The Cauchy–Kovalevskaya Theorem

Roger Cooke

1 Introduction

The Cauchy–Kovalevskaya theorem was proved in several stages, first for ordinary differential equations (1835) by Cauchy, then for a special quasi-linear partial differential equation (1842), also by Cauchy. Not knowing of Cauchy's work, Weierstrass proved a similar theorem for ordinary differential equations in 1842. Although both authors used (having apparently discovered independently) the method of majorants, an essential difference between their approaches to analytic function theory shows up very clearly in the kinds of theorems they were able to prove. The Cauchy approach for ordinary differential equations was expounded systematically by Briot and Bouquet in 1856, precisely at the time when Weierstrass was proving further theorems of the same type. After that, little seems to have been done for nearly two decades, when Kovalevskaya and Darboux almost simultaneously published extensions to partial differential equations, all the parties concerned being ignorant of Cauchy's extension to partial differential equations, published a generation earlier. The evolution of the theorem highlights a number of interesting issues in the history of mathematics, such as the allowable definitions of a function, questions of existence of mathematical objects in general, the meaning of complex numbers in physical problems, and what it means to solve a differential equation. In the following paragraphs we shall trace this evolution over the past two centuries in order to bring out these issues.

1.1 The context of the theorem: the method of undetermined coefficients

Ordinary differential equations arose almost as soon as there was a language (differential calculus) in which they could be expressed. These equations were used to formulate problems from geometry and physics in the late seventeenth century, and the natural approach to solving them was to apply the integral calculus, that is, to reduce a given equation to quadratures. It appeared at a very early stage that such a reduction is not in general possible, and both Newton and Leibniz were led early on to the use of infinite series with undetermined coefficients to solve such equations. The earliest such examples were cited by L. Feigenbaum ("Infinite series and solutions of ordinary differential equations, 1670–1770," in: *Companion Encyclopedia of the History and Philosophy* *of the Mathematical Sciences*, Vol. 1, Routledge, London and New York, 1994, pp. 504–519). In a set of notes entitled "Tractatus pro methodis serierum et fluxionum," written around 1670 and left unpublished during his lifetime (D. L. Whiteside, ed., *Mathematical Papers of Isaac Newton*, Vol. 3, p. 99), Newton considered the linear differential equation that we would now write as

$$\frac{dy}{dx} = 1 - 3x + x^2 + (1 + x)y.$$

(Newton wrote it as $n/m = 1 - 3x + y + xx + xy$.) He assumed tacitly that the constant term in the series expansion of y is zero, which, by the equation, gave a constant term 1 in the series for its derivative, thereby making the linear term in the series for y equal to x. The equation then shows that the linear term in the series for the derivative must be $-2x$, so that the quadratic term in the series for y must be $-x^2$. The equation then shows that the quadratic term in the series for the derivative $\frac{dy}{dx}$ is x^2, and so forth. Newton arranged this whole computation in an array that we would now call a matrix, without explicitly writing down any unknown coefficients. In this way he found

$$y = x - x^2 + \frac{1}{3}x^3 - \frac{1}{6}x^4 + \frac{1}{30}x^5 - \frac{1}{45}x^6 - \cdots.$$

Similarly, in a paper entitled "Supplementum geometriae practicae sese ad problemata transcendentia extendens, ope novae methodi generalissimae per series infinitas," published in *Acta Eruditorum* in 1693 (see *G. W. Leibniz, Mathematische Schriften*, Vol. V, Georg Olms Verlag, Hildesheim and New York, 1971, p. 287), Leibniz studied the differential equations for the logarithm and the arcsine in order to obtain what we now call the Maclaurin series of the logarithm, exponential, and sine functions. For example, he considered the equation $a^2\,dy^2 = a^2\,dx^2 + x^2\,dy^2$ and assumed $x = by + cy^3 + ey^5 + fy^7 + \cdots$, thereby obtaining the series that represents the function $x = a\sin\left(\frac{y}{a}\right)$.

Neither Newton nor Leibniz mentioned that the coefficients in these series were the derivatives of the functions represented by the series divided by the corresponding factorials. However, that realization came to Johann Bernoulli very soon after the publication of Leibniz' work. In a letter to Leibniz dated 2 September 1694 (see *G. W. Leibniz, Mathematische Schriften*, Vol. III/1, Georg Olms Verlag, Hildesheim and New York, 1971, p. 350), Bernoulli wrote:

> I take the opportunity here to sketch a universal series that occurred to me recently, which expresses any quadratures and rectifications in general, and which is very well suited to the inverse method of tangents: Suppose given an arbitrary differential that is to be integrated (I take n to be a quantity formed in an arbitrary manner from variables and constants), and let dx be taken constant. The integral will be equal to the following series: $nz - \frac{1}{1\cdot 2}zz\frac{dn}{dz} + \frac{1}{1\cdot 2\cdot 3}z^3\frac{ddn}{dz^2} - \frac{1}{1\cdot 2\cdot 3\cdot 4}z^4\frac{dddn}{dz^3}$, and so forth, which can be applied in a particular case by canceling dn, ddn, $dddn$, and so forth with dz, dz^2, dz^3, and so forth, so that the series consists entirely of purely algebraic terms.

> By using this universal series I easily find the right sine of a given arc and radius, and I solve other problems that you solved by your own method in the *Acta* in April 1693. The method by which I found this series will appear in the *Acta*, where you will see more extensive uses of it.

Obviously Bernoulli had obtained the series here through integration by parts. (The alternating signs are conclusive evidence of that.) Incidentally, his description of the integrand n came to be referred to later as "Bernoulli's definition of a function." In the seventeenth century the notion of a function was not yet clearly formed, although Leibniz had used the word as early as 1673. In an article in the 1694 *Acta Eruditorum* entitled "Nova calculi differentialis applicatio et usus ad multiplicem linearum constructionem ex data tangentium conditione" (see *G. W. Leibniz, Mathematische Schriften*, Vol. V, Georg Olms Verlag, Hildesheim and New York, 1971, pp. 301–306) he stated, "In general the problem can be thought of as follows: Given the ratio between two functions, find the curve…I define a *function*[1] to be the portion of a line cut off by lines drawn using only a fixed point and a given point lying on a curved line." As Leibniz explicitly said, a given curve defines a number of functions, namely its abscissas, its ordinates, its subtangents, and so on. The problem is to reconstruct the curve given the ratio between any two of these functions. (The "fixed" point in this definition is apparently allowed to move along a fixed straight line.)

This digression on the meaning of the word function is included here because it raises an important issue in the subject of differential equations. Classically the solution of a differential equation is a function or family of functions. Given that fact, the ways in which a function can be legally presented become an important issue. With the modern definition of a function and the familiar notation one might easily forget that in order to apply the theory of functions it is necessary to deal with particular functions, and these must be presented somehow. Bernoulli's description, which he explicitly used with the word *function* in 1718, addresses that issue, although it leaves open the question of what methods of combining variables and constants are legal.

As an illustration of the issues involved, consider the difference between Cauchy's definition of an analytic function and that of Weierstrass. Cauchy talks about functions in essentially the modern sense and defines a function to be analytic if it has a continuous complex derivative at every point, whereas Weierstrass defines a function to be analytic if it is defined by a power series in a neighborhood of each point. At first sight Cauchy's definition appears to be more general, and his theorem that an analytic function has a power-series expansion looks like a powerful theorem. However, in defense of Weierstrass, one must ask how we are to verify that a particular function has a continuous

[1] The use of the Latin word *functio* sounds peculiar in this context. The word means *performance* or *execution*, and is the noun derived from the participle of the deponent verb *fungor*, whose meaning is to discharge a duty or fulfill an office. Modern cognates are *functionary* in English and *fonctionnaire* in French, both meaning a bureaucrat.

derivative. The function must be presented to us in some manner. Quite often that presentation involves a formula, and a power series is one very common way of doing so. Indeed, in 1835 Cauchy proved the existence of a solution to an ordinary differential equation by demonstrating that the series generated a power series that he could prove convergent. It followed that the solution was analytic, but Cauchy's aim was clearly existence rather than analyticity. Thus, when Cauchy himself needed to exhibit a particular function, he resorted to power series to define it. Of course, functions can be presented without the use of power series. Examples would appear to be the elementary functions, which can be defined in terms of real and imaginary parts, for example, $e^{x+iy} = e^x \cos y + ie^x \sin y$. But the power series expansions of these functions can be developed from a real-variable point of view, *without* the general Cauchy theory. Thus, from a certain point of view (that of Weierstrass) there is no real loss in generality in taking power series as the starting point for the theory of analytic functions. Weierstrass took a great deal of satisfaction from the fact that both Cauchy and Riemann needed power-series representations. He was quoted as having said in 1884, "No matter how you twist and turn, you cannot avoid using some definite analytic representation, such as power series." (See "Zur Funktionentheorie," *Acta Mathematica*, **45** (1925), pp. 1–10. Quoted by R. Siegmund-Schultze, in *Ausgewählte Kapitel aus der Funktionenlehre*, Teubner, Leipzig, 1988, p. 253.)

Thus, one of the issues involved with the Cauchy–Kovalevskaya theorem is the variant approaches to complex analysis represented by the Cauchy integral theorem and the power-series approach. In either approach one can study functions defined by differential equations, and in that case a complex integral can be regarded merely as a function defined by a rather trivial type of differential equation. In many cases the same results were derived from these two different points of view. In particular, the result known nowadays as Mittag-Leffler's theorem, which gives a representation of a meromorphic function in terms of its principal parts at its poles, was proved almost simultaneously in the early 1880s by Mittag-Leffler and by Runge, the former using power series, the latter using the Cauchy integral. (Relations between these two men were very cordial, at least at the time. Mittag-Leffler published five of Runge's papers in his journal *Acta Mathematica* during 1885. Those were the only papers of Runge to appear in that journal. His interests were more in physics than mathematics, and none of his many subsequent papers were on topics suitable for publication in the *Acta*. Although Mittag-Leffler was ten years older than Runge, the two men died only six months apart, Runge in January 1927 and Mittag-Leffler in July of that year.)

If analytic functions are to be defined as the solutions of certain differential equations, it is essential to prove that the equations and initial or boundary conditions really do unambiguously define an analytic function. For ordinary differential equations Cauchy's disciples Briot and Bouquet took this approach in 1856, as Weierstrass had done earlier, in 1842. As mentioned above, here is a case where an analytic function is initially presented as a series whose convergence must be proved, rather than as a function whose partial derivatives satisfy the Cauchy–Riemann equations. One of the important results of moving the theory of differential equations into the complex domain was that an analytic function could be presented by giving a differential equation.

We are getting a bit ahead of the story, however. In the eighteenth century the use of power series to solve ordinary differential equations became quite familiar. These series were especially compatible with classical mechanics, since the Taylor series of a given function can be generated knowing the values of the function over any interval of the independent variable, no matter how short. Thus a quantity represented by such a series is determined for all values of the independent variable when the values are given on any interval at all. Given that the independent variable is usually time, that property corresponds to physical determinacy: Knowing the full state of a physical quantity for some interval of time determines its values for all time. Lagrange, in particular, was a proponent of power series, for which he invented the term *analytic function.* As we now know, the natural domain of analytic function theory is the complex numbers, or variants of it (Riemann surfaces, Weierstrass' *Gebilde*, and the like). What we now call analytic functions of a complex variable proved extremely useful in cartography because of their conformal mapping properties, and it turned out that the two components of velocity for a steady nonviscous planar fluid flow are precisely the real and imaginary parts of an analytic function (see Grattan-Guinness, *Convolutions in French Mathematics 1800–1840*, Birkhäuser, Basel, 1990, pp. 336–338 and 662–664).

In mechanics, however, the independent variable represents *time*, and that fact raises an important question: Why should time be a complex variable? To this question the eighteenth- and nineteenth- century mathematicians gave no answer. Indeed, it does not appear that many of them even asked the question. Was this omission due to blindness on their part? Since the relevant texts are silent on the subject, any answer must be conjectural. Here is a possible explanation, however. Complex numbers were first taken seriously in connection with cubic equations having real coefficients and real solutions. To solve even a simple cubic equation such as $x^3 - 6x + 3 = 0$, for example, the Cardano formula gives

$$x = -\big(\sqrt[3]{1.5 + \sqrt{-5.75}} + \sqrt[3]{1.5 - \sqrt{-5.75}}\big).$$

The three possible values of each of the two complex numbers on the right here are to be chosen as complex conjugates of each other, revealing three real solutions. Thus, even though the data and the solution are real numbers, the process of getting from the one to the other requires the use of complex numbers.[2] It was the inability to make the transition from coefficients to roots algebraically using only real numbers that led to the full theoretical development of the complex numbers. In the nineteenth century Galois investigated the question of which algebraic equations could be solved by radicals, laying the foundations of group theory in the process. Similarly, Liouville investigated which integrals could be

[2] François Viète, however, showed how to solve such a cubic equation using purely real-variable methods and trigonometric functions. Because trigonometric functions are involved, Viète's method is not algebraic.

