Ordinary Differential Equations – an intuitive introduction

Content of later chapters and figures to be filled in

Mark Levi
Department of Mathematics
Penn State University
levi@math.psu.edu.

August 22, 2016

Contents

Chapter 1

Background

The subject of ODE (ordinary differential equations) is really a combination of calculus, geometry, linear algebra, even number theory, and more – easy if you know some key ideas and impossible if you don't. These ideas are reviewed in this chapter. I tried to make it short, self–contained, and intuitive.

And I kept in mind the fact that there is something in common between a good proof and a good joke: both should have a punchline; the joke at the end, and the proof at the beginning (or close to it).

1.1 Functions as transformations.

The independent variable will be denoted by t (rather than x) throughout, since it arises most commonly as the time in differential equations.

The usual way to picture a function $x = f(t)$ as a graph is fine, but there is a more basic way that is simpler, no less important, no less useful, and easy to generalizable to higher dimensions. Without much ado, here it is.

By the definition, a function is a pair of sets (call them D and R, the domain and the range), and a rule (call it f) assigning to each element $t \in D$ a number $x \in R$, as shown in Figure **??**. In the first two calculus courses, both D and R were subsets of $\mathbb{R}$, usually left unmentioned. the function can be thought of as a transformation. For example, $y = 2x$ is the stretching by the factor of 2; $y = -x$ is the reflection with respect to 0; and $y = x^2$ folds the line and maps it to the half–line $y \geq 0$.

Some intuition Here is a thing they often don't tell you: it helps to think of a function dynamically: think of the t–axis as a trackpad (one dimensional), and of the x–axis as a screen (also one–dimensional. Then you can get a feel of a function by imagining moving the "finger" t and watching the cursor $x(t)$ respond. For example, if $x = t^2$, as the finger t moves from $-\infty$ to $+\infty$, the cursor t^2 moves from $+\infty$, touches 0 and goes back to $+\infty$.

Exercise. Build in your mind a dynamical visualization of the following familiar functions: $1/t$, $\sin t$, $\cos t$, $\tan t$, e^t.

1.2 Derivative as a stretching factor.

There is another meaning of the derivative, besides the slope or the velocity. The derivative

$$f'(t) = \lim_{h\to 0} \frac{f(t+h) - f(t)}{h} \tag{1.1}$$

is the limit of the ratio of the lengths of two segments, the image and the premage – in other words, $f'(t)$ is simply *the stretching coefficient* at t. To test your intuition, see if it is obvious that $(2t)' = 2$, or that $(t^2)'_{t=0} = 0$.

In the "trackpad–cursor" interpretation, $f'(t)$ is the sensitivity coefficient. When you set the trackpad sensitivity in the computer preferences, you are prescribing the value of the derivative.

1.3 The chain rule explained intitively

Let us start with the obvious: if I stretch a piece of rubber band by the factor of (say) 2, and then follow this by another stretching by the factor of 3, then the net result is the $2 \cdot 3 = 6$–fold stretching. This simple fact is behind the chain rule; here is how.

Think of $[t, t+h]$ as the original rubber band which we stretch* by applying g to it, and then stretch the resulting interval again by applying f (see these two steps in Figure). The stretching coefficients are first $g'(t)$ and then $f'(g(t))$, so that the net stretching is the product; this explains the chain rule

$$\frac{d}{dt} f(g(t)) = f'(g(t))g'(t). \tag{1.2}$$

*The word "stretch" is meant in the generalized sense: stretching by a factor < 1 is really a contraction. And if the factor is negative, it involves a reflection followed by contraction or stretching.

1.4 The definition of $\exp(t)$ and of e

When speaking the function e^t, many calculus texts define

$$e = \lim_{h \to 0} \big(1 + h\big)^{1/h}, \tag{1.3}$$

or sometimes in an alternative form (obtained by restricting $1/h = n$ to be an integer):

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

either which may seem to have been pulled out of the hat, and then prove that

$$\frac{d}{dt} e^t = e^t. \tag{1.4}$$

When I was a student, my (otherwise very nice) textbook left me with a bad aftertaste: why define e by (1.3), an unmotivated formula? And the long proof of (1.4) led to suspicion that things may be simpler than they were presented.

I think the reason for my dissatisfaction was that the story was told in the reverse order. Here is what I think is a more direct way:
1. Define the exponential function $\exp(t)$ as the one that grows at the rate equal to the amount present, i.e. satisfying (1.5) below, and
2. prove that $\exp(t) = e^t$ where e is given by (1.3).

1. The definition. $f(t) = \exp(t)$ is the function which satisfies two conditions:

$$\begin{aligned} f'(t) &= f(t) \\ f(0) &= 1 \end{aligned} \tag{1.5}$$

Informally, $\exp(t)$ is the amount of money at time t (measured in years) with the starting balance \$1 ($t = 0$) and compounded continuously at the 100% annual rate.

2. Proof of $\exp(t) = e^t$. Fixing t, we divide $[0, t]$ into n short subintervals of length $h = t/n$. Approximating the derivative in (1.5) by the ratio we get the discretized problem

$$\frac{f_{k+1} - f_k}{h} = f_k, \quad f_0 = 1, \tag{1.6}$$

for $k = 0, \ldots, n-1$. Here f_k approximates $\exp(k\frac{t}{n})$, and in particular f_n approximates $f(t)$: $\lim_{n\to\infty} f_n = f(t)$; this is a special case of a general statement whose prove we omit here.* From (1.6) we get the recursion relation

$$f_{k+1} = (1+h)f_k, \quad f_0 = 1,$$

implying that $f_n = (1+h)^n$. Since $n = t/h$ by the definition, we obtain

$$f_n = (1+h)^{t/h} = \left((1+h)^{1/h}\right)^t. \tag{1.7}$$

Finally,

$$f(t) = \lim_{h\to 0} f_n = \lim_{h\to 0}\left((1+h)^{1/h}\right)^t \stackrel{(1.3)}{=} e^t.$$

The following two problems outline an alternative approach to showing that $\exp(t)$ defined by the initial value problem (1.5) is an exponential function e^t.

Problem 1. Show that if $f(t)$ satisfies $f'(t) = f(t)$ (for all t), then

$$f(t+s) = f(t)f(s) \ \text{ for all } \ t,\ s \in \mathbb{R} \tag{1.8}$$

without using the fact that $f(t) = ce^t$. You can use the fact (proven later on) that two functions satisfying the same ODE $\dot{x} = x$ with the same initial condition are equal (this is a very special consequence of the uniqueness theorem proven later on).

Problem 2. Show that if the function $f(t)$ satisfies (1.8) and if $f(0) = f'(0) = 1$ then $f(t) = f(1)^t$.

Problem 3. Consider the direction field in the (t, x)–plane associated with the ODE $\dot{x} = x$; in other words, the slope of the field at the point (t, x) is x.

1. What happens to this field under the stretching $(t, x) \mapsto (t, ax)$ in the x–direction, i.e. what is the resulting field? Interpret this statement in terms of the ODE $\dot{x} = x$.

2. The same direction field slope$(t, x) = x$ is subjected to the t–contraction by factor b: $(t, x) \mapsto (t/b, x)$. What is the resulting field? Interpret this in terms of the ODE $\dot{x} = x$.

The preceding problem gives a geometrical view of the fact that (i) any solution of $\dot{x} = x$ multiplied by a constant a is still a solution, and (ii) if the time t is measured in new units $\tau = t/b$, then in the new units x evolves according to $\frac{d}{d\tau}x = bx$ (think of t measured in days and τ measured in years, i.e. $b = 365$, and it stands to reason that the per annum rate is 365 times greater than the daily rate.)

Problem 4. Write a difference equation which is the analog of the ODE $\dot{x} = ax$ and solve it.

*Namely, it is the statement on convergence of Euler's method of numerical solution of ODEs. Our approximation of the solution of (1.5) follows Euler's method.

1.5 The FTC

Theorem 1 (The fundamental theorem of calculus–version 1)**.** *If f is a differentiable function defined on $[a,b] \subset \mathbb{R}$, where $a < b$, then*

$$\int_a^b f'(s)\, ds = f(b) - f(a). \tag{1.9}$$

Here is a common cosmetic restatement we actually use in computing integrals: in (1.9), f is treated as the "main actor". Let us think of the integrand $f' \stackrel{def}{=} g$ as given (as usually happens in practice, e.g., if we need to find $\int_a^b t^2\, dt$); then $f \stackrel{def}{=} G$ the antiderivative: $G' = g$.

The FTC then takes an equivalent form:

Theorem 2 (The fundamental theorem of calculus–version 2)**.** *Let g be a continuous function an an interval $[a,b] \subset \mathbb{R}$, and let G be an(y) antiderivative: $G'(t) = g(t)$ for all $t \in [a,b]$. Then*

$$\int_a^b g(s)\, ds = G(b) - G(a). \tag{1.10}$$

Here is one more version of FTC:

Theorem 3 (The fundamental theorem of calculus–version 3)**.** *Let g be a continuous function an an interval $[a,b] \subset \mathbb{R}$. Then*

$$\frac{d}{dt} \int_a^t g(s)\, ds = g(t). \tag{1.11}$$

This simple theorem relates two seemingly unrelated operations: integration on the left (known already to Archimedes) and differentiation (or its inverse) on the right, (developed about 2,000 years later). The theorem is fundamental because it links two seemingly unrelated operations. Archimedes computed some integrals 2,400 years ago; it took his genius to do so. Now millions of people can now do more than Archimedes could, thanks to the link the FTC provides between integration and the much easier differentiation.

The equivalence of the three versions The first two are just notational restatements of one another.
To prove V.2⇒V.3, set $b = t$ and differentiate both sides, using $G' = g$ afterwards.

It remains to prove V.3⇒V.2, i.e. that (1.11)⇒(1.10). Let us treat b in (1.10) as variable, so that we have to show that

$$\int_a^t g(s)\,ds = G(t) - G(a) \tag{1.12}$$

follows from (1.11). But both sides of (1.12) are antiderivatives of g: the left by (1.11) and the right by the definition. Moreover, these two antiderivatives are equal for one value of t, namely $t = a$, and therefore for all t. $\diamondsuit$

Having shown the equivalence of the three forms, we now get to the real gist of FTC.

Proof of the FTC Since all the forms of FTC are equivalent, it suffices to prove (1.11), i.e.

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} g(s)\,ds = g(t). \tag{1.13}$$

Wishing to estimate the integral, we set $\underline{g} = \min_{s\in[t,t+h]} g(s)$ and $\overline{g} = \max_{s\in[t,t+h]} g(s)$ (see Figure), so that

$$\underline{g} \le g(s) \le \overline{g} \ \text{ for all } \ s \in [t, t+h],$$

so that by integration $\underline{g}\,h \le \int_t^{t+h} g(s)\,ds \le \overline{g}\,h$ (see Figure), and thus

$$\underline{f} \le \frac{1}{h}\int_t^{t+h} g(s)\,ds \le \overline{f}. \tag{1.14}$$

Intuitively, (1.14) says that the average height of the graph of g is bounded by the height of the highest and the lowest points. By continuity of g, $\underline{f} \to g(t)$ and $\underline{f} \to g(t)$ as $h \to 0$, and by the sandwich theorem (1.13) holds. $\diamondsuit$

A moving interpretation of (1.11) The FTC in its form (1.11) says the almost obvious: if the boundary AB of the region in Figure moves with speed 1, the area changes at the rate equal to the boundary's length. The longer the moving boundary, the faster the area grows.

This idea applies to other shapes. For example, if the radius of a circle grows with speed 1, its length grows with speed equal to the boundary's length $2\pi r$. This confirms what we know: $\frac{d}{dr}\pi r^2 = 2\pi r$. Or, the same idea applies to the volume: $\frac{d}{dr}\frac{4}{3}\pi r^3 = 4\pi r^2$: the volume of the sphere grows at the rate equal to its area, if the radius grows with speed 1. Another nice application of this idea is to find the rate of change of the volume of a cone, if its height grows at a constant rate. For more applications of this idea see [], [].

Problem 5. Find a geometrical explanation/interpretation of the familiar facts

1. $\frac{d}{dx}x^2 = 2x$
2. $\frac{d}{dx}x^3 = 3x^2$
3. $\frac{d}{dt}uv = u'v + uv'$
4. $\frac{d}{dt}uvw = u'vw + u'w + uvw'$

To conclude, another transparent way to prove form (1.10) of FTC is to observe that its discretization according to the table on page 14.

Problem 6. State and prove the discrete equivalent of the fundamental theorem of calculus.

The following problem involves the discrete analog of antiderivative: given the first difference of a sequence, we are asked to find the sequence itself.

Problem 7. Find A_n in finite form, i.e. in the form such that the number of terms does not grow with n in each of the following cases:

1. $A_{n+1} - A_n = \frac{1}{n(n+1)}$
2. $A_{n+1} - A_n = \frac{2n+1}{n^2(n+1)^2}$
3. $A_{n+1} - A_n = \frac{1}{n(n+1)}$
4. $A_{n+1} - A_n = \frac{1}{n(n+1)}$.

Problem 8. Use the discrete version of the FTC to find the sums

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}, \quad \sum_{k=1}^{\infty} \frac{2k+1}{n^2(k+1)^2}, \quad \sum_{k=1}^{n} k^2, \quad \sum_{k=1}^{n} k^3.$$

1.6 The discrete–continuous analogy

This short section gives the discrete analogs of some continuous concepts.

Continuous	**Discrete**
$t \in \mathbb{R}$ independent variable	$n \in \mathbb{Z}$ indpendent variable
$f(t)$ value of f	a_n element of a sequence
$f'(t) = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$	$a_{n+1} - a_n$ first difference
$f''(t)$ second derivative	$(a_{n+1} - a_n) - (a_n - a_{n-1}) = a_{n-1} - 2a_n + a_{n+1}$, second difference
$\int_0^t f(s)\,ds$	$\sum_{k=0}^n a_k$
$\int_0^t f'(s)\,ds = f(t) - f(0)$, the FTC	$\sum_{k=0}^n (a_{k+1} - a_k) = a_{n+1} - a_0$, a telescoping sum
$\frac{d}{dt}\int_0^t f(s)\,ds = f(t)$, the FTC	$\sum_{k=0}^n a_k - \sum_{k=0}^{n-1} a_k = a_n$
$F'(t) = f(t)$: F is an antiderivative of f	$A_{n+1} - A_n = a_n$; A_n is an "antidifference" of a_n

1.7 Parametric equations of basic curves: ellipses, spirals, hyperbolas.

Problem 9. Identify the curve $ax^2 + 2bxy + cy^2 = 1$ (e.g. if it's an ellipse, find its major and minor axes and their orientation).

