COURANT 12

JÜRGEN MOSER
EDUARD J. ZEHNDER

LECTURE NOTES

Notes on Dynamical Systems

American Mathematical Society
Courant Institute of Mathematical Sciences

AMS
AMERICAN MATHEMATICAL SOCIETY

Jürgen Moser

Eduard J. Zehnder
ETH Zurich

12 Notes on Dynamical Systems

Courant Institute of Mathematical Sciences
New York University
New York, New York

American Mathematical Society
Providence, Rhode Island

2000 *Mathematics Subject Classification.* Primary 37–01, 37Kxx, 53Dxx, 58Exx, 70Fxx, 70H05.

For additional information and updates on this book, visit
www.ams.org/bookpages/cln-12

Library of Congress Cataloging-in-Publication Data

Moser, Jürgen, 1928–
Notes on dynamical systems / Jürgen Moser, Eduard J. Zehnder.
p. cm. — (Courant lecture notes in mathematics ; 12)
Includes bibliographical references.
ISBN 0-8218-3577-7 (alk. paper)
1. Differentiable dynamical systems. 2. Hamiltonian systems. 3. Transformations (Mathematics) 4. Combinatorial dynamics. I. Zehnder, Eduard, 1940– II. Title. III. Series.

QA614.8.M67 2005
515′.39—dc22 2005055871

∞ The paper used in this book is acid-free and falls within the guidelines established to ensure permanence and durability.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1 10 09 08 07 06 05

Notes on Dynamical Systems

Preface

Written in 1979–80, these notes constitute the first three chapters of a book that was never finished. It was planned as an introduction to the field of dynamical systems, in particular, of the special class of Hamiltonian systems. We aimed at keeping the requirements of mathematical techniques minimal but giving detailed proofs and many examples and illustrations from physics and celestial mechanics. After all, the celestial N-body problem is the origin of dynamical systems and gave rise in the past to many mathematical developments.

The first chapter is about the transformation theory of systems and also contains the so-called Hamiltonian formalism. The second chapter is devoted to periodic phenomena and starts with perturbation methods going back to H. Poincaré and local existence results due to Lyapunov and E. Hopf. Classical periodic solutions are established in the restricted 3-body problem and the celestial 3- and 4-body problems. Variational techniques then allow searching for global periodic orbits like closed geodesics on Riemannian manifolds and closed orbits on convex energy surfaces of general Hamiltonian systems. The Poincaré-Birkhoff fixed point theorem of an area-preserving annulus map in the plane is also proven in the second chapter. This theorem led to the V. Arnold conjectures about forced oscillations of time-periodic Hamiltonian systems on symplectic manifolds. Incidentally, after these notes were written, the Arnold conjectures triggered new developments in symplectic geometry and Hamiltonian systems. Also, it turned out that the periodic phenomena of Hamiltonian systems are intimately related to symplectic invariants and surprising symplectic rigidity phenomena discovered by Y. Eliashberg and M. Gromov. These more recent developments are presented in the book *Symplectic Invariants and Hamiltonian Dynamics* by H. Hofer and E. Zehnder.

The third chapter is devoted to a special and interesting class of Hamiltonian systems possessing many integrals. Following the construction of the so-called action and angle variables, illustrated by the Delaunay variables, several examples of integrable systems are described in detail. Chapters 4 and 5 should have dealt with the analytically subtle stability problems in Hamiltonian systems close to integrable systems known as KAM theory, and with unstable hyperbolic solutions, which, in general, do coexist with the stable solutions. Unfortunately, these chapters were never completed.

These notes owe much to Jürgen Moser's deep insight into dynamical systems and his broad view of mathematics. They also reflect his specific approach to mathematics by singling out inspiring typical phenomena rather than designing abstract theories.

Finally, I would like to thank Paul Wright for carefully checking these notes.

