}(2)\}, \tag{2.9.71}$$

where $A \in \mathrm{SU}(2) \Leftrightarrow A \in S\ell(2, \mathbb{C})$ and $A^*A = I$. (Note that $L_A = L_{-A} \Rightarrow \widetilde{L}_A = \widetilde{L}_{-A}$.) See §5.3 of [**6**] for more on this. The sphere S^2 also has a non-Euclidean geometry, known as spherical geometry.

2.10. Higher dimensions

Beyond plane geometry, we have n-dimensional Euclidean geometry, for each $n \in \mathbb{N}$. We sketch some results on this subject here, referring to Chapter 3 of [**3**] for a systematic development of the algebraic aspects, and to [**5**] for the very rich theory of calculus of several variables.

We start with $\mathbb{R}^n$, which consists of n-tuples of real numbers,

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad x_j \in \mathbb{R},\ 1 \le j \le n. \tag{2.10.1}$$

The number x_j is called the jth component of x. We discuss some algebraic and metric structures on $\mathbb{R}^n$. first, there is addition. If x is as in (2.10.1) and also $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, we have

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \in \mathbb{R}^n. \tag{2.10.2}$$

Addition is done componentwise, Also, given $a \in \mathbb{R}^n$, we have

$$ax = (ax_1, \dots, ax_n) \in \mathbb{R}^n. \tag{2.10.3}$$

This is scalar multiplication. In (2.10.1), we represent x as a row vector. Sometimes it is convenient to represent x as a column vector,

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{2.10.4}$$

Then (2.10.2)–(2.10.3) are converted to

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad ax = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}. \tag{2.10.5}$$

As seen in Section 1.1 and subsequent material, the inner product played a central role in 2D geometry. In the setting of $\mathbb{R}^n$, we have the inner product

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j = x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}, \tag{2.10.6}$$

given $x, y \in \mathbb{R}^n$. This inner product has the properties

$$\begin{aligned} &\langle x, y \rangle = \langle y, x \rangle, \\ &\langle x, ay + bz \rangle = a\langle x, y \rangle + b\langle x, z \rangle, \\ &\langle x, x \rangle > 0 \ \text{ unless } \ x = 0. \end{aligned} \tag{2.10.7}$$

Note that

$$\langle x, x \rangle = x_1^2 + \cdots + x_n^2. \tag{2.10.8}$$

We set

$$|x| = \sqrt{\langle x, x \rangle}, \tag{2.10.9}$$

which we call the norm of x. Note that

$$\langle ax, ax\rangle = a^2\langle x, x\rangle, \tag{2.10.10}$$

hence

$$|ax| = |a|\cdot|x|, \quad \text{for} \ \ a\in\mathbb{R},\ x\in\mathbb{R}^n. \tag{2.10.11}$$

Extending the notion of distance in (1.1.13), we say that the distance from x to y in $\mathbb{R}^n$ is

$$d(x, y) = |x - y|. \tag{2.10.12}$$

We claim that the Euclidean distance, defined by (2.10.12), satisfies the "triangle inequality,"

$$d(x, y) \le d(x, z) + d(z, y), \quad \forall\, x, y, z \in \mathbb{R}^n \tag{2.10.13}$$

This in turn is a consequence of the following, also called the triangle inequality.

Proposition 2.10.1. *The norm* (2.10.9) *on* $\mathbb{R}^n$ *has the property*

$$|x + y| \le |x| + |y|, \quad \forall\, x, y\in\mathbb{R}^n. \tag{2.10.14}$$

Proof. We compare the squares of the two sides of (2.10.14). first,

$$\begin{aligned} |x+y|^2 &= \langle x+y, x+y\rangle \\ &= |x|^2 + |y|^2 + 2\langle x, y\rangle. \end{aligned} \tag{2.10.15}$$

Next,

$$(|x| + |y|)^2 = |x|^2 + |y|^2 + 2|x|\cdot|y|. \tag{2.10.16}$$

We see that (2.10.14) holds if and only if $\langle x, y\rangle \le |x|\cdot|y|$. Thus the proof of Proposition 2.10.1 is finished by the following result, known as Cauchy's inequality. □

Proposition 2.10.2. *For all* $x, y\in\mathbb{R}^n$*,*

$$|\langle x, y\rangle| \le |x|\cdot|y|. \tag{2.10.17}$$

Proof. We start with the chain

$$0 \le |x-y|^2 = \langle x-y, x-y\rangle = |x|^2 + |y|^2 - 2\langle x, y\rangle, \tag{2.10.18}$$

which implies

$$2\langle x, y\rangle \le |x|^2 + |y|^2, \quad \forall\, x, y\in\mathbb{R}^n. \tag{2.10.19}$$

If we replace x by tx and y by $t^{-1}y$, with $t > 0$, the left side of (2.10.19) is unchanged, so we have

$$2\langle x, y\rangle \le t^2|x|^2 + t^{-2}|y|^2, \quad \forall\, t > 0. \tag{2.10.20}$$