Problem 10. Consider the parametric curve

$$\begin{aligned} x &= a\cos t + b\sin t \\ y &= c\cos t + d\sin t. \end{aligned} \tag{1.15}$$

Show that:

1. the curve is an ellipse iff the coefficient matrix has a nonzero determinant.
2. Find the length and the direction of the semiaxes of the ellipse.
3. * Show that the directions of the ellipse's axes are the eigendirections of the matrix $A^T A$, where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1.8 The linearization of a function

In Calculus III the linearization $L(x, y)$ of a function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ at the point (x_0, y_0) was defined by the property that the graph of L is the tangent plane to the graph of f at $(x_0, y_0, f(x_0, y_0)) \in \mathbb{R}^3$, or equivalently, by

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{1.16}$$

Equivalently, L can be defined by the property that L is a linear function such that

$$|f(x, y) - L(x, y)| = o(|x - x_0| + |y - y_0|). \tag{1.17}$$

Using this as the definition is more honest because it makes the whole reason for introducing L explicit, namely that it approximates f. To re–emphasize this approximation property of L, one often writes (1.17) in this form:

$$f(x, y) = L(x, y) + o(|x - x_0| + |y - y_0|). \tag{1.18}$$

Problem 11. Show that if L satisfies (1.17) then it satisfies (1.16) and conversely, assume that f has continuous partial derivatives.

Hint: Prove that for a function g of a single variable $g(x+h) = g(x) + g'(x)h + o(h)$ and use this fact, along with the continuity of partial derivatives.

Recall that the differential of f is simply the "increment part" of L, i.e. the linear approximation to the change of f; in the commonly used notations $dx = x - x_0$, $dy = y - y_0$ for the deviations from x_0, y_0,

$$df = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy.$$

Problem 12. The dimensions of a rectangular solid were with dimensions a, b, c were changed by amounts da, db, dc. Use the differential to find an approximate resulting change of the volume, both absolute and relative.

The approximation idea (1.18) can be used to prove the chain rule.

Problem 13. Prove the chain rule $\frac{d}{dt} f(x(t), y(t)) = f_x x' + f_y y'$.

1.9 The directional derivative and the gradient

We are looking at scalar functions $f : \mathbb{R}^n \mapsto \mathbb{R}$, focusing on the planar case $n = 2$; since the higher dimensional case $n > 2$ offers no extra insights, it would be just a diversion from the substance.

The following problem captures two concepts in one formula: the directional derivative and the gradient.

Problem 14. Let $\mathbf{u}$ be a *unit* vector in $\mathbb{R}^2$, and $\mathbf{x} = \langle x, y\rangle$. Show:

$$\left.\frac{d}{ds}\right|_{s=0} f(\mathbf{x} + s\mathbf{u}) = \langle f_x, f_y\rangle \cdot \mathbf{u}, \tag{1.19}$$

where $\cdot$ denotes the dot product and where $f_x = f_x(x, y)$, etc. Hint: use the chain rule.

The left–hand side of (1.19) is the directional derivative of f at $\mathbf{x}$ in the direction $\mathbf{u}$, denoted by $D_{\mathbf{u}}f(\mathbf{x})$. It is the rate of change of f *per unit length* traveled along $\mathbf{u}$, when passing (x, y). The vector $\langle f_x, f_y\rangle$ in (1.19) is called the gradient of f at $\mathbf{x}$, denoted by $\nabla f(\mathbf{x})$, or simply ∇f if $\mathbf{x}$ is clear from context.

Here is a self–contained definition of ∇f, motivated by (1.19).

Definition. Given a function $f : \mathbb{R}^n \mapsto \mathbb{R}$, $\nabla f(\mathbf{x})$ is that vector $\mathbf{V} \in \mathbb{R}^n$ which satisfies

$$D_{\mathbf{u}}f(\mathbf{x}) = \mathbf{V} \cdot \mathbf{u} \quad \text{for all} \quad \mathbf{u} \in \mathbb{R}^n,\ |\mathbf{u}| = 1. \tag{1.20}$$

This definition is sometimes better than the formula $\nabla f = \langle f_x, f_y\rangle$, particularly if we want to keep things in vector form and do not want to stoop to discussing coordinates. Here is an example: find the gradient of $f(\mathbf{x}) = A\mathbf{x}\cdot\mathbf{x}$, where A is a $n \times n$ matrix (with any integer $n > 0$) and $\mathbf{x} \in \mathbb{R}^n$.* Solution: substituting f into (1.20), we obtain

$$\left.\frac{d}{ds}A(\mathbf{x} + s\mathbf{u}) \cdot (\mathbf{x} + s\mathbf{u})\right|_{s=0} = A\mathbf{x} \cdot \mathbf{u} + A\mathbf{u} \cdot \mathbf{x} = (A + A^T)\mathbf{x} \cdot \mathbf{u}.$$

According to the above definition, ∇f is the factor multiplying $\mathbf{u}$, so that

$$\nabla(A\mathbf{x} \cdot \mathbf{x}) = (A + A^T)\mathbf{x}. \tag{1.21}$$

We managed to avoid mentioning coordinates, and the proof works for any n, not just $n = 2$.

*matrix multiplication is taken before the dot product, i.e. $A\mathbf{x} \cdot \mathbf{x} \equiv (A\mathbf{x}) \cdot \mathbf{x}$.

Problem 15. Prove the characteristic properties of ∇f:

1. $|\nabla f(x,y)|$ gives the maximal rate of change of f per unit length at (x,y).
2. ∇f points in the direction of the maximal rate of increase of f per unit length at (x,y).
3. $\nabla f(x,y)$ is perpendicular to the level curve of f passing through (x,y).

1.10 Simplest linear vector fields.

The following vector fields are building blocks to which any linear vector field reduces, in a certain sense. This reduction is done later.

1. $\mathbf{F}_d(\mathbf{x}) = \langle x, y\rangle$ (dilation/contraction flow).
2. $\mathbf{F}_r(\mathbf{x}) = \langle -y, x\rangle$ (rotation flow)
3. $\lambda\mathbf{F}_d(\mathbf{x}) + \omega\mathbf{F}_r(\mathbf{x})$ for different combinations of signs of λ and ω.
4. $\mathbf{F}_h = \langle x, -y\rangle$ (hyperbolic vector field).
5. $\mathbf{F}_s = \langle y, 0\rangle$ (shear vector field).

1.11 Conservative vector fields; potential

Definition. A vector field $\mathbf{F}$ in $\mathbb{R}^n$ is said to be *conservative* if for any close curve γ one has

$$\oint_\gamma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = 0. \tag{1.22}$$

If the domain D of $\mathbf{F}$ is not simply connected, then one adds an extra condition to the definition, namely that γ be contractible in D.

Gravitational and electrostatic fields are conservative – otherwise we could build a perpetual motion machine by choosing γ for which the integral (1.22) is positive, and sell this positive work done by the field.

Definition. Let $\mathbf{F}$ be a conservative vector field on $\mathbb{R}^n$. Fix a point $\mathbf{x}_0 \in \mathbb{R}^n$. The function

$$P(\mathbf{x}) = -\int_\gamma \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} \tag{1.23}$$

where γ is a curve connecting $\mathbf{x}_0$ with $\mathbf{x}$, is called a potential of $\mathbf{F}$. (Note the indefinite article: there are infinitely many potentials, but they all differ by a constant.)

Problem 16.

1. Show that if the domain is simply connected, then $P(\mathbf{x})$ is independent of the choice of $\gamma_{\mathbf{x}_0\mathbf{x}}$.
2. Show that replacing a reference point $\mathbf{x}_0$ with a different point $\mathbf{x}_1$ changes the potential P by a constant.
3. Show that $\mathbf{F} = -\nabla P$.

If the domain is not simply connected, then P can be multiple–valued, as Problem **??** illustrates.

Problem 17. Show that the vector field $\mathbf{F} = \langle -y, x\rangle/(x^2 + y^2)$ is conservative, and that its potential P is a multiple–valued function. Can you find P?

Problem 18. Show that the vector field $\mathbf{F} = \langle x, y\rangle/(x^2+y^2)$ is conservative and find its potential.

A physical interpretation of P. Think of $\mathbf{F}(x)$ as the force (gravitational, say) acting on a point particle. To keep the particle in place, or to move it infinitesimally slowly, I have to apply force $-\mathbf{F}(x)$ (to cancel $\mathbf{F}$). Thus the work I have to to do to move along γ from $\mathbf{x}_0$ to $\mathbf{x}$ is precisely $\int_\gamma(-\mathbf{F}) \cdot d\mathbf{r} \overset{def}{=} P(\mathbf{x})$.

Theorem 4. *Any vector field given by a gradient: $\mathbf{F} = \nabla f$ is conservative. The converse is also true locally: if $\mathbf{F}$ is conservative then for any $\mathbf{x}$ in the domain of $\mathbf{F}$ there exists a neighborhood $\mathcal{N}$ of $\mathbf{x}$ and a function $f : \mathcal{N} \mapsto \mathbb{R}$ such that $\mathbf{F} = \nabla f$ on $\mathcal{N}$.*

Proof. Let us restrict attention to a neighborhood $\mathcal{N}$ of point $\mathbf{x}$ in the domain of $\mathbf{F}$. Since $\mathbf{F}$ is conservative on a simply connected set $\mathcal{N}$, the potential $P(\mathbf{x})$ of $\mathbf{F}$ is well defined on $\mathcal{N}$. And by a preceding problem,

$$\mathbf{F} = -\nabla P$$

for all $\mathbf{x} \in \mathcal{N}$. It remains to set $f = -P$, to conclude that $\mathbf{F} = \nabla f$. ♢

Problem 19. Show that the gravitational field $\mathbf{F}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|^3$, where $\mathbf{x} \in \mathbb{R}^3$ is conservative.

Problem 20. Show that the vector field $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$ is conservative if and only if the matrix A is symmetric.

Problem 21. Find the potential of the vector field $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$, where A is a symmetric matrix.

1.12 Divergence

The divergence of a vector field quantifies precisely what the word suggests: think of $\mathbf{F}$ as the velocity field (rather than the force field); the divergence at a point is the *rate of change of an infinitesimal area around this point, per unit area*, as the area is carried by the flow $\mathbf{F}$. Here is a precise definition.

Definition. The divergence of the vector field $\mathbf{F}$ in $\mathbb{R}^2$ is the limit*

$$\mathrm{div}\mathbf{F}(\mathbf{x}) = \lim_{|D|\to 0} \frac{1}{|D|} \oint_{\partial D} \mathbf{F} \cdot \mathbf{N} ds, \tag{1.24}$$

where D denotes a domain bounded by a closed curve ∂D, $|D|$ is the area of D and $\mathbf{N}$ is the outward normal to γ.

To see that this formula indeed expresses the idea of the first sentence of this section, note that $\mathbf{F} \cdot \mathbf{N}$ is the velocity component normal to γ, and thus $\mathbf{F} \cdot \mathbf{N}\, ds$ is the area swept per second by the moving arc ds. The integral in (1.24) thus expresses the rate of change of area of the region D moving according to $\mathbf{F}$.

An alternative interpretation of the integral in (1.24) is the *flux*, the amount of gas crossing the *fixed* curve γ, although what we mean by "amount" is vague unless we introduce the concept of density – a redundant thing at this stage.

Problem 22. Using the definition, find the divergence of each of the vector fields in Section 1.10

Problem 23. Using the definition, derive the formula $\mathrm{div}\mathbf{F} = P_x + Q_y$, where $P = P(x, y)$ and $Q(x, y)$ are the components of $\mathbf{F}$.

The following problem asks the same question but in polar coordinates.

Problem 24. The vector field in the plane is given in polar coordinates as follows: for a particle carried by the flow, its polar coordinates r, θ change at the prescribed rates $R(r, \theta),\ T(r, \theta))$. Find the expression of the divergence at r, θ using the definition. Verify the formula on the two examples from problems 17 and 18.

*We leave aside the question of existence of this limit and of its independence of the choice of D.

Derivative as divergence in 1D. Let us interpret $f(x)$ as the vector field on $\mathbb{R}$ (one can think of a highway with the car finding itself at x obliged to move precisely with speed $f(x)$. Then $f(x+h)-f(x)$ is the speed at which the space between two cars distance h apart grows, and so

$$\frac{f(x+h)-f(x)}{h}$$

is the rate of growth of this distance per unit distance (essentially, the exponential rate of growth of distance between two cars). Thus f' is exactly the one dimensional divergence.

Interest rate as divergence. Consider a bank account that is compounded continuously at the annual rate r, which by the definition can be written as $\frac{\dot{x}}{x}=r$. Geometrically representing x as an interval $[0,x]$, we see that the rate of its elongation per unit length (which rate is the 1D divergence) is r.

Problem 25. Find $\mathrm{div} A\mathbf{x}$, where A is an $n\times n$ matrix. What is the divergence if A is anti–symmetric? What is a geometrical explanation of this in $\mathbb{R}^2$?

Problem 26. In this section we have been interpreting $\mathbf{F}$ as the velocity field. What property does $\mathrm{div}\,\mathbf{F}$ characterize for a gravitational vector field $\mathbf{F}$, or for an electrostatic field $\mathbf{F}$?

Remark. Since $\mathrm{div}\,\mathbf{F}$ at a point depends only on the partial derivatives at that point, we conclude that $\mathrm{div}\,\mathbf{F}$ at a point depends only on the linearization of $\mathbf{F}$ at that point.

1.13 Curl in 2D

Let us think of a vector field $\mathbf{F}$ in $\mathbb{R}^2$ as the velocity field of imagined gas in the plane. The curl makes precise the vague concept of rotation, as follows.
Definition. Given a vector field $\mathbf{F}$ in $\mathbb{R}^2$, the curl is defined as

$$\mathrm{curl}\mathbf{F}(\mathbf{x})=\lim_{|D|\to 0}\frac{1}{|D|}\oint_{\gamma}\mathbf{F}(\mathbf{y})\cdot d\mathbf{y} \tag{1.25}$$

were D is a domain containing $\mathbf{x}$ and enclosed by a smooth closed curve γ and $|D|$ denotes the area of D. The integral of tangential velocity is called the circulation of $\mathbf{F}$ around γ.

If we think of $\mathbf{F}$ as the force field, rather than a velocity field, then the curl quantifies the non–conservativeness of the field.

Problem 27. 1. Show that the conservative fields have zero curl.

