AN ELEMENTARY PROOF OF THE CAUCHY-KOWALEVSKY THEOREM

WOLFGANG WALTER
Universität Karlsruhe, Mathematisches Institut I, Kaiserstr. 12, D-7500 Karlsruhe, West Germany

Dedicated to Professor Johannes Weissinger on his seventieth birthday*

0. Introduction. It is the main aim of this article to present an elementary proof of the Cauchy-Kowalevsky (= C – K) theorem based on the contraction principle. The theorem deals with a nonlinear partial differential equation

$$(1)\qquad D_t^p u = F(t, x_1, \ldots, x_n, u, \ldots, D_t^k D_x^\alpha u, \ldots) \qquad (k + |\alpha| \leq p;\ k < p)$$

or with a system of such equations, where $D_t = \partial/\partial t$, $D_x^\alpha = \partial^{|\alpha|}/\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}$, and $p \geq 1$. It states that if for $t = 0$ initial data $D_t^k u = \phi_k$ $(0 \leq k < p)$ are prescribed, where the ϕ_k are analytic functions of $x = (x_1, \ldots, x_n)$ in a neighborhood of $x_0 = (x_1^0, \ldots, x_n^0)$, and if the function F is analytic in a neighborhood of the point $(0, x_0, \ldots, D_x^\alpha \phi_k(0, x_0), \ldots)$, then there exists a unique solution $u(t, x)$ in a neighborhood of $(0, x_0)$, analytic in t and x and satisfying the initial conditions. The problem is called *Cauchy's problem*, the initial conditions *Cauchy data*. It was first solved by Cauchy [2], and, in a more general and simplified way, by Sophie v. Kowalevsky [9]. It represents the analogue for partial differential equations of the basic initial value problem for an ordinary differential equation $u^{(p)}(t) = f(t, u, u', \ldots, u^{(p-1)})$ with given initial data $u(0), \ldots, u^{(p-1)}(0)$. Obviously, this analogy motivated Cauchy to consider the problem.

As a preliminary step of the proof found, e.g., in [3; p. 39-56], the equation (or system) is first converted into a more convenient quasilinear first order system

$$(2)\qquad u_t = \sum_{j=1}^{n} B_j(t, x, u)\, u_{x_j} + c(t, x, u).$$

In the vector notation used here, $x = (x_1, \ldots, x_n)$, $u = (u_1, \ldots, u_m)$, the B_j are m by m matrices, and u, c are 1 by m matrices, i.e., column vectors. Some textbooks, e.g. [10] [18], consider only the linear case

$$(3)\qquad u_t = A(t, x)u + \sum_{j=1}^{n} B_j(t, x)\, u_{x_j} + c(t, x),$$

which is much easier to handle. In both cases the initial condition is given by $u(0, x) = \phi(x)$.

We show in the important special case of a nonlinear second order problem (with $n = 1$, for simplicity)

$$(1')\qquad u_{tt} = f(t, x, u, u_t, u_x, u_{tx}, u_{xx}),\ u(0, x) = \phi_0(x),\ u_t(0, x) = \phi_1(x),$$

Wolfgang Walter: I was born in Schwäbisch Gmünd/Germany. Since my early age, mathematics and music have been my main interests. Keeping the latter field for leisure, I studied mathematics and physics at the Universität Tübingen and received a Doctorate in 1956 under E. Kamke. In 1958 I followed an invitation by A. Weinstein and spent a year at the University of Maryland. Since 1963 I am Professor of Mathematics at the Universität Karlsruhe, spending a considerable part of my professional life at various American universities. My main interest is in ordinary and partial differential equations with a special emphasis in differential inequalities. I have written the monograph *Differential and Integral Inequalities* (Ergebnisse der Mathematik, Vol. 55) and several mathematical textbooks. The main ideas of the present article emerged in 1982/83 while I was a Visiting Professor at the University of Florida and the Georgia Institute of Technology.

*As early as 1952 Professor Weissinger suggested to use contraction proofs in the classroom. His article *Zur Theorie und Anwendung des Iterationsverfahrens* (Math. Nachr., 8, 193–212, 1952) contains the generalized contraction principle and numerous applications from different branches of pure and applied analysis.

how this transformation is accomplished. Setting $u_t = v$ and $u_x = w$, a first order problem

$$(1'')\qquad \begin{cases} u_t = v, & u(0,x) = \phi_0(x), \\ v_t = f(t,x,u,v,w,v_x,w_x), & v(0,x) = \phi_1(x), \\ w_t = v_x, & w(0,x) = \phi_0'(x), \end{cases}$$

is obtained. It is easily seen that u solves (1′) if and only if (u,v,w) solves (1″) (passing from the latter equation to the former one requires the equation $w = u_x$, which follows from $u_{tx} = w_t$ by integration in the t direction). Now, (1″) or, more generally, any first order system of the form

$$u_t = g(t,x,u,u_x) \text{ with } u = (u_1,\dots,u_m), \qquad g = (g_1,\dots,g_m)$$

transforms by differentiation with respect to x into a quasilinear system for $(u,v) = (u,u_x)$

$$(2')\qquad \begin{cases} u_t = g(t,x,u,v), \\ v_t = g_x + g_u v + g_v v_x, \end{cases}$$

the argument being (t,x,u,v) in all three terms. The reasoning above is easily adapted to the case where $x = (x_1,\dots,x_m)$ and/or $u = (u_1,\dots,u_m)$ in (1′).

In the classical method of proof, used by Cauchy **[2]** and Sophie v. Kowalevsky[1] **[9]** and presented in many textbooks, the solution is sought as a power series in t and x_i whose coefficients are determined by recursion formulas using the corresponding expansions of B_j and c. The centerpiece of the approach is the convergence proof for this series by the *method of majorants*. It consists of considering an auxiliary problem of the same structure (with the same recursion formulas) which majorizes the given problem in the sense that all coefficients involved are positive and larger than the absolute values of those in the given problem. Cauchy invented and applied the method in many instances. He called it, not very suggestively, *calcul des limites*.

