Lecture notes on multidimensional complex analysis

Harold P. Boas

Draft of January 6, 2020

Contents

1 Introduction

Karl Weierstrass (1815–1897), a giant of nineteenth-century analysis, took some steps toward a theory of functions of several complex variables. The "Weierstrass preparation theorem" and the "Weierstrass division theorem" are named after him. The modern theory of complex analysis in dimension 2 (and in higher dimensions) can reasonably be dated to the researches of Friedrich (Fritz) Hartogs (1874–1943) in the first decade of the twentieth century.[1] The so-called Hartogs Phenomenon, a fundamental property unrecognized by Weierstrass, reveals a dramatic difference between one-dimensional complex analysis and multidimensional complex analysis.

Some properties of holomorphic (complex analytic) functions, such as the maximum principle, are essentially the same in all dimensions. The most striking parts of the higher-dimensional theory are the features that differ from the one-dimensional theory.

Several complementary points of view illuminate the one-dimensional theory: power series expansions, integral representations, partial differential equations, and geometry. The higher-dimensional theory reveals new phenomena from each point of view. This chapter sketches some of the issues to be treated in detail later on.

A glance at the titles of articles listed in MathSciNet with primary classification number 32 (more than twenty-five thousand articles) or at postings in math.CV in the arXiv indicates the broad scope of interaction between complex analysis and other parts of mathematics, including algebra, functional analysis, geometry, mathematical physics, partial differential equations, and probability. One of the goals of this exposition is to glimpse some of these connections between different areas of mathematics.

1.1 A note on terminology

The unfortunate but standard name of the subject is "several complex variables" (abbreviated sometimes as "SCV"). The objects of study, of course, are not the variables but the functions. One might view the traditional terminology "several complex variables" as an elision of the more exact phrase "functions of several complex variables."

[1] A student of Alfred Pringsheim (1850–1941), Hartogs belonged to the Munich school of mathematicians. Because of their Jewish heritage, both Pringsheim and Hartogs suffered greatly under the Nazi regime in the 1930s. Pringsheim, a wealthy man, managed to buy his way out of Germany into Switzerland, where he died at an advanced age. The situation for Hartogs, however, grew ever more desperate, and he ultimately chose suicide rather than transportation to a death camp.

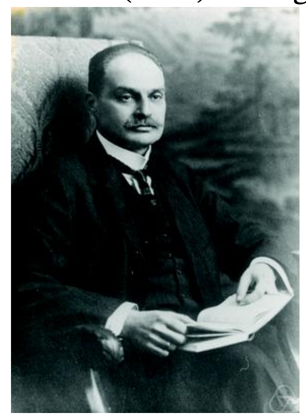
Karl Weierstrass

public domain photo

Alfred Pringsheim

source: *Brockhaus Enzyklopädie*

Friedrich (Fritz) Hartogs



Oberwolfach Photo Collection
Photo ID: 1567

In standard English, the common meaning of the word "several" is "more than two, but not many." In the setting of complex analysis, the big contrast is between functions of one variable and functions of two (or more) variables, so "several" is not really the right word. Nor will "multiple" or "many" fit the bill. One could conceivably say "functions of a plural number of complex variables," but this expression sounds weird.

The same problem exists in other languages. Both French and German use words equivalent to "several" to express "several complex variables," and Russian uses a word that translates as "many." A grammatical feature common to many languages is the inflection of nouns according to whether they are singular or plural,[2] yet there seems to be a paucity of words meaning "an unspecified number greater than one." Most terms conveying the notion of indefinite plurality have the connotation of "more than two," perhaps because every language has a dedicated word for the notion of precisely two.[3]

In English legal jargon, the word "several" does mean "two or more." One must accept that this technical meaning holds also in complex analysis: "several" means "n, where $n \geq 2$." The same principle applies to prefixes such as "multi" and "poly."

[2] Some languages, such as ancient Greek and modern Arabic, adjust the form of nouns according to whether they are singular, dual, or plural (in this context, meaning more than two).

[3] There is some debate about the extent to which numeration is innate in human culture. The aboriginal Australian language Warlpiri is often cited as an example of a language having the counting system "one, two, several, many." But modern Warlpiri, influenced by contact with the outside world, does have a full counting system.

1.2 Power series

A power series in one complex variable converges absolutely inside a certain disc[4] and diverges outside the closure of the disc. But the convergence region for a power series in two (or more) variables can have many different shapes (indeed, infinitely many). The largest open set in $\mathbb{C}^2$ in which the double series $\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} z_1^n z_2^m$ converges absolutely is the unit bidisc $\{(z_1, z_2) : |z_1| < 1 \text{ and } |z_2| < 1\}$; the series $\sum_{n=0}^{\infty} z_1^n z_2^n$ converges in the unbounded hyperbolic region where $|z_1 z_2| < 1$.

The theory of one-dimensional power series bifurcates into the case of entire functions (when the series has infinite radius of convergence) and the case of holomorphic functions on the unit disc (when the series has a finite radius of convergence—which might as well be normalized to the value 1). In dimensions greater than 1, the study of power series leads to function theory on infinitely many different types of domains. A natural problem, to be solved later, is to characterize the domains in $\mathbb{C}^n$ that are convergence domains for multivariable power series.

Exercise 1. Find a concrete power series whose domain of absolute convergence is the two-dimensional unit ball $\{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 < 1\}$.

While studying double power series, Hartogs discovered that every function holomorphic in a neighborhood of the boundary of the unit bidisc automatically extends to be holomorphic on the interior of the bidisc; a proof can be carried out by considering single-variable Laurent series on slices. Thus, in dramatic contrast to the situation in one variable, there are domains in $\mathbb{C}^2$ on which all the holomorphic functions extend to a larger domain. A natural question, to be answered later, is to characterize the *domains of holomorphy*, that is, the natural domains of existence of holomorphic functions.

The discovery of Hartogs shows additionally that holomorphic functions of several variables never have isolated singularities and never have isolated zeros, in contrast to the one-variable case. Moreover, zeros (and singularities) must propagate either to infinity or to the boundary of the domain on which the function is defined.

1.3 Integral representations

The one-variable Cauchy integral formula for a holomorphic function f on a domain bounded by a simple closed curve C says that[5]

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(w)}{w - z}\,dw \qquad \text{for } z \text{ inside } C.$$

[4] Researchers in the field of one-dimensional complex analysis usually use the spelling "disk," and researchers on higher-dimensional complex analysis traditionally use the spelling "disc."

[5] A prerequisite for reading further is that you can supply precise hypotheses for validity of this statement.

A remarkable feature of this formula is that the kernel $(w - z)^{-1}$ is both universal (independent of the curve C) and holomorphic in the free variable z. There is no such formula in higher dimensions! There are integral representations with a holomorphic kernel that depends on the domain, and there is a universal integral representation with a kernel that is not holomorphic. A huge literature addresses the problem of constructing and analyzing integral representations for various special types of domains.

There is an iterated Cauchy integral formula, namely,

$$f(z_1, z_2) = \left(\frac{1}{2\pi i}\right)^2 \int_{C_1} \int_{C_2} \frac{f(w_1, w_2)}{(w_1 - z_1)(w_2 - z_2)}\, dw_1\, dw_2$$

for z_1 in the region D_1 bounded by the simple closed curve C_1 and z_2 in the region D_2 bounded by the simple closed curve C_2. But this formula is special to a product domain $D_1 \times D_2$. Moreover, the integration here is over only a small portion of the boundary of the region, for the set $C_1 \times C_2$ (the "distinguished boundary") has real dimension 2, but the topological boundary of $D_1 \times D_2$ has real dimension 3. The iterated Cauchy integral is important and useful within its limited realm of applicability.

Exercise 2. Show that a function represented by the iterated Cauchy integral on the bidisc can be expanded inside the bidisc in an absolutely convergent double power series.

1.4 Partial differential equations

The Cauchy–Riemann equations for functions of one complex variable are a pair of real partial differential equations for a pair of real functions (namely, the real part and the imaginary part of a holomorphic function). In $\mathbb{C}^n$, there are still two functions (the real part and the imaginary part), but there are $2n$ equations (two equations for each of the n complex variables). Thus when $n > 1$, the inhomogeneous Cauchy–Riemann equations form an overdetermined system; there is a necessary compatibility condition for solvability of the Cauchy–Riemann equations. This feature is a significant difference from the one-variable theory.

When the inhomogeneous Cauchy–Riemann equations are solvable in $\mathbb{C}^2$ (or in $\mathbb{C}^n$ for some value of n larger than 1), there is (as will be shown later) a solution with compact support in the case of compactly supported data. When $n = 1$, however, it is not always possible to solve the inhomogeneous Cauchy–Riemann equations while maintaining compact support. The Hartogs phenomenon can be interpreted as one manifestation of this dimensional difference.

Exercise 3. Show that if u is the real part of a holomorphic function of two complex variables z_1 $(= x_1 + iy_1)$ and z_2 $(= x_2 + iy_2)$, then the function u must satisfy the following system of real second-order partial differential equations:

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial y_1^2} = 0, \qquad \frac{\partial^2 u}{\partial x_1 \partial x_2} + \frac{\partial^2 u}{\partial y_1 \partial y_2} = 0,$$
$$\frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial y_2^2} = 0, \qquad \frac{\partial^2 u}{\partial x_1 \partial y_2} - \frac{\partial^2 u}{\partial y_1 \partial x_2} = 0.$$

In other words, the real part of a holomorphic function of two complex variables not only is harmonic in each coordinate but also satisfies additional conditions that see an interaction between the two variables.

1.5 Geometry

The one-variable Riemann mapping theorem says that every simply connected planar domain other than $\mathbb{C}$ itself is biholomorphically equivalent to the unit disc. In higher dimension, no purely topological classification of biholomorphically equivalent domains can exist. Indeed, the unit ball in $\mathbb{C}^2$ and the unit bidisc in $\mathbb{C}^2$ are holomorphically inequivalent domains (as will be proved later).

An intuitive way to understand why the situation changes in dimension 2 is to realize that in $\mathbb{C}^2$, there is room for one-dimensional complex analysis to take place in the tangent space to the boundary of a domain. Indeed, the boundary of the bidisc contains pieces of one-dimensional affine complex subspaces, but the boundary of the two-dimensional ball does not contain any nontrivial *analytic disc* (the image of the unit disc under a holomorphic mapping).

Similarly, there is room for complex analysis to happen inside the zero set of a holomorphic function from $\mathbb{C}^2$ to $\mathbb{C}^1$. The zero set of a function such as $z_1 z_2$ is a one-dimensional complex variety inside $\mathbb{C}^2$. From this point of view, the reason that zeros of one-variable holomorphic functions are isolated is that the zero set of a nontrivial holomorphic function from $\mathbb{C}^1$ to $\mathbb{C}^1$ is a zero-dimensional variety (namely, a discrete set of points).

Notice that there is a mismatch between the dimension of the domain and the dimension of the range of a multivariable holomorphic function. Accordingly, one might expect the right analogue of a holomorphic function from $\mathbb{C}^1$ to $\mathbb{C}^1$ to be an equidimensional holomorphic *mapping* from $\mathbb{C}^n$ to $\mathbb{C}^n$. But here too there are surprises.

First of all, notice that biholomorphic mappings in dimension 2 (and higher) generally are not conformal (that is, angle-preserving). Even a linear transformation of $\mathbb{C}^2$, such as the mapping that sends (z_1, z_2) to $(z_1 + z_2, z_2)$, can change the angles at which lines meet.[6] Although conformal maps are plentiful in the setting of one complex variable, conformality is a quite rigid property in higher dimensions. A theorem of Joseph Liouville (1809–1882)

[6] Accordingly, biholomorphic mappings used to be called "pseudoconformal" mappings, but this word has gone out of fashion.

says[7] that when $n \geq 3$, the only conformal mappings from a domain in $\mathbb{R}^n$ into $\mathbb{R}^n$ are the (restrictions of) Möbius transformations: compositions of translations, dilations, orthogonal linear transformations, and inversions.

Remarkably, there exist biholomorphic mappings from all of $\mathbb{C}^2$ onto proper subsets of $\mathbb{C}^2$. There is active current research on such mappings, called *Fatou–Bieberbach maps*.[8]

[7] In 1850, Liouville published a fifth edition of *Application de l'analyse à la géométrie* by Gaspard Monge (1746–1818). An appendix includes seven long notes by Liouville. The sixth of these notes, bearing the title "Extension au cas des trois dimensions de la question du tracé géographique" and extending over pages 609–616, contains the proof of the theorem in dimension 3.

Two sources for modern treatments of this theorem are Chapters 5–6 of David E. Blair's *Inversion Theory and Conformal Mapping* [American Mathematical Society, 2000]; and Theorem 5.2 of Chapter 8 of Manfredo Perdigão do Carmo's *Riemannian Geometry* [Birkhäuser, 1992].

[8] The name recognizes constructions by the French mathematician Pierre Fatou (1878–1929), known also for the Fatou lemma in the theory of the Lebesgue integral; and by the German mathematician Ludwig Bieberbach (1886–1982), known also for the Bieberbach Conjecture about schlicht functions, posed in 1916 but proved much later (around 1984, by Louis de Branges). Bieberbach is infamous for having been an enthusiastic member of the Nazi Party in the 1930s.

2 Power series

Examples in the introduction show that domains of absolute convergence of multivariable power series can have a variety of shapes; in particular, the domain of convergence need not be a convex set. Nonetheless, there is a special kind of convexity property that does characterize convergence domains.

Developing the theory requires some notation. The Cartesian product of n copies of $\mathbb{C}$, the complex numbers, is denoted by $\mathbb{C}^n$. In contrast to the one-dimensional case, the space $\mathbb{C}^n$ is not an algebra when $n > 1$ (there is no multiplication operation). But the space $\mathbb{C}^n$ is a normed vector space, the usual norm being the Euclidean one:

$$\|(z_1, \dots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}\,.$$

A point $(z_1, \dots, z_n)$ in $\mathbb{C}^n$ is commonly denoted by a single letter z, a vector variable. When α is an n-dimensional vector all of whose coordinates are nonnegative integers, the symbol z^α means the product $z_1^{\alpha_1} \cdots z_n^{\alpha_n}$ (the quantity $z_1^{\alpha_1}$ being interpreted as 1 when z_1 and α_1 are simultaneously equal to 0). The notation $\alpha!$ abbreviates the product $\alpha_1! \cdots \alpha_n!$ (where $0! = 1$), and $|\alpha|$ means $\alpha_1 + \cdots + \alpha_n$. Thus $|\alpha|$ is the "total degree" of the monomial z^α. In this "multi-index" notation, a multivariable power series can be written in the form $\sum_\alpha c_\alpha z^\alpha$, which is an abbreviation for $\sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_n=0}^{\infty} c_{\alpha_1,\dots,\alpha_n} z_1^{\alpha_1} \cdots z_n^{\alpha_n}$.

There is some awkwardness in talking about convergence of a multivariable power series $\sum_\alpha c_\alpha z^\alpha$, because the value of a series depends (in general) on the order of summation, and there is no canonical ordering of n-tuples of nonnegative integers when $n > 1$.

Exercise 4. When $n = 2$, find complex numbers b_α such that the "square" sum

$$\lim_{k\to\infty} \sum_{\alpha_1=0}^{k} \sum_{\alpha_2=0}^{k} b_\alpha$$

and the "triangular" sum

$$\lim_{k\to\infty} \sum_{j=0}^{k} \sum_{|\alpha|=j} b_\alpha$$

have distinct finite values. (There is no power series in this exercise, but merely a series of complex numbers.)

Accordingly, it is convenient to restrict attention to *absolute* convergence, since the terms of an absolutely convergent series can be reordered arbitrarily without changing the value of the sum (or the convergence of the sum).

2.1 Domain of convergence

The *domain of convergence* of a power series means the interior of the set of points at which the series converges absolutely (that is, the largest open set on which the series converges absolutely). For example, the power series $\sum_{n=1}^{\infty} z_1^n z_2^{n!}$ converges absolutely on the union of three sets in $\mathbb{C}^2$: the points (z_1, z_2) for which $|z_2| < 1$ and z_1 is arbitrary; the points $(0, z_2)$ for arbitrary z_2; and the points (z_1, z_2) for which $|z_2| = 1$ and $|z_1| < 1$. The domain of convergence is the first of these three sets, for the other two sets contribute no additional interior points.

Being defined by absolute convergence, every convergence domain is *multicircular*: if a point $(z_1, \dots, z_n)$ lies in the domain, then so does the point $(\lambda_1 z_1, \dots, \lambda_n z_n)$ when $1 = |\lambda_1| = \cdots = |\lambda_n|$. Moreover, the comparison test for absolute convergence of series shows that the point $(\lambda_1 z_1, \dots, \lambda_n z_n)$ remains in the convergence domain when $|\lambda_j| \le 1$ for each j. Thus *every convergence domain can be expressed as a union of polydiscs centered at the origin*. (A polydisc means a Cartesian product of discs, possibly with different radii.)

Multicircular domains are often called *Reinhardt domains*. The name honors the German mathematician Karl Reinhardt[1] (1895–1941). A Reinhardt domain is said to be *complete* if whenever a point z lies in the domain, the whole polydisc $\{ w : |w_1| \le |z_1|, \dots, |w_n| \le |z_n| \}$ is contained in the domain. The conclusion of the preceding paragraph can be rephrased as saying that every convergence domain is a complete Reinhardt domain.

Convergence domains have an additional important property. If $\sum_\alpha |c_\alpha z^\alpha|$ converges, and $\sum_\alpha |c_\alpha w^\alpha|$ converges too, then Hölder's inequality implies that $\sum_\alpha |c_\alpha| |z^\alpha|^t |w^\alpha|^{1-t}$ converges when $0 \le t \le 1$. Indeed, the numbers $1/t$ and $1/(1-t)$ are conjugate indices for Hölder's inequality: the sum of their reciprocals evidently equals 1. Phrased in words, this deduction from Hölder's inequality says that if two points z and w lie in a convergence domain, then so does the point obtained by forming in each coordinate the geometric average of the moduli, with weights t and $1-t$. This property of a Reinhardt domain is called *logarithmic convexity*. Since a convergence domain is complete and multicircular, the domain is determined by the points with positive real coordinates; replacing the coordinates of each such point by their logarithms produces a convex domain in $\mathbb{R}^n$.

A special case of little practical importance is the empty set, which vacuously is complete and logarithmically convex. The series $\sum_{n=1}^{\infty} n!\, z_1^n z_2^n$ converges on no subset of $\mathbb{C}^2$ having

---

[1] A student of Ludwig Bieberbach (who solved the first part of Hilbert's 18th problem in 1910), Reinhardt has a place in mathematical history for solving the second part of Hilbert's 18th problem in 1928: he found a polyhedron that tiles three-dimensional Euclidean space but is not the fundamental domain of any group of isometries of $\mathbb{R}^3$. In other words, there is no group of motions such that the orbit of the polyhedron under the group covers $\mathbb{R}^3$, yet non-overlapping isometric images of the tile do cover $\mathbb{R}^3$. Later, Heinrich Heesch (1906–1995) found a two-dimensional example; Heesch is remembered too for developing computer methods to attack the four-color problem.

The date of Reinhardt's death does not mean that he was a war casualty: his obituary says to the contrary that he died after a long illness of unspecified nature. The University of Greifswald, where Reinhardt was a professor, is one of the oldest universities in Europe, having been founded in 1456. Located in northeastern Germany on the Baltic Sea, the city of Greifswald is a sister city of Bryan–College Station.

Karl Reinhardt

source: obituary

Greifswald

source: worldatlas.com

interior points, so the empty set is a convergence domain.

2.2 Characterization of domains of convergence

According to Section 2.1, every convergence domain is necessarily a Reinhardt domain that is both complete and logarithmically convex. The following theorem, obtained independently when $n = 2$ by Faber[2] and by Hartogs[3] in their respective habilitation theses,[4] states that this geometric property characterizes domains of convergence of power series.

Theorem 1. *A complete Reinhardt domain in* $\mathbb{C}^n$ *is the domain of convergence of some power series* $\sum_\alpha c_\alpha z^\alpha$ *if and only if the domain is logarithmically convex.*

Exercise 5. If D_1 and D_2 are convergence domains, are the intersection $D_1 \cap D_2$, the union $D_1 \cup D_2$, and the Cartesian product $D_1 \times D_2$ necessarily convergence domains too?

Proof of Theorem 1. When $n = 1$, the equivalence is a triviality. A complete Reinhardt domain in $\mathbb{C}^1$ is either a disc or the whole plane, hence is automatically logarithmically convex. Moreover, discs are convergence regions for geometric series, and the whole plane is the

[2] Georg Faber, Über die zusammengehörigen Konvergenzradien von Potenzreihen mehrerer Veränderlicher, *Mathematische Annalen* **61** (1905) 289–324. Faber (1877–1966) and Hartogs were PhD students of Alfred Pringsheim in Munich at the same time. Faber graduated in 1902, one year ahead of Hartogs, and went off to Würzburg. Faber's paper was finished in late 1904 and published in 1905.

[3] Fritz Hartogs, Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten, *Mathematische Annalen* **62** (1906) 1–88. This paper was finished in January 1905 and published in 1906. In a note added in page proof, Hartogs acknowledges the work of Faber, which had not yet appeared when Hartogs submitted his article. Both papers contain much more than the indicated theorem, and the two articles have a large symmetric difference.

[4] The habilitation, traditional in Europe, is a step beyond the doctoral dissertation.

convergence region for any entire function, such as e^z. Thus the two properties to be shown equivalent are universally satisfied for complete Reinhardt domains in $\mathbb{C}^1$. The following discussion therefore assumes implicitly that $n \geq 2$.

The part needing proof is the sufficiency: for every complete and logarithmically convex Reinhardt domain D, there exists some power series $\sum_\alpha c_\alpha z^\alpha$ whose domain of convergence is precisely D. The idea is to construct a power series that can be compared to a suitable geometric series.

Suppose first that the domain D is bounded (and nonvoid), for the construction is simpler to implement in this case. Let $N_\alpha(D)$ denote $\sup\{\, |z^\alpha| : z \in D \,\}$, the supremum norm on D of the monomial with exponent α. The quantity $N_\alpha(D)$ is finite under the hypothesis that D is bounded. The claim now is that $\sum_\alpha z^\alpha / N_\alpha(D)$ is the required power series whose domain of convergence is equal to D. What needs to be checked is that for each point w inside D, the series converges absolutely at w, and for each point w outside D, there is no neighborhood of w throughout which the series converges absolutely.

If w is a particular point in the interior of D, then there is a positive ε (depending on w) such that the scaled point $(1+\varepsilon)w$ still lies in D. Therefore $(1+\varepsilon)^{|\alpha|}|w^\alpha| \leq N_\alpha(D)$, so the series $\sum_\alpha w^\alpha / N_\alpha(D)$ converges absolutely by comparison with the convergent dominating series $\sum_\alpha (1+\varepsilon)^{-|\alpha|}$ (which is a product of n copies of $\sum_{k=0}^{\infty} (1+\varepsilon)^{-k}$, a convergent geometric series). Thus the first requirement is met.

Checking the second requirement involves showing that the series diverges at sufficiently many points outside D. The following argument demonstrates that $\sum_\alpha w^\alpha / N_\alpha(D)$ diverges whenever w is a point outside the closure of D whose coordinates are positive real numbers. Since the domain D is multicircular, this conclusion suffices. The strategy is to show that infinitely many terms of the series are greater than 1.

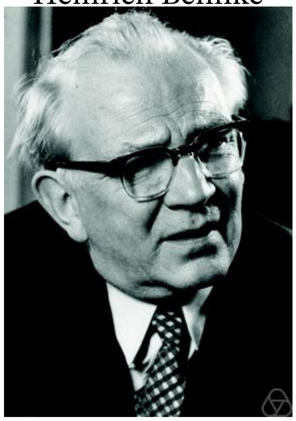
The hypothesis that D is logarithmically convex means precisely that the set

$$\{\, (u_1, \dots, u_n) \in \mathbb{R}^n : (e^{u_1}, \dots, e^{u_n}) \in D \,\}, \qquad \text{denoted by } \log D,$$

is a convex set in $\mathbb{R}^n$. By assumption, the point $(\log w_1, \dots, \log w_n)$ is a point of $\mathbb{R}^n$ outside the closure of the convex set $\log D$, so this point is separated from $\log D$ by a hyperplane. In other words, there is a linear function $\ell : \mathbb{R}^n \to \mathbb{R}$ whose value at the point $(\log w_1, \dots, \log w_n)$ exceeds the supremum of ℓ over the convex set $\log D$. (In particular, that supremum is finite.) Say $\ell(u_1, \dots, u_n) = \beta_1 u_1 + \cdots + \beta_n u_n$, where each coefficient β_j is a real number.

Every complete Reinhardt domain contains a neighborhood of the origin in $\mathbb{C}^n$, so there is a positive real constant m such that the convex set $\log D$ contains every point u in $\mathbb{R}^n$ for which $\max_{1 \leq j \leq n} u_j \leq -m$. Therefore none of the numbers β_j is negative, for otherwise the function ℓ would take arbitrarily large positive values on the set $\log D$. The assumption that D is bounded guarantees the existence of a positive real constant M such that $\max_{1 \leq j \leq n} u_j \leq M$ whenever $u \in \log D$. Consequently, if each number β_j is increased by some small positive amount ε, then the supremum of ℓ over $\log D$ increases by no more than $nM\varepsilon$. Therefore the coefficients of the linear function ℓ can be perturbed slightly without affecting the separating property of ℓ. Accordingly, there is no loss of generality in assuming that each β_j is a positive

Heinrich Behnke



Oberwolfach Photo Collection
Photo ID: 278

Karl Stein

Oberwolfach Photo Collection
Photo ID: 8848
© Gerd Fischer

rational number. Multiplying by a common denominator shows that the coefficients β_j can be taken to be positive integers.

Exponentiating reveals that $w^\beta > N_\beta(D)$ for the particular multi-index β just determined. (Since the coordinates of w are positive real numbers, no absolute-value signs are needed on the left-hand side of the inequality.) Consequently, if k is a positive integer, and $k\beta$ denotes the multi-index $(k\beta_1, \dots, k\beta_n)$, then $w^{k\beta} > N_{k\beta}(D)$. Therefore the series $\sum_\alpha w^\alpha / N_\alpha(D)$ of positive numbers diverges, for there are infinitely many terms larger than 1. This conclusion completes the proof of the theorem when the domain D is bounded.

When D is unbounded, let D_r denote the intersection of D with the ball of radius r centered at the origin. Then D_r is a bounded, complete, logarithmically convex Reinhardt domain, and the preceding analysis applies to D_r. The natural idea of splicing together power series of the type just constructed for an increasing sequence of values of r is too simplistic, for none of these series converges throughout the unbounded domain D.

One way to finish the argument is to invoke a theorem of Heinrich Behnke (1898–1979) and his student Karl Stein (1913–2000), usually called *the* Behnke–Stein theorem, according to which an *increasing* union of domains of holomorphy is again a domain of holomorphy.[5] A coming attraction (see Section 2.5) is to prove that a convergence domain for a power series supports some (other) power series that cannot be analytically continued across any boundary point whatsoever. Hence each D_r is a domain of holomorphy, so the Behnke–Stein theorem implies that D is a domain of holomorphy. Accordingly, D supports some holomorphic function that cannot be analytically continued across any boundary point of D. This holomorphic function is represented by a power series that converges in all of D, and

[5]H. Behnke and K. Stein, Konvergente Folgen von Regularitätsbereichen und die Meromorphiekonvexität, *Mathematische Annalen* **116** (1938) 204–216. Stein is the eponym of so-called Stein manifolds. After the war, Behnke had several other notable students who became prominent mathematicians, including Hans Grauert (1930–2011), Friedrich Hirzebruch (1927–2012), and Reinhold Remmert (1930–2016).

D is the convergence domain of this power series.

The discussion in the preceding paragraph is unsatisfying: besides being anachronistic and not self-contained, the argument provides no concrete construction of the required power series. What follows is a modification of the construction for bounded domains that makes the argument work for unbounded domains too.

Replace $N_\alpha(D)$ by the quantity $M_\alpha(D)$ defined as

$$\sup\{\,|z^\alpha|\, e^{-(|z_1|+\cdots+|z_n|)} \,:\, z \in D\,\}. \tag{2.1}$$

The decaying exponential factor guarantees that $M_\alpha(D)$ is finite for each multi-index α, even when the domain D is unbounded. The claim now is that the series $\sum_\alpha z^\alpha/M_\alpha(D)$ has the logarithmically convex, complete Reinhardt domain D as domain of convergence.

When $w \in D$, the proof that $\sum_\alpha |w^\alpha|/M_\alpha(D)$ converges is basically the same as in the case of bounded domains. Indeed, there is a positive ε such that the dilated point $(1+\varepsilon)w \in D$, and

$$(1+\varepsilon)^{|\alpha|}|w^\alpha|\, e^{-(1+\varepsilon)(|w_1|+\cdots+|w_n|)} \le M_\alpha(D).$$

The exponential factor is independent of α, so $\sum_\alpha |w^\alpha|/M_\alpha(D)$ converges by comparison with the convergent geometric series $\sum_\alpha (1+\varepsilon)^{-|\alpha|}$.

Next suppose that w is a point outside D whose coordinates are positive real numbers. The following argument shows that there are infinitely many choices of the multi-index α for which

$$w^\alpha/M_\alpha(D) \ge \prod_{j=1}^{n} \min(1, w_j).$$

Since the right-hand side is positive and independent of α, the series $\sum_\alpha w^\alpha/M_\alpha(D)$ diverges.

The first step is the same as in the case of bounded domains. The logarithmic convexity of D implies the existence of a vector $(\beta_1, \dots, \beta_n)$ of nonnegative real numbers (at least one of them different from 0) such that

$$\prod_{j=1}^{n} |z_j|^{\beta_j} \le \prod_{j=1}^{n} w_j^{\beta_j} \qquad \text{whenever } z \in D. \tag{2.2}$$

The inequality evidently still holds when the vector $(\beta_1, \dots, \beta_n)$ is replaced by $(k\beta_1, \dots, k\beta_n)$ for an arbitrary positive integer k (or even by an arbitrary positive real number). But a difficulty arises in the unbounded case, for a small perturbation of the exponents could destroy the inequality. If $\log D$ is the region on one side of a hyperplane, for example, then the vector $(\beta_1, \dots, \beta_n)$ can only be a multiple of the normal vector to that hyperplane. So it may not be possible to replace the exponents with rational numbers, let alone integers.

The following device overcomes this difficulty by introducing integer exponents in a new way. When x is a real number, the notation $\lceil x \rceil$ means the ceiling of x, that is, the least integer greater than or equal to x. If γ is a real number between 0 and 1, and x is nonnegative, then

$x^\gamma e^{-x} \le (1+x)^\gamma e^{-x} \le (1+x)e^{-x} \le 1$. Applying this inequality with x equal to $|z_j|$ and γ equal to $\lceil \beta_j \rceil - \beta_j$ shows that

$$|z_j|^{\lceil \beta_j \rceil - \beta_j} e^{-|z_j|} \le 1, \qquad \text{or} \qquad |z_j|^{\lceil \beta_j \rceil} e^{-|z_j|} \le |z_j|^{\beta_j}.$$

Substituting into (2.2) yields that

$$\left(\prod_{j=1}^{n} |z_j|^{\lceil \beta_j \rceil}\right) e^{-(|z_1|+\cdots+|z_n|)} \le \prod_{j=1}^{n} w_j^{\beta_j} \qquad \text{whenever } z \in D. \tag{2.3}$$

Moreover, $w_j^{\lceil \beta_j \rceil} = w_j^{\lceil \beta_j \rceil - \beta_j} w_j^{\beta_j} \ge \min(1, w_j)^{\lceil \beta_j \rceil - \beta_j} w_j^{\beta_j} \ge \min(1, w_j) w_j^{\beta_j}$, so (2.3) implies that

$$\left(\prod_{j=1}^{n} |z_j|^{\lceil \beta_j \rceil}\right) e^{-(|z_1|+\cdots+|z_n|)} \le \frac{\prod_{j=1}^{n} w_j^{\lceil \beta_j \rceil}}{\prod_{j=1}^{n} \min(1, w_j)} \qquad \text{whenever } z \in D.$$

Consequently, if α equals the multi-index $(\lceil \beta_1 \rceil, \dots, \lceil \beta_n \rceil)$, then $w^\alpha / M_\alpha(D) \ge \prod_{j=1}^{n} \min(1, w_j)$.

The same conclusion holds when α equals $(\lceil k\beta_1 \rceil, \dots, \lceil k\beta_n \rceil)$, where k is an arbitrary positive integer. Therefore the series $\sum_\alpha w^\alpha / M_\alpha(D)$ diverges, since infinitely many of the terms exceed a fixed positive number. This deduction completes the proof that D is the domain of convergence of the series $\sum_\alpha z^\alpha / M_\alpha(D)$.

In summary, every complete and logarithmically convex Reinhardt domain—bounded or unbounded—is the domain of convergence of some power series, and a suitable series can be written down concretely.[6] □

Exercise 6. Every bounded, complete Reinhardt domain in $\mathbb{C}^2$ can be described as the set of points (z_1, z_2) for which

$$|z_1| < r \qquad \text{and} \qquad |z_2| < e^{-\varphi(|z_1|)},$$

where r is some positive real number, and φ is some nondecreasing, real-valued function. Show that such a domain is logarithmically convex if and only if the function sending z_1 to $\varphi(|z_1|)$ is a subharmonic function on the disk where $|z_1| < r$.

In solving this exercise, keep in mind the analogy between convex functions on the real line and subharmonic functions on the plane.[7] In the following remarks about these two classes of functions, the symbol u denotes a real-valued, upper semicontinuous function defined either on an open interval in $\mathbb{R}$ or on an open subset of $\mathbb{C}$ (equivalently $\mathbb{R}^2$).

[6] The original proofs of Faber and of Hartogs both involve choosing certain countable dense sets, so do not represent the series as explicitly.