2. Let $A = (a_{ij})$ be any 2×2 matrix. Show that

$$\operatorname{curl}(A\mathbf{x}) = a_{21} - a_{12}, \quad \text{for any} \quad \mathbf{x} \in \mathbb{R}^2. \tag{1.26}$$

3. What is a necessary and sufficient condition on A for the vector field $A\mathbf{x}$ to be conservative?

Problem 28. Using the definition, find the curl of each of the vector fields in Section 1.10

Problem 29. Let $\mathbf{F}$ and $\mathbf{G}$ be two vector fields. Show that curl is a linear operation, i.e. that for any $a,\ b \in \mathbb{R}$ we have

$$\operatorname{curl}(a\,\mathbf{F} + b\,\mathbf{G}) = a\operatorname{curl}\mathbf{F} + b\operatorname{curl}\mathbf{G}. \tag{1.27}$$

Problem 30. 1. Show that the curl of a constant vector field is zero.

2. Show that if $\mathbf{r} = \langle p, q\rangle$ satisfies (1.30), and if p, q have continuous derivatives in a neighborhood of $\mathbf{x} = 0$, then $\operatorname{curl}\mathbf{r}\,(\mathbf{x})_{\mathbf{x}=\mathbf{0}} = 0$.

Theorem 5. *Assume that* $\mathbf{F}$ *is twice continuously differentiable.** *The curl of* $\mathbf{F} = \langle P, Q\rangle$ *is given by*

$$\operatorname{curl}\mathbf{F} = Q_x - P_y, \tag{1.28}$$

where the subscripts denote partial derivatives.

Problem 31. If $\mathbf{F} = \langle P, Q\rangle$ is a velocity field, find an interpretation of Q_x and P_y in terms of certain angular velocities, obtaining thereby an interpretation of $\operatorname{curl}\mathbf{F}$.

Proof of Theorem 5. The gist of the proof is to find the curl of the linear part of $\mathbf{F}$ – which we already did, in fact, in (1.26), and to show that this is the only part of $\mathbf{F}$ that contributes to curl.

Without the loss of generality, we take $\mathbf{x} = \mathbf{0}$. Let D be a rectangle with one vertex at $\mathbf{0}$, Figure. We extract the linear part of $\mathbf{F}$, as planned, by Taylor expanding the components of $\mathbf{F}(\mathbf{x})$ around $\mathbf{0}$:

$$P(x, y) = a_1 + a_{11}x + a_{12}y + p(x, y), \quad Q(x, y) = a_2 + a_{21}x + a_{22}y + q(x, y),$$

*The smoothness assumption can be weakened, but we have better ways to spend time.

with constants

$$a_1 = P(0,0), \quad a_{11} = P_x(0,0), \quad a_{12} = P_y(0,0), \quad \text{etc.},$$

and with the remainder terms satisfying $|p|,\ |q| = O(x^2+y^2)$ in a neighborhood of $\mathbf{x} = \mathbf{0}$. Summarizing this in vector form, we obtained

$$\mathbf{F}(\mathbf{x}) = \mathbf{c} + A\mathbf{x} + \mathbf{r}(\mathbf{x}), \tag{1.29}$$

where

$$\mathbf{c} = \mathbf{F}(\mathbf{0}), \quad A = \left.\begin{pmatrix} P_x & P_y \\ Q_x & Q_y \end{pmatrix}\right|_{\mathbf{x}=0},$$

and

$$|\mathbf{r}| \le C(x^2+y^2) \tag{1.30}$$

for some constant C and for all x, y in a neighborhood of $x = y = 0$.*

Applying curl to both sides of (1.29) we conclude (using additivity property of curl, (1.27)), that

$$\operatorname{curl}\mathbf{F}(\mathbf{x}) = \operatorname{curl}\mathbf{c} + \operatorname{curl}A\mathbf{x} + \operatorname{curl}\mathbf{r}(\mathbf{x}). \tag{1.31}$$

But $\operatorname{curl}\mathbf{c} = \operatorname{curl}\mathbf{r}(\mathbf{x})_{\mathbf{x}=0} = 0$ by Problem 30, while $\operatorname{curl}A\mathbf{x} \overset{(1.26)}{=} a_{21} - a_{12} = Q_x - P_y$. Therefore setting $\mathbf{x} = 0$ in the above formula gives (1.28), as claimed. ♢

1.14 Green's and Stokes' theorems in $\mathbb{R}^2$

The theorems, if viewed properly, are not much more than the restatements of the definitions of div and curl, as the proofs given below explain.

Theorem 6. *Given a smooth vector field* $\mathbf{F}$ *and a bounded region* D *enclosed by a piecewise smooth closed curve* ∂D*, one has*

$$\int\!\!\int_D \operatorname{div}\mathbf{F}\,dA = \int_{\partial D} \mathbf{F}\cdot\mathbf{N}\,ds, \tag{1.32}$$

in the same notations as used in the definition (1.24) of the divergence, and

$$\int\!\!\int_D \operatorname{curl}\mathbf{F}\,dA = \int_{\partial D} \mathbf{F}\,d\mathbf{r}. \tag{1.33}$$

*(1.29) is just the vector version of Taylor's formula, which we could have referred to instead of deriving, and which we could have written in the vector form

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{0}) + \mathbf{F}'(\mathbf{0})\mathbf{x} + \mathbf{r}(\mathbf{x})$$

which looks the same as the scalar case.

(1.32) is called the *divergence theorem*, and it holds for any dimension; (1.33) is the 2D version of Stokes's theorem. There is a screaming similarity of these theorems to the definitions of div and curl, a fact that makes the proofs of both theorems straightforwards.
Here are the two theorems in scalar form, substituting $\mathbf{F} = \langle P, Q\rangle$, $\mathbf{r} = \langle x, yl\rangle$ and $\mathbf{N}\,ds = Jd\mathbf{r} = \langle -dy, dx\rangle$:

$$\int\int_D (P_x + Q_y)\,dA = \int_{\partial D} -P\,dx + Q\,dy, \tag{1.34}$$

$$\int\int_D \mathrm{curl}\,(Q_y - P_x)\,dA = \int_{\partial D} P\,dx + Q\,dy. \tag{1.35}$$

Proof. The proofs of (1.32) and (1.33) are identical (or, one can derive one from the other by changing the names of P and Q), so we concentrate on (1.32).

Partitioning the region D into $O(N^2)$ small subregions D_i each of diameter $1/N$ as shown in the Figure, we have from (1.24):

$$\oint \partial D_i \mathbf{F}\cdot\mathbf{N}\,ds = (\mathrm{div}\,\mathbf{F}(\mathbf{x}_i) + r_i)|D_i|, \tag{1.36}$$

where the remainder is small uniformly in i, i.e. there exists $C > 0$ independent of N such that $|r_i| \le C/N$ for all i (the proof of this is left as an exercise). We now sum (1.36) for all i. Now *the integrals over shared edges cancel*, since the outward normals $\mathbf{N}$ at the shared boundary points of two neighboring subdomains are equal and opposite. Only the integrals over the unshared pieces of ∂D_i remain after summation, and we obtain

$$\oint_{\partial D} \mathbf{F}\cdot\mathbf{N}\,ds = \sum_i \mathrm{div}\,\mathbf{F}(\mathbf{x}_i)|D_i| + \sum_i r_i|D_i| \tag{1.37}$$

Now the first sum is a Riemann sum of the desired integral, while the second sum becomes small for large N:

$$\sum_i r_i|D_i| \le \sum_i \frac{C}{N}|D_i| = \frac{C|D|}{N}.$$

We conclude that in the limit $N \to \infty$ the right–hand side approaches $\int\int_D \mathrm{div}\,\mathbf{F}\,dA$. $\diamondsuit$

Problem 32. In this problem, a vector $\mathbf{v} \in \mathbb{R}^2$ has been rotated by $\pi/2$ counterclockwise, with the resulting vector denoted by $\mathbf{v}^{\perp}$.

1. If $\mathbf{v} = \langle x, y \rangle$, what are the coordinates of $\mathbf{v}^{\perp}$?
2. Find the matrix J that turns (pardon the pun) $\mathbf{v}$ into $\mathbf{v}^{\perp}$, i.e. such that $\mathbf{v}^{\perp} = J\mathbf{v}$.

Problem 33. Find the matrix $R(\theta)$ that rotates vectors in $\mathbb{R}^2$ through angle θ counterclockwise.

Hint: Write $\mathbf{v} = \langle x, y \rangle$ as a combination of the coordinate vectors $\mathbf{e}_1$, $\mathbf{e}_2$; it is much easier to see what happens to these under rotation.

Problem 34. Sketch the following parametric curves (here a, b, λ denote some given constants):

1. $x = a\cos t,\ y = b\sin t$, $0 < a < b$. (Hint: use a linear transformation.)
2. $x = e^{\lambda t}\cos t + e^{\lambda t}\sin t$, where $\lambda < 0$.
3. $x = 2\cos t{+}\sin t,\ \ y = \cos t{+}\sin t$. (Hint: apply a linear transformation to the circle.)
4. $x = e^{\lambda t}(2\cos t + \sin t) + e^{\lambda t}(\cos t + \sin t)$, $\lambda < 0$.
5. $x = \cos t,\ y = \cos 2t$.
6. $x = \cos t,\ y = \cos \sqrt{2}t$.
7. How will the curves in the above examples be affected if we replace $\sin t,\ \cos t$ with $\sin 100t,\ \cos 100t$?
8. How does the size and sign of λ affect the curves in examples 2 and 4?

Problem 35. All we know about a function $f(x, y)$ of two variables is that $f(1, 2) = 3$ and that its partial derivatives $f_x(1, 2) = 4,\ f_y(1, 2) = -3$. Find the approximate value of $f(1.1, 1.9)$, and give an intuitive explanation based on the definition of the derivative.

Problem 36. x, y are given functions of time, and f is a given function of x, y. Write $\frac{d}{dt}f(x(t), y(t))$ in terms of the derivatives of f, x and y.

Problem 37. A constant wind with velocity $\mathbf{v} = \langle v_1, v_2 \rangle$, same at all points, is blowing in the plane. What is the area of the air crossing the segment given by the vector* $\mathbf{g} = \langle g_1, g_2 \rangle$ per unit of time? (This is called the **flux** of $\mathbf{v}$ through the segment given by g, a special case of flux when $\mathbf{v}$ is constant and the "gate" is straight).

---

*g stands for "gate".

Problem 38. Write down the flux of the vector field $\mathbf{F}(x,y) = \langle P(x,y), Q(x,y)\rangle$ through the curve $\mathbf{r}(t) = \rangle x(t), y(t)\rangle$, $t \in [0,1]$ as a line integral, and also as a definite integral. The same question for the work done by the vector field along this curve.

Problem 39. Define the directional derivative, the gradient, the divergence, and the curl (the latter two only in 2D).

1.15 Matrices viewed geometrically.

Square $n \times n$ matrix A often arise in two settings: (i) linear vector fields, given by $\mathbf{F}(\mathbf{x}) \mapsto A\mathbf{x}$ and (ii) linear transformations $T : \mathbf{x} \mapsto A\mathbf{x}$.

Linear vector fields. The following simple examples are "building blocks" for understanding the picture of any linear vector field in $\mathbb{R}^2$.

1. Velocity field $A\mathbf{x}$ of rigid rotation: $A = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$
2. Velocity field $A\mathbf{x}$ of expansion at the exponential rate α: $A = \operatorname{diag}(\alpha, \alpha)$
3. Hyperbolic flow $A\mathbf{x}$, with $A = \operatorname{diag}(\alpha, -\alpha)$
4. Shear flow $A\mathbf{x}$, with $A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$

Linear transformations. Any linear transformation is a composition of the basic transformations $\mathbf{x} \mapsto A\mathbf{x}$ with the following matrices A.

1. $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, rotation through angle θ.
2. $A = \operatorname{diag}(\lambda, -\lambda)$, hyperbolic rotation.
3. $A = \lambda I = \lambda \operatorname{diag}(1,1)$, dilation by factor λ.
4. $A = \operatorname{diag}(1,0)$, projection onto the first coordinate axis.
5. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, shear along the first coordinate axis.

1.16 The determinant as a volume.

The determinant of a matrix is usually defined algebraically. This definition is equivalent to a geometrical one, namely the following: the determinant of a $n \times n$ matrix is the signed volume of the parallelepiped formed by the n column vectors of A. For the example of $n = 2$ the parallelepiped is the parallelogram, and the sign of the area depends on whether the two column vectors form the right–handed frame or the left–handed one.

Problem 40. Let $\mathrm{S}(\mathbf{a}, \mathbf{b})$ denote the (oriented) area of the parallelogram formed by vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^2$. Show *geometrically* that S is a bilinear anti–commutative function of 2 vectors, taking value 1 on the pair of unit coordinate vectors $\mathbf{e}_1$, $\mathbf{e}_2$. In other words, show that for any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ the following hold:

1. $S(\mathbf{e}_1, \mathbf{e}_2) = 1$, where $\mathbf{e}_1,\ \mathbf{e}_2$ are the unit vectors of the coordinate axes.
2. $S(k\mathbf{a}, \mathbf{b}) = -S(k\mathbf{b}, \mathbf{a})$.
3. $S(k\mathbf{a}, \mathbf{b}) = kS(\mathbf{a}, \mathbf{b})$ for any real k,
4. $S(\mathbf{a} + \mathbf{b}, \mathbf{c}) = S(\mathbf{a}, \mathbf{c}) + S(\mathbf{b}, \mathbf{c})$.

Instead of defining the determinant using the "base times height" idea that works for $\mathbb{R}^2$ and $\mathbb{R}^3$, we simply define the determinant as a function with the characteristic properties of the volume.*

Defnition. The determinant of order n is the real–valued function defined on the set of $n \times n$ matrices, and satisfying the following properties:

1. $\det I = 1$, where I is the identity matrix.
2. $\det\,(a\mathbf{v}_1 + b\mathbf{v}_1', \mathbf{v}_2, \ldots) = a \det\,(a\mathbf{v}_1, \mathbf{v}_2, \ldots) + b \det\,(\mathbf{v}_1', \mathbf{v}_2, \ldots)$,
3. $\det\,(\ldots \mathbf{v}_i \ldots \mathbf{v}_j \ldots) = -\det\,(\ldots \mathbf{v}_j \ldots \mathbf{v}_i \ldots)$.

To see that the a function det with properties $1 - 3$ indeed exists and is unique one simply has to expand each $\mathbf{v}_k$ in the coordinate basis and apply the rules $1 - 3$ to obtain the explicit formula (that is usually is given as the definition).