Contents

CHAPTER 1

Transformation Theory

1.1. Differential Equations and Vector Fields

(a) The flow of a system of differential equations. The object of these lecture notes are systems of ordinary differential equations of the form

$$\frac{dx}{dt} = f(x) \tag{1.1}$$

or in components,

$$\frac{dx_j}{dt} = f_j(x), \quad j = 1, \ldots, n, \tag{1.1'}$$

defined in an open domain $D \subset \mathbb{R}^n$. The right-hand side $f(x)$ is a vector-valued function mapping D into $\mathbb{R}^n$, belonging to $C^r(D, \mathbb{R}^n)$ for $r \geq 1$. We recall the well-known fact which will not be proven here that system (1.1) has a unique solution $x(t)$ for a given initial value $x(0) \in D$, where the solution $x(t)$ is defined on an interval $|t| < \delta$, $\delta > 0$. More precisely, if K is a compact subset of D, then there exists a $\delta > 0$ depending on K and f such that the solution $x(t)$ with initial values $x(0) \in K$ exists for the interval

$$I = \{t : |t| < \delta\}.$$

To indicate the dependence on the initial value $x(0)$ we denote this solution by

$$x(t) = \phi^t(x(0)).$$

Then, according to the standard existence theorem

$$\phi^t(x(0)) \in C^r(I \times K, D).$$

For fixed $t \in I$ we can view ϕ^t as a mapping of K into D, which satisfies

$$\phi^t \circ \phi^s = \phi^{t+s} \quad \text{for sufficiently small values of } |t| \text{ and } |s| \tag{1.2}$$

and

$$\phi^0 = \text{identity}. \tag{1.3}$$

This one parameter family of mappings ϕ^t is called the *flow* of system (1.1). Clearly we have

$$\frac{d\phi^t}{dt} = f(\phi^t) \quad \text{for sufficiently small } |t|.$$

Setting $t = 0$ we see that ϕ^t, in turn, determines f uniquely.

Setting $s = -t$ in (1.2) we see that ϕ^t has an inverse (defined on $\phi^t(K)$) which is also in C^r. We will call a C^r-mapping which has a C^r inverse a C^r-diffeomorphism. Thus ϕ^t is a C^r diffeomorphism, where defined.

We will also consider systems where f is C^∞ or real analytic (C^ω). The corresponding mapping ϕ^t is then C^∞ or C^ω, respectively.

Geometrically one interprets system (1.1) as a vector field, which assigns to each point $x \in D$ the vector $f(x)$. The solution $x(t) = \phi^t(x(0))$ is then a curve which at every point is tangent to this vector field. We will use the terms "system of differential equations" and "vector field" interchangeably.

(b) Transformation properties. We subject system (1.1) to an invertible coordinate transformation

$$x = u(y)$$

where we assume that the Jacobian matrix

$$\left(\frac{\partial u_i}{\partial y_j}\right) = u_y$$

is invertible. Then (1.1) goes over into a new system, say,

$$\frac{dy}{dt} = g(y)$$

where

$$g(y) = u_y^{-1} f(u(y)). \tag{1.4}$$

This is the transformation law for vector fields. If we denote by ψ^t the flow belonging to g, we have

$$\psi^t = u^{-1} \circ \phi^t \circ u \tag{1.5}$$

where $\circ$ indicates composition of the various diffeomorphisms and u^{-1} denotes the inverse map of u. Of course, the above relations have to be restricted to domains in which the mappings are defined.

To verify (1.5) we simply define ψ^t by (1.5) and then show that it agrees with the flow for g. Clearly $\psi^t =$ identity for $t = 0$ and differentiating the relation

$$u \circ \psi^t = \phi^t \circ u$$

we get

$$u_y \frac{d\psi^t}{dt} = f(\phi^t \circ u) = f(u \circ \psi^t) = u_y g(\psi^t)$$

hence

$$\frac{d\psi^t}{dt} = g(\psi^t).$$

Since ψ^t is uniquely determined by this relation and its initial value, (1.5) is proven.

The transformation law (1.4) is the same as that for the partial differential operator

$$X = \sum_{j=1}^{n} f_j(x) \frac{\partial}{\partial x_j}.$$

To describe the transformation laws under $x = u(y)$, observe that for any $h = h(x) \in C^1$ the expression

$$(Xh) \circ u$$

must be expressible in terms of a differential operator Y acting on $h \circ u$, i.e.,

$$(Xh) \circ u = Y(h \circ u).$$

We call Y the transformed differential operator and denote it by u^*X, so that

$$(u^*X)(h \circ u) = (Xh) \circ u. \tag{1.6}$$

If we write

$$u^*X = \sum g_k(y) \frac{\partial}{\partial y_k},$$

we find for the vector $g = (g_k)$ readily

$$g = u_y^{-1} f \circ u,$$

as we claimed.

There is a more direct relationship between the vector field (1.1) and X, namely

$$Xh = \frac{d}{dt} h(\phi^t)\Big|_{t=0},$$

i.e., Xh is the directional derivative of the function h along the vector field.

