I.7 A General Existence Theorem

> M. Cauchy annonce, que, pour se conformer au voeu du Conseil, il ne s'attachera plus à donner, comme il a fait jusqu'à présent, des démonstrations parfaitement rigoureuses.
>
> (Conseil d'instruction de l'Ecole polytechnique, 24 nov. 1825)

> You have all professional deformation of your minds; *convergence* does not matter here ... (P. Henrici 1985)

We now enter a new era for our subject, more theoretical than the preceding one. It was inaugurated by the work of Cauchy, who was not as fascinated by long numerical calculations as was, say, Euler, but merely a fanatic for perfect mathematical rigor and exactness. He criticized in the work of his predecessors the use of infinite series and other infinite processes without taking much account of error estimates or convergence results. He therefore established around 1820 a convergence theorem for the polygon method of Euler and, some 15 years later, for the power series method of Newton (see Section I.8). Beyond the estimation of errors, these results also allow the statement of *general existence theorems* for the solutions of arbitrary differential equations ("d'une équation différentielle quelconque"), whose solutions were only known before in a very few cases. A second important consequence is to provide results about the *uniqueness* of the solution, which allow one to conclude that the computed solution (numerically or analytically) is the only one with the same initial value and that there are no others. Only then we are allowed to speak of *the* solution of the problem.

His very first proof has recently been discovered on fragmentary notes (Cauchy 1824), which were never published in Cauchy's lifetime (did his notes not satisfy the Minister of education?: "... mais que le second professeur, M. Cauchy, n'a présenté que des feuilles qui n'ont pu satisfaire la commission, et qu'il a été jusqu'à présent impossible de l'amener à se rendre au voeu du Conseil et à exécuter la décision du Ministre").

Convergence of Euler's Method

Let us now, with bared head and trembling knees, follow the ideas of this historical proof. We formulate it in a way which generalizes directly to higher dimensional systems.

Starting with the one-dimensional differential equation

$$y' = f(x,y), \qquad y(x_0) = y_0, \qquad y(X) = ? \tag{7.1}$$

we make use of the method explained by Euler (1768) in the last section of his "Institutiones Calculi Integralis I" (Caput VII, p. 424), i.e., we consider a subdivision of the interval of integration

$$x_0, x_1, \ldots, x_{n-1}, x_n = X \tag{7.2}$$

and replace in each subinterval the solution by the first term of its Taylor series

$$\begin{aligned} y_1 - y_0 &= (x_1 - x_0) f(x_0, y_0) \\ y_2 - y_1 &= (x_2 - x_1) f(x_1, y_1) \\ &\ldots \\ y_n - y_{n-1} &= (x_n - x_{n-1}) f(x_{n-1}, y_{n-1}). \end{aligned} \tag{7.3}$$

For the subdivision above we also use the notation

$$h = (h_0, h_1, \ldots, h_{n-1})$$

where $h_i = x_{i+1} - x_i$. If we connect y_0 and y_1, y_1 and $y_2, \ldots$ etc by straight lines we obtain the *Euler polygon*

$$y_h(x) = y_i + (x - x_i) f(x_i, y_i) \qquad \text{for} \qquad x_i \le x \le x_{i+1}. \tag{7.3a}$$

Lemma 7.1. *Assume that* $|f|$ *is bounded by* A *on*

$$D = \Big\{ (x, y) \mid x_0 \le x \le X,\ |y - y_0| \le b \Big\}.$$

If $X - x_0 \le b/A$ *then the numerical solution* (x_i, y_i) *given by (7.3), remains in* D *for every subdivision (7.2) and we have*

$$|y_h(x) - y_0| \le A \cdot |x - x_0|, \tag{7.4}$$

$$\Big| y_h(x) - \Big(y_0 + (x - x_0) f(x_0, y_0) \Big) \Big| \le \varepsilon \cdot |x - x_0| \tag{7.5}$$

if $|f(x, y) - f(x_0, y_0)| \le \varepsilon$ *on D.*

Proof. Both inequalities are obtained by adding up the lines of (7.3) and using the triangle inequality. Formula (7.4) then shows immediately that for $A(x - x_0) \le b$ the polygon remains in D. □

Our next problem is to obtain an estimate for the change of $y_h(x)$, when the initial value y_0 is changed: let z_0 be another initial value and compute

$$z_1 - z_0 = (x_1 - x_0) f(x_0, z_0). \tag{7.6}$$

We need an estimate for $|z_1 - y_1|$. Subtracting (7.6) from the first line of (7.3) we obtain

$$z_1 - y_1 = z_0 - y_0 + (x_1 - x_0) \Big(f(x_0, z_0) - f(x_0, y_0) \Big).$$

This shows that we need an estimate for $f(x_0, z_0) - f(x_0, y_0)$. If we suppose

$$|f(x, z) - f(x, y)| \le L |z - y| \tag{7.7}$$

we obtain

$$|z_1 - y_1| \le \big(1 + (x_1 - x_0)L\big)|z_0 - y_0|. \tag{7.8}$$

Lemma 7.2. *For a fixed subdivision h let $y_h(x)$ and $z_h(x)$ be the Euler polygons corresponding to the initial values y_0 and z_0, respectively. If*

$$\Big|\frac{\partial f}{\partial y}(x, y)\Big| \le L \tag{7.9}$$

in a convex region which contains $(x, y_h(x))$ and $(x, z_h(x))$ for all $x_0 \le x \le X$, then

$$|z_h(x) - y_h(x)| \le e^{L(x-x_0)}|z_0 - y_0|. \tag{7.10}$$

Proof. (7.9) implies (7.7), (7.7) implies (7.8), (7.8) implies

$$|z_1 - y_1| \le e^{L(x_1 - x_0)}|z_0 - y_0|.$$

If we repeat the same argument for $z_2 - y_2$, $z_3 - y_3$, and so on, we finally obtain (7.10). □

Remark. Condition (7.7) is called a "Lipschitz condition". It was Lipschitz (1876) who rediscovered the theory (footnote in the paper of Lipschitz: "L'auteur ne connaît pas évidemment les travaux de Cauchy ...") and advocated the use of (7.7) instead of the more stringent hypothesis (7.9). Lipschitz's proof is also explained in the classical work of Picard (1891-96), Vol. II, Chap. XI, Sec. I.