Problem 41. Consider a time–dependent matrix $A(t)$, and let $\mathbf{v}_k = \mathbf{v}_k(t) \in \mathbb{R}^n$ be the column vectors. Show that

$$\frac{d}{dt}\det A(t) = \det(\dot{\mathbf{v}}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) + \ldots + \det(\mathbf{v}_1, \mathbf{v}_2, \ldots, \dot{\mathbf{v}}_n),$$

*One then must show that such function exists, and that it is unique.

and give a geometrical explanation of the answer in $\mathbb{R}^3$.

Hint: The proof is an almost verbatim copy of the proof of the usual product rule for

$$\frac{d}{dt}\big(f_1(t)f_2(t)\dots f_n(t)\big)$$

1.17 The determinant as the volume stretch

In the previous section we interpreted the determinant of a matrix as the (signed) volume of the parallelepiped built on the column–vectors.

Here is an equivalent interpretation of the determinant: Let Vol(S) denote the n–volume of a set $S \in \mathbb{R}^n$.* Then

$$\det A = \frac{\mathrm{Vol}(A(S))}{\mathrm{Vol}(S)}; \tag{1.38}$$

in other words, det A is the factor by which the transformation A changes the volume.

Proof. First, let S be the unit cube formed by the unit vectors $\mathbf{e}_i$ of the coordinate system. Now $A\mathbf{e}_i$ is the ith column of A, so that the transformed cube $A(S)$ is simply the parallelepiped of column–vectors of A, so that the numerator in (1.38) is just det A, while the denominator $= 1$. This proves the claim if S is a unit cube and hence for any dilated or translated unit cube. Now any set S we consider can be approximated with arbitrary precision by a disjoint union of such cubes, and thus (1.38) holds for such sets S. $\diamondsuit$

Problem 42. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis of eigenvectors of matrix A and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues. Use (1.38) to explain why det $A = \lambda_1 \dots \lambda_n$. Hint: pick the "right" set S.

The volume–stretching interpretation (1.38) of det A gives an extremely simple explanation of the following property of the determinant.

Theorem 7. *For any $n \times n$ matrices A, B one has*

$$\det\,(AB) = \det A \det B. \tag{1.39}$$

Proof. Applying the transformation AB to a set S of unit volume in two steps: first applying B, then applying A, we multiply Vol $(S) = 1$ first

*For the volume to be well defined, S has to be a measurable set; it suffices to think of a set bounded by a smooth or piecewise smooth surface, or even just a cube.

by det B, and then by det A, with the resulting volume = det A det B. Summarizing, we showed that

$$\frac{\text{Vol }((AB)(S))}{\text{Vol } S} = \det A \det B.$$

But the left–hand side is simply det AB according to (1.38). $\diamondsuit$

1.18 Eigenvalues and eigenvectors – two geometrical interpretations.

Recall that a nonzero vector $\mathbf{v}$ is said to be an *eigenvector* of a matrix A, and a number λ an eigenvalue of A if

$$A\mathbf{v} = \lambda\mathbf{v}. \tag{1.40}$$

The components of $\mathbf{v}$, as well as λ, may be complex numbers. The geometrical meaning of this will be explained separately; here we concentrate on the real case.

1. Matrix A as a transformation $\mathbf{x} \mapsto A\mathbf{x}$

In this case the line through the origin defined by $\mathbf{v}$ is invariant under the transformation A, and the vectors on this line are stretched by factor λ.

2. Matrix A defines the vector field $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$

In this case the line of $\mathbf{v}$ is invariant under the flow, i.e. the particles on the line stay on the line. If a particle's instantaneous position is $k\mathbf{v}$, then its velocity is $A(k\mathbf{v}) = \lambda(k\mathbf{v}) = \lambda k\mathbf{v}$. In other words, velocity = λ times position, which means exponential growth at the rate λ if $\lambda > 0$ or exponential decay if $\lambda < 0$. The particle stays on the line of $\mathbf{v}$, escaping to infinity or approaching the origin.

1.19 Symmetric matrices.

A square matrix $A = (a_{ij})$ is said to be *symmetric* if

$$a_{ij} = a_{ji}. \tag{1.41}$$

Here are a few condensed insights into the meaning of the symmetry of a matrix. A is symmetric if an only if any of the following hold.

1. The *force field* $A\mathbf{x}$ is conservative.

2. The *velocity field* $A\mathbf{x}$ has zero curl for $n = 3$ and zero 2D curl for $n = 2$.

3. A has an orthogonal basis of (real) eigenvectors.

Here is one more interpretation of (1.41):

1. The elegant but also a bit antiseptic definition of a symmetric $n \times n$ real matrix A as the one satisfying the identity

$$(A\mathbf{x}, \mathbf{y}) - (\mathbf{x}, A\mathbf{y}) = 0 \quad \text{for all} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \tag{1.42}$$

has a physical interpretation: this identity is equivalent to saying that the work done by the linear force field $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$ around the parallelogram generated by $\mathbf{x}$ and $\mathbf{y}$ vanishes.

Figure 2 illustrates the proof of this equivalence: the average forces on each of the sides of the parallelogram are equal to the forces $\mathbf{F}_i$ at the midpoints M_i; the total work W around the parallelogram, grouping parallel sides together, is

$$W = (\mathbf{F}_1 - \mathbf{F}_3, \mathbf{x}) + (\mathbf{F}_2 - \mathbf{F}_4, \mathbf{y});$$

and since $\mathbf{F}_1 - \mathbf{F}_3 = -A\mathbf{y}$ and $\mathbf{F}_2 - \mathbf{F}_4 = A\mathbf{x}$, this gives

$$W = (A\mathbf{x}, \mathbf{y}) - (\mathbf{x}, A\mathbf{y}).$$

In particular, (1.42) expresses the conservativeness of the vector field $A\mathbf{x}$.

2. Here is a physical reason why eigenvalues of a symmetric matrix are real. Assuming for a moment that they are not, consider the plane spanned by the real and the imaginary parts $\mathbf{u}$, $\mathbf{v}$ of the eigenvector $\mathbf{w} = \mathbf{u} + i\mathbf{v}$. At each point $\mathbf{x}$ in this plane the force $A\mathbf{x}$ lies in the plane (so that we can forget about the rest of $\mathbb{R}^n$). And since the work done by $A\mathbf{x}$ around a circle in this plane centered around the origin is zero, the tangential component of $A\mathbf{x}$ changes sign at some point(s) $\mathbf{x}_0$ on the circle – which is to say, $A\mathbf{x}_0$ is normal to the circle at $\mathbf{x}_0$, i.e. $\mathbf{x}_0$ is a (real) eigenvector.

3. Orthogonality of eigendirections of a real symmetric matrix can be seen by a "physical/geometrical" argument, where one can "feel" every step, not hidden by algebra. Let $\mathbf{u}$, $\mathbf{v}$ be two distinct eigenvectors of a symmetric $n \times n$ matrix A with the eigenvalues $\lambda \neq \mu = 0$ (the latter assumption involves no loss of generality since we can take $\mu = 0$ by replacing A with $A - \mu I$). Figure 3 shows the force field $A\mathbf{x}$ of such a matrix. Consider the

work of $A\mathbf{x}$ around the triangle OQP. The only contribution comes from PO since $A\mathbf{x}$ vanishes along OQ and is normal to QP. And if $A\mathbf{x}$ is conservative, then $W_{PO} = 0$ and hence $P = O$, implying $\mathbf{u} \perp \mathbf{v}$. This completes a "physical" proof of orthogonality of the eigenvectors of symmetric matrices.
4. The entry a_{ij}, $i \neq j$ of a square matrix $A = (a_{ij})$ has a dynamical interpretation: it is the angular velocity, in the (ij)–plane, of $\mathbf{e}_i$ moving with the vector field $A\mathbf{x}$.* Indeed, $a_{ij} = (A\mathbf{e}_i, \mathbf{e}_j)$, the projection of the velocity $A\mathbf{e}_i$ onto $\mathbf{e}_j$, Figure 4. And thus the symmetry condition $a_{ij} = a_{ji}$ illustrated in Figure 3 amounts also to stating that the 2D curl in every ij–plane vanishes. For 3×3 matrices the symmetry is equivalent to $\mathrm{curl} A\mathbf{x} = \mathbf{0}$. In fact, decomposition of a general square matrix into its symmetric and antisymmetric parts amounts to decomposing the vector field $A\mathbf{x}$ into the sum of a curl free and of a divergence free fields, a special case of Helmholtz's theorem, itself a special case of the Hodge decomposition theorem.

And the diagonal entries a_{ii} give the rate of elongation of $\mathbf{e}_i$; this explains geometrically why the cube formed at $t = 0$ by $\mathbf{e}_i$ and carried by the velocity field $A\mathbf{x}$ changes its volume at the rate $\mathrm{tr}\, A$ (at $t = 0$). This also gives a geometrical explanation of the matrix identity $\det e^A = e^{\mathrm{tr}\, A}$.

1.20 Geometrical meaning of complex eigenvalues and eigenvectors

1.21 Complex numbers

Definition. A *complex number* $z = x + iy$ is the point (x, y) in the plane, with the additional understanding that the points can be added by the parallelogram rule and multiplied by the following geometrical rule: in multiplying two complex numbers, the lengths multiply†, and the angles add.

This multiplication rule could have been discovered by observing that the familiar rules $1 \cdot 1 = 1$, $(-1) \cdot 1 = 1 \cdot (-1) = -1$, $(-1)(-1) = 1$ can be explained by one unifying principle of angle–addition. For the example $(-1)(-1) = 1$, -1 forms the angle π with the positive x–axis; the angle addition rule gives $\pi + \pi = 2\pi$, so that the product is aligned with the positive x–axis.

Incidentally, the angle–addition rule can also lead to the "discovery" of $i = (0, 1)$: if the product $i \cdot i = -1$ forms angle π with the positive x–axis, then i must

*to be more precise, we should be referring to the moving vector instantaneously aligned with $\mathbf{e}_i$.

†The length $|z|$, or the *absolute value* of z, is the distance to the origin, i.e. $|z| \stackrel{def}{=} \sqrt{x^2 + y^2}$.

form the angle $\pi/2$. Usually, the product $z_1 z_2$ of complex numbers is defined by the formula (amounting to the assumption $i^2 = -1$ and the distributive and commutative properties):

$$z_1 z_2 = x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1), \tag{1.43}$$

and the angle–addition form is proven afterwards. Here we went the other way, giving geometry the upper hand.

Problem 43. Prove (1.43) using the geometrical definition of multiplication and addition.

1.22 Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ – an intuitive derivation.

It is natural to define e^{it} as the solution of the IVP (initial value problem)

$$\dot{z} = iz, \quad z(0) = 1, \tag{1.44}$$

according to which *the velocity of the point z is normal to z and of the same magnitude,* Figure. It follows that

1. $\dot{z} \perp z$, so that z *stays on a circle*;
2. since $z(0) = 1$, that circle has radius 1;
3. the speed $|\dot{z}| = |z| = 1$, so the arc length from $(1, 0)$ to $z(t)$ equals t.

This description of $z(t)$ coincides with the *definition* of $(\cos t, \sin t) \equiv \cos t + i \sin t$, and proves Euler's formula. ♢

Remarkably, the complex equation (1.44) is easier to solve than the real one $\dot{x} = x$!

Chapter 2

An overview of ODEs.

Ordinary differential equations describe a striking variety of things, both moving and stationary, including

1. vibration of interconnected mass–spring systems, oscillations of pendulums, of ships and bridges
2. the motion of projectiles, planets, comets, asteroids and artificial satellites
3. the tumbling of gymnasts
4. the motion of spinning tops, of rolling coins, etc.
5. shapes of hanging chains, cables, of sagging beams.
6. dynamics of chemical reactions and of biological processes
7. population dynamics
8. spread of infectious diseases
9. change of currents and voltages in electric circuits
10. dynamics of airplanes
11. motion of charged particles in electomagnetic fields
12. climate models, etc, etc.

In fact, the entire Newtonian mechanics is really a branch of the theory of ODEs, since Newton's second law is an ODE. And since intuition, either geometrical or physical or both, plays a huge role in studying Newtonian mechanics, it is indispensable in many areas of ODEs.

Given this huge variety of applications, it is remarkable that *all* ODEs are equivalent to an object of one single kind: a vector field, although often in space of dimension other than 2 or 3, as will be explained shortly. And solving an ODE amounts, as we shall see, to finding paths of particles carried by a vector field.

The above motivates our nearest plan, which is to (i) practice formulating some "real life" problems as ODEs (without this skill we would be learning about the hammer without ever hitting the nails) – this is relegated to the Problems section; (ii) explaining how any ODE reduces to studying paths of particles in vector fields, and (iii) giving a bird's eye view of main classes of ODEs.

2.1 Definition and reduction to vector fields

Definition 1. *An ordinary differential equation for the unknown function $x = x(t)$ of the independent variable t is the relationship*

$$F(t, x, \dot{x}, \ddot{x}, \dots x^{(n)}) = 0, \tag{2.1}$$

where F is a given function of $n+1$ variables. One says that (2.1) is an ODE of order n, according to the order of the highest derivative.

Since the order is the most important attribute of the ODE (we will see why shortly), it is customary to express the highest derivative in terms of the rest:

$$x^{(n)} = f(t, x, \dots x^{(n-1)}). \tag{2.2}$$

As a side remark, for $x^{(n)}$ to be thus expressible, it suffices to require that F be "sensitive" to its last argument:

$$F(t, x_0, \dots x_0^{(n)}) = 0, \text{ and } \partial_{n+1} F(t, x_0, \dots x_0^{(n)}) \neq 0, \tag{2.3}$$

the condition of the implicit function theorem. Here ∂_{n+1} stands for the partial derivative with respect to the last argument of F.

Reducing any ODE to a vector field

The nth order ODE (2.2) can be rewritten as a first order ODE $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, but in higher dimension, $\mathbf{x} \in \mathbb{R}^n$, as follows. Introduce $x_1 = x$, $x_2 = \dot{x}_1$, $x_3 = \dot{x}_2, \ldots, x_n = \dot{x}_{n-1}$; in other words, we treat higher derivatives as new variables, interpreting them geometrically, as new coordinates in a higher dimensional space. Now $\dot{x}_n = x_1^{(n)} \overset{(2.2)}{=} f(t, x_1, \ldots, x_{n-1})$. In summary, Eq. becomes

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \ldots \\ \dot{x}_n = f(t, x_1,\ x_2, \ldots x_{n-1}), \end{cases} \tag{2.4}$$

or

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}),, \tag{2.5}$$

where $\mathbf{x} = (x_1, \ldots x_n)$, $\mathbf{f}(t, \mathbf{x}) = (x_2, \ldots, x_{n-2}, f(t, x_2, \ldots, x_{n-1}))$.

