PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 127, Number 7, Pages 2059–2063
S 0002-9939(99)04801-7
Article electronically published on February 26, 1999

C^r CONVERGENCE OF PICARD'S SUCCESSIVE APPROXIMATIONS

ALEXANDER J. IZZO

(Communicated by Hal L. Smith)

Abstract. A simple, elementary proof of the existence, uniqueness, and smoothness of solutions to ordinary differential equations is given. In fact, it is shown that for a differential equation of class C^r, the successive approximations of Picard converge in the C^r-sense.

It is a fundamental theorem of ordinary differential equations that an equation of the form $\alpha'(t) = f(\alpha(t))$, where f is a C^r function ($1 \le r \le \infty$) on an open set in euclidean space, has a unique solution for each initial condition and the solution is of class C^r as a function of both t and the initial condition. It is well known that the existence and uniqueness follow easily from the contraction mapping principle, but the usual proofs of the smoothness of the solution are difficult (see for instance [C–L], [D], [H–S], [L1], or [L2]). In fact, Michael Spivak [S] refers to the smoothness as "a very hard theorem". However, we will give below a simple, elementary proof of smoothness by extending the contraction mapping proof of existence and uniqueness. In fact, we will show that the successive approximations of Picard converge to the solution *in the C^r-sense*. The precise result is given in the theorem below. (As usual given a bounded function f defined on a set X, we will denote the supremum of f over X by $\|f\|_X$ or, when the set X is understood, by $\|f\|_\infty$. Given a function $g : U \times V \to \mathbb{R}^m$ where U and V are open sets in euclidean spaces, $D_j^k g$ will denote the kth partial derivative of g with respect to the jth variable ($j = 1, 2$).)

Theorem. *Let $f : U \to \mathbb{R}^m$ be a C^r function ($1 \le r \le \infty$), where $U \subset \mathbb{R}^m$ is open, and let x_0 be an arbitrary point in U. Choose $a > 0$ such that the closed ball $\overline{B}_{2a}(x_0)$, of radius $2a$ and center x_0, is contained in U, and choose $\eta > 0$ such that*

$$\eta \le a/\|f\|_{\overline{B}_{2a}(x_0)}$$

and

$$\eta < 1/\|Df\|_{\overline{B}_{2a}(x_0)}.$$

Then there is a unique function $\phi : (-\eta, \eta) \times B_a(x_0) \to \mathbb{R}^m$ satisfying the equations

$$\begin{cases} D_1\phi(t,x) = f(\phi(t,x)), \\ \quad \phi(0,x) = x. \end{cases}$$

Received by the editors October 6, 1997.
1991 *Mathematics Subject Classification.* Primary 34A12, 34A45.

©1999 American Mathematical Society

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use

Now let M denote the set of continuous maps of $(-\eta,\eta)\times B_a(x_0)$ into $\overline{B}_{2a}(x_0)$, and define a map $S : M \to M$ by

$$(S\alpha)(t,x) = x + \int_0^t f(\alpha(s,x))\,ds.$$

Then for every C^r function α_0 in M whose derivatives up to order r are bounded, the sequence $(S^n\alpha_0)_{n=0}^\infty$ obtained by iterating S converges in the C^r-sense to ϕ, and hence, ϕ is of class C^r.

Before proving the theorem, we need a simple lemma which is reminiscent of the contraction mapping principle. The proof of the lemma given below was shown to me by Andrew Browder and is a simplification of my original proof.

Lemma. *Suppose (X,d) is a metric space, and suppose $(Q_n)_{n=0}^\infty$ is a sequence of contractions on X for which there exists a number $c<1$ such that*

$$d(Q_n(x),Q_n(y)) \le c\,d(x,y)$$

for all $x,y\in X$ and all n. Suppose also that there is a point $x_\infty\in X$ such that $Q_n(x_\infty)\to x_\infty$ as $n\to\infty$. Let $x_0\in X$ be arbitrary, and define a sequence (x_n) recursively by setting

$$x_{n+1} = Q_n(x_n).$$

Then $x_n\to x_\infty$ as $n\to\infty$.