The operator X is determined by the vector field f and conversely X determines f; indeed for $h = x_j$ we find

$$Xx_j = f_j(x).$$

Therefore we will also use the operator X to describe the vector field. This is merely another notation which, however, has the advantage to reflect the transformation law under coordinate transformations. For this reason this notation is preferred in differential geometry and in the global study of vector fields on manifolds.

Incidentally, this notation shows that the vector fields in D form a Lie algebra since the commutator

$$XY - YX = [X, Y]$$

of two vector fields X, Y defines again a vector field. Indeed, if

$$X = \sum_{j=1}^{n} f_j(x) \frac{\partial}{\partial x_j}, \quad Y = \sum_{k=1}^{n} g_k(x) \frac{\partial}{\partial x_k},$$

then

$$[X, Y] = \sum_{j,k=1}^{n} \left(f_j \frac{\partial g_k}{\partial x_j} - g_j \frac{\partial f_k}{\partial x_j} \right) \frac{\partial}{\partial x_k}$$

since the second-order derivatives cancel. It is an almost obvious consequence of the definition (1.6) that

$$u^*[X, Y] = [u^*X, u^*Y]$$

so that the definition of $[X, Y]$ is independent of the choice of the coordinates.

(c) Local equivalence of vector fields. Two vector fields f, g which can be transformed into each other will be considered as equivalent; i.e., f, g are considered equivalent in some domains D_1, D_2, respectively, if there exists a diffeomorphism $u : D_2 \to D_1$ for which

$$g = u_y^{-1} f \circ u .$$

Only properties which are preserved under such transformation are of interest.

Therefore it is important to realize that locally, in the neighborhood of a point at which $f \neq 0$, it is equivalent to any other vector field with this property, e.g., to the vector field

$$\frac{dy}{dt} = e_1$$

where e_1 is the unit vector in the y_1 direction. Geometrically this statement simply means that in a small neighborhood of a point x^* where $f(x^*) \neq 0$ the flow can be mapped into the parallel flow

$$\psi^t(y) = y + te_1 .$$

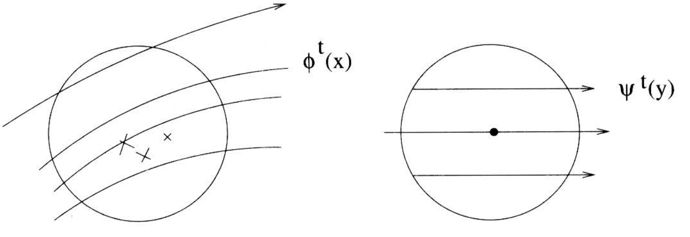

FIGURE 1.1

LEMMA 1.1 *If f, g define two vector fields with $f(x^*) \neq 0$, $g(y^*) \neq 0$, then there exist two neighborhoods D_1, D_2 of x^*, y^*, respectively, and a map $u : D_2 \to D_1$ such that $u_y^{-1} f \circ u = g$.*

PROOF: We may take $x^* = y^* = 0$ by applying a translation. By appropriate choice of the coordinate axes we may assume $f_1(0) = \langle f(0), e_1 \rangle \neq 0$. If $\phi^t(x)$ defines the flow of f, we set

$$u(y) = \phi^{y_1}(0, y_2, \dots, y_n) .$$

Then one computes readily

$$\det(u_y(0)) = f_1(0) \neq 0$$

so that $x = u(y)$ defines a diffeomorphism near $x = 0$. Moreover, with $\psi^t(y) = y + te_1$ we have

$$u \circ \psi^t(y) = \phi^{y_1 + t}(0, y_2, \dots, y_n) = \phi^t \circ u(y)$$

so that u maps ϕ^t into the parallel flow ψ^t. By differentiation we find

$$u_y^{-1} f \circ u = e_1 .$$

Similarly, we can construct a diffeomorphism v with

$$v_y^{-1} g \circ v = e_1 ,$$

thus $u \circ v^{-1}$ takes f into g. □

The assumption $f(x^*) \neq 0$ in the lemma is crucial. A point x^* at which $f(x^*) = 0$ is called a *singular point* (equilibrium point, stagnation point), of the vector field. If $x = u(y)$ maps a point y^* into $x^* = u(y^*)$, then

$$g = u_y^{-1} f \circ u$$

has a singular point at $y = y^*$ and the Jacobian is

$$g_y(y^*) = u_y^{-1} f_x(x^*) u_y$$

where $u_y = u_y(y^*)$. Thus the Jacobians $f_x(x^*)$, $g_y(y^*)$ at a stationary point are similar. Hence the eigenvalues of $f_x(x^*)$ are invariant and must be essential for the vector field. In fact, they are basic for the stability theory of vector fields at a singular point which was developed by Lyapunov.