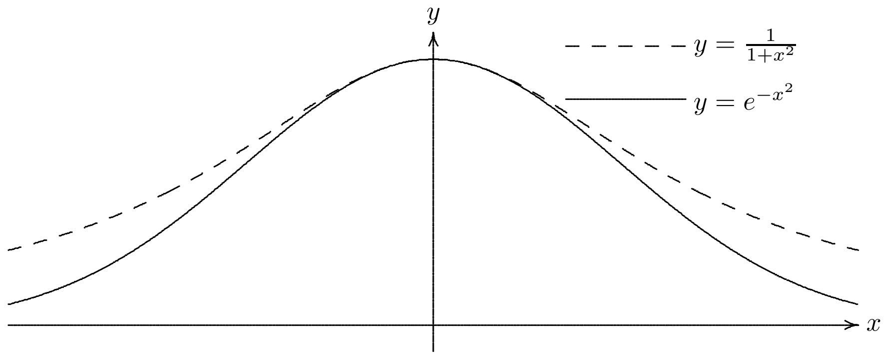

Figure 1: Two "bell-shaped" curves whose Maclaurin series exhibit different behaviors

evaluated in terms of algebraic functions, thereby beginning the subject of differential Galois theory. The two subjects were seen to exhibit certain analogies. These analogies were noticed by Cauchy, as will be seen below, and by Lie, who aimed at a group-theoretic discussion of differential equations in analogy with the Galois theory of algebraic equations. For that reason it would probably not have seemed strange that the complex-variable properties of solutions are relevant, even in the study of problems generated by physical considerations involving only real variables. For example, the two functions $f(x) = e^{-x^2}$ and $g(x) = \frac{1}{1+x^2}$ have very similar "bell-shaped" graphs, as Fig. 1 shows.

Yet their Maclaurin series behave very differently:

$$\begin{aligned} e^{-x^2} &= 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots + \frac{(-x^2)^n}{n!} + \cdots; \\ \frac{1}{1+x^2} &= 1 - x^2 + x^4 - x^6 + \cdots + (-x^2)^n + \cdots. \end{aligned} \tag{1}$$

The series for $f(x)$ converges for all x, while the series for $g(x)$ converges only for $-1 < x < 1$. This difference is impossible to account for on the basis of real values of x. But when x is replaced by the complex variable z, the difference appears immediately: $f(z)$ is analytic for all complex values of z, whereas $g(x)$ has poles at the points $z = \pm i$, which lie on the boundary of the circle of convergence of the Maclaurin series. We shall see below that the difference between these two functions is of great importance in explaining the limitations of the Cauchy–Kovalevskaya theorem.

Despite these considerations, mathematicians and physicists do not seem to have considered the implications of a complex-variable time at much length. Although the time coordinate in relativity is often regarded as imaginary, in order to get a positive-definite metric, time has always been considered as one real dimension, not two. Undoubtedly the reason is its directionality, from past to future, which would apparently be lost if it were made complex. On the other hand, important physical information can come from mathematical quantities that do not by themselves have physical meaning. For example, there does not seem to be any reason why the harmonic conjugate variable to a physically observable quantity needs to be observable, yet it may still contain important information about its physically interesting conjugate part. The difficulty in the case of time, however, is that it is nearly always an independent variable, so that its conjugate is trivial. Extensive searches of the literature by the present author have produced only the following comments on this interesting question, made by Weierstrass in 1885 (see his *Werke*, Bd. 3, S. 24):

> It is very remarkable that in a problem of mathematical physics where one seeks an unknown function of two variables that, in terms of their physical meaning, can have only real values and is such that for a particular value of one of the variables the function must equal a prescribed function of the other, an expression often results that is an analytic function of the variable and hence also has a meaning for complex values of the latter.

With this digression on the use of complex variables now complete, let us resume the story of differential equations.

In the middle of the eighteenth century mathematical physicists began to consider problems involving more than one independent variable. The most famous of these is the vibrating string problem discussed by Euler, d'Alembert, and Daniel Bernoulli during the 1740s and 1750s. At that point, partial differential equations arose, leading to new methods of solution (trigonometric series, in particular). Power series in several variables remained a viable technique, and the method of undetermined coefficients continued to be applied, even in the solution of partial differential equations. However, it turned out that certain complications arise and cause a kind of chaos that could be glimpsed only dimly in the study of ordinary differential equations. To explain the difficulties in the simplest case, consider Laplace's equation in polar coordinates:

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0. \tag{2}$$

Based on experience with the wave equation (which in rectangular coordinates looks almost the same as Laplace's equation), one ought to be able to prescribe initial values of u and $\frac{\partial u}{\partial r}$ at a value of r such as $r = 1$ and then solve the equation by undetermined coefficients. And indeed, if we assume

$$u(r,\theta) = \sum_{m,n=0}^{\infty} c_{m,n}(r-1)^m \theta^n, \tag{3}$$

the equation leads us to the recursive relation

$$\begin{aligned} c_{m,n+2} = -\tfrac{(m+2)(m+1)c_{m+2,n} + \left((m+1)c_{m+1,n} - mc_{m,n} + \cdots + (-1)^m c_{1,n}\right)}{(n+2)(n+1)} + \\ + 2c_{m-1,n+2} - 3c_{m-2,n+2} + \cdots + (-1)^{m+1}(m+1)c_{0,n+2}. \end{aligned} \tag{4}$$

which shows that the coefficients $c_{m,n}$ are uniquely determined once all the coefficients $c_{m,0}$ and $c_{m,1}$ are given. As far as the algebraic determination of the coefficients using Eq. (4) is concerned, these first two rows of coefficients, which determine $u(1, \theta)$ and $u_r(1, \theta)$, can be arbitrary.

On the other hand, there seems to be a paradox here. The solutions of Eq. (2) are harmonic functions, and it is well known that a harmonic function in the disk defined by $r < 1$ is uniquely determined by the values $u(1, \theta)$ alone. What goes wrong, as we now know, is that polar coordinates have a singularity at $r = 0$. Many different functions harmonic in an annulus containing the circle $r = 1$ may have the same values on that circle; at most one of these functions will be harmonic in the entire closed disk. In a sense, the method of undetermined coefficients works well for initial-value problems, as in the present example, where we prescribe the values and both the function and its the derivative in connection with a second-order differential equation; it does not work so well for boundary-value problems.

Guided by physical intuition, one can easily be led to the wrong conclusions. The one-dimensional heat equation $\frac{\partial u}{\partial t} = a\frac{\partial^2 u}{\partial x^2}$, whose solutions $u(x, t)$ are supposed to represent the temperature at time t at a point x units from a reference point in a long straight wire, revealed the difficulties with stark clarity, as we shall see below. One feels intuitively that the initial temperature distribution $u(x, 0)$ ought to determine the temperature for all time, and indeed the method of undetermined coefficients seems to justify this assumption. Assuming the expansion

$$u(x, t) = \sum_{m,n=0}^{\infty} c_{m,n} x^m t^n, \tag{5}$$

we find the recursive relation

$$(n + 1)c_{m,n+1} = (a(m + 2)(m + 1)c_{m+2,n}. \tag{6}$$

The relation (6) makes it possible to determine the temperature given the coefficients of the power series for the distribution $u(x, 0)$ at time $t = 0$.

On the other hand, the relation (6) can be used to generate recursive relations on either index. If the recursion is conducted with respect to the index n, then only the first row, that is, the coefficients $c_{m,1}$, can be prescribed arbitrarily, all others then being determined. But if it is conducted with respect to m, then two columns of coefficients, namely $c_{0,n}$ and $c_{1,n}$, can be prescribed arbitrarily. How does it happen that the same physical quantity requires only one set of initial values for its determination on one variable, but two sets on the other? This is one of the subtleties that arise when one seeks analytic solutions of partial differential equations. It was Kovalevskaya who first drew attention to these problems, as we shall see below.

The heat equation was the first partial differential equation to which the method of undetermined coefficients was applied, and its complexities show up very clearly in the contrasting approaches adopted in the early nineteenth century by Jean-Baptiste Joseph Fourier (1768–1830) and Siméon-Denis Poisson

(1781–1840). Poisson considered an initial-value problem on the spatial variable x, namely $u(0,t) = f(t)$, and arrived at the solution $u(x,t) = \sum_{r=0}^{\infty} \frac{x^{2r}}{(2r)!} f^{(r)}(t) + \sum_{r=0}^{\infty} \frac{x^{2r+1}}{(2r+1)!} g^{(r)}(t)$. Fourier, in contrast, followed Daniel Bernoulli's idea of superimposing the basic solutions $C_q e^{-q^2 t} \cos qx$, and was led to seek integral solutions of the form $u(t, x0 = \int_0^\infty Q(q) e^{-q^2 t} \cos qx \, dq$. In an effort to choose the arbitrary function $Q(q)$ so that the initial condition $u(x,0) = f(x)$ will hold, Fourier was led to discover what is now known as the Fourier integral formula:

$$f(x) = \frac{2}{\pi} \int_0^\infty \int_0^\infty f(t) \cos qt \cos qx \, dt \, dq.$$

Fourier also applied the method of undetermined coefficients to produce the solution

$$u(x,t) = \sum_{r=0}^{\infty} \frac{\varphi^{(2r)}(x)}{r!} t^r$$

for the heat equation, without realizing that this solution "usually" diverges.

This form of the solution seems to have been rediscovered a number of times. The limits of its validity were one of Koavalevskaya's best discoveries as a student of Weierstrass in the mid-1870s. As we shall see below, it makes a great deal of difference which of these two initial conditions one imposes on the equation. For more details, and the contributions of Laplace and Cauchy to the same problems and methods, see the book *Joseph Fourier, 1768–1830*, by I. Grattan-Guinness, The MIT Press, 1972, pp. 441–470.

2 Ordinary Differential Equations

The Cauchy–Kovalevskaya theorem is one of the fundamental results on the theory of partial differential equations in the complex domain. We discussed above the rather surprising aspects of the use of complex variables, considering that the variables that occurred in such equations were intuitively real. Since these variables represent time, mass, distance, and the like, we naturally raise two questions: First, how did differential equations come to be studied mathematically using the theory of functions of a complex variable? Second, what interpretation was given to complex variables in physical applications?

To answer the first question, there are two main reasons why complex variables came to be used in the study of differential equations. The first is that analytic functions have the kind of determinacy that one expects of classical physical systems, and their mathematical nature cannot be understood except in the context of complex variables. The second motive comes from the mathematical efficacy of the method of undetermined coefficients, which we mentioned above and which will be described in more detail below.

As for the second question, whether our conjecture about the mathematical attitude toward complex variables is correct or not, the theory of complex variables was often used in the nineteenth century to generate conjectures about systems of differential equations. For example, Kovalevskaya's other famous work, on the rotation of a rigid body, was a study of cases in which the components of the angular velocity of a rigid body rotating about a fixed point can be (what are now called) meromorphic functions of time.[3] Intriguing as this second question is, we must leave it aside. The first question will be addressed in the present essay, in which we examine the actual process of "complexification" of differential equations via the Cauchy–Kovalevskaya theorem.

2.1 Analytic function theory: Cauchy and Weierstrass

It was Cauchy who, in the 1830s, discovered the essential core of what is now complex analysis, including complex integration, and the expansion of an analytic function in a Taylor series. Developing the theory of power series more rigorously than had been done in the past, he investigated the convergence of such series and arrived at the crucial insight that a power series converges if and only if it is dominated by a geometric series. This fact amounts to what is today called Cauchy's inequality for the nth derivative of a function $f(z)$ that is analytic in a disk of radius r about a point z_0:

$$|f^{(n)}(z_0)| \le n!M(\rho)\rho^{-n}, \qquad n = 0, 1, 2, \dots \tag{7}$$

where ρ is any number satisfying $0 < \rho < r$, and $M(\rho)$ is the maximum value of $|f(z)|$ on the circle $|z - z_0| = \rho$. This result, along with many others, was presented at the Turin Academy in 1831. If a function $f(z)$ has a (complex) derivative in the disk of radius r about z_0, it does satisfy inequalities (7). It must therefore be analytic in the disk, since these inequalities imply that its Taylor series converges. Cauchy was not entirely clear about all this, since the distinction between real derivatives and complex derivatives was not yet as distinct as it has since become. At the time he believed, in fact, that a continuous function would have a (real) derivative at "most" points. By the early 1830s he had developed the heart of modern complex analysis and with more rigor than had been customary in the eighteenth century. The inequalities (7) are consequences of the fundamental fact about complex integrals of analytic functions, known as the Cauchy integral theorem. This theorem, in the case of a circle C, asserts that if the function $f(z)$ has a complex derivative on a disk that contains the circle C in its interior, then

$$\int_C f(z)\, dz = 0. \tag{8}$$

Cauchy assumed that the derivative was continuous, in which case the theorem is a consequence of what is now known as Green's theorem (which, however, was not well known when Cauchy proved his result). The assumption that $f'(z)$ is

[3] To be sure, she had originally set out to solve these equations in the general case. Only when she was unable to do that using abelian functions did she, at Weierstrass' urging, analyze the equations to see when a solution is possible using this technique.

continuous was retained until 1900, when Edmund Goursat proved the theorem without that extra assumption (in a paper entitled "Sur la définition générale des fonctions analytiques d'après Cauchy," *Transactions of the American Mathematical Society*, **1**, No. 1, 14–16).