We pick t so that the two terms on the right side of (2.10.20) are equal, namely,

$$t^2 = \frac{|y|}{|x|}, \quad t^{-2} = \frac{|x|}{|y|}. \tag{2.10.21}$$

Note that (2.10.17) is obvious if $x = 0$ or $y = 0$, so we will assume that $x \neq 0$ and $y \neq 0$. Plugging (2.10.21) into (2.10.20) gives

$$\langle x, y\rangle \le |x| \cdot |y|, \quad \forall\, x, y \in \mathbb{R}^n. \tag{2.10.22}$$

This is almost (2.10.17). To finish, we can replace x in (2.10.22) by $-x = (-1)x$, getting

$$-\langle x, y\rangle \le |x| \cdot |y|, \tag{2.10.23}$$

and together (2.10.22) and (2.10.23) give (2.10.17). □

Let us emphasize the conclusion of (2.10.15) as a very significant extension of (1.1.3):

$$|x + y|^2 = |x|^2 + |y|^2 + 2\langle x, y\rangle, \tag{2.10.24}$$

for all $x, y \in \mathbb{R}^n$. Parallel to (1.1.6), we write

$$x \perp y \Longleftrightarrow \langle x, y\rangle = 0, \tag{2.10.25}$$

and say x is orthogonal (or perpendicular) to y in $\mathbb{R}^n$. Parallel to Proposition 1.1.1 we have:

Proposition 2.10.3. *Given* $x, y \in \mathbb{R}^n$,

$$|x + y|^2 = |x|^2 + |y|^2 \Longleftrightarrow x \perp y. \tag{2.10.26}$$

To proceed, it is desirable to extend our scope, and consider more general inner product spaces V. First, V is a vector space. That is, given $x, y \in V,\ a \in \mathbb{R}$, we have defined vector addition, and multiplication by a scalar,

$$x, y \in V,\ a \in \mathbb{R} \Longrightarrow x + y,\ ax \in V. \tag{2.10.27}$$

These operations verify a standard set of commutative, associative, and distributive laws. See Chapter 1 of [**3**]. We mention that a map $T : V \to V$ is called a linear transformation provided T preserves these vector operations, i.e.,

$$T(ax + by) = aTx + bTy, \quad \forall\, x, y \in V,\ a, b \in \mathbb{R}. \tag{2.10.28}$$

As shown in Chapter 1 of [**3**], in case $V = \mathbb{R}^n$, matrix multiplication associates each $n \times n$ matrix $A \in M(n, \mathbb{R})$ to a linear transformation on $\mathbb{R}^n$.

An inner product space is a vector space V, equipped with an inner product:

$$x, y \in V \Longrightarrow \langle x, y\rangle \in \mathbb{R}, \tag{2.10.29}$$

satisfying (2.10.7).

Example. The integral formula

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt \tag{2.10.30}$$

defines an inner product on the space $C([a, b])$ of continuous functions on $[a, b]$. It also defines an inner product on $\mathcal{P}$, the space of polynomials in t.

For a general inner product space V, (2.10.9) defines a norm on V. Generally, a norm on V is a function $x \mapsto |x|$ satisfying

$$\begin{gathered} |ax| = |a|\cdot|x|, \quad \forall\, a \in \mathbb{R},\ x \in V, \\ |x| > 0 \ \text{ unless } \ x = 0, \\ |x + y| \le |x| + |y|. \end{gathered} \tag{2.10.31}$$

As before, the last of these results is called the triangle inequality. The proof that this holds for a general inner product space is identical to that of Proposition 2.10.1 (for $\mathbb{R}^n$). In particular, Cauchy's inequality (2.10.17) arises.

We specialize to finite-dimensional vector spaces V, i.e., to the case where there is a finite set $S = \{v_1, \dots, v_n\} \subset V$ such that $\operatorname{Span} S = V$, where

$$\operatorname{Span}\{v_1, \dots, v_n\} = \Bigl\{\sum_{j=1}^n a_j v_j : a_j \in \mathbb{R}\Bigr\}. \tag{2.10.32}$$

A minimal spanning set of V is called a basis. It is a fundamental fact that any two bases of V have the same number of elements (see [**3**], Corollary 1.3.3). This number is called the dimension of V, and denoted $\dim V$. In particular, of course, $\dim \mathbb{R}^n = n$.

If V is a finite-dimensional inner product space, a basis $\{u_1, \dots, u_n\}$ of V is called an orthonormal basis of V provided

$$\langle u_j, u_k\rangle = \delta_{jk}, \quad 1 \le j, k \le n, \tag{2.10.33}$$

i.e.,

$$|u_j| = 1, \quad j \ne k \Rightarrow \langle u_j, u_k\rangle = 0. \tag{2.10.34}$$

When (2.10.33) holds, we have

$$v = \sum_{j=1}^n a_j u_j,\ w = \sum_{j=1}^n b_j u_j \Longrightarrow \langle u, v\rangle = \sum_{j=1}^n a_j b_j. \tag{2.10.35}$$

It is very useful to be able to construct orthonormal bases. The construction we now describe is called the Gram-Schmidt construction.