If the subdivision (7.2) is refined more and more, so that

$$|h| := \max_{i=0,\dots,n-1} h_i \to 0,$$

we expect that the Euler polygons converge to a solution of (7.1). Indeed, we have

Theorem 7.3. *Let $f(x, y)$ be continuous, and $|f|$ be bounded by A and satisfy the Lipschitz condition (7.7) on*

$$D = \Big\{(x, y) \mid x_0 \le x \le X,\ |y - y_0| \le b\Big\}.$$

If $X - x_0 \le b/A$, then we have:

a) For $|h| \to 0$ the Euler polygons $y_h(x)$ converge uniformly to a continuous function $\varphi(x)$.

b) $\varphi(x)$ is continuously differentiable and solution of (7.1) on $x_0 \le x \le X$.

c) There exists no other solution of (7.1) on $x_0 \le x \le X$.

Proof. a) Take an $\varepsilon > 0$. Since f is uniformly continuous on the compact set D, there exists a $\delta > 0$ such that

$$|u_1 - u_2| \le \delta \qquad \text{and} \qquad |v_1 - v_2| \le A \cdot \delta$$

imply

$$|f(u_1,v_1)-f(u_2,v_2)|\le\varepsilon. \tag{7.11}$$

Suppose now that the subdivision (7.2) satisfies

$$|x_{i+1}-x_i|\le\delta, \qquad \text{i.e.,} \qquad |h|\le\delta. \tag{7.12}$$

We first study the effect of adding new mesh-points. In a first step, we consider a subdivision $h(1)$, which is obtained by adding new points only to the *first* subinterval (see Fig. 7.1). It follows from (7.5) (applied to this first subinterval) that for the new refined solution $y_{h(1)}(x_1)$ we have the estimate $|y_{h(1)}(x_1)-y_h(x_1)|\le\varepsilon|x_1-x_0|$. Since the subdivisions h and $h(1)$ are identical on $x_1\le x\le X$ we can apply Lemma 7.2 to obtain

$$|y_{h(1)}(x)-y_h(x)|\le e^{L(x-x_1)}(x_1-x_0)\varepsilon \qquad \text{for} \quad x_1\le x\le X.$$

We next add further points to the subinterval (x_1,x_2) and denote the new subdivision by $h(2)$. In the same way as above this leads to $|y_{h(2)}(x_2)-y_{h(1)}(x_2)|\le\varepsilon|x_2-x_1|$ and

$$|y_{h(2)}(x)-y_{h(1)}(x)|\le e^{L(x-x_2)}(x_2-x_1)\varepsilon \qquad \text{for} \quad x_2\le x\le X.$$

The entire situation is sketched in Fig. 7.1. If we denote by $\widehat{h}$ the final refinement, we obtain for $x_i<x\le x_{i+1}$

$$\begin{aligned}
&|y_{\widehat{h}}(x)-y_h(x)| \\
&\qquad\le\varepsilon\Big(e^{L(x-x_1)}(x_1-x_0)+\ldots+e^{L(x-x_i)}(x_i-x_{i-1})\Big)+\varepsilon(x-x_i) \\
&\qquad\le\varepsilon\int_{x_0}^{x}e^{L(x-s)}\,ds=\frac{\varepsilon}{L}\Big(e^{L(x-x_0)}-1\Big).
\end{aligned} \tag{7.13}$$

If we now have two different subdivisions h and $\widetilde{h}$, which both satisfy (7.12), we introduce a *third* subdivision $\widehat{h}$ which is a refinement of both subdivisions (just as is usually done in proving the existence of Riemann's integral), and apply (7.13) twice. We then obtain from (7.13) by the triangle inequality

$$|y_h(x)-y_{\widetilde{h}}(x)|\le 2\,\frac{\varepsilon}{L}\Big(e^{L(x-x_0)}-1\Big).$$

For $\varepsilon>0$ small enough, this becomes arbitrarily small and shows the uniform convergence of the Euler polygons to a continuous function $\varphi(x)$.

b) Let

$$\varepsilon(\delta):=\sup\Big\{\big|f(u_1,v_1)-f(u_2,v_2)\big|\;;\;|u_1-u_2|\le\delta,\;|v_1-v_2|\le A\delta,\;(u_i,v_i)\in D\Big\}$$

be the modulus of continuity. If x belongs to the subdivision h then we obtain from (7.5) (replace (x_0,y_0) by $(x,y_h(x))$ and x by $x+\delta$)

$$|y_h(x+\delta)-y_h(x)-\delta f\big(x,y_h(x)\big)|\le\varepsilon(\delta)\delta. \tag{7.14}$$

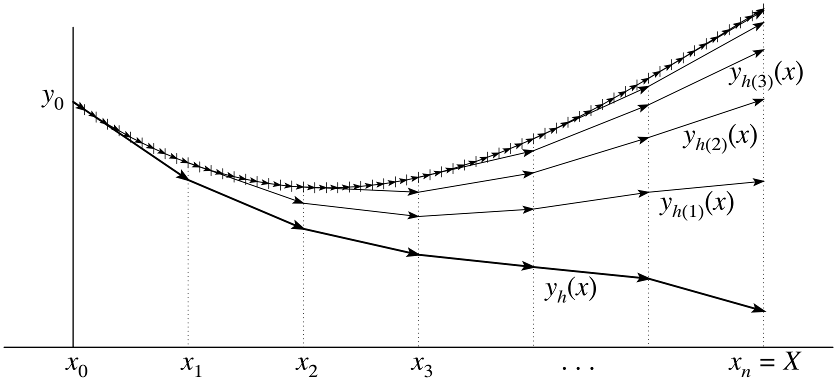

Fig. 7.1. Lady Windermere's Fan (O. Wilde 1892)

Taking the limit $|h| \to 0$ we get

$$|\varphi(x+\delta) - \varphi(x) - \delta f\big(x, \varphi(x)\big)| \le \varepsilon(\delta)\delta. \tag{7.15}$$

Since $\varepsilon(\delta) \to 0$ for $\delta \to 0$, this proves the differentiability of $\varphi(x)$ and $\varphi'(x) = f(x, \varphi(x))$.

c) Let $\psi(x)$ be a second solution of (7.1) and suppose that the subdivision h satisfies (7.12). We then denote by $y_h^{(i)}(x)$ the Euler polygon to the initial value $(x_i, \psi(x_i))$ (it is defined for $x_i \le x \le X$). It follows from

$$\psi(x) = \psi(x_i) + \int_{x_i}^{x} f\big(s, \psi(s)\big)\, ds$$

and (7.11) that

$$|\psi(x) - y_h^{(i)}(x)| \le \varepsilon |x - x_i| \qquad \text{for} \quad x_i \le x \le x_{i+1}.$$