Inequalities (7) are easily generalized to functions of two variables $f(z, w)$ that are analytic in an open domain R containing all points (z, w) satisfying $|z - z_0| \le r$, $|w - w_0| \le \rho$:

$$\left|\frac{\partial^{m+n} f}{\partial z^m \partial w^n}\right| \le m!n!M(r, \rho)r^{-m}\rho^{-n}, \tag{9}$$

where $M(r, \rho)$ is an upper bound for $|f(z, w)|$ in the region where $|z - z_0| \le r$, $|w - w_0| \le \rho$. Cauchy's notation was not quite the modern notation given here. Instead, he considered a function of a general set of variables $f(x, y, z, \ldots, t)$, denoting the complex increments of these variables by $\overline{x}$, $\overline{y}$, $\overline{z}, \ldots$, $\overline{t}$ and the absolute values of these increments by the corresponding letters in upright face: x, y, z,… t. The maximum of the absolute value of the function $f(x + \overline{x}, y + \overline{y}, z + \overline{z}, \ldots, t + \overline{t})$ when the increments vary over fixed circles (their arguments vary, but their absolute values remain equal to the constants indicated above) was denoted υ. Then the general case of inequality (9) was written

$$\text{mod.}\mathrm{D}_x^l \mathrm{D}_y^m \ldots \mathrm{D}_t^n f(x, y, z, \ldots, t) < N\frac{\upsilon}{\mathrm{x}^l \mathrm{y}^m \cdots \mathrm{t}^n}, \tag{10}$$

where $N = (1 \cdot 2 \cdots l)(1 \cdots 2 \cdots m) \cdots (1 \cdot 2 \cdots n)$ and υ is an upper bound for $f(x, y, z, \ldots t)$ in a region containing all points $(x+h, y+k, z+l, \ldots, t+r)$ such that $|h| \le \mathrm{x}$, $|k| \le \mathrm{y}$, $|l| \le \mathrm{z}, \ldots$, $|r| \le \mathrm{t}$.

The inequalities (10) gave Cauchy an idea for solving a problem that had occupied his thoughts even before he began developing complex analysis, namely the rigorous development of a theory of differential equations. The eighteenth-century mathematicians had studied a wide variety of ordinary and partial differential equations of interest in physics and geometry. They had performed valuable work in accumulating a stock of standard equations and methods of solving them and had carried out important taxonomic work as well, distinguishing between general, particular, and singular solutions of equations. The preferred method of solving differential equations was by integration (quadratures). However, many equations proved resistant to any manipulation that would allow them to be written as, say, exact differential equations. In such cases, both for ordinary and partial differential equations, series methods were used, as explained in the introduction above. The solution was assumed to be represented as a power series or a trigonometric series with undetermined coefficients. The equation itself and some auxiliary initial or boundary conditions were then applied to determine the coefficients. In order to explain the method of undetermined coefficients and see why it disturbed Cauchy, let us consider a simple first-order ordinary differential equation

$$y' = f(x, y) \tag{11}$$

with the initial condition $y = y_0$ when $x = x_0$, where $f(x, y)$ is an analytic function of its two variables in an open set containing the set $R = \{(x, y) : |x - x_0| \le r,\ |y - y_0| \le \rho\}$. Then

$$f(x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_{mn}(x - x_0)^m (y - y_0)^n \tag{12}$$

for $(x, y) \in R$. The coefficients c_{mn} are given by the usual formulas for a Taylor series:

$$c_{mn} = \frac{1}{m!n!} \frac{\partial^{m+n} f}{\partial x^m \partial y^n}(x_0, y_0). \tag{13}$$

Equation (13) is the crucial link that makes it possible to use Taylor series to solve differential equations such as (11). Suppose, for example, that there exists a solution of this equation satisfying the initial condition $y(x_0) = y_0$, and that the solution is an analytic function given by a Taylor series

$$y = y_0 + \sum_{n=1}^{\infty} y_k (x - x_0)^k. \tag{14}$$

It is easy to see that for each positive integer k there is a polynomial p_k depending on at most $k(k + 3)/2$ variables that expresses y_k in terms of $y_0, \ldots, y_{k-1}$ and the coefficients c_{mn} for which $m + n < k$. For example, by differentiating the relation

$$y'(x) = f\bigl(x, y(x)\bigr) \tag{15}$$

three times with respect to x, we find

$$y^{(iv)} = \frac{\partial^3 f}{\partial x^3} + 3\frac{\partial^3 f}{\partial x^2 \partial y} y' + 3\frac{\partial^3 f}{\partial x \partial y^2}(y')^2 + \frac{\partial^3 f}{\partial y^3}(y')^3 + 3\frac{\partial^2 f}{\partial x \partial y} y'' + 3\frac{\partial^2 f}{\partial y^2} y' y'' + \frac{\partial f}{\partial y} y'''. \tag{16}$$

When we set $x = x_0$ and $y = y_0$ in Eq. (16), we find an expression for $y^{(iv)}(x_0)$, and hence also for y_4, which equals $y^{(iv)}(x_0)/4!$. This expression is given in terms of the coefficients y_0, y_1, y_2, y_3, c_{00}, c_{01}, c_{10}, c_{02}, c_{11}, c_{20}, c_{03}, c_{12}, c_{21}, c_{30}, as follows:[4]

$$y_4 = \frac{1}{4}\bigl[c_{00} + c_{21} y_1 + c_{12} y_1^2 + c_{03} y_1^3 + c_{11} y_2 + 2c_{02} y_1 y_2 + c_{01} y_3\bigr]. \tag{17}$$

We remark for later reference that *all the coefficients in the polynomial* p_k *are nonnegative.*

In order to get the process started the initial point (x_0, y_0) has to be given. That is, the data for the problem consist of the function $f(x, y)$ and this point. An initial-value problem of this kind is called nowadays a Cauchy problem. Cauchy had recognized the fundamental importance of problems of this type a decade earlier, in his lectures at l'École Polytechnique. These lectures were

[4] As the expression (17) shows, the polynomial p_4 is independent of y_0, c_{00}, c_{10}, and c_{20}. This independence holds in general, so that p_k actually depends on only $k(k+1)/2$ variables.

known only at second hand for more than a century and a half, until 1980, when Dr. Christian Gilain found a printed copy of the first thirteen of them. (See Cauchy's *Équations Différentielles Ordinaires*, Johnson Reprint Corporation, 1981.) In the eighth of these lectures Cauchy had proved (by real-variable methods, using approximation by polygonal lines) that under certain conditions there exists a solution to the initial-value problem. He had discussed its uniqueness in Lecture 10. Both existence and uniqueness, of course, are relative to the domain of allowable solutions. However, in the domain of analytic functions, the method of undetermined coefficients itself guarantees that there can be only one solution of a given initial-value problem, since the method determines the coefficients uniquely.

The method of undetermined coefficients is described here in more general terms than any eighteenth-century mathematician seems to have done. Since it is sufficiently explained by giving just a few examples, it does not appear that anyone bothered to explain it in theoretical terms. However, the description given here would undoubtedly be recognizable to an eighteenth-century mathematician.

From a rigorous point of view, this method involved some wishful thinking. It changed the analytic problem of integrating the equation into an algebraic one: solving the equations that the differential equation generated for the coefficients. This step was a great simplification, but it raised three issues that disturbed Cauchy. First, even if the algebraic problem proved to be tractable, so that the coefficients could be computed one at a time (recursively, as we would now say), the convergence of the resulting series was not guaranteed. Second, although the method determined the coefficients of the series, in order to be sure that the resulting series (even assuming it convergent) represented a solution, it was necessary to know that a power series could be differentiated termwise, yielding the power series for the derivative of the function represented by the original series. Those points do not seem to have worried anyone very much before Cauchy. But the third issue was more serious: If the series is to be considered the solution of the differential equation, its partial sums are to be used as approximations to the solution and there was no estimate of the magnitude of the error resulting from truncating the series. By developing complex analysis, Cauchy disposed of the second of these problems. Provided one could verify that the series generated by a differential equation converged, there would be no difficulty with termwise differentiation. The remaining two problems (proving convergence and establishing the rate of convergence), Cauchy realized, might be handled by exploiting his fundamental insight into the behavior of power series, that is, the fact that a power series converges by virtue of being dominated by one particular power series: the geometric series

$$\frac{M}{1-z} = M + Mz + Mz^2 + \cdots Mz^n + \cdots .$$

To fill in the details of this reasoning, Cauchy developed in the mid-1830s a systematic procedure, which we are about to describe.

2.2 Cauchy's *calcul des limites* (method of majorants)

Although the basic idea of Cauchy's method is quite simple and can be explained in a few words, Cauchy's original explanation of it was technically somewhat more cumbersome than necessary. When applying it to partial differential equations in the 1842 paper entitled "Mémoire sur un théorème fondamental, dans le calcul intégral" published in *Comptes rendus*, **XIV** (1842), pp. 1020–1026, he stated the inequality (10) and noted that if the function was given by

$$u = f(x, y, z, \ldots, t) = ax^{-1}y^{-1}z^{-1} \cdots t^{-1}, \tag{18}$$

then

$$\mathrm{D}_x^l \mathrm{D}_y^m \ldots \mathrm{D}_t^n f(x, y, z, \ldots, t) = N \frac{u}{(-x)^l(-y)^m \cdots (-t)^n}, \tag{19}$$

so that an upper bound for the partial derivative occurring in (10) could be obtained by considering the corresponding derivative of the function u and replacing each variable by the negative of the absolute value of the increment of the corresponding variable in (10) and the function u in the numerator by v. From this fact Cauchy reaped a rich harvest, in the form of the following general corollary:

> The unknown function in an ordinary or partial differential equation can be expanded in a series exactly like [those considered here] by use of the formulas of Taylor or Maclaurin. Thus the principles we have just established can be applied to the integration of the equation by series; and, to prove the existence of a general integral in every case, it suffices to integrate the equation in the particular case when the functions that form the right-hand side are inversely proportional to the variables on which they depend, as we shall explain in more detail in later sessions.

The general method is clear from this paper, although its exact scope and the technical details are not. An elegant exposition of the method, roughly along the following lines, was given in 1856 by the French mathematicians Briot and Bouquet in a paper entitled "Recherches sur les propriétés des fonctions définies par les équations différentielles" (*Journal de l'École Polytechnique*, Cahier 36, pp. 133–198). Suppose, when confronted with the differential equation (11), we consider a second equation

$$z' = g(t, z), \tag{20}$$

where the function $g(t, z)$ is analytic in an open domain containing all points (t, z) with $|t| \leq r$ and $|z| \leq \rho$ and has nonnegative Taylor coefficients that majorize those of $f(x, y)$, that is,

$$g(t, z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} g_{mn} t^m z^n. \tag{21}$$

where $|c_{mn}| \leq g_{mn}$ for all m and n. Suppose further that Eq. (20) with the initial condition $z(0) = 0$ can be explicitly shown to have an analytic solution

$z = \sum_{k=1}^{\infty} z_k t^k$ (where this Taylor series converges for some non-zero value of t). The recursion relation that gives, for example, z_4 in terms of z_1, z_2, z_3, g_{01}, g_{12}, g_{21}, and g_{30} is precisely Eq. (17) with y_1, y_2, y_3 replaced by z_1, z_2, z_3 and c_{mn} replaced by g_{mn}. Consequently, it follows quickly by induction that the coefficients z_k are nonnegative. In the general case, because the coefficients of p_k and the coefficients z_k are all nonnegative, it follows by induction that replacing the coefficients $y_1, \ldots, y_{k-1}$ by $z_1, \ldots, z_{k-1}$ and the coefficient c_{mn} by g_{mn} results in a value for z_k that majorizes $|y_k|$. Since the series for z converges, it follows that the series for y also converges. Hence there really is an analytic solution of Eq. (11).

The trick needed to make the method work is a suitable choice of the majorant $g(t,z)$, and that is where Cauchy's inequality (9) comes in. This inequality shows that a suitable majorizing series is obtained by simply multiplying the two sides of the inequality by $t^m z^n$ and dividing out the two factorials. The resulting majorant is a geometric series whose sum can be explicitly computed as

$$g(t,z) = \frac{M}{\left(1 - \frac{t}{r}\right)\left(1 - \frac{z}{\rho}\right)}, \tag{22}$$

where $M = M(r, \rho)$. But with this choice of $g(t,z)$ the variables are separable in the differential equation (20), which has the explicit solution

$$z(t) = \rho\left[1 - \sqrt{1 + \frac{2Mr}{\rho}\ln\left(1 - \frac{t}{r}\right)}\right]. \tag{23}$$

The function $z(t)$ in Eq. (23) is analytic for t sufficiently close to 0, in fact for $|t| < r(1 - e^{-\rho/2Mr})$. The radius of convergence of the series generated by the original equation (11) must therefore be at least this large, and within that radius the convergence of the series generated by the original equation must be at least as rapid as the convergence of the series for this function. In particular, the truncation error is no greater for the original series than for the Taylor series of the function $z(t)$.