Proposition 2.10.4. *Let* $\{v_1, \dots v_n\}$ *be a basis of* V*, an inner product space. Then there is an orthonormal basis* $\{u_1, \dots, u_n\}$ *of* V *such that*

$$\operatorname{Span}\{u_j : 1 \le j \le \ell\} = \operatorname{Span}\{v_j : 1 \le j \le \ell\}, \quad 1 \le \ell \le n. \tag{2.10.36}$$

Proof. To begin, take

$$u_1 = \frac{1}{|v_1|} v_1. \tag{2.10.37}$$

Now define the linear transformation $P_1 : V \to V$ by $P_1 v = \langle v, u_1 \rangle u_1$ and set

$$\tilde{v}_2 = v_2 - P_1 v_2 = v_2 - \langle v_2, u_1 \rangle u_1. \tag{2.10.38}$$

We see that $\langle \tilde{v}_2, u_1 \rangle = \langle v_2, u_1 \rangle - \langle v_2, u_1 \rangle = 0$. Also $\tilde{v}_2 \ne 0$ since $v_2 \notin \operatorname{Span}(u_1)$. Hence we can set

$$u_2 = \frac{1}{|\tilde{v}_2|} \tilde{v}_2. \tag{2.10.39}$$

Inductively, suppose we have an orthonormal set $\{u_1, \dots, u_m\}$ with $m < n$ and (2.10.36) holding for $1 \le \ell \le m$. Then define $P_m : V \to V$ (the orthogonal projection of V onto $\operatorname{Span}(u_1, \dots, u_m)$) by

$$P_m v = \sum_{j=1}^{m} \langle v, u_j \rangle u_j, \tag{2.10.40}$$

and set

$$\begin{aligned} \tilde{v}_{m+1} &= v_{m+1} - P_m v_{m+1} \\ &= v_{m+1} - \sum_{j=1}^{m} \langle v_{m+1}, u_j \rangle u_j. \end{aligned} \tag{2.10.41}$$

We see that

$$j \le n \Rightarrow \langle \tilde{v}_{m+1}, u_j \rangle = \langle v_{m+1}, u_j \rangle - \langle v_{m+1}, u_j \rangle = 0. \tag{2.10.42}$$

Also, since $v_{m+1} \notin \operatorname{Span}(v_1, \dots, v_m) = \operatorname{Span}(u_1, \dots, u_m)$, it follows that $\tilde{v}_{m+1} \ne 0$. Hence we can set

$$u_{m+1} = \frac{1}{|\tilde{v}_{m+1}|} \tilde{v}_{m+1}. \tag{2.10.43}$$

This completes the construction. □

The following is a useful extension. Let V be an n-dimensional inner product space, and let $W \subset V$ be a linear subspace, i.e., a subset satisfying

$$w_j \in W,\ a_j \in \mathbb{R} \Longrightarrow \sum a_j w_j \in W. \tag{2.10.44}$$

Then W has an orthonormal basis $\{w_1, \dots, w_m\}$ ($m = \dim W$), and

$$P_W : V \longrightarrow W, \quad P_W v = \sum_{j=1}^{m} \langle v, w_j \rangle w_j \tag{2.10.45}$$

is the orthogonal projection of V onto W. Furthermore,

$$I - P_W : V \longrightarrow W^\perp \tag{2.10.46}$$

is the orthogonal projection of V onto

$$W^\perp = \{v \in V : v \perp w, \forall\, w \in W\}. \tag{2.10.47}$$

Given an n-dimensional inner product space V, with orthonormal basis $S = \{u_1, \dots, u_n\}$, we can define a linear map

$$\mathcal{J}_S : \mathbb{R}^n \to V, \quad \mathcal{J}_S \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{j=1}^{n} a_j u_j. \tag{2.10.48}$$

Equivalently,

$$\mathcal{J}_S e_j = u_j, \quad 1 \le j \le n, \tag{2.10.49}$$

where $\{e_1, \dots, e_n\}$ is the standard orthonormal basis of $\mathbb{R}^n$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{2.10.50}$$

The map $\mathcal{J}_S$ is bijective, with inverse

$$\mathcal{J}_S^{-1} : V \longrightarrow \mathbb{R}^n, \quad \mathcal{J}_S^{-1} \sum a_j u_j = \sum a_j e_j. \tag{2.10.51}$$

The map $\mathcal{J}_S$ also preserves inner products:

$$\langle \mathcal{J}_S x, \mathcal{J}_S y \rangle = \langle x, y \rangle, \quad \forall\, x, y \in \mathbb{R}^n, \tag{2.10.52}$$

where the inner product on the right is given by (2.10.6).

For a related concept, if V is an inner product space, and $T : V \to V$ is a linear transformation, we say

$$T \in \mathrm{O}(V) \Longleftrightarrow \langle Tu, Tv \rangle = \langle u, v \rangle, \quad \forall\, u, v, \in V. \tag{2.10.53}$$

We denote $\mathrm{O}(\mathbb{R}^n)$ by $\mathrm{O}(n)$. If $A \in M(n, \mathbb{R})$ defines a linear transformation,

$$A : \mathbb{R}^n \to \mathbb{R}^n, \quad (Ax)_j = \sum_k a_{jk} x_k, \; A = (a_{jk}), \tag{2.10.54}$$

then

$$A \in \mathrm{O}(n) \Longleftrightarrow A^t A = I, \; A^t = (a_{kj}). \tag{2.10.55}$$

Given an n-dimensional inner product space V, with distance $d(u,v) = |u-v|$, and a map $F : V \to V$, we say

$$F \in \operatorname{Isom}(V) \Longleftrightarrow d(F(u),F(v)) = d(u,v), \quad \forall\, u,v \in V. \tag{2.10.56}$$

We have the following extension of Corollary 1.2.4.