Using Lemma 7.2 we deduce in the same way as in part a) that

$$|\psi(x) - y_h(x)| \le \frac{\varepsilon}{L}\Big(e^{L(x-x_0)} - 1\Big). \tag{7.16}$$

Taking the limits $|h| \to 0$ and $\varepsilon \to 0$ we obtain $|\psi(x) - \varphi(x)| \le 0$, proving uniqueness. □

Theorem 7.3 is a *local* existence - and uniqueness - result. However, if we interpret the endpoint of the solution as a new initial value, we can apply Theorem 7.3 again and continue the solution. Repeating this procedure we obtain

Theorem 7.4. *Assume U to be an open set in $\mathbb{R}^2$ and let f and $\partial f/\partial y$ be continuous on U. Then, for every $(x_0, y_0) \in U$, there exists a unique solution of (7.1), which can be continued up to the boundary of U (in both directions).*

Proof. Clearly, Theorem 7.3 can be rewritten to give a local existence - and uniqueness - result for an interval (X, x_0) to the left of x_0. The rest follows from the fact that every point in U has a neighbourhood which satisfies the assumptions of Theorem 7.3. □

It is interesting to mention that formula (7.13) for $|\widehat{h}| \to 0$ gives the following *error estimate*

$$|y(x) - y_h(x)| \leq \frac{\varepsilon}{L}\Big(e^{L(x-x_0)} - 1\Big) \tag{7.17}$$

for the Euler polygon $(|h| \leq \delta)$. Here $y(x)$ stands for the exact solution of (7.1). The next theorem refines the above estimates for the case that $f(x,y)$ is also differentiable with respect to x.

Theorem 7.5. *Suppose that in a neighbourhood of the solution*

$$|f| \leq A, \qquad \left|\frac{\partial f}{\partial y}\right| \leq L, \qquad \left|\frac{\partial f}{\partial x}\right| \leq M.$$

We then have the following error estimate for the Euler polygons:

$$\big|y(x) - y_h(x)\big| \leq \frac{M + AL}{L}\Big(e^{L(x-x_0)} - 1\Big) \cdot |h|, \tag{7.18}$$

provided that $|h|$ is sufficiently small.

Proof. For $|u_1 - u_2| \leq |h|$ and $|v_1 - v_2| \leq A|h|$ we obtain, due to the differentiability of f, the estimate

$$|f(u_1, v_1) - f(u_2, v_2)| \leq (M + AL)|h|$$

instead of (7.11). When we insert this amount for ε into (7.16), we obtain the stated result. □

The estimate (7.18) shows that the global error of Euler's method is proportional to the maximal step size $|h|$. Thus, for an accuracy of, say, three decimal digits, we would need about a thousand steps; a precision of six digits will normally require a million steps etc. We see thus that the present method is not recommended for computations of high precision. In fact, the main subject of Chapter II will be to find methods which converge faster.

Existence Theorem of Peano

> Si a est un complexe d'ordre n, et b un nombre réel, alors on peut déterminer b' et f, où b' est une quantité plus grande que b, et f est un signe de fonction qui à chaque nombre de l'intervalle de b à b' fait correspondre un complexe (en d'autres mots, ft est un complexe fonction de la variable réelle t, définie pour toutes les valeurs de l'intervalle (b, b')); la valeur de ft pour $t = b$ est a; et dans tout l'intervalle (b, b') cette fonction ft satisfait à l'équation différentielle donnée. (Original version of Peano's Theorem)

The Lipschitz condition (7.7) is a crucial tool in the proof of (7.10) and finally of the Convergence Theorem. If we completely abandon condition (7.7) and only require that $f(x, y)$ be continuous, the convergence of the Euler polygons is no longer guaranteed.

An example, plotted in Fig. 7.2, is given by the equation

$$y' = 4\Big(\operatorname{sign}(y)\sqrt{|y|} + \max\Big(0, x - \frac{|y|}{x}\Big) \cdot \cos\Big(\frac{\pi \log x}{\log 2}\Big)\Big) \tag{7.19}$$

with $y(0) = 0$. It has been constructed such that

$$f(h, 0) = 4(-1)^i h \qquad \text{for } h = 2^{-i},$$
$$f(x, y) = 4 \operatorname{sign}(y) \cdot \sqrt{|y|} \qquad \text{for } |y| \geq x^2.$$

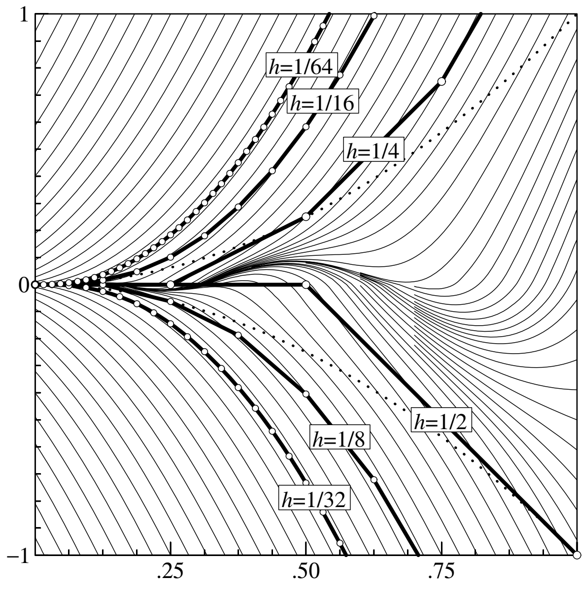

Fig. 7.2. Solution curves and Euler polygons for equation (7.19)

There is an infinity of solutions for this initial value, some of which are plotted in Fig. 7.2. The Euler polygons converge for $h = 2^{-i}$ and even i to the maximal solution $y = 4x^2$, and for odd i to $y = -4x^2$. For other sequences of h all intermediate solutions can be obtained as well.

Theorem 7.6 (Peano 1890). *Let $f(x,y)$ be continuous and $|f|$ be bounded by A on*

$$D = \Big\{(x,y) \mid x_0 \le x \le X,\ |y - y_0| \le b\Big\}.$$

If $X - x_0 \le b/A$, then there is a subsequence of the sequence of the Euler polygons which converges to a solution of the differential equation.