Cauchy's method is nowadays called the *method of majorants* in English. Cauchy himself called it the *calcul des limites* (calculus of bounds). He introduced it in a memoir to the Turin Academy in 1831 summarizing a paper that he published in lithographed form and bearing the title "Mémoire sur les rapports qui existent entre le calcul de residus et le calcul des limites". In a paper written in Prague in 1835 and bearing the title "Mémoire sur l'intégration des équations differentielles" Cauchy gave the application just discussed. He considered a system of n equations of the type (11), that is:

$$\frac{dy_i}{dx} = f_i(x, y_1, \ldots, y_n), \quad (i = 1, \ldots, n). \tag{24}$$

Besides the obvious merit that it gives a rigorous justification for the method of undetermined coefficients, Cauchy method of majorants had the additional advantage that it moved the theory of differential equations into the complex domain. Cauchy's estimates for the derivatives, on which the method is based, cannot be obtained in the real domain, since the constant $M(\rho)$ is the maximum value of the function on a circle in the complex plane. After obtaining Cauchy's existence theorem, one could of course restrict attention to real values of the independent variable, but doing so would only jettison a huge amount of information. The fact that complex function theory was needed to obtain the theorem is a good sign that the natural domain of existence of the solution is in the complex domain. Cauchy put the matter in clear perspective in the paper "Mémoire sur un théorème fondamental, dans le calcul intégral" cited above:

> In the theory of equations mathematicians have properly considered fundamental the question whether every equation has a root. Similarly in the integral calculus one of the most important questions, a fundamental question, is obviously whether every ordinary or partial differential equation can be integrated. But—and this is probably surprising—despite the numerous works of mathematicians on the integral calculus, this question, important as it is, is nowhere solved in full generality. To be sure, the existence of general integrals of ordinary differential equations, which contain only one independent variable, is now established by two different methods that I have given, the first in my lectures at l'École Polytechnique, the second in a lithographed memoir of 1835.

In this paper Cauchy set out the analogy between the role played by complex numbers in the theory of algebraic equations and the role to be played by functions of a complex variable in the theory of differential equations. It took some time for him to realize the gulf that exists between real derivatives and complex derivatives, and this confusion sometimes shows through in the hypotheses he imposes in order to get his theorems. In the present case Cauchy's emphasis was on exhibiting, one way or another, a solution to a class of differential equations, and the way to do that was, in this instance, to define the solution by a power series. Analyticity was useful because it provided a proof of existence, not because Cauchy particularly wanted only analytic solutions.

2.3 Weierstrass' early work

The existence of the majorant was deduced from Cauchy's inequalities for the derivatives, and these in turn were consequences of the Cauchy integral theorem, which gave an integral representation for the coefficients in the Taylor series of an analytic function. It would appear, therefore, that this method is very much dependent on the Cauchy approach to analytic function theory. For that reason it is somewhat surprising that Weierstrass, whose approach to complex analysis did not use the Cauchy integral, was able to use essentially the same method without referring to any complex integrals. The way to avoid integrals is to note that domination of a convergent power series by a geometric series can be deduced from properties of series alone, and it is not necessary to know the exact value of the bounding constant. Thus, given a *convergent* series

$$\sum_{n=0}^{\infty} a_n z_0^n, \tag{25}$$

where z_0 is a complex number, it follows that the terms are bounded, that is, there exists a positive number M such that $|a_n z_0^n| \leq M$ for all n. Then if z is any complex number whose absolute value is less than the absolute value of z_0, we have

$$\sum_{n=0}^{\infty} |a_n z^n| = \sum_{n=0}^{\infty} |a_n z_0^n| \left|\frac{z}{z_0}\right|^n \leq \sum_{n=0}^{\infty} M r^n = \frac{M}{1-r}, \tag{26}$$

where $r = \left|\frac{z}{z_0}\right|$, so that $0 \leq r < 1$. In other words, if the power series (25) converges, it has a majorant $\sum c_n z^n$ with positive coefficients c_n satisfying $|a_n| \leq c_n$, namely $c_n = \frac{M}{|z_0|^n}$, yielding the majorant function $g(z) = \frac{M}{1-\frac{z}{|z_0|}}$. It is not necessary to know that M is the maximal modulus of the function represented by the original series. In his first work, however, Weierstrass didn't phrase the argument quite this way, as we shall now see.

Although Cauchy referred to these papers on many later occasions and even included the result in his 1841 *Exercices de mathématiques*, they escaped the notice of Weierstrass, who in 1842 considered a system of ordinary differential equations formally less general than (24), but really of equal generality:

$$\frac{dx_j}{dt} = G_j(x_1, \ldots, x_n), \quad (j = 1, 2, \ldots, n). \tag{27}$$

The system (24) can be made into a system of $n+1$ equations of the form (27) by adjoining the equation $dy_0/dx = 1$ and the initial condition $y_0(x_0) = x_0$ and then replacing x with y_0 in all the functions on the right-hand side. Then, except for the trivial difference in the letters used, the two equations are of the same form.) Even though Weierstrass' system was formally equivalent to Cauchy's, he assumed that the functions $G_j(x_1, \ldots, x_n)$ were polynomials whose coefficients were analytic functions of other unspecified variables $u_1, \ldots, u_m$. Thus neither of the two mathematicians, strictly speaking, generalized the other's work. The reason why the functions G_j had to be polynomials is connected with Weierstrass' choice of majorant. As a majorant for any polynomial in the n variables Weierstrass chose the function $G(x_1, \ldots, x_n) = g(1 + x_1 + \cdots x_n)^m$. If the exponent m equals the highest-degree term that occurs among the polynomials on the right-hand side of (27), then G will have positive coefficients everywhere that any of the polynomials have nonzero coefficients. If g is then taken sufficiently large, the coefficients of G will majorize those of all the polynomials on the right-hand side of (27). Thus, Weierstrass was replacing this system with a very simple system, in which the right-hand side was the same in all equations. Although he did not spell out the details, they are easy to fill in. Because all the variables $x_1, \ldots, x_n$ have the same derivative in the new system, any two of them differ by a constant. One therefore needs to solve only for one of them, say x_1. For x_1 the equation for the majorant is

$$\frac{dx_1}{dt} = g(z + nx_1)^m, \tag{28}$$

whose solution is $x_1 = -\frac{a}{n} + Be^{gnt}$ if $m = 1$ and $x_1 = \frac{(C+gn(1-m)t)^{1/(1-m)}-a}{n}$ otherwise. The first of these is an entire function, and the second is analytic for t near zero if $C \neq 0$.

Weierstrass' point of view was different from Cauchy's, but similar to that of Briot and Bouquet in the work cited above, as shown by his title: "Definition analytischer Functionen einer Veränderlichen vermittlest algebraischer Differentialgleichungen." That is, he was less concerned with proving the existence and computability of the solution than in showing that the equation (or system of equations) really did unambiguously define locally an analytic function. In particular, he took the trouble to discuss the analytic continuation of the solution. For that purpose it was crucial to prove that the solution depended analytically on the initial value.

The idea of analytic continuation is to take a point near the circle of convergence of a power series and construct the Taylor series about that point. With luck, the new circle of convergence will extend outside the original circle, continuing the analytic function beyond its original domain. When the Taylor series is being generated by a differential equation, the new Taylor coefficients can be computed algebraically, as shown above, but the new initial values used to get the recursion started have to be computed from the first series. Thus, like Cauchy, Weierstrass recognized the central importance of the initial-value problem for ordinary differential equations. Weierstrass did not publish his work at the time. If he had, he would undoubtedly have been told of Cauchy's earlier work and would have been on the alert for Cauchy's simultaneous work in applying the same method to partial differential equations, which will be described below. In that case the human side of the rest of our story might have been less interesting than it turned out to be. As it happened, Weierstrass' paper was published only in his collected works, which began to appear in 1894, and he remained ignorant of Cauchy's work until 1875.

Weierstrass did publish a theorem of this type in the *Journal für die reine und angewandte Mathematik*, in 1856 (Band 51, pp. 1–60), in a paper entitled "Zur Theorie der analytischen Facultäten." (This theorem was omitted from the version of this paper published in Vol. 1 of his collected works.) Here is how he presented it:

> It should indeed be remarked that the preceding theorems do not always have converses, so that one might, for example, claim that the series for $f(x, y, \ldots)$ converges *only* for the values of $x, y, \ldots$ for which convergence occurs for the series that express $\varphi, \varphi_1, \ldots$. But they do not by any means give the true criteria by which one can determine the convergence of power series in one or several variables, even though they can be used in many investigations. These criteria

must rather be derived from some other source, of a type that I plan to exhibit on another occasion. At this juncture I permit myself only to add the following general theorem, as the principal result of my investigations:

> *Let $x_1, x_2, \ldots, x_n$ be defined as functions of a single independent variable x by means of a system of n differential equations*
>
> $$\begin{aligned} F\Big(x_1, x_2, \ldots, \frac{dx_1}{dx}, \frac{dx_2}{dx}, \ldots, \frac{d^2x_1}{dx}, \frac{d^2x_2}{dx} \ldots\Big) &= 0, \\ F_1\Big(x_1, x_2, \ldots, \frac{dx_1}{dx}, \frac{dx_2}{dx}, \ldots, \frac{d^2x_1}{dx}, \frac{d^2x_2}{dx} \ldots\Big) &= 0, \end{aligned}$$
>
> *and so forth, where F, F_1,... are entire functions of $x_1, x_2, \ldots$ whose derivatives up to any order can be written; and let $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ be infinite series of the form $\sum a_\alpha x^\alpha$, $\alpha = 0 \ldots \infty$, which satisfy the differential equations formally when substituted for $x_1, x_2, \ldots, x_n$. Then these series all converge unconditionally, perhaps not for all values of x, but at least for those whose absolute values do not exceed a certain limit g. They then represent continuous functions of x that actually do satisfy the differential equations. The limiting value g is the absolute value of a value of x, in any neighborhood of which one or more of the functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ or one of their derivatives becomes infinite.*[5]

If to this theorem, which finds application to all algebraic differential equations and remains valid with certain modifications when any single-valued functions of x_1, x_2,..., $\frac{dx_1}{dx}$, $\frac{dx_2}{dx}$,..., $\frac{d^2x_1}{dx^2}$, $\frac{d^2x_2}{dx^2}$,... that do not become infinite for any of these values are used in place of F_1, F_2,..., we now add a second theorem giving the conditions under which the series in question converge for values of x whose absolute value equals g, we obtain a way of finding the limits of convergence of an infinite series of the given form in many cases, even *before* they are actually represented, by means of a closer inspection of the character of the quantity to be developed. And this is especially important since one can then use any valid method of determining the coefficients, without having to determine the conditions for convergence afterwards using the law of formation of the series, which is a procedure that can be carried out only for the simplest series without great complications.

[5] This second phrase is puzzling: the derivative of a function that is analytic and single-valued in a deleted neighborhood of a point cannot have a pole at that point unless the function itself does. Perhaps Weierstrass has multivalued functions such as $\log z$ in mind.

The justification of this last assertion should be clear from the example of the wave equation given above, one of the simplest imaginable partial differential equations, yet for which the proof that an analytic solution exists is still rather messy. It is to be noted that Weierstrass regarded the differential equation as the definition of an analytic function. In fact, Weierstrass never taught a course in differential equations per se. For him they were apparently thought of as an application of analytic function theory.

2.4 Briot and Bouquet

Weierstrass' point of view was exactly that of Briot and Bouquet in the work mentioned above, as its title shows. The work of Briot and Bouquet was a continuation and application of analytic function theory, which they had expounded in an earlier paper in the same *Cahier* (No. 36). They considered a large variety of ordinary differential equations, extending the simple theorem for equations of the form (11) to cases in which the derivative has a pole. It is precisely these singular cases that, we now know, have particular interest. However, they made no mention of partial differential equations or of Cauchy's 1842 work. At least as far as the German mathematicians were concerned, these other works of Cauchy remained buried in the literature, even though the work of Briot and Bouquet was well known.

2.5 Jacobi

The works of Jacobi were enormously influential during the nineteenth century, and rightly so, due to their profundity and the clarity of Jacobi's exposition. Weierstrass, who, along with Borchardt and Clebsch helped to edit Jacobi's posthumous works, seems to have been particularly influenced by his analytic style. Indeed, it was Jacobi's formulation of the inversion problem for hyperelliptic integrals and his discovery of theta functions in connection with elliptic integrals that ultimately led Weierstrass to the brilliant paper on Abelian integrals that secured his reputation and gained him an invitation to Berlin. Among the posthumous works of Jacobi were several on the solution of differential equations, published in the *Journal für die reine und ungewandte Mathematik*, in which the concept of a "normal form" for a system of differential equations is given. In one such paper, entitled "De aequationum differentialium systemate non normali ad formam normalem revocando" (C.G.J. Jacobi, *Gesammelte Werke*, Chelsea Publishing Company, New York, 1969, Vol. 5, 483–513) Jacobi reflects on his earlier experience in solving such equations, saying:

> In my paper "Theoria novi multiplicatoris etc." I determined the multipliers of the isoperimetric differential equations, that is, the equations that occur in those isoperimetric problems in which the variation of a given integral containing one independent variable and several dependent ones is set equal to zero. I showed that the determination of these multipliers is beset with major difficulties if

> the highest-order derivatives of the dependent variables occurring in the integral are not all of the same order. For in that case the system of isoperimetric differential equations is not in a form for which the highest-order derivatives of the individual dependent variables can be obtained as unknowns whose values are determined by the differential equations themselves.