It has to be mentioned that the above lemma is valid only "in the small" and fails in large domains. This is illustrated by three simple examples in the plane:

$$\text{(i)}\ \dot{x}_1 = x_1 \quad \text{(ii)}\ \dot{x}_1 = -x_2 \quad \text{(iii)}\ \dot{x}_1 = x_1$$
$$\dot{x}_2 = x_2 \qquad \dot{x}_2 = x_1 \qquad \dot{x}_2 = -x_2 .$$

The corresponding flows are plotted in Figure 1.2 and it is obvious that there is no

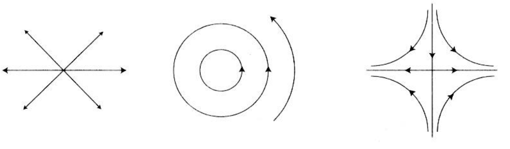

FIGURE 1.2

diffeomorphism taking any of these flows into any other—although this is possible locally near any point different from the origin.

The properties "in the large" are of principal interest. Examples of such properties are the existence of singular points, of periodic orbits, their stability or instability, etc., which will be investigated in these lecture notes.

Systems of differential equations of the form (1.1) are usually called *autonomous* to distinguish them from systems

$$\frac{dx}{dt} = f(x,t)$$

which depend on t as well, and are called *nonautonomous*. These systems can easily be reduced to (1.1) by introducing $x_0 = t + \text{constant}$ as an independent variable so that we obtain a system

$$\frac{dx_0}{dt} = 1, \quad \frac{dx_j}{dt} = f_j(x_0, x), \quad j = 1, 2, \ldots, n,$$

in $(n+1)$–dimensional space.

Also systems of second order

$$\frac{d^2x}{dt^2} = g\left(x, \frac{dx}{dt}\right)$$

can easily be reduced to (1.1), simply introducing x and dx/dt as independent variables. As a rule we will therefore assume that this reduction has been carried out and study systems of the form (1.1). The domain D is called the *phase space* in which we visualize the motion.

Examples. We illustrate this concept with some simple examples.

EXAMPLE 1. The differential equation

$$\frac{d^2x}{dt^2} + \sin x = 0, \quad x \in \mathbb{R}^1,$$

describe the motion of a pendulum, where x denotes the angle of deflection from the vertical.

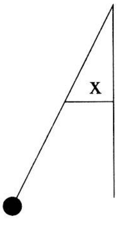

FIGURE 1.3

The phase space in this case is the plane with coordinates x and $\dot{x} = \frac{dx}{dt}$. Multiplying the equation by $\dot{x}$ and integrating we obtain the energy relation

$$\frac{1}{2}\dot{x}^2 - \cos x = E$$

where E is a constant along each orbit. This equation defines a set of curves (see Figure 1.4) on which the solutions travel.

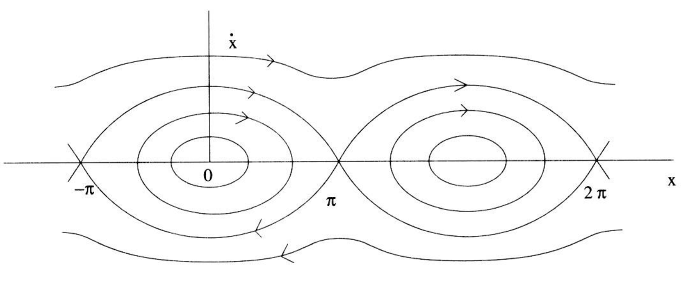

FIGURE 1.4

Without determining the solutions explicitly (they are given in terms of elliptic functions) we can read off the figure the nature of the motion: The oscillations about the down position is given by the islands, the motion of the pendulum swinging over the top by the wavy lines on top and bottom and the "separatrices" which describe a motion where the pendulum just goes from the top position, falls down, and asymptotically returns to the top position.

Since x is an angle we should identify the points $(x,\dot{x})$ and $(x_1,\dot{x}_1)$ if $x - x_1 = 2\pi j$, $\dot{x} - \dot{x}_1 = 0$ for any integer j. Thus the phase space becomes a cylinder and the many "islands" are identified to one.