[7] A foreshadowing of this analogy predates the invention of subharmonic functions by many years. See pages 43–44 of an article by O. Hölder, Ueber einen Mittelwerthssatz, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen* (1889) no. 2, 38–47. Page 44 of this paper is where the German mathematician Otto Hölder (1859–1937) derived "Hölder's inequality," although he points out that the inequality was previously obtained by L. J. Rogers, An extension of a certain theorem in inequalities, *Messenger of Mathematics* **17** (1888) 145–150 (page 149).

mean-value property A convex function has the property that the value at the midpoint of each subinterval of the domain is at most the average of the values of the function at the endpoints of the subinterval. In symbols, for every two points x_1 and x_2 in the domain,

$$u\left(\frac{x_1+x_2}{2}\right) \le \frac{u(x_1)+u(x_2)}{2}.$$

A subharmonic function has the property that for each closed disk in the domain of the function, the value at the center of the disk is at most the average of the values of the function around the boundary of the disk. In symbols, if the disk has center z_0 and radius r, then

$$u(z_0) \le \frac{1}{2\pi}\int_{-\pi}^{\pi} u(z_0+re^{i\theta})\,d\theta.$$

weighted averages If x_1 and x_2 are points in the domain of a convex function u, and $0 < t < 1$, then

$$u\left(tx_1+(1-t)x_2\right) \le tu(x_1)+(1-t)u(x_2).$$

More generally, if x_1, …, x_k are points in the domain of a convex function u, and a_1, …, a_k are positive numbers such that $a_1+\cdots+a_n=1$, then

$$u\left(a_1x_1+\cdots+a_kx_k\right) \le a_1u(x_1)+\cdots+a_ku(x_k).$$

This inequality for convex functions, due to Hölder,[8] was rediscovered by the Danish telephone engineer Johan Jensen[9] (1859–1925). Still more generally, if a is a positive integrable function such that $\int_0^1 a(x)\,dx=1$, and f is an integrable function, then

$$u\left(\int_0^1 a(x)f(x)\,dx\right) \le \int_0^1 a(x)u(f(x))\,dx.$$

This statement is Jensen's inequality.[10] The modern formulation is that when ν is a probability measure (total mass 1), and f is an integrable function,

$$u\left(\int f\,d\nu\right) \le \int u(f)\,d\nu.$$

In the setting of a disk, there is a natural weight whose integral over the boundary is equal to 1, namely, the Poisson kernel. Integrating some function against the Poisson kernel corresponds to forming a certain weighted average of the values of the function on the boundary. A subharmonic function has the property that the value at a point inside a disk is at most the value at the point of the Poisson integral of the boundary value of the function.

[8] Announced on page 39 of Hölder's previously cited paper, the inequality is the main result of that paper.

[9] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Mathematica* **30** (1906) 175–193. See page 180.

[10] See page 186 of the cited article of Jensen.

maximum principle A geometric interpretation of the convexity inequality is that the graph of a convex function lies below (or possibly on) each chord. An equivalent statement is that if ℓ is an arbitrary affine linear function (that is, $\ell(x)$ has the form $ax + b$), then the function $u - \ell$ cannot have a (weak) local maximum unless $u - \ell$ is constant in a neighborhood of the point where the maximum occurs.

A subharmonic function has the parallel property that if v is an arbitrary harmonic function on an arbitrary disk contained in the domain of u, then the function $u - v$ cannot have a (weak) local maximum in the disk unless $u - v$ is constant in a neighborhood of the point where the maximum occurs.

second derivatives If u is twice continuously differentiable, then u is convex if and only if the second derivative of u is nonnegative. Notice that functions with vanishing second derivative are the comparison functions in the maximum principle for convexity.

If the function u is twice continuously differentiable, then u is subharmonic if and only if the Laplacian of u is nonnegative. Notice that the functions with vanishing Laplacian are the comparison functions in the maximum principle for subharmonicity.

locality A function is convex on an open interval if and only if the function is convex on some neighborhood of each point of the interval. A function is subharmonic on an open set if and only if the function is subharmonic on a neighborhood of each point of the open set.

continuity An upper semicontinuous convex function is automatically continuous.[11] Now the analogy breaks down: the corresponding statement for subharmonic functions is *not* true. For instance, the function $\log|z|$ is subharmonic on the whole plane $\mathbb{C}$ but is not continuous at the origin.

Aside on infinite dimensions

The story changes when the finite-dimensional vector space $\mathbb{C}^n$ is replaced by an infinite-dimensional space. Consider, for example, the power series $\sum_{j=1}^{\infty} z_j^j$ depending on infinitely many variables $z_1, z_2, \ldots$. Where does this series converge?

Finitely many of the variables can be arbitrary, and the series will certainly converge if the remaining variables have absolute value less than a fixed number smaller than 1. On the other hand, the series will diverge if the variables do not eventually have absolute value less than 1. In particular, in the product of countably infinitely many copies of $\mathbb{C}$, there is no open set (with respect to the product topology) on which the series converges. (A basis for open sets in the product topology consists of sets for which each of finitely many variables is restricted to an open subset of $\mathbb{C}$, the remaining variables being unrestricted.) Holomorphic

[11] See page 189 of the cited article of Jensen.

functions ought to live on open sets, so apparently this power series in infinitely many variables does not represent a holomorphic function, even though the series converges at many points.

Perhaps an infinite-product space is not the right setting for this power series. The series could be considered instead on the Hilbert space of sequences $(z_1, z_2, \dots)$ for which $\sum_{j=1}^{\infty} |z_j|^2$ is finite. In this setting, the power series converges everywhere: the square-summability implies that $z_j \to 0$ when $j \to \infty$, so $|z_j^j|$ eventually is dominated by $1/2^j$. Similar reasoning shows that the power series converges uniformly on every ball of radius less than 1 (with an arbitrary center). Consequently, the series converges uniformly on every compact set. Yet the power series fails to converge uniformly on the closed unit ball centered at the origin (as follows by considering the standard unit basis vectors). In a finite-dimensional space, a series that is everywhere absolutely convergent must converge uniformly on every ball of every radius, but this convenient property breaks down when the dimension is infinite.

The preceding remarks indicate that the theory of holomorphic functions needs to be rethought when the dimension is infinite.[12] Two noteworthy changes in infinite dimension are the existence of inequivalent norms (all norms on a finite-dimensional vector space are equivalent) and the nonexistence of interior points of compact sets (closed balls are never compact in infinite-dimensional Banach spaces).

Incidentally, the notion of convergence of infinite series in Banach spaces involves some subtleties. In finite dimensions, the concepts of absolute convergence and unconditional convergence are equivalent; when the dimension is infinite, absolute convergence implies unconditional convergence but not conversely. For example, let e_n denote the nth unit basis element in the space of square-summable sequences (all entries of e_n are equal to 0 except the nth one, which equals 1), and consider the infinite series $\sum_{n=1}^{\infty} \frac{1}{n} e_n$. This series converges unconditionally (in other words, without regard to the order of summation) to the square-summable sequence $(1, \frac{1}{2}, \frac{1}{3}, \dots)$, yet the series fails to converge absolutely (since the sum of the norms of the terms is the divergent harmonic series). A famous theorem[13] due to Aryeh Dvoretzky (1916–2008) and C. Ambrose Rogers (1920–2005) says that this phenomenon is general: in every infinite-dimensional Banach space, there is an unconditionally convergent series $\sum_{n=1}^{\infty} x_n$ such that $\|x_n\| = 1/n$ (whence the series fails to converge absolutely).

2.3 Local properties of holomorphic functions

Convergent power series are local models for holomorphic functions. Power series converge uniformly on compact sets, so they represent continuous functions that are holomorphic in each variable separately (when the other variables are held fixed). Thus a reasonable working

[12] One book on the subject is Jorge Mujica's *Complex Analysis in Banach Spaces*, originally published by North-Holland in 1986 and reprinted by Dover in 2010.

[13] A. Dvoretzky and C. A. Rogers, Absolute and unconditional convergence in normed linear spaces, *Proceedings of the National Academy of Sciences of the United States of America* **36** (1950) 192–197.

definition of a holomorphic function of several complex variables is a function (on an open set) that is holomorphic in each variable separately and continuous in all variables jointly.[14]

If D is a polydisc in $\mathbb{C}^n$ with center at the origin and with polyradius $(r_1, \dots, r_n)$, and if f is holomorphic on a neighborhood of the closure of D, then iterating the one-dimensional Cauchy integral formula shows that

$$f(z) = \left(\frac{1}{2\pi i}\right)^n \int_{|w_1|=r_1} \cdots \int_{|w_n|=r_n} \frac{f(w_1, \dots, w_n)}{(w_1 - z_1) \cdots (w_n - z_n)} \, dw_1 \cdots dw_n$$

when the point z with coordinates $(z_1, \dots, z_n)$ is in the interior of the polydisc. The assumed continuity of f guarantees that this iterated integral makes sense and can be evaluated in any order by Fubini's theorem.

Expanding the Cauchy kernel in a geometric series, just as in the one-variable case, shows that $f(z)$ admits a power series expansion $\sum_\alpha c_\alpha z^\alpha$ that converges in the (open) polydisc. The coefficient c_α is uniquely determined as

$$\left(\frac{1}{2\pi i}\right)^n \int_{|w_1|=r_1} \cdots \int_{|w_n|=r_n} \frac{f(w_1, \dots, w_n)}{w_1^{1+\alpha_1} \cdots w_n^{1+\alpha_n}} \, dw_1 \cdots dw_n,$$

or equivalently $f^{(\alpha)}(0)/\alpha!$, where the symbol $f^{(\alpha)}$ abbreviates the derivative $\partial^{|\alpha|} f / \partial z_1^{\alpha_1} \cdots \partial z_n^{\alpha_n}$. Placing the center of the polydisc at the origin is merely a notational convenience. If instead the polydisc has center ζ, then the series expansion has the form $\sum_\alpha c_\alpha (z - \zeta)^\alpha$, where now $c_\alpha = f^{(\alpha)}(\zeta)/\alpha!$.

Every complete Reinhardt domain is a union of concentric polydiscs, so the uniqueness of the coefficients c_α implies that every holomorphic function in a complete Reinhardt domain admits a power series expansion that converges in the whole domain. Thus holomorphic functions and convergent power series are identical notions in complete Reinhardt domains.

The iterated Cauchy integral formula can be used to establish the basic local properties of holomorphic functions by the same arguments as in the single-variable case. For example, holomorphic functions are infinitely differentiable, satisfy the Cauchy–Riemann equations in each variable, and obey a local maximum principle. The multivariable Cauchy estimates for derivatives say that if f is holomorphic on a polydisc of polyradius $(r_1, \dots, r_n)$, and if $|f|$ is bounded above by a constant M in the polydisc, then

$$|f^{(\alpha)}(\text{center})| \le \frac{\alpha! \, M}{r^\alpha}.$$

Holomorphic functions of several variables satisfy an identity principle, but the statement is different from the standard single-variable formulation. In dimension 1, an accumulation point of zeros forces a holomorphic function to be identically zero, but in higher dimension, zeros are never isolated. A statement valid in all dimensions is that if a holomorphic function on a connected open set is identically equal to 0 on some (small) ball or polydisc, then the

[14] A remarkable theorem of Hartogs states that the continuity hypothesis is superfluous. See Section 2.7.

function is identically equal to 0. To prove this statement via a connectedness argument, consider one-dimensional slices to see that if a holomorphic function is identically equal to 0 in a neighborhood of a point, then the function is identically equal to 0 in every polydisc centered at the point and contained in the domain of the function. And every two points in the domain can be joined by a chain of overlapping polydiscs.

The iterated Cauchy integral also implies that if a sequence of holomorphic functions converges normally (uniformly on compact sets), then the limit function is holomorphic. Indeed, the conclusion is a local property that can be checked on small polydiscs, and the locally uniform convergence implies that the limit of the iterated Cauchy integrals equals the iterated Cauchy integral of the limit function. On the other hand, the one-variable integral that counts zeros inside a curve lacks an obvious multivariable analogue (since zeros are not isolated), so a special argument is needed to verify that Hurwitz's theorem generalizes from one variable to several variables.

Exercise 7. Prove a multidimensional version of Hurwitz's theorem: On a connected open set, the normal limit of zero-free holomorphic functions is either zero-free or identically equal to zero.

2.4 The Hartogs phenomenon

So far the power series under consideration have been Maclaurin series. Studying Laurent series reveals a phenomenon of automatic analytic continuation, a discovery of Hartogs in his 1903 dissertation.[15]

Theorem 2 (Hartogs). *Suppose r is a positive number less than 1. If f is holomorphic in $\{\, (z_1, z_2) \in \mathbb{C}^2 : |z_1| < 1$ and $r < |z_2| < 1 \,\} \cup \{\, (z_1, z_2) : |z_2| < 1$ and $r < |z_1| < 1 \,\}$, then f extends (uniquely) to be holomorphic on the unit bidisc, $\{\, (z_1, z_2) : |z_1| < 1$ and $|z_2| < 1 \,\}$.*

The initial domain of definition of f is a Reinhardt domain, but not a *complete* Reinhardt domain. The theorem implies that if a function is holomorphic in a neighborhood of the boundary of a bidisc, then the function extends to be holomorphic in the whole bidisc. The result carries over to higher dimension with an analogous proof. Hartogs himself pointed out the important corollary that holomorphic functions of two (or more) complex variables cannot have isolated singularities.[16]

Proof. For each fixed z_1 in the unit disc, the function sending z_2 to $f(z_1, z_2)$ is holomorphic in the annulus where $r < |z_2| < 1$, so has a Laurent expansion valid in this annulus. In other

[15] *Beiträge zur elementaren Theorie der Potenzreihen und der eindeutigen analytischen Funktionen zweier Veränderlichen*, published in 1904 by Teubner. A scan of the publication can be found at Google Books. The theorem appears in §17 (page 55).

[16] "Als spezieller Fall ergibt sich daraus ohne weiteres, daß eine *eindeutige analytische Funktion $f(x, y)$ keine isolierten singulären Stellen besitzen kann*" [emphasis in original, page 55 of the dissertation].

words, for each integer j there is a coefficient $c_j(z_1)$ such that

$$f(z_1, z_2) = \sum_{j=-\infty}^{\infty} c_j(z_1) z_2^j \qquad \text{when } |z_1| < 1 \text{ and } r < |z_2| < 1. \tag{2.4}$$

Moreover, $c_j(z_1)$ has an integral representation. If s is a radius such that $r < s < 1$, then

$$c_j(z_1) = \frac{1}{2\pi i} \int_{|w_2|=s} \frac{f(z_1, w_2)}{w_2^{1+j}}\, dw_2. \tag{2.5}$$

When $|z_1| < 1$ and $|w_2| = s$, the function $f(z_1, w_2)$ is jointly continuous in both variables and holomorphic in z_1, so this integral representation shows (by Morera's theorem, say) that each coefficient $c_j(z_1)$ is a holomorphic function of z_1 in the unit disc.

But when $r < |z_1| < 1$, the function sending z_2 to $f(z_1, z_2)$ is holomorphic in the unit disk, so the Laurent series (2.4) reduces to a Taylor series. In other words, if $j < 0$, then $c_j(z_1)$ is identically equal to 0 when $r < |z_1| < 1$. By the one-variable identity theorem, the holomorphic function $c_j(z_1)$ remains identically 0 in the whole disc where $|z_1| < 1$. In other words, the Laurent series (2.4) reduces to a Taylor series for every value of z_1. This series, if uniformly convergent on compact subsets of $\{(z_1, z_2) : |z_1| < 1 \text{ and } |z_2| < s\}$, defines the required holomorphic extension of f.

To verify this normal convergence, fix an arbitrary compact subset K of the unit disc in the space of the variable z_1. The continuous function $|f(z_1, w_2)|$ has some finite upper bound M on the compact set where $z_1 \in K$ and $|w_2| = s$. Estimating the integral representation (2.5) for the series coefficient shows that $|c_j(z_1)| \le M/s^j$ when $z_1 \in K$. Consequently, if t is an arbitrary positive number less than s, then the series $\sum_{j=0}^{\infty} c_j(z_1) z_2^j$ converges absolutely when $z_1 \in K$ and $|z_2| \le t$ by comparison with the convergent geometric series $\sum_{j=0}^{\infty} M(t/s)^j$. Since the required locally uniform convergence holds, the series $\sum_{j=0}^{\infty} c_j(z_1) z_2^j$ does define the required holomorphic extension of f to the whole bidisc. □

The method can be adjusted to apply to more general geometry. Here is one example, which is the basic version of what is sometimes called the Kugelsatz ("sphere theorem") of Hartogs.

Exercise 8. If r is a positive radius less than 1, and f is a holomorphic function on the spherical shell $\{(z_1, z_2) \in \mathbb{C}^2 : r^2 < |z_1|^2 + |z_2|^2 < 1\}$, then f extends to be a holomorphic function on the whole unit ball.

The ultimate theorem of this type says that if D is an open subset of $\mathbb{C}^n$, where $n \ge 2$, and if K is a compact subset of D such that the set difference $D \setminus K$ is connected, and if f is holomorphic on $D \setminus K$, then f extends to be holomorphic on all of D. Roughly speaking, holomorphic functions of several variables extend across compact holes. This general result is not easy to prove directly using the tools available at this point in the exposition,[17] but a short proof will be possible later, after some additional theory is developed.

The preceding results can be viewed as demonstrating "internal" analytic continuation. Hartogs observed in his habilitation that "external" analytic continuation can occur too.

Theorem 3 (Hartogs). *Suppose r is a positive number less than 1. If f is holomorphic on $\{(z_1, z_2) \in \mathbb{C}^2 : |z_1| < r$ and $|z_2| < 1\} \cup \{(z_1, z_2) : |z_1| < 1$ and $1 - r < |z_2| < 1\}$, then f extends to be holomorphic on the unit bidisc, $\{(z_1, z_2) : |z_1| < 1$ and $|z_2| < 1\}$.*

The proof is no different from the proof of Theorem 2, and an analogous theorem holds in higher dimension. Biholomorphic images of regions of the form indicated in the hypothesis of the theorem are known as "Hartogs figures." Such theorems are known collectively as "the Hartogs phenomenon."

Exercise 9. Suppose D is a complete Reinhardt domain, and f is holomorphic on D. Show that f extends to be holomorphic on the smallest logarithmically convex complete Reinhardt domain containing D.

2.5 Natural boundaries

The one-dimensional power series $\sum_{k=0}^{\infty} z^k$ has the unit disc as domain of convergence, yet the function represented by the series, which equals $1/(1-z)$, extends holomorphically to $\mathbb{C} \setminus \{1\}$. On the other hand, there exist power series that converge in the unit disc and have the unit circle as "natural boundary," meaning that the function represented by the series does not continue analytically across any boundary point of the disc whatsoever. One concrete example is the gap series $\sum_{k=0}^{\infty} z^{2^k}$, which has an infinite radial limit at the boundary for a dense set of angles. More generally, the Hadamard gap theorem[18] says that if $\{n_k\}_{k=0}^{\infty}$ is an increasing sequence of natural numbers, if the series $\sum_{k=0}^{\infty} a_k z^{n_k}$ has radius of convergence equal to 1, and if there exists a positive number s such that $n_{k+1} \geq (1+s)n_k$ for every k, then the series has the unit circle as natural boundary.

Convergence regions for power series in two or more variables can have infinitely many possible shapes. Is every convergence domain (that is, every complete and logarithmically convex Reinhardt domain) the natural domain of existence of some holomorphic function? The following theorem[19] provides an affirmative answer.

[17] For a proof using only geometric tools, see Joël Merker and Egmont Porten, A Morse-theoretical proof of the Hartogs extension theorem, *Journal of Geometric Analysis* **17** (2007) no. 3, 513–546.

[18] The gap theorem of Jacques Hadamard (1865–1963) appears in his dissertation, Essai sur l'étude des fonctions données par leur développement de Taylor, *Journal de mathématiques pures et appliquées (4)* **VIII** (1892) 101–186. See page 116. The elegant standard proof is due to the famous number theorist Louis J. Mordell (1888–1972): On power series with the circle of convergence as a line of essential singularities, *Journal of the London Mathematical Society* **2** (1927) 146–148. One textbook where you can find an exposition of the proof is *Invitation to Complex Analysis* by Ralph P. Boas, second edition revised by Harold P. Boas, Mathematical

H. Cartan and Thullen
half a century later

Oberwolfach Photo Collection
Photo ID: 12444
photo by Burchard Kaup

Theorem 4 (Cartan–Thullen)**.** *The domain of convergence of a multivariable power series is a domain of holomorphy. In other words, for every domain of convergence there exists some power series that converges in the domain and that is singular at every boundary point.*

The word "singular" does not necessarily mean that the function blows up. To say that a power series is singular at a boundary point of the domain of convergence means that the series does not admit a direct analytic continuation to a neighborhood of the point. A function whose absolute value tends to infinity at a boundary point is singular at that point, but so is a function whose absolute value tends to zero at a nonpolynomial rate.

Here are two proofs of Theorem 4, both different from the original proof. The first proof constructs a noncontinuable multivariable gap series, an idea that goes back to Faber's 1905 habilitation. The second proof demonstrates the existence of many noncontinuable series without actually exhibiting one.

Proof of Theorem 4 using the Hadamard gap theorem. When the convergence domain D is the whole space $\mathbb{C}^n$, there is nothing to prove. So suppose that D is not the whole space. The complement of D then has nonvoid interior (since D is a complete Reinhardt domain). Choose a countable dense subset $\{w(k)\}_{k=1}^{\infty}$ of the interior of the complement of D such that

Association of America, 2010.

[19]Henri Cartan and Peter Thullen, Zur Theorie der Singularitäten der Funktionen mehrerer komplexen Veränderlichen: Regularitäts- und Konvergenzbereiche, *Mathematische Annalen* **106** (1932) number 1, 617–647. See Corollary 1 on page 637.

One of the leading mathematicians of the twentieth century, Henri Cartan (1904–2008) was a major force in the development of multidimensional complex analysis. His father was the influential mathematician Élie Cartan (1869–1951). Peter Thullen (1907–1996) collaborated with his teacher, Heinrich Behnke, on the first book about multidimensional complex analysis (*Theorie der Funktionen mehrerer komplexer Veränderlichen*, 1934; an updated version appeared in 1970). When Hitler came to power in 1933, Thullen left Germany on principle and subsequently emigrated to Ecuador. Later on, Thullen had a career in political economics and worked for the United Nations International Labour Organization.

the coordinates of each point $w(k)$ are nonzero. The reason for writing $w(k)$ instead of w_k is that the latter notation is reserved for the kth coordinate of a vector w. Thus $w(k) = (w_1(k), \dots, w_n(k))$.

Making use of the quantity $M_\alpha(D)$ defined in (2.1), the proof of Theorem 1 provides a sequence $\{\alpha(k)\}_{k=1}^\infty$ of multi-indices such that

$$\frac{|w(k)^{\alpha(k)}|}{M_{\alpha(k)}(D)} \geq \prod_{j=1}^n \min(1, w_j(k)), \qquad \text{and } |\alpha(k+1)| \geq 2|\alpha(k)| \text{ for every } k.$$

The gap series

$$\sum_{k=1}^\infty \frac{z^{\alpha(k)}}{M_{\alpha(k)}(D)} \tag{2.6}$$

is a subseries of $\sum_\alpha z^\alpha / M_{\alpha(D)}$. That series was shown in the proof of Theorem 1 to converge absolutely inside D, so the series (2.6) converges absolutely inside D too. On the other hand, if ω is a point outside the closure of D having no coordinate equal to 0, then density of the sequence $\{w(k)\}_{k=1}^\infty$ implies the existence of infinitely many values of k for which

$$|w_j(k)| \leq |\omega_j| \quad \text{when } 1 \leq j \leq n \quad \text{and} \quad \prod_{j=1}^n \min(1, w_j(k)) \geq \frac{1}{2} \prod_{j=1}^n \min(1, \omega_j).$$

Accordingly, the series (2.6) evaluated at ω has infinitely many terms with absolute value at least $\frac{1}{2}\prod_{j=1}^n \min(1, \omega_j)$, hence diverges. In summary, the series (2.6) has D as domain of convergence.

What remains to show is that the power series (2.6) cannot be extended holomorphically to a neighborhood of any boundary point of D. Seeking a contradiction, suppose that the series (2.6) admits a holomorphic extension $f(z)$ to a neighborhood of some boundary point of D. This neighborhood necessarily contains some (other) boundary point of D that has no coordinate equal to 0. Call this point w.

The idea now is to restrict to the complex line through w. If $\lambda \in \mathbb{C}$, and $|\lambda| < 1$, then $\lambda w \in D$ (since the convergence domain D is a complete Reinhardt domain). The series (2.6) therefore converges absolutely at λw, that is, the series

$$\sum_{k=1}^\infty \frac{w^{\alpha(k)}}{M_{\alpha(k)}(D)} \lambda^{|\alpha(k)|} \tag{2.7}$$

converges absolutely. If $|\lambda| > 1$, on the other hand, then λw is a point in the exterior of D where (2.6) diverges, so the series (2.7) diverges. Viewed as a power series in the complex variable λ, the series (2.7) thus has radius of convergence equal to 1.

Now $|\alpha(k+1)| \geq 2|\alpha(k)|$ by construction, so the series (2.7) is a gap series with respect to the variable λ. By Hadamard's gap theorem, the series (2.7) cannot be analytically continued to any neighborhood of the point where $\lambda = 1$. On the other hand, the function $f(z)$ is a holomorphic extension of the series (2.6) to a neighborhood of w, so the function sending λ to $f(\lambda w_1, \dots, \lambda w_n)$ is a holomorphic extension of the series (2.7) to a neighborhood of the point 1. The required contradiction has been reached, so the validity of Theorem 4 is established. □

Proof of Theorem 4 using the Baire category theorem. For the same reason as in dimension 1, the space of holomorphic functions on an open set in $\mathbb{C}^n$ is a metrizable space when equipped with the topology of uniform convergence on compact sets. One way to define a suitable metric is to exhaust the open set by an increasing sequence $\{K_j\}_{j=1}^{\infty}$ of compact sets and to declare the distance between two functions f and g to be

$$\sum_{j=1}^{\infty} \min\left(\frac{1}{2^j}, \max\{|f(z) - g(z)| \, : \, z \in K_j\}\right).$$

Since the normal limit of holomorphic functions is still holomorphic, this metric space is complete. So the Baire category theorem is applicable to the space of holomorphic functions.

In modern formulation, the theorem says that a complete metric space is not the union of a countable number of nowhere dense subsets.[20] In the terminology of Baire (1874–1932), a countable union of nowhere dense sets is "a set of first category," and all other sets are sets of second category. The theorem indicates that in a complete metric space, a set of first category is a "small" set, since the complementary set evidently cannot be a set of first category.

The notion of "small" depends on the context. For example, in the metric space $\mathbb{R}$ with the usual absolute-value distance, the rational numbers form a dense subset of first category.

Consider a specific boundary point of a convergence domain D and the set of holomorphic functions on D that extend holomorphically to a neighborhood of this point. The main goal is to prove that this set has first category in the metric space of all holomorphic functions on D. Considering a countable dense set in the boundary of D will then show the existence of a power series that is singular at every boundary point of D. Indeed, most power series that converge in D are singular at every boundary point.

A first step toward the goal is a multidimensional version of a power-series lemma that dates back to the end of the nineteenth century. The correct attribution of the single-variable statement is problematic, but attaching the name of Pringsheim seems appropriate. The generalization to higher dimensions seems not to have been made explicit in the literature until the twenty-first century.[21]

Lemma 1 (Pringsheim lemma in arbitrary dimension). If the coefficients of a power series are real and nonnegative, then the series is singular at every boundary point of the domain of convergence for which all the coordinates are nonnegative real numbers.

Proof. Pringsheim's proof in dimension 1 goes as follows. Suppose that the series $\sum_{j=0}^{\infty} a_j z^j$ has nonnegative coefficients, and let $f(z)$ denote the corresponding holomorphic function.

[20] The theorem was originally formulated on the real line in the doctoral thesis of René Baire, Sur les fonctions de variables réelles, *Annali di Matematica Pura ed Applicata (3)* **3** (1899) 1–123. See page 65.

[21] Alexander D. Scott and Alan D. Sokal, The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma, *Journal of Statistical Physics* **118** (2005) 1151–1261. See Proposition 2.11 on page 1170.

There is no loss of generality in supposing that the radius of convergence of the series is equal to 1. The function f must be singular at *some* point of the unit circle, for otherwise the radius of convergence would be greater than 1. Let $e^{i\theta}$ be a boundary point at which f is singular.

The Taylor series of f centered at the point $e^{i\theta}/2$ must then have radius of convergence equal to $1/2$. Now

$$f^{(k)}(e^{i\theta}/2) = \sum_{j=k}^{\infty} a_j\, j(j-1)\cdots(j-k+1)\left(e^{i\theta}/2\right)^{j-k},$$

and the positivity of a_j for every j implies that $|f^{(k)}(e^{i\theta}/2)| \le f^{(k)}(1/2)$ for every k. Therefore the Taylor series of f centered at the point $1/2$ cannot converge on a disk of radius greater than $1/2$. Accordingly, the function f is singular at the point 1.

This lovely proof does not seem to generalize to establish the lemma in higher dimension. The following argument instead adapts a variant proof invented in dimension 1 by Edmund Landau (1877–1938).

Seeking a contradiction, suppose that the holomorphic function $f(z)$ represented by the power series $\sum_\alpha c_\alpha z^\alpha$ (where $z \in \mathbb{C}^n$) does extend holomorphically to a neighborhood of some boundary point w of the domain of convergence having nonnegative real coordinates. Bumping w reduces to the case that the coordinates of w are strictly positive. A dilation of coordinates modifies the coefficients of the series by positive factors, so there is no loss of generality in supposing additionally that $\|w\| = 1$ (where $\|\cdot\|$ denotes the usual Euclidean norm on the vector space $\mathbb{C}^n$). Let ε be a positive number less than $1/3$ such that the closed ball with center w and radius 3ε lies inside the neighborhood of w to which f extends holomorphically.

The closed ball of radius 2ε centered at the point $(1-\varepsilon)w$ lies inside the indicated neighborhood of w, for if

$$\|z-(1-\varepsilon)w\| \le 2\varepsilon,$$

then the triangle inequality implies that

$$\|z-w\| = \|z-(1-\varepsilon)w-\varepsilon w\| \le \|z-(1-\varepsilon)w\| + \varepsilon\|w\| \le 2\varepsilon+\varepsilon = 3\varepsilon.$$

Consequently, the Taylor series of f about the center $(1-\varepsilon)w$ converges absolutely throughout the closed ball of radius 2ε centered at this point, and in particular at the point $(1+\varepsilon)w$. The value of this Taylor series at the point $(1+\varepsilon)w$ equals

$$\sum_\alpha \frac{1}{\alpha!} f^{(\alpha)}\left((1-\varepsilon)w\right)(2\varepsilon w)^\alpha.$$

The point $(1-\varepsilon)w$ lies inside the domain of convergence of the original power series $\sum_\beta c_\beta z^\beta$, so derivatives of f at $(1-\varepsilon)w$ can be computed by differentiating that series. Let $\beta-\alpha$ denote the multi-index having jth component equal to $\beta_j-\alpha_j$, and say that $\beta \ge \alpha$

when all components of $\beta - \alpha$ are nonnegative. Then

$$f^{(\alpha)}((1-\varepsilon)w) = \sum_{\beta \geq \alpha} \frac{\beta!}{(\beta-\alpha)!} c_\beta \left((1-\varepsilon)w\right)^{\beta-\alpha}.$$

Combining the preceding two expressions shows that the series

$$\sum_\alpha \Bigl(\sum_{\beta \geq \alpha} \frac{\beta!}{\alpha!\,(\beta-\alpha)!} c_\beta \left((1-\varepsilon)w\right)^{\beta-\alpha} \Bigr) (2\varepsilon w)^\alpha$$

converges. All the quantities involved in the sum are nonnegative real numbers, so the parentheses can be removed and the order of summation can be reversed without affecting the convergence. The sum then simplifies (via the binomial expansion) to the series

$$\sum_\beta c_\beta \left((1+\varepsilon)w\right)^\beta.$$

This convergent series is the original series for f evaluated at the point $(1+\varepsilon)w$.

The comparison test implies that the series $\sum_\alpha c_\alpha z^\alpha$ converges absolutely when z lies inside the polydisc determined by the point $(1+\varepsilon)w$, and in particular throughout an open neighborhood of w. (This step uses the supposition that the coordinates of w are nonzero.) Thus w is not a boundary point of the domain of convergence, contrary to hypothesis. This contradiction shows that f must be singular at w after all. □

Now suppose that D is the domain of (absolute) convergence of a power series $\sum_\alpha c_\alpha z^\alpha$. Then D is also the domain of convergence of the series $\sum_\alpha |c_\alpha|\, z^\alpha$. By the lemma, this series is singular at every boundary point of D having positive real coordinates. An arbitrary boundary point of D can be written in the form $(r_1 e^{i\theta_1}, \dots, r_n e^{i\theta_n})$, where each r_j is nonnegative, and the lemma implies that the power series $\sum_\alpha c_\alpha e^{-i(\alpha_1\theta_1+\cdots+\alpha_n\theta_n)} z^\alpha$ is singular at this boundary point. In other words, for every boundary point of D there exists some power series that converges in D but is singular at the specified boundary point.

Choose a countable dense subset $\{w(j)\}_{j=1}^\infty$ of the boundary of D. For each natural number k, let $B_{j,k}$ denote the ball of radius $1/k$ with center $w(j)$. The space of holomorphic functions on $D \cup B_{j,k}$ embeds continuously into the space of holomorphic functions on D via the restriction map. The image of this embedding is not the whole space of holomorphic functions on D, for the preceding paragraph produces a convergent power series on D that cannot be holomorphically extended to the ball $B_{j,k}$. By a corollary of the Baire category theorem (dating back to Banach's famous book[22]), the image of the embedding must be of

[22] Stefan Banach, *Théorie des opérations linéaires*, 1932, second edition 1978, currently available through AMS Chelsea Publishing; an English translation, *Theory of Linear Operations*, is currently available through Dover Publications. The relevant statement is the first theorem in Chapter 3. For a modern treatment, see section 2.11 of Walter Rudin's *Functional Analysis*; a specialization of the theorem proved there is that a continuous linear map between Fréchet spaces (locally convex topological vector spaces equipped with complete translation-invariant metrics) either is an open surjection or has image of first category. In the present context, the restriction map from the metric space of holomorphic functions on $D \cup B_{j,k}$ to the metric space of holomorphic functions on D either is a homeomorphism or has range of first category.

first category (the cited theorem says that if the image were of second category, then it would be the whole space, which it is not). Accordingly, the set of power series on D that extend some distance across some boundary point can be realized as a countable union of sets of first category, hence itself is a set of first category. Therefore the set of nowhere extendable functions is a set of second category.[23] Thus most power series that converge in D have the boundary of D as natural boundary. □

Proof of Theorem 4 using probability. The idea of the second proof is to show that with probability 1, a randomly chosen power series that converges in D is noncontinuable.[24] As a warm-up, consider the case of the unit disc in $\mathbb{C}$. Suppose that the series $\sum_{n=0}^{\infty} c_n z^n$ has radius of convergence equal to 1. The claim is that $\sum_{n=0}^{\infty} \pm c_n z^n$ has the unit circle as natural boundary for almost all choices of the plus-or-minus signs.

The statement can be made precise by introducing the Rademacher functions. When n is a nonnegative integer, the Rademacher function $\varepsilon_n(t)$ can be defined on the interval $[0, 1]$ as follows:

$$\varepsilon_n(t) = \operatorname{sgn}\sin(2^n \pi t) = \begin{cases} 1, & \text{if } \sin(2^n\pi t) > 0, \\ -1, & \text{if } \sin(2^n\pi t) < 0, \\ 0, & \text{if } \sin(2^n\pi t) = 0. \end{cases}$$

Alternatively, the Rademacher functions can be described in terms of binary expansions. If a number t between 0 and 1 is written in binary form as $\sum_{n=1}^{\infty} a_n(t)/2^n$, then $\varepsilon_n(t) = 1 - 2a_n(t)$, except for the finitely many rational values of t that can be written with denominator 2^n (which in any case are values of t for which $a_n(t)$ is not well defined).

