

SMOOTHING DERIVATIVES OF FUNCTIONS AND APPLICATIONS

BY

F. WESLEY WILSON, JR.⁽¹⁾

Introduction. The primary purpose of this note is to construct a C^∞ Lyapunov function for a continuous vector field which has an asymptotically stable invariant set. Since there already exist techniques for constructing continuous Lyapunov functions with continuous derivatives by the vector field, our problem is simply to smooth such a function in such a way that the desired derivative properties are preserved. An extremely simple smoothing procedure is to average the function by a convolution integral. Unfortunately, while it is well known that such averaging provides a uniform approximation for continuous directional derivatives, it seems very difficult to show that the continuous derivatives by an arbitrary continuous vector field are preserved (if they are). In the first section, we show that the derivative is preserved in the case where the function satisfies a Lipschitz condition.

Thus the problem comes to approximating our given function by one which satisfies a Lipschitz condition, still preserving the derivative of course. This problem is solved in the second section by using an averaging technique which was developed by Kurzweil [3]. Unfortunately, while the convolution provides a close approximation for the derivative, the Kurzweil averaging only allows for a one-sided approximation of the derivative. However, this is enough for our purposes, and in the third section we apply these results to the converse Lyapunov problem. In the final section of this paper we apply these results to obtain a topological version of the Hirsch-Cairns smoothing theorem.

We owe special gratitude to Charles C. Pugh for his careful reading of the manuscript, especially §2. Because of his diligence, the reader will be spared many ambiguities and typographical errors, and in addition one rather annoying technical error which appeared in the original manuscript.

1. Smoothing Lipschitzian Functions. Let M be a paracompact C^∞ manifold, and let X be a continuous nonsingular vector field on M whose trajectories are defined for all time and depend uniquely upon their initial conditions. Suppose that $f: M \rightarrow R$ is a given continuous function such that the derivative Xf is defined and

Received by the editors October 12, 1967 and, in revised form, July 16, 1968.

⁽¹⁾ This research was supported by the United States Army Research Office, Durham, under Contract DA-31-124-ARO-D-270, by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant 1243-67, and by the National Science Foundation under NSF GP 8965.

satisfies a boundedness condition, i.e., there are continuous functions $\mu, \nu: M \rightarrow R$ such that $\mu(p) \leq Xf(p) \leq \nu(p)$ for all $p \in M$. Let $\varepsilon: M \rightarrow R$ be any positive function. We shall say that a function g ε -approximates f in its X -derivative (relative to the stated boundedness condition) provided

$$(i) \quad |f(p) - g(p)| < \varepsilon(p),$$

$$(ii) \quad \mu(p) - \varepsilon(p) \leq Xg(p) \leq \nu(p) + \varepsilon(p)$$

for all $p \in M$. In the case where Xf is itself continuous, we will normally take $\mu(p) = \nu(p) = Xf(p)$. We shall be interested in the problem of deciding when f can be approximated in its X -derivative by a C^∞ function.

We shall now discuss a result which is well known (see [2, Chapter 1], for instance), and which motivates this section. Let $\phi: R^n \rightarrow R$ be a C^∞ nonnegative function which vanishes outside of the unit disk and satisfies $\int \phi = 1$. For any continuous function $f: R^n \rightarrow R$, define

$$f_\varepsilon(p) = \int f(p + \varepsilon x) \phi(x) dx.$$

PROPOSITION 1.1. *The family $\{f_\varepsilon\}$ consists of C^∞ functions which approximate f uniformly on compact subsets as ε tends to 0. Moreover, if D is any directional derivative operator, then $Df_\varepsilon = (Df)_\varepsilon$ where Df is interpreted in the distribution sense, if the usual derivative is not defined. Moreover, if Df is continuous, then Df_ε approximates Df uniformly on compact subsets.*

If it were true that $Xf_\varepsilon = (Xf)_\varepsilon$, for any continuous vector field, then our problem would have a very simple solution. However, the left side may be C^0 while the right side is C^∞ , so equality does not hold. In fact, even if $(Xf)_\varepsilon$ exists, it does not seem to be known whether or not $Xf_\varepsilon - (Xf)_\varepsilon$ must tend uniformly to zero with ε in general (which again would be sufficient for our purposes).

REMARK. One must be very careful with the symbol $(Xf)_\varepsilon$. If X and f are only continuous, then there are two serious problems which arise, even if Xf is defined. First of all, although Xf is measurable, it may not be integrable. The assumption of a uniform Lipschitz condition for f on compact subsets provides essential bounds for Xf on compact subsets which yields integrability. The second problem is that the symbol Xf represents the derivative of f along X -trajectories. If X is only continuous, then the X -trajectories need not satisfy a uniqueness condition and $Xf(p)$ may take on several values depending on the choice of trajectory. This is illustrated by the following example:

$$\begin{aligned} f(x, y) &= (x^3 - y)^{1/3} & \text{if } x^3 \geq |y|, \\ &= 0 & \text{if } x^3 \leq |y|, \end{aligned}$$

defined on R^2 , has different derivatives along the tangent curves through the origin

$$\begin{aligned} \phi(t) &= (t, 0), & f\phi(t) &= t, \\ \psi(t) &= (t, t^3), & f\psi(t) &= 0. \end{aligned}$$

Hence $Xf(p)$ is not necessarily well defined, even though f is differentiable along X -trajectories. However, if f is Lipschitzian (uniform Lipschitz condition on compact subsets), then f has a differential almost everywhere [8], and so Xf is well defined almost everywhere. Therefore, if f is Lipschitzian, then $(Xf)_\varepsilon$ is always defined.

The next proposition is the analogue to 1.1 for Lipschitzian functions and continuous vector fields (cf. [3] and [5]).