Proposition 2.10.5. *Given $F \in \operatorname{Isom}(V)$, then F has the form*

$$F(v) = Av + y, \quad A \in \mathrm{O}(V),\ y \in V. \tag{2.10.57}$$

In fact, as in §1.2, there is a decomposition $\operatorname{Isom}(V) = \operatorname{Isom}^+(V) \cup \operatorname{Isom}^-(V)$. To describe this, it suffices to take $V = \mathbb{R}^n$, so (2.10.57) becomes

$$F(v) = Av + y, \quad A \in \mathrm{O}(n),\ y \in \mathbb{R}^n, \tag{2.10.58}$$

and to see that

$$\mathrm{O}(n) = \mathrm{O}^+(n) \cup \mathrm{O}^-(n). \tag{2.10.59}$$

To do this, we bring in the determinant,

$$\det : M(n,\mathbb{R}) \longrightarrow \mathbb{R}. \tag{2.10.60}$$

To characterize $\det A$ for $A \in M(n,\mathbb{R})$, we write $A = (a_1,\dots,a_n)$, where each $a_j \in \mathbb{R}^n$ is a column vector. The following result is the key to determinants.

Proposition 2.10.6. *There is a unique function $\vartheta : M(n,\mathbb{R}) \to \mathbb{R}$ such that*

(a) $\vartheta(A)$ *is linear in each column of A,*
(b) $\vartheta(\widetilde{A}) = -\vartheta(A)$ *if $\widetilde{A}$ is obtained by switching two columns of A,*
(c) $\vartheta(I) = 1$.

We denote this function by $\det A$. *If (c) is changed to*

(c′) $\vartheta(I) = r$,

then $\vartheta(A) = r \det A$.

For a proof, see [**3**], §1.5. The analysis there gives also

$$\det AB = (\det A)(\det B), \quad \det A^t = \det A. \tag{2.10.61}$$

Then (2.10.55) yields

$$A \in \mathrm{O}(n) \Rightarrow (\det A)^2 = 1 \Rightarrow \det A = \pm 1. \tag{2.10.62}$$

To define $O^{\pm}(n)$, we set

$$\mathrm{O}^{\pm}(n) = \{A \in \mathrm{O}(n) : \det A = \pm 1\}. \tag{2.10.63}$$

We also use the notation

$$\mathrm{SO}(n) = \mathrm{O}^{+}(n). \tag{2.10.64}$$

Returning to orthogonal complements, we mention an alternative way to find $W^{\perp}$ when W is a 2-dimensional subspace of $\mathbb{R}^3$, using the cross product, an $\mathbb{R}$-bilinear map $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$i \times j = k, \quad j \times k = i, \quad k \times i = j, \quad u \times v = -v \times u. \tag{2.10.65}$$

Here i, j, k is a typical notation for the standard basis e_1, e_2, e_3 of $\mathbb{R}^3$. A convenient alternative characterization is the following.

Proposition 2.10.7. *Given* $u, v, w \in \mathbb{R}^3$,

$$\langle w, u \times v \rangle = \det \begin{pmatrix} w_1 & u_1 & v_1 \\ w_2 & u_2 & v_2 \\ w_3 & u_3 & v_3 \end{pmatrix}. \tag{2.10.66}$$

Proof. The linearity of the right side in u, v, w and the antisymmetry follow from Proposition 2.10.6. It is straightforward to check the identities $i \times j = k$, etc., listed in (2.10.65). □

The antisymmetry of the determinant also implies

$$w \in \mathrm{Span}(u, v) \Longrightarrow \langle w, u \times v \rangle = 0, \tag{2.10.67}$$

hence

$$u \times v \perp \mathrm{Span}(u, v). \tag{2.10.68}$$

Consequently, if $\{u, v\}$ span the 2-dimensional space $W \subset \mathbb{R}^3$, then

$$N = \frac{1}{|u \times v|}(u \times v) \tag{2.10.69}$$

is a unit vector spanning $W^{\perp}$. Thus

$$P_{W^{\perp}} w = \langle w, N \rangle N. \tag{2.10.70}$$

We mention another useful fact about the cross product.

Proposition 2.10.8. *Given* $u, v \in \mathbb{R}^3$,

$$T \in \mathrm{SO}(3) \Longrightarrow Tu \times Tv = T(u \times v). \tag{2.10.71}$$

Proof. Denote the matrix on the right side of (2.10.66) by M. Then, if $T \in \mathrm{SO}(3)$,

$$\langle Tw, Tu \times Tv \rangle = \det TM = (\det T)(\det M) = \det M, \tag{2.10.72}$$

hence $T^t(Tu \times Tv) = u \times v$, yielding (2.10.71). □

We briefly mention the regular solids in $\mathbb{R}^3$, also called the Platonic solids. In contrast to the infinite family of regular polygons, there are only five:

tetrahedron,

cube,

octahedron,

dodecahedron,

icosahedron,

having, respectively, 4,6,8,12, and 20 faces. Constructions of the first three is elementary. The last two are discussed in [**9**].

