JOURNAL OF APPROXIMATION THEORY **30**, 203–207 (1980)

Metric Entropy and Ordinary Differential Equations

ROBERT KAUFMAN

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

Communicated by G. G. Lorentz

Received May 8, 1979

1

Let C_1 be the set of vector fields $\Phi = (\varphi_1, \varphi_2)$ in the plane, each of whose components fulfills the boundedness conditions

$$|\varphi_i| \leqslant 1, \qquad |\varphi_i(z_1) - \varphi_i(z_2)| \leqslant |z_i - z_2|.$$

To each vector field Φ is associated a family of mappings (a flow) F_t:

$$F_0(z) = z, \qquad \frac{\partial}{\partial t} F_t(z) = \Phi(F_t(z)), \qquad -\infty < t < \infty.$$

We denote by S_1 the set of all mappings F_1, realized in this way from fields Φ in C_1, and by $S_1(\delta)$ the set of all mappings obtained under the further condition $\varphi_1 \geqslant \delta > 0$. In the following statement $N_\epsilon(S)$ is the number of sets required to cover a set of functions S, each set having diameter $< \varepsilon$ in the uniform memtric on the square Q: $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$.

THEOREM. $N_\epsilon(S_1) \geqslant \exp c_1 \varepsilon^{-2}$, $0 < \varepsilon < 1$,

$$N_\epsilon(S_1(\delta)) \leqslant \exp c_2(\delta)\, \varepsilon^{-4/3}, \qquad 0 < \varepsilon < 1.$$

2

The lower bound is very elementary and is included for completeness. Let $g(z) = (1 - |z|^2)^+$ and let $z_1, \ldots, z_N$ be points in Q, with $|z_i - z_j| \geqslant \varepsilon$ when $i \neq j$. This can be done so that $N \geqslant c\varepsilon^{-2}$. (From this point onward we use c to denote positive constants.) Now we set $\varphi_2 = 0$, $\varphi_1 = \Sigma \pm \varepsilon g(\varepsilon^{-1} z - \varepsilon^{-1} z_i)$. Each center z_i moves left or right at least $c\varepsilon$ up to time $y = 1$, according to

0021-9045/80/110203-05$02.00/0

640/30/3 4

the choice of $\pm$: we note that φ_1 never changes sign along any orbit $F_t(z)$. This leads to the lower bound $\exp c_1\varepsilon^{-2}$.

3

In obtaining the upper bound we make several reductions, but these have clearly no effect on the generality. First we suppose that φ_1 and φ_2 are of class $C^1(R^2)$ and then we suppose that $\varphi_1 = 1$ and $\varphi_2 = 0$ outside the large square $-2 \leqslant x \leqslant 3$, $-2 \leqslant y \leqslant 3$. (The second adjustment may require an increase in the Lipschitz constants.) Each mapping F_t is differentiable and its Jacobian (or differential) is a matrix $J(t, z)$. The entries of $J(t, z)$ for $0 \leqslant t \leqslant 1$ are uniformly bounded for all Φ in C_1. Moreover $J(t, z)$ depends differentiably on t, according to the equation of variation (compare [1, p. 96])

$$(\partial/\partial t)\, J(t, z) = A(F_t(z)) \cdot J(t, z)$$

with

$$A_{ij}(z) = \partial\Phi_i/\partial x_j \qquad (x_1 = x,\ x_2 = y).$$

From the identity

$$J(1, F_s(z)) \cdot J(s, z) = J(s + 1, z)$$

or

$$J(1, F_s(z)) = J(s + 1, z)\, J(s, z)^{-1},$$

we see that

$$\left| \frac{\partial}{\partial t} J(1, F_t(z)) \right| \leqslant c, \qquad \text{at} \quad t = 0.$$

It should be noted that the operator $(\partial/\partial t) f(F_t(z))$ at $t = 0$ equals $L(f)$, and $L = \varphi_1(\partial/\partial x) + \varphi_2(\partial/\partial y)$. Our inequality can be stated as follows for each component u of $F_1(z)$, and each operator $D = \partial/\partial x$ or $\partial/\partial y$ $|L(Du)| \leqslant c$, where L must be defined using the flow F_t. Now, formally, $LDu - DLu = g_1(x, y)\, u_x + g_2(x, y)\, u_y$, with $|g_1| + |g_2| \leqslant c$, so hat Lu satisfies a uniform Lipschitz condition in the plane. (The proof of the identity for $LD - DL$ proceeds by a smoothing of Φ so that u is C^2 and all derivatives have their usual interpretation.)