PROPOSITION 1.2. *If $f: R^n \rightarrow R$ is a Lipschitzian function, and X is any continuous vector field, then $Xf_\varepsilon - (Xf)_\varepsilon$ tends to zero with ε , uniformly on compact subsets.*

Proof. For each $p \in R^n$, we can choose a curve $p(t)$ which is a trajectory of X and satisfies $p(0)=p$. Let K be a given compact subset of R^n , let U be a relatively compact neighborhood of K , and let V be a relatively compact neighborhood of \bar{U} (closure). Then there is a $t' > 0$ such that $p(t) \in V(U)$ whenever $p \in U(K)$ and $t \leq t'$. Let k denote a Lipschitz constant for f on V . Now

$$Xf_\varepsilon(p) - (Xf)_\varepsilon(p) = \lim_{t \rightarrow 0} \frac{1}{t} \int \{f[p(t) + \varepsilon x] - f[(p + \varepsilon x)(t)]\} \phi(x) dx.$$

It suffices to show that for each $\delta > 0$, there exists $\varepsilon^*, t^* > 0$ such that the above integral expression is bounded by δ whenever $p \in K$, $\varepsilon < \varepsilon^*$, and $t < t^*$. Define $\psi: [0, \infty) \rightarrow R$ by

$$\psi(\tau) = \sup \{ \|X_p - X_{p(t)}\| \mid 0 \leq t \leq \tau, p \in K \}.$$

Then ψ is a nondecreasing function, $\psi(0)=0$, and ψ is continuous at 0. Moreover, for each $q \in K$,

$$\|q + tX_q - q(t)\| \leq \int_0^t \psi(\tau) d\tau \leq t\psi(t).$$

Now choose $\varepsilon^* > 0$ to be smaller than $d(K, V - U)$ and $d(U, R^n - V)$, and so that $\|X_p - X_{p+\varepsilon^*x}\| < \delta/3k$ whenever $p \in U$ and $\|x\| \leq 1$. Choose $t^* > 0$ so that $t^* \leq t'$ and so that $\psi(t^*) < \delta/3k$. Then

$$\begin{aligned} & \frac{1}{t} \int \{f[p(t) + \varepsilon x] - f[(p + \varepsilon x)(t)]\} \phi(x) dx \\ & \leq \frac{k}{t} \int \|p(t) + \varepsilon x - p - tX_p - \varepsilon x\| \phi(x) dx \\ & \quad + \frac{k}{t} \int \|p + tX_p + \varepsilon x - (p + \varepsilon x) - tX_{p+\varepsilon x}\| \phi(x) dx \\ & \quad + \frac{k}{t} \int \|(p + \varepsilon x) + tX_{p+\varepsilon x} - (p + \varepsilon x)(t)\| \phi(x) dx \\ & \leq 2k\psi(t) + k \int \|X_p - X_{p+\varepsilon x}\| \phi(x) dx \\ & \leq \delta \end{aligned}$$

whenever $0 < t < t^*$, $0 < \varepsilon < \varepsilon^*$. This completes the proof.

Note that if Xf satisfies a boundedness condition, then Xf_ε satisfies a similar boundedness condition. Thus we are ready to prove the main result of this section.

THEOREM 1.3. *If $f: M \rightarrow R$ is Lipschitzian and if X is a continuous vector field such that Xf satisfies a boundedness condition relative to functions $\mu, \nu: M \rightarrow R$, then for each positive function $\varepsilon: M \rightarrow R$, there is a C^∞ function $g: M \rightarrow R$ which ε -approximates f in its X -derivative.*

Proof. Let $\{U_i\}$ be a locally finite covering of M by relatively compact coordinate neighborhoods, and let $\{\alpha_i\}$ be a subordinate C^∞ partition of unity. By 1.2, there is a function $g_i: U_i \rightarrow R$ which approximates $f|U_i$ in the X -derivative to any desired degree. Define $g = \sum \alpha_i g_i$. Then

$$Xg = Xf + \sum (X\alpha_i)(g_i - f) + \sum \alpha_i (Xg_i - Xf).$$

Thus we can compute the preciseness with which g_i must approximate f . Set $\varepsilon_i = \inf \{\varepsilon|U_i\}$. The desired degree of approximation for g_i is

$$\frac{\varepsilon_i}{2\beta_i[1 + \max \{ |X\alpha_j(x)| \mid j = 1, 2, \dots, \beta, \text{ and } x \in U_i \}]}$$

where $\alpha_j \equiv 0$ on U_i for all $j > \beta_i$.

Thus our problem reduces to approximating f in its X -derivative by a function which is Lipschitzian. We shall deal with this problem in §2. As we have already noted, we will not obtain a full approximation for Xf , but only a one-sided approximation. The next theorem is preparatory for this goal, and is of interest in its own right for two reasons. First of all, it provides a shortcut for obtaining a Lipschitzian approximation whenever the vector field X is itself Lipschitzian. Secondly, under the hypothesis that X is C^∞ it provides the full approximation of Xf .

THEOREM 1.4. *Let $f: M \rightarrow R$ be a continuous function and let X be a continuous vector field (with unique trajectories) so that Xf is defined and satisfies a boundedness condition (possibly trivial). Then f can be approximated by a continuous function g such that Xg is continuous. Moreover, (g, Xg) will satisfy any Lipschitzian or differentiability conditions that X satisfies.*

Proof. Let I^n denote the unit cube in R^n with the vector field Y defined on I^n by $Y(x) = (1, 0, \dots, 0)$. Since X is nonsingular, we can find a neighborhood U of each point $p \in M$ and a homeomorphism $h: U \rightarrow I^n$ which carries the X -trajectories onto the Y -trajectories (preserving parameterization) (cf. [4, p. 52]). Moreover, h will satisfy whatever Lipschitzian or differentiability conditions that X satisfies. Note that $Y(fh^{-1}) = (Xf)h^{-1}$ and that Y is a directional derivative. Therefore, by 1.1 we have $Y(fh^{-1})_\varepsilon = [Y(fh^{-1})]_\varepsilon$. Define $g = (fh^{-1})_\varepsilon h$. g approximates $f|U$, and $Xg = [Y(fh^{-1})]_\varepsilon h$; so Xg approximates Xf . Since $(fh^{-1})_\varepsilon$ is a C^∞ function, it is clear

that (g, Xg) is continuous and satisfies whatever Lipschitzian and differentiability properties which h satisfies, i.e., which X satisfies.

Now that we have established an approximation on the flow box neighborhoods U , we can obtain the global approximation as in the proof of 1.3.

REMARK I. Combining 1.3, 1.4, and §3, we can construct a C^∞ Lyapunov function in the special case where X is Lipschitzian. This case was first treated by Massera [5], [6] in the case of a uniformly asymptotically stable critical point.