The original proof of Peano is, in its crucial part on the convergence result, very brief and not clear to unexperienced readers such as us. Arzelà (1895), who took up the subject again, explains his ideas in more detail and emphasizes the need for an *equicontinuity* of the sequence. The proof usually given nowadays (for what has become the theorem of Arzelà-Ascoli), was only introduced later (see e.g. Perron (1918), Hahn (1921), p. 303) and is sketched as follows:

Proof. Let

$$v_1(x), v_2(x), v_3(x), \ldots \tag{7.20}$$

be a sequence of Euler polygons for decreasing step sizes. It follows from (7.4) that for fixed x this sequence is bounded. We choose a sequence of numbers $r_1, r_2, r_3, \ldots$ dense in the interval (x_0, X). There is now a subsequence of (7.20) which converges for $x = r_1$ (Bolzano-Weierstrass), say

$$v_1^{(1)}(x), v_2^{(1)}(x), v_3^{(1)}(x), \ldots \tag{7.21}$$

We next select a subsequence of (7.21) which converges for $x = r_2$

$$v_1^{(2)}(x), v_2^{(2)}(x), v_3^{(2)}(x), \ldots \tag{7.22}$$

and so on. Then take the "diagonal" sequence

$$v_1^{(1)}(x), v_2^{(2)}(x), v_3^{(3)}(x), \ldots \tag{7.23}$$

which, apart from a finite number of terms, is a subsequence of each of these sequences, and thus converges for all r_i. Finally, with the estimate

$$|v_n^{(n)}(x) - v_n^{(n)}(r_j)| \le A|x - r_j|$$

(see (7.4)), which expresses the equicontinuity of the sequence, we obtain

$$\begin{aligned} &|v_n^{(n)}(x) - v_m^{(m)}(x)| \\ &\qquad \le |v_n^{(n)}(x) - v_n^{(n)}(r_j)| + |v_n^{(n)}(r_j) - v_m^{(m)}(r_j)| + |v_m^{(m)}(r_j) - v_m^{(m)}(x)| \\ &\qquad \le 2A|x - r_j| + |v_n^{(n)}(r_j) - v_m^{(m)}(r_j)|. \end{aligned}$$

For fixed $\varepsilon > 0$ we then choose a finite subset R of $\{r_1, r_2, \ldots\}$ satisfying

$$\min\{|x - r_j|\ ;\ r_j \in R,\ x_0 \le x \le X\} \le \varepsilon/A$$

and secondly we choose N such that

$$|v_n^{(n)}(r_j) - v_m^{(m)}(r_j)| \le \varepsilon \qquad \text{for} \quad n, m \ge N \quad \text{and} \quad r_j \in R.$$

This shows the uniform convergence of (7.23). In the same way as in part b) of the proof of Theorem 7.3 it follows that the limit function is a solution of (7.1). One only has to add an $\mathcal{O}(|h|)$-term in (7.14), if x is not a subdivision point. □

Exercises

1. Apply Euler's method with constant step size $x_{i+1} - x_i = 1/n$ to the differential equation $y' = ky$, $y(0) = 1$ and obtain a classical approximation for the solution $y(1) = e^k$. Give an estimate of the error.

2. Apply Euler's method with constant step size to

   a) $y' = y^2$, $y(0) = 1$, $y(1/2) =?$

   b) $y' = x^2 + y^2$, $y(0) = 0$, $y(1/2) =?$

   Make rigorous error estimates using Theorem 7.4 and compare these estimates with the actual errors. The main difficulty is to find a suitable region in which the estimates of Theorem 7.4 hold, without making the constants A, L, M too large and, at the same time, ensuring that the solution curves remain inside this region (see also I.8, Exercise 3).

3. Prove the result: if the differential equation $y' = f(x, y)$, $y(x_0) = y_0$ with f continuous, possesses a unique solution, then the Euler polygons converge to this solution.

4. "There is an elementary proof of Peano's existence theorem" (Walter 1971). Suppose that A is a bound for $|f|$. Then the sequence

   $$y_{i+1} = y_i + h \cdot \max\{f(x, y) | x_i \le x \le x_{i+1}, y_i - 3Ah \le y \le y_i + Ah\}$$

   converges for all continuous f to a (the maximal) solution. Try to prove this. Unfortunately, this proof does not extend to systems of equations, unless they are "quasimonotone" (see Section I.10, Exercise 3).

I.8 Existence Theory using Iteration Methods and Taylor Series

A second approach to existence theory is possible with the help of an iterative refinement of approximate solutions. The first appearances of the idea are very old. For instance many examples of this type can be found in the work of Lagrange, above all in his astronomical calculations. Let us consider here the following illustrative example of a Riccati equation

$$y' = x^2 + y + 0.1y^2, \qquad y(0) = 0. \tag{8.1}$$

Because of the quadratic term, there is no elementary solution. A very natural idea is therefore to neglect this term, which is in fact very small at the beginning, and to solve for the moment

$$y_1' = x^2 + y_1, \qquad y_1(0) = 0. \tag{8.2}$$

This gives, with formula (3.3), a first approximation

$$y_1(x) = 2e^x - (x^2 + 2x + 2). \tag{8.3}$$

With the help of this solution, we now know more about the initially neglected term $0.1y^2$; it will be close to $0.1y_1^2$. So the idea lies at hand to reintroduce this solution into (8.1) and solve now the differential equation

$$y_2' = x^2 + y_2 + 0.1 \cdot \big(y_1(x)\big)^2, \qquad y_2(0) = 0. \tag{8.4}$$

We can use formula (3.3) again and obtain after some calculations

$$\begin{aligned} y_2(x) = y_1(x) + \frac{2}{5}e^{2x} - \frac{2}{15}e^x(x^3 + 3x^2 + 6x - 54) \\ - \frac{1}{10}(x^4 + 8x^3 + 32x^2 + 72x + 76). \end{aligned}$$

This is already much closer to the correct solution, as can be seen from the following comparison of the errors $e_1 = y(x) - y_1(x)$ and $e_2 = y(x) - y_2(x)$:

$$\begin{array}{lll} x = 0.2 & e_1 = 0.228 \times 10^{-07} & e_2 = 0.233 \times 10^{-12} \\ x = 0.4 & e_1 = 0.327 \times 10^{-05} & e_2 = 0.566 \times 10^{-09} \\ x = 0.8 & e_1 = 0.534 \times 10^{-03} & e_2 = 0.165 \times 10^{-05}. \end{array}$$

It looks promising to continue this process, but the computations soon become very tedious.