Proof. For each n,

$$\begin{aligned} d(x_{n+1},x_\infty) &= d(Q_n(x_n),x_\infty)\\ &\le d(Q_n(x_n),Q_n(x_\infty)) + d(Q_n(x_\infty),x_\infty)\\ &\le c\,d(x_n,x_\infty) + d(Q_n(x_\infty),x_\infty), \end{aligned}$$

so since $Q_n(x_\infty)\to x_\infty$, we obtain

$$\limsup_{n\to\infty} d(x_n,x_\infty) \le c \limsup_{n\to\infty} d(x_n,x_\infty).$$

It follows that $\limsup d(x_n,x_\infty)=0$, so $x_n\to x_\infty$. □

Proof of the Theorem. We will establish existence and uniqueness in the usual way by applying the contraction mapping principle to S. First we should check that S is well-defined, that is, that $S\alpha$ really is an element of M for each α in M. It is easy to see that $S\alpha$ is continuous. Moreover, for any $t\in(-\eta,\eta)$ and $x\in B_a(x_0)$ we have

$$\begin{aligned} \big|(S\alpha)(t,x)-x\big| &= \left|\int_0^t f(\alpha(s,x))\,ds\right|\\ &\le \eta\|f\|_{\overline{B}_{2a}(x_0)}\\ &\le a, \end{aligned}$$

so $\big|(S\alpha)(t,x)-x_0\big| < 2a$. Thus $S\alpha$ is in M.

Note that if $\alpha : (-\eta,\eta)\times B_a(x_0)\to\mathbb{R}^m$ satisfies

$$\text{(1)}\qquad \begin{cases} D_1\alpha(t,x) = f(\alpha(t,x)),\\ \alpha(0,x) = x, \end{cases}$$

then the image of α must in fact lie in $B_{2a}(x_0)$. (If $\alpha(t,x)$ lies outside $B_{2a}(x_0)$ for some $t>0$ and $x\in B_a(x_0)$, then there must be a smallest positive number τ

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use

for which $\alpha(\tau, x)$ lies outside $B_{2a}(x_0)$. Then $\left|\alpha(\tau,x)-\alpha(0,x)\right| > a$, whereas the vector-valued form of the mean value theorem shows that we should have

$$\left|\alpha(\tau,x)-\alpha(0,x)\right| \leq \tau\|D_1\alpha\|_{(0,\tau)\times\{x\}} \leq \eta\|f\|_{\overline{B}_{2a}(x_0)} \leq a.$$

A similar argument applies if $\alpha(t,x) \notin B_{2a}(x_0)$ for some $t < 0$ and $x \in B_a(x_0)$.) Now the fundamental theorem of calculus shows that a function $\alpha : (-\eta,\eta) \times B_a(x_0) \to \mathbb{R}^m$ satisfies (1) if and only if α is a fixed point of S.

Note that M is a complete metric space under the supremum metric. Moreover, given $\alpha, \beta \in M$ we have

$$\begin{aligned}\|S\alpha - S\beta\|_\infty &= \sup_{(t,x)}\left|\int_0^t [f(\alpha(s,x)) - f(\beta(s,x))]\,ds\right| \\ &\leq \eta\|Df\|_{\overline{B}_{2a}(x_0)}\,\|\alpha-\beta\|_\infty,\end{aligned}$$

as

$$\left|f(\alpha(s,x)) - f(\beta(s,x))\right| \leq \|Df\|_{\overline{B}_{2a}(x_0)}\,\|\alpha-\beta\|_\infty$$

for all s and x. Since $\eta\|Df\|_{\overline{B}_{2a}(x_0)} < 1$, this shows that $S : M \to M$ is a contraction. Therefore, by the contraction mapping principle, S has a unique fixed point ϕ, and moreover, if $\alpha_0 \in M$ is arbitrary, then the sequence $(S^n\alpha_0)_{n=0}^\infty$ converges uniformly to ϕ.

Now let α_0 be an arbitrary C^r function in M whose derivatives up to order r are bounded. Since the limit of a sequence that converges in the C^r-sense must be of class C^r, the proof of the theorem will be complete once we have shown that the sequence $(S^n\alpha_0)$ converges in the C^r-sense to ϕ. For the moment we consider only the case $r = 1$. First we note that if α is a C^1 function in M whose derivative is bounded, then the same is true of $S\alpha$, for applying the fundamental theorem of calculus to the equation defining S gives

$$D_1(S\alpha)(t,x) = f(\alpha(t,x)) \tag{2}$$

and differentiating under the integral sign gives

$$D_2(S\alpha)(t,x) = I + \int_0^t (Df)(\alpha(s,x)) \cdot D_2\alpha(s,x)\,ds \tag{3}$$

where I is the identity on $\mathbb{R}^m$.