A new proof based on entirely different ideas was discovered in 1941 by M. Nagumo **[11]**. Nagumo considers the quasilinear system (2) with zero initial values (no loss of generality) and transforms it into an equivalent equation for $v = u_t$

$$v(t,z) = \sum_{j=1}^{n} B_j(t,z,Iv)\,Iv_{z_j} + c(t,z,Iv),$$

where $(Iv)(t,z) = \int_0^t v(\tau,z)\,d\tau\ (= u)$. This equation is treated as an operator equation $v = \Phi(v)$ in an appropriate Banach space E, and a solution is obtained as a fixed point of Φ, using the Schauder fixed point theorem. An essential tool is a lemma, reproduced with proof in Section 2, which provides an estimate of the derivative of a holomorphic function in terms of the function itself (in **[11]** it is given for the case where Ω is a ball $|z| < R$). This lemma has proved to be of importance in many modern developments of the Cauchy problem and is called here *Nagumo's lemma*. As a consequence, the variables x_i are now complex by necessity, and for this reason we have written z instead of x. On the other hand, the variable t is assumed to be real, and the solution is shown to exist in a conical region $0 \leqslant t < l$, $|z| < R - Lt$. The functions B_j and c are assumed to be continuous in all variables and holomorphic in z and u, and v is continuous in (t,z) and holomorphic in z. Strictly speaking, Nagumo's theorem does not cover the C-K theorem. But its proof is also valid in the case where t is a complex variable and the functions involved are holomorphic in all variables, a fact which Nagumo does not mention. In short, Nagumo obtained a new proof and, by allowing t to be a real variable, an important generalization of the C-K theorem. But this approach has also a drawback: uniqueness does not follow from the Schauder fixed point theorem and has to be proved separately. Another point of comparison

[1] Different ways of spelling the name have appeared in the more recent literature (Kowalewska, Kowalewskaya, . . .). The author, who has neither the historic nor the linguistic expertise to make a reasonable choice, has been led by the doctoral dissertation [9] which was written by "Frau Sophie von Kowalevsky".

regards the "real" version of the C-K theorem, which states that the solution is real-valued for real values of the variables if the coefficients in the differential equation have this property. Nagumo does not prove that fact, yet it is easily obtained in the framework of his proof. If one restricts either the whole Banach space or the convex, compact subset which is mapped into itself by Φ, to functions which are real-valued for real z_j, then his proof gives the desired result. Recently, Keller and Schneider [8] published a proof of the classical C-K theorem by reduction to Schauder's fixed point theorem, which bears a likeness to Nagumo's proof (the latter is not cited).

Another approach which gives rise to various generalized versions of the C-K theorem was developed in the sixties. The basic notion is that of a *scale of Banach spaces*, that is a collection $(B_\rho)_{\rho>0}$ of Banach spaces with the property that $0<\rho<\sigma$ implies $B_\sigma \subset B_\rho$ and $\|u\|_\rho \leqslant \|u\|_\sigma$ for $u \in B_\sigma$. One considers bounded linear operators L from B_σ into B_ρ, $\|Lu\|_\rho \leqslant \|L\|_\rho^\sigma \|u\|_\sigma$ for any $0<\rho<\sigma$, which satisfy an estimate

$$\|L\|_\rho^\sigma \leqslant \frac{c}{\sigma-\rho}.$$

These notions are already used in Vol. III of Gelfand and Silov's treatise [4], A. 2.1-2.3, in connection with the linear Cauchy problem for constant coefficients and analytic initial values. In 1965, Ovsjannikov [14] proved an existence and uniqueness theorem, often quoted as *Ovsjannikov's theorem*, for the Cauchy problem

$$\frac{du}{dt} = L_t u, \qquad u(0) = u_0 \in B_\rho$$

in this abstract framework, where (L_t) is a family of linear operators with the properties stated above for L. Further research along these lines by Treves [20] [21], Nirenberg [12], Ovsjannikov [15] [16], Nishida [13], Pate [17] and others (see also the bibliography of these articles for further reading) led to linear and nonlinear "abstract" Cauchy-Kowalevsky theorems. The connection with the classical theorem is established by the example

$$B_\rho = \left\{ u\colon u(z) \text{ is holomorphic for } \max|z_i| < \rho,\ \|u\|_\rho = \sup_{|z|<\rho} |u(z)| < \infty \right\}$$

and the operators $L = \partial/\partial z_i\colon B_\sigma \to B_\rho$ which satisfy $\|\partial/\partial z_i\|_\rho^\sigma \leqslant 1/(\sigma-\rho)$ for $0<\rho<\sigma$. An easily accessible exposition of the linear case is given in Chapter 17 of Treves' book [22]. This theory shares with Nagumo's proof the property that t can be a real or complex variable.

The proof presented in this paper is elementary inasmuch as it uses Banach's contraction principle as the only tool from functional analysis. Basic properties of holomorphic functions $u(\zeta)$ depending on a single complex variable ζ such as Cauchy's integral formula are used. As far as holomorphic functions of several complex variables are concerned, it is sufficient to know the definition (a function $f\colon G \to \mathbb{C}$, where G is an open subset of $\mathbb{C}^n$, is holomorphic in G, if f and the derivatives $\partial f/\partial z_i$ $(i=1,\ldots,n)$ are continuous in G) and the fact that $f(u_1(\zeta),\ldots,u_n(\zeta))$ is holomorphic in ζ, if f and the u_i are holomorphic. Power series expansions, which can be used to give an equivalent definition of holomorphy, are not used. The exposition is written with an eye to classroom use. For this reason, the linear case is treated first, where the proof is remarkably simple and short. Apart from didactic advantages, the proof has an important mathematical consequence: it gives not only existence and uniqueness, but also estimates between the solution and an approximate solution, in particular, continuous dependence on the initial values and the right hand side of the differential equation. In short: the proof by contraction shows that the analytic Cauchy problem is a well-posed (or correctly set) problem.