Let $0 < \varepsilon < 1$ and let C_1 be devided into sets B_r of diameter $\varepsilon^{2/3}$ in the uniform metric of R^2 (or, what is the same, on the square $-2 \leqslant x \leqslant 3$,

$-2 \leqslant y \leqslant 3$). The number of sets necessary here is $\exp c\varepsilon^{-4/3}$ (by [2, p. 153]). We choose one vector field ψ_1, ψ_2 from some B_r and try to estimate the functions u for vector fields Φ in B_r; $|\varphi_1 - \psi_1| < \varepsilon^{2/3}$, $|\varphi_2 - \psi_2| < \varepsilon^{2/3}$. Henceforth $L = \psi_1(\partial/\partial x) + \psi_2(\partial/\partial y)$, and $L(\varphi)$ is the operator with φ_1, φ_2.

We proved before that

$$|(L_\varphi u)(z_1) - (L_\varphi u)(z_2)| \leqslant c\,|z_1 - z_2|,$$

whence $|(L_\varphi u)(z_1) - (Lu)(z_2)| \leqslant c\,|z_1 - z_2| + c\varepsilon^{2/3}$. We map the half-plane $t \geqslant 0$, $-\infty < y < \infty$ to the half-plane $x \geqslant 0$ by the formula $H(t, y) = F_t(0, y)$. The rectangle $R: 0 \leqslant t \leqslant \delta^{-1}$, $|y| \leqslant 1 + 2\delta^{-1}$ covers Q, and t, y are coordinates of class C^1; the partials of t, y and of the inverse transformation are bounded by a function δ alone. Henceforth we write $u(t, y)$ in place of $u(H(t, y))$, observing that $\partial/\partial t = L$.

The functions $u(t, y)$ are divided into subsets of diameter ε in the uniform metric over the set composed of lines $t = m\varepsilon^{1/3}$, or rather the segements of these lines contained in R; the number of sets needed is $\exp c\varepsilon^{-4/3}$ by [2, Chap. 10] and elementary properties of the uniform metric.

Suppose now that $m\varepsilon^{1/3} < t < (m+1)\,\varepsilon^{1/3}$ and that t has the special form

$$m\varepsilon^{1/3} + \left(\sum_1^N a_j 2^{-j}\right)\varepsilon^{1/3}, \qquad a_j = 0 \text{ or } 1.$$

We write $t_0 = m\varepsilon^{1/3}$, $t_j = m\varepsilon^{1/3} + (\sum_1^j a_j\, 2^{-1})\,\varepsilon^{1/3}$ and use finite differences in the formula

$$u(t_N, y) = u(t_0, y) + \sum_1^N u(t_j, y) - u(t_{j-1}, y).$$

To represent all functions $u(t_N, y)$ we use $< c\varepsilon^{-1/3}\, 2^j$ differences $u(t_j, y) - u(t_{j1}, y)$. We are going to find the metrical properties of these differences as functions of y.

By the mean-value theorem

$$\begin{aligned} &u(t+h, y+k) - u(t+h, y) - u(t, y+k) + u(t, y) \\ &\quad = h[u_t(t+\theta h, y+k) - u_t(t+\theta h, y)], \qquad 0 < \theta < 1, \end{aligned}$$

whence the symmetric difference is at most $c\,|h|\,|k| + c\,|h|\,\varepsilon^{2/3}$. We suppose that $h = 2^{-j}\varepsilon^{1/3}$ ($j = 1, 2, 3,\ldots$); if $k = 2^j j^{-2}\varepsilon^{2/3}$ the estimate for the symmetric difference is less than $cj^{-2}\varepsilon + c\,2^{-j}\varepsilon \leqslant cj^{-2}\varepsilon$. Thus each function $u(t+h, y) - u(t, y)$ varies by at most $cj^{-2}\varepsilon$ when y varies on an interval of length $2^j j^{-2}\varepsilon^{2/3}$; clearly $|u(t+h, y) - u(t, y)| \leqslant c\,|h|$. By the procedure already cited [2], we can cover the set of all such functions by sets of diameter $cj^{-2}\varepsilon$, using no more than $\exp cj^2 2^{-j}\varepsilon^{-2/3} \exp c$ sets. As explained above, this number must be raised to the power $C\, 2^j\varepsilon^{-1/3}$ yielding $\exp c\, 2^j\varepsilon^{-1/3} \exp cj^2\varepsilon^{-1}$. This type of estimation is continued until $h = 2^{-j}\varepsilon^{1/3} < \varepsilon$, or $2^{-j} < \varepsilon^{2/3}$, so the distance between the lines, on which u is estimated, is at most ε. Thus $2^j \leqslant 2\varepsilon^{-2/3}$, so that multiplication of the estimates for these numbers j yields $< \exp c\varepsilon^{-4/3}$ sets. Because $\sum j^{-2}$ converges we obtain a covering by sets of diameter $c\varepsilon$.