We move on to a discussion of volume. There is a notion of n-dimensional volume for certain bounded sets $S \subset \mathbb{R}^n$, defined in a fashion parallel to the definition of area for the planar case done in §2.6. To start, we define a standard n-dimensional cube Q of side-length ε parallel to (2.6.2),

$$\operatorname{Vol} Q = \varepsilon^n. \tag{2.10.73}$$

and then define $\mathrm{Cont}^{\pm}(S)$ and $\mathrm{Vol}(S)$ in a fashion parallel to (2.6.3)–(2.6.7). Parallel to Proposition 2.6.5, we have

Proposition 2.10.9. *If $S \subset \mathbb{R}^n$ is bounded, then*

$$\mathrm{Cont}^{\pm}(F(S)) = \mathrm{Cont}^{\pm}(S), \quad \forall\, F \in \mathrm{Isom}(\mathbb{R}^n). \tag{2.10.74}$$

However, we need a different proof of this result, since it is not so convenient to extend Lemma 2.6.4. In fact, it is convenient to take a different approach, yielding the following more general result.

Proposition 2.10.10. *Let $S \subset \mathbb{R}^n$ be bounded and assume $A \in M(n, \mathbb{R})$ is invertible. Then*

$$\mathrm{Cont}^{\pm}(A(S)) = |\det A|\, \mathrm{Cont}^{\pm}(S). \tag{2.10.75}$$

Here is an application of Proposition 2.10.10 to the volume of a parallelopiped in $\mathbb{R}^3$. To set it up, consider the standard unit cube in $\mathbb{R}^3$,

$$Q = \{x \in \mathbb{R}^3 : 0 \le x_j \le 1,\ 1 \le j \le 3\}, \quad \operatorname{Vol} Q = 1. \tag{2.10.76}$$

Let M denote the matrix appearing on the right side of (2.10.66), so $w = Mi,\ u = Mj,\ v = Mk$. The image

$$P = M(Q) \tag{2.10.77}$$

is a parallelopiped, with vertices $0, u, v, w, u+v, u+w, v+w$, and $u+v+w$. If we combine Proposition 2.10.10 with (2.10.66), we obtain the volume calculation

$$\operatorname{Vol} P = |\langle w, u \times v\rangle|. \tag{2.10.78}$$

A proof of Proposition 2.10.10 is sketched in Proposition 1.6.5 of [**3**]. A detailed proof is given in Proposition 3.1.12 of [**5**]. This proof is done in the setting of Riemann integrals of functions on $\mathbb{R}^n$. This result is extended in Proposition 3.1.14 of [**5**] to the identity

$$\int_{\Omega} f(y)\,dy = \int_{\mathcal{O}} f(G(x))|\det DG(x)|\,dx, \tag{2.10.79}$$

where $G : \mathcal{O} \to \Omega$ is a C^1 map with C^1 inverse and $DG(x)$ is the $n \times n$ matrix of first order partial derivatives of the components of $G(x)$. This is an n-dimensional extension of Proposition 2.1.5.

Extending the notion of arc-length of a planar curve, we have the notion of $(n-1)$-dimensional area of an $(n-1)$-dimensional surface $M \subset \mathbb{R}^n$. This material is developed in Chapter 3 of [**5**] (along with the k-dimensional area of a k-dimensional surface in $\mathbb{R}^n$, for $1 \le k \le n-1$). We restrict ourselves to a description of a selection of results established there, with a focus on the computation of volumes of balls and areas of spheres in $\mathbb{R}^n$,

$$B^n = \{x \in \mathbb{R}^n : |x| \le 1\}, \quad S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}. \tag{2.10.80}$$

To start, suppose M is the image of a C^1 map

$$\varphi : \mathcal{O} \longrightarrow \mathbb{R}^n, \tag{2.10.81}$$

where $\mathcal{O} \subset \mathbb{R}^{n-1}$ is smoothly bounded. Let $D\varphi(x)$ denote the $n \times (n-1)$ matrix of partial derivatives of the components of φ, and set

$$G(x) = D\varphi(x)^t D\varphi(x), \quad g(x) = \det G(x). \tag{2.10.82}$$

Then

$$\operatorname{Area} M = \int_{\mathcal{O}} \sqrt{g(x)}\,dx. \tag{2.10.83}$$

More generally, if $f : M \to \mathbb{R}$ is continuous,

$$\int_M f\,dS = \int_{\mathcal{O}} f(\varphi(x))\sqrt{g(x)}\,dx. \tag{2.10.84}$$

See (3.2.12) of [**5**]. In case $n = 3$, this becomes

$$\int_M f\,dS = \int_{\mathcal{O}} f(\varphi(x,y))|\partial_x\varphi \times \partial_y\varphi|\,dx\,dy. \tag{2.10.85}$$

See (3.2.18) of [**5**].