Picard-Lindelöf Iteration

The general formulation of the method is the following: we try, if possible, to split up the function $f(x,y)$ of the differential equation

$$y' = f(x,y) = f_1(x,y) + f_2(x,y), \qquad y(x_0) = y_0 \tag{8.5}$$

so that any differential equation of the form $y' = f_1(x,y) + g(x)$ can be solved analytically and so that $f_2(x,y)$ is small. Then we start with a first approximation $y_0(x)$ and compute successively $y_1(x),\ y_2(x), \ldots$ by solving

$$y'_{i+1} = f_1(x, y_{i+1}) + f_2\big(x, y_i(x)\big), \qquad y_{i+1}(x_0) = y_0. \tag{8.6}$$

The most primitive form of this process is obtained by choosing $f_1 = 0,\ f_2 = f$, in which case (8.6) is immediately integrated and becomes

$$y_{i+1}(x) = y_0 + \int_{x_0}^{x} f\big(s, y_i(s)\big)\, ds. \tag{8.7}$$

This is called the *Picard-Lindelöf iteration method.* It appeared several times in the literature, e.g., in Liouville (1838), Cauchy, Peano (1888), Lindelöf (1894), Bendixson (1893). Picard (1890) considered it merely as a by-product of a similar idea for partial differential equations and analyzed it thoroughly in his famous treatise Picard (1891-96), Vol. II, Chap. XI, Sect. III.

The fast *convergence* of the method, for $|x - x_0|$ small, is readily seen: if we subtract formula (8.7) from the same with i replaced by $i-1$, we have

$$y_{i+1}(x) - y_i(x) = \int_{x_0}^{x} \Big(f\big(s, y_i(s)\big) - f\big(s, y_{i-1}(s)\big) \Big)\, ds. \tag{8.8}$$

We now apply the Lipschitz condition (7.7) and the triangle inequality to obtain

$$|y_{i+1}(x) - y_i(x)| \le L \int_{x_0}^{x} |y_i(s) - y_{i-1}(s)|\, ds. \tag{8.9}$$

When we assume $y_0(x) \equiv y_0$, the triangle inequality applied to (8.7) with $i = 0$ yields the estimate

$$|y_1(x) - y_0(x)| \le A|x - x_0|$$

where A is a bound for $|f|$ as in Section I.7. We next insert this into the right hand side of (8.9) repeatedly to obtain finally the estimate (Lindelöf 1894)

$$|y_i(x) - y_{i-1}(x)| \le A L^{i-1} \frac{|x - x_0|^i}{i!}. \tag{8.10}$$

The right-hand side is a term of the Taylor series for $e^{L|x-x_0|}$, which converges for all x; we therefore conclude that $|y_{i+k}-y_i|$ becomes arbitrarily small when i is large. The error is bounded by the remainder of the above exponential series. So the sequence $y_i(x)$ converges uniformly to the solution $y(x)$. For example, if $L|x-x_0|\le 1/10$ and the constant A is moderate, 10 iterations would provide a numerical solution with about 17 correct digits.

The main practical drawback of the method is the need for repeated computation of integrals, which is usually not very convenient, if at all analytically possible, and soon becomes very tedious. However, its fast convergence and new machine architectures (parallelism) coupled with numerical evaluations of the integrals have made the approach interesting for large problems (see Nevanlinna 1989).

Taylor Series

Après avoir montré l'insuffisance des méthodes d'intégration fondées sur le développement en séries, il me reste à dire en peu de mots ce qu'on peut leur substituer. (Cauchy)

A third existence proof can be based on a study of the convergence of the Taylor series of the solutions. This was mentioned in a footnote of Liouville (1836, p. 255), and brought to perfection by Cauchy (1839-42).

We have already seen the recursive computation of the Taylor coefficients in the work of Newton (see Section I.2). Euler (1768) then formulated the general procedure for the higher derivatives of the solution of

$$y'=f(x,y),\quad y(x_0)=y_0 \tag{8.11}$$

which, by successive differentiation, are obtained as

$$\begin{aligned} y'' &= f_x+f_yy'=f_x+f_yf \\ y''' &= f_{xx}+2f_{xy}f+f_{yy}f^2+f_y(f_x+f_yf) \end{aligned} \tag{8.12}$$

etc. Then the solution is

$$y(x_0+h)=y(x_0)+y'(x_0)h+y''(x_0)\frac{h^2}{2!}+\ldots . \tag{8.13}$$

The formulas (8.12) for higher derivatives soon become very complicated. Euler therefore proposed to use only a few terms of this series with h sufficiently small and to repeat the computations from the point $x_1=x_0+h$ ("analytic continuation").

We shall now outline the main ideas of Cauchy's *convergence proof* for the series (8.13). We suppose that $f(x,y)$ is *analytic* in the neighbourhood of the initial value x_0,y_0, which for simplicity of notation we assume located at the origin $x_0=y_0=0$:

$$f(x,y)=\sum_{i,j\ge 0} a_{ij}x^iy^j, \tag{8.14}$$

where the a_{ij} are multiples of the partial derivatives occurring in (8.12). If the series (8.14) is assumed to converge for $|x| \le r$, $|y| \le r$, then the Cauchy inequalities from classical complex analysis give

$$|a_{ij}| \le \frac{M}{r^{i+j}}, \qquad \text{where} \qquad M = \max_{|x|\le r, |y|\le r} |f(x,y)|. \tag{8.15}$$

The idea is now the following: since all signs in (8.12) are positive, we obtain the worst possible result if we replace in (8.14) all a_{ij} by the largest possible values (8.15) ("method of majorants"):

$$f(x,y) \;\to\; \sum_{i,j\ge 0} M\,\frac{x^i y^j}{r^{i+j}} = \frac{M}{(1-x/r)(1-y/r)}.$$

However, the majorizing differential equation

$$y' = \frac{M}{(1-x/r)(1-y/r)}, \qquad y(0)=0$$

is readily integrated by separation of variables (see Section I.3) and has the solution

$$y = r\left(1 - \sqrt{1 + 2M\log\left(1-\frac{x}{r}\right)}\;\right). \tag{8.16}$$

This solution has a power series expansion which converges for all x such that $|2M\log(1-x/r)| < 1$. Therefore, the series (8.13) also converges at least for all $|h| < r\big(1-\exp(-1/2M)\big)$. □

Recursive Computation of Taylor Coefficients

. . . dieses Verfahren praktisch nicht in Frage kommen kann.
(Runge & König 1924)

The exact opposite is true, if we use the right approach . . .
(R.E. Moore 1979)

The "right approach" is, in fact, an extension of Newton's approach and has been rediscovered several times (e.g,. Steffensen 1956) and implemented into computer programs by Gibbons (1960) and Moore (1966). For a more extensive bibliography see the references in Wanner (1969), p. 10-20.