In other words, the equations cannot be solved for the highest-order derivatives of the dependent variables. A simple example will exhibit the phenomenon Jacobi is referring to. Consider the system of two differential equations in two unknowns

$$x' - y' = 0, \quad x''y'' = 0, \tag{29}$$

which appears to be of order 2 in both unknowns. One would therefore expect to find four arbitrary constants in its general solution. However, the explicit solution is easy to find. The first equation implies that $x'' = y''$. When this equation is substituted in the second, one finds $x'' = 0 = y''$, so that both x and y are linear functions of the independent variable t: $x = at + b$, $y = ct + d$. It appears that the four arbitrary constants are a, b, c, and d. However, the first differential equation now implies that $a = c$, so that there are really only three arbitrary constants in the general solution. By changing variables and writing $x = u + v$, $y = u - v$, one can convert this system into the equivalent system

$$v' = 0, \quad u'' = 0, \tag{30}$$

which is obviously of order 2 in u and order 1 in v, so that we should expect three arbitrary constants.

In this paper Jacobi considers a system of differential equations

$$u_1 = 0,\ \ u_2 = 0,\ \ ,\ldots, u_n = 0, \tag{31}$$

where each function u_i depends on variables $x_1, \ldots, x_m$ and their derivatives with respect to a single independent variable t and shows by combinatorial work on what we would now call a matrix of order m (he called it a square scheme of m^2 quantities) how to replace the system (31) by an equivalent system that is solved for the highest-order derivatives of the variables, that is, a system of the form

$$x_1^{(a_{n_1,1})} = X_1,\ \ x_2^{(a_{n_2,2})} = X_2, \ldots, x_m^{(a_{n_m,m})} = X_m, \tag{32}$$

where the left-hand sides are the highest-order derivatives of those variables (the right-hand sides contain only derivatives of lower order). This concept of *normal form* turned out to be crucial in bringing order to the Cauchy problem for both ordinary and partial differential equations.

If we are to judge by the number of times this concept occurs in his posthumous works, Jacobi must have thought about the problem fairly often. In another paper, entitled "De investigando ordine systematis aequationum differentialium vulgarium cujuscunque," (*Journal für die reine und angewandte Mathematik*, **64**, 1865, 297–320; *Gesammelte Werke*, Vol. 5, 191–216) Jacobi gave the full proofs of the results illustrated by examples in the abovementioned paper. He also gave an explanation of what normal form is, calling it *canonical form* on this occasion:

> A system of ordinary differential equations is *non-canonical* if the equations contain the highest-order derivatives of the dependent variables in such a way that their values cannot be obtained. [We would say that the equations cannot be solved for the highest-order derivatives.] *In that case the number of arbitrary constants in the general solution, otherwise known as the order of the system, is less than the sum of the highest orders of the derivatives in the system.*

In a footnote the editor (Borchardt) remarks, "A system that is here called *canonical* or *of canonical form*, is the same that is said to be *of normal form* in the paper 'Theoria novi multiplicatoris,' but differs considerably from that to which Jacobi gave the name canonical in his paper 'Nova methodus, aequationes differentiales partiales primi ordinis inter numerum variabilium quemcunque propositas integrandi'." And indeed, in that paper, which occupied pages 1 through 181 of Band 60 of the *Journal für die reine und angewandte Mathematik*, and pages 3 through 189 of Volume 5 of the *Gesammelte Werke*, the name *canonical* is given to a system of equations of the form now known to physicists as Hamilton's equations of motion.

In yet a third of these posthumous works, entitled "De aequationum differentialium isoperimetricarum transformationibus earumque reductione ad aequationem differentialem partialem primi ordinis non linearem" (*Gesammelte Werke*, Vol. 5, 465–482) Jacobi reduces the Euler equations for a variational problem to a system of equations, remarking, "And these equations have quasi-canonical form, in that the right-hand sides involve only derivatives of lower order than the derivatives that constitute the left-hand sides." As he mentioned in the passage quoted earlier, Jacobi had considered the same problem in Section 31 of his long paper on multipliers ("Theoria novi multiplicatoris systematis æquationum differentialium vulgarium applicandi," *Journal für die reine und angewandte Mathematik*, **29**, 333–376; a continuation of earlier work, all reproduced in *Gesammelte Werke*, Vol. 4, 317–509). Thus, by repeated emphasis, Jacobi makes it clear that the secret of bringing order to a system of ordinary differential equations is to solve explicitly for the highest-order derivatives of the dependent variables. Now Jacobi also worked frequently with partial differential equations, although he does not seem to have used the method of undetermined coefficients to do so. But in his *Vorlesungen über Dynamik* (Volume 8 of his *Gesammelte Werke*) he once again encountered the need for a canonical form, this time for a partial differential equation. He first showed how to reduce a general first-order partial differential equation to a form in which the independent variable does not appear, so that it is necessary to consider only an equation of the form

$$\varphi\Big(\frac{\partial V}{\partial q_1}, \frac{\partial V}{\partial q_2}, \dots, \frac{\partial V}{\partial q_n}, q_1, q_2, \dots, q_n\Big) = 0. \tag{33}$$

By setting $\frac{\partial V}{\partial q_i} = p_i$, as Lagrange had once done in the case of two variables, one gets an exact differential $p_1\,dq_1 + p_2\,dq_2 + \cdots + p_n\,dq_n$, which can then be integrated to recover V. The problem, then, is to determine n functions p_1, p_2,..., p_n of the variables q_1, q_2,..., q_n in such a way that the equation

$$\varphi(p_1, \ldots, p_n, q_1, \ldots, q_n) = 0 \tag{34}$$

holds and (simultaneously) the differential $p_1\,dq_1 + p_2\,dq_2 + \cdots + p_n\,dq_n$ is exact. Actually the distinction between closed and exact differentials was not of concern to Jacobi, who regarded the problem as being a matter of getting Eq. (34) satisfied simultaneously with the $\frac{n(n-1)}{2}$ conditions for a closed differential form, namely

$$\frac{\partial p_i}{\partial q_j} = \frac{\partial p_j}{\partial q_i}, \quad (1 \le i < j \le n). \tag{35}$$

When $n = 2$, this problem thus gives two equations to determine two functions of two variables. Lagrange had considered this case. When n is greater than 2, the problem is more complicated. As Jacobi commented,

> Up to now this difficulty has kept analysts from extending the Lagrange method to a larger number of variables. We shall not be intimidated by this fact. On the contrary, since we know a priori that, even though the problem appears to be overdetermined, it nevertheless has a solution, we shall investigate all the more how it happens that one can satisfy $\frac{n(n-1)}{2}$ conditions using $n-1$ functions.

This investigation led Jacobi once more to a system of ordinary differential equations, written out in detail in an appendix by Clebsch (Jacobi's *Gesammelte Werke*, Vol. 8, p. 296) when this work was published posthumously, in which each equation was solved for the highest-order derivative of a function. Anyone who read Jacobi's works would naturally think in terms of a similar hypothesis when trying to solve partial differential equations. Kovalevskaya, in particular, recognized this principle (and referred to Jacobi's work in her paper).

3 Partial Differential Equations

While Weierstrass was independently constructing the theory of analytic solutions of ordinary differential equations, in ignorance of Cauchy's work, Cauchy himself was busy extending it to partial differential equations. Although the method possesses great generality in application to ordinary differential equations, it is by no means a trivial task to judge the extent to which it can be applied to partial differential equations. Initial-value problems for such equations are not always the most natural. For example, for Laplace's equation in two variables,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{36}$$

a boundary condition is both more natural and more useful than initial conditions would be. Indeed, Hadamard showed that solutions of this equation do not depend stably on initial values. If the initial conditions are $u_n = 0$, $\frac{\partial u_n}{\partial y} = e^{-\sqrt{n}} \sin nx$ when $y = 0$, the solution is unique, namely

$$u_n(x, y) = \frac{1}{n} e^{-\sqrt{n}} \sinh ny \sin nx.$$

(See Hadamard, J. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Dover, New York, 1952, p. 33). But the initial values both tend to zero uniformly in x as n tends to infinity, while $u_n(0, y)$ tends to infinity for any nonzero y.

Similarly, as we saw above in the case of the wave equation (a hyperbolic equation), an initial-value problem may work well with the method of undetermined coefficients when a boundary-value problem does not. The problem is so complicated that it was necessary to consider differential equations of special forms in order to get started proving theorems. Cauchy considered one such special form in a paper entitled "Mémoire sur l'emploi du calcul des limites dans l'intégration des équations aux dérivées partielles" (*Comptes rendus*, **XV**, 1842, pp. 44–59), which was the promised continuation of the paper on the method of majorants discussed above. In this paper he considered a quasi-linear equation (which he called "linear")

$$D_t \varpi = A D_x \varpi + B D_y \varpi + \cdots + K, \tag{37}$$

where the coefficients $A, B, \ldots, K$ depend on the independent variables $t, x, y, \ldots$ and also on the unknown function ϖ. (It is the dependence on ϖ that makes Eq. (37) quasi-linear, rather than linear.) Cauchy demonstrated that this equation has an analytic solution.

3.1 Kovalevskaya's dissertation

Of the results that Kovalevskaya achieved while a student of Weierstrass from 1870 to 1874, the most significant by far was her proof that a system of partial differential equations with analytic data has (under certain hypotheses) an analytic solution. Fortunately certain paragraphs in letters written by Weierstrass during this time enable us to trace the chronology of her work in some detail. To orient ourselves, we must exclude knowledge of the work of Cauchy just described, since neither Weierstrass nor Kovalevskaya was aware of it. As far as Kovalevskaya knows in 1873, she is just a 23-year-old graduate student beginning a dissertation by trying to show that analytic functions of several complex variables can be defined through partial differential equations, and no one has proved anything in this area up to now. Where should she begin? Obviously by examining a number of examples of analytic functions defined in this way and seeing how the series converge. That is what she did, and the first result she achieved was a counterexample! That is, she discovered a partial differential equation (lost to history) that generated a multiple power series that does not converge in any region. Very soon she was able to exhibit such an example for the heat equation, as described below. This result appeared to her to be rather discouraging, since it meant that new hypotheses would be needed to prove any results in general. Had she known of Cauchy's work, of course, she might have been less discouraged, but in that case she might not have been working on the topic at all. Weierstrass, however, found this result both profound and encouraging, and urged her to look for the required hypotheses. No doubt, it was the need for such hypotheses that led her to look closely at the work of Jacobi described above, with its many references to "normal form" for a system of ordinary differential equations. Jacobi had given an indication of what "normal form" meant for partial differential equations. But was this concept a sufficient condition for the existence of analytic solutions? That was the question she answered, and which became the main part of her dissertation. Let us now describe it.

In the paper that she published on this topic after receiving her degree from Göttingen ("Zur Theorie der partiellen Differentialgleichungen," *Journal für die reine und angewandte Mathematik*, **80**, 1–32) Kovalevskaya summarized the earlier work of Briot and Bouquet, Weierstrass, and Jacobi on ordinary differential equations, and then stated her goal, as follows:

> I have taken these theorems in the formulation above from the lectures of my esteemed teacher Herr Weierstrass.[6] In the present work I undertake to study the question of whether and to what extent these theorems can be extended to the case when algebraic partial differential equations are given for the determination of analytic functions of several variables.

Kovalevskaya began with a fairly "clean" quasi-linear system of n partial differential equations in independent variables $x, x_1, \ldots, x_r$, which she wrote in the form

$$\begin{aligned} \frac{\partial \varphi_1}{\partial x} &= \sum_{\beta=1}^{n} G_{1\beta}^{(1)}(\varphi_1, \ldots, \varphi_n)\frac{\partial \varphi_\beta}{\partial x_1} + \cdots + \sum_{\beta=1}^{n} G_{r\beta}^{(1)}(\varphi_1, \ldots, \varphi_n)\frac{\partial \varphi_\beta}{\partial x_r}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (38) \\ \frac{\partial \varphi_n}{\partial x} &= \sum_{\beta=1}^{n} G_{1\beta}^{(n)}(\varphi_1, \ldots, \varphi_n)\frac{\partial \varphi_\beta}{\partial x_1} + \cdots + \sum_{\beta=1}^{n} G_{r\beta}^{(n)}(\varphi_1, \ldots, \varphi_n)\frac{\partial \varphi_\beta}{\partial x_r}. \end{aligned}$$

Here the coefficients $G_{\alpha\beta}^{(\gamma)}$ were assumed to be analytic functions of n variables in a certain domain in what is now called $\mathbb{C}^n$, the vector space of n complex variables (Kovalevskaya referred to it simplye as "a certain region"). Except for the relatively unimportant generalization from a single equation to a system of

[6]It would be interesting to know what lectures she was referring to. As mentioned above, Weierstrass did not ever give a course bearing the title "Differentialgleichungen." For him all these theorems seem to have been a part of analytic function theory.

more than one, this problem was exactly the problem Cauchy had considered three decades earlier. Indeed, the derivative-free term K in Eq. (37) makes Cauchy's result appear to be the more general of the two.