Since we already know that $(S^n\alpha_0)$ converges uniformly to ϕ, to show that $(S^n\alpha_0)$ converges in the C^1-sense to ϕ it suffices to show that $(D_1(S^n\alpha_0))$ and $(D_2(S^n\alpha_0))$ converge uniformly. Since equation (2) gives $D_1(S^n\alpha_0) = f\circ(S^{n-1}\alpha_0)$, the uniform convergence of $(S^n\alpha_0)$ and the uniform continuity of f on $\overline{B}_{2a}(x_0)$ together show that $(D_1(S^n\alpha_0))$ converges uniformly.

For notational convenience, let $\alpha_n = S^n\alpha_0$ and $\alpha_\infty = \phi$. Let $\mathcal{L}(\mathbb{R}^m)$ denote the space of linear operators on $\mathbb{R}^m$ with the usual operator norm, and let L denote the space of bounded continuous maps from $(-\eta,\eta) \times B_a(x_0)$ to $\mathcal{L}(\mathbb{R}^m)$ under the supremum norm. For each $n \in \mathbb{Z}_+ \cup \{0,\infty\}$ define $Q_n : L \to L$ by

$$(Q_n l)(t,x) = I + \int_0^t (Df)(\alpha_n(s,x)) \cdot l(s,x)\,ds.$$

It is easy to see that the formula on the right-hand side defines an element of L so that Q_n is well-defined.

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use

Now note that equation (3) shows that

$$D_2\alpha_{n+1} = Q_n D_2\alpha_n \tag{4}$$

for all $n \in \mathbb{Z}_+ \cup \{0\}$. Moreover, for every $l_1, l_2 \in L$ and $n \in \mathbb{Z}_+ \cup \{0, \infty\}$ we have

$$\begin{aligned}\|Q_n l_1 - Q_n l_2\|_\infty &= \sup_{(t,x)} \left\| \int_0^t (Df)(\alpha_n(s,x)) \cdot (l_1(s,x) - l_2(s,x))\, ds \right\| \\ &\leq \eta \|Df\|_{\overline{B}_{2a}(x_0)} \|l_1 - l_2\|_\infty\end{aligned}$$

as $\|(Df)(\alpha_n(s,x))\| \, \|l_1(s,x) - l_2(s,x)\| \leq \|Df\|_{\overline{B}_{2a}(x_0)} \|l_1 - l_2\|_\infty$ for all s and x. Since $\eta\|Df\|_{\overline{B}_{2a}(x_0)} < 1$, this shows that the sequence (Q_n) satisfies the first condition in the lemma. Of course it also shows that Q_∞ is a contraction.

Now note that for every $l \in L$ and $n \in \mathbb{Z}_+$ we have

$$\begin{aligned}&\|(Q_n l)(t,x) - (Q_\infty l)(t,x)\| \\ &\qquad = \left\| \int_0^t [(Df)(\alpha_n(s,x)) - (Df)(\alpha_\infty(s,x))] \cdot l(s,x) ds \right\| \\ &\qquad \leq \eta \left[\sup_{(s,x)} \|(Df)(\alpha_n(s,x)) - (Df)(\alpha_\infty(s,x))\| \right] \|l\|.\end{aligned} \tag{5}$$

Since $\alpha_n \to \alpha_\infty$ uniformly and Df is uniformly continuous on $\overline{B}_{2a}(x_0)$, the expression in brackets on the last line goes to 0 as $n \to \infty$. Hence $Q_n l \to Q_\infty l$ in L. In particular, if we let l_∞ denote the fixed point of Q_∞, then $Q_n l_\infty \to l_\infty$.

Recalling equation (4) and applying the lemma, we conclude that $D_2\alpha_n \to l_\infty$ as $n \to \infty$. This completes the proof of the theorem in the case $r = 1$.