REMARK II. As another application, we can complete the proof of the existence of C^∞ "nice functions" [9]. In Smale's case he has both a C^∞ vector field and a Lipschitzian function. Therefore, we can use either 1.3 or 1.4 to approximate Smale's function by a C^∞ function, preserving the X derivative (of course, we would not bother Smale's function near the critical points).

2. Approximating by Lipschitzian functions. Let the manifold M and the vector field X be as described in §1. We note that everything which we shall do in this section depends on the fact that X has unique trajectories. Suppose that the continuous function f has been given and that Xf is continuous (cf. 1.4). Then for each $p \in M$ and $t \in R$, we have

$$f[p(t)] = f(p) + \int_0^t Xf[p(\tau)] d\tau.$$

Our goal is to define a Lipschitzian function which approximates f in its X -derivative relative either to the boundedness condition $[-\infty; Xf(p)]$ or to $[Xf(p); +\infty]$. While we would prefer to obtain a sharper approximation of the X -derivative, the above bounds seem to be all that the Kurzweil process will yield. Such an approximation will be called one-sided.

We shall begin by using the Kurzweil process to obtain local approximations. Then global approximations can be obtained as in 1.3. Since X is nonsingular and has unique trajectories, we can find a flow box neighborhood at each point of M . Let U, W be open flow box neighborhoods such that $\bar{U} \subset W$, and let V be another neighborhood such that $\bar{U} \subset V$ and $\bar{V} \subset W$. We shall construct an approximation for f on U . Since W is a flow box, we can define a system of coordinates (p, t) on \bar{W} where $p \in W_0$ (the leading edge of W) and $t \in [0, T]$ (T is the time length of W). It will be useful to have notation $W_t = \{(p, t) \in W \mid t \text{ fixed}\}$. Then we have

$$f(p, t) = f(p, 0) + \int_0^t Xf(p, \tau) d\tau = f(p, T) - \int_t^T Xf(p, \tau) d\tau.$$

The key to Kurzweil's process is that one is allowed to integrate along a more favorable path, provided that he is willing to pay the cost. The cost which must be paid is defined by means of an integral. We shall now define this integral and study some of its basic properties.

Let $\phi: W \rightarrow [1, \infty)$ be a C^∞ function which has value 1 on V and which tends to

∞ at ∂W . If $\gamma: [0, T_\gamma] \rightarrow W$ is any piecewise differentiable curve, then we define the cost of γ by

$$\Phi(\gamma) = \int_0^{T_\gamma} \phi(\gamma) \|\dot{\gamma} - X_\gamma\|$$

(in what follows, it will be convenient to suppress the variable of integration from the notation). Define Γ_ε to be the set of all piecewise differentiable arcs $\gamma: [0, T_\gamma] \rightarrow \bar{W}$ such that

- (i) $\gamma(t) \in W$ if $0 \leq t < T_\gamma$ and $\gamma(T_\gamma) \in \partial W$,
- (ii) $\Phi(\gamma) < \varepsilon$.

It will follow from 2.1 that no arc which satisfies (ii) can stay in W for infinite time, and so the fact that every arc has finite length is not an additional assumption on Γ_ε . Let $\Gamma_\varepsilon(p, t)$ denote the subset of Γ_ε whose elements begin at (p, t) .

PROPOSITION 2.1. *If $\gamma \in \Gamma_\varepsilon$ and $\gamma(t) \in W_{t'}$, $\gamma(t+\tau) \in W_{t''}$, then*

$$t'' - t' \geq \tau - \left(\varepsilon / \inf_W \{ \|X\| \} \right).$$

In particular, each $\gamma \in \Gamma_\varepsilon$ must reach ∂W in at least time $T + (\varepsilon / \inf_W \{ \|X\| \})$.

Proof. $\varepsilon \geq \int_{t'}^{t''} \|\dot{\gamma} - X_\gamma\| \geq (\tau + t' - t'') \inf_W \{ \|X\| \}$.

PROPOSITION 2.2. *Each element $\gamma \in \Gamma_\varepsilon$ approaches $\gamma(T_\gamma)$ asymptotic to the trajectory through $\gamma(T_\gamma)$. Consequently, γ can leave W only through W_T .*

Proof. This follows immediately from the fact that $\phi \rightarrow \infty$ at ∂W .

PROPOSITION 2.3. *Let $\delta > 0$ be given. Then for sufficiently small values of $\varepsilon > 0$, we have for any $\gamma \in \Gamma_\varepsilon$, with $\gamma(t_0) = (p', t')$, that*

$$(1) \quad d[\gamma(t_0 + \tau), (p', t' + \tau)] < \delta, \quad 0 \leq \tau \leq T - t',$$

$$(2) \quad \left| \int_0^{T_\gamma} Xf(\gamma) - \int_{t'}^T Xf(p, \cdot) \right| < \delta, \quad (t_0 = 0).$$

Proof. We first prove (2), assuming that (1) holds. By (1), and the uniform continuity of Xf on \bar{W} , we see that $\sup_t \{ |Xf[\gamma(t)] - Xf(p, t' + \tau)| \}$ can be made as small as we please by restricting ε . Therefore,

$$\begin{aligned} \left| \int_0^{T_\gamma} Xf(\gamma) - \int_{t'}^T Xf(p, \cdot) \right| &\leq \int_0^{T-t'} |Xf(\gamma) - Xf(p, t' + \cdot)| + \int_{T-t'}^{T-t'+(\varepsilon/\inf\{\|X\|\})} |Xf(\gamma)| \\ &\leq T \sup \{ |Xf(\gamma) - Xf(p, t' + \cdot)| \} + \varepsilon \frac{\sup \{ |Xf| \}}{\inf \{ \|X\| \}}. \end{aligned}$$

We can choose $\varepsilon > 0$ so small that this last expression is less than δ .

Assume that (1) is not satisfied, i.e., there is a $\delta > 0$, a point (p', t') , and a sequence $\varepsilon_k \rightarrow 0$ with related curves γ_k such that $\gamma_k(t_k) = (p', t')$ and yet there are times τ_k such that (1) is violated. Expressing

$$\gamma_k(t) = \gamma_k(t_k) + \int_{t_k}^t \dot{\gamma} = \gamma_k(t_k) + \int_{t_k}^t X_\gamma + u_k(t),$$

where $\|u_k(t)\| < \varepsilon_k$ (here we assume that $W \subset R^n$), we prove easily that $\{\gamma_k\}$ is an equicontinuous family which has a trajectory for a limit point. But this trajectory must be $(p', t' + t)$, and this leads immediately to a contradiction.