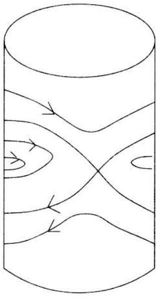

FIGURE 1.5

EXAMPLE 2 (Geodesics on $\mathbb{S}^2$). The two-dimensional sphere $\mathbb{S}^2$ can be given by the equation

$$|x|^2 = x_1^2 + x_2^2 + x_3^2 = 1$$

and the geodesics on it are the greatest circles. They are described by the differential equation

$$\frac{d^2x}{dt^2} = \lambda x$$

where the scalar λ is determined so that the equation $|x| = 1$ remains valid for all t. This requires $\lambda = -|\dot{x}|^2$ since

$$0 = \left(\frac{d}{dt}\right)^2 |x|^2 = 2(\langle x,\ddot{x}\rangle + |\dot{x}|^2) = 2(\lambda + |\dot{x}|^2),$$

and the differential equation becomes

$$\frac{d^2x}{dt^2} + |\dot{x}|^2 x = 0,$$

where we have to restrict ourselves to

$$|x|^2 = 1, \quad \langle x,\dot{x}\rangle = 0.$$

More precisely, if the last two conditions hold for $t = 0$ then they hold for all t. We see this by applying the uniqueness theorem for the initial value problem to the system

$$\frac{d^2z}{dt^2} = -2|\dot{x}|^2 z, \quad z(0) = 0, \quad \dot{z}(0) = 0,$$

where we take $z = \frac{|x|^2-1}{2}$.

What is the phase space in this case? If we set $y = \dot{x}$, we have a system of first order

$$\dot{x} = y, \quad \dot{y} = -|y|^2 x, \tag{1.7}$$

where $|x| = 1$, $\langle x, y\rangle = 0$. Thus the elements of the phase space are the vectors y attached at the points $x \in \mathbb{S}^2$. The set of vectors

$$\{(x, y) \in \mathbb{R}^6 : |x| = 1,\ \langle x, y\rangle = 0\}$$

form a manifold which is called the *tangent bundle* $T(\mathbb{S}^2)$ *of the sphere*. Thus $T(\mathbb{S}^2)$ is the phase space in this case.

The speed $|\dot{x}| = |y|$ is a constant along any solution since

$$\frac{d|y|^2}{dt} = 2\langle y, \dot{y}\rangle = -2|y|^2\langle y, x\rangle = 0$$

and we may restrict ourselves to the case $|\dot{x}| = |y| = 1$ in which case t is the arc length. Then the phase space is given by

$$\{(x, y) \in \mathbb{R}^6 : |x| = 1,\ \langle x, y\rangle = 0,\ |y| = 1\},$$

the unit tangent bundle $T_1(\mathbb{S}^2)$ of the sphere. Clearly all solutions are periodic of period 2π.

To give a better picture of this flow and its phase space we show that $T_1(\mathbb{S}^2)$ can be mapped one to one onto SO(3), the group of 3-by-3 orthogonal matrices U with determinant $+1$ and the differential equation becomes

$$\frac{dU}{dt} = UA \quad \text{where} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.8}$$

The solutions of this system are clearly

$$U(t) = U(0)e^{tA} = U(0)\begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.9}$$

The required mapping is obtained as follows: For $(x, y) \in T_1(\mathbb{S}^2)$ we construct the orthonormal frame $x, y, z = x \wedge y$[1]

$$Ue_1 = x, \quad Ue_2 = y, \quad Ue_3 = z;$$

i.e., x, y, z can be taken as the column vectors of U. Then writing $U = (x, y, z)$ we have

$$\frac{dU}{dt} = (\dot{x}, \dot{y}, \dot{z}) = (y, -x, 0) = UA.$$

Thus both the representation of the differential equation (1.7) on the unit tangent bundle $T_1(\mathbb{S}^2)$ and (1.8) on SO(3) are equivalent, in the sense that one can be transformed into the other. This illustrates the concept of equivalence of vector fields, but shows the lack of our previous definition, since we have to extend our concepts from the local representation to the global one on manifolds. We return to the definition of vector fields on manifolds later.

---

[1] $x \wedge y$ denotes the vector product in $\mathbb{R}^3$.