Exercise 10. Show that the Rademacher functions form an orthonormal system in the space $L^2[0, 1]$ of square-integrable, real-valued functions. Do the Rademacher functions a *complete* orthonormal system?

The Rademacher functions provide a mathematical model for the notion of "random plus and minus signs." In the language of probability theory, the Rademacher functions are independent and identically distributed symmetric random variables. Each function takes the value $+1$ with probability $1/2$, the value -1 with probability $1/2$, and the value 0 on a set of measure zero (in fact, on a finite set). The intuitive meaning of "independence" is that knowing the value of one particular Rademacher function gives no information about the value of any other Rademacher function.

Here is a precise version of the statement about random series being noncontinuable.[25]

---

[23] Applying Banach's theorem to deduce the noncontinuability of most functions from the existence of a single noncontinuable function is an idea that goes back to Pierre Lelong, Fonctions plurisousharmoniques dans les espaces vectoriels topologiques, *Séminaire Pierre Lelong (Analyse) (1967–1968)*, pp. 167–189, *Lecture Notes in Mathematics*, Vol. 71, Springer, Berlin, 1968. See pages 184–185.

[24] A reference for this section is Jean-Pierre Kahane, *Some Random Series of Functions*, Cambridge University Press; see especially Chapter 4.

[25] R. E. A. C. Paley and A. Zygmund, On some series of functions, (1), *Proceedings of the Cambridge Philosophical Society* **26** (1930), number 3, 337–357 (announcement of the theorem without proof); On some series of

Theorem 5 (Paley–Zygmund). *If the power series $\sum_{n=0}^{\infty} c_n z^n$ has radius of convergence equal to* 1*, then for almost every value of t in* $[0,1]$*, the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ has the unit circle as natural boundary.*

The words "almost every" mean, as usual, that the exceptional set is a subset of $[0,1]$ having measure zero. In probabilists' language, one says that the power series "almost surely" has the unit circle as natural boundary. Implicit in the conclusion is that the radius of convergence of the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ is almost surely equal to 1; this property is evident since the radius of convergence depends only on the moduli of the coefficients in the series, and almost surely $|\varepsilon_n(t) c_n| = |c_n|$ for every n.

Proof. It suffices to show for an arbitrary point p on the unit circle that the set of points t in the unit interval for which the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ continues analytically across p is a set of measure zero. Indeed, take a countable set of points $\{p_j\}_{j=1}^{\infty}$ that is dense in the unit circle: the union over j of the corresponding exceptional sets of measure zero is still a set of measure zero, and when t is in the complement of this set, the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ is nowhere continuable.

So fix a point p on the unit circle. A technicality needs to be checked: is the set of values of t for which the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ continues analytically to a neighborhood of the point p a *measurable* subset of the interval $[0,1]$? In probabilists' language, the question is whether continuability across p is an *event*. The answer is affirmative for the following reason.

A holomorphic function f on the unit disc extends analytically across the boundary point p if and only if there is some rational number r greater than $1/2$ such that the Taylor series of f centered at the point $p/2$ has radius of convergence greater than r. An equivalent statement is that

$$\limsup_{k\to\infty} (|f^{(k)}(p/2)|/k!)^{1/k} < 1/r,$$

or that there exists a positive rational number s less than 2 and a natural number N such that

$$|f^{(k)}(p/2)| < k!\, s^k \qquad \text{whenever } k > N.$$

If $f_t(z)$ denotes the series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$, then

$$|f_t^{(k)}(p/2)| = \left| \sum_{n=k}^{\infty} \varepsilon_n(t) c_n \frac{n!}{(n-k)!} (p/2)^{n-k} \right|.$$

The absolutely convergent series on the right-hand side is a measurable function of t since each $\varepsilon_n(t)$ is a measurable function, so the set of t in the interval $[0,1]$ for which $|f_t^{(k)}(p/2)| <$

functions, (3), *Proceedings of the Cambridge Philosophical Society* **28** (1932), number 2, 190–205 (proof of the theorem).

$k!\, s^k$ is a measurable set, say E_k. The set of points t for which the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ extends across the point p is then

$$\bigcup_{\substack{0<s<2 \\ s\in\mathbb{Q}}} \bigcup_{N\geq 1} \bigcap_{k>N} E_k,$$

which again is a measurable set, being obtained from measurable sets by countably many operations of taking intersections and unions.

Notice too that extendability of $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ across the boundary point p is a "tail event": the property is insensitive to changing any finite number of terms of the series. A standard result from probability known as *Kolmogorov's zero–one law* implies that this event either has probability 0 or has probability 1.

Moreover, each Rademacher function has the same distribution as its negative (both ε_n and $-\varepsilon_n$ take the value 1 with probability $1/2$ and the value -1 with probability $1/2$), so a property that is almost sure for the series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ is almost sure for the series $\sum_{n=0}^{\infty} (-1)^n \varepsilon_n(t) c_n z^n$ or for any similar series obtained by changing the signs according to a fixed pattern that is independent of t. The intuition is that if S is a measurable subset of $[0,1]$, and each element t of S is represented as a binary expansion $\sum_{n=1}^{\infty} a_n(t)/2^n$, then the set S' obtained by systematically flipping the bit $a_5(t)$ from 0 to 1 or from 1 to 0 has the same measure as the original set S; and similarly if multiple bits are flipped simultaneously.

Now suppose, seeking a contradiction, that there is a neighborhood U of p to which the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ continues analytically with positive probability, hence with probability 1 by the zero–one law. This neighborhood contains, for some natural number k, an arc of the unit circle of length greater than $2\pi/k$. For each nonnegative integer n, set b_n equal to -1 if n is a multiple of k and $+1$ otherwise. By the preceding observation, the power series $\sum_{n=0}^{\infty} b_n \varepsilon_n(t) c_n z^n$ continues analytically to U with probability 1. The difference of two continuable series is continuable, so the power series $\sum_{j=0}^{\infty} \varepsilon_{jk}(t) c_{jk} z^{jk}$ (containing only those powers of z that are divisible by k) continues to the neighborhood U with probability 1. This new series is invariant under rotation by every integral multiple of angle $2\pi/k$, so this series almost surely continues analytically to a neighborhood of the whole unit circle. In other words, the power series $\sum_{j=0}^{\infty} \varepsilon_{jk}(t) c_{jk} z^{jk}$ almost surely has radius of convergence greater than 1. Fix a natural number ℓ between 1 and $k-1$ and repeat the argument, changing b_n to be equal to -1 if n is congruent to ℓ modulo k and 1 otherwise. It follows that the power series $\sum_{j=0}^{\infty} \varepsilon_{jk+\ell}(t) c_{jk+\ell} z^{jk+\ell}$, which equals z^ℓ times the rotationally invariant series $\sum_{j=0}^{\infty} \varepsilon_{jk+\ell}(t) c_{jk+\ell} z^{jk}$, almost surely has radius of convergence greater than 1. Adding these series for the different residue classes modulo k recovers the original random series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$, which therefore has radius of convergence greater than 1 almost surely. But as observed just before the proof, the radius of convergence of $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ is almost surely equal to 1. The contradiction shows that the power series $\sum_{n=0}^{\infty} \varepsilon_n(t) c_n z^n$ does, after all, have the unit circle as natural boundary almost surely. □

Now consider the multidimensional situation: suppose that D is the domain of convergence in $\mathbb{C}^n$ of the power series $\sum_\alpha c_\alpha z^\alpha$. Let ε_α denote one of the Rademacher functions, a different one for each multi-index α. The goal is to show that almost surely, the power series $\sum_\alpha \varepsilon_\alpha(t) c_\alpha z^\alpha$ continues analytically across no boundary point of D. It suffices to show for one fixed boundary point p with nonzero coordinates that the series almost surely is singular at p; one gets the full conclusion as before by considering a countable dense sequence in the boundary.

Having fixed such a boundary point p, observe that if δ is an arbitrary positive number, then the power series $\sum_\alpha c_\alpha z^\alpha$ fails to converge absolutely at the dilated point $(1+\delta)p$; for in the contrary case, the series would converge absolutely in the whole polydisc centered at 0 determined by the point $(1+\delta)p$, so p would be in the interior of the convergence domain D instead of on the boundary. (The assumption that all coordinates of p are nonzero is used here.) Consequently, there are infinitely many values of the multi-index α for which $|c_\alpha[(1+2\delta)p]^\alpha| > 1$; for otherwise, the series $\sum_\alpha c_\alpha[(1+\delta)p]^\alpha$ would converge absolutely by comparison with the convergent geometric series $\sum_\alpha[(1+\delta)/(1+2\delta)]^{|\alpha|}$. In other words, there are infinitely many values of α for which $|c_\alpha p^\alpha| > 1/(1+2\delta)^{|\alpha|}$.

Now consider the single-variable random power series obtained by restricting the multivariable random power series to the complex line through p. This series, as a function of λ in the unit disc in $\mathbb{C}$, is $\sum_{k=0}^\infty \left(\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right) \lambda^k$. The goal is to show that this single-variable power series almost surely has radius of convergence equal to 1 and almost surely is singular at the point on the unit circle where $\lambda = 1$. It then follows that the multivariable random series $\sum_\alpha \varepsilon_\alpha(t) c_\alpha z^\alpha$ almost surely is singular at p.

The deduction that the one-variable series almost surely is singular at 1 follows from the same argument used in the proof of the Paley–Zygmund theorem. Although the series coefficient $\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha$ is no longer a Rademacher funtion, it is still a symmetric random variable (symmetric means that the variable is equally distributed with its negative), and the coefficients for different values of k are independent, so the same proof applies.

What remains to show, then, is that the single-variable power series almost surely has radius of convergence equal to 1. The verification of this property requires deducing information about the size of the coefficients $\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha$ from the knowledge that $|c_\alpha p^\alpha| > 1/(1+2\delta)^{|\alpha|}$ for infinitely many values of α.

The orthonormality of the Rademacher functions implies that

$$\int_0^1 \left| \sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha \right|^2 dt = \sum_{|\alpha|=k} |c_\alpha p^\alpha|^2.$$

The sum on the right-hand side is at least as large as any single term, so there are infinitely many values of k for which

$$\int_0^1 \left| \sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha \right|^2 dt > \frac{1}{(1+2\delta)^{2k}}.$$

The issue now is to obtain some control on the function $\left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^2$ from the lower bound on its integral.

A technique for obtaining this control is due to Paley and Zygmund.[26] The following lemma, implicit in the cited paper, is sometimes called *the Paley–Zygmund inequality*.

Lemma 2. If $g\colon [0,1] \to \mathbb{R}$ is a nonnegative, square-integrable function, then the Lebesgue measure of the set of points at which the value of g is greater than or equal to $\frac{1}{2}\int_0^1 g(t)\,dt$ is at least

$$\frac{\left(\int_0^1 g(t)\,dt\right)^2}{4\int_0^1 g(t)^2\,dt}. \tag{2.8}$$

Proof. Let S denote the indicated subset of $[0,1]$ and μ its measure. On the set $[0,1]\setminus S$, the function g is bounded above by the constant $\frac{1}{2}\int_0^1 g(t)\,dt$, so

$$\begin{aligned}\int_0^1 g(t)\,dt &= \int_S g(t)\,dt + \int_{[0,1]\setminus S} g(t)\,dt \\ &\le \int_S g(t)\,dt + (1-\mu)\cdot\frac{1}{2}\int_0^1 g(t)\,dt \\ &\le \int_S g(t)\,dt + \frac{1}{2}\int_0^1 g(t)\,dt.\end{aligned}$$

Therefore

$$\frac{1}{4}\left(\int_0^1 g(t)\,dt\right)^2 \le \left(\int_S g(t)\,dt\right)^2.$$

By the Cauchy–Schwarz inequality,

$$\left(\int_S g(t)\,dt\right)^2 \le \mu\int_S g(t)^2\,dt \le \mu\int_0^1 g(t)^2\,dt.$$

Combining the preceding two inequalities yields the desired conclusion (2.8). □

Now apply the lemma with $g(t)$ equal to $\left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^2$. The integral in the denominator of (2.8) equals

$$\int_0^1 \left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^4 dt. \tag{2.9}$$

Exercise 11. The integral of the product of four Rademacher functions equals 0 unless the four functions are equal in pairs (possibly all four functions are equal).

[26] See Lemma 19 on page 192 of R. E. A. C. Paley and A. Zygmund, On some series of functions, (3), *Proceedings of the Cambridge Philosophical Society* **28** (1932), number 2, 190–205.

There are three ways to group four items into two pairs, so the integral (2.9) equals

$$\sum_{|\alpha|=k} |c_\alpha p^\alpha|^4 + 3 \sum_{\substack{|\alpha|=k \\ |\beta|=k \\ \alpha\neq\beta}} |c_\alpha p^\alpha|^2\, |c_\beta p^\beta|^2.$$

This expression is no more than $3\big(\sum_{|\alpha|=k} |c_\alpha p^\alpha|^2\big)^2$, or $3\big(\int_0^1 g(t)\,dt\big)^2$. Accordingly, the quotient in (2.8) is bounded below by $1/12$ for the indicated choice of g. (The specific value $1/12$ is not significant; what matters is the positivity of this constant.)

The upshot is that there are infinitely many values of k for which there exists a subset of the interval $[0, 1]$ of measure at least $1/12$ such that

$$\left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^{1/k} > \frac{1}{2^{1/\{2k\}}(1+2\delta)}$$

for every t in this subset. The right-hand side exceeds $1/(1+3\delta)$ when k is sufficiently large. For different values of k, the expressions on the left-hand side are independent functions. The probability that two independent events occur simultaneously is the product of their probabilities, so if m is a natural number, and m of the indicated values of k are selected, then the probability that there is *none* for which

$$\left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^{1/k} > \frac{1}{(1+3\delta)} \tag{2.10}$$

is at most $(11/12)^m$. Since $(11/12)^m$ tends to 0 as m tends to infinity, the probability is 1 that inequality (2.10) holds for *some* value of k. For an arbitrary natural number N, the same conclusion holds (for the same reason) for some value of k larger than N. The intersection of countably many sets of probability 1 is again a set of probability 1, so

$$\limsup_{k\to\infty}\left|\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\right|^{1/k} \geq \frac{1}{(1+3\delta)}$$

with probability 1. (The argument in this paragraph is nothing but the proof of the standard *Borel–Cantelli lemma* from probability theory.)

Thus the one-variable power series $\sum_{k=0}^\infty \Big(\sum_{|\alpha|=k} \varepsilon_\alpha(t) c_\alpha p^\alpha\Big) \lambda^k$ almost surely has radius of convergence bounded above by $1 + 3\delta$. But δ is an arbitrary positive number, so the radius of convergence is almost surely bounded above by 1. The radius of convergence is surely no smaller than 1, for the series converges absolutely when $|\lambda| < 1$. Therefore the radius of convergence is almost surely equal to 1. This conclusion completes the proof. □

2.6 Summary: domains of convergence

The preceding discussion shows that for complete Reinhardt domains, the following properties are all equivalent.

- The domain is logarithmically convex.
- The domain is the domain of convergence of some power series.
- The domain is a domain of holomorphy.

In other words, the problem of characterizing domains of holomorphy is solved for the special case of complete Reinhardt domains.

2.7 Separate holomorphicity implies joint holomorphicity

The working definition of a holomorphic function of two (or more) variables is a continuous function that is holomorphic in each variable separately. Hartogs proved that the hypothesis of continuity is superfluous.[27]

Theorem 6 (Hartogs)**.** *Suppose $f(z_1, z_2)$ is holomorphic in z_1 for each fixed z_2 and holomorphic in z_2 for each fixed z_1. Then $f(z_1, z_2)$ is holomorphic jointly in the two variables. In other words, $f(z_1, z_2)$ can be represented locally as a convergent power series in two variables.*

An analogous theorem holds for functions of n complex variables with minor adjustments to the proof. But there is no corresponding theorem for functions of real variables. Indeed, the function on $\mathbb{R}^2$ that equals 0 at the origin and equals $xy/(x^2 + y^2)$ when $(x, y) \neq (0, 0)$ is real-analytic in each variable separately but is not even continuous as a function of the two variables jointly. A large literature exists about deducing properties that hold in all variables jointly from properties that hold in each variable separately.[28]

The proof of Hartogs depends on some prior work of the American complex analyst William Fogg Osgood (1864–1943). Here is the initial step.[29]

Theorem 7 (Osgood)**.** *If $f(z_1, z_2)$ is holomorphic in each variable separately and is bounded (locally) in both variables jointly, then $f(z_1, z_2)$ is holomorphic in both variables jointly.*

[27] See §3 of his 1906 habilitation thesis.

[28] See a survey article by Marek Jarnicki and Peter Pflug, Directional regularity vs. joint regularity, *Notices of the American Mathematical Society* **58** (2011), number 7, 896–904. For more detail, see the same authors' book *Separately Analytic Functions*, European Mathematical Society, 2011.

[29] W. F. Osgood, Note über analytische Functionen mehrerer Veränderlichen, *Mathematische Annalen* **52** (1899), number 2–3, 462–464.

Proof. The conclusion is local and is invariant under translations and dilations of the coordinates, so there is no loss of generality in supposing that the domain of definition of f is the unit bidisc and that the absolute value of f is bounded above by 1 in the bidisc.

There are two natural ways to proceed. Evidently the product of a holomorphic function of z_1 and a holomorphic function of z_2 is jointly holomorphic, so one strategy is to show that $f(z_1, z_2)$ can be realized as the limit of a normally convergent series of product functions. An alternative method is to show directly that f is jointly continuous, whence f can be represented locally by the iterated Cauchy integral formula.

Method 1 For each fixed value of z_1, the function sending z_2 to $f(z_1, z_2)$ is holomorphic, hence can be expanded in a power series $\sum_{k=0}^{\infty} c_k(z_1) z_2^k$ that converges for z_2 in the unit disc. Moreover, the uniform bound on f implies that $|c_k(z_1)| \le 1$ for each k by Cauchy's estimate for derivatives. Accordingly, the series $\sum_{k=0}^{\infty} c_k(z_1) z_2^k$ converges uniformly in both variables jointly in an arbitrary compact subset of the open unit bidisc. All that remains to show, then, is that the coefficient function $c_k(z_1)$ is a holomorphic function of z_1 in the unit disc for each k.

Proceed by induction on k. For the basis step, observe that $c_0(z_1) = f(z_1, 0)$, so $c_0(z_1)$ is a holomorphic function of z_1 in the unit disc by the hypothesis of separate holomorphicity. Now make the induction hypothesis that for some natural number k, the function $c_j(z_1)$ is holomorphic whenever $j < k$. Observe that

$$\frac{f(z_1, z_2) - \sum_{j=0}^{k-1} c_j(z_1) z_2^j}{z_2^k} = c_k(z_1) + \sum_{m=1}^{\infty} c_{k+m}(z_1) z_2^m \qquad \text{when } z_2 \neq 0.$$

When z_2 tends to 0, the right-hand side converges to $c_k(z_1)$ uniformly with respect to z_1, hence so does the left-hand side. For every fixed nonzero value of z_2, the left-hand side is a holomorphic function of z_1 by the induction hypothesis and the hypothesis of separate holomorphicity. So when z_2 tends to 0, the function $c_k(z_1)$ arises as the normal limit of holomorphic functions, hence is holomorphic. This conclusion completes the induction argument and also the proof of the theorem.

Method 2 In view of the local nature of the problem, checking continuity at the origin will suffice. By the triangle inequality,

$$|f(z_1, z_2) - f(0,0)| \le |f(z_1, z_2) - f(z_1, 0)| + |f(z_1, 0) - f(0,0)|.$$

When z_1 is held fixed, the function that sends z_2 to $f(z_1, z_2) - f(z_1, 0)$ is holomorphic in the unit disc, has absolute value bounded above by 2, and is equal to 0 at the origin. Accordingly, the Schwarz lemma implies that $|f(z_1, z_2) - f(z_1, 0)| \le 2|z_2|$. Parallel reasoning implies that $|f(z_1, 0) - f(0,0)| \le 2|z_1|$. Thus

$$|f(z_1, z_2) - f(0,0)| \le 2(|z_1| + |z_2|),$$

so f is indeed jointly continuous at the origin. □

Subsequently, Osgood made further progress but failed to achieve the ultimate result.[30]

Theorem 8 (Osgood)**.** *If $f(z_1,z_2)$ is holomorphic in each variable separately on a domain D, then f is holomorphic in both variables jointly on a dense open subset of D.*

Proof. The goal is to show that if $D_1\times D_2$ is an arbitrary closed bidisc contained in the domain of definition of f, then there is an open subset of $D_1\times D_2$ on which f is jointly holomorphic. For each natural number k, let E_k denote the set of values of the variable z_1 in D_1 such that $|f(z_1,z_2)| \le k$ whenever $z_2 \in D_2$. The continuity of $|f(z_1,z_2)|$ in z_1 for fixed z_2 implies that the set $\{\, z_1 \in D_1 : |f(z_1,z_2)| \le k\,\}$ is closed, and E_k is the intersection of these closed sets as z_2 runs over D_2. So E_k is a closed subset of D_1. The continuity of $|f(z_1,z_2)|$ in z_2 for fixed z_1 implies that $\bigcup_{k=1}^{\infty} E_k = D_1$. By the Baire category theorem, there is some value of k for which the closed set E_k has nonvoid interior. Consequently, there is an open subset of $D_1\times D_2$ on which the separately holomorphic function f is bounded, hence holomorphic by Theorem 7. ☐

Proof of Theorem 6 on separate holomorphicity. The theorem is essentially local, so there is no loss of generality in supposing that the domain of the function is a bidisc. By Theorem 8, there is some smaller bidisc on which the separately holomorphic function is jointly holomorphic. Expand this smaller bidisc as much as possible until singularities are encountered.

To show that the imagined singularities are not actually present, it suffices to prove that if $f(z_1,z_2)$ is separately holomorphic on a neighborhood of the closed unit bidisc, and if there exists a positive δ less than 1 such that f is jointly holomorphic in a neighborhood of the smaller bidisc where $|z_2| \le \delta$ and $|z_1| \le 1$, then f is jointly holomorphic on the open unit bidisc.

In this situation, write $f(z_1,z_2)$ as a series $\sum_{k=0}^{\infty} c_k(z_1)z_2^k$. Each coefficient function $c_k(z_1)$ can be written as an integral

$$\frac{1}{2\pi i}\int_{|z_2|=\delta} \frac{f(z_1,z_2)}{z_2^{k+1}}\,dz_2,$$

so the joint holomorphicity of f on the small bidisc implies that $c_k(z_1)$ is a holomorphic function of z_1 in the unit disc. If M is an upper bound for $|f(z_1,z_2)|$ when $|z_2| \le \delta$ and $|z_1| \le 1$, then $|c_k(z_1)| \le M/\delta^k$ for every k. Accordingly,

$$|c_k(z_1)|^{1/k} \le \frac{\max\{1,M\}}{\delta} \qquad \text{for every value of } k.$$

The constant on the right-hand side is independent of k. Moreover, for each fixed z_1, the series $\sum_{k=0}^{\infty} c_k(z_1)z_2^k$ converges for z_2 in the unit disc, so $\limsup_{k\to\infty} |c_k(z_1)|^{1/k} \le 1$ for every z_1 by the formula for the radius of convergence.

The goal now is to show that if ε is an arbitrary (small) positive number and r is an arbitrary radius slightly less than 1, then there exists a natural number N such that $|c_k(z_1)|^{1/k} < 1+\varepsilon$

[30] W. F. Osgood, Zweite Note über analytische Functionen mehrerer Veränderlichen, *Mathematische Annalen* **53** (1900), number 3, 461–464.

when $k \geq N$ and $|z_1| \leq r$. This property implies that the series $\sum_{k=0}^{\infty} c_k(z_1)z_2^k$ converges uniformly on the set where $|z_1| \leq r$ and $|z_2| \leq 1/(1+2\varepsilon)$, so $f(z_1, z_2)$ is holomorphic on the interior of this set. Since r and ε are arbitrary, the function $f(z_1, z_2)$ is jointly holomorphic on the open unit bidisc.

Letting $u_k(z_1)$ denote the subharmonic function $|c_k(z_1)|^{1/k}$ reduces the problem to the following technical lemma, after which the proof will be complete. □

Lemma 3. Suppose $\{u_k\}_{k=1}^{\infty}$ is a sequence of subharmonic functions on the open unit disc in $\mathbb{C}^1$ that are uniformly bounded above by a (large) constant B, and suppose $\limsup_{k\to\infty} u_k(z) \leq 1$ for every z in the unit disc. Then for every positive ε and every radius r less than 1, there exists a natural number N such that $u_k(z) \leq 1 + \varepsilon$ when $|z| \leq r$ and $k \geq N$.

Proof. A compactness argument reduces the problem to showing that for each point z_0 in the closed disk of radius r, there is a neighborhood U of z_0 and a natural number N such that $u_k(z) \leq 1 + \varepsilon$ when $z \in U$ and $k \geq N$. The definition of lim sup provides only a natural number N depending on z such that $u_k(z) \leq 1 + \varepsilon$ when $k \geq N$. The goal is to obtain an analogous inequality that is locally uniform (in other words, N should be independent of the point z).

Subharmonic functions are upper semicontinuous, so there is a neighborhood of z_0 in which the values of u_k are not much bigger than $u_k(z_0)$, but the size of this neighborhood in principle could depend on k. The key idea for proving a locally uniform estimate is to apply the subaveraging property of subharmonic functions, observing that integrals over discs are stable under small perturbations of the center point because of the uniform bound on the functions. Here are the details.

Fix a positive number δ less than $(1-r)/3$. Fatou's lemma about integrals of nonnegative functions implies that

$$\int_{|z-z_0|<\delta} \liminf_{k\to\infty} (B - u_k(z))\, d\text{Area}_z \leq \liminf_{k\to\infty} \int_{|z-z_0|<\delta} (B - u_k(z))\, d\text{Area}_z.$$

This step uses the hypothesis that there is some (large) uniform upper bound B for the sequence of subharmonic functions. Subtracting $B\pi\delta^2$ from both sides and changing the signs (which reverses the direction of the inequality) shows that

$$\int_{|z-z_0|<\delta} \limsup_{k\to\infty} u_k(z)\, d\text{Area}_z \geq \limsup_{k\to\infty} \int_{|z-z_0|<\delta} u_k(z)\, d\text{Area}_z.$$

The hypothesis that $\limsup_{k\to\infty} u_k(z) \leq 1$ implies that the left-hand side is bounded above by $\pi\delta^2$. Accordingly, there is a natural number N such that

$$\int_{|z-z_0|<\delta} u_k(z)\, d\text{Area}_z < \left(1 + \tfrac{1}{2}\varepsilon\right)\pi\delta^2 \qquad \text{when } k \geq N.$$

If γ is a positive number less than δ, and z_1 is a point such that $|z_1 - z_0| < \gamma$, then the disc of radius $\delta + \gamma$ centered at z_1 contains the disc of radius δ centered at z_0, with an excess of area equal to $\pi(\gamma^2 + 2\gamma\delta)$. The subaveraging property of subharmonic functions implies that

$$\pi(\delta+\gamma)^2 u_k(z_1) \le \int_{|z-z_1|<\delta+\gamma} u_k(z)\, d\mathrm{Area}_z < \left(1 + \tfrac{1}{2}\varepsilon\right)\pi\delta^2 + B\pi\left(\gamma^2 + 2\gamma\delta\right)$$

when $k \ge N$, or

$$u_k(z_1) < \frac{\left(1 + \frac{1}{2}\varepsilon\right)\pi\delta^2 + B\pi(\gamma^2 + 2\gamma\delta)}{\pi(\delta+\gamma)^2}.$$

The limit of the right-hand side when $\gamma \to 0$ equals $1 + \frac{1}{2}\varepsilon$, so there is a small positive value of γ such that $u_k(z_1) < 1 + \varepsilon$ when $k \ge N$ and z_1 is an arbitrary point in the disk of radius γ centered at z_0. This locally uniform estimate completes the proof of the lemma. □

Exercise 12. Find a counterexample showing that the conclusion of the lemma can fail if the hypothesis of a uniform upper bound B is omitted.

Exercise 13. What adjustments are needed in the proof to obtain the analogue of Theorem 6 in dimension n?
(Hartogs addresses this question in §4 of his habilitation.)

Exercise 14. Prove that a separately polynomial function on $\mathbb{C}^2$ is necessarily a jointly polynomial function.[31]

Exercise 15. Define $f : \mathbb{C}^2 \to \mathbb{C} \cup \{\infty\}$ as follows:

$$f(z_1, z_2) = \begin{cases} (z_1 + z_2)^2/(z_1 - z_2), & \text{when } z_1 \ne z_2; \\ \infty, & \text{when } z_1 = z_2 \text{ but } (z_1, z_2) \ne (0, 0); \\ 0, & \text{when } (z_1, z_2) = (0, 0). \end{cases}$$

Show that f is separately meromorphic, yet f is not jointly continuous at $(0, 0)$ with respect to the spherical metric on the extended complex numbers.[32]

[31] The corresponding statement for functions on $\mathbb{R}^2$ was proved by F. W. Carroll, A polynomial in each variable separately is a polynomial, *American Mathematical Monthly* **68** (1961) 42.

[32] This example is due to Theodore J. Barth, Families of holomorphic maps into Riemann surfaces, *Transactions of the American Mathematical Society* **207** (1975) 175–187. Barth's interpretation of the example is that f is a *mapping* from $\mathbb{C}^2$ into the Riemann sphere (a one-dimensional, compact, complex manifold), and f is separately holomorphic but not jointly holomorphic. In other words, Theorem 6 about separately holomorphic functions being holomorphic can fail when the target space $\mathbb{C}$ is replaced by a complex manifold.

3 Convexity

From one point of view, convexity is an unnatural property in complex analysis. The Riemann mapping theorem shows that already in dimension 1, convexity is not preserved by biholomorphic mappings: indeed, every nonconvex but simply connected domain in the plane is conformally equivalent to the unit disc.

On the other hand, Section 2.2 reveals that a certain kind of convexity property—logarithmic convexity—appears naturally in studying convergence domains of power series. The theme of this chapter is that some other analogues of convexity are fundamental in multidimensional complex analysis.

3.1 Real convexity

Ordinary geometric convexity in $\mathbb{R}^n$ can be described either through an internal property (the line segment joining two points of the set stays within the set) or through an external property (every point exterior to the set can be separated from the set by a hyperplane). The latter geometric property can be rephrased in the language of analysis by saying that every point in the exterior of the convex set can be separated from the set by a linear function; that is, there is a linear function that is larger at the specified exterior point than anywhere on the convex set.

More precisely, the internal property says that if x and y are two arbitrary points of the set, and t is an arbitrary real number between 0 and 1, then the point $tx + (1-t)y$ lies in the set. One can deduce by induction that if $x(1), \dots, x(k)$ are points of the set, and $t(1), \dots, t(k)$ are nonnegative real numbers that sum to 1, then the point $t(1)x(1) + \cdots + t(k)x(k)$ lies in the set. The external property is most conveniently formulated for *closed* sets. In this setting, the property says that for each point outside the set, there exists a hyperplane passing through the point and leaving the set on one side. Equivalently, there exists an affine linear function that equals 0 at the specified point and is negative on the set.

To see that the two properties are equivalent for a closed set C, suppose first that the external property holds. To see that the internal property holds, let x and y be two arbitrary points of C. Observe that if ℓ is an arbitrary affine linear function, then $\ell(tx + (1-t)y)$ is an affine linear function of t, hence is monotonic. Accordingly, an affine linear function that has negative values at x and y also has a negative value at every point of the line segment joining x to y. Therefore every point between x and y lies in C. Thus the exterior property implies the interior property.

In the other direction, suppose that the internal property holds for the closed set C. To verify the external property, let q be an arbitrary point in the complement of C. There exists a point x in C at minimal distance from q. The open ball B centered at q with radius equal to the distance $\|q-x\|$ is then disjoint from the set C. Every open half-line with endpoint x that intersects B must be disjoint from C by the internal convexity property of C. The union of all these open half-lines is an open half-space containing q and having boundary equal to the hyperplane through x orthogonal to the line segment joining q to x. The parallel hyperplane through q is the required hyperplane disjoint from C. Thus the interior property implies the exterior property.

For an arbitrary set C, not necessarily closed or convex, the *convex hull* of C is the smallest convex set containing C, namely, the intersection of all convex sets containing C. An equivalent characterization of the convex hull of a set C in $\mathbb{R}^n$ is the collection of all points of the form $t(1)x(1) + \cdots + t(n+1)x(n+1)$, where $x(1), \ldots, x(n+1)$ are arbitrary points of C, and $t(1), \ldots, t(n+1)$ are arbitrary nonnegative real numbers that sum to 1 (Carathéodory's theorem). This description easily implies that the convex hull of an open set is open, and the convex hull of a compact set is compact.[1]

Exercise 16. Find an example of a closed subset of $\mathbb{R}^2$ whose convex hull is not closed.

Observe that an open set G in $\mathbb{R}^n$ is convex if and only if the convex hull of every compact subset K is again a compact subset of G. Indeed, if K is a subset of G, then the convex hull of K is a subset of the convex hull of G, so if G is already convex, then the convex hull of K is both compact and a subset of G. Conversely, if G is not convex, then there are two points of G such that the line segment joining them intersects the complement of G; take K to be the union of the two points.

3.2 Convexity with respect to a class of functions

The analytic description of convexity has a natural generalization. Suppose that $\mathcal{F}$ is a class of upper semicontinuous[2] real-valued functions on an open set G in $\mathbb{C}^n$ (which might be $\mathbb{C}^n$ itself). A compact subset K of G is said to be convex with respect to the class $\mathcal{F}$ if for every point q in $G \setminus K$ there exists an element f of $\mathcal{F}$ for which $f(q) > \max_{z\in K} f(z)$; in other words, every point outside K can be separated from K by a function in $\mathcal{F}$. If $\mathcal{F}$ is a class of functions that are complex-valued but not real-valued (holomorphic functions, say), then

[1] In an infinite-dimensional Hilbert space, however, the convex hull of a compact set is not necessarily closed, let alone compact. But the *closure* of the convex hull of a compact set is compact in every Hilbert space and in every Banach space. See, for example, Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third edition, Springer, 2006, section 5.6.

A standard reference for the finite-dimensional theory of real convexity is R. Tyrrell Rockafellar, *Convex Analysis*, Princeton University Press, 1970 (reprinted 1997). A less comprehensive but more accessible book is *Convexity* by H. G. Eggleston, Cambridge University Press, 1958.

[2] Recall that a real-valued function f is upper semicontinuous if $f^{-1}(-\infty, a)$ is an open set for every real number a. Upper semicontinuity guarantees that f attains a maximum on each compact set.

one can consider convexity with respect to the class of absolute values of the functions in $\mathcal{F}$ (so that inequalities are meaningful). In this context, one typically says "$\mathcal{F}$-convex" for short when the meaning is really "$\mathcal{G}$-convex, where $\mathcal{G} = \{\, |f| : f \in \mathcal{F} \,\}$."

The $\mathcal{F}$-convex hull of a compact set K, denoted by $\widehat{K}_{\mathcal{F}}$ (or simply by $\widehat{K}$ in contexts where the class $\mathcal{F}$ is understood), is the set of all points of G that cannot be separated from K by a function in the class $\mathcal{F}$. The open set G itself is called $\mathcal{F}$-convex if for every compact subset K of G, the $\mathcal{F}$-convex hull $\widehat{K}_{\mathcal{F}}$ (by definition a subset of G) is a compact subset of G.