COROLLARY. *Let $\delta > 0$ be given. Then for sufficiently small $\varepsilon > 0$,*

$$\left| f(p, t) - f[\gamma(T_\gamma)] + \int_0^{T_\gamma} Xf(\gamma) \right| < \delta$$

for all $\gamma \in \Gamma_\varepsilon(p, t)$.

We can now define our candidates for the approximating functions. Take

$$f^\varepsilon(p, t) = \sup_{\gamma \in \Gamma_\varepsilon(p, t)} \left\{ \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) \right] \left[1 - \frac{1}{\varepsilon} \Phi(\gamma) \right] \right\}.$$

Using the above corollary and the fact that $\Phi(p, \cdot) = 0$, we see that

$$(1) \quad f(p, t) \leq f^\varepsilon(p, t) \leq f(p, t) + \xi_\varepsilon$$

for some constant ξ_ε which goes to zero with ε . Thus we see that $f^\varepsilon \rightarrow f$ uniformly on U as $\varepsilon \rightarrow 0$. We first verify that f^ε is Lipschitzian. Since W is relatively compact, there is no loss of generality in assuming that f is strictly positive on W . Let constants ζ, η be chosen so that $0 < \zeta < \frac{1}{2}$ and $0 < \eta < \zeta f(p, t)$. Then there is a curve $\gamma \in \Gamma_\varepsilon(p, t)$ such that

$$(2) \quad \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) \right] \left[1 - \frac{1}{\varepsilon} \Phi(\gamma) \right] > f(p, t) - \eta.$$

Then by (1) and (2), we have

$$(3) \quad \left[1 - \frac{1}{\varepsilon} \Phi(\gamma) \right] \geq (1 - \zeta) \frac{f(p, t)}{f(p, t) + \xi_\varepsilon}.$$

Let $r = d(U, W - V)$ and choose $\varepsilon \leq r/8$. Let $\delta' = \min \{r/2, r/(8 \sup_W \{ \|X\| \})\}$.

For these choices, we have

LEMMA. *If $d[(p, t), U] < \delta'$, and $\gamma \in \Gamma_\varepsilon(p, t)$ then $\gamma(t) \in V$ for $0 \leq t \leq \delta'$.*

Proof. Suppose that there exists such (p, t) and a $\gamma \in \Gamma_\varepsilon(p, t)$ such that $\gamma(t) \in W - V$ for some $t \leq \delta'$. Since $\delta' < r/2$, there are times t_1, t_2 such that $d[\gamma(t_1), W - V] = r/2$, $d[\gamma(t_2), W - V] = r/4$, and $r/4 < d[\gamma(t), W - V] < r/2$ for $t_1 < t < t_2$. Now $\varepsilon \geq \int_{t_1}^{t_2} \|\dot{\gamma} - X_\gamma\|$, and so

$$d[\gamma(t_1), \gamma(t_2)] \leq \int_{t_1}^{t_2} \|\dot{\gamma}\| \leq \varepsilon + \int_{t_1}^{t_2} \|X\| < r/4.$$

But this contradicts the choices of t_1 and t_2 .

Select constants

$$K = \sup \{1, \|X(p, t)\|, f(p, t) + \xi_\varepsilon, |Xf(p, t)| \mid (p, t) \in V\},$$

$$L = \inf \{\tfrac{1}{2}f(p, t) \mid (p, t) \in W\} \quad (L > 0).$$

Set $\delta = \min \{\delta', \varepsilon L/3K^2\}$. We shall show that whenever $(p, t), (p', t')$ are points in the δ -neighborhood of U such that $d[(p, t), (p', t')] < \delta$, then

$$(4) \quad |f^\varepsilon(p, t) - f^\varepsilon(p', t')| < 5K^2\varepsilon^{-1}d[(p, t), (p', t')].$$

Let $0 < \eta < \varepsilon L$. Then there is a $\gamma \in \Gamma_\varepsilon(p, t)$ such that (2) is satisfied. For this γ we have

$$(5) \quad \Phi(\gamma) \leq [1 - (L/K)]\varepsilon.$$

Assuming that W is a subset of R^n , we take v to be the vector from (p, t) to (p', t') . Define a curve $\beta: [0, T_\gamma] \rightarrow W$ by

$$\begin{aligned} \beta(s) &= \gamma(s) + [1 - (s/\|v\|)]v & \text{if } 0 \leq s \leq \|v\|, \\ &= \gamma(s) & \text{if } \|v\| \leq s. \end{aligned}$$

Thus β is a piecewise differentiable curve starting at (p', t') . It has derivative

$$(6) \quad \dot{\beta}(t) = \dot{\gamma}(t) - v/\|v\| \quad \text{for } 0 \leq t \leq \|v\|.$$

If we can show $\Phi(\beta) < \varepsilon$, then $\beta \in \Gamma_\varepsilon(p', t')$. Set $\Delta = \Phi(\beta) - \Phi(\gamma)$,

$$(7) \quad \Delta = \int_0^{T_\beta} \phi(\beta)\|\dot{\beta} - X_\beta\| - \int_0^{T_\gamma} \phi(\gamma)\|\dot{\gamma} - X_\gamma\| = \int_0^{\|v\|} \|\dot{\beta} - X_\beta\| - \|\dot{\gamma} - X_\gamma\|.$$

But $\gamma(t) \in V$ for $0 \leq t \leq \|v\|$ and $d[\beta(t), \gamma(t)] < \|v\|$ for all t . Thus, $\beta(t)$ remains in the δ -neighborhood of V for $0 \leq t \leq \|v\|$. By (6), (7), and the choices of K, L ,

$$(8) \quad |\Delta| \leq \|\dot{\gamma} - X_\gamma\| - \|(\dot{\gamma} - X_\gamma - v/\|v\|)\| \cdot \|v\| < 3K\|v\|.$$

But then by (5) and (8),

$$\Phi(\beta) \leq \Phi(\gamma) + |\Delta| < (1 - L/K)\varepsilon + 3K\|v\| < \varepsilon.$$