EXAMPLE 3 (Kepler problem in the plane). It is described by system

$$\frac{d^2x}{dt^2} = -\frac{x}{r^3}, \quad \frac{d^2y}{dt^2} = -\frac{y}{r^3}, \quad r^2 = x^2 + y^2. \tag{1.10}$$

This system of second-order differential equations has as its phase space the four-dimensional space $\mathbb{R}^4$ (with coordinates $x, y, \dot{x}, \dot{y}$) minus the plane $x = y = 0$, where the system is singular. This system possesses the energy integral

$$\frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \frac{1}{r}$$

which is constant, say E, along each orbit. It is well-known that the solutions correspond to conic sections in the x, y plane; hyperbola for $E > 0$, parabolas for $E = 0$ and ellipses for $E < 0$. Thus if we consider the energy surface

$$\frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \frac{1}{r} = E < 0$$

for a fixed negative E, all solutions are periodic and have, as it turns out, a fixed period (namely $2\pi(-2E)^{-3/2}$).

In the course of this chapter we will show that after an appropriate change of t and an appropriate compactification this flow of the Kepler problem on a fixed negative energy surface is equivalent to the flow (1.9) of the geodesics on $\mathbb{S}^2$. In particular, it will follow that the energy surface properly compactified is equivalent to SO(3).

For the following we will extend the concept of equivalence of two vector fields

$$\frac{dx}{dt} = f(x) \quad \text{and} \quad \frac{dy}{ds} = g(y)$$

in domains D_1 and D_2, respectively. We will say that f is *equivalent in the extended sense* if there exists a diffeomorphism $u : D_2 \to D_1$ and a positive function $\lambda = \lambda(y) \in C^r(D_2, R)$ such that

$$g = \lambda u_y^{-1} f \circ u. \tag{1.11}$$

The factor $\lambda = \lambda(y)$ corresponds to a change of the independent variable. More precisely, if the independent variable for the g-vector field is called s, i.e., if

$$\frac{dy}{ds} = g(y)$$

and ψ^s is the corresponding flow, then we have

$$\psi^s = u^{-1} \circ \phi^t \circ u \tag{1.12}$$

where s and t are related by

$$t = v(s, y) = \int_0^s \lambda(\psi^\sigma(y))d\sigma.$$

This shows that the solutions of one system are mapped into those of the other with a change of parametrization. We have clearly

$$u_y g(\psi^s) = u_y \frac{d\psi^s}{ds} = \frac{d}{ds}(\phi^t \circ u) = \lambda \frac{d}{dt}(\phi^t \circ u) = \lambda f \circ \phi^t \circ u$$

showing (1.11). In other words, here we subject the vector field to the transformation

$$x = u(y), \quad t = v(s, y),$$

of the $n+1$ dimension space $\mathbb{R}^n \times I$. We will simply write

$$\frac{dt}{ds} = \lambda .$$

REMARK. We assume throughout that $u \in C^r$, $r \geq 1$, but frequently one studies also "topological equivalence" of vector fields when u is assumed to be a homeomorphism only. In that case (1.4) losses meaning and topological equivalence is defined through that of the flows (see (1.12)) and an appropriate t-transformation. We will hardly be concerned with this concept and discuss it when it comes up.

Exercises

EXERCISE 1.

(a) Show with the example $\frac{dx}{dt} = x^2$ for $x \in \mathbb{R}^1$ that the flow ϕ^t is not defined for all t.

(b) Show if in system (1.1) with $D = \mathbb{R}^n$ and

$$|f(x)| < M \quad \text{in } \mathbb{R}^n,$$

then $\phi^t(x)$ is defined for all real t.

(c) Show, if in system (1.1), with $D = \mathbb{R}^n$

$$|f_x(x)| < M \quad \text{in } \mathbb{R}^n,$$

then ϕ^t is again defined for all real t.

EXERCISE 2.

(a) Let $f(x)$ be a C^1-vector field satisfying

$$f(x) = c \quad \text{for } |x| > r,$$
$$\langle f(x), c\rangle > 0 \quad \text{for all } x \in \mathbb{R}^n,$$

where c is a constant vector in $\mathbb{R}^n$ and r a positive constant. Let ϕ^t, ψ^t denote the flows corresponding to the vector fields $\dot{x} = f(x)$, $\dot{y} = g(y) = c$. Show that ϕ^t, ψ^t are defined for all t and that

$$u = \lim_{t\to\infty} \phi^{-t} \circ \psi^t$$

is a diffeomorphism satisfying

$$g = u_y^{-1} f \circ u .$$

(b) Use this result to give a proof of Lemma 1.1 by setting $g(x) = c$ and modifying f outside a small ball so that $f = c$ there.