Autonomous vs. non–autonomous ODEs

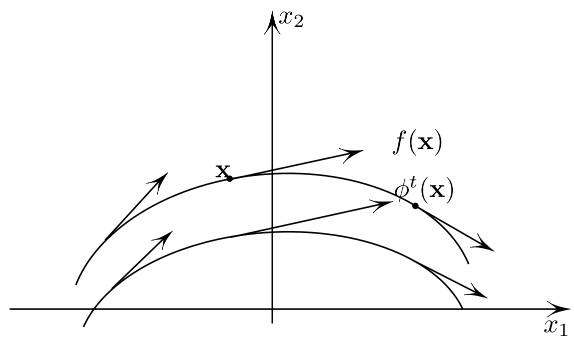
Definition. The ODE $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ with the vector field $\mathbf{f}$ independent of time is referred to as *autonomous.*

The term "autonomous" suggests the absence of external influence upon the system. For example, the ODE describing the exponential grows $\dot{x} = ax$ is autonomous, but if the "interest rate" a depends on time: $a = a(t)$, then the equation is non–autonomous.

The non-autonomous ODE Eq. (2.5) is equivalent to an autonomous ODE in dimension $n + 1$. Indeed, introduce the new dimension $x_0 = t$; then $\dot{x}_0 = 1$, combined with $\dot{\mathbf{x}} = \mathbf{f}(x_0, \mathbf{x})$ can be written as a single ODE for the vector $\mathbf{X} = (x_0, \mathbf{x}) \in \mathbb{R}^{n+1}$:

$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}),\ \text{where } \mathbf{F} = (1, \mathbf{f}). \tag{2.6}$$

We rewrote Eq. (2.5) as an autonomous system at the cost or raising the number of dependent variables to $n + 1$. Note that the velocity field in the x_0–direction has velocity 1.



The imaginary gas interpretation of ODEs. As mentioned before, the vector field $\mathbf{f}$ can be interpreted as velocity field of an imagined gas in $\mathbb{R}^n$; a gas particle carried by this flow then satisfies our ODE $\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x})$. To solve the ODE means to determine the motion $\mathbf{x}(t)$ (as a function of time and of initial position). It is remarkable that this gas interpretation works for any ODE, whatever it's origin: an RLC circuit, a pendulum, an ecosystem with competing species, a planetary motion, etc.

Definitions.

1. A *solution* of the differential equation is any function $\mathbf{x} = \mathbf{x}(t)$ which satisfies the equation.

2. The space $\{\mathbf{x}\}$ is called the *phase space* of the system.

3. The space $\{(t, \mathbf{x})\}$ is called an *extended phase space.*

4. The curve $\mathbf{x}(t)$ (where $\mathbf{x}$ is a solution of the ODE) is called the *trajectory.* The trajectory is thus a projection of the solution curve $\langle t, \mathbf{x}(t)\rangle$ from the extended phase space onto the phase space; Figure shows an example.

Problem 44. Sketch the vector field of the ODE $\dot{x} = x(x-1)(x-2)$ in the extended phase space and sketch the trajectories in the extended phase space.

Problem 45. Write the ODE $\ddot{x} = x$ as a first order system, and sketch trajectories in $\mathbb{R}^2$ and the solution curves in the extended phase space $\mathbb{R}^3$.

2.2 Time–shift invariance in autonomous systems.

Autonomous systems $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ have the following characteristic property: if $\mathbf{x}(t)$ is a solution, then so is its time shift $\mathbf{x}(t-c)$ for any constant $c \in \mathbb{R}$. Indeed, the vector field $\langle 1, \mathbf{f} \rangle$ is invariant under t–translations, and thus if a curve is tangent to this field then so is the c–translation of the curve. In other words, *solutions inherit the invariance of the system.* This invariance under time translation fits common sense: if the wind $\mathbf{f}(\mathbf{x})$ is steady, then the travel history of a particle does not depend on the starting time – something that is patently false if the wind changes with time. A more formal proof of the invariance under time translation will be given once we introduce an important notation.

2.3 The flow of an autonomous ODE

A fundamental object of ODE is a solution to the initial value problem

$$\begin{aligned} &\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \\ &\mathbf{x}(t_0) = \mathbf{x}_0, \end{aligned} \tag{2.7}$$

also called the *Cauchy problem.* The notation $\mathbf{x}(t)$ for the solution is inadequate since it hides the initial data. To fix this drawback, one denotes the solution of the IVP (2.7) by $\boldsymbol{\phi}(t; t_0, \mathbf{x}_0)$, showing both the starting time and the starting position; to summarize,

$$\begin{aligned} &\tfrac{\partial}{\partial t}\boldsymbol{\phi}(t; t_0, \mathbf{x}_0) = \mathbf{f}(t, \boldsymbol{\phi}(t; t_0, \mathbf{x}_0) \\ &\boldsymbol{\phi}(t_0; t_0, \mathbf{x}_0) = \mathbf{x}_0. \end{aligned} \tag{2.8}$$

The *theorem on existence and uniqueness* states that $\boldsymbol{\phi}$ is a well defined function (for a certain range of arguments), under some mild assumptions on the vector field $\mathbf{f}$.

In the autonomous case, which is the only case we consider through the end of this section, the solution operator $\boldsymbol{\phi}$ has the following property:

$$\boldsymbol{\phi}(t; t_0, \mathbf{x}_0) = \boldsymbol{\phi}(t - t_0; 0, \mathbf{x}_0); \tag{2.9}$$

in other words, only the starting point $\mathbf{x}_0$ and the *duration* of travel, but not the starting time, affects the destination. To prove (2.9), we observe that (i) both sides satisfy the same ODE: the left by the definition, the right because a time–shifted solution is still a solution (according to the preceding section), and (ii) the two sides coincide for $t = t_0$; by uniqueness of solutions, they coincide for all t.

According to (2.9) no generality is lost by choosing $t_0 = 0$. This suggests a shorter notation

$$\boldsymbol{\phi}^t \mathbf{x}_0 \stackrel{def}{=} \boldsymbol{\phi}(t; 0, \mathbf{x}_0). \tag{2.10}$$

In short, $\boldsymbol{\phi}^t \mathbf{x}_0$ denotes the solution of the IVP $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0$. $\boldsymbol{\phi}^t$ is referred to as the t–advance map, or as the *flow* associated with the ODE $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$.

So far we may have thought of fixed $\mathbf{x}_0$ of varying t. But for each fixed t, $\boldsymbol{\phi}^t : \mathbb{R}^n \mapsto \mathbb{R}^n$ is an operator. We thus have a one parameter family of time–advance operators. These operators form a group, according to the following theorem.

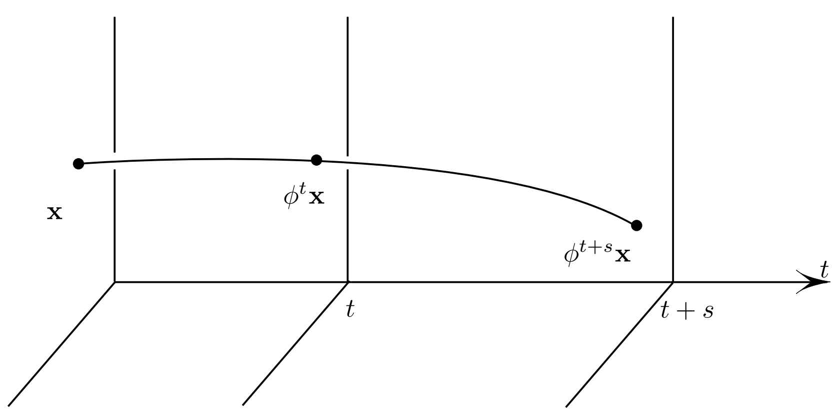
Theorem 8. *Assume that the autonomous vectorfield* $\mathbf{f}$ *on* $\mathbb{R}^n$ *is such that the flow* ϕ^t *is well defined for all* $t \in \mathbb{R}$ *and for all* $\mathbf{x}_0 \in \mathbb{R}^n$. *Then the one-parameter family of maps* $\boldsymbol{\phi}^t$ *forms a group under composition:*

$$\boldsymbol{\phi}^t \circ \boldsymbol{\phi}^s = \boldsymbol{\phi}^{t+s} \quad \textit{for all} \quad t, s \in \mathbb{R}, \tag{2.11}$$

and

$$\boldsymbol{\phi}^0 = id. \tag{2.12}$$

Proof. It suffices to prove that $\mathbf{x}(t) = \boldsymbol{\phi}^t(\boldsymbol{\phi}^s \mathbf{x}_0)$ and $\mathbf{y}(t) = \boldsymbol{\phi}^{t+s} \mathbf{x}_0$ are equal for all t and all $\mathbf{x}_0$. Now $\mathbf{x}(t)$ is a solution of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ by the definition of $\boldsymbol{\phi}$, while $\mathbf{y}(t)$ is the solution of the same ODE because it is a time–shift of the solution $\boldsymbol{\phi}^t \mathbf{x}_0$. Moreover, these solutions coincide for $t = 0$ (with $\boldsymbol{\phi}^s \mathbf{x}_0$), and thus for all t by the uniqueness of solutions. $\diamondsuit$



2.4 More on applications of ODEs.

A major reason why the ODEs are so common is that Newton's second law $\mathbf{F} = m\mathbf{a}$, which governs classical mechanics, usually gives rise to an ODE. Hence the entire subject of classical mechanics is a branch of the theory of ordinary differential equations. Here are a few examples.

The Kepler problem

Kepler's problem deals with the motion of a planet in the gravitational field of the Sun. Effects of other celestial bodies are ignored, as are the tidal effect, solar wind, relativistic effects, etc. The position vector $\mathbf{r}$ of the planet with the Sun at the origin* satisfies a vector ODE

$$\ddot{\mathbf{r}} = -k\,\frac{\mathbf{r}}{r^3}, \quad r = |\mathbf{r}|, \tag{2.13}$$

where k is a constant.

Problem 46. Derive (2.13) from two basic principles: (i) Newton's second law and (ii) Newton's law of gravitational attraction, according to which two point masses attract with the force inverse proportional to their distance from each other.

Newton showed that all trajectories of this differential equation are conic sections (ellipses, parabolas or hyperbolas, depending on the initial data), and thereby explained Kepler's laws from two very simple principles, the inverse square gravitational law and the law $F = ma$, both rolled into one differential equation (2.13).† This was probably the greatest success of science since the antiquity up to that time.

Later developments

Since Newton's time, the theory of ODEs underwent great developments at the hands of some of the greatest mathematicians, including Laplace, Legendre, Hamilton, Jacobi, Lyapunov, Poincaré, and more recently Kolmogorov, Arnold and Moser. In the late 1800s Poincaré discovered "chaos" (although the term "chaos" came into use almost 100 years later).

*Or, one can choose the common center of mass as the origin–the equation is affected only by a constant factor which can be eliminated by a change of units.

†Actually, Newton showed the converse: if the motion satisfies Kepler's laws then the attraction force varies as inverse square of the distance. But Newton's argument is reversible.

Most developments in the theory of ODEs were stimulated by applications – first by celestial mechanics, and later by the problems of space exploration, by electric circuits, and, more recently, by problems in biology, climate and population dynamics. New questions were asked and new phenomena were discovered. The term of *dynamical systems* – another name for an ODE – came into increasing use starting in the late 1930s.

The theory of ODEs uses methods from many different areas of mathematics, including analysis, differential geometry, topology, number theory, functional analysis, algebra.

2.5 A birds' eye view.

A first ODE courses usually emphasize explicit solutions, and usually do not emphasize that such solutions are overrated, for the reasons including:

1. Fragility: the tiniest change of the equation may destroy the formula: $\dot{x} = x$ is easy to solve, but a tiny change: $\dot{x} = x + .00001(\sin x + \sin t)$ is no longer solvable by a formula.

2. An explicit answer, even if possible, may be useless, since it may be hard to recover what we really want to know: the "qualitative" behavior: for instance, what happens to different solutions as $t \to \infty$?

3. Explicitly solvable equations are extremely rare, even among the relatively simple ODEs which can be understood by other means.

The volume of literature on ODEs may give an illusion that we understand a lot. In reality only a sliver of a huge unexplored universe of ODEs is understood qualitatively. And, as mentioned before, only a tiny subset of these are solvable analytically. And since the majority of systems is not discussed-simply because there is nothing to say about them - we get a somewhat distorted view of reality

In the retrospect it is not surprising at all that most equations are not solvable explicitly: it would be presumptuous to think that the small collection of elementary functions is sufficient to model the enormous complexity of the physical world. Even a simple–looking pendulum equation exhibits chaotic behavior, as next described, and thus has no hope of being explicitly solvable.

2.6 Chaos and the lack of explicit solutions.

Consider a pendulum consisting of a mass on a weightless rod pivoting on a hinge, and subject to a sinusoidally varying torque at the hinge. With proper scaling, the angle x of the pendulum with the vertical satisfies

$$\ddot{x} + \sin x = \sin t. \tag{2.14}$$

This equation has no solutions expressible by any elementary functions. This may seem surprising until one tries to solve it. What is much more surprising is that this simple-looking equation has "chaotic" solutions: It can be proven that given **any infinite** sequence of positive integers

$$\mathbf{m} = (m_1, m_2, \dots m_k, \dots)$$

there exists an initial condition $x(0) = x_0, \dot{x}(0) = v_0$ depending on $\mathbf{m}$ such that the pendulum will tumble m_1 times clockwise, then (after hesitating near the top equilibrium) it will tumble m_2 times counterclockwise, etc, ad infinitim – with no interference apart from our choice of the initial conditions!

Note that I can prescribe m_{n+1} independently of m_n – I could make the choices by tossing a coin! – for example, $m_n = +1$ or -1 depending on whether I have heads or tails on the nth toss. All this despite uniqueness theorem (which says that the initial data determine the solution uniquely).