At first glance this statement seems to contradict well-known facts about second order partial differential equations of mathematical physics. In explaining the situation, we consider the simplest prototypes, (a) the hyperbolic (vibrating string) equation $u_{tt} = u_{xx}$, (b) the elliptic (potential) equation $u_{tt} + u_{xx} = 0$, and (c) the parabolic (heat) equation $u_t = u_{xx}$. In the hyperbolic case, the problem models a simple physical situation. It is correctly set, and it has a solution not only for analytic, but also for much more general (e.g., continuous) initial values. In contrast, the classical, well-posed elliptic problems are not Cauchy problems, but boundary value problems where only one function (e.g., u) is prescribed on the whole boundary. Consider, for example, problem (b′) with initial values $u(0,x) = 0$, $u_t(0,x) = \phi_1(x)$. If a solution exists, it is a harmonic and hence (real) analytic function for $t > 0$. Furthermore, the solution allows an analytic continuation for $t \leqslant 0$ given by $u(-t,x) = -u(t,x)$. Consequently, ϕ_1 is analytic. Thus problem (b'), while possessing a solution for analytic ϕ_1, by virtue of the C-K theorem, never has a solution for non-analytic ϕ_1. A more detailed account is given in Hadamard's classical *Lectures on Cauchy's problem* [5, Chapter II]. Here one finds also the famous example (p. 33):

$$u_{tt} + u_{xx} = 0, \quad u(0,x) = 0, \quad u_t(0,x) = a \sin kx,$$

with the solution $u(t,x) = (a/k)(\sin kx)(\sinh kt)$. If $a > 0$ is small, then $|u(0,x)| \leqslant a$ is small, but for any fixed $t \neq 0$ the solution can be made arbitrarily large by choosing k large. Hence the zero solution ($a = 0$) does not depend continuously on the initial values. These considerations give the impression that the applicability of the C-K theorem to elliptic problems is a strange curiosity: the solution exists, but it behaves oddly.

But the odd behavior appears only as long as we consider analytic functions for *real* variables (one might add that there are other cases of odd behavior in this setting). If we put holomorphic functions in their natural habitat, which is $\mathbb{C}$ or $\mathbb{C}^n$, the oddities vanish. Theorem 2 will make this point more precise. It is followed by a discussion of Hadamard's example.

The heat equation of problem (c) is not of the type of equation (1); in this equation we have $p = 1$, but $|\alpha| = 2$ in violation of the condition $|\alpha| \leqslant p$. The C-K theorem becomes false for the heat equation. The following counterexample is already found in [9, p. 22]. The problem $u_t = u_{xx}$, $u(0,x) = 1/(1-x)$ has a unique formal power series solution

$$u(t,x) = \sum_{i=0}^{\infty} \frac{(2i)!\,t^i}{i!(1-x)^{2i+1}} = \sum_{i,k=0}^{\infty} \frac{(2i)!}{i!k!} t^i x^k,$$

but the series diverges for any $t \neq 0$.

In this connection we remark that well-posedness, while being a standard subject with regard to initial and boundary value problems for elliptic, hyperbolic, ... partial differential equations, is only rarely discussed in the Cauchy problem with analytic data. Also, the notion of well-posedness is often defined in a narrow sense, meaning continuous dependence on initial values only, while the physical meaning of this notion requires that continuous dependence on the right hand side of the differential equation be included.

Let us finally note that there is a more general Cauchy problem for equation (1), where Cauchy data (i.e., initial values) are given on a noncharacteristic hypersurface Ω defined by an equation $\phi(t, z_1, \ldots, z_n) = 0$. This case can be reduced—at least locally—to the case treated here where Ω is lying in the hyperplane $t = 0$. The notion of noncharacteristic surface and the reduction procedure are described in many textbooks, e.g., in [7; Chapter 3] and [**18**; § 3].

The author has looked for a proof of the C-K theorem by contraction for some time. His conviction that such a proof must exist was based on Bessaga's noteworthy theorem [**1**]:

If M is a set, $\alpha \in (0,1)$ and T: $M \to M$ a map with the property that T^n has one and only one fixed point for $n = 1, 2, \ldots$, then there exists a metric d on M which makes M a complete metric space and for which $d(Tx, Ty) \leqslant \alpha d(x,y)$.

The Cauchy problem in question can be written as a fixed point equation $u = Tu$ (see (9) for the linear case and (18) for the nonlinear case), and the C-K theorem states that T has a unique fixed point. Moreover, it is not difficult to show that T^n also has a unique fixed point for any $n \geqslant 1$. Bessaga's theorem now implies that T is a contraction in a suitable metric space. The problem was to find the metric.

1. Geometrical preliminaries. Since we want to construct a solution in the large, some

geometrical c
$\Gamma = \partial\Omega$, and
norm $|z| =$ n
consists of al
Ω_t be the set
(4)
is the distanc
base Ω and s
($t < 0$). This
later:
(5)
This follows
If one is
$B_R(z_0)$, the
$\Omega_t = B_t(z_0)$
2. The lin
(6)
(7)
which can be
(8)
where
Here, $u = (u$
and f are co
and c are co
holomorphic
case it is unc
cases, a solut
derivatives w
(7), and vice
To avoid
sponding ope
with the prop
The follow
and is essenti
NAGUMO'S
where $C_p = ($

geometrical considerations are necessary. Let Ω be an open set in $\mathbb{C}^n$ with a nonempty boundary $\Gamma = \partial\Omega$, and let $d(z) = \operatorname{dist}(z, \Gamma)$ be the distance from $z \in \Omega$ to Γ, measured in the maximum norm $|z| = \max_{i=1,\dots,n}|z_i|$. The set G, a subset of $\mathbb{R} \times \mathbb{C}^n$ (real case) or $\mathbb{C}^{n+1}$ (complex case), consists of all points (t, z) with $z \in \Omega$ and $|t| < \eta d(z)$, where $\eta > 0$ will be specified below. Let Ω_t be the set of all $z \in \Omega$ such that $(t, z) \in G$ or, equivalently, $d(z) > |t|/\eta$. It is easily seen that

$$(4) \qquad d(t, z) = d(z) - \frac{|t|}{\eta} > 0 \qquad (z \in \Omega_t)$$

is the distance from z to the boundary $\Gamma_t = \partial\Omega_t$. In geometrical terms, G is the double cone with base Ω and slope η, and Ω_t is the base of that part of G which lies above t ($t > 0$) or below t ($t < 0$). This description applies to the real case. The following property of $d(t, z)$ will be needed later:

$$(5) \qquad z \in \Omega_t, |z - z'| = r < d(t, z) \Rightarrow z' \in \Omega_t \text{ and } d(t, z') \geq d(t, z) - r.$$

This follows readily from the triangle inequality $d(z) \leq |z - z'| + d(z')$.