Finally, we compute (4),

$$\begin{aligned} f^\varepsilon(p', t') &\geq \left[f(\beta(T_\beta)) - \int_0^{T_\beta} Xf(\beta) \right] [1 - \varepsilon^{-1}\Phi(\beta)] \\ &\geq \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) - 2K\|v\| \right] [1 - \varepsilon^{-1}\Phi(\beta)] \\ &\geq \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) \right] [1 - \varepsilon^{-1}\Phi(\gamma) - \varepsilon^{-1}|\Delta|] - 2K\|v\| \\ &\geq f^\varepsilon(p, t) - \eta - \varepsilon^{-1}K|\Delta| - 2K\|v\| \\ &\geq f^\varepsilon(p, t) - \eta - 5K^2\varepsilon^{-1}d[(p, t), (p', t')]. \end{aligned}$$

(4) now follows by symmetry and the fact that $\eta > 0$ was arbitrary. Now that we have established that f^ε is Lipschitzian, it follows that Xf^ε is defined almost everywhere and that it is integrable on compact subsets (cf. Remark following 1.1). We

will now show that the X -derivates of f^ε provide a one-sided approximation of the form $(-\infty, Xf)$ for the X -derivatives of f . Precisely, we show

$$(9) \quad \limsup_{h \rightarrow 0^+} -\frac{1}{h} [f^\varepsilon(p, t-h) - f^\varepsilon(p, t)] \leq Xf(p, t).$$

Choose constants ζ, η as before and choose $\gamma \in \Gamma_\varepsilon(p, t)$ so that (2) and (3) are satisfied. Define a curve $\beta: [0, T_\gamma + h] \rightarrow W$ as follows:

$$\begin{aligned} \beta(s) &= (p, t-h+s) & \text{if } 0 \leq s \leq h, \\ &= \gamma(s-h) & \text{if } h \leq s \leq T_\gamma + h. \end{aligned}$$

Then β follows the X -trajectory from $(p, t-h)$ to (p, t) and γ thereafter. Now $\Phi(\beta) = \Phi(\gamma) < \varepsilon$; so $\beta \in \Gamma_\varepsilon(p, t-h)$. Also,

$$(10) \quad \int_0^{T_\beta} Xf(\beta) = \int_{-h}^0 Xf(p, t+\cdot) + \int_0^{T_\gamma} Xf(\gamma).$$

Combining (2), (3), and (10) we conclude that

$$\begin{aligned} f^\varepsilon(p, t-h) &\geq \left[f(\beta(T_\beta)) - \int_0^{T_\beta} Xf(\beta) \right] \left[1 - \frac{1}{\varepsilon} \Phi(\beta) \right] \\ &\geq \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) \right] \left[1 - \frac{1}{\varepsilon} \Phi(\gamma) \right] \\ &\quad - \left[\int_{-h}^0 Xf(p, t+\cdot) \right] \left[1 - \frac{1}{\varepsilon} \Phi(\gamma) \right] \\ &\geq f^\varepsilon(p, t) - \eta - \int_{-h}^0 Xf(p, t+\cdot) \end{aligned}$$

since $1 - \Phi(\gamma)/\varepsilon < 1$. Hence

$$-\frac{1}{h} [f^\varepsilon(p, t-h) - f^\varepsilon(p, t)] \leq \frac{\eta}{h} + \frac{1}{h} \int_{-h}^0 Xf(p, t+\cdot).$$

Since η is independent of h , (9) follows.

We state the following theorem as a summary of what has been proved thus far in this section.

THEOREM 2.4. *Let f be a continuous function which has a continuous derivative by the nonsingular vector field X (continuous, unique trajectories). Then for each positive function $\varepsilon: M \rightarrow R$, there are Lipschitzian functions $g, h: M \rightarrow R$ which are ε -approximations to f in value and which provide one-sided approximations to the X -derivative of the form*

$$Xg \geq Xf - \varepsilon, \quad Xh \leq Xf + \varepsilon.$$

The upper estimate can be obtained by defining

$$f^\varepsilon(p, t) = \inf_{\gamma \in \Gamma_\varepsilon(p, t)} \left\{ \left[f(\gamma(T_\gamma)) - \int_0^{T_\gamma} Xf(\gamma) \right] \left[1 + \frac{1}{\varepsilon} \Phi(\gamma) \right] \right\}.$$

Combining 1.3, 1.4, and 2.4, we obtain the full approximation theorem.

THEOREM 2.5. *If $f: M \rightarrow R$ is a continuous function, X is a continuous, nonsingular vector field with unique trajectories on M , and if Xf is defined and satisfies a boundedness condition, then for any positive function $\varepsilon: M \rightarrow R$, we can find a C^∞ function which is a one-sided ε -approximation to f in the X -derivative. [This approximation can be chosen to approximate Xf from either side, but not from both sides simultaneously.]*

REMARK. While everything which we have done depends strongly on the existence of unique trajectories for X , the same is not true of Kurzweil's work [3]. Rather than a full approximation theorem, he is only interested in proving the existence of a Lyapunov function for a uniformly asymptotically stable critical point of a time varying system (see §3 for these definitions). For this, he was able to dispense with the uniqueness assumption. He defined Γ to be the class of all piecewise differentiable curves $\gamma: [0, \infty) \rightarrow M$ such that the penalty function (integrated for infinite time) has value < 1 . Then he found a positive function $\chi: M \rightarrow R$ such that $\int \chi < \infty$ for all $\gamma \in \Gamma$. Defining f by

$$f(x) = \sup_{\gamma \in \Gamma(x)} \left\{ \int_0^\infty \chi(\gamma) [1 - \Phi(\gamma)] \right\}$$

he obtained the desired Lyapunov function. (Here he is not integrating along trajectories, and it is not even necessary that trajectories exist. For instance in situations where existence is a problem, this procedure might give rise to energy functions which would allow us to make qualitative statements regarding the nature of solutions whenever they exist.) In our case, where we consider large attracting sets, it is not clear that the technique of integrating over whole trajectories would work in general. Moreover by restricting to the case of unique trajectories, we have obtained a much shorter exposition, and an approximation theorem for general attracting sets. It is possible that such an approximation theorem could be proved without the necessity of uniqueness, and we will now make a few comments on the difficulties which would be involved. We have used uniqueness in two ways:

1. The flow box construction.
2. The formula $f(p, t) = f(p, 0) + \int_0^t Xf(p, t) dt$.

Suppose that we had f continuous, $\mu \leq Xf \leq \nu$ (for all possibilities; cf. Remark following 1.1), but f not necessarily absolutely continuous along trajectories. We used (1.4) the flow box construction to make Xf continuous, and hence we obtained the above formula. Without flow boxes, we might still hope to find some neighborhood on which we could do the averaging, and some function χ which we could integrate in place of Xf . The choice of the neighborhood is not difficult; using the local triviality of the tangent bundle and the continuity of X , we still have the vectors locally "parallel" and so we can construct "generalized flow boxes", i.e., neighborhoods W homeomorphic to the unit cube, and so that the vector field points inward at the bottom, outward at the top, and slightly inward along the sides. Such

a neighborhood is an adequate place to do the Kurzweil averaging. We would also want to choose these neighborhoods so that the tops and bottoms were contained in level surfaces of f . Let W be such a neighborhood. Suppose that for each $\varepsilon > 0$, we could find an integrable function $\chi_\varepsilon: W \rightarrow R$ such that

$$\mu - \varepsilon \leq \chi_\varepsilon \leq \nu + \varepsilon,$$

$$f(p) - \xi_\varepsilon \leq c - \int_0^{T_\gamma} \chi_\varepsilon(\gamma) \leq f(p) + \xi_\varepsilon$$

for all $\gamma \in \Gamma_\varepsilon(p)$ (c is the value of f at the top of W). Given such a family $\{\chi_\varepsilon\}$, it is easy to see how to proceed to obtain the desired approximation for f . However, it may be difficult to find such functions χ_ε .

3. C^∞ converse Lyapunov theorems. As the first application of the smoothing operations which we have described, we shall study the converse Lyapunov problem. We shall assume that the reader is familiar with the basic ideas of the Lyapunov theory (see [4], for instance).

Let M be a C^∞ paracompact Riemannian manifold, and let X be a continuous vector field on M which has unique trajectories, i.e., X induces a continuous dynamical system $f: M \times R \rightarrow M$ (for the definitions and basic properties of dynamical systems, see [8]). We shall denote the distance function on M by $d(\cdot, \cdot)$ and the related Hausdorff distance between subsets of M by $d_H(\cdot, \cdot)$.

Let A be a closed, invariant set of X . We say that (X, A) is uniformly asymptotically stable if it satisfies:

Stability. There is a continuous increasing function δ carrying R^+ (positive reals) onto itself such that $d[f(p, t), A] < \varepsilon$ whenever $d(p, A) < \delta(\varepsilon)$, and $t \geq 0$.

Uniform Attraction. There is a constant $\delta_0 > 0$ and a continuous decreasing function T carrying R^+ onto itself such that $d[f(p, t), A] < \varepsilon$ whenever $d(p, A) < \delta_0$, and $t \geq T(\varepsilon)$.

We define the domain of attraction D to be the largest invariant set whose trajectories all tend to A . D is an open set.

REMARK. If A is compact, then the condition of uniform attraction can be given more simply:

Attraction. There is a neighborhood U of A in which every point has its ω -limit set contained in A .

A necessary and sufficient condition for (X, A) to be uniformly asymptotically stable is the existence of a special kind of function, a Lyapunov function ([3], [4], or [6]). A Lyapunov function is a continuous nonnegative function $V: D \rightarrow R$ which satisfies

1. $V|_A = 0$, and $d_H[V^{-1}(c), A] > 0$ for all $c > 0$.
2. There is a positive definite function $a: R \rightarrow R$ and a positive function b defined on R^+ such that

$$-XV(p) \geq a[d(p, A)] \cdot b[V(p)].$$

3. $d_H[V^{-1}(c), A]$ is finite for all $c \geq 0$, and $\lim_{c \rightarrow 0} d_H[V^{-1}(c), A] = 0$.

4. If $\{p_n\}$ is any sequence in D which has its limit in the boundary of D , or whose distance from A becomes infinite, then $\lim V(p_n) = \infty$.

We will say that a Lyapunov function for (X, A) is Lipschitzian at A provided that there is a neighborhood U of A and a constant K such that $V(p) < K d(p, A)$ for all $p \in U$.

THEOREM 3.1. *Suppose that (X, A) is uniformly asymptotically stable with domain of attraction D . Then there is a continuous Lyapunov function for (X, A) which is Lipschitzian at A .*

Proof. Define $g: D \rightarrow R$ by $g(p) = \inf \{d[f(p, t), A] \mid t \geq 0\}$. Then $g[f(p, t)] \leq g(p)$ if $t \geq 0$, and $\delta[d(p, A)] \leq g(p) \leq d(p, A)$.

LEMMA. g is continuous on $D_0 = \{p \mid d(p, A) \leq \delta(\delta_0)\}$.

Proof. For each $p \in D$, $r > 0$, define

$$S_{p,r} = \{x \in D \mid d(x, A) < g(p) - r\}.$$

Then $p \notin f(S_{p,r}, t)$ for all $t \geq 0$. Note that if $p_n \rightarrow p$, then $\lim g(p_n) \leq g(p)$. We shall show that equality holds. For suppose that $\{p_n\}$ is a convergent sequence and $\limsup g(p_n) = g' < g(p)$. Then $g' < g(p) - r$ for some $r > 0$, and so there are negative times $\{t_n\}$ such that $f(p_n, t_n) \in S_{p,r}$ for n large. But then $\{p_n\}$ cannot have p for its limit.