Example 1. Let G be $\mathbb{R}^n$, and let $\mathcal{F}$ be the set of all continuous functions on $\mathbb{R}^n$. Every compact set K is $\mathcal{F}$-convex because, by Urysohn's lemma, every point not in K can be separated from K by a continuous function. (There is a continuous function that is equal to 0 on K and equal to 1 at a specified point not in K.)

Example 2. Let G be $\mathbb{C}^n$, and let $\mathcal{F}$ be the set of absolute values of the coordinate functions, $\{|z_1|, \dots, |z_n|\}$. The $\mathcal{F}$-convex hull of a single point w is the set of all points z for which $|z_j| \le |w_j|$ for all j, that is, the polydisc determined by the point w. (If some coordinate of w is equal to 0, then this polydisc is degenerate.) More generally, the $\mathcal{F}$-convex hull of a compact set K is the set of points z for which $|z_j| \le \max\{\, |\zeta_j| : \zeta \in K \,\}$ for every j. The $\mathcal{F}$-convex open sets are precisely the open polydiscs centered at the origin.

Exercise 17. Show that a domain in $\mathbb{C}^n$ is convex with respect to the class $\mathcal{F}$ consisting of all the absolute values of monomials z^α if and only if the domain is a logarithmically convex, complete Reinhardt domain.

A useful observation is that increasing the class of functions $\mathcal{F}$ makes separation of points easier, so the collection of $\mathcal{F}$-convex sets becomes larger. In other words, if $\mathcal{F}_1 \subset \mathcal{F}_2$, then every $\mathcal{F}_1$-convex set is also $\mathcal{F}_2$-convex.

Exercise 18. As indicated above, ordinary geometric convexity in $\mathbb{R}^n$ is the same as convexity with respect to the class of linear functions $a_1x_1+\dots+a_nx_n$; convexity with respect to the class of affine linear functions $a_0 + a_1x_1 + \dots + a_nx_n$ is the same notion. The aim of this exercise is to determine what happens if the functions are replaced with their absolute values.

1. Suppose $\mathcal{F}$ is the set $\{\, |a_1x_1 + \dots + a_nx_n| \,\}$ of absolute values of linear functions on $\mathbb{R}^n$. Describe the $\mathcal{F}$-convex hull of a general compact set.
2. Suppose $\mathcal{F}$ is the set $\{\, |a_0 + a_1x_1 + \dots + a_nx_n| \,\}$ of absolute values of affine linear functions. Describe the $\mathcal{F}$-convex hull of a general compact set.

Exercise 19. Repeat the preceding exercise in the setting of $\mathbb{C}^n$ and functions having complex coefficients:

1. Suppose $\mathcal{F}$ is the set $\{\, |c_1z_1+\dots+c_nz_n| \,\}$ of absolute values of complex linear functions. Describe the $\mathcal{F}$-convex hull of a general compact set.
2. Suppose $\mathcal{F}$ is the set $\{\, |c_0+c_1z_1+\dots+c_nz_n| \,\}$ of absolute values of affine complex linear functions. Describe the $\mathcal{F}$-convex hull of a general compact set.

A point and a compact set can be separated by $|f|$ if and only they can be separated by $|f|^2$ or more generally by $|f|^k$ for some positive exponent k. Accordingly, if $\mathcal{F}$ is a class of holomorphic functions, then (from the point of view of $\mathcal{F}$-convexity) there is no loss of generality in assuming that $\mathcal{F}$ is closed under the formation of positive integer powers. Many interesting classes of functions have this property or some additional algebraic structure. For example, the algebra generated by the coordinate functions is the class of polynomials, which is the next topic.

3.2.1 Polynomial convexity

Let the domain G be all of $\mathbb{C}^n$, and let $\mathcal{F}$ be the set of (absolute values of) polynomials (in the complex variables). Then $\mathcal{F}$-convexity is called *polynomial convexity*. (In the setting of $\mathbb{C}^n$, the word "polynomial" is usually understood to mean "holomorphic polynomial," in other words, a polynomial in the complex coordinates $z_1, \ldots, z_n$ rather than a polynomial in the underlying real coordinates of $\mathbb{R}^{2n}$.)

A first observation is that polynomial convexity is no different from convexity with respect to (absolute values of) entire functions. Indeed, an entire function can be approximated uniformly on each compact set by polynomials (namely, by partial sums of the Maclaurin series), so a point can be separated from a compact set by an entire function if and only if the separation can be achieved by a polynomial.

A second observation is that the polynomial hull of a compact set is a subset of the ordinary convex hull. Indeed, if a point is separated from a compact set by a real-linear function $\operatorname{Re} \ell(z)$, then the point is separated equally well by $e^{\operatorname{Re} \ell(z)}$ and hence by $|e^{\ell(z)}|$. Apply the first observation to the entire function $e^{\ell(z)}$. (Alternatively, apply the solution of Exercise 19.)

When $n = 1$, polynomial convexity is a topological property. The basic version of Runge's approximation theorem says that if K is a compact subset of $\mathbb{C}$ (possibly disconnected), and if K has no holes (meaning that $\mathbb{C} \setminus K$ is connected), then every function holomorphic in a neighborhood of K can be approximated uniformly on K by (holomorphic) polynomials.[3] So if K has no holes, and q is a point outside K, then Runge's theorem implies that the function equal to 0 in a neighborhood of K and equal to 1 in a neighborhood of q can be arbitrarily well approximated on $K \cup \{q\}$ by polynomials; hence q is not in the polynomial hull of K. On the other hand, if K has a hole, then the maximum principle implies that points inside the hole belong to the polynomial hull of K. In other words, a compact set K in $\mathbb{C}$ is polynomially convex if and only if K has no holes. For a connected *open* subset of $\mathbb{C}$, polynomial convexity is equivalent to simple connectivity (meaning that the complement of the set with respect to the extended complex numbers is connected).

The situation is much more complicated when the dimension n exceeds 1, for polynomial convexity is no longer determined by a topological condition. For instance, whether or not

[3]There is a deeper approximation theorem, due to S. N. Mergelyan (1928–2008), stating that the conclusion follows from the weaker hypothesis that the function to be approximated is continuous on K and holomorphic on the interior of K.

a circle is polynomially convex depends on how that curve is situated with respect to the complex structure of $\mathbb{C}^n$, as the following example shows.

Example 3. (a) In $\mathbb{C}^2$, the circle $\{ (\cos\theta + i\sin\theta, 0) : 0 \le \theta \le 2\pi \}$ is not polynomially convex. This circle lies in the complex subspace where the second complex coordinate is equal to 0. The one-dimensional maximum principle implies that the polynomial hull of this curve is the disc $\{ (z_1, 0) : |z_1| \le 1 \}$.

(b) In $\mathbb{C}^2$, the circle $\{ (\cos\theta + 0i, \sin\theta + 0i) : 0 \le \theta \le 2\pi \}$ is polynomially convex. This circle lies in the real subspace $\mathbb{R}^2$ where both complex coordinates happen to be real numbers. Since the polynomial hull is a subset of the ordinary convex hull, all that needs to be shown is that points inside the disc bounded by the circle in $\mathbb{R}^2$ can be separated from the circle by polynomials in the complex coordinates. The polynomial $1 - z_1^2 - z_2^2$ is identically equal to 0 on the circle and takes positive real values at points inside the circle, so this polynomial exhibits the required separation.

The next example generalizes the preceding idea to produce a general class of polynomially convex sets.

Example 4. If K is a compact subset of the standard real subspace of $\mathbb{C}^n$ (that is, $K \subset \mathbb{R}^n \subset \mathbb{C}^n$), then K is polynomially convex.

This statement could be proved by invoking the Weierstrass approximation theorem in $\mathbb{R}^n$. Indeed, a point in $\mathbb{R}^n \setminus K$ can be separated from K by a continuous real-valued function, hence by a real polynomial; now replace the variables in the polynomial by complex variables. But the big hammer of the Weierstrass approximation theorem is not really needed. Here is an elementary constructive argument that yields the required conclusion.

Let q be a point outside the compact set K. The goal is to find an entire function that separates q from K (since, as observed above, such an entire function can be approximated on $K \cup \{q\}$ by a polynomial). A suitable separating function is the Gaussian

$$\exp \sum_{j=1}^{n} -(z_j - \operatorname{Re} q_j)^2.$$

To see why this function has the required property, let $M(z)$ denote the absolute value,

$$\exp \sum_{j=1}^{n} \left[(\operatorname{Im} z_j)^2 - (\operatorname{Re} z_j - \operatorname{Re} q_j)^2 \right].$$

If $q \in \mathbb{C}^n \setminus \mathbb{R}^n$, then $M(q) = \exp \sum_{j=1}^{n} (\operatorname{Im} q_j)^2 > 1$, and

$$\max_{z\in K} M(z) = \max_{z \in K} \exp \sum_{j=1}^{n} -(\operatorname{Re} z_j - \operatorname{Re} q_j)^2 \le 1.$$

If $q \in \mathbb{R}^n \setminus K$, then $M(q) = 1$, and the continuous real-valued expression $\sum_{j=1}^{n} (\operatorname{Re} z_j - \operatorname{Re} q_j)^2$ has a strictly positive lower bound on the compact set K, so $\max_{z\in K} M(z) < 1$. The required separation holds in both cases. (Checking the second case suffices, for the polynomial hull of K is a subset of the convex hull of K, hence a subset of $\mathbb{R}^n$.)

A further generalization is possible. A real subspace of $\mathbb{C}^n$ is called *totally real* if it contains no nontrivial complex subspace, that is, if every nonzero point z in the subspace has the property that the point iz lies outside the subspace. If x_1 and x_2 denote the real parts of the variables z_1 and z_2, and y_1 and y_2 denote the imaginary parts, then the x_1x_2 subspace is totally real, as are the x_1y_2, y_1x_2, and y_1y_2 subspaces. The real subspace $\{ (\lambda, \overline{\lambda}) : \lambda \in \mathbb{C} \}$ is another example of a totally real subspace of $\mathbb{C}^2$.

For dimensional reasons, every one-dimensional real subspace of $\mathbb{C}^n$ is totally real, but no real subspace of $\mathbb{C}^n$ of real dimension greater than n can be totally real. A "typical" real subspace of $\mathbb{C}^n$ of real dimension n or less is totally real, in the sense that a generic small perturbation of a complex line is totally real.[4]

Exercise 20. Show that every compact subset of a totally real subspace of $\mathbb{C}^n$ is polynomially convex.

Additional polynomially convex sets can be obtained from ones already in hand by applying the following observation.

Example 5. If K is a polynomially convex compact subset of $\mathbb{C}^n$, and p is a polynomial, then the graph $\{ (z, p(z)) \in \mathbb{C}^{n+1} : z \in K \}$ is a polynomially convex compact subset of $\mathbb{C}^{n+1}$.

To see why, suppose that $\alpha \in \mathbb{C}^n$ and $\beta \in \mathbb{C}$, and the point (α, β) in $\mathbb{C}^{n+1}$ is not in the graph of p over K. To verify that the point (α, β) can be separated from the graph by a polynomial, consider two cases. If $\alpha \notin K$, then there is a polynomial of n variables that separates α from K in $\mathbb{C}^n$; the same polynomial, viewed as a polynomial on $\mathbb{C}^{n+1}$ that is independent of z_{n+1}, separates the point (α, β) from the graph of p. On the other hand, if $\alpha \in K$, but $\beta \neq p(\alpha)$, then the polynomial $z_{n+1} - p(z)$ is identically equal to 0 on the graph and is not equal to 0 at (α, β), so this polynomial separates (α, β) from the graph.

Some basic examples of polynomially convex sets in $\mathbb{C}^n$ are the *polynomial polyhedra*, which are intersections of sublevel sets of absolute values of polynomials. The model case is the unit polydisc $\{ z \in \mathbb{C}^n : |z_1| \le 1, \dots, |z_n| \le 1 \}$. More generally, if $p_1, \dots, p_k$ are polynomials, then $\{ z \in \mathbb{C}^n : |p_1(z)| \le 1, \dots, |p_k(z)| \le 1 \}$ is a closed polynomial polyhedron, and $\{ z \in \mathbb{C}^n : |p_1(z)| < 1, \dots, |p_k(z)| < 1 \}$ is an open polynomial polyhedron. A concrete example is the logarithmically convex, complete Reinhardt domain $\{ (z_1, z_2) \in \mathbb{C}^2 : |z_1| < 1$, $|z_2| < 1$, and $|2z_1z_2| < 1 \}$.

Instead of constraining a polynomial p by requiring that $|p(z)| \le 1$, one could require that $|p(z)| \le r$ for some constant r, but nothing is gained by this apparently more general condition. Indeed, if r is strictly positive, then $|p(z)| \le r$ if and only if $\left|\frac{p(z)}{r}\right| \le 1$, so one can

[4]In the language of algebraic geometry, the Grassmannian of real two-dimensional subspaces of $\mathbb{C}^2$ ($= \mathbb{R}^4$) is a manifold of real dimension 4. The real two-planes that happen to be one-dimensional complex subspaces are the same as the complex lines in $\mathbb{C}^2$. The set of complex lines in $\mathbb{C}^2$ is one-dimensional complex projective space, a manifold of real dimension 2. Thus the real two-planes that fail to be totally real form a codimension 2 submanifold of the Grassmannian of all real two-planes in $\mathbb{C}^2$.

make the upper bound equal to 1 simply by rescaling the polynomial. And if $r = 0$, then $|p(z)| \le r$ if and only if both $|p(z) - 1| \le 1$ and $|p(z) + 1| \le 1$, so here too the upper bound can be taken to be 1 at the expense of increasing the number of polynomials.

The observation in the preceding sentence can be rephrased in the language of algebraic geometry. An (affine) *algebraic variety* in $\mathbb{C}^n$ is the set of common zeros of a finite[5] number of polynomials. Since the zero set of a polynomial can be described by two inequalities for absolute values of polynomials, every algebraic variety can be realized as a closed polynomial polyhedron (closed but not in general compact when $n > 1$).

A question arose in class of whether an arbitrary finite set of points in $\mathbb{C}^n$ can be realized as a closed polynomial polyhedron. An affirmative answer follows by showing that every finite set is an algebraic variety. A finite set in $\mathbb{C}^1$ is the zero set of a single polynomial p, hence is an algebraic variety. If this set in $\mathbb{C}^1$ is viewed as a subset of $\mathbb{C}^n$ by making the extra coordinates equal to 0, then the set is still an algebraic variety in $\mathbb{C}^n$, being the zero set of the polynomials $p(z_1), z_2, \ldots, z_n$.

One way to show that an arbitrary finite subset of $\mathbb{C}^n$ is an algebraic variety is to reduce to the preceding case by applying the following lemma about polynomial automorphisms of $\mathbb{C}^n$. In this context, the word "automorphism" means a bijective mapping from $\mathbb{C}^n$ onto itself having holomorphic coordinate functions and a holomorphic inverse mapping. (There is a holomorphic version of the inverse-function theorem, so the inverse mapping actually is automatically holomorphic.) A polynomial automorphism has coordinate functions that are polynomials. (The coordinate functions of the inverse mapping are automatically polynomials.) The polynomial automorphisms evidently form a group under composition.

Lemma 4. Suppose $n > 1$, and suppose given in $\mathbb{C}^n$ two finite sets of the same cardinality, say $w(1), \ldots, w(k)$ and $w'(1), \ldots, w'(k)$. There exists a polynomial automorphism of $\mathbb{C}^n$ that maps $w(j)$ to $w'(j)$ for every j between 1 and k.

The lemma implies that when $n > 1$, an arbitrary finite subset of $\mathbb{C}^n$ can be mapped by a polynomial automorphism to a subset of a one-dimensional complex subspace. By the special case considered above, this finite set is an algebraic variety. The inverse image of an algebraic variety under a polynomial automorphism evidently is again an algebraic variety, because composing two polynomials gives another polynomial.

Proof of the lemma. A simple induction reduces the problem to the following claim:

> If $n > 1$, and E is a finite subset of $\mathbb{C}^n$, and w and w' are points outside of E, then there exists a polynomial automorphism of $\mathbb{C}^n$ that fixes every point of E and maps w to w'.

The proof of this claim is a small modification of an argument in a famous article[6] by the

[5] The set of common zeros of an infinite number polynomials can always be expressed as the set of common zeros of a finite number of polynomials because, according to Hilbert's basis theorem, every ideal in the polynomial ring $\mathbb{C}[z_1, z_2, \ldots, z_n]$ is finitely generated.

[6] Jean-Pierre Rosay and Walter Rudin, Holomorphic maps from $\mathbb{C}^n$ to $\mathbb{C}^n$, *Transactions of the American Mathematical Society* **310** (1988), no. 1, 47–86. See page 50.

French-American mathematician Jean-Pierre Rosay and the Austrian-American mathematician Walter Rudin (1921–2010). Here are the details.

Since the set E is finite, there is a complex hyperplane passing through w and disjoint from E. This hyperplane is a level set of some linear function $L\colon \mathbb{C}^n \to \mathbb{C}$, and the finite set $L(E)$ in $\mathbb{C}$ does not contain the value $L(w)$. Let p be a polynomial of one variable whose zeros coincide with the set $L(E)$ and such that $p(L(w)) = 1$. Let w'' be an arbitrary point in $\mathbb{C}^n$ other than w that lies in the indicated hyperplane; that is, $L(w'' - w) = 0$. The polynomial map sending a point z in $\mathbb{C}^n$ to the image $z + p(L(z))(w'' - w)$ is an automorphism (namely, a polynomial shear); the inverse map is obtained by changing the sign in the sum. The properties of the polynomial p imply that the indicated polynomial automorphism fixes every point of E and sends w to w''.

Now consider two hyperplanes disjoint from E, one passing through the point w and the other passing through the point w'. If necessary, rotate one of the hyperplanes slightly to ensure that the two hyperplanes intersect, and let w'' be some point in the intersection. By the construction in the preceding paragraph, there is a polynomial automorphism of $\mathbb{C}^n$ that fixes every point of E and maps w to w'', and there is a second polynomial automorphism that fixes every point of E and maps w'' to w'. Compose these two polynomial automorphisms to obtain the required one. □

The preceding argument shows that a finite set in $\mathbb{C}^n$ can be realized as the common zero set of n polynomials. The fact that every finite set is an algebraic variety can alternatively be established without invoking the lemma, but at the expense of introducing an unnecessarily large number of polynomials. Indeed, a single point w in $\mathbb{C}^n$ is the common zero set of the polynomials $z_1 - w_1, z_2 - w_2, \ldots, z_n - w_n$, hence is an algebraic variety. View this observation as the basis step of an induction. It suffices to show now that the union of two algebraic varieties is an algebraic variety. If the first variety is the common zero set of polynomials $p_1, \ldots, p_j$, and the second variety is the common zero set of polynomials $q_1, \ldots, q_k$, then the union of the varieties is the common zero set of the product polynomials $p_\ell q_m$, where $1 \le \ell \le j$ and $1 \le m \le k$.

A polynomial polyhedron in $\mathbb{C}^n$ is polynomially convex, since a point in the complement is separated from the polyhedron by at least one of the defining polynomials. Notice that the number of polynomials defining the polyhedron is often larger than the dimension n. (If the polyhedron is compact and nonvoid, then the number k of polynomials cannot be less than n, but proving this property requires some tools not yet introduced.[7]) A standard way to force a polynomial polyhedron to be bounded is take the intersection with a polydisc (that is, include in the set of defining polynomials the function z_j/R for some large R and for each j from 1 to n).

[7] If w is a point of the polyhedron, then the k sets $\{z \in \mathbb{C}^n : p_j(z) - p_j(w) = 0\}$ are analytic varieties of codimension 1 that intersect in an analytic variety of dimension at least $n - k$ that is contained in the polyhedron. If $k < n$, then this analytic variety has positive dimension, but there are no compact analytic varieties of positive dimension.

A proposition from single-variable complex analysis, *Hilbert's lemniscate theorem*,[8] says that every simple closed curve in $\mathbb{C}$ can be approximated by a level curve of the absolute value of a polynomial (that is, by a lemniscate[9]). More precisely, Hilbert's statement is that if γ_1 and γ_2 are two simple closed curves, with γ_2 inside γ_1, then there is a lemniscate inside γ_1 that surrounds γ_2. Equivalently, if K is a compact, polynomially convex subset of $\mathbb{C}$, and U is an open neighborhood of K, then there is a polynomial p such that $|p(z)| < 1$ when $z \in K$ and $|p(z)| > 1$ when $z \in \mathbb{C} \setminus U$. The theorem generalizes to higher dimension as follows.

Theorem 9. *A polynomially convex set in $\mathbb{C}^n$ can be approximated by polynomial polyhedra:*

(a) If K is a compact polynomially convex set, and U is an open neighborhood of K, then there is an open polynomial polyhedron P such that $K \subset P \subset U$.

(b) If G is a polynomially convex open set, then G can be expressed as the union of an increasing sequence of open polynomial polyhedra.

Proof. (a) Being bounded, the set K is contained in the interior of some closed polydisc D. If D is a subset of U, then the interior of D is already the required polyhedron. On the other hand, if $D \setminus U$ is nonvoid, then for each point w in $D \setminus U$, there is a polynomial p that separates w from K. This polynomial can be multiplied by a suitable constant to guarantee that $\max\{\, |p(z)| : z \in K \,\} < 1 < |p(w)|$. Hence the set $\{\, z : |p(z)| < 1 \,\}$ contains K and is disjoint from a neighborhood of w. Since the set $D \setminus U$ is compact, there are finitely many polynomials $p_1, \ldots, p_k$ such that the polyhedron $\bigcap_{j=1}^{k} \{\, z : |p_j(z)| < 1 \,\}$ contains K and does not intersect $D \setminus U$. Cutting down this polyhedron by intersecting with the interior of D gives a new polyhedron that contains K and is contained in U.

(b) Exhaust G by an increasing sequence of compact sets. The polynomial hulls of these sets form another increasing sequence of compact subsets of G (under the hypothesis that G is polynomially convex). Discarding some of the sets (if necessary) and renumbering produces an exhaustion of G by a sequence $\{K_j\}_{j=1}^{\infty}$ of polynomially convex compact sets such that each K_j is contained in the interior of K_{j+1}. The first part of the theorem then provides a sequence $\{P_j\}_{j=1}^{\infty}$ of open polynomial polyhedra such that $K_j \subset P_j \subset K_{j+1}$ for every j. □

In Hilbert's lemniscate theorem in $\mathbb{C}^1$, only a single polynomial is needed to determine a polynomial polyhedron that approximates a prescribed polynomially convex set. An interesting question is whether n polynomials suffice to define an approximating polyhedron in $\mathbb{C}^n$.

---

[8] D. Hilbert, Ueber die Entwickelung einer beliebigen analytischen Function einer Variabeln in eine unendliche nach ganzen rationalen Functionen fortschreitende Reihe, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* (1897) 63–70. See pages 67–68.

[9] The Latin adjective *lemniscatus* has the meaning of "decorated with ribbons." The derived English word "lemniscate" originally applied to a ribbon-like figure-eight curve, such as the set of values of the complex variable z for which $|z^2 - 1| = 1$. The word subsequently acquired the more general meaning of the level set of the absolute value of a polynomial of arbitrary degree.

A partial result in this direction has been known for over half a century. A polyhedron can be a disconnected set, and Errett Bishop (1928–1983) showed[10] that the approximation can be accomplished by a set that is the union of some of the connected components of a polyhedron defined by n polynomials. Additional quantitative information about the polynomials is known.[11] For sets with enough symmetry in two dimensions[12] and in n dimensions,[13] an affirmative answer is known, but the general case remains open.

The theory of polynomial convexity is sufficiently mature that there exists a good reference book,[14] but determining the polynomial hull of even quite simple sets in $\mathbb{C}^2$ remains a difficult problem. The following example shows that the union of two disjoint, compact, polynomially convex sets in $\mathbb{C}^2$ need not be polynomially convex (in contrast to the situation in $\mathbb{C}^1$).

Example 6 (Kallin,[15] 1965). Let K_1 be $\{(e^{i\theta}, e^{-i\theta}) \in \mathbb{C}^2 : 0 \le \theta < 2\pi\}$, and let K_2 be $\{(2e^{i\theta}, \frac{1}{2}e^{-i\theta}) \in \mathbb{C}^2 : 0 \le \theta < 2\pi\}$. Both of the sets K_1 and K_2 are polynomially convex in view of Exercise 20, since K_1 lies in the totally real subspace of $\mathbb{C}^2$ in which $z_1 = \overline{z_2}$, and K_2 lies in the totally real subspace in which $z_1/4 = \overline{z_2}$. The union $K_1 \cup K_2$ is not polynomially convex, for the polynomial hull contains the set $\{(\lambda, 1/\lambda) \in \mathbb{C}^2 : 1 < |\lambda| < 2\}$, which can be thought of as an "analytic annulus." Indeed, if $p(z_1, z_2)$ is a polynomial on $\mathbb{C}^2$ whose absolute value is less than 1 on $K_1 \cup K_2$, then $p(\lambda, 1/\lambda)$ is a holomorphic function on $\mathbb{C} \setminus \{0\}$ whose absolute value is less than 1 on the boundary of the annulus $\{\lambda \in \mathbb{C} : 1 < |\lambda| < 2\}$ and hence (by the one-dimensional maximum principle) on the interior of the annulus.

Moreover, the polynomial hull of $K_1 \cup K_2$ is precisely the set $\{(\lambda, 1/\lambda) : 1 \le |\lambda| \le 2\}$. To see why, consider the polynomial $1 - z_1 z_2$. Since this polynomial is identically equal to 0 on $K_1 \cup K_2$, the only points that have a chance to lie in the polynomial hull of $K_1 \cup K_2$ are points where $1 - z_1 z_2 = 0$. If such a point additionally has first coordinate with absolute value greater than 2, then the polynomial z_1 separates that point from $K_1 \cup K_2$. On the other hand, if a point in the zero set of $1 - z_1 z_2$ has first coordinate with absolute value less than 1, then the second coordinate has absolute value greater than 1, so the polynomial z_2 separates the point from $K_1 \cup K_2$.

The preceding example shows that in general, polynomial convexity is not preserved by taking unions. Nonetheless, there are some useful situations in which new polynomially convex sets can be obtained by taking unions.

[10] Errett Bishop, Mappings of partially analytic spaces, *American Journal of Mathematics* **83** (1961), number 2, 209–242.

[11] Stéphanie Nivoche, Polynomial convexity, special polynomial polyhedra and the pluricomplex Green function for a compact set in $\mathbb{C}^n$, *Journal de Mathématiques Pures et Appliquées (9)* **91** (2009), no. 4, 364–383.

[12] Thomas Bloom, Norman Levenberg, and Yu. Lyubarskii, A Hilbert lemniscate theorem in $\mathbb{C}^2$, *Annales de l'Institut Fourier* **58** (2008), no. 6, 2191–2220.

[13] Alexander Rashkovskii and Vyacheslav Zakharyuta, Special polyhedra for Reinhardt domains, *Comptes Rendus Mathématique, Académie des Sciences, Paris* **349**, issues 17–18, (2011) 965–968.

[14] Edgar Lee Stout, *Polynomial Convexity*, Birkhäuser Boston, 2007.

[15] Eva Kallin, Polynomial convexity: The three spheres problem, *Proceedings of the Conference in Complex Analysis (Minneapolis, 1964)*, pp. 301–304, Springer, Berlin, 1965.

Proposition 10. *If K is a polynomially convex, compact set in $\mathbb{C}^n$, then the union of K and a finite set of points is polynomially convex.*

Proof. The result follows from an evident induction on the number of points as soon as the basis step is accomplished. Suppose, then, that w is a specified point outside of K. What will be shown is that if w' is some other point outside of K, then there exists a polynomial p such that

$$\max\{\,|p(z)| : z \in K \cup \{w\}\,\} < 1 < |p(w')|.$$

In other words, an arbitrary point w' outside $K \cup \{w\}$ can be separated from $K \cup \{w\}$ by a polynomial.

There certainly exists a polynomial p_1 such that $p_1(w') = 1$ and $p_1(w) = 0$. Let M denote $\max\{\,|p_1(z)| : z \in K\,\}$. By the polynomial convexity of K, there is a polynomial p_2 such that

$$\max\{\,|p_2(z)| : z \in K\,\} < |p_2(w')|.$$

Multiplying p_2 by a suitable constant shows that there is no loss of generality in supposing that

$$\max\{\,|p_2(z)| : z \in K\,\} < 1 < |p_2(w')|.$$

After making that normalization, choose a natural number k such that

$$\max\{\,|p_2^k(z)| : z \in K\,\} < \frac{1}{2M}.$$

The required polynomial p is the product $p_1 p_2^k$. Indeed, $|p(w')| = |p_2^k(w')| > 1$, and $p(w) = 0$, and $\max\{\,|p(z)| : z \in K\,\} < 1/2$. □

Exercise 21. Show that if K is a polynomially convex compact set in $\mathbb{C}^n$, and $w(1), \dots, w(k)$ are distinct points in the complement of K, and $c(1), \dots, c(k)$ are complex numbers, and ε is a positive real number, then there exists a polynomial p such that when $1 \le j \le k$, the polynomial p takes the value $c(j)$ at the point $w(j)$, and $\max\{\,|p(z)| : z \in K\,\} < \varepsilon$.

Proposition 11. *If K_1 and K_2 are disjoint, compact, convex sets in $\mathbb{C}^n$, then the union $K_1 \cup K_2$ is polynomially convex.*

Proof. The disjoint convex sets K_1 and K_2 can be separated by a real hyperplane, or equivalently by the real part of a complex linear function L. The geometric picture is that L projects $\mathbb{C}^n$ onto a complex line (a one-dimensional complex subspace). The sets $L(K_1)$ and $L(K_2)$ are then disjoint, compact, convex sets in $\mathbb{C}$.

Suppose now that w is a point outside of $K_1 \cup K_2$. The goal is to separate w from $K_1 \cup K_2$ by a polynomial. There are two cases, depending on the location of $L(w)$.

If $L(w) \notin L(K_1) \cup L(K_2)$, then Runge's theorem provides a polynomial p of one complex variable such that $|p(L(w))| > 1$, and $|p(z)| < 1$ when $z \in L(K_1) \cup L(K_2)$. In other words, the composite polynomial function $p \circ L$ separates w from $K_1 \cup K_2$ in $\mathbb{C}^n$.

If $L(w) \in L(K_1) \cup L(K_2)$, then suppose without loss of generality that $L(w) \in L(K_1)$. Since $w \notin K_1$, and K_1 is polynomially convex, there is a polynomial p on $\mathbb{C}^n$ such that $|p(w)| > 1$ and $|p(z)| < 1/3$ when $z \in K_1$. Let M be an upper bound for $|p|$ on the compact set K_2. Applying Runge's theorem in $\mathbb{C}$ produces a polynomial q of one variable such that $|q| < 1/(3M)$ on $L(K_2)$ and $2/3 \le |q| \le 1$ on $L(K_1)$. The claim now is that the product polynomial $p \cdot (q \circ L)$ separates w from $K_1 \cup K_2$. Indeed, on K_1, the factor p has absolute value less than $1/3$, and the factor $q \circ L$ has absolute value no greater than 1; on K_2, the factor p has absolute value at most M, and the factor $q \circ L$ has absolute value less than $1/(3M)$; and at w, the factor p has absolute value exceeding 1, and the factor $q \circ L$ has absolute value at least $2/3$. □

The preceding proposition is a special case of a separation lemma of Eva Kallin, who showed in the paper cited above that the union of *three* pairwise disjoint closed balls in $\mathbb{C}^n$ is always polynomially convex. The question of whether the union of *four* pairwise disjoint closed balls is always polynomially convex remains open after more than half a century. The problem is a subtle one, for Kallin constructed an example of three pairwise disjoint closed *polydiscs* in $\mathbb{C}^3$ whose union is *not* polynomially convex.

Runge's theorem in dimension 1 reveals that polynomial convexity is closely connected with the approximation of holomorphic functions by polynomials. The following theorem, an analogue of Runge's theorem in higher dimension, is known as the Oka–Weil theorem [named after the Japanese mathematician Kiyoshi Oka (1901–1978) and the French mathematician André Weil (1906–1998)].

Theorem 12 (Oka–Weil)**.** *If K is a compact, polynomially convex set in $\mathbb{C}^n$, then every function holomorphic in a neighborhood of K can be approximated uniformly on K by (holomorphic) polynomials.*

The proof has to be deferred until later, after some theory is developed about solvability of the inhomogeneous multivariable Cauchy–Riemann equations.

Exercise 22. Give an example of a compact set K in $\mathbb{C}^2$ that is not polynomially convex, yet every function holomorphic in a neighborhood of K can be approximated uniformly on K by polynomials.

3.2.2 Linear and rational convexity

The preceding examples involve functions that are globally defined on the whole of $\mathbb{C}^n$. But in many interesting cases, the class of functions depends on the region under consideration.

For instance, suppose that G is an open set in $\mathbb{C}^n$, and $\mathcal{F}$ is the class of those linear fractional functions

$$\frac{a_0 + a_1 z_1 + \cdots + a_n z_n}{b_0 + b_1 z_1 + \cdots + b_n z_n}$$

that happen to be holomorphic on G (in other words, the denominator is nonzero at all points inside G). A more precise notation is $\mathcal{F}_G$, but typically the open set G is clear from context. By the solution of Exercise 19, every convex set is $\mathcal{F}$-convex. A simple example of a nonconvex but $\mathcal{F}$-convex open set is $\mathbb{C}^2 \setminus \{ (z_1, z_2) \in \mathbb{C}^2 : z_2 = 0 \}$. Indeed, if K is a compact subset of this open subset of $\mathbb{C}^2$, then the function $1/z_2$ is bounded on K, so $\widehat{K}_{\mathcal{F}}$ stays away from the boundary of the open set.

An open set G is convex in the ordinary geometric sense if through each boundary point of G there passes a real hyperplane that does not intersect G (a *supporting* hyperplane). The claim now is that an open set G in $\mathbb{C}^n$ is $\mathcal{F}$-convex if and only if through each boundary point of G there passes a complex hyperplane that does not intersect G.

For the proof, suppose first that G is $\mathcal{F}$-convex, and let w be a point in the boundary of G. If K is a compact subset of G, then $\widehat{K}_{\mathcal{F}}$ is again a compact subset of G, so to every point w' in G sufficiently close to w there corresponds a linear fractional function f in $\mathcal{F}$ such that $f(w') = 1 > \max\{ |f(z)| : z \in K \}$. If L denotes the difference between the numerator of f and the denominator of f, then $L(z) = 0$ at a point z in G if and only if $f(z) = 1$. Hence the zero set of L, which is a complex hyperplane, passes through w' and does not intersect K. Multiply L by a suitable constant to ensure that the vector consisting of the coefficients of L has length 1.