Kovalevskaya showed that, given n "initial" functions of the last r independent variables $x_1, \ldots, x_r$, which she wrote as

$$\varphi(x_1, \ldots, x_r)_{10}, \ldots, \varphi(x_1, \ldots, x_r)_{n0} \tag{39}$$

and which were all assumed to be analytic in some region of (what is now called) $\mathbb{C}^r$, all vanishing at the origin, there exist n formal power series $\varphi_1, \ldots, \varphi_n$ in the $r+1$ variables $x, x_1, \ldots, x_r$ that satisfy the system (38) formally (that is, when all the operations on the two sides of the equations are performed termwise, the two sides are power series having exactly the same coefficients) and which reduce to the n Maclaurin series (39) in (what is now called) the complex hyperplane $x = 0$. And, she said,

> All n of these series converge unconditionally in some region [of $\mathbb{C}^{r+1}$] and represent functions that actually do satisfy the differential equations [(38)].

There is no need to repeat the technical details of Kovalevskaya's proof. It suffices to say that she adapted Weierstrass' majorant $g(1 + x_1 + \cdots + x_n)^m$ for use with infinite series instead of polynomials, getting a majorant of the form

$$\frac{G}{1 - \frac{\psi_1 + \psi_2 + \cdots + \psi_n}{g}}, \tag{40}$$

where the analytic functions $\psi_1, \ldots, \psi_n$ were to be chosen having nonnegative Maclaurin coefficients at least as large as the corresponding coefficients in $\varphi(x_1, \ldots, x_r)_{10}, \ldots, \varphi(x_1, \ldots, x_r)_{n0}$ respectively. This condition was satisfied by taking them all equal to

$$\frac{g'(x_1 + x_2 + \cdots + x_r)}{1 - \frac{x_1 + x_2 + \cdots + x_r}{\rho}} \tag{41}$$

for suitable positive constants g' and ρ, the extra factor $(x_1 + x_2 + \cdots x_r)$ in the numerator being introduced to assure that all these functions vanish at the origin.

Although this result would not have been seen as a large advance on Cauchy's work, had anyone who looked at it remembered the latter, Kovalevskaya proceeded from this point into genuinely new territory. First, she showed how to introduce extra variables so as to obtain the same conclusion for the apparently more general system

$$\begin{array}{rcl}
G^{(1)}(\varphi_1, \ldots, \varphi_n)\dfrac{\partial \varphi_1}{\partial x} & = & \sum G^{(1)}_{\nu\beta}(\varphi_1, \ldots, \varphi_n)\dfrac{\partial \varphi_\beta}{\partial x_\nu} + G^{(1)}_{00}(\varphi_1, \ldots, \varphi_n), \\
\cdots\cdots & \cdots & \cdots\cdots\cdots\cdots \\
G^{(n)}(\varphi_n, \ldots, \varphi_n)\dfrac{\partial \varphi_1}{\partial x} & = & \sum G^{(n)}_{\nu\beta}(\varphi_1, \ldots, \varphi_n)\dfrac{\partial \varphi_\beta}{\partial x_\nu} + G^{(n)}_{00}(\varphi_1, \ldots, \varphi_n), \\
\nu = 1, \ldots, r, & & \beta = 1, \ldots, n.
\end{array} \tag{42}$$

She then considered an algebraic partial differential equation of order n

$$G\Big(x, x_1, \ldots, x_r, \varphi, \ldots, \frac{\partial^{\alpha+\alpha_1+\cdots+\alpha_r}\varphi}{\partial x^{\alpha}\partial x_1^{\alpha_1}\cdots\partial x_r^{\alpha_r}}\Big) = 0, \tag{43}$$

where the function G is a polynomial in its arguments, which include the $r+1$ variables $x, x_1, \ldots, x_r$, the dependent variable φ, and its partial derivatives $\frac{\partial^{\alpha+\alpha_1+\cdots+\alpha_r}\varphi}{\partial x^{\alpha}\partial x_1^{\alpha_1}\cdots\partial x_r^{\alpha_r}}$. She found the key to extending of all the preceding work to partial differential equations when she formulated the proper notion of Jacobi's normal form. In her words,

> I begin with the case that I shall call *normal* when at least one of the derivatives
> $$\frac{\partial^n\varphi}{\partial x^n}, \ \frac{\partial^n\varphi}{\partial x_1^n}, \ldots, \frac{\partial^n\varphi}{\partial x_r^n}$$
> say $\frac{\partial^n\varphi}{\partial x^n}$, actually occurs in G.

The phrase "actually occurs" simply means that the derivative of G with respect to the argument containing this variable is not identically zero, which is *almost* the same thing as saying that the equation $G = 0$ can be solved for this variable near the initial point. Obviously the variable x was left unsubscripted precisely to make the notation simpler, since it was to be distinguished from the others as the variable that "actually occurs."

Later in the paper Kovalevskaya referred to equations in normal form with respect to a particular variable, which is a clearer use of the term. The complication that can arise can easily be seen in the case of the function $G(x, y) = x^2 + y^2 - 1$. Although the variable y "actually occurs" in this function, the equation $G(x, y) = 0$ cannot be solved for y near the points $x = \pm 1$, $y = 0$, and the ordinary differential equation $G(x, x') = 0$, whose general solution is $x = \sin(t+C)$, also has the constant solutions $x = \pm 1$, which cannot be obtained by assigning a value to the parameter of the general solution. Solutions of this kind, called *singular solutions*, had been known for a century or more. Such solutions could be obtained by solving the simultaneous equations $G = 0 = G'$, where G' was Kovalevskaya's notation for the derivative of G with respect to the argument containing $\frac{\partial^n\varphi}{\partial x^n}$. In the present illustration, that means eliminating y between the equations $x^2 + y^2 - 1 = 0$ and $2y = 0$, leading to the equation $x^2 = 1$. Kovalevskaya pointed out that the result of doing this is either an algebraic equation for the solution (as in this case) or a differential equation for which the solution is not singular. An eighteenth-century way of finding singular solutions, still taught nowadays, was to write the general solution as an equation containing an arbitrary parameter, differentiate with respect to the parameter, and then eliminate the parameter between the two equations. In the example given here, that requires eliminating C between the equations and $x = \sin(t + C)$ and $0 = \cos(t + C)$. Obviously the result of doing so is, once again, the equation $x^2 = 1$. The topic of singularities for differential equations was soon to become a major area of study, but it had not yet reached that state, and Kovalevskaya referred to it only in passing.

Kovalevskaya's theorem is of great importance, since it applies to nearly all of the partial differential equations of mathematical physics: the heat equation, the wave equation, Laplace's equation, the Maxwell equations, and many others. The key to verifying that analytic solutions exist for initial-value problems is to solve for a pure derivative of maximal order, that is, put the equation in normal form. In the final section of her paper Kovalevskaya proved the natural extension of this theorem for a system of n such equations in n dependent variables, where the ith equation is in normal form with respect to the ith dependent variable. Before doing so, however, she gave an extended discussion of the meaning of normal form.

The introduction of the concept of normal form was a significant step forward in determining what could and could not be said in relation to analytic solutions of partial differential equations. Jacobi had realized that when a system of ordinary differential equations is not in normal form, its general solution contains fewer arbitrary constants than the degrees of the individual equations would lead one to expect; his method of reducing such a system to normal form was designed to bring out the actual order of the system. Kovalevskaya showed how to bring a partial differential equation in non-normal form to normal form via a "linear" (what we now call affine) change of variables. In that connection she exhibited the example that had led her to seek additional hypotheses. This example was an equation for which some initial-value problems do not have an analytic solution. To be specific, she considered the well-studied equation of heat conduction

$$\frac{\partial\varphi}{\partial x} = \frac{\partial^2\varphi}{\partial y^2} \tag{44}$$

with any prescribed initial function $\varphi(x, y) = \varphi_0(y)$ when $x = a$. (Actually she wrote $\varphi(y|b)$, since she was interested in power series expansions about a general value (a, b).) She pointed out that the series

$$\sum_{\nu=0}^{\infty} \frac{d^{2\nu}\varphi_0(y|b)}{dy^{2\nu}} \frac{(x-a)^\nu}{\nu!} \tag{45}$$

formally satisfies the differential equation (44) and reduces to $\varphi_0(y|b)$ at $x = a$. However, the simple choice of

$$\varphi_0(y|0) = \frac{1}{1-y} \tag{46}$$

leads to the double power series for $\varphi(x, y)$ about the point $(0, 0)$ given by

$$\sum_{\nu=0}^{\infty} \frac{(2\nu)!}{\nu!} \frac{x^\nu}{(1-y)^{2\nu+1}}, \tag{47}$$

and this series converges only for $x = 0$. She went on to show that this series will diverge whenever the initial function $\varphi_0(y)$ has a singularity anywhere in the plane. If this function is an entire function, the series for $\varphi(x, y)$ *may* have a nontrivial region of convergence, as in the case of $\varphi_0(y) = e^{-y^2}$, but Kovalevskaya pointed out that even in this case such a region was not guaranteed,[7] as she showed by the example

$$\varphi_0(y|b) = \sum_{\nu=0}^{\infty} \frac{(y-b)^\nu}{(\nu!)^{1/3}}. \tag{48}$$

This initial function, with $b = 0$, generates the double series

$$\varphi(x,y) = \sum_{\mu,\nu=0}^{\infty} \frac{\big((2\mu+\nu)!\big)^{2/3}}{\mu!\nu!} x^\mu y^\nu, \tag{49}$$

which has positive coefficients. But one can easily see that this series cannot converge for any non-zero value of x. In fact, even in the best-possible case when $y = 0$, the series becomes

$$\sum_{\mu=0}^{\infty} \frac{\big((2\mu)!\big)^{2/3}}{\mu!} x^\mu, \tag{50}$$

whose radius of convergence the ratio test reveals to be zero.

The heat equation on which this counterexample is based was of course studied extensively by Fourier in his *Analytical Theory of Heat* (Sect. 374), where he gave the equivalent solutions

$$v = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-q^2} f(x + 2q\sqrt{t})\, dq \quad \text{and} \quad v = \int_{-\infty}^{+\infty} f(\alpha) \frac{Ae^{-\frac{(x-\alpha)^2}{4t}}}{\sqrt{t}}\, d\alpha$$

of the equation that he wrote as $\frac{dv}{dt} = \frac{d^2v}{dx^2}$. Guided by physical intuition, Fourier said the following:

> Since the function v... represents the initial state when in it we make $t = 0$, and since it satisfies the differential equation of the propagation of heat, it represents also that state of the solid which exists at the commencement... and we cannot suppose that there is a more general expression.

Kovalevskaya's counterexample stimulated Weierstrass' interest in this problem. In particular, he wanted to know how the solutions presented by Fourier could be represented analytically, now that he realized that they didn't automatically generate convergent Taylor series.

Today, a century and a quarter later, it is difficult to appreciate how surprising such examples were in 1874. The discovery of them often requires a young and fresh mind like Kovalevskaya's, to wander more freely along routes that experience had trained older minds to ignore as unpromising. As just mentioned, while Kovalevskaya had been busy constructing a positive theory of

[7] In my 1984 monograph *The Mathematics of Sonya Kovalevskaya* (Springer-Verlag, Berlin - Heidelberg - New York), p. 34, I erroneously asserted that she had shown that the radius of convergence would be positive in this case.

analytic solutions to partial differential equations, Weierstrass, spurred by her counterexample with the heat equation, had been re-examining this equation. Teacher and pupil seem to have finished their respective labors about the same time. Weierstrass communicated his results to Kovalevskaya in a letter of 6 May 1874, a few months before she received her doctoral degree, but when she had obviously finished the main work. Here is what he wrote:

> Let λ be a real variable and $f(\lambda)$ a function of this variable subject only to the following conditions:
>
> 1) it remains finite for values of λ between any two finite limits λ_1 and λ_2;
>
> 2) it may be discontinuous or undefined at arbitrarily many points, even infinitely many, provided that the integral $\int\limits_{\lambda_1}^{\lambda_2}$ is defined in the sense given by Riemann (see his paper on Fourier series);
>
> 3) the limit of $f(\lambda)$ as λ approaches $-\infty$ or $+\infty$ need not be finite, but the function
>
> $$\frac{\log|f(\lambda)|}{\lambda^2}$$
>
> must vanish for $\lambda = \pm\infty$. (Such is the case, for example, when $f(\lambda)$ becomes infinite like a positive power of λ, or like $e^{\alpha\lambda^\beta}$, where α and β are positive quantities and $\beta < 2$.)
>
> The following theorem holds. If u, v, and w are complex variables and the real part of u is <u>positive</u>, then the integral
>
> $$\int_{-\infty}^{+\infty} f(\lambda)e^{-(u\lambda^2+v\lambda+w)}\,d\lambda$$
>
> <u>also</u> always has a definite finite value wherever u, v, and w do and is a <u>regular function of u, v, and w</u>. This function can therefore be expanded in a series of positive powers of
>
> $$u-u', \quad v-v', \quad w-w'$$
>
> that always converges provided u is taken sufficiently close to u' so that for all other values of these quantities having the same distance from u' the real part is positive.
>
> <u>This series can also be obtained by expanding</u>
>
> $$f(\lambda)e^{-(u\lambda^2+v\lambda+w)}$$
>
> <u>in powers of</u> $u-u'$, $v-v'$, $w-w'$, <u>and then integrating each coefficient of the resulting series from</u> $-\infty$ to $+\infty$.
>
> Applied to the integral
>
> $$\varphi(x,t)\varphi(x,t) = \frac{1}{2\sqrt{\pi}}\frac{1}{\sqrt{t}}\int_{-\infty}^{\infty} f(\lambda)e^{-\frac{(\lambda-x)^2}{4t}}\,d\lambda,$$