We turn to calculations involving spherical polar coordinates, which will result in calculations of

$$V_n = \operatorname{Vol} B^n, \quad A_{n-1} = \operatorname{Area} S^{n-1}. \tag{2.10.86}$$

As seen in (3.2.28) of [**5**], we can use (2.10.79) to show that if f is a continuous function on $\mathbb{R}^n$ that decreases sufficiently rapidly at infinity, then

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{S^{n-1}} \Bigl\{\int_0^\infty f(r\omega)r^{n-1}\,dr\Bigr\}\,dS(\omega). \tag{2.10.87}$$

In particular, for $f(x) = g(|x|)$, we have

$$\int_{\mathbb{R}^n} g(|x|)\,dx = A_{n-1}\int_0^\infty g(r)r^{n-1}\,dr. \tag{2.10.88}$$

By a limiting argument, (2.10.87) applies to $g(r) = 1$ for $0 \le r \le 1$, 0 otherwise, yielding

$$V_n = A_{n-1}\int_0^1 r^{n-1}\,dr = \frac{1}{n}A_{n-1}. \tag{2.10.89}$$

Recall we already have this for $n = 2$:

$$A_1 = \ell(S^1) = 2\pi, \quad V_2 = \operatorname{Area}\mathcal{D} = \pi. \tag{2.10.90}$$

We desire to compute A_{n-1} for $n \ge 3$. Along the way we pick up some further marvelous identities by applying (2.10.87) to the Gaussian, $f(x) = e^{-|x|^2}$. We are looking at

$$I_n = \int_{\mathbb{R}^n} e^{-|x|^2}\,dx. \tag{2.10.91}$$

Since $e^{-|x|^2} = e^{-x_1^2}\cdots e^{-x_n^2}$, we have

$$I_n = I_1^n. \tag{2.10.92}$$

Meanwhile, (2.10.88) implies

$$\begin{aligned} I_n &= A_{n-2}\int_0^\infty e^{-r^2}r^{n-1}\,dr \\ &= \frac{1}{2}A_{n-1}\int_0^\infty e^{-s}s^{n/2-1}\,ds. \end{aligned} \tag{2.10.93}$$

The case $n = 2$ yields an elementary integral:

$$I_2 = \frac{1}{2}A_1\int_0^\infty e^{-s}\,ds = \frac{1}{2}A_1 = \pi, \tag{2.10.94}$$

the last identity thanks to (2.10.90). Hence, by (2.10.92),

$$I_1 = \pi^{1/2}, \quad I_n = \pi^{n/2}. \tag{2.10.95}$$

Therefore (2.10.93) yields

$$A_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \tag{2.10.96}$$

where

$$\Gamma(z) = \int_0^\infty e^{-s} s^{z-1}\, ds, \quad z > 0, \tag{2.10.97}$$

defines Euler's gamma function. We record our result.

Proposition 2.10.11. *For $n \ge 1$, the area of S^{n-1} is given by* (2.10.96).

Note that

$$A_0 = 2 \Longrightarrow \Gamma\Big(\frac{1}{2}\Big) = \pi^{1/2}, \tag{2.10.98}$$

and (2.10.94) yields

$$\Gamma(1) = 1. \tag{2.10.99}$$

We can evaluate $\Gamma(n/2)$ for other $n \in \mathbb{N}$ via the following result.

Proposition 2.10.12. *For $z > 0$,*

$$\Gamma(z+1) = z\Gamma(z). \tag{2.10.100}$$

Proof. Indeed,

$$\begin{aligned} \Gamma(z+1) &= \int_0^\infty e^{-s} s^z\, ds = -\int_0^\infty \Big(\frac{d}{ds} e^{-s}\Big) s^z\, ds \\ &= \int_0^\infty e^{-s} \frac{d}{ds} s^z\, ds \\ &= \int_0^\infty e^{-s} z s^{z-1}\, ds, \end{aligned} \tag{2.10.101}$$

the third identity involving integration by parts. □

For $n = 3$, we have

$$\begin{aligned} A_2 &= \frac{2\pi^{3/2}}{\Gamma(3/2)} = \frac{2\pi^{3/2}}{(1/2)\Gamma(1/2)} = 4\pi, \\ V_3 &= \frac{1}{3} A_2 = \frac{4\pi}{3}, \end{aligned} \tag{2.10.102}$$

results obtained by Archimedes. Moving up a dimension, for $n = 4$ we have

$$\begin{aligned} A_3 &= \frac{2\pi^2}{\Gamma(2)} = 2\pi^2, \\ V_4 &= \frac{1}{4} A_3 = \frac{\pi^2}{2}. \end{aligned} \tag{2.10.103}$$

Having discussed various aspects of Euclidean geometry on $\mathbb{R}^n$, with emphasis on $n > 2$ in this section, we connect with the planar case, which was the focus of Sections 1–20. One clear difference is that in the plane we used multiplication in $\mathbb{C}$. For dimension $n \ge 3$, this role of multiplication shifted to multiplication of a matrix $A \in M(n, \mathbb{R})$ by a vector $x \in \mathbb{R}^n$. Indeed, $\mathbb{R}^n$ does not generally have the structure of an algebra with such nice properties that we exploited on $\mathbb{C}$, such as multiplicative inverses and a multiplicative norm. W. Hamilton spent some time trying to find such a structure on $\mathbb{R}^3$. Eventually he succeeded, not for $\mathbb{R}^3$, but for $\mathbb{R}^4$, with the quaternions, which we briefly discuss. Detailed treatments are given in Chapter 8 of [**3**] and Chapter 10 of [**8**].