Now exhaust G by an increasing sequence $\{K_j\}$ of compact sets, and apply the preceding construction to obtain a sequence $\{w_j\}$ of points in G converging to w and a sequence $\{L_j\}$ of normalized first-degree polynomials such that $L_j(w_j) = 0$, the zero set of L_j being disjoint from K_j. Since the set of vectors in $\mathbb{C}^n$ of length 1 is compact, taking the limit of a suitable subsequence produces a complex hyperplane that passes through the boundary point w and does not intersect the open set G.

Conversely, a supporting complex hyperplane at a boundary point w is the zero set of a certain first-degree polynomial L, and $1/L$ is then a linear fractional function that blows up at w and is holomorphic on G. Therefore the $\mathcal{F}$-convex hull of a compact set K in G stays away from w. Since w is arbitrary, the hull $\widehat{K}_{\mathcal{F}}$ stays away from the whole boundary of G. The coordinate functions are elements of $\mathcal{F}$, so $\widehat{K}_{\mathcal{F}}$ is bounded too. Consequently, the hull $\widehat{K}_{\mathcal{F}}$ is compact. Since K is arbitrary, the domain G is $\mathcal{F}$-convex.

The property of $\mathcal{F}_G$-convexity when $\mathcal{F}_G$ is the indicated class of linear fractional functions that are holomorphic on G is known as *weak linear convexity*. The more restrictive property of *linear convexity* means that the complement of the domain can be written as a union of complex hyperplanes. The terminology is not completely standardized, however, so one has to check each author's definitions.

There exist weakly linearly convex domains that are not linearly convex. The idea can be seen already for the notion of ordinary convexity in $\mathbb{R}^2$. Take an equilateral triangle of side length 1 and erase a middle portion, leaving in the corners three equilateral triangles of side length slightly less than $1/2$. There is a supporting line through each boundary point of this disconnected set, but there is no line through the centroid that is disjoint from the three triangles. This idea can be implemented in $\mathbb{C}^2$ to construct a connected, weakly linearly convex domain that is not linearly convex.[16]

[16] A reference is Mats Andersson, Mikael Passare, and Ragnar Sigurdsson, *Complex Convexity and Analytic*

Next consider general rational functions (quotients of polynomials). A compact set K in $\mathbb{C}^n$ is called *rationally convex* if every point w outside K can be separated from K by a rational function that is holomorphic on $K \cup \{w\}$, that is, if there is a rational function f such that $|f(w)| > \max\{\,|f(z)| : z \in K\,\}$. In this definition, the holomorphicity of f at the point w is unimportant, for if $f(w)$ is undefined, then one can slightly perturb the coefficients of f to make $|f(w)|$ a large finite number without changing the values of f on K very much.

Example 7. Every compact set K in $\mathbb{C}$ is rationally convex. Indeed, if w is a point outside K, then the rational function $1/(z-w)$ blows up at w, so w is not in the rationally convex hull of K. Indeed, for a suitably small positive ε, the rational function $1/(z-w-\varepsilon)$ has larger absolute value at w than anywhere on K.

There is some awkwardness in talking about rational functions of two or more variables, because the singularities can be either poles (where the absolute value blows up) or points of indeterminacy (like the origin for the function z_1/z_2). A convenient device is to rephrase the notion of rational convexity by using only polynomials, as follows.

The notion of polynomial convexity is based on separating a point w from a compact set K by the absolute value of a polynomial; introducing the absolute value is natural in order to write inequalities. One can, however, consider the weaker separation property that a point w is separated from a compact set K if there is a polynomial p such that the image of w under p is not contained in the image of K under p. This weaker separation property turns out to be identical to the notion of rational convexity.

Indeed, if the point $p(w)$ does not belong to the set $p(K)$, then for every sufficiently small positive ε, the function $1/(p(z)-p(w)-\varepsilon)$ is a rational function of z that is holomorphic in a neighborhood of K and has larger absolute value at w than anywhere on K. Conversely, if f is a rational function, holomorphic on $K \cup \{w\}$, whose absolute value separates w from K, then the function $1/(f(z)-f(w))$ is a rational function of z that is holomorphic on K and singular at w. This function can be rewritten as a quotient of polynomials, and the denominator will be a polynomial that is zero at w and nonzero on K. Thus, a point w is in the rationally convex hull of a compact set K if and only if every polynomial that is equal to zero at w also has a zero on K.

Exercise 23. The rationally convex hull of a compact subset of $\mathbb{C}^n$ is again a compact subset of $\mathbb{C}^n$.

Example 8 (the Hartogs triangle). The open set $\{\,(z_1, z_2) \in \mathbb{C}^2 : |z_1| < |z_2| < 1\,\}$ is convex with respect to the linear fractional functions, because through each boundary point passes a complex line that does not intersect the domain. Indeed, the line on which $z_2 = 0$ serves at the origin $(0, 0)$; at any other boundary point where the two coordinates have equal absolute value, there is some value of θ for which a suitable line is the one that sends the complex parameter λ to $(\lambda, e^{i\theta}\lambda)$; and at a boundary point where the second coordinate has absolute value equal to 1, there is some value of θ such that a suitable line is the one on which $z_2 = e^{i\theta}$.

Functionals, Birkhäuser, 2004, Example 2.1.7.

In particular, the open Hartogs triangle is a rationally convex domain, since there are more rational functions than there are linear fractions. The open Hartogs triangle is certainly not polynomially convex. Indeed, consider the circle $\{(0, \frac{1}{2}e^{i\theta}) : 0 \leq \theta < 2\pi\}$. No point of the disc bounded by this circle can be separated from the circle by a polynomial, so the polynomial hull of the circle with respect to the open Hartogs triangle is not a compact subset of the triangle.

Next consider the *closed* Hartogs triangle, the set where $|z_1| \leq |z_2| \leq 1$. The rationally convex hull of this compact set is the whole closed bidisc. Indeed, if p is a polynomial having no zero on the closed Hartogs triangle, then by continuity, p has no zero in an open neighborhood of the closed triangle. Consequently, the reciprocal $1/p$ is holomorphic in some Hartogs figure, so by Theorem 3, the function $1/p$ extends to be holomorphic on the whole (closed) bidisc. Therefore the polynomial p cannot have any zeros in the bidisc. Accordingly, the rational hull of the closed Hartogs triangle contains the whole bidisc. The rational hull cannot contain any other points, since the rational hull is a subset of the convex hull.

3.2.3 Holomorphic convexity

Suppose that G is a domain in $\mathbb{C}^n$, and $\mathcal{F}$ is the class of holomorphic functions on G, which is commonly denoted $\mathcal{O}(G)$. Then $\mathcal{F}$-convexity, that is, $\mathcal{O}(G)$-convexity, is called *holomorphic convexity* (with respect to G).

Example 9. When $G = \mathbb{C}^n$, holomorphic convexity is equivalent to polynomial convexity, since every entire function can be approximated uniformly on compact sets by polynomials (namely, by partial sums of the Maclaurin series).

If $G_1 \subset G_2$, and K is a compact subset of G_1, then the holomorphically convex hull of K with respect to G_1 evidently is a subset of the holomorphically convex hull of K with respect to G_2 (because there are more holomorphic functions on G_1 than there are on the larger domain G_2). So if K is $\mathcal{O}(G_2)$ convex, then K is $\mathcal{O}(G_1)$ convex. In particular, a polynomially convex compact set is holomorphically convex with respect to every domain G that contains it; so is a convex set.

Example 10. Let K be the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ in the complex plane.

(a) Suppose that G is the whole plane. The $\mathcal{O}(G)$-convex hull of K is the closed unit disc (by the maximum principle).

(b) Suppose that G is the punctured plane $\{z \in \mathbb{C} : z \neq 0\}$. Then K is an $\mathcal{O}(G)$-convex set (because the function $1/z$, which is holomorphic on G, separates points inside the circle from points on the circle).

This example demonstrates that the notion of holomorphic convexity of a compact subset K of G depends both on K and on G.

Exercise 24. Show that if K is a holomorphically convex compact subset of G, and p is an arbitrary point of G, then the union $K \cup \{p\}$ is a holomorphically convex compact subset of G.

The next theorem solves the problem of characterizing holomorphically convex domains in $\mathbb{C}^n$. Although the theorem is valid for all values of n, the interesting case occurs when $n > 1$, for Example 7 implies that every domain in the complex plane is holomorphically convex. The theory of holomorphic convexity is due to Henri Cartan and Peter Thullen.[17]

Theorem 13. *The following properties of a domain G in $\mathbb{C}^n$ are equivalent.*

1. *The domain G is holomorphically convex (that is, for every compact set K contained in G, the holomorphically convex hull $\widehat{K}_{\mathcal{O}(G)}$ is again a compact subset of G).*
2. *For every sequence $\{p_j\}$ of points in G having no accumulation point inside G, there exists a holomorphic function f on G such that $\lim_{j\to\infty} |f(p_j)| = \infty$.*
3. *For every sequence $\{p_j\}$ of points in G having no accumulation point inside G, there exists a holomorphic function f on G such that $\sup_j |f(p_j)| = \infty$.*
4. *For every compact set K contained in G and for every unit vector v in $\mathbb{C}^n$, the distance from K to the boundary of G in the direction v is equal to the distance from $\widehat{K}_{\mathcal{O}(G)}$ to the boundary of G in the direction v.*
5. *For every compact set K contained in G, the distance from K to the boundary of G is equal to the distance from $\widehat{K}_{\mathcal{O}(G)}$ to the boundary of G.*
6. *The domain G is a weak domain of holomorphy.*
7. *The domain G is a domain of holomorphy.*

Precise definitions of the final two items are needed before proving the theorem.

Domains of holomorphy

Cartan and Thullen define a domain of holomorphy to be a domain on which there exists a holomorphic function that does not extend holomorphically to any larger domain.[18] But to Cartan and Thullen, the word "domain" means a Riemann domain spread over $\mathbb{C}^n$ (the higher-dimensional analogue of a Riemann surface). To formulate the concept of domain of holomorphy without introducing the machinery of manifolds requires some acrobatics. The next two examples illustrate why the definition is necessarily convoluted.

---

[17] See the article cited on page 20.

[18] "Einen Bereich $\mathfrak{B}$ nennen wir einen *Regularitätsbereich (domaine d'holomorphie)*, falls es eine in $\mathfrak{B}$ eindeutige und reguläre Funktion $f(z_1, \dots, z_n)$ gibt derart, daß jeder $\mathfrak{B}$ enthaltende Bereich $\mathfrak{B}'$, in dem $f(z_1, \dots, z_n)$ eindeutig und regulär ist, notwendig mit $\mathfrak{B}$ identisch ist." Cartan and Thullen, loc. cit., p. 618.

Example 11. There is a holomorphic branch of the single-variable function $\sqrt{z}$ on the slit plane $\mathbb{C} \setminus \{ z : \operatorname{Im} z = 0 \text{ and } \operatorname{Re} z \le 0 \}$. This function is discontinuous at all points of the negative part of the real axis, so the function certainly does not extend to be holomorphic in a neighborhood of any of these points. But the function $\sqrt{z}$ does continue holomorphically across each nonzero boundary point *from one side* (indeed, from either side). The natural domain of definition of $\sqrt{z}$ is not the slit plane but rather a two-sheeted Riemann surface.

In the preceding example in $\mathbb{C}^1$, there are functions that fail to admit extension to a full neighborhood of any boundary point yet admit one-sided extensions across some boundary points. Nonetheless, there are holomorphic functions on the slit plane that fail to admit even one-sided extensions. (Apply the Weierstrass theorem to construct a holomorphic function on the slit plane whose zeros accumulate at every point of the slit from both sides.) So the example is not decisive. In the following example[19] in $\mathbb{C}^2$, there is a part of the boundary across which *all* holomorphic functions admit one-sided extension, yet some holomorphic function fails to admit extension to a full neighborhood of those boundary points.

Example 12. Consider the following three product domains. Let G_1 be the Cartesian product of the open unit disk in the z_2 plane with the z_1 plane slit along the negative part of the real axis. Let G_2 be the Cartesian product of the complement of the closed unit disk in the z_2 plane with the z_1 plane slit along the positive part of the real axis. Let G_3 be the Cartesian product of the whole z_2 plane with the open upper half of the z_1 plane. Notice that G_1 is disjoint from G_2, but G_3 intersects both G_1 and G_2. Accordingly, the union $G_1 \cup G_2 \cup G_3$ is a connected open subset of $\mathbb{C}^2$. An alternative description of this domain is the complement in $\mathbb{C}^2$ of the union of the following three closed subsets:

$$\begin{aligned} &\{ (z_1, z_2) : \operatorname{Im} z_1 = 0 \text{ and } \operatorname{Re} z_1 \le 0 \text{ and } |z_2| \le 1 \} \\ \text{and } &\{ (z_1, z_2) : \operatorname{Im} z_1 = 0 \text{ and } \operatorname{Re} z_1 \ge 0 \text{ and } |z_2| \ge 1 \} \\ \text{and } &\{ (z_1, z_2) : \operatorname{Im} z_1 \le 0 \text{ and } |z_2| = 1 \}. \end{aligned}$$

The first claim is that every holomorphic function on $G_1 \cup G_2 \cup G_3$ extends holomorphically from the open subset where $\operatorname{Re} z_1 < 0$ and $\operatorname{Im} z_1 > 0$ and z_2 is arbitrary to the open half-space where $\operatorname{Re} z_1 < 0$ and $\operatorname{Im} z_1$ is arbitrary and z_2 is arbitrary. In particular, all holomorphic functions admit extension from one side across the boundary points where $\operatorname{Re} z_1 < 0$ and $\operatorname{Im} z_1 = 0$ and $|z_2| \le 1$. Indeed, suppose f is a holomorphic function on $G_1 \cup G_2 \cup G_3$. If the value of z_2 is arbitrary, and the radius r is strictly greater than $\max\{1, |z_2|\}$, and $\operatorname{Re} z_1 < 0$, then the integral

$$\frac{1}{2\pi i} \int_{|\zeta|=r} \frac{f(z_1, \zeta)}{\zeta - z_2} \, d\zeta$$

makes sense and determines a holomorphic function of z_1 and z_2. Since the value of the integral does not change when r increases, the integral defines a holomorphic function in

[19]The example is modified from one in the book of B. V. Shabat, *Introduction to Complex Analysis: Part II, Functions of Several Variables*, translated from the Russian by J. S. Joel, American Mathematical Society, 1992. See Chapter III, §12, subsection 33, pages 177–178.

the half-space where $\operatorname{Re} z_1 < 0$. In the part of the half-space where $\operatorname{Im} z_1 > 0$, the integral recovers the value $f(z_1, z_2)$ of the original function (by the single-variable Cauchy integral formula). Therefore the integral defines the required extension of f to the half-space. (This argument is the same as in the discussion of the Hartogs phenomenon, and Theorem 3 could be quoted instead of repeating the argument.)

The second claim is the existence of a holomorphic function on $G_1 \cup G_2 \cup G_3$ that does not admit a two-sided extension across any boundary point where $\operatorname{Re} z_1 < 0$ and $\operatorname{Im} z_1 = 0$ and $|z_2| < 1$. To construct such a function, start with the principal branch of $\sqrt{z_1}$ on the z_1 plane slit along the negative part of the real axis. (In other words, choose the value of $\sqrt{z_1}$ to lie in the right-hand half of the z_1 plane.) Extend this function to be independent of z_2 on G_1. Next consider on the z_1 plane slit along the positive part of the real axis the branch of $\sqrt{z_1}$ for which the argument of z_1 is chosen to lie between 0 and 2π. (In other words, choose the value of $\sqrt{z_1}$ to lie in the upper half of the z_1 plane.) Extend this function to be independent of z_2 on G_2. Observe that these two branches of $\sqrt{z_1}$ agree when $\operatorname{Im} z_1 > 0$. Extend this common branch to be independent of z_2 on G_3. Since the indicated holomorphic functions on G_1 and G_3 match on the intersection $G_1 \cap G_3$, and the functions on G_2 and G_3 match on $G_2 \cap G_3$, the construction provides a well-defined holomorphic function on $G_1 \cup G_2 \cup G_3$. When this holomorphic function is extended across boundary points where $\operatorname{Re} z_1 < 0$ and $\operatorname{Im} z_1 = 0$ and $|z_2| < 1$ from the side where $\operatorname{Im} z_1 > 0$, the result is a branch of $\sqrt{z_1}$ taking values in the second quadrant, hence equal to the negative of the original function. Thus the extension is not two-sided.

Of course, the boundary of a general domain can be much more complicated than the one in the preceding example. For instance, the boundary need not be locally connected (think of a comb). So the notion of one-sided extension is too simple to cover the general situation. One way to capture the full complexity without entering into the world of manifolds is the following definition.

A holomorphic function f on a domain G is called *completely singular* at a boundary point p if for every connected open neighborhood U of p, there does *not* exist a holomorphic function F on U that agrees with f on *some* nonvoid open subset of $U \cap G$ (equivalently, on some connected component of $U \cap G$). A completely singular function[20] is "holomorphically non-extendable" in the strongest possible way.

A domain G in $\mathbb{C}^n$ is called a *domain of holomorphy* if there exists some holomorphic function on G that is completely singular at every boundary point of G. This property appears to be hard to verify in concrete cases. A presumably easier property to check is the existence for each boundary point p of a holomorphic function on G that is completely singular at p (possibly a different function for each choice of the boundary point p). A domain satisfying

[20]The terminology "completely singular" is not completely standard. Two books using this terminology are *Holomorphic Functions and Integral Representations in Several Complex Variables* by R. Michael Range (Springer, 1986); and *From Holomorphic Functions to Complex Manifolds* by Klaus Fritzsche and Hans Grauert (Springer, 2002).

this (apparently less restrictive) property is called[21] a *weak domain of holomorphy*.

Example 13. Convex domains are weak domains of holomorphy. Indeed, at each boundary point there is an affine complex linear function that is zero at the boundary point but nonzero inside the domain. The reciprocal of the function is then holomorphic inside and completely singular at the specified boundary point. Exhibiting a holomorphic function that is singular at *every* boundary point of a convex domain presents a more difficult problem.

Proof of Theorem 13. All of the properties hold for elementary reasons when $G = \mathbb{C}^n$, so assume that the boundary of G is not empty. Much of the proof is merely point-set topology. Complex analysis enters through the representation of functions by power series.

Certain implications are easy. Evidently (2) $\Longrightarrow$ (3), and (7) $\Longrightarrow$ (6).

Some notation is useful to discuss properties (4) and (5). When z is a point in G, let $d(z)$ denote $\inf_{w\in\mathbb{C}^n\setminus G} \|z-w\|$, the distance from z to the boundary of G. Similarly let $d(S)$ denote the distance from a subset S of G to the boundary of G, namely, $\inf\{\, d(z) : z \in S\,\}$. When v is a unit vector in $\mathbb{C}^n$, and z is a point of G, let $d_v(z)$ denote

$$\sup\{\, r \in \mathbb{R} : z + \lambda v \in G \text{ when } \lambda \in \mathbb{C} \text{ and } |\lambda| < r\,\}.$$

The quantity $d_v(z)$ (which could be infinite when G is unbounded) represents the radius of the "largest" one-dimensional complex disc with center z and direction v that fits inside G. (The quotation marks are present because the supremum is not attained.) When S is a subset of G, let $d_v(S)$ denote $\inf\{\, d_v(z) : z \in S\,\}$. This quantity represents the distance from S to the boundary of G in the (complex) direction of the unit vector v. Evidently $d(z) = \inf\{\, d_v(z) : \|v\| = 1\,\}$, so $d(S) = \inf\{\, d_v(S) : \|v\| = 1\,\}$ for every set S. Therefore (4) $\Longrightarrow$ (5).

Also elementary is the implication that (5) $\Longrightarrow$ (1). For if property (5) holds, then $\widehat{K}_{\mathcal{O}(G)}$ is a relatively closed subset of G that has positive distance from the boundary of G. Consequently, the set $\widehat{K}_{\mathcal{O}(G)}$ is closed as a subset of $\mathbb{C}^n$. Moreover, the holomorphically convex hull $\widehat{K}_{\mathcal{O}(G)}$ is a subset of the ordinary convex hull of K, hence is a bounded subset of $\mathbb{C}^n$. Being both closed and bounded, the set $\widehat{K}_{\mathcal{O}(G)}$ is compact.

To show that (1) $\Longrightarrow$ (2), suppose that G is holomorphically convex, and let $\{p_j\}$ be a sequence of points of G having no accumulation point inside G. The first goal is to construct an increasing sequence $\{K_j\}$ of holomorphically convex compact sets that exhausts G and a sequence $\{q_j\}$ of distinct points of G such that $q_j \in K_{j+1} \setminus K_j$ for each j, and the point sets $\{p_j\}$ and $\{q_j\}$ are identical. (The sequence $\{q_j\}$ is a reordering of the sequence $\{p_j\}$ after any repeated points in the sequence $\{p_j\}$ are removed.)

Suppose for a moment that this construction has been accomplished. The definition of holomorphic convexity guarantees (by a routine induction) the existence for each j of a holomorphic function f_j on G such that $|f_j(z)| < 2^{-j}$ when $z \in K_j$, and $|f_j(q_j)| > j + \sum_{k=1}^{j-1} |f_k(q_j)|$. The infinite series $\sum_{j=1}^{\infty} f_j$ converges uniformly on each compact subset of G to a holomorphic function f such that $|f(q_j)| > j-1$ for each j. Thus $\lim_{j\to\infty} |f(q_j)| = \infty$,

[21]The notion of "weak domain of holomorphy" appears in the two books just cited.

so $\lim_{j\to\infty} |f(p_j)| = \infty$ as well, since the two sequences are essentially the same except for the order of terms.

The necessary construction can be accomplished as follows. For each positive integer m, the set

$$\{ z \in G : \|z\| \leq m \text{ and } d(z) \geq 1/m \}$$

is a compact subset of G. Denote the holomorphically convex hull of this set by L_m, which is again a compact subset of G by hypothesis. Let K_1 be the empty set. Let m_1 be an index for which the set L_{m_1} contains some points of the sequence $\{p_j\}$ (necessarily finitely many, since the sequence has no accumulation point in G). Arrange these points in a list (ignoring repetitions), say $q_1, \ldots, q_{k_1}$. For each j between 1 and k_1, let K_{j+1} be the finite set $\{q_1, \dots, q_j\}$. Choose an index m_2 larger than m_1 for which the set $L_{m_2} \setminus L_{m_1}$ contains points of the sequence $\{p_j\}$. Label these points $q_{k_1+1}, \dots, q_{k_2}$. For each j between $k_1 + 1$ and k_2, let K_{j+1} be the set $L_{m_1} \cup \{q_{k_1+1}, \dots, q_j\}$ (which is holomorphically convex by Exercise 24). Continue recursively to obtain the required sequences $\{K_j\}$ and $\{q_j\}$. Thus (1) $\implies$ (2).

To prove that (3) $\implies$ (1), let K be an arbitrary compact subset of G. Every holomorphic function on G is bounded on K, hence on $\widehat{K}_{\mathcal{O}(G)}$. Consequently, property (3) implies that every sequence in $\widehat{K}_{\mathcal{O}(G)}$ must have an accumulation point in G. But $\widehat{K}_{\mathcal{O}(G)}$ is relatively closed in G by definition, so the accumulation point lies in $\widehat{K}_{\mathcal{O}(G)}$. Thus $\widehat{K}_{\mathcal{O}(G)}$ is sequentially compact.

The proof that (2) $\implies$ (7) is purely point-set topology. For each positive integer k, the set

$$\{ z \in G : 2^{-(k+1)} \leq d(z) \leq 2^{-k} \text{ and } \|z\| \leq 2^k \}$$

is compact, so this set can be covered by a finite number of open balls of radius $2^{-(k+2)}$ with centers in the set. Collect these centers for every k and arrange them in a sequence $\{p_j\}$. Every compact subset of G has positive distance from the boundary of G and so contains only finitely many points of this sequence. Thus the sequence has no accumulation point inside G. On the other hand, the sequence evidently has every boundary point of G as an accumulation point. The claim is that even more is true: if U is an arbitrary connected open set that intersects the boundary of G, and V is a component of $U \cap G$, then the points of the sequence $\{p_j\}$ that lie in V accumulate at every boundary point of V that is contained in U. (There *are* boundary points of V inside U, for in the contrary case, the open set V would be relatively closed in U, contradicting the connectedness of U. Boundary points of V that are inside U are necessarily boundary points of G as well.)

To verify the claim, let q be a point of U on the boundary of V. Choose N so large that U contains the ball centered at q of radius 2^{-N}, and also $\|q\| < 2^N$. Suppose m is an arbitrary integer larger than N. Let q' be a point in the intersection of V and the ball of radius 2^{-m} centered at q. There is an integer k (at least as large as m) for which

$$2^{-(k+1)} \leq d(q') \leq 2^{-k}.$$

By construction, some point of the sequence $\{p_j\}$ has distance from q' less than $2^{-(k+2)}$. This point lies in V because the open ball centered at q' of radius $2^{-(k+2)}$ is entirely contained in V

(since this ball is contained in $U \cap G$ and intersects V, which is a component of $U \cap G$). Thus there is a subsequence of $\{p_j\}$ that lies in V and converges to the arbitrary boundary point q.

Property (2) provides a holomorphic function that blows up along the sequence $\{p_j\}$. This function evidently is completely singular at every boundary point of G. Thus (2) $\Longrightarrow$ (7).

Next consider the implication that (6) $\Longrightarrow$ (5). Evidently $d\big(\widehat{K}_{\mathcal{O}(G)}\big) \le d(K)$, since $K \subseteq \widehat{K}_{\mathcal{O}(G)}$. Seeking a contradiction, suppose that $d\big(\widehat{K}_{\mathcal{O}(G)}\big)$ is strictly less than $d(K)$. Then there is a point w in $\widehat{K}_{\mathcal{O}(G)}$ and a point p in the boundary of G such that $\|w-p\| < d(K)$. Consequently, there is an n-tuple $(r_1, \dots, r_n)$ of positive radii such that the open polydisc centered at w with polyradius r equal to $(r_1, \dots, r_n)$ contains the point p, yet for every point z in K, the closed polydisc centered at z with polyradius r is contained in G.

Under the hypothesis that G is a weak domain of holomorphy, there is a holomorphic function f on G that is completely singular at p. The union of the closed polydiscs with polyradius r centered at points of the compact set K is a compact subset of G, so the function f is bounded on this set by some constant M. By Cauchy's estimates for derivatives (these inequalities follow from the iterated Cauchy integral on polydiscs),

$$|f^{(\alpha)}(z)| \le \frac{M\alpha!}{r^\alpha} \qquad \text{for } z \text{ in } K \text{ and for every multi-index } \alpha.$$

Since $w \in \widehat{K}_{\mathcal{O}(G)}$, the same inequalities hold with z replaced by w. Therefore the Taylor series for f centered at w converges in the interior of the polydisc centered at w with polyradius r (by comparison with a product of convergent geometric series).

Thus f fails to be completely singular at p. In view of this contradiction, the supposition that $d\big(\widehat{K}_{\mathcal{O}(G)}\big) < d(K)$ is untenable. Accordingly, (6) $\Longrightarrow$ (5).

The proof that (3) $\Longrightarrow$ (4) is similar. Seeking a contradiction, suppose there is a unit vector v in $\mathbb{C}^n$ and a point w in $\widehat{K}_{\mathcal{O}(G)}$ such that $d_v(w) < d_v(K)$. Let λ_0 be a complex number of absolute value equal to $d_v(w)$ such that $w + \lambda_0 v$ lies in the boundary of G. Apply property (3) to produce a holomorphic function f on G that is unbounded on the sequence $\{w + \frac{j}{j+1}\lambda_0 v\}$. Then the function of one complex variable that sends λ to $f(w + \lambda v)$ has a Maclaurin series with radius of convergence equal to $d_v(w)$. The goal now is to obtain a contradiction by showing that the radius of convergence actually is larger than $d_v(w)$.

Choose a number r strictly between $d_v(w)$ and $d_v(K)$. The set of points

$$\{\, z + \lambda v : z \in K \text{ and } |\lambda| \le r \,\}$$

is a compact subset of G, so the function f is bounded on this set, say by M. Cauchy's estimate for derivatives implies that when $z \in K$, the kth Maclaurin coefficient of the single-variable function sending λ to $f(z + \lambda v)$ is bounded by M/r^k. By the chain rule, the Maclaurin coefficient is the value at z of a linear combination of partial derivatives of f, hence is a holomorphic function on G. Since $w \in \widehat{K}$, the corresponding Maclaurin coefficient of the function that sends λ to $f(w + \lambda v)$ admits the same bound M/r^k. But k is arbitrary, so $f(w + \lambda v)$ is a holomorphic function of λ at least in an open disk of radius r. This deduction contradicts that the radius of convergence is equal to $d_v(w)$. Thus (3) $\Longrightarrow$ (4).

Putting together the above deductions shows that (1) $\Longrightarrow$ (2) $\Longrightarrow$ (7) $\Longrightarrow$ (6) $\Longrightarrow$ (5) $\Longrightarrow$ (1), and also (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) $\Longrightarrow$ (5) $\Longrightarrow$ (1). Thus all seven statements are equivalent. □

After one singular function is known to exist, Baire's magic wand can be applied to show that most functions are singular. The next result is essentially a corollary of Theorem 13.

Theorem 14. *The following properties of a domain G in $\mathbb{C}^n$ are equivalent.*

1. *For every boundary point p, there exists a holomorphic function on G that is completely singular at p. (In other words, G is a weak domain of holomorphy.)*
2. *For every boundary point p, most (in the sense of Baire category) holomorphic functions on G are completely singular at p.*
3. *There exists a holomorphic function on G that is completely singular at every boundary point of G. (In other words, G is a domain of holomorphy.)*
4. *Most (in the sense of Baire category) holomorphic functions on G are completely singular at every boundary point.*

As observed on page 23 in the proof of Theorem 4, the space of holomorphic functions on a domain G is a complete metric space. (Convergence with respect to the metric is the same as uniform convergence on compact sets.) The word "most" in the statement of Theorem 14 means "a residual set, that is, the complement of a set of first Baire category."

Proof of Theorem 14. Evidently (4) $\Longrightarrow$ (3) $\Longrightarrow$ (1) and (4) $\Longrightarrow$ (2) $\Longrightarrow$ (1). What remains to show is that (1) $\Longrightarrow$ (4). A particular consequence is that properties (1) and (3) are equivalent (which was already demonstrated in the proof of Theorem 13).

Suppose, then, that property (1) holds. Let U be a connected open set that intersects the boundary of G, and let V be a component of the intersection $G \cap U$. By hypothesis, there exists a holomorphic function on G that cannot be extended holomorphically from V to U.

The first claim is that *most* holomorphic functions on G do not extend holomorphically from V to U. The vector space of holomorphic functions on G is not only a complete metric space but also an F-space or Fréchet space (meaning that the vector space operations are continuous, and the metric is invariant under translation). A standard notation for the space of holomorphic functions on G is $\mathcal{O}(G)$. The subspace of $\mathcal{O}(G)$ consisting of functions that extend holomorphically from V to U can be viewed as a Fréchet space whose metric is the sum of the metrics from $\mathcal{O}(G)$ and $\mathcal{O}(U)$; this subspace is embedded continuously into $\mathcal{O}(G)$. By the hypothesis from the preceding paragraph, the image of the embedding is not the whole of $\mathcal{O}(G)$, so by a theorem from functional analysis, the image is of first Baire category.[22] Thus the functions in a residual set in $\mathcal{O}(G)$ cannot be extended holomorphically from V to U.

[22]The argument is the same as the one on page 26. The theorem from Banach's book cited there applies, or one could invoke the version of the open mapping theorem from Walter Rudin's book *Functional Analysis* (section 2.11, page 48 of the second edition): a continuous linear mapping between Fréchet spaces either is a surjective open map or has image of first category.

To strengthen the conclusion, choose a countable dense set of points in the boundary of G. For each point, choose a countable neighborhood basis of open balls centered at the point, say the balls whose radii are reciprocals of positive integers. The intersection of each ball with G has either a finite or a countably infinite number of connected components. Arrange the collection of components from all balls and all points into a countable list $\{V_j\}_{j=1}^{\infty}$. According to what was just shown, the set of holomorphic functions on G that extend holomorphically from a particular V_j to the corresponding ball is a set of first category in $\mathcal{O}(G)$. Therefore the set of holomorphic functions on G that extend from any V_j whatsoever is a countable union of sets of first category, hence still a set of first category. In other words, the complementary set of holomorphic functions on G that extend from no V_j to the corresponding ball is a residual set.

What remains to check is the plausible assertion that every member of this residual set of holomorphic functions is completely singular at every boundary point. Indeed, if U is an arbitrary connected open set that intersects the boundary of G, and V is a component of $G \cap U$, then some ball in the constructed sequence simultaneously is contained in U and is centered at a boundary point of V. Some V_j corresponding to this ball is a subset of V, so all of the functions in the indicated residual set fail to extend holomorphically from V to U. Thus every function in the residual set is completely singular at every boundary point of G. □

Exercise 25. For each of the following subsets of $\mathbb{C}^2$, determine if the subset is a domain of holomorphy.

(a) The complement of the point $(0, 0)$.

(b) The complement of the real line $\{ (x, 0) \in \mathbb{C}^2 : x \in \mathbb{R} \}$.

(c) The complement of the complex line $\{ (z, 0) \in \mathbb{C}^2 : z \in \mathbb{C} \}$.

(d) The complement of the totally real 2-plane $\{ (x_1, x_2) \in \mathbb{C}^2 : x_1 \in \mathbb{R} \text{ and } x_2 \in \mathbb{R} \}$.

(e) The complement of the complex half-line $\{ (x + iy, 0) : x \in \mathbb{R} \text{ and } y \geq 0 \}$.

Exercise 26. (a) Is the union of two domains of holomorphy again a domain of holomorphy?

(b) Is (each connected component of) the intersection of two domains of holomorphy again a domain of holomorphy?

(c) If G_1 is a domain of holomorphy in $\mathbb{C}^{n_1}$, and G_2 is a domain of holomorphy in $\mathbb{C}^{n_2}$, is the Cartesian product $G_1 \times G_2$ a domain of holomorphy in $C^{n_1+n_2}$?

(d) Show that holomorphic convexity is a biholomorphically invariant property. In other words, if $f : G_1 \to G_2$ is a bijective holomorphic map, then G_1 is a domain of holomorphy if and only if G_2 is a domain of holomorphy.

(e) Suppose G is a domain of holomorphy in $\mathbb{C}^n$, and $f : G \to \mathbb{C}^n$ is a holomorphic map (not necessarily either injective or surjective). If the image $f(G)$ is an open set in $\mathbb{C}^n$, must the image be a domain of holomorphy?

(f) Suppose G is a domain of holomorphy in $\mathbb{C}^n$, and $f : G \to \mathbb{C}^k$ is a holomorphic map (not necessarily either injective or surjective). Show that if D is a domain of holomorphy in $\mathbb{C}^k$, then (each connected component of) the inverse image $f^{-1}(D)$ is a domain of holomorphy in $\mathbb{C}^n$.

(g) Show that if $f_1, \ldots, f_k$ are holomorphic functions defined on a holomorphically convex domain G, then each connected component of $\{ z \in G : \sum_{j=1}^k |f_j(z)| < 1 \}$ is a domain of holomorphy.