> the theorem asserts that this integral can be expanded in a convergent series of the form
>
> $$\frac{e^{-\frac{x^2}{4t}}}{\sqrt{t}}\left\{\varphi_0(t) + \varphi_1(t)x + \cdots + \varphi_n(t)x^n + \cdots\right\}$$
>
> when t is subject to the condition that its real part is always positive (so that the same is true of $1/t$), while x varies unrestrictedly. Here $\varphi_0(t)$, $\varphi_1(t)$,... are regular functions of t. The function $\varphi(x,t)$ also satisfies the differential equation
>
> $$\frac{\partial\varphi(x,t)}{\partial t} = \frac{\partial^2\varphi(x,t)}{\partial x^2}$$
>
> in the entire region of value of (x,t) for which it is defined.
>
> The proof of this theorem is very simple; in fact a much more general result can be proved, asserting that analytic functions of complex quantities can be derived from completely arbitrary functions of a real variable λ.
>
> I'll talk face to face with you about all this and many related things.
>
> So you see, dearest Sonia, how your remark—which seemed so simple to you—on the peculiar property of partial differential equations, that an infinite series may satisfy such an equation *formally* without converging for any system of values, became for me the point of departure for research of great interest, that clarifies many things. I hope my pupil may continue to thank her teacher and friend in this way.

Weierstrass was inspired to undertake this research precisely because he had not expected to find such counterexamples. Six months later he wrote to Paul Du Bois-Reymond about Kovalevskaya's dissertation:

> I had been of the opinion that a power series in several variables that satisfies a partial differential equation formally must always converge in some region, and hence represent a function that actually satisfies the equation. That such is not the case, as you will see from the example of the equation $\frac{\partial\varphi}{\partial t} = \frac{\partial^2\varphi}{\partial x^2}$ discussed in the dissertation, was discovered completely independently by my student, to my great astonishment, and indeed originally with a much more complicated differential equation than this one, so that at that point she was in great doubt whether she would be able to prove any general result.

(Quoted by R. Bölling in *Briefwechsel zwischen Karl Weierstraß und Sofja Kowalewskaja*, Akademie-Verlag, Berlin, 1993, p. 126. First published by Mittag-Leffler as "Briefe von Weierstrass an Paul du Bois-Reymond," *Acta Mathematica*, **39** (1923), 199–225.)

3.2 Co-discovery by Darboux

Kovalevskaya's dissertation appeared in Band 80 of the *Journal für die reine und angewandte Mathematik* (1875). In April of that year Weierstrass received the *Comptes rendus* rather late, as he had allowed his subscription to lapse. To his amazement, when he finally received them, he found two notes by Gaston Darboux (1842–1917) bearing the titles "Mémoire sur l'existence de l'intégrale dans les équations aux dérivées partielles contenant un nombre quelconque de fonctions et de variables indépendants," (*Comptes rendus*, **LXXX**, 1875, 101–104) and "Sur l'existence de l'intégrale dans les équations aux dérivées partielles d'ordre quelconque," (*Comptes rendus*, **LXXX**, 1875, 317–319). Of course, proofs were not usually given in full in the *Comptes rendus*, only the results, so that Weierstrass could not be quite sure what Darboux had done. In the first of these papers Darboux stated that the initial-value problem

$$\frac{\partial V}{\partial t} = F(p_1, p_2, \ldots, p_n, q_1, \ldots, q_n, t), \quad \Big(p_j = \frac{\partial V}{\partial q_j}\Big),$$
$$V(q_1, \ldots, q_n, 0) = f(q_1, \ldots, q_n),$$

where f and F are analytic functions of their arguments, has an analytic solution $V(q_1, \ldots, q_n, t)$. It is not clear how Darboux came to be studying this problem. His notation is identical to that of Jacobi and Lagrange, so that we may suspect some inspiration from that quarter. Whatever his motivation, his studying this problem at the same time as Kovalevskaya shows that the time was ripe for extending the well-known existence theorems for ordinary differential equations to partial differential equations. Writing to Kovalevskaya on 21 April, Weierstrass took comfort from the fact that Kovalevskaya had been one of the leaders in this effort. He noted that Darboux also mentioned some "exceptional cases," where the integral does not exist and surmised that Darboux also had noticed problems with equations like the heat equation. (In fact, Darboux had not reached the same generality as Kovalevskaya, incorporating the appropriate concept of "normal form" into the hypothesis.)

What might have been a priority dispute, leading to hard feelings, turned out quite amicably. Weierstrass sent a copy of Kovalevskaya's dissertation to Hermite and asked Borchardt (editor of the *Journal für die reine und angewandte Mathematik* at the time) to state that her article had been received in August 1874. Kovalevskaya had no dispute with Darboux or Hermite; indeed they became her closest friends in Paris and were instrumental in arranging the Bordin Prize competition that she won 13 years later. The appearance of the work of Darboux prompted Angelo Genocchi to write to the *Comptes rendus* pointing out Cauchy's 1842 work. Kovalevskaya did have some hard feelings, however, but not against the French mathematicians.

4 Priority Questions: the Personal Aspect

The difficulty of tracking priority for a mathematical discovery, even in the comparatively small world of nineteenth-century Europe, is considerable. Results were discovered, published, forgotten, rediscovered, and republished, and priority was difficult to establish in some cases. This phenomenon manifested itself in regard to the work of Cauchy on partial differential equations, in that no reference was made to his work by Briot and Bouquet, Kovalevskaya, Weierstrass, or Darboux. It appears that neither the German nor the French mathematicians remembered having seen this work, if indeed they did see it; and it took the letter of Genocchi to the *Comptes rendus*, mentioned above, to call attention to it. Controversy was avoided at the time, as already stated. However, the issue of priority arose again a decade later, and may have caused the rupture of a friendship. In 1884, just after Kovalevskaya's appointment as docent in Stockholm, both she and Mittag-Leffler were eager to do two things: (1) get Kovalevskaya a permanent appointment at Stockholm; (2) bring some outstanding foreign and Swedish mathematicians to Stockholm for collaboration. Both efforts, along with the founding of the *Acta Mathematica*, were part of Mittag-Leffler's ambition to make Stockholm into an outstanding center of mathematical research (an effort, by the way, that succeeded admirably). One excellent prospect, they thought, was the young Carl Runge (1856–1927), who had received the doctorate from Berlin in 1880 under the direction of Kummer and Weierstrass. He was a friend of Kovalevskaya's compatriot Dmitrii Selivanov (1855–1932), his fellow student in Berlin, and became acquainted with Mittag-Leffler through Kovalevskaya. Relations among the three mathematicians were at first most cordial, as shown by Kovalevskaya's letter to Mittag-Leffler on 3 September 1884:

> A few days ago I received a letter from Runge, in which he indicated his intention of visiting us during the holidays. He didn't say exactly when he plans to come, but judging from what he said earlier, I think it will be during the second half of September. I think it would be very pleasant for all of us to spend several weeks together in Södertälje, where it is so much nicer than in Stockholm. Just think how wonderful that would be!

Runge was asked to proofread Kovalevskaya's paper on the solution of the Lamé equations in elasticity theory (which contained a serious error that Runge failed to catch). Writing to Kovalevskaya with his comments on 17 November 1884, Runge finished his letter with the following lighthearted banter. The italicized phrases are in English in the original:

> I must unfortunately put an end to your joy over my ailing foot. It is better. And I have not missed two further social dances. Sunday I even resumed playing tennis. Herr Hansemann, whom I met at the Pringsheims, asked very particularly about you. Now that's a fine thing, that you want "*small talk*" in my letters. You? Oh, but did you not upbraid poor Selivanoff for writing only about personal things? Well, I'm in a very bad mood *and won't inflict myself any longer upon you.*

After getting Kovalevskaya a five-year appointment to a professorship based partly on the paper Runge had proofread, Mittag-Leffler was keen to bring this talented young German mathematician to Stockholm. However, one small incident may have spoiled his plans. On July 9, 1885, while vacationing in Switzerland, he wrote a long letter to Kovalevskaya, ending with the following unpleasant paragraph:

> Did you know that your doctoral dissertation can be found in Cauchy's papers in the *Comptes rendus*, Vols. 14, 15, and 16? Runge has discovered this. Weierstrass was very amazed. It's a good thing we didn't know about this when we were arranging the professorship in Stockholm. It doesn't matter, now that you have published much more significant papers.
>
> I think you should study Cauchy's papers and write a new paper in which you explain the situation fully and show what is indisputably your own original research, also including your beautiful joint work with Poincaré.[8] If you do nothing, Schwarz will sooner or later attack you and accuse you of plagiarism.

Of course, as we know, Mittag-Leffler need not have worried. Kovalevskaya wrote back only 6 days later, from Russia, where she had gone for the holidays.

> That Cauchy worked on partial differential equations before me and that he proved that an equation of this kind in general has an integral is not at all news to me, and the priority for this discovery does not belong to Runge. The subject was thoroughly discussed in the French Academy when Darboux published his paper on the same subject, some months after mine. There was also a long correspondence between Weierstrass and Hermite. When I wrote my paper, I did not indeed know, nor did Weierstrass, that Cauchy had already studied the question, and that was naturally a big mistake on our part. However, Weierstrass remarked at the time to Hermite that neither Cauchy nor Darboux had pointed out the essential conditions for a partial differential equation to have an integral (in other words, to be reducible to normal form). Neither of them had suspected that there are cases such as, for example,
>
> $$\frac{\partial \varphi}{\partial t} = a\frac{\partial^2 \varphi}{\partial x^2}$$
>
> where there exists a series that formally satisfies the equation and assumes an arbitrary value $\varphi(x)$ for $t = 0$, but converges only for very special choices of $\varphi(x)$. In the general case it is divergent.

[8] It is not clear what work Mittag-Leffler is referring to. Possibly he means Kovalevskaya's work on hyperelliptic integrals that degenerate to elliptic integrals for some values of their parameters. This work had formed a minor part of her dissertation, and she had just published it in *Acta Mathematica*. Poincaré was impressed by this work and commented on it in print. But that hardly amounts to "joint work."

> That is the result of my research to which Weierstrass has always attached the greatest importance, and, unless Runge has discovered a new paper by Cauchy that escaped the attention of the French mathematicians, I believe that result belongs exclusively to me. Not having the *Comptes rendus* at hand at the moment, I cannot be sure that the papers cited by Runge are the same ones that were discussed when the papers of Darboux and myself were discussed at the Academy. I am almost certain that I told this story to both you and Runge. You probably weren't listening.

Whether from this letter alone, or from other hints Kovalevskaya may have dropped, Mittag-Leffler began to have doubts about her desire to see Runge in Stockholm. Six weeks later he wrote to her from the Hotel Victoria in Switzerland concerning a misunderstanding about the Oscar II mathematical competition that he was trying to arrange (it was won by Poincaré):

> Poor Weierstrass! I feel genuinely sorry for him. He is a great and magnanimous man, and the world has so few places for such people, especially in Germany... Did I perhaps do wrong in telling Weierstrass about the disagreements with Kronecker? I considered it my duty to do so, since Weierstrass was a member of the committee, and I was afraid Kronecker might come to Weierstrass with some other version of what happened.
>
> Runge has acted like a gentleman. Kronecker disparaged me to him in the strongest terms. He wrote me about that immediately to warn me, and I received his letter only a few days after Kronecker's letter. This feature of Runge's touched me, not because it was in any way unusual in itself, but because I don't know any other German who would have done it. ...
>
> But tell me now clearly and frankly, do you want us to talk with Runge or not? It would be comparatively easy to get Runge an appointment, and I have the impression he would accept gladly. His chances in Germany are small, since he has too independent a personality...
>
> If you have any particular reason why you wouldn't want us to choose Runge, please tell me as frankly as possible. You can rely on my discretion completely. If you would rather not give the reason (if there is a reason), at least give me your frank opinion whether we should consider him or not. You can simply tell me clearly and definitely that you don't want us to consider him, and I'll drop the matter.
>
> Runge has sent me several articles and descriptions of various new research. There are some excellent things among them. For example, he can express any real-valued function (integrable, I assume) of one variable as a series of continued fractions.