With such chaotic behavior, there little hope for formula solving Eq. (2.14). No formula behaves in such a chaotic way! But this lack of formula should not depress us. Formulas are sometimes overrated; in fact, a formula may look messier and be harder to understand than the original equation!*

2.7 The Cauchy Problem and the phase flow.

The *Cauchy problem*, or synonymously, the *inital value problem*, asks for the solution of

where $t_0 \in \mathbb{R}$ and $\mathbf{x}_0 \in \mathbb{R}^n$ are prescribed initial time and initial position. If $\mathbf{f}$ is sufficiently "nice" then the Cauchy problem has a unique solution which depends nicely on the initial condition $\mathbf{x}_0$. Later we will prove this nice dependence, but for now we assume it. Namely, we assume that a

*This is why some calculus texts are so thick: instead of giving a short idea and the thought, they attempt to give a collection of recipes. And recipes tend to be longer than the simple ideas they are based on.

solution exists for all $\mathbf{x}_0$, and that it is defined for all* $t \in \mathbb{R}$. Moreover, we assume that the solution is unique, i.e. that any two solutions of Eq. (2.7) which share an initial condition actually coincide. Finally, we assume the solution depends smoothly on $\mathbf{x}_0$. We denote the solution of Eq. (2.7) by $\phi^t \mathbf{x}_0$; according to our assumption, the map $\phi^t : \mathbb{R}^n \mapsto \mathbb{R}^n$ is smooth for any t. By the definition, $\phi^t \mathbf{x}$ is differentiable in t as well.

Definition 2. *The family of maps $\phi^t : \mathbb{R}^n \mapsto \mathbb{R}^n$ parametrized by t is called the* **phase flow**. *The map ϕ^t is referred to as a* t**-advance map**.

2.8 Limitations of the theory.

Solvable equations are exceptional. About the simplest example of the differential equation is $\dot{x} = t$, which is solved by simple antidifferentiation: $x(t) = \frac{t^2}{2} + c$, where c is an arbitrary constant. A slightly more complicated example is $\dot{x} = 2x$, with exponential solutions $x = cd^{2t}$.

The vast majority of equations do not admit an explicit solution. This is to be expected; it would be naive to expect that a small supply of functions we learn in school can describe the complexity observed in life. Even a simple-looking equation

$$\ddot{x} + \sin x = \sin t, \tag{2.15}$$

is not solvable by any known functions. This equation describes the angle x of the frictionless pendulum – a unit mass on a stick of unit length, in gravitational field $g = 1$, subject to a sinusoidally varying torque.

In fact, this simple equation exhibits "chaos"! – given any infinite sequence of integers $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ (take the digits of π, for example), there exists an initial condition such that the pendulum will execute m_1 tumbles clockwise, then (after hesitating near the top equilibrium) it will tumble m_2 turns clockwise, etc., ad infinitum. This happens with no external interference, once the initial conditions were imparted! In other words, by the careful choice of initial data we can control future sequence $\mathbf{m}$. The sequence $\mathbf{m}$ is encoded in the initial data, and as the time goes on, further and further "digits" in the initial data "come to the surface" in the form of integers m_k. We term the motion chaotic because the consecutive integers m_k can be prescribed completely independently of one another. Of course, the solution is deterministic: initial data determine the solution uniquely.

---

*See problem 54 below.

Why there is no hope for an explicit solution. With such chaotic behavior there is no hope for a formula for the general solution of Eq. (2.15): elementary functions or anything we can build out of them (by a finite number of steps) don't behave so chaotically.

2.9 Problems

Problem 47. Find the mistake in the following "solution" of the differential equation $\dot{x} = x$: integrating, we obtain $x = x^2/2 + c$. Solving this algebraic equation, we get x.

Problem 48. A mass of algae in a jar grows at the rate which is proportional to both the volume of the algael mass and to the volume that is algae-free. Write down the equation governing the time-evolution of the volume of algae.

Problem 49. Consider the one-dimesional ODE $\dot{x} = x(x-1)(x-2)$.

1. Sketch the vector field on $\mathbb{R}$ corresponding to of this ODE, so that the pattern is clear.
2. Consider the solution with $x(0) = \frac{1}{2}$. Find $\lim_{t\to\infty} x(t)$ without solving the equation explicitly.

Problem 50. Consider the ODE $\dot{x} = f(x)$, $x \in \mathbb{R}$. Let $a < b$ be two isolated zeros of $f(x)$, with $f(x) > 0$ for all $a < x < b$. Prove that for any solution with $x(0) \in (a, b)$ we have $\lim_{t\to\infty} x(t) = b$, $\lim_{t\to-\infty} x(t) = a$.

Problem 51. Write the following equations as first-order systems and sketch the corresponding vectorfields.

1. $\ddot{x} + x = 0$.
2. $\ddot{x} - x = 0$.
3. $\ddot{x} = 0$.
4. $\ddot{x} = -x^{-2}$.

Solution. The systems corresponding to the equations are, respectively:

1. $\dot{x} = y, \ \dot{y} = -x$
2. $\dot{x} = y, \ \dot{y} = x$
3. $\dot{x} = y, \ \dot{y} = 0$

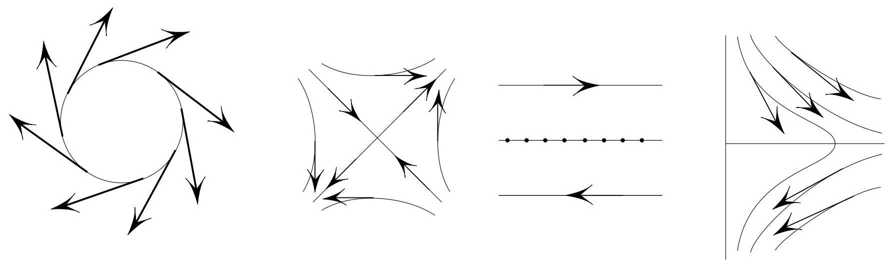


Figure 2.1: For Problem 51

4. $\dot{x} = y, \ \dot{y} = -x^{-2}$

The vectorfields corresponding to these systems are sketched in figure 2.1.

Problem 52. The temperature inside the house changes at the rate proportional to the difference in temperatures between the outside and the inside. The outside temperature is changing sinusoidally with time: $T_o(t) = \sin t$. Write the differential equation for the temperature T inside the house, and sketch the vectorfield in the extended phase space.

Problem 53. Find the flows ϕ^t for the following ODEs: $\dot{x} = x$, $\dot{x} = x^2$, $\dot{x} = x(1-x)$. Is φ^t defined for all t?

Problem 54.

1. Show that any solution of $\dot{x} = 1 + x^2$ blows up in finite time.

2. Show that some solutions of the logistic equation $\dot{x} = x(1-x)$ blow up in finite time when followed backwards in time.

3. Show that all nontrivial solutions of the second-order equation $\ddot{x} = x^2$ blow up in finite time. Physically, a particle subject to a quadratic repelling force escapes to infinity in finite time*.

4. Show that solutions cannot reach infinitiy in finite time if $|\mathbf{f}(\mathbf{x})| \leq C|\mathbf{x}|$ for all $\mathbf{x} \in \mathbb{R}^n$. Here C is a constant independent of $\mathbf{x}$.

   Hint: note that $\frac{d}{dt}\left(\frac{\dot{x}^2}{2} - \frac{x^3}{3}\right) = \text{const.}$ for any solution.

Solution. Blow-up of solutions of $\dot{x} = 1 + x^2$. Two methods:

*This ODE ignores relativistic effects and is, of course, non-realistic.

1. the equation is equivalent to $\frac{\dot{x}}{1+x^2} = 1$, i.e. $\frac{d}{dt}\tan^{-1} x = 1$, i.e. $\tan^{-1} x - \tan^{-1} x_0 = t$. Thus $x = \tan(t + c)$ ($c = \tan^{-1} x_0$.) This function is defined on the maximal interval of length π and blows up to $\pm\infty$ as t approaches the ends of the interval.

2. The solution of $\dot{y} = y^2$ with $y(t_0) = 1$ blows up in finite time (for any t_0, as we showed earlier (by explicitly writing $x(t)$). But since $\dot{x} = x^2 + 1 \geq 1$, for any solution $x(t)$ we have $x(t_0) > 1$ for some t_0. By comparison theorem, $x(t) \geq y(t)$ for all $t > t_0$ for as long as the two solutions exist. Since y escapes to infinity in finite time, so does x, and no later than y.

Problem 55. Derive the differential equation for the shape of the hanging chain. This curve is referred to as the *catenary.* Hint: the net force acting on an infinitesimal segment of the chain is zero, both in the horizontal and in the vertical directions.

Solution. Let $T(x)$ be the tension of the chain, see figure. The sum of all forces on any arc of the chain is zero in the projection on the x-direction:

$$T(x + \Delta x)\cos\theta(x + \Delta x) = T(x)\cos\theta(x) \tag{2.16}$$

and in the y -direction:

$$T(x + \Delta x)\sin\theta(x + \Delta x) - T(x)\sin\theta(x) = \rho g ds, \tag{2.17}$$

where ρ the linear density of the chain and $ds = \sqrt{1 + y'^2}dx$ is the length of the small arc. Dividing the second equation by $T(x)\cos\theta(x) = T(x + \Delta x)\cos\theta(x + \Delta x) \equiv T_0$, we obtain

$$\frac{\tan\theta(x + \Delta x) - \tan\theta(x)}{dx} = \frac{\rho g}{T_0}\frac{ds}{dx},$$

or, since $\tan\theta(x) = y'$:

$$y'' = k\sqrt{1 + y'^2}, \; k = \frac{\rho g}{T_0}. \tag{2.18}$$

Notes:

1. T_0 is the tension of the chain at the point where the chain bottoms out, and that this is equal to the horizontal component of the tension holding the ends of the chain.

2. By rescaling x and y it suffices to consider the case of $k = 1$.

3. The hanging chain has the shape of the hyperbolic cosine.

Problem 56. Consider the pendulum - a point mass on a weightless rod of length ℓ attached to a pivot point by a frictionless hinge. Derive the equation governing the time-evolution of the angle θ between the rod and the vertical.

Answer. $\ddot{\theta} = -\frac{g}{L}\sin\theta$ comes from Newton's second law: $ma = F$ projected on the direction tangent to the circle. Note that $L\theta$ is the distance along the circle to the equilibrium point, so that the acceleration in the tangential direction is $a = \frac{d^2}{dt^2}(L\theta)$. The sum of all forces acting on the mass is $-mg\sin\theta$ (one must draw a figure to see all this). Substituting the above expressions for a and F gives the equation above. Solving this equation is a different matter which will be addressed later.

Problem 57. A concrete column with a statue of weight W on top is tapered so that the weight per unit area of a horizontal cross–section is the same for all cross–sections (this way all parts of the column are equally stressed). Find the dependence of the radius on height.

Problem 58. A ping-pong ball is subject to the air resistance force linearly proportional to the ball's airspeed. A ball is dropped from height H with zero initial speed. Derive the equation governing the time-evolution of the ball's height off the ground during its fall.

Problem 59. * An object is tossed directly upwards from the ground level. Which is greater: the time of ascent or the time of descent? What if the initial velocity is not vertical? Assume that the air resistance is an increasing function of the speed.

Problem 60. A toy boat's "wave engine" extracts energy from water waves. The force produced by the "wave engine" is linearly proportional to the speed of the waves relative to the boat. The drag on the boat is linearly proportional to the square of the speed of the boat relative to the water. Write the differential equation for the speed of the boat. The speed w of the waves relative to the water is given.

Solution. $m\ddot{x} = a(\dot{x} - w) - b|\dot{x}|\dot{x}$.

Problem 61. A bullet is fired from the surface of the Moon vertically upwards. Write down the differential equation governing the distance of the bullet to the Moon's surface. The gravitational acceleration on Moon's surface is g_m, and the radius of the Moon is R.

Solution. $\ddot{x} = -\frac{k}{(x+R)^2}$. To find k note that at the surface ($x = 0$) we have $\frac{k}{(R)^2} = g_m$, so that $k = g_m R^2$, and the ODE becomes $\ddot{x} = -g_m \frac{R^2}{(x+R)^2}$.

Problem 62. If a plane flying straight up with speed of Max 3 (roughly $1\,km/sec$) suddenly turned off its engine, and if the air resistance vanished, how high would it get? Take $g = 10m/sec^2$.

Problem 63. A capacitor* starts out with charge q_0 is then shorted by a resistor R. Try to guess the formula for the charge $q(t)$ on the capacitor at time t. Then write the differential equation for $q(t)$ and compare with the guess.

Solution. The guessing process: it feels like the charge should be decaying exponentially, i.e. $q(t) = ce^{-at}$; what are c and a? First, $q(0) = q_0$ gives $c = q(0)$. Now a large R makes the decay of $q(t)$ slow, so R should be in the denominator of a. Similarly, a large capacitance C makes for a slower discharge as well, and so C also belongs in the denominator, so the guess is $c = \frac{1}{RC}$.

Here is a rigorous solution: $V_C + V_R = 0$ (Kirkhoff's First Law, see Chapter **??** for all the necessary circuits background). Substituting $V_C = q/C$ (definition of capacitance) and $V_R = IR$ (Ohm's law), where $I = \dot{q}$ (definition of the current) we get $q/C + \dot{q}R = 0$, or

$$\dot{q} = -kq,\ k = \frac{1}{RC}. \tag{2.19}$$

We conclude that the charge on the capacitor $q(t) = q_0 e^{-\frac{t}{CR}}$ decays exponentially. This agrees with the guess.

Problem 64. Derive the differential equation for the time-evolution of the current in an RLC circuit.

Problem 65. A bank account grows with the instantaneous speed proportional to the amount present, doubling after a year. (a) What happens to the amount after half a year? (b) Prove that the length of time it takes for the money to (say) triple does not depend on the starting date.

Problem 66. An island has a population of rabbits and foxes. From one day to the next, the population of rabbits increases by the amount proportional to the number of rabbits present and decreases by the amount proportional to the number of foxes present. The population of foxes increases

*For the necessary background on electric circuits see Chapter **??**

in proportion of the number present, and also increases in proportion to the population of rabbits. Write the system of differential equations approximating the evolution of the two populations with time, treating these populations as continuous functions of time.

Problem 67. Consider the IVP

$$\dot{y} = a(t)y + b(t), \quad y(0) = 0, \tag{2.20}$$

where the function a is such that the solution of the IVP $\dot{x} = a(t)x$, $x(0) = 1$ satisfies $0 \leq x \leq e^{-t}$ for all $t \geq 0$, and where b satisfies $0 \leq b(t) \leq e^{-t}$. *True or false:* the solution of (2.20) is bounded for all $t \geq 0$.

Problem 68. Two identical glasses A and B of water are at different temperatures: $0°C$ and $100°C$ Figure **??**. Cold water is clean; the hot water is dirty. We want to heat the clean water using dirty water without mixing them. To that end, we pump cold water from A through a thin tube into into the third glass C. Part of the tube is submerged into the hot and dirty glass B, so that the water picks up some heat from B and emerges from the tube at the same temperature as B. The dirty glass B is gradually cooling, as cold water passing through the pipe picks up some of the heat. Find the eventual temperature of the clean water after all all of it ends up in the third glass (and mixed, so that the water temperature is the same throughout the glass. Assume that no heat is lost to the surrounding medium, that the vessels are perfectly insulating, that the water in each glass has the same temperature throughout, and also assume that the pipe conducts the heat perfectly.