3.2.4 Pseudoconvexity

Pseudoconvexity means convexity with respect to a certain class of real-valued functions that Kiyoshi Oka[23] called "pseudoconvex functions." Pierre Lelong[24] called these functions "plurisubharmonic functions," and this terminology is the one that has become standard. The discussion had better start with the base case of dimension 1.

Subharmonic functions

A function u that is defined on an open subset of the complex plane $\mathbb{C}$ and that takes values in $[-\infty, \infty)$ is called *subharmonic* if firstly u is upper semicontinuous, and secondly u satisfies one of the following equivalent properties.

1. For every point a in the domain of u, there is a radius $r(a)$ such that u satisfies the sub-mean-value property on every disc of radius ρ less than $r(a)$:

$$u(a) \le \frac{1}{2\pi} \int_0^{2\pi} u(a + \rho e^{i\theta})\, d\theta.$$

2. The function u satisfies the sub-mean-value property on every closed disc contained in the domain of u.

3. For every closed disc D contained in the domain of u and every function h that is harmonic on a neighborhood of D, if $u \le h$ on the boundary of D, then $u \le h$ on all of D.

[23] Kiyoshi Oka, Sur les fonctions analytiques de plusieurs variables. VI. Domaines pseudoconvexes, *Tôhoku Mathematical Journal* **49** (1942) 15–52. See section 11 of the paper.

[24] Pierre Lelong, Définition des fonctions plurisousharmoniques, *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **215** (1942) 398–400; Sur les suites de fonctions plurisousharmoniques, *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **215** (1942) 454–456.

4. For every compact set K contained in the domain of u and for every function h that is harmonic on a neighborhood of K, if $u \le h$ on the boundary of K, then $u \le h$ on all of K.

5. If Δ denotes the Laplace operator $\frac{d^2}{dx^2} + \frac{d^2}{dy^2}$, then $\Delta u \ge 0$. (If u does not have second derivatives in the classical sense, then Δu is understood in the sense of distributions.)

That these properties are equivalent is shown in textbooks on the theory of functions of one complex variable. Some authors disallow subharmonic functions to be constantly equal to $-\infty$ on a component of the domain of the function.

A fundamental example of a subharmonic function is $|f|$ when f is holomorphic. Since a holomorphic function has the mean-value property, the absolute value of the function has the sub-mean-value property because the absolute value of an integral does not exceed the integral of the absolute value.

Another elementary example of a subharmonic function in $\mathbb{C}$ is $\log|z|$. This function is even harmonic when $z \ne 0$, so the mean-value property holds on small discs centered at nonzero points; and the sub-mean-value property holds automatically at 0, because the function takes the value $-\infty$ at 0. Since the class of harmonic functions is preserved under composition with a holomorphic function, property 4 implies that the class of subharmonic functions is preserved too. In particular, $\log|f|$ is subharmonic when f is holomorphic.

The following two useful lemmas about subharmonic functions can be proved from first principles.

Lemma 5. If u is subharmonic, then the integral of u on a circle is a (weakly) increasing function of the radius. In other words,

$$\int_0^{2\pi} u(a + r_1 e^{i\theta})\, d\theta \le \int_0^{2\pi} u(a + r_2 e^{i\theta})\, d\theta \quad \text{when} \quad 0 < r_1 \le r_2.$$

Lemma 6. A subharmonic function defined on a connected open subset of $\mathbb{C}$ is either locally integrable or identically equal to $-\infty$.

Proof of Lemma 5. Since u is upper semicontinuous, there is for each positive ε a continuous function h on the circle of radius r_2 such that $u < h < u + \varepsilon$ on this circle. By solving a Dirichlet problem, one may assume that h is harmonic inside the disc of radius r_2, or, after slightly dilating the coordinates, in a neighborhood of the closed disc. Then $u < h$ on the circle of radius r_1, since u is subharmonic, so $\int_0^{2\pi} u(a + r_1 e^{i\theta})\, d\theta < \int_0^{2\pi} h(a + r_1 e^{i\theta})\, d\theta = 2\pi h(a) = \int_0^{2\pi} h(a + r_2 e^{i\theta})\, d\theta < 2\pi\varepsilon + \int_0^{2\pi} u(a + r_2 e^{i\theta})\, d\theta$. Let ε go to 0 to obtain the required inequality. □

Proof of Lemma 6. An upper semicontinuous function is locally bounded above, so what needs to be proved is that the integral of a subharmonic function u on a disc is not equal to $-\infty$ unless the function is identically equal to $-\infty$.

Suppose a is a point at which $u(a) \neq -\infty$. The sub-mean-value property implies that

$$\frac{1}{2\pi} \int_0^{2\pi} u(a + re^{i\theta})\, d\theta \geq u(a)$$

when the closed disc centered at a of radius r is contained in the domain of u. Multiplying by r and integrating with respect to r shows that $|D|^{-1} \int_D u \geq u(a)$ for every disc D centered at a (where $|D|$ denotes the area of D). Hence u is locally integrable in a neighborhood of every point of the disc D.

On the other hand, if b is a point such that $u(b) = -\infty$, but u is not identically equal to $-\infty$ in a neighborhood of b, then there is a point a closer to b than to the boundary of the domain of definition of u and with the property that $u(a) \neq -\infty$. By the previous observation, the function u is integrable in a neighborhood of b.

The preceding two paragraphs show that the set of points such that u is integrable in a neighborhood of the point is both open and relatively closed. Therefore the function u, if not identically equal to $-\infty$, is locally integrable in a neighborhood of every point of the domain. □

Exercise 27. (a) The sum of two subharmonic functions is subharmonic.

(b) If u is subharmonic and c is a positive constant, then cu is subharmonic.

(c) If u_1 and u_2 are subharmonic, then so is the pointwise maximum of u_1 and u_2.

(d) If φ is an increasing convex function on the range of a subharmonic function u, then the composite function $\varphi \circ u$ is subharmonic. [A useful special case is $\varphi(|z|)$.]

Some care is needed in handling infinite processes involving subharmonic functions. Easy examples show that two things could go wrong when taking the pointwise supremum of an infinite family of subharmonic functions. If $f_k(z)$ is the constant function k, then the sequence $\{f_k\}$ of subharmonic functions has limit $+\infty$, which is not an allowed value for an upper semicontinuous function. If $f_k(z) = \frac{1}{k} \log |z|$, then the sequence $\{f_k\}$ of subharmonic functions on the unit disc has pointwise supremum equal to 0 when $z \neq 0$ and equal to $-\infty$ when $z = 0$; this limit function is not upper semicontinuous. The following exercise says that these two kinds of difficulties are the only obstructions to subharmonicity of a pointwise supremum.

Exercise 28. If u_α is a subharmonic function for each α in some index set A (possibly an uncountable set), and the pointwise supremum $\sup_{\alpha \in A} u_\alpha$ is a measurable function, then this pointwise supremum satisfies the sub-mean-value property. Consequently, the pointwise supremum is subharmonic if it is upper semicontinuous (which entails, in particular, that the supremum is nowhere equal to $+\infty$).

Taking a pointwise supremum of subharmonic functions is a process used in Perron's method for solving the Dirichlet problem in a planar domain.

Taking the maximum of two subharmonic functions always produces another one, but taking the minimum does not. For instance, $\min(1, |z|)$ does not have the sub-mean-value property at the point where $z = 1$. Nonetheless, monotonically decreasing sequences of subharmonic functions have subharmonic limits.

Theorem 15. *The pointwise limit of a decreasing sequence of subharmonic functions is again subharmonic. Moreover, every subharmonic function on an open set is, on each compact subset, the limit of a decreasing sequence of infinitely differentiable subharmonic functions.*

Proof. Observe that the limit u of a decreasing sequence $\{u_k\}_{k=1}^{\infty}$ of upper semicontinuous functions is still upper semicontinuous, because $\{\, z : u(z) < a \,\} = \bigcup_{k=1}^{\infty} \{\, z : u_k(z) < a \,\}$, and the union of open sets is open. Now if K is a compact subset of the domain of definition of the functions, and h is a harmonic function on K such that $u \le h$ on the boundary of K, then certainly $u < h + \varepsilon$ on the boundary of K for every positive ε. If z is a point of bK, then $u_k(z) < h(z) + \varepsilon$ for all sufficiently large k. The upper semicontinuity of u_k implies that $u_k(w) \le h(w) + 2\varepsilon$ for every w in a neighborhood of z. Since bK is compact, and the sequence of functions is decreasing, there is some k such that $u_k \le h + 2\varepsilon$ on all of bK. The subharmonicity of u_k then implies that $u_k \le h + 2\varepsilon$ on all of K. Therefore $u \le h + 2\varepsilon$ on K, and letting ε go to 0 shows that $u \le h$ on K. Hence the limit function u is subharmonic.

To prove the second part of the theorem, let u be a subharmonic function on a domain G in $\mathbb{C}$, and extend u to be identically equal to 0 outside G. Let φ be an infinitely differentiable, nonnegative function, with integral 1, supported in the unit ball, and depending only on the radius; and let $\varphi_\varepsilon(x)$ denote $\varepsilon^{-2}\varphi(x/\varepsilon)$. Let u_ε denote the convolution of u and φ_ε: namely, $u_\varepsilon(z) = \int_{\mathbb{C}} \varphi_\varepsilon(z-w)u(w)\,dA_w = \int_{\mathbb{C}} u(z-w)\varphi_\varepsilon(w)\,dA_w$, where dA denotes Lebesgue area measure in the plane. Thus the value of u_ε at a point is a weighted average of the values of u in an ε-neighborhood of the point.

The sub-mean-value property of subharmonic functions implies that $u(z) \le u_\varepsilon(z)$ at every point z whose distance from the boundary of G is at least ε. Moreover, Lemma 5 implies that on a compact subset of G, the functions u_ε decrease when ε decreases, once ε is smaller than the distance from the compact set to bG. Since u is upper semicontinuous, the average of u over a sufficiently small disc is arbitrarily little more than the value of u at the center of the disc; the decreasing limit of $u_\varepsilon(z)$ is therefore equal to $u(z)$. The first expression for the convolution shows that the functions u_ε are infinitely differentiable, for one can differentiate under the integral sign, letting the derivatives act on φ_ε. That u_ε is subharmonic follows by integrating u_ε on a circle, interchanging the order of integration, and invoking the subharmonicity of u. □

Here are two interesting examples that follow from the preceding considerations.

Example 14. Let $\{a_k\}_{k=1}^{\infty}$ be a bounded sequence of distinct points of $\mathbb{C}$, and define $u(z)$ to be $\sum_{k=1}^{\infty} 2^{-k} \log|z - a_k|$. Then u is a subharmonic function on the whole plane. Notice that the sequence $\{a_k\}$ could be dense in some compact set. For instance, the sequence could be the set of points in the unit square having rational coordinates.

To see why u is subharmonic, first suppose that z_0 is neither a point of the sequence nor a limit point of the sequence. Then the function $\log|z - a_k|$ is bounded above and below for z in a neighborhood of z_0, independently of k, so the series defining $u(z)$ converges uniformly in the neighborhood. The limit of a uniformly convergent series of harmonic functions is harmonic, so u is harmonic on the complement of the closure of the sequence $\{a_k\}$.

Now suppose that z_0 is a point in the closure of the sequence $\{a_k\}$. Split the sum defining $u(z)$ into two parts: the sum of the terms for which $|a_k - z_0| < 1/2$ and the sum of the terms for which $|a_k - z_0| \geq 1/2$. The second sum converges uniformly for z in a neighborhood of z_0 (as in the preceding paragraph) and represents a harmonic function there. The first sum is a sum of negative terms (for z in a neighborhood of z_0), so the partial sums form a decreasing sequence of subharmonic functions. By Theorem 15, the partial sums converge to a subharmonic function.

Thus u is subharmonic in the whole plane $\mathbb{C}$, and u takes the value $-\infty$ at every point of the sequence $\{a_k\}$. The set where u equals $-\infty$ necessarily is a set of measure zero, since the subharmonic function u is locally integrable by Lemma 6.

Example 15. Let G be a proper subdomain of the complex plane. If $d(z)$ denotes the distance from the point z to the boundary of G, then $-\log d(z)$ is a subharmonic function of z in G. Indeed, if a is a point of bG (the boundary of G), then $-\log|z - a|$ is a harmonic function on G. Since

$$\sup\{\,-\log|z-a| \,:\, a \in bG\,\} = -\inf\{\,\log|z-a| \,:\, a \in bG\,\} = -\log d(z),$$

the subharmonicity follows from Exercise 28.

Plurisubharmonic functions

Introduction An upper semicontinuous function defined on an open subset of $\mathbb{C}^n$ is called a *plurisubharmonic* function if the restriction to every complex line is subharmonic. Both the name and the fundamental properties of plurisubharmonic functions are due to Lelong.[25]

A domain in $\mathbb{C}^n$ that is convex with respect to the class of plurisubharmonic functions is called a *pseudoconvex* domain. It will be shown later that Example 15 generalizes to higher dimension: a proper subdomain of $\mathbb{C}^n$ is pseudoconvex if and only if the function $-\log d(z)$ is plurisubharmonic in the domain.

If f is holomorphic, then $|f|$ is plurisubharmonic. Consequently, the hull of a compact set with respect to the class of plurisubharmonic functions is no larger than the holomorphically convex hull. Therefore every holomorphically convex domain is pseudoconvex. The famous *Levi problem*, to be solved later, is to prove the converse: every pseudoconvex domain is a domain of holomorphy.

[25] The following seminal article develops the basic theory: Pierre Lelong, Les fonctions plurisousharmoniques, *Annales scientifiques de l'École Normale Supérieure, Sér. 3*, **62** (1945) 301–338.

Equivalent definitions Suppose that u is an upper semicontinuous function on a domain G in $\mathbb{C}^n$. Each of the following properties is equivalent to u being a plurisubharmonic function on G.

1. For every point z in G and every nonzero vector w in $\mathbb{C}^n$, the function $\lambda \mapsto u(z+\lambda w)$ is a subharmonic function of λ in the open subset of $\mathbb{C}$ where this function is defined. (This statement formalizes the meaning of u being subharmonic on every complex line.)

2. For every holomorphic mapping f from the unit disc into G, the composite function $u \circ f$ is subharmonic on the unit disc. (In other words, the restriction of u to every "analytic disc" is subharmonic.)

3. In the special case that u is twice continuously differentiable,

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 u}{\partial z_j \partial \overline{z}_k} w_j \overline{w}_k \geq 0 \qquad \text{for every vector } w \text{ in } \mathbb{C}^n.$$

   The standard notation $\partial/\partial z_j$ means $\frac{1}{2}(\partial/\partial x_j - i\partial/\partial y_j)$ in terms of the underlying real coordinates for which $z_j = x_j + iy_j$. Similarly, $\partial/\partial \overline{z}_j$ means $\frac{1}{2}(\partial/\partial x_j + i\partial/\partial y_j)$. The Austrian mathematician Wilhelm Wirtinger (1865–1945) introduced this notation for complex partial derivatives.[26]

   If u is not twice differentiable, then one can interpret the preceding inequality in the sense of distributions. Alternatively, one can say that u is the limit of a decreasing sequence of infinitely differentiable functions satisfying the inequality.

4. For every closed polydisc of arbitrary orientation contained in G, the value of u at the center of the polydisc is at most the average of u on the torus in the boundary of the polydisc. It is equivalent to say that each point of G has a neighborhood such that the indicated property holds for polydiscs contained in the neighborhood.

The last property needs some explanation. A polydisc is a product of one-dimensional discs. One part of the boundary of the polydisc is the Cartesian product of the boundaries of the one-dimensional discs. This Cartesian product of circles is a multidimensional torus. There is no standard designation for this torus, which different authors call by various names, including "spine," "skeleton," and "distinguished boundary." Lelong uses the French word "arête," which means "edge" in mathematical contexts and more generally can refer to a mountain ridge, the bridge of a nose, and a fishbone. The words "arbitrary orientation" mean that the polydisc need not have its sides parallel to the coordinate axes: the polydisc could be rotated by a unitary transformation.

[26] W. Wirtinger, Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen, *Mathematische Annalen* **97** (1927) 357–376.

It is useful to look at some examples of plurisubharmonic functions before proving the equivalence of the various properties. If f is a holomorphic function, then both $|f|$ and $\log|f|$ are plurisubharmonic, because the restriction of a holomorphic function to a complex line is holomorphic as a function of one variable. Notice that $\log|z_1|$ is identically equal to $-\infty$ on the complex line where $z_1 = 0$, so the indicated definition of plurisubharmonic function works only if one allows subharmonic functions in the plane to be identically equal to $-\infty$.

Less obvious examples are the functions $\log(|z_1|^2 + |z_2|^2)$ and $\log(1 + |z_1|^2 + |z_2|^2)$. The plurisubharmonicity can be verified by computing second derivatives and checking that the complex Hessian matrix is nonnegative, but an alternative approach handles the higher-dimensional analogue with no extra work. Observe that $|z_1|^2 + |z_2|^2 = \sup\{\, |z_1\overline{w}_1 + z_2\overline{w}_2| : |w_1|^2 + |w_2|^2 = 1\,\}$. For fixed values of w_1 and w_2, the function $z_1\overline{w}_1 + z_2\overline{w}_2$ is a holomorphic function of z_1 and z_2, so $\log|z_1\overline{w}_1 + z_2\overline{w}_2|$ is plurisubharmonic. The pointwise supremum of a family of plurisubharmonic functions, if upper semicontinuous, is plurisubharmonic (just as in the one-dimensional case), so $\log(|z_1|^2 + |z_2|^2)$ is plurisubharmonic. The same argument shows that $\log(|z_1|^2 + |z_2|^2 + |z_3|^2)$ is a plurisubharmonic function in $\mathbb{C}^3$, and fixing z_3 equal to 1 shows that $\log(1 + |z_1|^2 + |z_2|^2)$ is a plurisubharmonic function in $\mathbb{C}^2$.

Exercise 29. Show that $\log(|z_1| + |z_2|)$ and $\log(1 + |z_1| + |z_2|)$ are plurisubharmonic functions in $\mathbb{C}^2$.

The function $\max(|z_1|, |z_2|)$ is plurisubharmonic in $\mathbb{C}^2$ because the pointwise maximum of two (pluri)subharmonic functions is again (pluri)subharmonic. Accordingly, the bidisc in $\mathbb{C}^2$, whose definition as a polynomial polyhedron requires two functions, can be defined as a sub-level set of one plurisubharmonic function. Moreover, every polynomial polyhedron can be defined by a single plurisubharmonic function. This example shows that compared to holomorphic functions, plurisubharmonic functions have an advantageous flexibility.

Exercise 30. There is no identity principle for plurisubharmonic functions: Give an example of an everywhere defined plurisubharmonic function, not identically equal to zero, that is nonetheless identically equal to zero on a nonvoid open set.

The example $(\log|z_1|)(\log \frac{1}{|z_2|})$ on the open set where $z_1 z_2 \neq 0$ shows that a function can be subharmonic in each variable separately without being plurisubharmonic. On this open set, the function is even harmonic in each variable separately, but the determinant of the complex Hessian is negative (as a routine calculation shows), so the function is not plurisubharmonic. A routine calculation shows too that the restriction of the function to a complex line on which $z_1 = cz_2$ (where c is an arbitrary nonzero complex number) has negative Laplacian, which is another way to see that the function is not plurisubharmonic. A function that is subharmonic in each variable separately does have the sub-mean-value property on polydiscs with faces parallel to the coordinate axes, which explains why property (4) needs to allow polydiscs of arbitrary orientation.

Exercise 31. Show that the product function $(\log|z_1|)(\log|z_2|)$ is not plurisubharmonic in any neighborhood of a point where $z_1 z_2 \neq 0$.

It is tempting to try to extend the list of equivalent properties for plurisubharmonicity in parallel with the analogous properties for subharmonicity. One might define a function to be "subpluriharmonic" [a nonstandard term] if whenever the function is bounded above on the boundary of a compact set by a pluriharmonic function, the bound propagates to the whole set. (The standard term "pluriharmonic" means that the restriction of the function to each complex line is harmonic. Equivalently, a function is pluriharmonic if locally the function equals the real part of a holomorphic function.) A plurisubharmonic function is subpluriharmonic, but the converse is false. Indeed, every plurisubharmonic function on $\mathbb{C}^n$ is subharmonic as a function on $\mathbb{R}^{2n}$, every pluriharmonic function is harmonic, and every subharmonic function is subpluriharmonic. But the function $|z_1|^2 - |z_2|^2$ is harmonic as a function on $\mathbb{R}^4$, hence subpluriharmonic, but not plurisubharmonic.

Proof of the equivalence of definitions of plurisubharmonicity. Suppose that the function u is twice continuously differentiable. Then saying that $u(z + \lambda w)$ is subharmonic as a function of λ in $\mathbb{C}$ is the same as saying that the Laplacian is nonnegative. The chain rule implies that this Laplacian equals $4\sum_{j=1}^{n}\sum_{k=1}^{n} u_{j\bar{k}} w_j \overline{w}_k$, where $u_{j\bar{k}}$ means $\partial^2 u/\partial z_j \partial \overline{z}_k$. Therefore property (1) implies property (3).

To see that property (3) implies (2), observe that the composite function $u \circ f$ of one variable is subharmonic precisely when the Laplacian is nonnegative. The chain rule implies that when f is a holomorphic mapping, the Laplacian of $u \circ f$ equals $4\sum_{j=1}^{n}\sum_{k=1}^{n} u_{j\bar{k}} f_j \overline{f_k}$.

Evidently property (2) implies property (1). Accordingly, properties (1), (2), and (3) are all equivalent when u is sufficiently smooth.

When u is not smooth, property (2) still implies property (1), which is the special case of a holomorphic mapping that is a first-degree polynomial. To prove the converse, take a decreasing sequence of smooth plurisubharmonic functions converging to u (by convolving u with smooth mollifying functions, just as in one variable). For the smooth approximants, property (2) holds by the first part of the proof, and this property evidently continues to hold in the limit.

It remains to show that property (1) is equivalent to property (4). Property (1) is invariant under composition with a complex-linear transformation (since such transformations take complex lines to complex lines), so it suffices to show that (1) $\implies$ (4) for polydiscs with faces parallel to the coordinate axes. Integrate on the torus by integrating over each circle separately. Applying (1) for each integral shows that (4) holds. Conversely, suppose (4) holds. Since upper semicontinuous functions are bounded above on compact sets, subtracting a constant from u reduces to the case that u is negative. Integrate on a polydisc and let $n-1$ of the radii tend to 0. Apply Fatou's lemma to deduce that the restriction of u to a disc in a complex line satisfies the sub-mean-value property. Thus (1) holds. □

Pseudoconvexity can be characterized by convexity with respect to the plurisubharmonic functions but alternatively by a certain geometric property. Ordinary convexity says that if the boundary of a line segment lies in the domain, then the whole line segment lies in the domain. A natural way to generalize this notion to multidimensional complex analysis is to ask for an analytic disc to lie in the domain whenever the boundary of the disc lies in the domain. An *analytic disc* means a continuous mapping from the closed unit disc D in $\mathbb{C}$ into $\mathbb{C}^n$ that is holomorphic on the interior of the disc. Often an analytic disc is identified with the image (at least when the mapping is one-to-one). Here are two versions of the so-called continuity principle for analytic discs, also known by the German name, *Kontinuitätssatz*. The principle may or may not hold for a particular domain G in $\mathbb{C}^n$.

(a) If for each α in some index set A, the mapping $f_\alpha : D \to G$ is an analytic disc whose image is contained in the domain G, and if there is a compact subset of G that contains $\bigcup_{\alpha\in A} f_\alpha(bD)$ (the "boundaries" of the analytic discs), then there is a compact subset of G that contains $\bigcup_{\alpha\in A} f_\alpha(D)$.

(b) If $f_t : D \to \mathbb{C}^n$ is a family of analytic discs varying continuously with respect to the parameter t in the interval $[0, 1]$, if $\bigcup_{0\le t\le 1} f_t(bD)$ is contained in the domain G (hence automatically contained in a compact subset of G), and if $f_0(D)$ is contained in G, then $\bigcup_{0\le t\le 1} f_t(D)$ is contained in G (hence in a compact subset of G).

The following theorem gives four equivalent ways to characterize pseudoconvex domains. A sufficiently smooth function is called *strictly* (or *strongly*) plurisubharmonic if the complex Hessian matrix is positive definite (rather than positive semi-definite). A function $u : G \to [-\infty, \infty)$ is an *exhaustion function* for G if for every real number a, the set $u^{-1}[-\infty, a)$ is contained in a compact subset of G. The intuitive meaning of an exhaustion function is that the function blows up at the boundary of G (and also at infinity, if G is unbounded).

Theorem 16. *The following properties of a domain G in $\mathbb{C}^n$ are equivalent.*

1. *There exists an infinitely differentiable, strictly plurisubharmonic exhaustion function for G.*

2. *The domain G is convex with respect to the plurisubharmonic functions (that is, G is a pseudoconvex domain).*

3. *The continuity principle (Kontinuitätssatz) holds for G.*

4. *The function $-\log d(z)$ is plurisubharmonic, where $d(z)$ denotes the distance from z to the boundary of G.*

Exercise 32. The unit ball $\{ z \in \mathbb{C}^n : \|z\| < 1 \}$ (where $\|z\|^2 = |z_1|^2 + \cdots + |z_n|^2$) is convex, hence convex with respect to the holomorphic functions, hence convex with respect to the plurisubharmonic functions. The distance from z to the boundary equals $1 - \|z\|$. Verify that $-\log(1 - \|z\|)$ is plurisubharmonic and that $-\log(1 - \|z\|^2)$ is an infinitely differentiable, plurisubharmonic exhaustion function.

Proof of Theorem 16. The plan of the proof is (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) $\Longrightarrow$ (1).

Suppose (1) holds: let u be a plurisubharmonic exhaustion function for G. If K is a compact subset of G, then u is bounded above on K by some constant M. The plurisubharmonic hull of K is contained in $\{ z \in G : u(z) \leq M \}$ by the definition of the hull. This set is contained in a compact subset of G by the definition of exhaustion function. Being a relatively closed subset of G that is contained in a compact subset of G, the plurisubharmonic hull of K is compact. Thus (1) $\Longrightarrow$ (2).

Suppose (2) holds. If $f : D \to G$ is an analytic disc, and u is a plurisubharmonic function on G, then $u \circ f$ is a subharmonic function on the unit disc, and the maximum principle for subharmonic functions implies that $u(f(\lambda)) \leq \max\{ u(f(e^{i\theta})) : 0 \leq \theta \leq 2\pi \}$ for every point λ in D. In other words, $f(D)$ is contained in the plurisubharmonic hull of $f(bD)$. Hence version (a) of the continuity principle holds: the plurisubharmonic hull of a compact set containing $\bigcup_{\alpha \in A} f_\alpha(bD)$ is a compact set containing $\bigcup_{\alpha \in A} f_\alpha(D)$. To get version (b) of the continuity principle, consider the set S of points t in the interval $[0, 1]$ for which $f_t(D) \subset G$. This set is nonvoid, since S contains 0 by hypothesis. If $t \in S$, then $f_t(D)$ is a compact subset of G (since f_t is continuous on the closed disc D), so $f_s(D) \subset G$ for s near t (since the discs vary continuously with respect to the parameter). Thus S is an open set. Version (a) of the continuity principle implies that the set S is closed. Hence S is all of $[0, 1]$, which is what needed to be shown. Consequently, (2) implies (3).

Suppose that (3) holds. To see that $-\log d(z)$ is plurisubharmonic, fix a point z_0 in G and a vector w_0 in $\mathbb{C}^n$ such that the closed disc $\{ z_0 + \lambda w_0 : |\lambda| \leq 1 \}$ lies in G. To show that $-\log d(z_0 + \lambda w_0)$ is subharmonic as a function of λ, it suffices to fix a polynomial p of one complex variable such that

$$-\log d(z_0 + \lambda w_0) \leq \operatorname{Re} p(\lambda) \qquad \text{when } |\lambda| = 1$$

and to show that the same inequality holds when $|\lambda| < 1$. This problem translates directly into the equivalent problem of showing that if

$$d(z_0 + \lambda w_0) \geq |e^{-p(\lambda)}| \qquad \text{when } |\lambda| = 1,$$

then the same inequality holds when $|\lambda| < 1$. A further reformulation is to show that if, for every point ζ in the open unit ball of $\mathbb{C}^n$, the point $z_0 + \lambda w_0 + \zeta e^{-p(\lambda)}$ lies in G when $|\lambda| = 1$, then the same property holds when $|\lambda| < 1$.

Now the map taking λ to $z_0 + \lambda w_0 + \zeta e^{-p(\lambda)}$ is an analytic disc in $\mathbb{C}^n$ depending continuously on the parameter ζ. When $\zeta = 0$, the analytic disc lies in G by hypothesis. Also by hypothesis, the boundaries of all these analytic discs lie in G. Hence version (b) of the continuity principle (applied along the line segment joining 0 to ζ) implies that all the analytic discs lie in G. Thus (3) implies (4).

Finally, suppose (4) holds, that is, $-\log d(z)$ is plurisubharmonic. This function evidently blows up at the boundary of G. The method for obtaining property (1) is to modify this function to make the function both smooth and *strictly* plurisubharmonic. Here are the technical details.

To start, let $u(z)$ denote $\max(\|z\|^2, -\log d(z))$. Evidently u is a continuous, plurisubharmonic exhaustion function for G. Add a suitable constant to u to ensure that the minimum value of u on G is equal to 0. For each positive integer j, let G_j denote the subset of G on which $u < j$. These sets form an increasing sequence of relatively compact open subsets of G.

Extend u to be equal to 0 outside G. For each j, convolve u with a smooth, radially symmetric mollifying function having small support to obtain an infinitely differentiable function on $\mathbb{C}^n$ that is plurisubharmonic on a neighborhood of the closure of G_j and that closely approximates u from above on that neighborhood. Adding $\varepsilon_j\|z\|^2$ for a suitably small positive constant ε_j gives a smooth function u_j on $\mathbb{C}^n$, strictly plurisubharmonic on a neighborhood of the closure of G_j, such that $u < u_j < u + 1$ on that neighborhood. It remains to splice the functions u_j together to get the required smooth, strictly plurisubharmonic exhaustion function for G.

A natural way to build the final function is to use an infinite series. A simple way to guarantee that the sum remains infinitely differentiable is to make the series locally finite. To carry out this plan, let χ be an infinitely differentiable, convex function of one real variable such that $\chi(t) = 0$ when $t \leq 0$, and both χ' and χ'' are positive when $t > 0$.

Exercise 33. Verify that an example of such a function χ is

$$\begin{cases} 0, & \text{if } t \leq 0, \\ e^t e^{-1/t}, & \text{if } t > 0. \end{cases}$$

Exercise 34. Show that if φ is an increasing convex function of one real variable, and v is a plurisubharmonic function, then the composite function $\varphi \circ v$ is plurisubharmonic. Moreover, if φ is strictly convex and v is strictly plurisubharmonic, then $\varphi \circ v$ is strictly plurisubharmonic.

The remainder of the proof consists of inductively choosing c_j to be a suitable positive constant to make the series $\sum_{j=1}^{\infty} c_j \chi(u_j(z) - j + 1)$ have the required properties. The induction statement is that on the set G_k, the sum $\sum_{j=1}^{k} c_j \chi(u_j(z) - j + 1)$ is strictly plurisubharmonic and larger than $u(z)$.

For the basis step ($k = 1$), observe that u_1 is strictly larger than u on a neighborhood of the closure of G_1 and hence is strictly positive there. By Exercise 34, the composite function $\chi \circ u_1$ is strictly plurisubharmonic on the neighborhood. Take the constant c_1 large enough that $c_1 \chi \circ u_1$ exceeds u on G_1.

Suppose now that the induction statement holds for an integer k. There is a neighborhood of the closure of G_{k+1} such that if z is in that neighborhood but outside G_k, then $k \leq u(z) < u_{k+1}(z)$. For such z, the function $\chi(u_{k+1}(z) - k)$ is positive and strictly plurisubharmonic. Multiply by a sufficiently large constant c_{k+1} to guarantee that $\sum_{j=1}^{k+1} c_j \chi(u_j(z) - j + 1)$ is both strictly plurisubharmonic and larger than $u(z)$ when z is in G_{k+1} but outside G_k. Since the function $\chi(u_{k+1}(z) - k)$ is nonnegative and (weakly) plurisubharmonic on all of G_{k+1}, the induction hypothesis implies that the sum of all $k+1$ terms is strictly plurisubharmonic and larger than u on all of G_{k+1}.

It remains to check that the infinite series does converge to an infinitely differentiable function on G. This property is local, so it is enough to check the property on a ball whose closure is contained in G and hence in some G_m. If $j \geq m+2$, then $\chi(u_j(z) - j + 1) = 0$ when $z \in G_m$ (since $u_j < u + 1$ on G_j), so only finitely many terms contribute to the sum on the ball. Hence the series converges to an infinitely differentiable function. The preceding paragraph shows that the limit function is strictly plurisubharmonic. Since the sum exceeds the exhaustion function u, the sum is an exhaustion function too. □

3.3 The Levi problem

The characterizations of pseudoconvexity considered so far are essentially internal to the domain. Eugenio Elia Levi (1883–1917) discovered[27] a way to characterize pseudoconvexity through the differential geometry of the boundary of the domain. This condition requires the boundary to be a twice continuously differentiable manifold. Since Levi's condition is local, one ought first to observe that pseudoconvexity is indeed a local property of the boundary.

Theorem 17. *A domain G in $\mathbb{C}^n$ is pseudoconvex if and only if each boundary point of G has an open neighborhood U in $\mathbb{C}^n$ such that (each component of) the intersection $U \cap G$ is pseudoconvex.*

Proof. If G is pseudoconvex, and B is a ball centered at a boundary point, then the maximum of $-\log d_B(z)$ and $-\log d_G(z)$ is a plurisubharmonic exhaustion function for $B \cap G$, so (each component of) the intersection $B \cap G$ is pseudoconvex.

Conversely, suppose each boundary point p of G has a neighborhood U such that $-\log d_{U\cap G}$ is a plurisubharmonic function on $U \cap G$. The neighborhood U contains a ball centered at p, and if z lies in the concentric ball B of half the radius, then the distance from z to the boundary of G is less than the distance from z to the boundary of U. Therefore the function $-\log d_G$ is plurisubharmonic on $B \cap G$, being equal on this set to the function $-\log d_{U\cap G}$. The union of such balls for all boundary points of G is an open neighborhood V of the boundary of G such that $-\log d_G$ is plurisubharmonic on $V \cap G$. What remains to accomplish is to modify this function to get a plurisubharmonic exhaustion function defined on all of G. If G is bounded, then $G \setminus V$ is a compact set, and $-\log d_G$ has an upper bound M on $G \setminus V$. The continuous function $-\log d_G$ is less than $M+1$ on an open neighborhood of $G \setminus V$, and $\max\{M+1, -\log d_G\}$ is a plurisubharmonic exhaustion function for G.