Continuing with the proof of 3.1, we define $U: D_0 \rightarrow R$ by

$$U(p) = \max \{g[f(p, t)](1+2t)(1+t)^{-1} \mid t \geq 0\}.$$

Then $\delta[d(p, A)] \leq U(p) \leq 2d(p, A)$ and so U satisfies 1, 3, and is Lipschitzian at A . That U is continuous on D_0 follows immediately from the uniformity condition (we leave the verification to the reader). It remains to verify 2 and to extend U to D . Let $p \in D_0 - A$ and $h > 0$. Define $p' = f(p, h)$. Then

$$\begin{aligned} U(p') &= g[f(p', t')](1+2t')(1+t')^{-1} \quad (\text{some } t' > 0) \\ &= g[f(p, t)](1+2t)(1+t)^{-1}[1-h(1+2t')^{-1}] \\ &\leq U(p)[1-h(1+2T_p)^{-2}] \end{aligned}$$

where $t = t' + h$ and $T_p = \max_{0 \leq t \leq h} \{T(\frac{1}{2}\delta[d(f(p, t), A)])\}$. Therefore,

$$-XU(p) \geq U(p)(1+2T_p)^{-2}$$

and so 2 is satisfied on D_0 , if we define $a(x) = \delta(x)(1+2T[\delta(x)])^{-2}$ and $b(x) = x$. We shall now extend U to a suitable function on D . There is a $c_0 > 0$ such that $U^{-1}(c_0) \subset D_0$. Now every trajectory of X in $D - A$ intersects $U^{-1}(c_0)$ in a single point, and

so we can define a continuous function $t: D - A \rightarrow R$ by $U[f(p, t[p])] = c_0$. Define $V: D \rightarrow R$ by

$$\begin{aligned} V(p) &= U(p) && \text{if } t(p) \leq 0, \\ &= c_0 + t(p) && \text{if } t(p) \geq 0. \end{aligned}$$

Then $XV(p) = -1$ when $V(p) > c_0$, and V is continuous. Condition 4 follows easily from the fact that D is open (we leave the proof to the reader). Thus V is the desired Lyapunov function.

THEOREM 3.2. *Suppose that (X, A) is uniformly asymptotically stable with domain of attraction D . Then there is a C^∞ Lyapunov function for (X, A) defined on D .*

Proof. Let U denote a continuous Lyapunov function which is Lipschitzian at A . By 2.5, we can define a C^∞ function $W: D - A \rightarrow R$ satisfying $|W(p) - U(p)| < \frac{1}{2}U(p)$, and $XW(p) \leq \frac{1}{2}XU(p)$ for $p \in D - A$. W extends to all of D by $W|_A = 0$ (we call the extension W also). Now W is a continuous Lyapunov function which is Lipschitzian at A and C^∞ on $D - A$. Defining

$$V(p) = W(p) \exp(-[W(p)]^{-2}),$$

we obtain the desired C^∞ Lyapunov function.

Although we have considered only one of the several Lyapunov-type situations, it is clear that this general approach will apply to a variety of situations. In fact, one would expect that whenever there is given a C^0 Lyapunov function, then it can be smoothed by our methods. As an application of 3.2, we will now investigate the nonautonomous situation.

Let X be a continuous, time varying vector field on M which has unique trajectories, i.e., X can be considered as a continuous, time invariant vector field on $M \times R$ which has unique trajectories. If $A \subset M$ is a closed invariant set, then the definition of uniform asymptotic stability of (X, A) is slightly weaker than requiring $A \times R \subset M \times R$ to be uniformly asymptotically stable. The usual requirement is as follows:

There is a $t_0 \in R$ so that $A \times [t_0, \infty)$ is uniformly asymptotically stable.

Thus we would only expect a Lyapunov function satisfying 1, 2, 3, 4 on some domain of the form $D \times [t_0, \infty)$. We will leave the possible variations of the shape of the full domain of attraction to the imagination of the reader, and content ourselves with the following theorem.

THEOREM 3.3. *Suppose that $(X, A; t_0)$ is uniformly asymptotically stable. Let D be a neighborhood of A such that $D \times [t_0, \infty)$ is contained in the domain of attraction of $A \times [t_0, \infty)$. Then there is a C^∞ Lyapunov function $V: D \times [t_0, \infty) \rightarrow R$.*

In concluding this section, we make a few comments regarding the difficulties which arise in the autonomous case because we have allowed the attracting set A to be noncompact. The primary difficulty is that some of the trajectories which tend

towards A may have empty ω -limit sets. Such trajectories are annoying since "tending towards A " is then dependent on the particular choice of Riemannian metric. In fact, we can think of the Lyapunov function as the choice of a particular metric in which asymptotic stability is satisfied. Such a function (or metric) defines a set of neighborhoods for A which has a countable basis, and since A (noncompact) has an uncountable neighborhood basis, it seems that the study of Lyapunov functions for the case of noncompact attractors is rather meaningless, except in the presence of a specific metric. In particular, stating the definitions of stability and attraction in terms of neighborhoods instead of distance functions seems pointless.

4. Smoothing topological manifolds. We shall now turn our attention to the problem of when a topological manifold admits a differentiable structure. Our goal is a topological version of the Hirsch-Cairns smoothing theorem [1].

Let M^n be a topological manifold, and consider the topological vector bundle $M^n \times R^k$. The choice of k linearly independent cross sections fixes a continuously varying basis for the fibers, with respect to which the vector space operations are continuous. Having such a basis in mind, we define a *continuous inner product structure* for $M^n \times R^k$ to be a continuous mapping from M^n into $S_{k,0}$ (the space of positive definite, symmetric, $k \times k$ matrices). Since $S_{k,0}$ is contractible, any two inner product structures are homotopic.

We also define a *continuous family of l -planes* in $M^n \times R^k$ to be a continuous mapping from M^n into $G_{l,k-l}$. Relative to a fixed inner product structure, it is equivalent to specify either the continuous family of l -planes, or the continuous family of orthogonal complements.

We shall say that a differentiable structure for $M^n \times R^k$ is *linearly compatible* if the restriction to each fiber is compatible with the linear structure. We shall say that a differentiable structure is *uniformly linearly compatible* if for each $p \in M^n$ there is a neighborhood U of p in M^n and a diffeomorphism $h: U \times D^k \rightarrow R^{n+k}$ which is linear on each fiber (D^k is a k -disk). For example if M^n is a combinatorial manifold, then M^n can be triangulated as a Brouwer manifold, and so any combinatorially compatible differentiable structure for $M^n \times R^k$ is also uniformly linearly compatible.

Let X be the vector field defined on $M^n \times R^k$ by taking the field of inward radial vectors in each fiber.

LEMMA 4.1. *If $M^n \times R^k$ has a uniformly linearly compatible differentiable structure, then X has unique trajectories, X is uniformly asymptotically stable at $M^n \times 0$, and X is continuous on a neighborhood of $M^n \times 0$.*

Proof. The first two assertions are clear. We must show continuity, and for this we can work locally. By the uniform linear compatibility of the differentiable structure, it suffices to work in R^{n+k} with linear fibers, and with the trajectories being radial lines in the fibers. In this case, continuity follows by an easy sequential argument.