Problem 69. A tick jumps on a passing person. Assume that he/she jumps the distance $1m$, landing with speed of $.1m/sec$. What is his takeoff speed, given that his free fall speed is $.5m/sec$. Neglect the gravity, i.e. assume that the tick moves horizontally.

Problem 70. I am twirling a stone on a rope in a circle. Both ends of the rope go in concentric circles of radii $r < R$. The rope forms angle θ with the radius–vector of the fingertips holding the rope. Explain why such concentric motion implies accelerating spin. Write the ODE for the speed v of the stone. Find $v(1)$ given that $v(0) = 1$, $\theta = \pi/4$, and $R = 1$. (all quantities are measured in the units of the same system.

Problem 71. A bicycle is guided by hand so that its front wheel goes straight along the curb with speed v. The frame of the bike forms angle θ

with the direction of the curb. Write the differential equation satisfied by θ. In this problem the bike is a segment RF, with R and F representing the points of contact of the wheels with the ground, with $|RF| = \text{const.}$, and with the velocity of R pointing at F at all times.

2.10 English-to-Math Translation Problems

The following problems exercise your translation skills. No physics besides what is in our freshman/sophomore calculus book is required (just Newton's second law and torque).

Problem 72. Let $x(t)$ denote the amount of money (in dollars, say) in an account at time t. The amount grows with the speed (measured in dollars per year) equal to 5% of the amount present. Express this as an ODE for $x(t)$.

Problem 73. A bullet is shot into water vertically down. Let $x(t)$ be the penetration distance at time t. The bullet is subject to two forces: water resistance, directly proportional to the square of the velocity, and gravity. Express this as an ODE for $x(t)$.

Problem 74. Let $x(t)$ be the angle formed by the string of the pendulum with the downward vertical direction. The angular acceleration of the pendulum is in direct proportion to the torque of the gravity relative to the pivot.* Express this as an ODE for $x(t)$.

Problem 75. Let $x(t)$ be the position at time t of a bug crawling on the x–axis. The bug's velocity is in direct proportion to his distance to the point $x = 1$. Express this as an ODE for $x(t)$.

Problem 76. Let $x(t)$ denote the position at time t of a point mass m on the x–axis. The mass is subject to the force directly proportional to the distance of the mass to the point $x = 2$, and pointing towards that point. Express this as an ODE for $x(t)$.

Problem 77. A point mass m in the (x, y)–plane is subject to a constant force pulling it directly down in the direction of the negative y–axis. Express this as a pair of ODEs for x and for y.

*Recall the definition of torque from Math 230 or 231: it is the "intensity of turning", given by the product of the lever and the component of the force perpendicular to the lever.

Problem 78. A point mass m in the (x, y)–plane is subject to the force pulling it directly into the origin, and of magnitude proportional to the distance to the origin. Express this as an ODE for the vector $\mathbf{r}(t) = \langle x(t), y(t)\rangle$, or equivalently (i.e. if you prefer) as a pair of ODEs for $x(t)$, $y(t)$.

Problem 79. A point mass m in the (x, y)–plane is subject to the gravitational force of the star located at the origin. Express this as an ODE for the vector $\mathbf{r}(t) = \langle x(t), y(t)\rangle$, or equivalently (i.e. if you prefer) as a pair of ODEs for $x(t)$, $y(t)$.

Problem 80. Let $x(t)$ be the temperature of a cup of coffee, and let T_0 be the room temperature. The temperature of coffee approaches that of the room at the rate directly proportional to the mismatch between the two temperatures. Express this as an ODE for $x(t)$.

Problem 81. A full pot with hot water is standing in the sink. Water is pouring into the pot at the rate of a gallons per minute and spills over the edge (at the same rate a, of course) after being thoroughly mixed. Denoting by $x(t)$ the temperature of the water in the pot at time t, write the ODE for $x(t)$.

Problem 82. A rock, thrown upwards, is subject to two forces: gravity and air resistance, assumed to be directly proportional to the speed. Write the ODE for $x(t)$, the height of the rock at time t.

Problem 83.

1. Consider a smooth hill with the given elevation function $h(x, y)$, where x and y are the latitude and longitude of the point above which the elevation is measured. A hiker is descending the hill, always choosing the steepest direction down, and so that his speed *projected onto the*(x, y)*–plane* equals the slope at his location. Write the above as the differential equation for the coordinates $(x(t), y(t))$ of the hiker.

2. The same question as above, with the only difference that now he follows the level line, and travels in the direction pointing to the right of the downhill direction.

The following problem explains, in a remarkably transparent way, Euler's formula $e^{it} = \cos t + i \sin t$.

Problem 84. A bug travels in the (x, y)–plane, keeping his velocity perpendicular to his position vector at all times and his speed equal to his distance to the origin. Write the ODE for the bug's position vector, using the complex notation, i.e. by writing $z(t) = (x(t), y(t)) = x(t) + iy(t)$ to denote the bug's position.*

Problem 85. A projectile in the vertical (x, y)—plane is subject to two forces: gravitational, and the air drag, the latter directly proportional to the speed. Write the ODE for the in the projectile's coordinates. Gravity points down along the y–axis.

Problem 86. Translate the following sentences into formulas.

1. Vector $\mathbf{a}$ is parallel to vector $\mathbf{b}$ and is twice as long.
2. Vector $\mathbf{c}$ is a linear combination of the vectors $\mathbf{a}$ and $\mathbf{b}$ (and draw a sketch).
3. Vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are linearly dependent.

*Please keep in mind that there is nothing imaginary about $x + iy$; it is simply another way to write (x, y).

Chapter 3

First Order Systems

Some highlights:

1. Any solution of $\dot{x} = f(x)$ with $f(x_0) \neq 0$ is unique if f is merely continuous - no Lipschitz assumption is necessary.

2. Some paradoxes: Newton's law is not (always) deterministic: a particle in *some* force fields may start moving spontaneously at any time of your choice, without violating Newton's law!

3.

3.1 Classification

The most general first order ODE has the form

$$\dot{x} = f(t, x). \tag{3.1}$$

Even such a simple–looking form allows too much variety to be captured by elementary functions. Therefore we list important special cases and leave the vast sea of all other possibilities unexplored.*

Here the list of the special types of (**??**) we will discuss:

Here they are:

1. *Autonomous:* $\dot{x} = f(x)$.

*Of course, any *individual* ODE can be solved numerically with some degree of approximation – but we are looking for something meaningful to say about a class of ODEs, not about a single one.

2. *Linear:* f is linear in x: $\dot{x} = a(t)x + b(t)$.*

3. *Separable:* f separates into a product, each depending only on one variable: $\dot{x} = a(t)b(x)$.

4. *Homogeneous:* f is a homogeneous function, i.e. it depends on the ratio of its variables only: $\dot{x} = f(\frac{x}{t})$.

5. *Riccati's equation:* $\dot{x} = a(t) + b(t)x + c(t)x^2$.

A note on autonomous systems. Autonomous systems $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ in $\mathbb{R}^n$ can be classified by the dimension n. The scalar case $n = 1$ is solvable analytically (the solution boils down to integration). For $n = 2$ analytic solution is impossible except in special cases, but still, the *qualitative* behavior is well understood. And the case $n \geq 3$ is so rich that only special cases have been analyzed, and it is safe to say that this case will never be fully understood.†

In the next four sections we deal with each of the above types.

3.2 Linear ODEs

Linear first order ODEs are of the form

$$\dot{x} = a(t)x + b(t); \tag{3.2}$$

recall that *linear* refers to the x–dependence in f; the t–dependence can be nonlinear, with no restrictions on the functions a and b (apart from the properties needed as we go).

Lagrange's method of variation of constant. To solve this ODE, we first solve the homogeneous case

$$\dot{y} = a(t)y$$

by separation of variables. Rewriting this as

$$\frac{dy}{y} = a(t)\,dt$$

*nonlinearity in t is allowed; what really matters is the linearity in x.

†This reminds of Stanislaw Lem's parody [**?**] where he classifies the types of genius: the first type is "run of the mill", recognized during his lifetime; the second type of genius is too far ahead of other and is recognized only posthumously; and the pinnacle type #3 is so far advanced that he is never recognized.

we integrate, obtaining $\ln y = \int a(t)\,dt + c$, or

$$y = Ce^{A(t)}, \quad \text{where} \quad A(t) = \int a(t)\,dt.$$

Lagrange had many great ideas, one of which was to allow C to depend on t in the hope that then $C(t)e^{A(t)}$ may satisfy (3.2). To find the right choice of $C(t)$ he substituted the guess into (3.2) and see what we must ask of $C(t)$ (from now on we write $C(t) = C, a(t) = a$, etc.:

$$\frac{d}{dt}(Ce^A) = aCe^A + b, \quad \text{or} \quad \dot{C}e^A + \cancel{Cae^A} = \cancel{aCe^A} + b,$$

using $\dot{A} = a$. The ODE reduces to $\dot{C}e^A = b$, and this dictates the choice of $C = \int e^{-A}b\,dt + c$. This choice guarantees that Ce^A is a solution; we found a solution of (3.2):

$$x = e^A\left(c + \int e^{-A}b\,dt\right).$$

To show that this solution will satisfy any initial condition $x(0) = x_0$ for any x_0, we replace the antiderivatives with the properly chosen definite integrals, setting

$$A(t) = \int_0^t a(s)\,ds.$$

Choosing $c = x_0$, we obtain the solution to the IVP:

$$x = x(0)e^{A(t)} + e^{A(t)}\int_0^t e^{-A(\tau)}b(\tau)\,d\tau. \tag{3.3}$$

A heuristic derivation of (3.3). It helps to think of x as the amount of money in an account which grows at the variable rate $a(t)$ and which is subject to continuous deposit at the rate $b(t)$. First, for the homogeneous equation, $e^{A(t)}$ is the factor by which the account grew from time 0 to t. Now for the inhomogeneous equation, the amount $x(t)$ consists of the part that grew from the initial deposit, namely $x(0)e^{A(t)}$, plus the part that is due to continuous deposit.

To calculate the continuous deposit part we must convert the deposit made at each time τ to the present t. We do this by first converting everything to $t = 0$, and then to t. During time $[\tau, \tau + d\tau]$ the amount $b(\tau)\,d\tau$ was deposited; in the dollars of $t = 0$ this amount is $e^{-A(\tau)}b(\tau)\,d\tau$. And from $t = 0$ to present it grows to $e^{A(t)}e^{-A(\tau)}b(\tau)\,d\tau$. Finally, we must add up effect of deposits for each of the intervals $[\tau, \tau + d\tau]$. This explains (3.3).

3.3 Separable ODEs

Separable ODEs

$$\dot{x} = a(t)b(x) \tag{3.4}$$

includes two even more special cases: $\dot{x} = a(t)$ and $\dot{x} = b(x)$. The name of this class suggests also the solution: we separate the variables to the opposite sides of the equation, writing it as

$$\frac{dx}{b(x)} = a(t)\ dt \tag{3.5}$$

and integrate, getting $\int \frac{\dot{x}}{b(x)} = \int a(t)\ dt$ – the desired solution, implicit in x and containing an arbitrary constant which can be adjusted to satisfy the initial condition, if given.

Now the meaning of (3.5) is a bit vague, since we did not clarify what is meant by, say, $a(t)\,dt$. Here is a more honest way. We separate the variables in (3.4):

$$\frac{\dot{x}}{b(x)} = a(t) \tag{3.6}$$

and note that the left–hand side is the t–derivative of $B(x) = \int \frac{dx}{b(x)}$, e.g. of

$$B(x) = \int_0^x \frac{ds}{b(s)}.$$

Now (3.6) turns into

$$\frac{d}{dt}B(x(t)) = a(t),$$

and we integrate, using FTC on the left:

$$B(x(t)) - B(x(0)) = \int_0^t a(\tau)\,d\tau. \tag{3.7}$$

Example. Solve $\dot{x} = \frac{1}{3}x^{-2}$ with the initial condition $x(0) = 1$.
Solution. Rewrite the equation as $3x^2\dot{x} = 1$ (this is what one calls separation of variables), and notice that the left-hand side is the time-derivative of x^3. Integrating both sides from $t = 0$ to t we get

$$\int_0^t 3x^2(\tau)\dot{x}(\tau)d\tau = \int_0^t d\tau,$$

or $x^3(t) - x^3(0) = t$. Substituting the initial condition $x(0) = 1$, we obtain $x^3 = t + 1$, and solving for x we obtain the solution explicitly:

$$x(t) = (t+1)^{\frac{1}{3}}.$$

This is the solution, given implicitly.

Problem 87. Solve the IVP $\dot{x} = x^2 t$, $x(0) = 1$.

Problem 88. Under what conditions on b does (3.7) determine $x(t)$ uniquely?

Problem 89. 1. Show that the solution of the IVP $\dot{x} = f(x)$, $x(0) = x_0$ exists (for t in a certain interval) and is unique provided that f is merely continuous and $f(x) \neq 0$ for all x.

2. Show that without the assumption $f(x) \neq 0$ the solution may not be unique.

3.4 Homogeneous ODEs

The ODE reducible to the form

$$\dot{x} = f\left(\frac{x}{t}\right) \tag{3.8}$$

is called *homogeneous*. Homogeneity may not be immediately obvious, as in the example

$$\dot{x} = \frac{at + bx}{ct + dx}.$$

Geometrically, homogeneity amounts to the following statement: the slope of the direction field along each ray $x = kt$ is fixed; in other words, each such ray is an *isocline* (unfortunately no one uses the term *isoslope* instead of isocline).

Geometrically, homogeneity amounts to the invariance of the direction field under dilations $(t, x) \mapsto (kt, kx)$; under such dilations x/t doesn't change, which suggests using $s = x/t$ as the new unknown. To rewrite (3.8) with the new unknown s we note that $\dot{x} = \frac{d}{dt}(st) = \dot{s}t + s$, and (3.8) becomes

$$\dot{s}t + s = f(s),$$

which is separable:

$$\frac{\dot{s}}{f(s) - s} = \frac{1}{t},$$

which has been solved in the preceding section.