The idea is the following: let

$$Y_i = \frac{1}{i!}\,y^{(i)}(x_0), \qquad F_i = \frac{1}{i!}\Big(f\big(x,y(x)\big)\Big)^{(i)}\Big|_{x=x_0} \tag{8.17}$$

be the Taylor coefficients of $y(x)$ and of $f\big(x,y(x)\big)$, so that (8.13) becomes

$$y(x_0+h) = \sum_{i=0}^{\infty} h^i Y_i.$$

Then, from (8.11),

$$Y_{i+1} = \frac{1}{i+1} F_i. \tag{8.18}$$

Now suppose that $f(x,y)$ is the composition of a sequence of algebraic operations and elementary functions. This leads to a sequence of items,

$$x, y, p, q, r, \ldots, \text{ and finally } f. \tag{8.19}$$

For each of these items we find formulas for generating the ith Taylor coefficient from the preceding ones as follows:

a) $r = p \pm q$:

$$R_i = P_i \pm Q_i, \qquad i = 0, 1, \ldots \tag{8.20a}$$

b) $r = pq$: the Cauchy product yields

$$R_i = \sum_{j=0}^{i} P_j Q_{i-j}, \qquad i = 0, 1, \ldots \tag{8.20b}$$

c) $r = p/q$: write $p = rq$, use formula b) and solve for R_i:

$$R_i = \frac{1}{Q_0}\Big(P_i - \sum_{j=0}^{i-1} R_j Q_{i-j}\Big), \qquad i = 0, 1, \ldots \tag{8.20c}$$

There also exist formulas for many elementary functions (in fact, because these functions are themselves solutions of rational differential equations).

d) $r = \exp(p)$: use $r' = p' \cdot r$ and apply (8.20b). This gives for $i = 1, 2, \ldots$

$$R_0 = \exp(P_0), \qquad R_i = \frac{1}{i} \sum_{j=0}^{i-1} (i-j) R_j P_{i-j}. \tag{8.20d}$$

e) $r = \log(p)$: use $p = \exp(r)$ and rearrange formula d). This gives

$$R_0 = \log(P_0), \qquad R_i = \frac{1}{P_0}\Big(P_i - \frac{1}{i} \sum_{j=1}^{i-1} (i-j) P_j R_{i-j}\Big). \tag{8.20e}$$

f) $r = p^c, c \neq 1$ constant. Use $pr' = crp'$ and apply (8.20b):

$$R_0 = P_0^c, \qquad R_i = \frac{1}{iP_0}\Big(\sum_{j=0}^{i-1} \big(ci - (c+1)j\big) R_j P_{i-j}\Big). \tag{8.20f}$$

g) $r = \cos(p)$, $s = \sin(p)$: as in d) we have

$$
\begin{aligned}
R_0 &= \cos P_0, \qquad & R_i &= -\frac{1}{i}\sum_{j=0}^{i-1}(i-j)S_j P_{i-j}, \\
S_0 &= \sin P_0, \qquad & S_i &= \frac{1}{i}\sum_{j=0}^{i-1}(i-j)R_j P_{i-j}.
\end{aligned}
\tag{8.20g}
$$

The alternating use of (8.20) and (8.18) then allows us to compute the Taylor coefficients for (8.17) to any wanted order in a very economical way. It is not difficult to write subroutines for the above formulas, which have to be called in the same order as the differential equation (8.11) is composed of elementary operations. There also exist computer programs which "compile" Fortran statements for $f(x,y)$ into this list of subroutine calls. One has been written by T. Szymanski and J.H. Gray (see Knapp & Wanner 1969).

Example. The differential equation $y' = x^2 + y^2$ leads to the recursion

$$
Y_0 = y(0), \qquad Y_{i+1} = \frac{1}{i+1}\Big(P_i + \sum_{j=0}^{i} Y_j Y_{i-j}\Big), \qquad i = 0, 1, \ldots
$$

where $P_i = 1$ for $i = 2$ and $P_i = 0$ for $i \neq 2$ are the coefficients for x^2. One can imagine how much easier this is than formulas (8.12).

An important property of this approach is that it can be executed in *interval analysis* and thus allows us to obtain *reliable error bounds* by the use of Lagrange's error formula for Taylor series. We refer to the books by R.E. Moore (1966) and (1979) for more details.

Exercises

1. Obtain from (8.10) the estimate

$$
|y_i(x) - y_0| \leq \frac{A}{L}\Big(e^{L(x-x_0)} - 1\Big)
$$

and explain the similarity of this result with (7.16).

2. Apply the method of Picard to the problem $y' = Ky$, $y(0) = 1$.

3. Compute three Picard iterations for the problem $y' = x^2 + y^2$, $y(0) = 0$, $y(1/2) = ?$ and make a rigorous error estimate. Compare the result with the correct solution $y(1/2) = 0.041791146154681863220768806849179$.

4. Compute with an iteration method the solution of

$$y' = \sqrt{x} + \sqrt{y}, \qquad y(0) = 0$$

and observe that the method can work well for equations which pose serious problems with other methods. An even greater difference occurs for the equations

$$y' = \sqrt{x} + y^2, \quad y(0) = 0 \qquad \text{and} \qquad y' = \frac{1}{\sqrt{x}} + y^2, \quad y(0) = 0.$$

5. Define $f(x,y)$ by

$$f(x,y) = \begin{cases} 0 & \text{for } x \le 0 \\ 2x & \text{for } x > 0,\ y < 0 \\ 2x - \dfrac{4y}{x} & \text{for } 0 \le y \le x^2 \\ -2x & \text{for } x > 0,\ x^2 < y. \end{cases}$$

   a) Show that $f(x,y)$ is continuous, but not Lipschitz.
   b) Show that for the problem $y' = f(x,y)$, $y(0) = 0$ the Picard iteration method does not converge.
   c) Show that there is a unique solution and that the Euler polygons converge.