If one is interested only in a local solution, it suffices to consider the case where Ω equals $B_R(z_0)$, the open ball in $\mathbb{C}^n$ with center at z_0 and radius R. In this case $d(z) = R - |z - z_0|$, $\Omega_t = B_r(z_0)$ with $r = R - |t|/\eta$ and $d(t, z) = R - |z - z_0| - |t|/\eta$.

2. The linear case. The linear C-K theorem deals with the following initial value problem

$$(6) \qquad u_t = A(t, z)u + \sum_{j=1}^{n} B_j(t, z)u_{z_j} + c(t, z) \text{ in } G,$$

$$(7) \qquad u(0, z) = \phi(z) \text{ in } \Omega,$$

which can be stated as an equivalent integral equation of Volterra type

$$(8) \qquad u(t, z) = g(t, z) + \int_0^t \left[A(\tau, z)u(\tau, z) + \sum_{j=1}^{n} B_j(\tau, z)u_{z_j}(\tau, z) \right] d\tau,$$

where

$$g(t, z) = \phi(z) + \int_0^t c(\tau, z)\, d\tau.$$

Here, $u = (u_1, \dots, u_m)$ has values in $\mathbb{C}^m$, A and B_j are complex-valued m by m matrices, and c and ϕ are complex-valued m-vectors. In our matrix notation, u, $u_t = \partial u/\partial t$, $u_{z_j} = \partial u/\partial z_j$, ϕ and c are considered column vectors. A solution of (8) is, by definition, continuous in G and holomorphic in z for fixed t (real case) or holomorphic in t and z (complex case). In the complex case it is understood that integration is performed along the straight line from 0 to t. In both cases, a solution of (8) is, under proper assumptions regarding $A, \dots, \phi$, of class $\mathbb{C}$ in G (complex derivatives with respect to the complex variables) and a solution of the initial value problem (6) (7), and vice versa.

To avoid ambiguity, we use the maximum norm $|u| = \max|u_j|$ $(1 \leq j \leq m)$ and the corresponding operator norm for matrices $A = (a_{jk})$, $|A| = \max_k \sum_j |a_{jk}|$, but any other pair of norms with the property $|Au| \leq |A||u|$ would do as well.

The following lemma, which goes back to Nagumo [11], gives a bound for $\partial f/\partial z_j$ in terms of f and is essential for the proof.

NAGUMO'S LEMMA. *Let $f\colon \Omega \to \mathbb{C}^m$ be holomorphic and $p \geq 0$. Then*

$$|f(z)| \leq \frac{c}{d^p(z)} \Rightarrow |f_{z_j}(z)| \leq C_p \frac{c}{d^{p+1}(z)},$$

where $C_p = (1 + p)\left(1 + \frac{1}{p}\right)^p < e(p + 1)$, $C_0 = 1$.

Proof. For a function ψ of a single complex variable ζ, which is holomorphic in the circle $|\zeta - \zeta'| \leqslant r$, the estimate

$$|\psi'(\zeta)| \leqslant \frac{1}{r} \max_{|\zeta - \zeta'| = r} |\psi(\zeta')|$$

holds. It follows immediately from the Cauchy integral representation formula

$$\psi'(\zeta) = \int \psi(\zeta')/\left[2\pi i(\zeta' - \zeta)^2\right] d\zeta',$$

in which the integral extends over the circle $|\zeta - \zeta'| = r$. Applying this result with $\zeta = z_i$, we obtain the inequalities

$$|f_{z_i}(z)| \leqslant \frac{1}{r} \max_{|z - z'| = r} |f(z')| \leqslant \frac{c}{r} \max \frac{1}{d^p(z')} \leqslant \frac{c}{r(d-r)^p}.$$

Here, $0 < r < d \equiv d(z)$ and $d(z') \geqslant d - r$ because of (2). The choice $r = d/(p+1)$ gives the (optimal) value stated in the lemma. □

To be specific, the following theorem is formulated for the real case.

THEOREM 1. *Assume that*

(i) *the functions $A(t,z)$, $B_i(t,z)$, $c(t,z)$ are continuous in G and holomorphic in z for fixed t, the function $f(z)$ is holomorphic in z;*

(ii) *there exist positive constants α, β_i, γ, δ and p such that*

$$|A(t,z)| \leqslant \frac{\alpha}{d(t,z)}, \qquad |B_i(t,z)| \leqslant \beta_i,$$

$$|c(t,z)| \leqslant \frac{\gamma}{d^{p+1}(t,z)}, \qquad |\phi(z)| \leqslant \frac{\delta}{d^p(z)} \quad \textit{in } G.$$

(iii) $\alpha/p + (1 + (1/p))^{p+1}\Sigma\beta_i < 1/\eta$ *(this can always be achieved by diminishing η, if necessary).*

Then Equation (8) *has a unique solution u in G, and it satisfies $|u(t,z)| \leqslant C/d^p(t,z)$ in G.*

Proof. Let E be the Banach space of all functions $u \in C^0(G, \mathbb{C}^m)$, which are holomorphic in z and have a finite norm

$$\|u\| := \sup_G |u(t,z)| d^p(t,z).$$

Note that convergence in the norm implies uniform convergence in compact subsets of G; hence the limit is holomorphic in z and E is complete. We write Equation (8) in the form

$$u = g + Tu,$$

where T is a linear operator given by

(9) $$(Tu)(t,z) = \int_0^t \left[A(\tau,z)u(\tau,z) + \sum_i B_i(\tau,z) u_{z_i}(\tau,z) \right] d\tau.$$