A person doesn't write things like this unless he has some idea of what the response will be. He obviously has a strong suspicion that Kovalevskaya doesn't want Runge around Stockholm. Most likely he had been trying unsuccessfully to draw her out on this subject for some time. Well, he asked for her frank opinion, and Kovalevskaya gave it in her next letter, dated 30 August:

> In regard to Runge, I can only tell you that a number of facts have revealed to me that the character flaw of vanity is just as overdeveloped in him as in the majority of Germans. That is what has made me somewhat cool towards him and makes me hesitate to have to work with him. I will confess frankly that for me personally it would not be pleasant to see him in Stockholm, precisely because I had been very friendly towards him and then changed my opinion. In such cases personal relations always become difficult and awkward.
>
> On the other hand, I recognize fully that Runge has a great deal of talent, and I do not feel I have the right to stand in the way of your proposing him as a candidate if (after what I have just said) you are convinced that he would be more useful than anyone else in that position. For me personally any other candidate would be more agreeable. It seems to me that of the 4 you named Mellin has the most talent and would be the pleasantest to have in Stockholm. It is also a very good thing that Mellin is one of your students, and that will be encouraging to the others.

One cannot help wondering what the facts were that convinced Kovalevskaya of Runge's vanity; since we have no other information, suspicion rests strongly on his completely unnecessary and misguided attack on her dissertation. Whether because of Kovalevskaya's objections or for some other reason, Runge didn't get the job in Stockholm. It went instead to Anders Lindstedt (1854–1939), a Swede who was at the time professor of applied mathematics at the University of Dorpat. (Lindstedt served as rector of the Stockholm Högskola from 1902 to 1909.) On 12 October Mittag-Leffler wrote to Runge:

> As you see, under the circumstances I cannot advise you to send in your application. I don't have to tell you how deeply it pains me to have to give up the possibility of greeting you as a colleague here in Stockholm. It is a long-cherished plan, one that I formed with great enthusiasm, which I must relinquish. Of course, I know that Lindstedt is competent; but for me he is not enough of a mathematician, and with you here in Stockholm we could have done much more for mathematics than with Lindstedt.

Runge's biographer Richenhagen stated that no foreigner could have hoped to compete with a native Swede for this position, so that Runge's chances were nil from the outset. At the very least, we may note, having a Swede appointed gave Mittag-Leffler a diplomatic reason to tell Runge not to apply; perhaps that is why Lindstedt won out, not only over Runge but also over Mellin. One suspects that Mittag-Leffler would not have wanted to explain to Runge that he was choosing his own student, who happened to be Finnish, in preference to Runge. However that may be, Runge found a position in Hannover the following year.

5 The Continuing Significance of the Theorem

The Cauchy–Kovalevskaya theorem has had remarkable endurance, surviving large changes in styles and interests in partial differential equations and appearing in almost unchanged form in the majority of standard textbooks throughout the twentieth century. It appeared, for example, in Ludwig Bieberbach's monograph *Differentialgleichungen*, zweite Auflage (Die Grundlehren der mathematischen Wissenschaften, Band VI, Springer-Verlag, Berlin, 1926, pp. 300–305), without being given any name. It was proved by the method of majorants, which the author correctly attributed to Cauchy. He did not mention Kovalevskaya, however. The first use of the name "Cauchy–Kovalevskaya theorem" that I have found is in the lectures by J. Hadamard, some given as early as 1911 at Columbia University, and printed as *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover, New York, 1952, originally published in 1923). Hadamard devotes about twenty pages to a thorough discussion of the theorem, giving the proof with some improvements in the majorant suggested by Goursat. He also mentions Darboux in this connection. After giving the proof, he comments:

> The reasonings of Cauchy, S. Kowalewsky, and Darboux, the equivalent of which has been given above, are perfectly rigorous; nevertheless, their conclusion must not be considered as an entirely general one. The reason for this lies in the hypothesis, made above, that Cauchy's data, as well as the coefficients of the equations, are expressed by analytic functions; and the theorem is very often likely to be false when this hypothesis is not satisfied.
>
> We say "often" and not "always," for it may also happen that the statement of Cauchy–Kowalewsky given above should prove to be accurate for a quite general choice of data; and indeed, one of the most curious facts in this theory is that apparently very slightly different equations behave in quite opposite ways in this matter.

Hadamard once again looked at the heat equation, noting that Sergei Bernshtein had proved that any solution of a parabolic equation (of which the heat equation is one) with analytic coefficients is analytic with respect to the variable we have called x.[9] He also noted that Holmgren had shown that the solution to the Cauchy problem is unique whether analytic or not.

[9]Bernshtein was working on Hilbert's nineteenth problem, which asked if the solutions to differential equations arising in the calculus of variations (which were usually elliptic, as Hilbert pointed out) were necessary analytic. For his positive solution to this problem in 1903 Bernshtein was awarded the doctoral degree by an examining committee containing Hadamard, Picard, and Poincaré.

To cite just a few of the many appearances of the Cauchy–Kovalevskaya theorem in twentieth-century literature, we note that it features prominently in the monograph of I. G. Petrovsky (*Partial Differential Equations*, Interscience Publishers, New York, 1952, pp. 14–26). In the 1975 monograph by E. T. Copson (*Partial Differential Equations*, Cambridge University Press) the theorem is given early on (pp. 9–27), and proved by the method of majorants. Then, late in the book (pp. 244–246), a section is devoted to an application of it. This application turns out to be the fact that if $u(x,t)$ is an analytic function of two variables satisfying the heat equation $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$, then $F(x) = u(x,0)$ must be an entire function of x and

$$u(x,t) = F(x) + \sum_{1}^{\infty} \frac{t^n}{n!} F_{2n}(x), \tag{51}$$

where $F_k(x)$ is the kth derivative of F, so that (51) is the solution of the heat equation satisfying the initial condition $u(x,0) = F(x)$. As shown above, this result is found almost word for word in Kovalevskaya's dissertation and in her 1875 paper, although no reference is given to either of those works. Probably they had become "folk theorems" after a whole century, known to all the experts, but whose source had been lost.

Although the theorem appears to have evolved only slightly over the last 125 years, there have been a few improvements on it, noted in the latest textbooks. In the monograph *Linear Partial Differential Equations. Foundations of the Classical Theory* (*Encyclopaedia of Mathematical Sciences*, Vol. 30), Springer-Verlag, Berlin - Heidelberg - New York, 1993) Yu. V. Egorov and M. A. Shubin mention a 1973 result of Hörmander, which asserts that an initial-value problem for the equation

$$\frac{\partial^{\beta_1+\cdots+\beta_n} u}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} = \sum_{\alpha_1+\cdots+\alpha_n \le \beta_1+\cdots+\beta_n} a_{\alpha_1,\ldots,\alpha_n} \frac{\partial^{\alpha_1+\cdots+\alpha_n} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} + f$$

with analytic coefficients and analytic data has an analytic solution if the n-tuple $\beta_1,\ldots,\beta_n)$ is not a convex combination of n-tuples $(\alpha_1,\ldots,\alpha_n)$, or if the sum of the absolute values of the coefficients $a_{\alpha_1,\ldots,\alpha_n}$ for which the equality $\alpha_1 + \cdots + \alpha_n = \beta_1 + \cdots + \beta_n$ holds is bounded at the origin. They further point out that it is not necessary for the function G on the right-hand side of Eq. (43), which Kovalevskaya had assumed was a polynomial, to be analytic in all its arguments. It suffices for it to be analytic in all arguments except x and continuous in x. They also note Holmgren's remark, mentioned by Hadamard, to the effect that the Cauchy–Kovalevskaya theorem makes it possible to prove uniqueness of the solution in some cases, even when the data and the solution are not analytic.

Acknowledgement. I am grateful to I. Grattan-Guinness for helpful information in the writing of this paper.

6 Literature

Belhoste, Bruno. *Cauchy, un mathématicien légitimiste au XIXe siècle*, Belin, Paris, 1985.

Bieberbach, L. *Differentialgleichungen*, zweite Auflage (Die Grundlehren der mathematischen Wissenschaften, Band VI), Springer-Verlag, Berlin, 1926.

Bölling, R., ed. *Briefwechsel zwischen Karl Weierstraß und Sofja Kowalewskaja*, Akademie-Verlag, Berlin, 1993.

Briot, Ch. and Bouquet, J.-K. "Étude des fonctions d'une variable imaginaire," *Journal de l'École Polytechnique*, Cahier 36 (1856), 85–131.

Briot, Ch. and Bouquet, J.-K. "Recherches sur les propriétés des fonctions définies par des équations différentielles," *Journal de l'École Polytechnique*, Cahier 36 (1856), 133–198.

Cauchy, A.-L. "Mémoire sur un théorème fondamental, dans le calcul intégral," *Comptes rendus*, **XIV** (1842), 1020–1026.

Cauchy, A.-L. "Mémoire sur l'emploi du calcul des limites dans l'intégration des équations aux dérivées partielles, *Comptes rendus*, **XV** (1842), 44–59.

Cauchy, A.-L. *Équations différentielles ordinaires: Résumé des leçons données à l'École Royale Polytechnique*, Johnson Reprint Corporation, New York, 1981.

Cooke, R. *The Mathematics of Sonya Kovalevskaya*, Springer-Verlag, Berlin - Heidelberg - New York, 1984.

Copson, E. T. *Partial Differential Equations*, Cambridge University Press, 1975.

Darboux, G. "Mémoire sur l'existence de l'intégrale dans les équations aux dérivées partielles contenant un nombre quelconque de fonctions et de variables indépendants," *Comptes rendus*, **LXXX** (1875), 101–104.

Darboux, G. "Sur l'existence de l'intégrale dans les équations aux dérivées partielles d'ordre quelconque," *Comptes rendus*, **LXXX** (1875), 317–319.

Egorov, Yu. V. and Shubin, M. A., eds. *Partial Differential Equations 1. Foundations of the Classical Theory* (Encyclopaedia of Mathematical Sciences, Vol. 30), Springer-Verlag, Berlin - Heidelberg - New York, 1992.

Feigenbaum, L. "Infinite series and solutions of ordinary differential equations, 1670–1770," in: *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, I. Grattan- Guinness, ed., Routledge, London and New York, 1994, Vol. 1, pp. 504–519.

Fourier, J. J. *The Analytical Theory of Heat*, Stechert, New York, 1888.

Gerhardt, C. I. *G. W. Leibniz, Mathematische Schriften* (7 volumes), Georg Olms Verlag, Hildesheim and New York, 1971.

Goursat, E. "Sur la définition générale des fonctions analytiques d'après Cauchy," *Transactions of the American Mathematical Society*, **1**, No. 1, 14–16.

Grattan-Guinness, I. *Convolutions in French Mathematics, 1800–1840* (Science Networks, Historical Studies, Vols. 2–4), Birkhäuser, Basel, 1990.

Grattan-Guinness, I. *Joseph Fourier, 1768–1830*, The MIT Press, Cambridge, 1972.

Hadamard, J. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Dover, New York, 1952

Jacobi, C. G. J. *Gesammelte Werke* (8 volumes). Chelsea Publishing Company, New York, 1969.

Klein, F. *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea, New York, 1967.

Kolmogorov, A. N. and Yushkevich, A. P. *Mathematics of the 19th Century. Volume 3: Constructive Function Theory, Ordinary Differential Equations, Calculus of Variations, Theory of Finite Differences*, Birkhäuser Verlag, Basel, 1998.

Lützen, J. "The solution of partial differential equations by separation of variables: a historical survey," in: *Studies in the History of Mathematics* (Vol. 26 of *Studies in Mathematics*), Mathematical Association of America, 1987, pp. 242–277.

Lützen, J. "Partial differential equations," in: *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, I. Grattan-Guinness, ed., Routledge, London and New York, 1994, Vol. 1, pp. 452–469.

Petrovsky, I. G. *Partial Differential Equations*, Interscience Publishers, New York, 1952.

Richenhagen, G. *Carl Runge (1856–1927): Von der reinen Mathematik zur Numerik* (Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte der Mathematik, Band 1), Vandenhoeck & Ruprecht, Göttingen, 1985.

Siegmund-Schultze, R. *K. Weierstraß, Ausgewählte Kapitel aus der Funktionenlehre* (Teubner-Archiv zur Mathematik, Band 9), Teubner, Leipzig, 1988.

Weierstrass, K. *Mathematische Werke* (7 volumes). Johnson Reprint Corporation, New York, 1927.

Whiteside, D. T., ed. *The Mathematical Papers of Isaac Newton* (8 volumes), Cambridge University Press, 1967–1981.

Ковалевская, С. В. *Научные Работы*, Издательство Академии Наук СССР, 1948.

Кочина, П. Я. *Софья Васильевна Ковалевская, 1850–1891*, Издательство «Наука», Москва 1981.

Кочина, П. Я. *Переписка С. В. Ковалевской и Г. Миттаг-Лефлера*, (*Научное Наследство*, том седьмой), Издательство «Наука», Москва, 1984.

Олейник, О. А., “Теорема С. В. Ковалевской и ее роль в современной теории уравнений с частными производными,” *Математика в Школе*, **5** (1975), 5–9.

Петрова, С. С. и Демидов, С. С., “Развитие математического анализа,” в кн.: *Очерки по Истории Математики*, Издательство Московского Университета, 1997, с. 7–93.