The algebra $\mathbb{H}$ of quaternions is a 4-dimensional real vector space, with elements of the form

$$\xi = a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R}. \tag{2.10.104}$$

Here, i, j, and k play a role similar to the standard basis of $\mathbb{R}^3$ that arose above. The product on $\mathbb{H}$ is an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$, with $1 \in \mathbb{H}$ acting as the multiplicative identity. Products involving i, j, and k have some features in common with the cross product:

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik. \tag{2.10.105}$$

On the other hand, the squares of these objects follow a new rule:

$$i^2 = j^2 = k^2 = -1. \tag{2.10.106}$$

Otherwise stated, if we write

$$\xi = a + u, \quad a \in \mathbb{R}, \ u \in \mathbb{R}^3, \tag{2.10.107}$$

and similarly write $\eta = b + v, \ b \in \mathbb{R}, \ v \in \mathbb{R}^3$, the product is given by

$$\xi\eta = (a + u)(b + v) = (ab - u \cdot v) + av + bu + u \times v. \tag{2.10.108}$$

Here $u \cdot v = \langle u, v \rangle$ is the inner product on $\mathbb{R}^3$, and $u \times v$ is the cross product. The quantity $ab - u \cdot v$ is the real part of $\xi\eta$ and $av + bu + u \times v$ is the vector part.

Multiplication on $\mathbb{H}$ is not commutative, but we have the following basic result.

Proposition 2.10.13. *Multiplication on $\mathbb{H}$ is associative, i.e.,*

$$\zeta(\xi\eta) = (\zeta\xi)\eta, \quad \forall\, \zeta, \xi, \eta \in \mathbb{H}. \tag{2.10.109}$$

The proof is a straightforward check. See Proposition 8.1.2 in [**3**].

In addition to the product, we have a conjugation operator:

$$\overline{\xi} = a - bi - cj - dk = a - u. \tag{2.10.110}$$

A calculation gives

$$\xi\overline{\eta} = (ab + u \cdot v) - av + bu - u \times v. \tag{2.10.111}$$

In particular,

$$\mathrm{Re}(\xi\overline{\eta}) = \mathrm{Re}(\overline{\eta}\xi) = \langle \xi, \eta \rangle, \tag{2.10.112}$$

the right side denoting the inner product on $\mathbb{R}^4$. Setting $\eta = \xi$ in (2.10.111) gives

$$\xi\overline{\xi} = |\xi|^2, \tag{2.10.113}$$

the Euclidean square-norm of ξ. Hence, whenever $\xi \in \mathbb{H}$ is nonzero, it has a multiplicative inverse,

$$\xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{2.10.114}$$

This structure makes $\mathbb{H}$ what is called a division ring.

To continue with products and conjugates, a routine calculation gives

$$\overline{\xi\eta} = \overline{\eta}\,\overline{\xi}. \tag{2.10.115}$$

Hence, via the associative law,

$$\begin{aligned} |\xi\eta|^2 &= (\xi\eta)(\overline{\xi\eta}) = \xi\eta\overline{\eta}\,\overline{\xi} \\ &= |\eta|^2\xi\overline{\xi} = |\xi|^2|\eta|^2, \end{aligned} \tag{2.10.116}$$

so

$$|\xi\eta| = |\xi|\,|\eta|. \tag{2.10.117}$$

This makes $\mathbb{H}$ a normed division ring.

Let us examine (2.10.117) when $\xi = u$ and $\eta = v$ are purely vectorial. We have

$$uv = -u \cdot v + u \times v. \tag{2.10.118}$$

Hence, directly,

$$|uv|^2 = (u \cdot v)^2 + |u \times v|^2, \tag{2.10.119}$$

while (2.10.117) implies

$$|uv|^2 = |u|^2|v|^2. \tag{2.10.120}$$

Now the law of cosines gives

$$u \cdot v = |u|\,|v| \cos\theta, \tag{2.10.121}$$

where θ is the angle between u and v in the plane in $\mathbb{R}^3$ spanned by these two vectors. See (2.2.42). Hence (2.10.119) implies

$$|u \times v|^2 = |u|^2|v|^2 \sin^2\theta. \tag{2.10.122}$$

We note that the set of unit-norm elements of $\mathbb{H}$,

$$\mathrm{Sp}(1) = \{\xi \in \mathbb{H} : |\xi| = 1\} = S^3, \tag{2.10.123}$$

forms a multiplicative group. Furthermore, by (2.10.117), we have the map

$$\lambda : \mathrm{Sp}(1) \times \mathrm{Sp}(1) \to \mathrm{SO}(4), \quad \lambda(\xi, \eta)\zeta = \xi\zeta\overline{\eta}, \tag{2.10.124}$$

which is a two-to-one homomorphism of $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$ onto $\mathrm{SO}(4)$, and