6. Use the method of Picard iteration to prove: if $f(x,y)$ is continuous and satisfies a Lipschitz condition (7.7) on the infinite strip $D = \{(x,y)\,;\, x_0 \le x \le X\}$, then the initial value problem $y' = f(x,y)$, $y(x_0) = y_0$ possesses a unique solution on $x_0 \le x \le X$.

   Compare this global result with Theorem 7.3.

7. Define a function $y(x)$ (the "inverse error function") by the relation

$$x = \frac{2}{\sqrt{\pi}} \int_0^y e^{-t^2}\, dt$$

and show that it satisfies the differential equation

$$y' = \frac{\sqrt{\pi}}{2} e^{y^2}, \qquad y(0) = 0.$$

Obtain recursion formulas for its Taylor coefficients.

I.9 Existence Theory for Systems of Equations

The first treatment of an existence theory for simultaneous systems of differential equations was undertaken in the last existing pages (p. 123-136) of Cauchy (1824). We write the equations as

$$\begin{array}{lll} y_1' = f_1(x, y_1, \ldots, y_n), & y_1(x_0) = y_{10}, & y_1(X) = ? \\ \ldots & \ldots & \ldots \\ y_n' = f_n(x, y_1, \ldots, y_n), & y_n(x_0) = y_{n0}, & y_n(X) = ? \end{array} \tag{9.1}$$

and ask for the existence of the n solutions $y_1(x), \ldots, y_n(x)$. It is again natural to consider, in analogy to (7.3), the method of Euler

$$y_{k,i+1} = y_{ki} + (x_{i+1} - x_i) \cdot f_k(x_i, y_{1i}, \ldots, y_{ni}) \tag{9.2}$$

(for $k = 1, \ldots, n$ and $i = 0, 1, 2, \ldots$). Here y_{ki} is intended to approximate $y_k(x_i)$, where $x_0 < x_1 < x_2 \ldots$ is a subdivision of the interval of integration as in (7.2).

We now try to carry over everything we have done in Section I.7 to the new situation. Although we have no problem in extending (7.4) to the estimate

$$|y_{ki} - y_{k0}| \le A_k |x_i - x_0| \qquad \text{if} \qquad |f_k(x, y_1, \ldots, y_n)| \le A_k, \tag{9.3}$$

things become a little more complicated for (7.7): we have to estimate

$$f_k(x, z_1, \ldots, z_n) - f_k(x, y_1, \ldots, y_n) = \frac{\partial f_k}{\partial y_1} \cdot (z_1 - y_1) + \ldots + \frac{\partial f_k}{\partial y_n} \cdot (z_n - y_n), \tag{9.4}$$

where the derivatives $\partial f_k / \partial y_i$ are taken at suitable intermediate points. Here Cauchy uses the inequality now called the "Cauchy-Schwarz inequality" ("Enfin, il résulte de la formule (13) de la 11e leçon du calcul différentiel ...") to obtain

$$\begin{aligned} &|f_k(x, z_1, \ldots, z_n) - f_k(x, y_1, \ldots, y_n)| \\ &\quad \le \sqrt{\Big(\frac{\partial f_k}{\partial y_1}\Big)^2 + \ldots + \Big(\frac{\partial f_k}{\partial y_n}\Big)^2} \cdot \sqrt{(z_1 - y_1)^2 + \ldots + (z_n - y_n)^2}. \end{aligned} \tag{9.5}$$

At this stage, we begin to feel that further development is advisable only after the introduction of vector notation.

Vector Notation

This was promoted in our subject by the papers of Peano, (1888) and (1890), who was influenced, as he says, by the famous "Ausdehnungslehre" of Grassmann and the work of Hamilton, Cayley, and Sylvester. We introduce the vectors (Peano called them "complexes")

$$y = (y_1, \ldots, y_n)^T, \quad y_i = (y_{1i}, \ldots, y_{ni})^T, \quad z = (z_1, \ldots, z_n)^T \quad \text{etc},$$

and hope that the reader will not confuse the components y_i of a vector y with vectors with indices. We consider the "vector function"

$$f(x, y) = \big(f_1(x, y), \ldots, f_n(x, y)\big)^T,$$

so that equations (9.1) become

$$y' = f(x, y), \qquad y(x_0) = y_0, \qquad y(X) = ?, \tag{9.1'}$$

Euler's method (9.2) is

$$y_{i+1} = y_i + (x_{i+1} - x_i) f(x_i, y_i), \qquad i = 0, 1, 2, \ldots \tag{9.2'}$$

and the Euler polygon is given by

$$y_h(x) = y_i + (x - x_i) f(x_i, y_i) \qquad \text{for} \qquad x_i \le x \le x_{i+1}.$$

There is no longer any difference in notation with the one-dimensional cases (7.1), (7.3) and (7.3a).

In view of estimate (9.5), we introduce for a vector $y = (y_1, \ldots, y_n)^T$ the *norm* (originally "modulus")

$$\|y\| = \sqrt{y_1^2 + \ldots + y_n^2} \tag{9.6}$$

which satisfies all the usual properties of a norm, for example the triangle inequality

$$\|y + z\| \le \|y\| + \|z\|, \qquad \Big\|\sum_{i=1}^{n} y_i\Big\| \le \sum_{i=1}^{n} \|y_i\|. \tag{9.7}$$

The Euclidean norm (9.6) is not the only one possible, we also use ("on pourrait aussi définir par mx la plus grande des valeurs absolues des élements de x; alors les propriétes des modules sont presqu'évidentes.", Peano)

$$\|y\| = \max(|y_1|, \ldots, |y_n|), \tag{9.6'}$$

$$\|y\| = |y_1| + \ldots + |y_n|. \tag{9.6''}$$

We are now able to formulate estimate (9.3) as follows, in perfect analogy with (7.4): if for some norm $\|f(x, y)\| \le A$ on $D = \{(x, y) \mid x_0 \le x \le X, \|y - y_0\| \le b\}$ and if $X - x_0 \le b/A$ then the numerical solution (x_i, y_i), given by (9.2'), remains in D and we have

$$\|y_h(x) - y_0\| \le A \cdot |x - x_0|. \tag{9.8}$$

The analogue of estimate (7.5) can be obtained similarly.