REMARK. This lemma does not hold if we only assume linear compatibility of the differential structure. The vector field X still has unique trajectories and is still asymptotically stable, but it may fail to be continuous, as is demonstrated by the following example:

Fill the plane with the family of smooth curves

$$\begin{aligned}\phi_x(t) &= (x, t), & -\infty < t < \infty \text{ if } x \leq 0, \\ &= \left(t + x, \varepsilon + \frac{x}{\pi} \tan \frac{\pi t}{x}\right), & -\frac{x}{2} < t < \frac{x}{2} \text{ if } x > 0.\end{aligned}$$

The homeomorphism which carries the x -axis onto itself and the vertical line over $(x, 0)$ onto ϕ_x by arc length defines a linear structure on each ϕ_x under which the plane is homeomorphic to $R^1 \times R^1$. The usual differentiable structure on the plane is linearly compatible with this product structure. However, along the line $y = \varepsilon$, the tangent vector to ϕ_x is vertical if $x \leq 0$ and has slope 1 if $x > 0$. Therefore, the vector field X has a discontinuity at $(0, \varepsilon)$. Since any diffeomorphism must preserve this discontinuity, there can be no admissible coordinate chart for a neighborhood of $(0, \varepsilon)$ which satisfies the uniform linear compatibility condition. While this example still leaves a neighborhood of the x -axis on which X is continuous, it is easy to see that a full counterexample can be constructed by putting small kinks of this kind arbitrarily close to the X -axis.

THEOREM 4.2. *Suppose that M^n is a topological manifold and that $M^n \times R^k$ has a uniformly linearly compatible differentiable structure α (for some k). Then there is a differentiable structure β for M^n and a smooth embedding of M^n_β into $(M^n \times R^k)_\alpha$ as a cross section which is transverse to the fibers.*

Proof. We proceed by induction on k . When $k = 1$, a noncritical level surface of a smooth Lyapunov function for the vector field X has two components, each of which will do for the desired smooth embedding of M^n . In general, the level surface of a smooth Lyapunov function provides a smooth submanifold which is transverse to the fibers and homeomorphic with $M^n \times S^{k-1}$. Let $(M^n \times S^{k-1})_\alpha$ denote such a differentiable submanifold and let M^n_0 denote any continuous cross section. X defines a continuous family of lines X_0 over M^n_0 and the family of orthogonal complements (relative to some inner product structure) is a manifold which is homeomorphic to $M^n \times R^{k-1}$. Projection parallel to X_0 gives a fiber preserving homeomorphism of a neighborhood of M^n_0 in $(M^n \times S^{k-1})_\alpha$ onto a neighborhood of M^n_0 in $M^n \times R^{k-1}$. The induced differentiable structure on $M^n \times R^{k-1}$ is uniformly linearly compatible, completing the induction. The desired smooth embedding is obvious in case $k = 1$.

Otherwise, we compose the homeomorphisms

$$M^n_\beta \rightarrow M^n \times R^{k-1} \rightarrow (M^n \times S^{k-1})_\alpha \rightarrow (M^n \times R^k)_\alpha.$$

Since these are precisely the homeomorphisms that we have used to obtain our

differentiable structure, the result is a differentiable embedding. M_β^n is transverse to the fibers since the level surfaces of the Lyapunov functions are.

REMARKS. 1. If M^n had a differentiable structure, we cannot guarantee to preserve it by our smoothing process unless $M^n \times 0$ is transverse to the fibers. For example, take $M^1 = \{(x, y) \mid y^3 = x\}$ and consider the uniformly linearly compatible differentiable structure which R^2 imparts on $M^1 \times R$ (the fibers are vertical lines). Our method would smooth M^1 to a horizontal line.

2. Using the viewpoint of the preceding remark, our theorem is a generalization of the Weierstrass approximation theorem, which in fact provided the motivation for this theorem. The author was disturbed by the fact that while combinatorial smoothings à la Cairns-Whitehead always require a local transverse condition (essentially a Lipschitz condition, cf. [9]) analytic smoothings of graphs of functions are never bothered by cusps, etc.

3. The simplest noncombinatorial example of a uniformly linearly compatible differentiable structure on $M^n \times R^k$ occurs when $M^n \subset R^{n+k}$ has a product normal microbundle neighborhood defined by a continuous field of topologically transverse k -planes (cf. [10]). We expect that these results can be generalized to the case where the bundle is not a product. For instance, our methods do produce a smooth D^k -bundle neighborhood for M^n whenever M^n has a normal microbundle neighborhood defined by a field of topologically transverse k -planes.

REFERENCES

1. M. W. Hirsch, *On combinatorial submanifolds of differentiable manifolds*, Comment Math. Helv. **36** (1962), 103–111.
2. L. Hörmander, *Linear partial differential operators*, Academic Press, New York, 1963.
3. J. Kurzweil, *On the inversion of Lyapunov's second theorem on stability of motion*, Czechoslovak Math. J. **81** (1956), 217–259, 455–484; English transl., Amer. Math. Soc. Transl. (2) **24** (1963), 19–77.
4. S. Lefschetz, *Differential equations: Geometric theory*, Interscience, New York, 1957.
5. J. Massera, *On the existence of Lyapunov functions*, Fac. Ingen. Agrimens. Montevideo Publ. Didact. Inst. Mat. Estadíst. **3** (1960), 111–124.
6. ———, *Contributions to stability theory*, Ann. of Math. **64** (1956), 182–206.
7. V. V. Nemytskii and V. V. Stepanov, *Qualitative theory of differential equations*, Princeton Univ. Press, Princeton, N. J., 1960.
8. H. Rademacher, *Über partielle und totale Differenzierbarkeit von Functionen mehrerer Variablen, und über die Transformation der Doppelintegrale*, Math. Ann. **79** (1918), 340–359.
9. S. Smale, *On gradient dynamical systems*, Ann. of Math. **74** (1961), 199–206.
10. F. W. Wilson, *Implicit submanifolds*, J. Math. Mech. **17** (1968), 229–236.

UNIVERSITY OF COLORADO,
BOULDER, COLORADO