The case $2 \leq r < \infty$ we handle by induction. Assume that (α_n) converges to ϕ in the C^{r-1}-sense whenever f is of class C^{r-1}. Now suppose that f is of class C^r. Note that if α is a C^r function in M whose derivatives up to order r are bounded, then by induction, formulas (2) and (3) imply that the same is true of $S\alpha$. By our assumption, (α_n) converges in the C^{r-1}-sense, and since by equation (2) we have $D_1\alpha_{n+1} = f \circ \alpha_n$, it follows that $(D_1\alpha_n)$ converges in the C^{r-1}-sense. Therefore, to prove that (α_n) converges in the C^r-sense, all we need to show is that $(D_2^r\alpha_n)$ converges uniformly. If α is a C^r function, then repeated application of the chain rule shows that $D_2^r(f \circ \alpha)$ is a polynomial in $(D^1 f) \circ \alpha, \ldots, (D^r f) \circ \alpha$ and $D_2^1\alpha, \ldots, D_2^r\alpha$ whose only dependence on $D_2^r\alpha$ is the term $((Df) \circ \alpha) \cdot (D_2^r\alpha)$. In particular, if we let $p_\alpha = D_2^r(f \circ \alpha) - ((Df) \circ \alpha) \cdot (D_2^r\alpha)$, then p_α depends continuously on $D_2^1\alpha, \ldots, D_2^{r-1}\alpha$ and is independent of $D_2^r\alpha$. Thus since (α_n) converges to α_∞ in the C^{r-1}-sense, (p_{α_n}) converges uniformly to p_{α_∞}. Let $\mathcal{L}^r(\mathbb{R}^m)$ denote the space of r-multilinear maps on $\mathbb{R}^m$ with the usual norm, and let L^r denote the space of bounded continuous maps from $(-\eta, \eta) \times B_a(x_0)$ to $\mathcal{L}^r(\mathbb{R}^m)$ under the supremum norm. For each $n \in \mathbb{Z}_+ \cup \{0, \infty\}$ define $\tilde{Q}_n : L^r \to L^r$ by

$$(\tilde{Q}_n l)(t,x) = \int_0^t p_{\alpha_n}(s,x)\, ds + \int_0^t (Df)(\alpha_n(s,x)) \cdot l(s,x)\, ds.$$

Applying D_2^r to the formula defining $S\alpha$, we see that

$$D_2^r\alpha_{n+1} = \tilde{Q}_n D_2^r\alpha_n.$$

Note that since (p_{α_n}) converges uniformly to p_{α_∞}, the first term in the definition of $\tilde{Q}_n$ converges in L^r to $\int_0^t p_{\alpha_\infty}(s,x)\, ds$. Now we can apply to $(\tilde{Q}_n)$ the same

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use

reasoning as we applied to (Q_n), to conclude that $(D_2^r \alpha_n)$ converges. This completes the proof in the case $r < \infty$. The case $r = \infty$ is an immediate consequence of the case $r < \infty$. □

If one is interested only in showing that ϕ is of class C^r, then the induction argument given above can be eliminated as there is a standard argument (which also uses induction) showing that the C^r smoothness of solutions to C^r differential equations follows from the C^1 smoothness of solutions to C^1 differential equations. See [H–S], p. 302. However, if we attempt to use the standard method to show C^r convergence of $(S^n \alpha_0)$, then the set on which we get convergence shrinks each time we apply induction. In particular, we would not obtain C^∞ convergence when f is of class C^∞. The theorem above gives C^r convergence on a set that depends on $\|f\|_{\overline{B}_{2a}(x_0)}$ and $\|Df\|_{\overline{B}_{2a}(x_0)}$ but is *independent* of higher order derivatives of f.

Although we have considered above only differential equations in $\mathbb{R}^m$, our method can also be applied to differential equations in a Banach space. If we replace $\mathbb{R}^m$ by a Banach space throughout, everything goes through as before (provided a is chosen such that f and Df are bounded on $\overline{B}_{2a}(x_0)$) except that we can no longer conclude that the expression in brackets on the last line of (5) goes to 0 as $n \to \infty$ since we are no longer assured that Df is uniformly continuous on $\overline{B}_{2a}(x_0)$. However, the continuity of Df still implies that $(Df) \circ \alpha_n \to (Df) \circ \alpha_\infty$ uniformly *on compact subsets* of $(-\eta, \eta) \times B_a(x_0)$. We thus conclude that $(S^n \alpha_0)$ converges to ϕ in the C^r-sense *on compact subsets* of $(-\eta, \eta) \times B_a(x_0)$. This is sufficient to ensure that ϕ is of class C^r.

References

[C–L] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, New York, 1955. MR **16:**1022b

[D] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1960. MR **50:**1782

[H–S] M. W. Hirsh and S. Smale, *Differential equations, dynamical systems, and linear algebra*, Academic Press, New York, 1974. MR **58:**6484

[L1] S. Lang, *Analysis II*, Addison-Wesley, Menlo Park, California, 1969.

[L2] ———, *Differentiable manifolds*, Springer-Verlag, New York, 1985.

[S] M. Spivak, *A comprehensive introduction to differential geometry*, 2nd ed., vol. 1, Publish or Perish, Wilmington, Delaware, 1979. MR **42:**2369

Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403

E-mail address: aizzo@math.bgsu.edu

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use