In order to prove the implication “(7.9) $\Rightarrow$ (7.7)” for vector-valued functions it is convenient to work with norms of matrices.

Subordinate Matrix Norms

The relation (9.4) shows that the difference $f(x,z)-f(x,y)$ can be written as the product of a matrix with the vector $z-y$. It is therefore of interest to estimate $\|Qv\|$ and to find the best possible estimate of the form $\|Qv\| \le \beta\|v\|$.

Definition 9.1. Let Q be a matrix (n columns, m rows) and $\|\ldots\|$ be one of the norms defined in (9.6), (9.6') or (9.6"). The *subordinate matrix norm* of Q is then defined by

$$\|Q\| = \sup_{v\neq 0} \frac{\|Qv\|}{\|v\|} = \sup_{\|u\|=1} \|Qu\|. \tag{9.9}$$

By definition, $\|Q\|$ is the smallest number such that

$$\|Qv\| \le \|Q\|\cdot\|v\| \qquad \text{for all } v \tag{9.10}$$

holds. The following theorem gives explicit formulas for the computation of (9.9).

Theorem 9.2. *The norm of a matrix Q is given by the following formulas: for the Euclidean norm (9.6),*

$$\|Q\| = \sqrt{\textit{largest eigenvalue of } Q^TQ}\,; \tag{9.11}$$

for the max-norm (9.6'),

$$\|Q\| = \max_{k=1,\ldots,m}\Big(\sum_{i=1}^{n}|q_{ki}|\Big); \tag{9.11'}$$

for the norm (9.6"),

$$\|Q\| = \max_{i=1,\ldots,n}\Big(\sum_{k=1}^{m}|q_{ki}|\Big). \tag{9.11''}$$

Proof. Formula (9.11) can be seen from $\|Qv\|^2 = v^TQ^TQv$ with the help of an orthogonal transformation of Q^TQ to diagonal form.

Formula (9.11') is obtained as follows (we denote (9.6') by $\|\ldots\|_\infty$):

$$\|Qv\|_\infty = \max_{k=1,\ldots,m}\Big|\sum_{i=1}^{n} q_{ki}v_i\Big| \le \Big(\max_{k=1,\ldots,m}\sum_{i=1}^{n}|q_{ki}|\Big)\cdot\|v\|_\infty \tag{9.12}$$

shows that $\|Q\| \le \max_k \sum_i |q_{ki}|$. The equality in (9.11') is then seen by choosing a vector of the form $v=(\pm1,\pm1,\ldots,\pm1)^T$ for which equality holds in (9.12). The formula (9.11") is proved along the same lines. □

All these formulas remain valid for *complex matrices.* Q^T has only to be replaced by Q^* (transposed and complex conjugate). See e.g., Wilkinson (1965), p. 55-61, Bakhvalov (1976), Chap. VI, Par. 3. With these preparations it is possible to formulate the desired estimate.

Theorem 9.3. *If $f(x,y)$ is differentiable with respect to y in an open convex region U and if*

$$\left\| \frac{\partial f}{\partial y}(x,y) \right\| \le L \qquad \textit{for} \qquad (x,y) \in U \tag{9.13}$$

then

$$\|f(x,z) - f(x,y)\| \le L\,\|z-y\| \qquad \textit{for} \qquad (x,y),(x,z) \in U. \tag{9.14}$$

(Obviously, the matrix norm in (9.13) is subordinate to the norm used in (9.14).)

Proof. This is the "mean value theorem" and its proof can be found in every textbook on calculus. In the case where $\partial f/\partial y$ is continuous, the following simple proof is possible. We consider $\varphi(t) = f\big(x, y + t(z-y)\big)$ and integrate its derivative (componentwise) from 0 to 1

$$\begin{aligned} f(x,z) - f(x,y) = \varphi(1) - \varphi(0) &= \int_0^1 \varphi'(t)\,dt \\ &= \int_0^1 \frac{\partial f}{\partial y}\big(x, y + t(z-y)\big) \cdot (z-y)\,dt. \end{aligned} \tag{9.15}$$

Taking the norm of (9.15), using

$$\left\| \int_0^1 g(t)\,dt \right\| \le \int_0^1 \|g(t)\|\,dt, \tag{9.16}$$

and applying (9.10) and (9.13) yields the estimate (9.14). The relation (9.16) is proved by applying the triangle inequality (9.7) to the finite Riemann sums which define the two integrals. □

We thus have obtained the analogue of (7.7). All that remains to do is, *Da capo al fine,* to read Sections I.7 and I.8 again: *Lemma 7.2, Theorems 7.3, 7.4, 7.5, and 7.6 together with their proofs and the estimates (7.10), (7.13), (7.15), (7.16), (7.17), and (7.18) carry over to the more general case with the only changes that some absolute values are to be replaced by norms.*

The *Picard-Lindelöf iteration* also carries over to systems of equations when in (8.7) we interpret $y_{i+1}(x), y_0$ and $f(s, y_i(s))$ as vectors, integrated componentwise. The convergence result with the estimate (8.10) also remains the same; for its proof we have to use, between (8.8) and (8.9), the inequality (9.16).

The Taylor series method, its convergence proof, and the recursive generation of the Taylor coefficients also generalize in a straightforward manner to systems of equations.

Exercises

1. Solve the system

$$y_1' = -y_2, \qquad y_1(0) = 1$$
$$y_2' = +y_1, \qquad y_2(0) = 0$$

   by the methods of Euler and Picard, establish rigorous error estimates for all three norms mentioned. Verify the results using the correct solution $y_1(x) = \cos x$, $y_2(x) = \sin x$.

2. Consider the differential equations

$$y_1' = -100y_1 + y_2, \qquad y_1(0) = 1, \qquad y_1(1) = ?$$
$$y_2' = y_1 - 100y_2, \qquad y_2(0) = 0, \qquad y_2(1) = ?$$

   a) Compute the exact solution $y(x)$ by the method explained in Section I.6.
   b) Compute the error bound for $\|z(x) - y(x)\|$, where $z(x) = 0$, obtained from (7.10).
   c) Apply the method of Euler to this equation with $h = 1/10$.
   d) Apply Picard's iteration method.

3. Compute the Taylor series solution of the system with constant coefficients $y' = Ay$, $y(0) = y_0$. Prove that this series converges for all x. Apply this series to the equation of Exercise 1.

   Result.

$$y(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!} A^i y_0 =: e^{Ax} y_0.$$