It follows easily from the inequality

(10) $$\left| \int_0^t \frac{d\tau}{d^{p+1}(\tau,z)} \right| = \int_0^{|t|} \frac{ds}{(d(z) - s/\eta)^{p+1}} < \frac{\eta}{p d^p(t,z)}$$

that $g \in E$. According to the definition of the norm, we have

(11) $$|u(t,z)| \leqslant \frac{\|u\|}{d^p(t,x)}.$$

Nagumo's lemma, applied to the region Ω_t with distance function $d(t,z)$ instead of Ω and $d(z)$, gives the esimtate

$$|u_{z_j}(t,z)| \leqslant C_p \frac{\|u\|}{d^{p+1}(t,z)}. \tag{12}$$

It follows from the assumption (ii) together with (11) and (12) that

$$|Au| \leqslant \frac{\alpha\|u\|}{d^{p+1}(t,z)}, \; |B_j u_{z_j}| \leqslant \frac{\|u\|}{d^{p+1}(t,z)} \beta_j C_p,$$

hence, with $\beta = \Sigma\beta_j$,

$$|(Tu)(t,z)| \leqslant \|u\|(\alpha + \beta C_p)\left|\int_0^t \frac{d\tau}{d^{p+1}(\tau,z)}\right|$$
$$\leqslant \frac{1}{p}(\alpha + \beta C_p)\eta\frac{\|u\|}{d^p(t,z)}.$$

In the last step, inequality (10) was used. Going back to the definition of the norm, we arrive at the final estimate

$$\|Tu\| \leqslant q\|u\|, \qquad \text{where } q = \left(\frac{\alpha}{p} + \beta\left(1 + \frac{1}{p}\right)^{p+1}\right)\eta < 1, \tag{13}$$

according to assertion (iii). Hence the contraction principle applies to the equation $u = g + Tu$, and it follows that this equation has exactly one solution in E. This proves the theorem except for the possibility (which will be excluded below) that there are other solutions which do not belong to E. □

REMARK 1. It is seen from (iii) that the constant η, which describes the region of existence of the solution, depends only on $\beta = \Sigma_{j=1}^n \beta_j$. Indeed, for p large, α/p is small and $\left(1 + \frac{1}{p}\right)^{p+1}$ is close to e (note that α and β do not depend on p). Hence (iii) is satisfied (for large p) if $\eta < 1/(e\beta)$. Very simple examples show that this result is, except for the factor e, optimal. Let $u_t = bu_z$ and $u(0,z) = \phi(z)$ ($n = 1$, b constant). Then the solution is given by $u(t,z) = \phi(bt + z)$. If ϕ is holomorphic, say, in the circle $|z| < 1$, then the solution exists for $|z + bt| < 1$, and this is best possible when ϕ cannot be continued analytically beyond the unit circle. If we vary b, keeping $|b| = \beta$ fixed, then the largest region common to all those regions of existence is the circular cone $\beta|t| < 1 - |z|$, i.e., $\eta = 1/\beta$ is best possible.

REMARK 2. In order to understand the nature of the assumptions (ii), let us assume that $n = 1$ and that z_0 is an isolated boundary point of Ω, i.e., that $z \in \Omega$ and $d(z) = |z - z_0|$ for z near z_0. Then A is allowed to have a pole of the first order, while f and c may have a pole of any order at z_0. Now let Ω be unbounded (this case is usually not considered in the context of the C-K theorem). Then $d(t,z) \to \infty$ (perhaps) as $|z| \to \infty$, which implies that A, c and ϕ must vanish at infinity. If, on the other hand, assumption (ii) is satisfied, then the region of existence G extends to infinity in the t direction. By a slight change in the proof one can handle the case where A, c and ϕ do not vanish at infinity at the price of a smaller region of existence which is bounded in t. The change concerns the definition of $d(z)$. Instead of the distance to Γ we may take $d(z) = \min(\text{dist}(z,\Gamma), d_0)$, where d_0 is a positive constant, or more generally, any function with the properties

$$0 < d(z) \leqslant \text{dist}(z,\Gamma) \quad \text{and} \quad |d(z) - d(z')| \leqslant |z - z'| \quad \text{for } z, z' \in \Omega. \tag{14}$$

If G, Ω_t and $d(t,z)$ are defined as before, using the new function $d(z)$, then property (5) and its one-line proof remain valid, and $0 < d(t,z) \leqslant \text{dist}(z,\Gamma_t)$ for $z \in \Omega_t$. As a consequence, the lemma remains true, and the proof of the theorem goes through. If we choose a bounded function $d(z)$, then the inequalities in assumption (ii) can be satisfied as long as A, c and f remain bounded for $|z| \to \infty$.

REMARK 3. In the complex case it is assumed that the functions A, B_j and c are holomorphic in t and z. Otherwise the assumptions are unchanged. Theorem 1 remains true, and the proof goes through with the obvious change that E is now the Banach space of all functions u which are holomorphic in t and z and have a finite norm. Note that (10) remains true as it stands.

REMARK 4. In order to exclude the possibility that there are other solutions not belonging to E, let us assume that u^* is a solution to (6) (7) in some open domain $D \subset G$. Let $(0, z_0) \in D$, let $\Omega^* = B_r(z_0)$ be a ball in $\mathbb{C}^n$ with small radius r, and let G^* be the corresponding set in $\mathbb{R} \times \mathbb{C}$ defined by the inequalities $|t| < \eta d^*(z)$, where now $d^*(z) = r - |z - z_0|$. Applying our theorem to the set Ω^* we see that there is exactly one solution in G^* belonging to the corresponding Banach space E^*. But since our original solution u as well as the solution u^* (strictly speaking, their restrictions to G^*) belong to E^*, we arrive at $u = u^*$ in G^*. This argument, together with the fact that solutions are holomorphic in z, leads to $u = u^*$ at least in a small strip, say, for $(t, z) \in D$ and $|t| \leq \alpha$. But now the reasoning can be repeated with initial values given for $t = \pm\alpha$ instead of $t = 0$. In this way one obtains $u = u^*$ in D. In the complex case the last step is not necessary, since $u = u^*$ in the small implies equality in the large.