If G is unbounded, then the set $G \setminus V$ is closed but not necessarily compact. For each nonnegative real number r, the continuous function $-\log d_G$ has a maximum value on the intersection of $G \setminus V$ with the closed ball of radius r centered at 0. By Exercise 35 below, there

[27] E. E. Levi, Studii sui punti singolari essenziali delle funzioni analitiche di due o più variabili complesse, *Annali di Matematica Pura ed Applicata (3)* **17** (1910) 61–87.

is a continuous function $\varphi(\|z\|)$ that is plurisubharmonic on $\mathbb{C}^n$, exceeds $-\log d_G(z)$ when z is in $G \setminus V$, and blows up at infinity. Then $\max\{\varphi(\|z\|), -\log d_G(z)\}$ is a plurisubharmonic exhaustion function for G, so G is pseudoconvex. □

Although pseudoconvexity is a local property of the boundary of a domain, none of the properties so far shown to be equivalent to holomorphic convexity appears to be local. Given a locally defined holomorphic function that is singular at a boundary point of a domain, there is no obvious way to create a globally defined holomorphic function that is singular at the point. The essence of the Levi problem—the equivalence between pseudoconvexity and holomorphic convexity—is to show that being a domain of holomorphy actually is a local property of the boundary of the domain.

Exercise 35. If g is a continuous function on $[0, \infty)$, then there exists an increasing convex function φ such that $\varphi(t) > g(t)$ for every t.

3.3.1 The Levi form

Suppose that in a neighborhood of a boundary point of a domain there exists a real-valued defining function ρ. In other words, the boundary of the domain is the set where $\rho = 0$, the interior of the domain is the set where $\rho < 0$, and the exterior of the domain is the set where $\rho > 0$. Suppose additionally that ρ has continuous partial derivatives of second order and that the gradient of ρ is nowhere equal to 0 on the boundary of the domain. The implicit function theorem then implies that the boundary of the domain (in the specified neighborhood) is a twice differentiable real manifold. The abbreviation for this set of conditions is that the domain has "class C^2 boundary" or "class C^2 smooth boundary."

Levi's local condition is that at boundary points of the domain,

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 \rho}{\partial z_j \partial \overline{z}_k} w_j \overline{w}_k \geq 0 \qquad \text{whenever} \qquad \sum_{j=1}^{n} \frac{\partial \rho}{\partial z_j} w_j = 0.$$

This condition is weaker than plurisubharmonicity of the defining function ρ in two ways: the inequality holds only at boundary points of the domain, not on an open set; and the inequality holds only for certain vectors w in $\mathbb{C}^n$, namely, for *complex tangent vectors* (vectors satisfying the side condition). The indicated Hermitian quadratic form, restricted to act on the complex tangent space, is known as the *Levi form*. If the Levi form is *strictly* positive definite everywhere on the boundary, then the domain is called *strictly pseudoconvex* (or strongly pseudoconvex).

Exercise 36. Even though the Levi form depends on the choice of the defining function ρ, the positivity (or nonnegativity) of the Levi form is independent of the choice of defining function. Moreover, positivity (or nonnegativity) of the Levi form is invariant under local biholomorphic changes of coordinates.

Levi's condition can be rephrased as the existence of a positive constant C such that

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k + C\|w\|\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right| \ge 0 \qquad \text{for every vector } w \text{ in } \mathbb{C}^n.$$

The constant C can be taken to be locally independent of the point where the derivatives are being evaluated. An advantage of this reformulation is the elimination of the side condition about the complex tangent space: the inequality now holds for *every* vector w. The second formulation evidently implies the first statement of Levi's condition. To see, conversely, that Levi's condition implies the reformulation, decompose an arbitrary vector w into an orthogonal sum $w' + w''$, where $\sum_{j=1}^{n}\rho_j w'_j = 0$ (here ρ_j is a typographically convenient abbreviation for $\partial\rho/\partial z_j$), and $\sum_{j=1}^{n}\rho_j w''_j = \sum_{j=1}^{n}\rho_j w_j$. By hypothesis, the length of the gradient of ρ is locally bounded away from 0, so the length of the vector w'' is comparable to $\sum_{j=1}^{n}\rho_j w_j$. Levi's condition implies that

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k = \sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w'_j\overline{w}'_k + O(\|w\|\,\|w''\|) \ge -C\|w\|\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right|$$

for some constant C, which is the reformulated version of the Levi condition.

Theorem 18. *A domain with class C^2 smooth boundary is pseudoconvex if and only if the Levi form is positive semi-definite at each boundary point.*

Proof. First suppose that the domain G is pseudoconvex in the sense that the negative of the logarithm of the distance to the boundary of G is plurisubharmonic. A convenient function ρ to use as defining function is the signed distance to the boundary:

$$\rho(z) = \begin{cases} -\operatorname{dist}(z, bG), & z \in G, \\ +\operatorname{dist}(z, bG), & z \notin G. \end{cases}$$

The implicit function theorem implies that this defining function is class C^2 in a neighborhood of the boundary of G. By hypothesis, the complex Hessian of $-\log|\rho|$ is nonnegative in the part of this neighborhood inside G:

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\left(-\frac{1}{\rho}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k} + \frac{1}{\rho^2}\frac{\partial\rho}{\partial z_j}\frac{\partial\rho}{\partial\overline{z}_k}\right)w_j\overline{w}_k \ge 0 \qquad \text{for every } w \text{ in } \mathbb{C}^n.$$

But $-1/\rho$ is positive at points inside the domain, so

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k \ge 0 \qquad \text{when} \qquad \sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j = 0.$$

As observed just before the statement of the theorem, this Levi condition is equivalent to the existence of a positive constant C such that

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k + C\|w\|\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right| \geq 0 \qquad \text{for every vector } w \text{ in } \mathbb{C}^n.$$

The constant C depends on the maximum of the second derivatives of ρ and the maximum of $1/|\nabla\rho|$, and these quantities are bounded near the boundary of G by hypothesis. The continuity of the second derivatives of ρ implies that the inequality persists on the boundary of G, and Levi's condition follows.

Conversely, suppose that Levi's condition holds. What is required is to construct a plurisubharmonic exhaustion function for the domain G. In view of Theorem 17, a local construction suffices.

The implicit function theorem implies that a boundary of class C^2 is locally the graph of a twice continuously differentiable real-valued function. After a complex-linear change of coordinates, a local defining function ρ takes the form $\varphi(\operatorname{Re} z_1, \operatorname{Im} z_1, \dots, \operatorname{Re} z_{n-1}, \operatorname{Im} z_{n-1}, \operatorname{Re} z_n) - \operatorname{Im} z_n$. The hypothesis is the existence of a positive constant C such that at boundary points in a local neighborhood,

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k + 2C\|w\|\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right| \geq 0 \qquad \text{for every vector } w \text{ in } \mathbb{C}^n$$

(where a factor of 2 has been inserted for later convenience). Since ρ depends linearly on $\operatorname{Im} z_n$, all derivatives of ρ are independent of $\operatorname{Im} z_n$. Thus the preceding condition holds not only locally on the boundary of G but also locally off the boundary, say in some ball in $\mathbb{C}^n$.

Let u denote $-\log|\rho|$. The goal is to modify the function u to get a local plurisubharmonic function in G that blows up at the boundary. At points inside G, the same calculation as above shows that

$$\begin{aligned}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 u}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k &= \frac{1}{|\rho|}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2\rho}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k + \frac{1}{\rho^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial\rho}{\partial z_j}\frac{\partial\rho}{\partial\overline{z}_k}w_j\overline{w}_k \\ &\geq -\frac{2C}{|\rho|}\|w\|\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right| + \frac{1}{\rho^2}\left|\sum_{j=1}^{n}\frac{\partial\rho}{\partial z_j}w_j\right|^2\end{aligned}$$

for every vector w in $\mathbb{C}^n$. But $-2ab \geq -a^2 - b^2$ for real numbers a and b, so

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 u}{\partial z_j\partial\overline{z}_k}w_j\overline{w}_k \geq -C^2\|w\|^2.$$

The preceding inequality implies that $u(z) + C^2\|z\|^2$ is a plurisubharmonic function in the intersection of G with a small ball B centered at a boundary point a, and this function blows up at the boundary of G. If the ball B has radius r, then $\max\{-\log(r-\|z-a\|), u(z)+C^2\|z\|^2\}$ is a plurisubharmonic exhaustion function for $B \cap G$. Thus G is locally pseudoconvex near every boundary point, so by Theorem 17, the domain G is pseudoconvex. □

In view of Levi's condition, the notion of pseudoconvexity can be rephrased as follows.

Theorem 19. *A domain is pseudoconvex if and only if the domain can be expressed as the union of an increasing sequence of class C^∞ smooth domains each of which is locally biholomorphically equivalent to a strongly convex domain.*

Proof. A convex domain is pseudoconvex, and pseudoconvexity is a local property that is biholomorphically invariant, so a domain that is locally equivalent to a convex domain is pseudoconvex. Version (b) of the Kontinuitätssatz implies that an increasing union of pseudoconvex domains is pseudoconvex. Thus one direction of the theorem follows by putting together prior results.

Conversely, suppose that G is a pseudoconvex domain. Then G admits an infinitely differentiable, strictly plurisubharmonic exhaustion function u. Fix a base point in G, let c be a large real number, and consider the connected component of the set where $u < c$ that contains the base point. By Sard's theorem,[28] the gradient of u is nonzero on the set where $u = c$ for almost every value of c (all but a set of measure zero in $\mathbb{R}$). Thus G is exhausted by an increasing sequence of C^∞ smooth, strictly pseudoconvex domains.

What remains to show is that each smooth level set where the strictly plurisubharmonic function u equals a value c is locally equivalent to a strongly convex domain via a local biholomorphic mapping. Fix a point a such that $u(a) = c$, and consider the Taylor expansion of $u(z) - c$ in a neighborhood of a: namely,

$$\begin{aligned} 2\operatorname{Re}\Bigg[\sum_{j=1}^{n} \frac{\partial u}{\partial z_j}(a)(z_j - a_j) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 u}{\partial z_j \partial z_k}(a)(z_j - a_j)(z_k - a_k)\Bigg] \\ + \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 u}{\partial z_j \partial \overline{z}_k}(a)(z_j - a_j)(\overline{z}_k - \overline{a}_k) + O(\|z-a\|^3). \end{aligned} \tag{3.1}$$

The expression whose real part appears on the first line of (3.1) is a holomorphic function of z with nonzero gradient at the point a. This function will serve as the first coordinate w_1 of a local biholomorphic change of coordinates $w(z)$ such that $w(a) = 0$. In a neighborhood of the point a, the level surface on which $u(z) - c = 0$ has a defining function $\rho(w)$ in the new coordinates of the form

$$2\operatorname{Re} w_1 + \sum_{j=1}^{n}\sum_{k=1}^{n} L_{jk} w_j \overline{w}_k + O(\|w\|^3),$$

[28] Arthur Sard, The measure of the critical values of differentiable maps, *Bulletin of the American Mathematical Society* **48** (1942) 883–890. Since the function u takes values in $\mathbb{R}^1$, the claim already follows from an earlier result of Anthony P. Morse, The behavior of a function on its critical set, *Annals of Mathematics (2)* **40** (1939), number 1, 62–70.

where the matrix L_{jk} is a positive definite Hermitian matrix corresponding to the positive definite matrix $u_{j\bar{k}}$ in the new coordinates. Thus the quadratic part of the real Taylor expansion of ρ in the real coordinates corresponding to w is positive definite, which means that the level set on which $\rho = 0$ is strongly convex in the real sense. □

Exercise 37. Solve the Levi problem for complete Reinhardt domains in $\mathbb{C}^2$ by showing that Levi's condition in this setting is equivalent to logarithmic convexity.

3.3.2 Applications of the $\bar{\partial}$ problem

Theorem 17 shows that pseudoconvexity is a local property of the boundary of a domain, but the corresponding local nature of holomorphic convexity is far from obvious. To solve the Levi problem for a general pseudoconvex domain, one needs some technical machinery to forge the connection between the local and the global. One approach is sheaf theory, another is integral representations, and a third is the $\bar{\partial}$-equation. The following discussion uses the third method, which seems the most intuitive.

Some notation is needed. If f is a function, then $\bar{\partial} f$ denotes $\sum_{j=1}^{n} (\partial f/\partial \bar{z}_j)\, d\bar{z}_j$, a so-called $(0,1)$-form. A function f is holomorphic precisely when $\bar{\partial} f = 0$. The question of interest here is whether a given $(0,1)$-form β, say $\sum_{j=1}^{n} b_j(z)\, d\bar{z}_j$, can be written as $\bar{\partial} f$ for some function f. Necessary conditions for the mixed second-order partial derivatives of f to match are that $\partial b_j/\partial \bar{z}_k = \partial b_k/\partial \bar{z}_j$ for every j and k. These conditions are abbreviated by writing that $\bar{\partial}\beta = 0$; in words, the form β is $\bar{\partial}$-closed.

The key ingredient for solving the Levi problem is the following theorem about solvability of the inhomogeneous Cauchy–Riemann equations.

Theorem 20. *Let G be a bounded pseudoconvex domain in $\mathbb{C}^n$ with C^∞ smooth boundary. If β is a $\bar{\partial}$-closed $(0,1)$-form with C^∞ coefficients in G, then there exists a C^∞ function f in G such that $\bar{\partial} f = \beta$.*

The conclusion holds without any hypothesis about boundary smoothness, but then the proof is more technical. For present purposes, proving the theorem under the additional hypothesis of strong pseudoconvexity suffices.

Solution of the Levi problem for bounded strongly pseudoconvex domains

Granted Theorem 20, one can easily solve the Levi problem for the approximating strongly pseudoconvex domains arising in the proof of Theorem 19. Indeed, let G be a bounded domain with boundary defined by an infinitely differentiable, strictly plurisubharmonic function. (One need not assume here that the gradient of the defining function is nonzero on the boundary of G, for Theorem 20 will be applied not on G but on a smooth domain approximating G from outside.) Showing that G is a (weak) domain of holomorphy requires producing a global holomorphic function on G that is singular at a specified boundary point p. The proof of Theorem 19 provides a holomorphic function f_p defined in a neighborhood of p, equal to 0 at p, and zero-free on the part of $\overline{G} \setminus \{p\}$ in the neighborhood. Indeed, the polynomial whose real part appears in the first line of formula (3.1) will serve for f_p, since the expression on the second line of (3.1) is strictly positive on a small punctured neighborhood of p. The goal is to modify $1/f_p$, the locally defined reciprocal function, to obtain a globally defined function on G that is singular at p.

Let χ be a smooth, real-valued, nonnegative cut-off function that is identically equal to 1 in a neighborhood of p and identically equal to 0 outside a larger neighborhood (contained in the set where f_p is defined). The function χ/f_p is defined globally on G and blows up at p, but χ/f_p is not globally holomorphic. The tool for adjusting this function to get a holomorphic function is the theorem on solvability of the $\bar{\partial}$-equation.

The $(0,1)$ form $(\bar{\partial}\chi)/f_p$ is identically equal to 0 in a neighborhood of p, and the zero set of f_p touches $\overline{G}$ only at p inside the support of χ. Therefore the form $(\bar{\partial}\chi)/f_p$ has C^∞ coefficients in a neighborhood D of $\overline{G}$, which may be taken to be a strictly pseudoconvex domain with C^∞ smooth boundary. The form $(\bar{\partial}\chi)/f_p$ is $\bar{\partial}$-closed on this domain D, since $\bar{\partial}\chi$ is $\bar{\partial}$-closed, and $1/f_p$ is holomorphic away from the zeros of f_p. Theorem 20 produces an infinitely differentiable function v on D such that $\bar{\partial}v = (\bar{\partial}\chi)/f_p$.

The function $v - (\chi/f_p)$ is a holomorphic function on the set where $f_p \neq 0$, hence on $\overline{G} \setminus \{p\}$. The function v, being smooth on $\overline{G}$, is bounded there, so the holomorphic function $v - (\chi/f_p)$ is singular at p. Since there exists a holomorphic function on all of G that is singular at a prescribed boundary point, the domain G is a weak domain of holomorphy, hence (by Theorem 13) a domain of holomorphy.

The preceding argument solves the Levi problem for bounded strictly pseudoconvex domains, modulo the proof of solvability of the $\bar{\partial}$-equation.

It is worthwhile noticing that this argument implies the existence of a *peak function* at an arbitrary boundary point p of a strictly pseudoconvex domain G: namely, a holomorphic function h on a neighborhood of $\overline{G}$ that takes the value 1 at p and has absolute value strictly less than 1 everywhere on $\overline{G} \setminus \{p\}$. Indeed, to construct a peak function, first observe that the function f_p obtained from formula (3.1) has negative real part on the intersection of $\overline{G} \setminus \{p\}$ with a small neighborhood of p. Let c be a real constant larger than the maximum of $|v|$ on $\overline{G}$, and let g denote the function $1/[c + v - (\chi/f_p)]$. Then g is well defined and holomorphic on $\overline{G} \setminus \{p\}$, because the denominator has positive real part (and so is nonzero). On the other hand, in the neighborhood of p where the cut-off function χ is identically equal to 1, the function g equals $f_p/[(c+v)f_p - 1]$, so g is holomorphic in a small neighborhood of p and equals 0 at p. Thus g is holomorphic in a neighborhood of $\overline{G}$, has positive real part on $\overline{G} \setminus \{p\}$, and equals 0 at p. Therefore e^{-g} serves as the required holomorphic peak function h.

Proof of the Oka–Weil theorem

Another application of the solvability of the $\bar{\partial}$-equation on strongly pseudoconvex domains is the Oka–Weil theorem (Theorem 12). Indeed, the tools are at hand to prove the following generalization.

Theorem 21. *If G is a domain of holomorphy in $\mathbb{C}^n$, and K is a compact subset of G that is convex with respect to the holomorphic functions on G, then every function holomorphic in a neighborhood of K can be approximated uniformly on K by functions holomorphic on G.*

Theorem 12 follows by taking G equal to $\mathbb{C}^n$, because convexity with respect to entire functions is the same as polynomial convexity, and approximation by entire functions is equivalent to approximation by polynomials.

Proof of Theorem 21. Suppose f is holomorphic in an open neighborhood U of K, and ε is a specified positive number. The goal is to approximate f on K within ε by functions that are holomorphic on the domain G. There is no loss of generality in supposing that the closure of the neighborhood U is a compact subset of G.

Let L be a compact subset of G containing U and convex with respect to $\mathcal{O}(G)$. The initial goal is to show that f can be approximated on K within ε by functions that are holomorphic in a neighborhood of L. Then a limiting argument as L expands will finish the proof.

Fix an open neighborhood V of L having compact closure in G. The first observation is that there are finitely many holomorphic functions $f_1, \ldots, f_k$ on G such that

$$K \subseteq \{\, z \in V : |f_1(z)| \leq 1, \ldots, |f_k(z)| \leq 1 \,\} \subset U.$$

In other words, the compact set K can be closely approximated from outside by a compact analytic polyhedron defined by functions that are holomorphic on G. The reason is similar to the proof of Theorem 9: the set $\overline{V} \setminus U$ is compact, and each point of this set can be separated from the holomorphically convex set K by a function holomorphic on G, so a compactness argument furnishes a finite number of separating functions. Hence there is no loss of generality in assuming from the start that K is equal to the indicated analytic polyhedron.

For the same reason, the holomorphically convex compact set L can be approximated from outside by a compact analytic polyhedron contained in V and defined by a finite number of functions holomorphic on G. Again, one might as well assume that L equals that analytic polyhedron.

The main step in the proof is to show that functions holomorphic in a neighborhood of L are dense in the functions holomorphic in a neighborhood of $\{\, z \in L : |f_1(z)| \leq 1 \,\}$. An evident induction on the number of functions defining the polyhedron K then implies that $\mathcal{O}(L)$ is dense in $\mathcal{O}(K)$.

At this point, Oka's great insight enters: Oka had the idea that raising the dimension by looking at the graph of f_1 can simplify matters. Let L_1 denote $\{\, z \in L : |f_1(z)| \leq 1 \,\}$, and let D denote the closed unit disc in $\mathbb{C}$. The claim is that if g is a holomorphic function in a neighborhood of L_1, then there is a corresponding function $F(z, w)$ in $\mathbb{C}^n \times \mathbb{C}$, holomorphic in a neighborhood of $L \times D$, such that $g(z) = F(z, f_1(z))$ when z is in a neighborhood of L_1. In other words, there is a holomorphic function on all of $L \times D$ whose restriction to the graph of f_1 recovers g on L_1.

How does this construction help? The point is that F can be expanded in a Maclaurin series in the last variable, $F(z, w) = \sum_{j=0}^{\infty} a_j(z)w^j$, in which the coefficient functions a_j are holomorphic on L. Then $g(z) = \sum_{j=0}^{\infty} a_j(z) f_1(z)^j$ in a neighborhood of L_1, and the partial sums of this series are holomorphic functions on L that uniformly approximate g on L_1.

To construct F, take an infinitely differentiable cut-off function χ in $\mathbb{C}^n$ that is identically equal to 1 in a neighborhood of L_1 and that is identically equal to 0 outside a slightly larger neighborhood (contained in the set where g is defined). Consider in $\mathbb{C}^{n+1}$ the $(0,1)$-form

$$\frac{g(z)\,\bar{\partial}\chi(z)}{f_1(z) - w}, \qquad \text{where } z \in \mathbb{C}^n, \text{ and } w \in \mathbb{C}.$$

This $(0,1)$-form is well defined and smooth on a neighborhood of $L \times D$, for if z lies in the support of $\bar{\partial}\chi$, then z lies outside a neighborhood of L_1, whence $|f_1(z)| > 1$, and the denominator of the form is nonzero. Evidently the form is $\bar{\partial}$-closed. The compact analytic polyhedron L can be approximated from outside by open analytic polyhedra, so $L \times D$ can be approximated from outside by domains of holomorphy (more precisely, each connected component can be so approximated). By Theorem 19, the compact set $L \times D$ can be approximated from outside by bounded, smooth, strongly pseudoconvex open sets. Consequently, the solvability of the $\bar{\partial}$-equation guarantees the existence of a smooth function v in a neighborhood of $L \times D$ such that $g(z)\chi(z) - v(z, w)(f_1(z) - w)$ is holomorphic on $L \times D$. The latter function is the required holomorphic function $F(z, w)$ on $L \times D$ such that $F(z, f_1(z)) = g(z)$ on L_1.

The proof is now complete that $\mathcal{O}(L)$ is dense in $\mathcal{O}(K)$. What remains is to approximate a function holomorphic in a neighborhood of K by a function holomorphic on all of G. To this end, let $\{K_j\}_{j=0}^{\infty}$ be an exhaustion of G by an increasing sequence of holomorphically convex, compact subsets of G, each containing an open neighborhood of the preceding one, where the initial set K_0 may be taken equal to K. By what has already been proved, $\mathcal{O}(K_j)$ is dense in $\mathcal{O}(K_{j-1})$ for every positive integer j. Suppose given a function f holomorphic in a neighborhood of K_0 and a positive ε. There is a function h_1 holomorphic in a neighborhood of K_1 such that $|f - h_1| < \varepsilon/2$ on K_0. For $j > 1$, inductively choose h_j holomorphic on K_j such that $|h_j - h_{j-1}| < \varepsilon/2^j$ on K_{j-1}. The telescoping series $h_1 + \sum_{j=2}^{\infty}(h_j - h_{j-1})$ then converges uniformly on every compact subset of G to a holomorphic function that approximates f within ε on K_0. □

Solution of the Levi problem for arbitrary pseudoconvex domains

What has been shown so far is that if G is a pseudoconvex domain, then there exists an infinitely differentiable, strictly plurisubharmonic exhaustion function u, and the Levi problem is solvable for the sublevel sets of u, which are thus domains of holomorphy. A limiting argument is needed to show that G itself is a domain of holomorphy.

For each real number r, let G_r denote the sublevel set $\{ z \in G : u(z) < r \}$, and let $\overline{G}_r$ denote the closure, the set $\{ z \in G : u(z) \le r \}$. The key lemma is that $\mathcal{O}(G_t)$ is dense in $\mathcal{O}(\overline{G}_r)$ when $t > r$.

In view of the preceding approximation result, Theorem 21, what needs to be shown is that the compact set $\overline{G}_r$ is convex with respect to the holomorphic functions on G_t. Since G_t is a domain of holomorphy, the hull of $\overline{G}_r$ with respect to the holomorphic functions on G_t is a compact subset of G_t. Seeking a contradiction, suppose that this hull properly contains $\overline{G}_r$. Then the exhaustion function u attains a maximal value s on the hull, where $r < s < t$, and this maximal value is assumed at some point p on the boundary of G_s. As observed on page 77, there is a holomorphic peak function h for G_s at p such that $h(p) = 1$, and $|h(z)| < 1$ when $z \in G_s$. Since h is holomorphic in a neighborhood of the holomorphically convex hull of $\overline{G}_r$ with respect to G_t, the function h can be approximated on this hull by functions holomorphic on G_t (by Theorem 21). Since h separates p from $\overline{G}_r$, so do holomorphic functions on G_t, and therefore p is not in the holomorphic hull of $\overline{G}_r$ after all. The contradiction shows that $\overline{G}_r$ is $\mathcal{O}(G_t)$-convex, so Theorem 21 implies that $\mathcal{O}(G_t)$ is dense in $\mathcal{O}(\overline{G}_r)$.

The same argument as in the final paragraph of the proof of Theorem 21 (with a telescoping series) now shows that $\mathcal{O}(G)$ is dense in $\mathcal{O}(G_r)$ for every r.

To prove that G is a domain of holomorphy, fix a compact subset K. What needs to be shown is that $\widehat{K}_G$ is a compact subset of G. Fix a real number r so large that K is a compact subset of G_r. Since G_r is a domain of holomorphy, the hull $\widehat{K}_{G_r}$ is a compact subset of G_r. The claim now is that $\widehat{K}_G \subseteq \widehat{K}_{G_r}$ (whence $\widehat{K}_G = \widehat{K}_{G_r}$, since $\widehat{K}_{G_r}$ automatically is a subset of $\widehat{K}_G$). In other words, the claim is that if $p \notin \widehat{K}_{G_r}$, then there is a holomorphic function on G that separates p from K.

If $p \in G_r$, then there is no difficulty, because there is a holomorphic function on G_r that separates p from K, and $\mathcal{O}(G)$ is dense in $\mathcal{O}(G_r)$. If $p \notin G_r$, then choose some s larger than r for which $p \in G_s$. Since $\mathcal{O}(G_s)$ is dense in $\mathcal{O}(G_r)$, the intersection of $\widehat{K}_{G_s}$ with G_r equals $\widehat{K}_{G_r}$. Therefore the function that is identically equal to 0 in a neighborhood of $\widehat{K}_{G_r}$ and identically equal to 1 in a neighborhood of $G \setminus G_r$ is holomorphic on $\widehat{K}_{G_s}$. By Theorem 21, this function can be approximated on $\widehat{K}_{G_s}$ by functions holomorphic on G_s and hence (since $\mathcal{O}(G)$ is dense in $\mathcal{O}(G_s)$) by functions holomorphic on G. Thus the point p can be separated from K by functions holomorphic on G, so p is not in $\widehat{K}_G$.

The argument has shown that the pseudoconvex domain G is holomorphically convex, so G is a domain of holomorphy. The solution of the Levi problem for pseudoconvex domains is now complete, except for proving the solvability of the $\overline{\partial}$-equation on bounded strongly pseudoconvex domains with smooth boundary.

3.3.3 Solution of the $\overline{\partial}$-equation on smooth pseudoconvex domains

The above resolution of the Levi problem requires knowing that the $\overline{\partial}$-equation is solvable on a bounded, strongly pseudoconvex domain G with smooth boundary. The following discussion proves this solvability by using ideas developed in the 1950s and 1960s by Charles B. Morrey, Donald C. Spencer, Joseph J. Kohn, and Lars Hörmander.

The method is based on Hilbert-space techniques. The relevant Hilbert space is $L^2(G)$, the space of square-integrable functions on G with inner product $\langle f, g\rangle$ equal to $\int_G f\,\overline{g}\,dV$, where dV denotes Lebesgue volume measure. The inner product extends to differential forms by summing the inner products of components of the forms.

The operator $\overline{\partial}$ acts on square-integrable functions in the sense of distributions, so one can view $\overline{\partial}$ as an unbounded operator from the space $L^2(G)$ to the space of $(0,1)$-forms with coefficients in $L^2(G)$. A function f lies in the domain of the operator $\overline{\partial}$ when the distributional coefficients of $\overline{\partial} f$ are represented by square-integrable functions. Since the compactly supported, infinitely differentiable functions are dense in $L^2(G)$, the operator $\overline{\partial}$ is a densely defined operator, and routine considerations show that this operator is a closed operator. Consequently, there is a Hilbert-space adjoint $\overline{\partial}^*$, which too is a closed, densely defined operator.

If f is a $(0,1)$-form $\sum_{j=1}^n f_j\, d\overline{z}_j$, then

$$\overline{\partial} f = \sum_{1\le j<k\le n} \left(\frac{\partial f_j}{\partial \overline{z}_k} - \frac{\partial f_k}{\partial \overline{z}_j}\right) d\overline{z}_k \wedge d\overline{z}_j.$$

If you are unfamiliar with the machinery of differential forms, then you can view the preceding expression as simply a formal gadget that is a convenient notation for stating the necessary condition for solvability of the equation $\overline{\partial} u = f$: namely, that $\overline{\partial} f = 0$. The goal is to show that this necessary condition is sufficient on bounded pseudoconvex domains with C^∞ smooth boundary. Moreover, the solution u is infinitely differentiable if the coefficients of f are infinitely differentiable. (The solution u is not unique, because any holomorphic function can be subtracted from u, but if one solution is an infinitely differentiable function in G, then every solution is infinitely differentiable.)

Reduction to an estimate

The claim is that the whole problem boils down to proving the following *basic estimate*: there exists a constant C such that

$$\|f\|^2 \le C(\|\overline{\partial} f\|^2 + \|\overline{\partial}^* f\|^2) \tag{3.2}$$

for every $(0,1)$-form f that belongs to both the domain of $\overline{\partial}$ and the domain of $\overline{\partial}^*$. The constant C turns out to depend on the diameter of the domain, so boundedness of the domain is important. Why does this estimate imply solvability of the $\overline{\partial}$-equation?

Suppose that g is a specified $\bar{\partial}$-closed $(0,1)$-form with coefficients in $L^2(G)$. Consider the mapping that sends $\bar{\partial}^* f$ to $\langle f, g\rangle$ when f is a $(0,1)$-form belonging to both the domain of $\bar{\partial}^*$ and the kernel of $\bar{\partial}$. The basic estimate implies that $\|\bar{\partial}^* f\|$ dominates $\|f\|$, so this mapping is a well-defined bounded linear operator on the subspace $\bar{\partial}^*(\operatorname{dom}\bar{\partial}^* \cap \ker\bar{\partial})$ of $L^2(G)$.

The Riesz representation theorem produces a square-integrable function u such that $\langle\bar{\partial}^* f, u\rangle = \langle f, g\rangle$ for every f in the intersection of the domain of $\bar{\partial}^*$ and the kernel of $\bar{\partial}$. On the other hand, if f is in the intersection of the domain of $\bar{\partial}^*$ and the orthogonal complement of the kernel of $\bar{\partial}$, then the same equality holds trivially because sides vanish (namely, $\langle f, g\rangle = 0$ because g is in the kernel of $\bar{\partial}$; similarly $\langle f, \bar{\partial}\varphi\rangle = 0$ for every infinitely differentiable, compactly supported function φ, so $\langle\bar{\partial}^* f, \varphi\rangle = 0$, and therefore $\bar{\partial}^* f = 0$). Consequently, u is in the domain of the adjoint of $\bar{\partial}^*$, hence in the domain of $\bar{\partial}$, and $\langle f, \bar{\partial}u\rangle = \langle f, g\rangle$ for every f in the domain of $\bar{\partial}^*$. The domain of $\bar{\partial}^*$ is dense, so $\bar{\partial}u = g$.

Thus the basic estimate implies the existence of a solution of the $\bar{\partial}$-equation in $L^2(G)$. Why is the solution u infinitely differentiable in G when g has coefficients that are infinitely differentiable functions in G? For each index j, the function $\partial u/\partial\bar{z}_j$ is a component of g and hence is infinitely differentiable. The question, then, is whether the distributional derivative $\partial^{|\beta|}u/\partial z^\beta$ exists as a continuous function for every multi-index β. In view of Sobolev's lemma (or the Sobolev embedding theorem) from functional analysis, what needs to be shown is that such derivatives of u are locally square-integrable. An equivalent problem is to show that for every infinitely differentiable, real-valued function φ having compact support in G, the integral

$$\int_G \varphi \frac{\partial^{|\beta|}u}{\partial z^\beta}\,\overline{\frac{\partial^{|\beta|}u}{\partial z^\beta}}\,dV$$

is finite. Integrating all the derivatives by parts results in a sum of integrals involving only barred derivatives of u, and these derivatives are already known to be smooth functions (and hence locally square-integrable). Thus all derivatives of u are square-integrable on compact subsets of G, and the solution u is infinitely differentiable in G when g is. (The catchphrase here from the theory of partial differential equations is "interior elliptic regularity.")

The much more difficult question of whether the derivatives of the solution u extend smoothly to the boundary of the domain G when g has this property is beyond the scope of these notes. This question of boundary regularity is the subject of current research, and the situation is not completely understood.

Proof of the basic estimate

The cognoscenti sometimes describe the proof of the basic estimate as "an exercise in integration by parts." This characterization becomes less of an exaggeration if you admit Stokes's theorem as an instance of integration by parts.

The plan is to work on the right-hand side of the basic estimate, assuming that the differential forms have coefficients that are sufficiently smooth functions on the closure of G. There is a technical point that needs attention here: namely, to prove that reasonably smooth forms are dense in the intersection of the domains of $\bar\partial$ and $\bar\partial^*$. That the necessary density does hold is a special case of the so-called Friedrichs lemma, a general construction of Kurt Friedrichs.[29]

Suppose, then, that $f = \sum_{j=1}^n f_j \, d\bar z_j$, and each f_j is a smooth function on the closure of G. Since

$$|\bar\partial f|^2 = \sum_{1\le j<k\le n} \left| \frac{\partial f_j}{\partial \bar z_k} - \frac{\partial f_k}{\partial \bar z_j} \right|^2 = \frac{1}{2}\sum_{j=1}^n \sum_{k=1}^n \left| \frac{\partial f_j}{\partial \bar z_k} - \frac{\partial f_k}{\partial \bar z_j} \right|^2$$
$$= \sum_{j=1}^n \sum_{k=1}^n \left(\left| \frac{\partial f_j}{\partial \bar z_k} \right|^2 - \frac{\partial f_j}{\partial \bar z_k} \overline{\frac{\partial f_k}{\partial \bar z_j}} \right),$$

integrating over G shows that

$$\|\bar\partial f\|^2 = \sum_{j=1}^n \sum_{k=1}^n \int_G \left(\left| \frac{\partial f_j}{\partial \bar z_k} \right|^2 - \frac{\partial f_j}{\partial \bar z_k} \overline{\frac{\partial f_k}{\partial \bar z_j}} \right) dV. \tag{3.3}$$

To analyze $\|\bar\partial^* f\|$ requires a formula for $\bar\partial^* f$. If u is a smooth function on the closure of G, and ρ is a defining function for G normalized such that $|\nabla\rho| = 1$ on the boundary of G, then

$$\langle \bar\partial^* f, u\rangle = \langle f, \bar\partial u\rangle = \int_G \sum_{j=1}^n f_j \overline{\frac{\partial u}{\partial \bar z_j}} \, dV$$
$$= \int_G \sum_{j=1}^n -\frac{\partial f_j}{\partial z_j} \bar u \, dV + \int_{bG} \sum_{j=1}^n f_j \frac{\partial \rho}{\partial z_j} \bar u \, dS,$$

where dS denotes $(2n-1)$-dimensional Lebesgue measure on the boundary of G. Since u is arbitrary, the $(0,1)$-form f is in the domain of $\bar\partial^*$ if and only if

$$\sum_{j=1}^n f_j \frac{\partial \rho}{\partial z_j} = 0 \qquad \text{on the boundary of } G, \tag{3.4}$$

and then $\bar\partial^* f = -\sum_{j=1}^n \partial f_j / \partial z_j$.