$$\pi : \mathrm{Sp}(1) \to \mathrm{SO}(3), \quad \pi(\xi)v = \xi v \bar{\xi}, \tag{2.10.125}$$

which is a two-to-one homomorphism of $\mathrm{Sp}(1)$ onto $\mathrm{SO}(3)$. We also have the following variant of Euler's formula. Assume $u \in \mathbb{R}^3$, $|u| = 1$, and regard $u \in \mathbb{H}$. Then, for $t \in \mathbb{R}$,

$$e^{tu} = \cos t + (\sin t)u, \quad e^{tu} \in \mathrm{Sp}(1). \tag{2.10.126}$$

Further results on the group $\mathrm{Sp}(1)$ can be found in Chapter 8 of [**3**] and in Chapter 10 of [**8**].

Bibliography

[1] S. Lang, *Algebra*, Addison-Wesley, Reading MA, 1965.

[2] E. Maor, *The Pythagorean Theorem – A 4,000-Year History*, Princeton Univ. Press, Princeton NJ, 2007.

[3] M. Taylor, *Linear Algebra*, Undergraduate Texts #45, Amer. Math. Soc., Providence RI, 2020.

[4] M. Taylor, *Introduction to Analysis in One Variable*, Undergraduate Texts #47, Amer. Math. Soc., Providence RI, 2020.

[5] M. Taylor, *Introduction to Analysis in Several Variables*, Undergraduate Texts #46, Amer. Math. Soc., Providence RI, 2020.

[6] M. Taylor, *Introduction to Complex Analysis*, GSM #202, Amer. Math. Soc., Providence RI, 2019.

[7] M. Taylor, *Measure Theory and Integration*, GSM #76, Amer. Math. Soc., Providence RI, 2006.

[8] M. Taylor, *Introduction to Lie Groups*, Open Math Notes, Amer. Math. Soc., Providence RI, 2021.

[9] M. Taylor, *The Pentagon, the Dodecahedron, and the Icosahedron*, Notes, available at https://mtaylor.web.unc.edu/notes, Item 18, Elementary geometry notes.

Index

acute angle, 29
angle bisection, 24
angle measurement, 3, 24, 26
angle sum for triangles, 71
arclength, 5, 63
Area, 7, 84
area of a disk, 89
areas of spheres, 118

basis, 110

calculus, 5, 56
Cauchy's inequality, 15, 108, 110
chain rule, 56
change of variable formula, 58
circle, 4, 46
  circumscribed, 51
  inscribed, 52
cis, 5, 64
compass and straightedge construction, 94
congruence, 3, 19, 31
conjugation, 2
Cont, 7, 84
cos, 5, 64
cosh, 101
cross product, 9, 114, 119
cube, 115
curve, 61

derivative, 56
derivative of a power series, 59
determinant, 9, 113
dimension, 110
disk, 46
distance, 16
drop a perpendicular, 40

equilateral triangle, 4, 38
Euclidean n-dimensional space, 107
Euclidean numbers, 8, 94
Euler formula, 6, 11, 69, 73, 121
exponential function, 5, 67

field, 2
fundamental theorem of calculus, 57

gamma function, 118
Gaussian integrals, 117
geometric construction of $\sqrt{2a}$, 49
Gram-Schmidt construction, 110
group, 97

Heron's formula, 8, 91
homomorphism, 97
hyperbolic functions, 101
hyperbolic geometry, 106

inner product, 3, 15
inner product space, 9, 109
integral, 56
integral formula for area, 88
inverse function theorem, 57
Isom, 3, 17, 113
Isom$^+$, 19
Isom$^-$, 19
isometry, 17
isosceles triangle, 4, 37

law of cosines, 66, 120
line, 20
line segment, 22
linear fractional transformation, 8, 97
linear transformation, 109
log, 69

Machin's formula, 77
mean value theorem, 56

n-gon, 42
non-euclidean geometry, 106
normed division ring, 120
normed field, 2
normed space, 110

obtuse angle, 29
orthogonal, 15
orthogonal projection, 112
orthonormal basis, 110

parallel, 3, 21
parallelogram, 43
parallelogram law, 45
parametrization by arclength, 63
perpendicular, 3, 15, 22
pi, 5, 65, 70, 89
Poincaré disk, 106
Poincaré upper half-plane, 106
power series, 58
Pythagorean theorem, 3, 29, 41, 88
Pythagorean triples, 50

quaternions, 10, 119

ray, 24
real angle measurement, 64
rectangle, 42
regular hexagon, 79
regular pentagon, 80
regular polygons, 7, 79
reparametrization of a curve, 61
rhombus, 45
right angle, 29
right triangle, 29
right triangle in a circle, 47
rigid motion, 17

Sim, 39
similarity, 4, 39
sin, 5, 64
sinh, 101
sphere, 116
square, 43
surface area, 116

tan, 75
tangent, 52
tanh, 103
three circles problem, 104
triangle, 3, 26
triangle inequality, 3, 15, 108, 110
trigonometric functions, 73
trisecting an angle, 82

vector space, 109
volume, 10, 115
volumes of balls, 118

wedge, 24