REMARK 5. Assume that the functions in Theorem 1 have the property that their components are real-valued for real values of z_j and t. Then the solution u has the same property. This well-known fact follows immediately from our proof. Indeed, the function $g(t, z) = \phi(z) + \int_0^t c(s, z)\,ds$ has the property in question. If v has the property, then also Tv. Hence, all members of the sequence (u_k) defined by $u_0 = g$, $u_{k+1} = g + Tu_k$ $(k = 0, 1, \dots)$ have the property, and so does the solution $u = \lim u_k$.

The fixed point of a contractive operator S is not only unique; it also depends continuously on S. Our next theorem deals with this important consequence of the contraction principle.

THEOREM 2. *Under the assumptions of Theorem* 1, *the solution u of the Cauchy problem* (6) (7) *depends continuously on the given functions A, B_j, c and f: For any given $\varepsilon > 0$, there exists $\delta > 0$ such that the solution u^* of a corresponding problem* (6) (7) *with A, B_j, c and ϕ replaced by corresponding starred quantities which satisfy the inequalities*

$$|A(t,z) - A^*(t,z)| \leq \frac{\delta}{d(t,z)}, \qquad |B_j(t,z) - B_j^*(t,z)| \leq \delta,$$

$$|c(t,z) - c^*(t,z)| \leq \frac{\delta}{d^{p+1}(t,z)}, \qquad |\phi(z) - \phi^*(z)| \leq \frac{\delta}{d^p(z)},$$

belongs to the ε-neighborhood of u (in E),

$$|u(t,z) - u^*(t,z)| \leq \frac{\varepsilon}{d^p(t,z)}.$$

Proof. The assertion follows easily from a standard estimate for contractive mappings. Let S be any contraction, $\|Su - Sv\| \leq q\|u - v\|$, and let u be the fixed point of S. Then

$$\|u - v\| \leq \|Su - Sv\| + \|Sv - v\| \leq q\|u - v\| + \|Sv - v\|,$$

which implies

$$\|u - v\| \leq \frac{1}{1-q}\|v - Sv\|. \tag{15}$$

As in the proof of Theorem 1, we write $u = g + Tu$ and analogously $u^* = g^* + T^*u^*$. Then (15) reads

$$\|u - u^*\| \leqslant \frac{1}{1-q}\left\{\|g - g^*\| + \|(T - T^*)u^*\|\right\}. \tag{16}$$

It follows easily from our assumptions and (10) that $\|g - g^*\| \leqslant \delta C$, $C = 1 + \eta/p$. The estimate (13), applied to $T - T^*$ (which means $\alpha = \beta = \delta$) shows that $\|T - T^*\| \leqslant \delta C_1$. It also follows from (13) that for $0 \leqslant \delta < \delta_0$, where δ_0 is small, $\|T^*\| \leqslant q^* < 1$ with q^* independent of δ. Hence u^* is bounded independent of δ (take, for example, (15) with u, v, q and S replaced by u^*, 0, q^* and $g^* + T^*$). Summing up, we have obtained an estimate $\|u - u^*\| \leqslant \delta C_2$ with C_2 independent of δ, which proves the theorem. □

REMARK. We come back to Hadamard's example ($a = 1$)

$$u_{tt} + u_{xx} = 0, \qquad u(0,x) = 0, \qquad u_t(0,x) = \sin kx,$$

or, in the form of a first order system ($v = u_t, w = u_x$),

$$u_t = v, \qquad v_t = -w_x, \qquad w_t = v_x,$$

$$u(0,x) = w(0,x) = 0, \qquad v(0,x) = \sin kx.$$

We replace x by z, choose the ball $|z| < \rho$ for Ω, and consider this example with regard to Theorems 1 and 2. It is easily seen that in Theorem 1 we have $|A| = 1$, $|B| = 1$, $\alpha = \beta = 1$; hence G is defined by $|t| < \eta(\rho - |z|)$ with $\eta < 1/e$. The solution is given by

$$u(t,z) = \frac{1}{k}\cdot \sin kz \cdot \sinh kt.$$

The maximum of $|\sin kz|$ or $|\cos kz|$ for $|z| \leqslant \sigma$ is approximately $\frac{1}{2}e^{k\sigma}$, the maximum of $|\sinh kt|$ or $|\cosh kt|$ for $|t| \leqslant \tau$ is approximately $\frac{1}{2}e^{k\tau}$, and in G we have $|t| + \eta|z| < \rho$, which implies $|t| + |z| < \rho$. Hence for $k \to \infty$ the solution (u, u_t, u_x) does not grow faster in G than on Ω.

3. The quasilinear case. As pointed out in the introduction, the case of a general nonlinear system reduces to that of a quasilinear system of the first order

$$u_t = \sum_{j=1}^{n} B_j(t,z,u)u_{z_j} + c(t,z,u). \tag{17}$$

As before, the B_j are m by m matrices, and c is a column vector of order m. This system is dealt with in our next theorem. As in the linear case, we describe a region of ascertained existence for the solution in terms of the given functions. But it is not our aim to give an optimal region, which would be somewhat cumbersome. The domain $G = G_\eta$ is defined as before by the inquality $|t| < \eta d(z)$, and B_R is the open ball $|z| < R$ in $\mathbb{C}^m$. In order to avoid complications as described in Remark 2, it is assumed from the beginning that the "distance function" $d(z)$ is bounded and, of course, satisfies (14) (for example, Ω bounded). If the initial condition is given by $u(0,z) = \phi(z)$ in Ω, we write the solution in the form $u = \phi + v$. This gives a new differential equation for v, which is of the same type as (17), and reduces the initial values to zero. We shall therefore restrict consideration to the latter case.