4

The upper bound is somewhat surprising, in as much as F_1 need not belong to C^r for any $r > 1$. (When $1 < r < 2$, C^r is the class of function with derivatives in the Holder-class of exponent $r - 1$.) This, however, is not the main obstable to improving the exponent 4/3; if we study fields in which $\varphi_2 = 0$, $\varphi_1 \geqslant \delta > 0$, we can obtain any exponent > 1, by the interpolation process used above.

It is natural to look for co-ordinates u, v of class C^2, in which $\dot{v} = 0$, that is $\varphi_1(x, y)\, v_x + \varphi_2(x, y)\, v_y = 0$. This is a first-order partial differential equation, to which we must add $\nabla v \neq 0$. In contrast to the behavior of ordinary differential equations, solutions v are generally no smoother than Φ. To illustrate this we take the simplest example: $\varphi_1 \equiv 1$, $\varphi_2 = \varphi_2(y)$. We suppose that $\varphi_2(y) = 0$ for $y \leqslant 0$ and $\varphi_2 \in C^1(-\infty, \infty)$, while φ_2' increases on $(0, 1)$. Because $v_x = 0$ for $y \leqslant 0$, $v_y(0, y) = a \neq 0$. Proceeding formally from the relation $v_x + \varphi_2(y)\, v_y = 0$, we obtain $(v_y)_x + \varphi_2'(y)\, v_y + \varphi_2(y)\, v_{yy} = 0$. This has to be interpreted as a derivative along the trajectories of the system $\dot{x} = 1$, $\dot{y} = \varphi_2(y)$, as in Section 3, and can be justified if we first treat $(v_y)_x$ and v_{yy} as distributions, or generalized derivatives, [3, p. 49]. For economy in writing, we set $\psi = v_y$, so $\psi(0, y) = a \neq 0$; we assume as we can, that $a > 0$. Let $(x(t), y(t))$ be a solution of the differential system, with initial position $x(0) = 0$, $y(t) = \bar{y} > 0$. If $\bar{y}$ is sufficiently small, then $\bar{y} \leqslant y(t) \leqslant 2\bar{y} \leqslant 1$ over $0 \leqslant t \leqslant 1$, because $\varphi_2(y) = o(y)$ near $y = 0$. Along the solution $(x(t), y(t))$, we have $\dot{\psi} = -\varphi_2'(y)\, \psi \leqslant -\varphi_2'(\bar{y})\psi$, so that $\psi(x(1), y(1)) \leqslant \psi(0, \bar{y})$ $\exp -\varphi_2'(\bar{y})$. Both the initial point $(0, \bar{y})$ and the point $(1, y(1))$ have distance at most $2\bar{y}$ from the x-axis, on which $\psi = a$. The modulus of continuity of ψ, in the square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$, admits a lower bound $w(\bar{y}) \geqslant c\varphi_2'(\bar{y})$, and this means that v_y is no smoother than φ_2'. In particular, when $\varphi_2(y) = y^2$, for $y \geqslant 0$, v cannot be of class C^2.

The inequality $N_\varepsilon(S_1(\delta)) \geqslant \exp c_3(\delta)\, \varepsilon^{-1}$ is not difficult; we omit the proof because we do not know the best exponent.

REFERENCES

1. P. HARTMAN, "Ordinry Differential Equations," Wiley, New York, 1964.
2. G. G. LORENTZ, "Approximation of Functions," Holt, Rinehart & Winston, New York, 1966.
3. K. YOSIDA, "Functional Analysis," Springer--Verlag, Berlin, 1965.