Now integrate by parts:

$$\|\bar\partial^* f\|^2 = \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial f_j}{\partial z_j} \overline{\frac{\partial f_k}{\partial z_k}} \, dV = -\sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial^2 f_j}{\partial z_j \partial \bar z_k} \bar f_k \, dV,$$

[29] K. O. Friedrichs, The identity of weak and strong extensions of differential operators, *Transactions of the American Mathematical Society* **55**, number 1, (1944) 132–151.

where the boundary term vanishes because f satisfies the boundary condition (3.4) for membership in the domain of $\bar\partial^*$. Integrating by parts a second time shows that

$$\|\bar\partial^* f\|^2 = \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial f_j}{\partial \bar z_k} \overline{\frac{\partial f_k}{\partial \bar z_j}}\, dV - \sum_{j=1}^n \sum_{k=1}^n \int_{bG} \frac{\partial f_j}{\partial \bar z_k} \bar f_k \frac{\partial \rho}{\partial z_j}\, dS.$$

The boundary condition (3.4) implies that the differential operator $\sum_{k=1}^n (\bar f_k)(\partial/\partial \bar z_k)$ is a tangential differential operator, so applying this operator to (3.4) shows that on the boundary,

$$0 = \sum_{k=1}^n \bar f_k \frac{\partial}{\partial \bar z_k}\Big(\sum_{j=1}^n f_j \frac{\partial \rho}{\partial z_j}\Big) = \sum_{j=1}^n \sum_{k=1}^n \Big(\bar f_k \frac{\partial f_j}{\partial \bar z_k}\frac{\partial \rho}{\partial z_j} + f_j \bar f_k \frac{\partial^2 \rho}{\partial z_j \partial \bar z_k}\Big).$$

Combining this identity with the preceding equation shows that

$$\begin{aligned} \|\bar\partial^* f\|^2 &= \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial f_j}{\partial \bar z_k} \overline{\frac{\partial f_k}{\partial \bar z_j}}\, dV + \sum_{j=1}^n \sum_{k=1}^n \int_{bG} \frac{\partial^2 \rho}{\partial z_j \partial \bar z_k} f_j \bar f_k\, dS \\ &\ge \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial f_j}{\partial \bar z_k} \overline{\frac{\partial f_k}{\partial \bar z_j}}\, dV, \end{aligned} \tag{3.5}$$

where the final inequality uses for the first (and only) time that the domain G is pseudoconvex (which implies nonnegativity of the boundary term).

Combining (3.3) and (3.5) shows that

$$\|\bar\partial f\|^2 + \|\bar\partial^* f\|^2 \ge \sum_{j=1}^n \sum_{k=1}^n \int_G \left|\frac{\partial f_j}{\partial \bar z_k}\right|^2 dV.$$

Actually, the preceding inequality is not the one that is needed, but if you followed the calculation, then you should be able to keep track of some extra terms in the integrations by parts to solve the following exercise.

Exercise 38. If a is an infinitely differentiable positive weight function, then

$$\begin{aligned} \int_G (|\bar\partial f|^2 + |\bar\partial^* f|^2)\, a\, dV &= \sum_{j=1}^n \sum_{k=1}^n \int_G \left|\frac{\partial f_j}{\partial \bar z_k}\right|^2 a\, dV + \int_{bG} \frac{\partial^2 \rho}{\partial z_j \partial \bar z_k} f_j \bar f_k\, a\, dS \\ &\quad - \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial^2 a}{\partial z_j \partial \bar z_k} f_j \bar f_k\, dV + 2\,\mathrm{Re}\Big\langle \sum_{k=1}^n f_k \frac{\partial a}{\partial z_k}, \bar\partial^* f\Big\rangle \\ &\ge - \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial^2 a}{\partial z_j \partial \bar z_k} f_j \bar f_k\, dV + 2\,\mathrm{Re}\Big\langle \sum_{k=1}^n f_k \frac{\partial a}{\partial z_k}, \bar\partial^* f\Big\rangle. \end{aligned}$$

In the preceding exercise, replace the positive weight function a by $1 - e^b$, where b is a smooth negative function. Then

$$\frac{\partial^2 a}{\partial z_j \partial \overline{z}_k} = -e^b \frac{\partial^2 b}{\partial z_j \partial \overline{z}_k} - e^b \frac{\partial b}{\partial z_j} \frac{\partial b}{\partial \overline{z}_k},$$

so it follows that

$$\begin{aligned}&\int_G (|\overline{\partial} f|^2 + |\overline{\partial}^* f|^2)\, a\, dV \\ &\quad \geq \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial^2 b}{\partial z_j \partial \overline{z}_k} f_j \overline{f}_k\, e^b\, dV + \int_G \left| \sum_{k=1}^n \frac{\partial b}{\partial z_k} f_k \right|^2 e^b\, dV - 2\,\mathrm{Re} \left\langle \sum_{k=1}^n f_k \frac{\partial b}{\partial z_k} e^{b/2}, e^{b/2}\, \overline{\partial}^* f \right\rangle.\end{aligned}$$

Applying the Cauchy-Schwarz inequality to the last term on the right-hand side and using that $a + e^b = 1$ shows that

$$\|\overline{\partial}^* f\|^2 + \|\overline{\partial} f\|^2 \geq \int_G |\overline{\partial}^* f|^2 + a|\overline{\partial} f|^2\, dV \geq \sum_{j=1}^n \sum_{k=1}^n \int_G \frac{\partial^2 b}{\partial z_j \partial \overline{z}_k} f_j \overline{f}_k e^b\, dV.$$

Now choose a point p in G, let δ denote the diameter of G, and set the negative function b equal to $-1 + |z - p|^2/\delta^2$. The preceding inequality then implies that

$$\|\overline{\partial} f\|^2 + \|\overline{\partial}^* f|^2 \geq \|f\|^2/(\delta^2 e).$$

Thus the basic estimate (3.2) holds with the constant C equal to e times the square of the diameter of the domain G.

4 Holomorphic mappings

Conformal mapping is a key topic in the theory of holomorphic functions of one complex variable. As already observed, holomorphic mappings of more variables—indeed, even linear mappings—rarely preserve angles. This chapter explores some new phenomena that arise in the study of multidimensional holomorphic mappings.

4.1 Fatou–Bieberbach domains

In one variable, there are few biholomorphisms (bijective holomorphic mappings) of the whole plane. All such mappings arise by composing rotations, dilations, and translations. In other words, the group of holomorphic automorphisms of the plane consists of those functions that transform z into $az + b$, where b is an arbitrary complex number and a is an arbitrary nonzero complex number.

When $n \geq 2$, there is a huge group of automorphisms of $\mathbb{C}^n$. Indeed, if f is an arbitrary holomorphic function of one complex variable, then the mapping that sends a point (z_1, z_2) of $\mathbb{C}^2$ to the image point $(z_1 + f(z_2), z_2)$ is an automorphism of $\mathbb{C}^2$. (The coordinate functions evidently are holomorphic, and the mapping that sends (z_1, z_2) to $(z_1 - f(z_2), z_2)$ is the inverse transformation.) An automorphism of this special type is called a *shear*.[1]

The vastness of the group of automorphisms of $\mathbb{C}^n$ when $n \geq 2$ can be viewed as an explanation of the following surprising phenomenon. When $n \geq 2$, there exist biholomorphic mappings from the whole of $\mathbb{C}^n$ onto a proper subset of $\mathbb{C}^n$. Such mappings are known as Fatou–Bieberbach mappings.

In 1922, Fatou gave an example[2] of a nondegenerate entire mapping of $\mathbb{C}^2$ whose range is not dense in $\mathbb{C}^2$. In a footnote, Fatou pointed out that Poincaré had already projected the existence of such mappings, but without offering an example.[3] Bieberbach gave the first injective example.[4]

[1] The terminology is due to Jean-Pierre Rosay and Walter Rudin, Holomorphic maps from $\mathbb{C}^n$ to $\mathbb{C}^n$, *Transactions of the American Mathematical Society* **310** (1988), number 1, 47–86.

[2] P. Fatou, Sur certaines fonctions uniformes de deux variables, *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **175** (1922) 1030–1033.

[3] H. Poincaré, Sur une classe nouvelle de transcendants uniformes, *Journal de mathématiques pures et appliquées (4)* **6** (1890) 313–365.

[4] L. Bieberbach, Beispiel zweier ganzer Funktionen zweier komplexer Variablen, welche eine schlichte volumtreue Abbildung des R_4 auf einen Teil seiner selbst vermitteln, *Sitzungsberichte Preussische Akademie der Wissenschaften* (1933) 476–479.

4.1.1 Example

The following construction, based on an example in the cited paper of Rosay and Rudin, provides a concrete example of a proper subset of $\mathbb{C}^2$ that is biholomorphically equivalent to $\mathbb{C}^2$. This Fatou–Bieberbach domain arises as the basin of attraction of an attracting fixed point of an automorphism of $\mathbb{C}^2$.

The starting point is a particular auxiliary polynomial mapping F: namely,

$$F(z_1, z_2) = \frac{1}{2}\left(z_2, z_1 + z_2^2\right).$$

Obtained by composing a shear, a transposition of variables, and a dilation by a factor of $1/2$, the mapping F evidently is an automorphism of $\mathbb{C}^2$. Moreover, the inverse mapping is easy to compute:

$$F^{-1}(z_1, z_2) = 2\left(z_2 - 2z_1^2, z_1\right).$$

Exercise 39. The origin is a fixed point of F. Show that F has exactly one other fixed point.

The complex Jacobian matrix of F equals

$$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 2z_2 \end{pmatrix},$$

so the Jacobian determinant is identically equal to the constant value $-1/4$. At the origin, the linear approximation of F corresponds to a transposition of variables composed with a dilation by a factor of $1/2$, so the origin is an *attracting fixed point* of F.

The rate of attraction can be quantified as follows. When $z = (z_1, z_2)$, let $\|z\|_\infty$ denote $\max\{|z_1|, |z_2|\}$, the maximum norm on $\mathbb{C}^2$. The unit bidisc is the "unit ball" with respect to this norm. The second component of $F(z)$ admits the following estimate:

$$\left|\tfrac{1}{2}(z_1 + z_2^2)\right| \le \|z\|_\infty \cdot \frac{1 + \|z\|_\infty}{2}.$$

Since the first component of $F(z)$ admits an even better estimate, the mapping F is a contraction on the unit bidisc: namely, if $0 < \|z\|_\infty < 1$, then $\|F(z)\|_\infty < \|z\|_\infty$. The contraction is strict on smaller bidiscs. For instance,

$$\|F(z)\|_\infty \le \frac{2}{3}\|z\|_\infty \qquad \text{when} \qquad \|z\|_\infty \le \frac{1}{3}. \tag{4.1}$$

In summary, if $F^{[k]}$ denotes $\underbrace{F\circ\cdots\circ F}_{k \text{ times}}$, the kth iterate of F, then

$$\lim_{k\to\infty} F^{[k]}(z) = 0 \qquad \text{when} \qquad \|z\|_\infty < 1.$$

Thus the open unit bidisc D lies inside the *basin of attraction* of the fixed point 0.

Moreover, every point that is attracted to the origin under iteration of the mapping F has the property that some iterate of F maps the point to an image point inside D. Accordingly, the basin of attraction of the origin is precisely the union

$$\bigcup_{k=0}^{\infty} F^{[-k]}(D),$$

where $F^{[-k]}$ denotes the kth iterate of the inverse transformation F^{-1}, and $F^{[0]}$ means the identity transformation. Since F^{-1} is an open map, the basin of attraction is an open set.

Exercise 40. Show that the basin of attraction is a connected set.

Let $\mathcal{B}$ denote this basin of attraction. The indicated representation of the basin reveals that the automorphism F of $\mathbb{C}^2$ maps $\mathcal{B}$ onto itself. In other words, not only is F an automorphism of $\mathbb{C}^2$, but the restriction of F to $\mathcal{B}$ is an automorphism of $\mathcal{B}$.

The basin $\mathcal{B}$ does not contain the second fixed point of F and so is not all of $\mathbb{C}^2$. Moreover, the complement of $\mathcal{B}$ is sizeable: the complement contains the set $\{\,(z_1, z_2) : |z_2| \geq 3+|z_1|\,\}$, for the following reason. This set is disjoint from the unit bidisc, so if it can be shown that this set is mapped into itself by F, then no point of the set is attracted to the origin. What needs to be shown, then, is that if $|z_2| \geq 3+|z_1|$, then $|\frac{1}{2}z_1 + \frac{1}{2}z_2^2| \geq 3+\frac{1}{2}|z_2|$, or equivalently, $|z_1 + z_2^2| \geq 6 + |z_2|$. The assumption that $|z_2| \geq 3 + |z_1|$ implies, in particular, that $|z_2| \geq 3$ and $2|z_2| \geq 3 + |z_2|$, so

$$|z_1 + z_2^2| \geq |z_2|^2 - |z_1| \geq |z_2|^2 + 3 - |z_2| \geq 3|z_2| + 3 - |z_2| \geq 6 + |z_2|,$$

as claimed. Thus the basin $\mathcal{B}$ is an open subset of $\mathbb{C}^2$ whose complement contains a nontrivial open set: namely, the set $\{\,(z_1, z_2) : |z_2| > 3 + |z_1|\,\}$.

Although the indicated set in the complement of $\mathcal{B}$ is multicircular, the basin $\mathcal{B}$ itself is not a Reinhardt domain. For example, the point $(4, 2i)$ belongs to $\mathcal{B}$ [since $F(4, 2i) = (i, 0)$, and $F^{[2]}(4, 2i) = (0, i/2)$, which is a point in the interior of the unit bidisc, hence belongs to $\mathcal{B}$], but the point $(4, 2)$ does not belong to $\mathcal{B}$ [since $F(4, 2) = (1, 4)$, and this point belongs to the set considered in the preceding paragraph that lies in the complement of $\mathcal{B}$].

Notice too that $\mathcal{B}$ is identical to the basin of attraction of the origin for the mapping $F^{[2]}$. Indeed, the iterates of F eventually map a specified point into the unit bidisc if and only if the iterates of $F^{[2]}$ do. Let Φ denote $F^{[2]}$. A routine computation shows that

$$4\Phi(z) = z + (z_2^2, \tfrac{1}{2}z_1^2 + z_1 z_2^2 + \tfrac{1}{2}z_2^4),$$

where $z = (z_1, z_2)$. In particular,

$$\text{if} \qquad \|z\|_\infty \leq 1 \qquad \text{then} \qquad \|4\Phi(z) - z\|_\infty \leq 2\|z\|_\infty^2. \tag{4.2}$$

The claim now is that $\mathcal{B}$, a proper subdomain of $\mathbb{C}^2$, is biholomorphically equivalent to $\mathbb{C}^2$; that is, this basin of attraction is a Fatou–Bieberbach domain. More precisely, the claim is that when $j \to \infty$, the mapping $4^j\Phi^{[j]}$ converges uniformly on compact subsets of $\mathcal{B}$ to a biholomorphic mapping from $\mathcal{B}$ onto $\mathbb{C}^2$.

The immediate goal is to show that the sequence $\{4^j\Phi^{[j]}\}$ is a Cauchy sequence, uniformly on an arbitrary compact subset of $\mathcal{B}$. The strategy is to prove that consecutive terms of this sequence get exponentially close together.

Fix a compact subset of $\mathcal{B}$ and a natural number N such that the iterate $\Phi^{[N]}$ maps the specified compact set into the bidisc of radius $1/3$. If k is a positive integer, and z lies in the specified compact set, then

$$\begin{aligned}\left\|4^{N+k+1}\Phi^{[N+k+1]}(z) - 4^{N+k}\Phi^{[N+k]}(z)\right\|_\infty &= 4^{N+k}\left\|4\Phi\big(\Phi^{[N+k]}(z)\big) - \Phi^{[N+k]}(z)\right\|_\infty \\ &\le 4^{N+k}\cdot 2\cdot \|\Phi^{[N+k]}(z)\|_\infty^2 \qquad \text{by (4.2)} \\ &\le 4^{N+k}\cdot 2\cdot\left(\left(\tfrac{2}{3}\right)^{2k}\cdot\tfrac{1}{3}\right)^2 \qquad \text{by (4.1)} \\ &= \frac{2^{2N+1}}{9}\cdot\left(\frac{8}{9}\right)^{2k}.\end{aligned}$$

This geometric decay implies that the sequence $\{4^j\Phi^{[j]}\}$ is indeed a Cauchy sequence (uniformly on every compact subset of $\mathcal{B}$). Accordingly, a holomorphic limit mapping G appears that maps $\mathcal{B}$ into $\mathbb{C}^2$.

What remains to show is that G is both injective and surjective. Since the Jacobian determinant of F is identically equal to $-1/4$, the Jacobian determinant of Φ is identically equal to $1/16$, the Jacobian determinant of 4Φ is identically equal to 1, and the Jacobian determinant of $4^j\Phi^j$ is identically equal to 1 for every j. Hence the Jacobian determinant of G is identically equal to 1. Therefore the holomorphic mapping G is at least locally injective.

Exercise 41. Show that if a normal limit of injective holomorphic mappings (from a neighborhood in $\mathbb{C}^n$ into $\mathbb{C}^n$) is locally injective, then the limit is globally injective.

The surjectivity of G follows from the observation that $4G\circ\Phi = G$ (which is true because both sides of this equation represent the limit of the sequence $\{4^j\Phi^{[j]}\}$). Since Φ is a bijection of the basin $\mathcal{B}$, the range of G is equal to the range of $4G$. But the range of G contains a neighborhood of the origin in $\mathbb{C}^2$, and the only neighborhood of the origin that is invariant under dilation by a factor of 4 is the whole space.

Thus G is a holomorphic bijection from the basin $\mathcal{B}$ onto $\mathbb{C}^2$. Accordingly, the basin $\mathcal{B}$ is a Fatou–Bieberbach domain, as claimed.

4.1.2 Theorem

The general result of which the preceding example is a special case states that if F is an arbitrary automorphism of $\mathbb{C}^n$ with an attracting fixed point at the origin (meaning that every eigenvalue of the Jacobian matrix at the origin has absolute value less than 1), then the basin of attraction of the fixed point is biholomorphically equivalent to $\mathbb{C}^n$. Of course, the basin might be all of $\mathbb{C}^n$, but if the basin is a proper subset of $\mathbb{C}^n$, then the basin is a Fatou–Bieberbach domain. The appendix of the cited paper of Rosay and Rudin (pages 80–85) provides a proof of the theorem.

Research on Fatou–Bieberbach domains continues. One result from the twenty-first century is the existence of Fatou–Bieberbach domains that do not arise as the basin of attraction of an automorphism.[5]

4.2 Inequivalence of the ball and the bidisc

The Riemann mapping theorem implies that every bounded, simply connected domain in $\mathbb{C}^1$ can be mapped biholomorphically to the unit disc. In higher dimension, there is no such topological characterization of biholomorphic equivalence. Indeed, the open unit ball in $\mathbb{C}^2$ is not biholomorphically equivalent to the bidisc even though these two sets, viewed as subdomains of $\mathbb{R}^4$, are topologically (even diffeomorphically) equivalent.

The usual shorthand for this observation is that "there is no Riemann mapping theorem in higher dimension." But the story has another chapter. If a bounded, simply connected domain in $\mathbb{C}^n$ has connected, smooth boundary that is *spherical* (locally biholomorphically equivalent to a piece of the boundary of a ball), then indeed the domain is biholomorphically equivalent to a ball.[6]

The proof of the positive result is beyond the scope of this document. But the proof of the inequivalence of the ball and the bidisc is relatively easy. Here is one argument, based on the intuitive idea that the boundary of the bidisc contains one-dimensional complex discs, but the boundary of the ball does not. (Since holomorphic maps live on open sets and do not a priori extend to the boundary, a bit of trickery is needed to turn this intuition into a proof.)

Seeking a contradiction, suppose that there does exist a biholomorphic mapping from the unit bidisc to the unit ball. Let $\{a_j\}$ be an arbitrary sequence of points in the unit disc tending to the boundary, and consider the restriction of the alleged biholomorphic map to the sequence of one-dimensional discs $\{\, (a_j, z_2) : |z_2| < 1 \,\}$. The sequence of restrictions is a normal family (since the image is bounded), so there is a subsequence converging normally to a limit mapping from the unit disc into the boundary of the ball. Projecting onto a one-dimensional complex subspace through a point in the image produces a holomorphic function that realizes its maximum absolute value at an interior point of the unit disc, so the maximum principle implies that the limit mapping is constant.

Consequently, the z_2-derivative of each component of the original biholomorphic mapping tends to zero along the sequence of discs. For every point b in the unit disc, then, the z_2-derivative of each component of the mapping tends to zero along the sequence $\{(a_j, b)\}$. The sequence $\{a_j\}$ is arbitrary, and a holomorphic function in the unit disc that tends to zero

[5] Erlend Fornæss Wold, Fatou–Bieberbach domains, *International Journal of Mathematics* **16** (2005), number 10, 1119–1130.

[6] Shanyu Ji and Shiing-Shen Chern, On the Riemann mapping theorem, *Annals of Mathematics (2)* **144** (1996), number 2, 421–439.

along *every* sequence approaching the boundary reduces to the identically zero function. Thus the z_2-derivative of each component of the mapping is identically equal to zero in the whole bidisc. The same conclusion holds by symmetry for the z_1-derivative, so both components of the mapping reduce to constants. This conclusion contradicts the assumption that the mapping is biholomorphic. The contradiction shows that no biholomorphism from the bidisc to the ball can exist.

Many authors attribute to Henri Poincaré (1854–1912) the proposition that the ball and the bidisc are holomorphically inequivalent. Although Poincaré wrote an influential paper[7] about the holomorphic equivalence problem, there is no explicit statement of the proposition in the paper. Poincaré did compute the group of holomorphic automorphisms of the ball in $\mathbb{C}^2$. The automorphism group of the bidisc is easy to determine (the group is generated by Möbius transformations in each variable separately together with transposition of the variables) and is clearly not isomorphic to the automorphism group of the ball, so a straightforward deduction from Poincaré's paper does yield the proposition.

4.3 Injectivity and the Jacobian

A fundamental proposition from real calculus states that if a (continuously differentiable) mapping from a domain in $\mathbb{R}^N$ into $\mathbb{R}^N$ has Jacobian determinant different from zero at a point, then the mapping is injective in a neighborhood of the point. (Indeed, the mapping is a local diffeomorphism.) If a holomorphic mapping from a domain in $\mathbb{C}^n$ into $\mathbb{C}^n$ has complex Jacobian determinant different from zero at a point, is the mapping necessarily injective in a neighborhood of the point? An affirmative answer follows immediately from the following exercise.

Exercise 42. A holomorphic mapping from $\mathbb{C}^n$ to $\mathbb{C}^n$ induces a real transformation from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ (through suppression of the complex structure). Show that the determinant of the $2n \times 2n$ real Jacobian matrix equals the square of the absolute value of the determinant of the $n \times n$ complex Jacobian matrix.
Hint: When $n = 1$, writing a holomorphic function $f(z)$ as $u(x, y) + iv(x, y)$ and applying the Cauchy–Riemann equations shows that

$$|f'|^2 = u_x^2 + v_x^2 = \det \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix}.$$

Can you generalize to higher dimension?

(Moreover, when the complex Jacobian determinant is nonzero, the mapping is a local biholomorphism. Indeed, the chain rule implies that the first-order real partial derivatives of the inverse mapping satisfy the Cauchy–Riemann equations.)

[7] Henri Poincaré, Les fonctions analytiques de deux variables et la représentation conforme, *Rendiconti del Circolo Matematico di Palermo* **23** (1907) 185–220.

In real calculus, the Jacobian determinant of an injective mapping can have zeros. Indeed, the function of one real variable that sends x to x^3 is injective but has derivative equal to zero at the origin. The story changes for holomorphic mappings.

A standard property from the theory of functions of one complex variable states that a holomorphic function is locally injective in a neighborhood of a point if and only if the derivative is nonzero at the point. A corresponding statement holds in higher dimension but is not obvious. The goal of this section is to provide a proof that a locally injective holomorphic mapping from a domain in $\mathbb{C}^n$ into $\mathbb{C}^n$ has nonzero Jacobian determinant.

This theorem appeared in the 1913 PhD dissertation of Guy Roger Clements (1885–1956) at Harvard University. An announcement[8] appeared in 1912 with the details published[9] the following year. When I was a graduate student, an elegant short proof appeared in the now standard textbook on algebraic geometry by Griffiths and Harris.[10] The idea was rediscovered by Rosay[11] and embellished by Range.[12] The proof below is my implementation of the method.

Exercise 43. The story changes when the dimension of the domain does not match the dimension of the range. Find an example of an injective holomorphic mapping from $\mathbb{C}^1$ to $\mathbb{C}^2$ whose derivative vanishes at the origin. What about mappings from $\mathbb{C}^2$ to $\mathbb{C}^1$?

The proof of the proposition is straightforward for readers who know the concept of a variety. The goal of the following lemma is to provide an "elementary" and self-contained proof that avoids explicit mention of varieties.

Lemma 7. Suppose f is a holomorphic function defined on a neighborhood of a point p in $\mathbb{C}^n$, where $n \geq 2$, and $f(p) = 0$.

1. If the gradient of f at p is not the zero vector, then there is a local biholomorphic change of coordinates near p after which p becomes the origin and the zero set of f becomes the subspace $\mathbb{C}^{n-1} \times \{0\}$ in a neighborhood of the origin.

2. If the gradient of f at p is the zero vector, but the function f is not identically equal to 0 in a neighborhood of p, then there is a nearby point q and a local biholomorphic change of coordinates near q after which q becomes the origin and the zero set of f becomes the subspace $\mathbb{C}^{n-1} \times \{0\}$ in a neighborhood of the origin.

[8] G. R. Clements, Implicit functions defined by equations with vanishing Jacobian, *Bulletin of the American Mathematical Society* **18** (1912) 451–456.

[9] Guy Roger Clements, Implicit functions defined by equations with vanishing Jacobian, *Transactions of the American Mathematical Society*, **14** (1913) 325–342.

[10] Phillip Griffiths and Joseph Harris, *Principles of Algebraic Geometry*, Wiley, 1978, pages 19–20. The book was reprinted in 1994.

[11] Jean-Pierre Rosay, Injective holomorphic mappings, *American Mathematical Monthly* **89** (1982), number 8, 587–588.

[12] R. Michael Range, *Holomorphic Functions and Integral Representations in Several Complex Variables*, Springer, 1986, Chapter I, Section 2.8.

Proof. Readers who already know the basic notions of analytic varieties can view the zero set of f as an $(n-1)$-dimensional variety. The first statement says that the point p is nonsingular, so the variety can be straightened locally by a holomorphic change of coordinates. The second statement follows by taking q to be a regular point of the variety close to p. Readers unfamiliar with the technology of varieties can proceed as follows.

To prove the first statement, start by translating p to 0, and then permute the coordinates to arrange that $\partial f/\partial z_n(0) \neq 0$. Now consider the map sending a general point $(z_1, \dots, z_n)$ to the image point $(z_1, \dots, z_{n-1}, f(z_1, \dots, z_n))$. The Jacobian determinant of this mapping equals the nonzero value $\partial f/\partial z_n(0)$, so the transformation has a local holomorphic inverse G near the origin. Since $f \circ G$ vanishes precisely when the nth coordinate is equal to 0, the mapping G provides the required change of coordinates. (This argument essentially says that f itself can be chosen as one of the n local complex coordinates.)

The second statement of the lemma follows from the first statement whenever there is a nearby point q such that q lies in the zero set of f and the gradient of f at q is not the zero vector. In general, however, such a point q need not exist, for f might be the square of another holomorphic function, in which case the gradient of f vanishes wherever f does. To handle this situation, fix a neighborhood of p (as small as desired) and observe that (since f is not identically equal to zero) there is a minimal natural number k such that all derivatives of f of order k or less vanish identically on the zero set of f in the specified neighborhood, but some derivative of f of order $k+1$ is nonzero at some point q in the zero set of f in the neighborhood. Apply the first case of the lemma to the appropriate kth-order derivative of f whose gradient is not zero.

After a suitable holomorphic change of coordinates, the point q becomes the origin, and the zero set of the specified kth-order derivative of f becomes the subspace $\mathbb{C}^{n-1} \times \{0\}$ in a neighborhood of the origin in $\mathbb{C}^n$, say in a polydisc of radius ε centered at 0. By construction, the zero set of f in the polydisc is a nonvoid *subset* of the subspace $\mathbb{C}^{n-1} \times \{0\}$. It remains to show that the zero set of f is identical to this subspace in some neighborhood of the origin.

In the contrary case, the closedness of the zero set of f implies the existence of an open polydisc D in $\mathbb{C}^{n-1}$ centered at some point w of $\mathbb{C}^{n-1} \times \{0\}$ such that $\max_{1 \le j \le n-1} |w_j| < \varepsilon/2$ and the zero set of f is disjoint from D. Shrink D, if necessary, to ensure that D is entirely contained in the polydisc of radius $\varepsilon/2$ centered at 0. On the Hartogs figure

$$D \times \{ z_n \in \mathbb{C} : |z_n| < \varepsilon \} \bigcup \{ z \in \mathbb{C}^n : \varepsilon/2 < |z_n| < \varepsilon \text{ and } \max_{1 \le j \le n-1} |z_j - w_j| < \varepsilon/2 \},$$

the function f is holomorphic and nowhere equal to 0, so the function $1/f$ is holomorphic. By the Hartogs phenomenon, the function $1/f$ extends to be holomorphic on the polydisc

$$\{ (z_1, \dots, z_{n-1}) : \max_{1 \le j \le n-1} |z_j - w_j| < \varepsilon/2 \} \times \{ z_n : |z_n| < \varepsilon \}.$$

Since this polydisc contains the origin, which lies in the zero set of f, a contradiction arises. The contradiction means that the zero set of f must be identical to the subspace $\mathbb{C}^{n-1} \times \{0\}$ in some neighborhood of the origin. □

The proof of the proposition about nonvanishing of the Jacobian determinant of an injective holomorphic mapping uses induction on the dimension, the one-dimensional case being known. Suppose, then, that the proposition has been established for dimension $n-1$, and consider a local injective holomorphic mapping $(f_1, \ldots, f_n)$ that (without loss of generality) fixes the origin.

If the gradient of the coordinate function f_n at the origin is nonzero, then the first case of the lemma reduces the problem (via a local biholomorphic change of coordinates) to the situation that f_n vanishes precisely on the subspace where $z_n = 0$. Consequently, $\partial f_n/\partial z_j(0) = 0$ when $j \neq n$, and $\partial f_n/\partial z_n(0) \neq 0$. Moreover, the n-dimensional mapping now takes $\mathbb{C}^{n-1} \times \{0\}$ into itself. The mapping of $\mathbb{C}^{n-1}$ obtained by restricting to the subspace where $z_n = 0$ has nonzero Jacobian determinant at the origin by the induction hypothesis. The Jacobian determinant of the n-dimensional mapping at the origin equals this nonzero Jacobian determinant of the $(n-1)$-dimensional mapping multiplied by the nonzero factor $\partial f_n/\partial z_n(0)$. Thus the required conclusion holds when the gradient of f_n at the origin is nonzero.

Making a permutation of the variables in the range shows, by the same argument, that the Jacobian determinant of the mapping at the origin is nonzero if any one of the coordinate functions has nonzero gradient at the origin. Accordingly, the problem reduces to showing that a contradiction arises if the Jacobian matrix at the origin has all entries equal to zero.

The Jacobian determinant is then a holomorphic function that equals zero at the origin. Making a suitable holomorphic change of coordinates at a nearby point via the second part of the lemma reduces to the situation that the zero set of the Jacobian determinant near the origin is precisely $\mathbb{C}^{n-1} \times \{0\}$. By the first part of the argument, each coordinate function of the mapping has vanishing gradient on this complex subspace. Consequently, the coordinate functions are constant along $\mathbb{C}^{n-1} \times \{0\}$, contradicting the injectivity of the mapping. This contradiction implies that the Jacobian determinant of the injective holomorphic mapping cannot vanish after all, thus completing the proof by induction.

4.4 The Jacobian conjecture

The Fatou–Bieberbach example discussed in Section 4.1 is a holomorphic map $G : \mathbb{C}^2 \to \mathbb{C}^2$ whose Jacobian determinant is identically equal to 1 (whence G is everywhere locally invertible), yet G is not surjective (whence G is not globally invertible as a map from $\mathbb{C}^2$ to $\mathbb{C}^2$). The map G appears as a normal limit of polynomial maps, but G itself is not a polynomial map. An unresolved problem of long standing is the nonexistence of weird polynomial maps.

Open Problem (Jacobian conjecture)**.** *For every positive integer n, if $F : \mathbb{C}^n \to \mathbb{C}^n$ is a polynomial mapping whose Jacobian determinant is identically equal to* 1*, then F is a polynomial automorphism of $\mathbb{C}^n$.*

The conclusion entails that F is both injective and surjective, and the inverse of F is a polynomial mapping. It is known that if F is globally injective, then the other two conclusions follow. The conjecture is known to be true for mappings involving only polynomials of degree at most 2. Moreover, the conjecture holds in general if it holds for mappings involving only polynomials of degree at most 3. Many alleged proofs of the conjecture have been published, none correct (so far). It is unclear whether the conjecture is essentially a problem of algebra, analysis, combinatorics, or geometry.[13]

A natural real analogue of the Jacobian conjecture is known to be false. Pinchuk[14] produced a remarkable example of a polynomial map $F: \mathbb{R}^2 \to \mathbb{R}^2$ whose Jacobian determinant is everywhere positive, yet F is not a global diffeomorphism from $\mathbb{R}^2$ onto $\mathbb{R}^2$.

[13]The indicated facts and much more can be found in the following article: Hyman Bass, Edwin H. Connell, and David Wright, The Jacobian conjecture: Reduction of degree and formal expansion of the inverse, *Bulletin of the American Mathematical Society* **7** (1982), number 2, 287–330.

[14]Sergey Pinchuk, A counterexample to the strong real Jacobian conjecture, *Mathematische Zeitschrift* **217** (1994), number 1, 1–4.