THEOREM 3. *Assume that the functions $B_j(t,z,u)$ and $c(t,z,u)$ are continuous in $G_\eta \times B_R$ and holomorphic with respect to z and u (real case), or holomorphic in $G_\eta \times B_R$ with respect to all variables (complex case), and that the following estimates hold there:*

$$|c| \leqslant \frac{\gamma}{\sqrt{d(t,z)}}, \qquad d(t,z)|c(t,z,u) - c(t,z,v)| \leqslant \gamma'|u - v|,$$

$$|B_j| \leqslant \beta_j, \qquad \sqrt{d(t,z)}\,|B_j(t,z,u) - B_j(t,z,v)| \leqslant \beta_j'|u - v|$$

($j = 1, \ldots, n$). If $\eta > 0$ is such that

$$2\eta\sqrt{d_0}(\beta + \gamma) < R, \qquad \eta\big(3\sqrt{3}(\beta + \gamma) + 2\beta\big) \leqslant 1 \quad \text{and} \quad \eta e\beta' < 1$$

with $d_0 = \sup d(z) < \infty$, $\beta = \sum \beta_j$ *and* $\beta' = \sum \beta_j'$, *then the quasilinear system* (17) *has a unique solution* u *satisfying the initial condition* $u(0, z) = 0$ *for* $z \in \Omega$ *and existing in* G_η.

Proof. As in the linear case we consider the corresponding integral equation

$$u(t,z) = \int_0^t \Big[\sum_j B_j(\tau, z, u) u_{z_j}(\tau, z) + c(\tau, z, u) \Big] d\tau \tag{18}$$

and treat it as a fixed point equation $u = Su$ to which the contraction principle will be applied. The underlying Banach space E is the same as in the proof of Theorem 1, and the nonlinear operator S is given by the right hand side of (18). In contrast to the linear case the operator S is not defined on all of E. In the first part of the proof we shall describe a proper closed subset F of E which is mapped into itself by S. We define F as the set of all functions $u \in E$ which satisfy the inequalities

$$|u(t,z)| \leqslant \rho \text{ and } |u_{z_k}(t,z)| \leqslant \frac{1}{\sqrt{d(t,z)}},$$

where $\rho = 2\eta\sqrt{d_0}(\beta + \gamma) < R$. Let $u \in F$ and $v = Su$. Then

$$|v_t(t,z)| \leqslant \sum_j |B_j||u_{z_j}| + |c| \leqslant \frac{\beta + \gamma}{\sqrt{d(t,z)}}.$$

Using the estimate

$$\int_0^{|t|} \frac{ds}{\sqrt{d(s,z)}} = -2\eta \sqrt{d(z) - \frac{s}{\eta}} \,\Bigg|_0^{|t|} \leqslant 2\eta\sqrt{d_0},$$

integration of v_t with respect to t yields

$$|v(t,z)| \leqslant (\beta + \gamma) 2\eta\sqrt{d_0} = \rho.$$

Now we estimate the derivatives of v, using the notation $c(u) = c(t, z, u(t,z))$ and similarly for B_j. Nagumo's lemma implies

$$\left| \frac{\partial}{\partial z_k} c(u) \right| \leqslant \frac{\gamma C}{d^{3/2}}, \qquad |u_{z_j z_k}| \leqslant \frac{C}{d^{3/2}}$$

and

$$\left| \frac{\partial}{\partial z_k} B_j(u) \right| \leqslant \frac{\beta_j}{d},$$

where $d = d(t,z)$ and $C = C_{1/2} = 3\sqrt{3}/2$. Using the product rule of differentiation, we obtain

$$|v_{t, z_k}| \leqslant \frac{1}{d^{3/2}} (\beta C + \beta + \gamma C)$$

and, after integration with respect to t (see (10))

$$|v_{z_k}| \leqslant \frac{2\eta}{\sqrt{d}} (\beta C + \beta + \gamma C) \leqslant \frac{1}{\sqrt{d}}.$$

The estimations above show that $u \in F$ implies $v = Su \in F$.

In the second part of the proof the difference $Su - Sv$ will be estimated for $u, v \in F$. In an abbreviated notation, we get

$$(Su - Sv)_t = \sum_j \big(B_j(u) - B_j(v)\big) u_{z_j} + \sum_j B_j(v)(u_{z_j} - v_{z_j}) + c(u) - c(v)$$

and

Since $|u - v| \leqslant$

and, after integ

are obtained. V
estimate

This inequality

REMARK 1.
there are other
considering su

REMARK 2.
of t, z and u.
approximation

REMARK 3.
the right hanc
follows the sar

REMARK 4.
compact. Hen
Schauder's fixe
in [11] and [8].

Let us note fi
theorem. It c
differential equ
with delay in
may consider

In this space p
find that the C
for more detai

1. C. Bessag
2. A. Cauch
partielles. -Mém
-Mémoire sur les
systèmes d'équati
17-33, 33-49 and
3. R. Courar

and

$$|(Su - Sv)_{,t}| \leqslant \sum_j \beta_j' \frac{|u - v|}{d} + \sum_j \beta_j |u_{z_j} - v_{z_j}| + \frac{\gamma'}{d}|u - v|.$$

Since $|u - v| \leqslant \|u - v\|/d^p$ and $|u_{z_j} - v_{z_j}| \leqslant C_p\|u - v\|/d^{p+1}$, the estimates

$$|(Su - Sv)_{,t}| \leqslant \frac{\|u - v\|}{d^{p+1}}\left(\beta' + \beta C_p + \gamma'\right)$$

and, after integration,

$$|Su - Sv| \leqslant \frac{\|u - v\|}{d^p}\frac{\eta}{p}\left(\beta' + C_p\beta + \gamma'\right)$$

are obtained. We multiply by d^p, take the supremum with respect to G_η, and arrive at the final estimate

$$\|Su - Sv\| \leqslant \frac{\eta}{p}\left(\beta' + C_p\beta + \gamma'\right)\|u - v\|.$$

This inequality shows that (for large p) S is a contraction and hence concludes the proof. □

REMARK 1. Remark 4 of Section 2 carries over to the quasilinear case. The possibility that there are other solutions to the Cauchy problem not belonging to E is carried ad absurdum by considering subregions of G_η.

REMARK 2. In a similar way, Remark 5 carries over. If B_j and c are real-valued for real values of t, z and u, then the solution u is real-valued for real t and z. This is shown by successive approximation $u_{k+1} = Su_k$, starting with $u_0 = 0$.

REMARK 3. As in the linear case, the solution depends continuously on the initial value and on the right hand side of the differential equation. The proof is based on the estimate (15) and follows the same lines as before.

REMARK 4. The subset F of E constructed in the first part of the proof is easily seen to be compact. Hence the existence of a solution follows already from the first part of the proof, if Schauder's fixed point theorem is called upon. This proof would have some similarity with those in [**11**] and [**8**].

Let us note finally that the method used in this article is not confined to the classical C-K theorem. It can easily be adapted to abstract versions and, more generally, to functional differential equations of the Cauchy-Kowalevsky type, for example to delay-differential equations with delay in the t-direction (real case). In the case where the set Ω is the whole space $\mathbb{C}^n$, one may consider the space of entire holomorphic functions with finite norm

$$\|f\| = \sup|f(z)|e^{-\alpha|z|} \qquad (z \in \mathbb{C}^n, \alpha > 0).$$

In this space partial differentiation is a bounded operation, $\|\partial f/\partial z_i\| \leqslant \alpha e\|f\|$. Using this fact, we find that the Cauchy problem can be treated in a very simple way; see Redheffer and Walter [**19**] for more details.

References

1. C. Bessaga, On the converse of the Banach "fixed-point principle", Colloq. Math., 7 (1959) 41–43.

2. A. Cauchy, Mémoire sur l'emploi du calcul des limites dans l'intégration des équations aux dérivées partielles. -Mémoire sur l'application du calcul des limites à l'intégration d'équations aux dérivées partielles. -Mémoire sur les systèmes d'équations aux dérivées partielles d'ordres quelconques, et sur leur réduction à des systèmes d'équations linéaires du premier ordre. Comptes Rend., 15 (1842) (Oeuvres complètes, 1. Série, Tome VII, 17-33, 33-49 and 52-58).

3. R. Courant and D. Hilbert, Methods of Mathematical Physics, vol. II, Interscience Publishers, 1962.

4. I. M. Gelfand and G. E. Silov, Generalized Functions, vol. 3: Theory of Differential Equations, Academic Press, New York, 1967.

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

6. L. Hörmander, Linear Partial Differential Operators, Springer Verlag, Berlin-Heidelberg-New York, 1963.

7. F. John, Partial Differential Equations, 3rd ed., Springer Verlag, New York-Heidelberg-Berlin, 1978.

8. K. Keller and A. Schneider, Ein funktionalanalytischer Beweis des Satzes von Cauchy-Kowalewsky, Manuscripta math., 39 (1982) 31–37.

9. S. v. Kowalevsky, Zur Theorie der partiellen Differentialgleichungen, J. Reine Angew. Math., 80 (1) (1875) 1–32.

10. S. Mizohata, The Theory of Partial Differential Equations, Cambridge Univ. Press, 1973.

11. M. Nagumo, Über das Anfangswertproblem Partieller Differentialgleichungen, Japan. J. Math., 18 (1941-43) 41–47.

12. L. Nirenberg, An abstract form of the nonlinear Cauchy-Kowalewski theorem, J. Differential Geom., 6 (1972) 561–576.

13. T. Nishida, A note on a theorem of Nirenberg, J. Differential Geom., 12 (1977) 629–633.

14. L. V. Ovsjannikov, Singular operators in Banach scales, Dokl. Akad. Nauk SSSR, 163 (1965) 819–822; Soviet Math. Dokl., 6 (1965) 1025–1028.

15. ______, A nonlinear Cauchy problem in a scale of Banach spaces, Dokl. Akad. Nauk SSSR, 200 (1971) 789–792; Soviet Math. Dokl., 12 (1971) 1497–1502.

16. ______, Abstract form of the Cauchy-Kowalewski theorem and its applications (Russian), Partial Differential Equations (Proc. Conf. Novosibirsk, 1978), 88–94, 250 "Nauka" Sibirsk, Otdel., Novosibirsk, 1980.

17. T. H. Pate, A Direct Iterative Method for an Abstract Cauchy-Kowalewsky Theorem, Indiana Univ. Math. J., 30 (3) (1981) 415–425.

18. I. G. Petrovski, Lectures on Partial Differential Equations, Interscience Publishers 1961 (third printing).

19. R. Redheffer and W. Walter, Existence theorems for strongly coupled systems of partial differential equations over Bernstein classes, Bull. Amer. Math. Soc., 82 (1976) 899–902.

20. F. Treves, On the theory of linear partial differential operators with analytic coefficients, Trans. Amer. Math. Soc., 137 (1969) 1–20.

21. ______, An abstract nonlinear Cauchy-Kovalevska theorem, Trans. Amer. Math. Soc., 150 (1970) 77–92.

22. ______, Basic Linear Partial Differential Equations, Academic Press, New York-San Francisco-London, 1975.

NOTES

EDITED BY SABRA S. ANDERSON, SHELDON AXLER, AND J. ARTHUR SEEBACH, JR.

For instructions about submitting Notes for publication in this department see the inside front cover.

THE GAMMA FUNCTION AND THE HURWITZ ZETA-FUNCTION

BRUCE C. BERNDT*

Department of Mathematics, University of Illinois, Urbana, IL 61801

The gamma function $\Gamma(x)$ may be defined by

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!(n+1)^x}{x(x+1)\cdots(x+n)}, \tag{1}$$

where x is any complex number, while the Hurwitz zeta-function $\zeta(s,x)$ is defined by

$$\zeta(s,x) = \sum_{k=0}^{\infty} (k+x)^{-s}, \tag{2}$$

*Research partially supported by National Science Foundation Grant No. MCS 7903359.