3.5 Riccati's equation

Riccati's equation is quadratic in the unknown function:

$$\dot{x} = a(t) + b(t)x + c(t)x^2. \tag{3.9}$$

Perhaps the main reason this equation is of interest is that it describes the time–evolution of the slope of solution vectors of linear ODEs in $\mathbb{R}^2$. In other words, for any linear (possibly time–dependent) vector field $A(t)\mathbf{x}$ in the plane, the slope of the radius–vector of each moving particle satisfies reef (**??**), see Theorem.

This equation cannot in general be solved in elementary functions. The reason I bring this equation up is that it describes linear systems in $\mathbb{R}^2$; this is described in the following theorem.

Theorem 9. *Consider a linear ODE $\dot{\mathbf{x}} = A(t)\mathbf{x}$ in $\mathbb{R}^2$:*

$$\begin{cases} \dot{x}_1 = a_{11}x_1 + a_{12}x_2 \\ \dot{x}_2 = a_{21}x_1 + a_{22}x_2 \end{cases} \tag{3.10}$$

*For any solution (x_1, x_2), the ratio $x = x_2/x_1$ satisfies the Riccati equation (**??**) with*

$$a = a_{21}, \;\; s\; b = a_{22} - a_{11}, \quad c = -a_{12}.$$

Proof amounts to differentiating $x = \frac{x_2}{x_2}$, replacing $\dot{x}_1$, $\dot{x}_2$ with their expressions from (3.10) and grouping the powers of $x = x_2/x_1$. ♢

Problem 90.

1. What linear systems does the ODE $\dot{x} = x^2$ correspond to?
2. Can you explain geometrically the blow–up of solutions of this ODE in finite time, and to find the blow–up time geometrically?
3. Can you solve the IVP $\dot{x} = x^2$, $x(0) = x_0$ by using the corresponding linear system?

3.6 Geometry of first order autonomous ODEs.

The ODE $\dot{x} = f(x)$ can be interpreted as follows: the x–axis is a "highway" with velocity $f(x)$ prescribed at every single location x, Figure. The zero velocity points, i.e. the zeros of f, organize the behavior of the system. If $a < b$ are two neighboring zeros of f then f has a fixed sign on the interval

Figure 3.1: Interpreting the ODE $\dot{x} = f(x)$ as a vector field on the line.

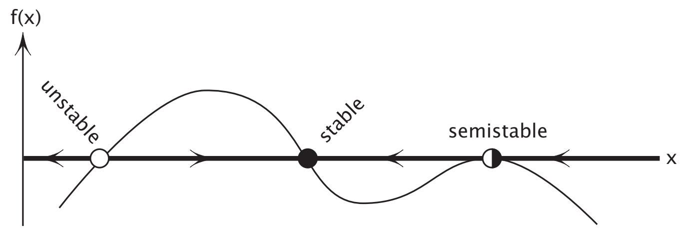


Figure 3.2: Qualitative picture of the flow on the line.

(a, b), and thus every "car" on (a, b) moves either to the right or to the left. Figure **??** shows the flow on a line; the graph of $f(x)$ is plotted in the same figure; the points move right along the x–axis where $f(x) > 0$.

Figure 3.3

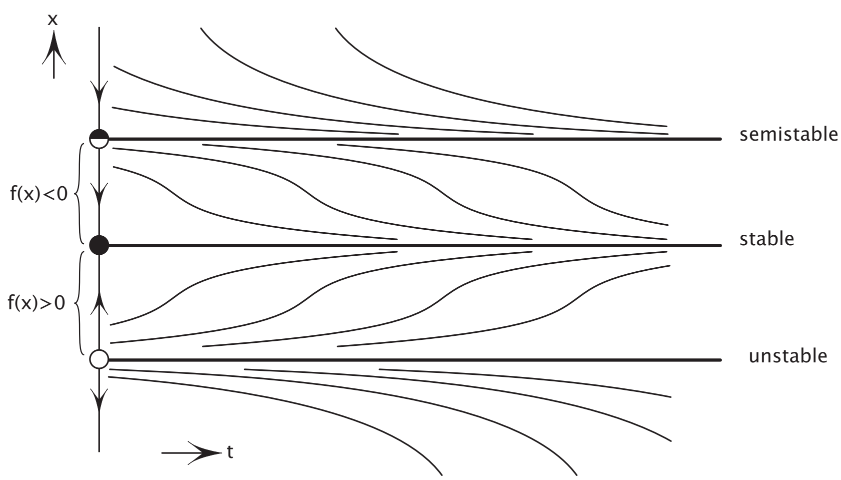


Figure 3.3: Solution curves in the (t, x)–plane.

Definition 3. *1. A constant solution $x(t) = a = const.$ of an ODE is called an equilibrium, an equilibrium solution, or a rest point.*

2. An equilibrium $x = a$ is said to be stable if all solutions starting near it approach it as $t \to \infty$, or formally, if there exists $\delta > 0$ such that $\varphi^t x \to a$ as $t \to \infty$ for all $x \in (a - \delta, a + \delta)$.

The following theorem captures the intuitively obvious fact that constant solutions have zero velocity, and that zero velocity points are equilibrium solutions.

Theorem 10. *If $x(t) = a$ is an equilibrium solution of $\dot{x} = f(x)$, then $f(a) = 0$, and vice versa.*

Proof. Assuming $x(t) = a$ is a solution, i.e. that $\dot{a} = f(a)$ implies $f(a) = 0$. Conversely, if $f(a) = 0$, then $x(t) = a$ satisfies the ODE: $\dot{a} = f(a)$, since both sides vanish. $\diamondsuit$

It is also intuitively obvious that the constant solution is stable if its neighbor's velocities point towards it, Figure. Here is a precise statement.

Theorem 11. *Assume that $f(x) > 0$ for all $x \in (a, b)$, and that $f(a) = f(b) = 0$. Then for any $x \in (a, b)$ we have $\lim_{t\to\infty} \phi^t x \to b$ and $\lim_{t\to-\infty} \phi^t x \to a$.*

Proof. Fix any $x \in (a, b)$. By the uniqueness theorem, $a < \phi^t x < b$ for all $t \in \mathbb{R}$. But then $f(\phi^t x) = \frac{d}{dt}\phi^t x > 0$ for all t, so that $\phi^t x$ is a monotone increasing function. Since it is also bounded from above (by b), the limit $\lim_{t\to\infty} \phi^t x_0 = b^* \leq b$ exists. We must only show that $b^* = b$, which amounts to proving that $f(b^*) = 0$. Assuming the contrary: $f(b^*) > 0$, by the continuity assumption f is positive also on an interval: for some $\varepsilon > 0$, $\delta > 0$ we have $f(x) \geq \varepsilon > 0$ for $(b^* - \delta, b^* + \delta)$. Hence in time less than $2\delta/\varepsilon$ the solution would pass through the interval, never to return to it because of monotonicity. Thus the solution cannot have b^* as its limit. The contradiction completes the proof. $\diamondsuit$

Corollary 12. *The equilibrium $x = a$ is stable if $f'(a) < 0$ and backwards stable if $f'(a) > 0$.*

In the higher dimensional system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ the above stability condition of an equilibrium is replaced by the condition that the matrix $\mathbf{f}'(\mathbf{a})$ has eigenvalues in the left half plane.

3.7 Comparison Theorems for $\dot{x} = f(t, x)$

Sometimes the only way to get information on an otherwise intractable ODE $\dot{x} = \mathbf{f}(t, x)$ is to compare it to a simpler one. The following theorem is the tool for such comparison.

Theorem 13. *Consider two ODEs $\dot{x} = f(t, x)$ and $\dot{y} = g(t, y)$, with $g(t, x) \geq f(t, x)$ for all t, x and with the initial conditions $y(0) \geq x(0)$. Assume that the solutions depend continuously on initial conditions*. Then $y(t) \geq x(t)$ for all $t \geq 0$ for which both solutions are defined.*

Proof. Consider at first the strict inequality case: $g(t, x) > f(t, x)$, $y(0) > x(0)$. We claim that

$$y(t) > x(t) \text{ for all } t \geq 0.^{\dagger} \tag{3.11}$$

This is clear from Figure: the graph of y starts above that of x (since $y(0) > x(0)$), and must remain above, since otherwise at the first time $t = t^*$ of crossing we would have had $\dot{y}(t^*) \leq \dot{x}(t^*)$, i.e. $g(t^*, y(t^*) \leq f(t^*, x(t^*))$, where $y(t^*) = x(t^*)$, contradicting the assumption $g(t, y) > f(t, y)$.

It remains to remove the strictness assumptions in (3.11), which we do by a perturbation argument, making the inequality strict and then taking a limit. Let us perturb one ODE: $\dot{y}_\varepsilon = g(t, y_\varepsilon) + \varepsilon$ and its initial condition: $y_\varepsilon(0) = y_0 + \varepsilon$. The inequalities between this perturbed IVP and that for x now strict, and the result of the preceding paragraph applies:

$$y_\varepsilon(t) > x(t) \text{ for all } t \geq 0 \tag{3.12}$$

for all $\varepsilon > 0$. Since $y_\varepsilon(t)$ depends continuously on the parameter ε, $\lim_{\varepsilon \to 0} y_\varepsilon(t) = y(t)$. Taking limit as $\varepsilon \downarrow 0$ in Eq. (3.12) gives $y(t) \geq x(t)$ for $t \geq 0$. ♢

Problem 91. Give a rigorous proof of the fact that if $y(t) > x(t)$ for $t \in [0, t^*)$ and if $y(t^*) = x(t^*)$, then $\dot{y}(t^*) \leq \dot{x}(t^*)$. Must the last inequality be strict?

3.8 Numerical solutions of $\dot{x} = f(t, x)$

I will describe two methods for solving the IVP

$$\begin{aligned} \dot{x} &= f(t, x) \\ x(t_0) &= x_0. \end{aligned} \tag{3.13}$$

The first method (Euler's) is very simple but not very accurate; the second one (Runge–Kutta's) is the other way around: very accurate but not very simple.

*it suffices to require that f and g be continuously differentiable in x and measurable in t.

†We assume that the solutions are defined for all $t \geq 0$; otherwise t must be additionally restricted to the common interval of existence of the two solutions.

Euler's method

Figure explains the method: We compute the slope $f(t_0, x_0)$ at the starting point (t_0, x_0), and follow the straight line with this slope to the intersection with the line $t = t_0 + h$, where h is a small step size we choose, and repeat the process, treating the point of intersection as the new starting point. In other words, we define the iteration process

$$\begin{aligned} t_{n+1} &= t_n + h \\ x_{n+1} &= x_n + f(t_n, x_n)h. \end{aligned}$$

Example. Let us find the approximation to the solution of $\dot{x} = x, \;\; x(0) = 1$ at $t = 1$ using Euler's method. Dividing the unit time interval into N pieces, we get the step size $h = \frac{1}{N}$ (for an integer n). Iteration step is $x_{n+1} = x_n + f(t_n, x_n)h = \big(1 + \frac{1}{N}\big)x_n$. Using $x_0 = 1$ we obtain, after N steps:

$$x_N = \left(1 + \frac{1}{N}\right)^N,$$

the approximation to e, as expected.

Accuracy of Euler's method. The error of Euler's step is, by the definition, the difference between the true solution and its estimate after one step h. Since the a curve deviates from its tangent by amount quadratic in the distance from the tangency point, we expect $O(h^2)$ for the error – this is provided the curvature of the curve is bounded. To prove this expectation amounts to showing that there exists a constant $C > 0$ such that the solution x with the initial condition $x(a) = b$ satisfies

$$|x(a + h) - x(a) - f(a, b)h| \leq Ch^2.$$

By Taylor's formula with the Lagrange remainder, $x(a+h) - x(a) = \dot{x}(a)h + \frac{1}{2}\ddot{x}(a + \hat{h}$, for some $0 < \hat{h} < h$. But $\dot{x}(a) = f(a, x(a))$, and

$$\ddot{x}(t) = \frac{d}{dt} f(t, x(t)) = f_t + f_x\dot{x} = f_t + f_x f.$$

Substituting this into the above,

3.9 Existence, uniqueness and regularity.

What conditions must $f(x)$ satisfy for the initial value problem to have a unique solution? According to the general theory (Chapter **??**) differentiability of f is sufficient; mere continuity is not, as many examples show (see problem **??** below). The one-dimensional case, however, is special: as long as $\frac{1}{2}$ is integrable, the solution is unique! In particular, if $f(x)$ is merely continuous and $f(x) \neq 0$, we have uniqueness. This was proven in section

Consider the Cauchy problem

$$\dot{x} = f(x, \lambda), \ x(0) = x_0, \ x, \ \lambda \in \mathbb{R} \tag{3.14}$$

where the right-hand side is allowed to depend on a parameter. We already showed by explicit method that the solution exists if $f(x) \neq 0$. We observe that the same method shows that the solution depends continuously on x_0 and on the parameter λ.

Theorem 14. *Assume that $f(x, \lambda) \neq 0$ for all $x, \ \lambda$ and that f is continuous in both x and λ. Then Eq. (3.14) has a unique solution which, moreover, depends continuously both on x_0 and λ.*

Remark 1. *Mere continuity of f along with the condition $f(x) \neq 0$ imply existence and uniqueness (see Problem 104). This is in sharp contrast to the case of higher dimension $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$. There are examples of continuous vectorfields $\mathbf{f}(\mathbf{x})$ for which any Cauchy problem has multiple solutions: through each point in the plane there pass infinitely many integral curves! This is quite striking: despite the fact that the direction of at each point is strictly prescribed, one still can, starting from the same point, follow different paths. We give an intuitive explanation of this in Section* **??**. *This explanation boils down to the fact that a quantity which decays at a super-exponential rate can become zero in finite time. The quantity mentioned is the distance between two nearby solutions.*

Remark 2. *The assumption $f(x) \neq 0$ is crucial. Wherever this assumption breaks down, the uniqueness may fail, as the example $\dot{x} = x^{\frac{1}{3}}$, $x(0) = 0$ shows. Problem 103 discusses the case when f vanishes at a point.*

We now prove the theorem.

♢ Following the separations of variables procedure, we rewrite the Cauchy problem Eq. (3.14) in an equivalent form:

$$F(x, \lambda) = t, \tag{3.15}$$

where

$$F(x, \lambda) = \int_{x_0}^{x} \frac{dy}{f(y, \lambda)}$$

Now, F is monotone in x, and therefore (3.15) defines $x = x(t, \lambda, x_0)$ as a function of t, λ, x_0. Furthermore, F is continuous in λ and x_0. By the implicit function theorem, $x(t, \lambda, x_0)$ is a a continous function of λ and x_0 (and t). $\diamondsuit$

From now on we will assume that $f'(x)$ exists for all x.

3.10 Linearizing transformation.

In the preceding section we gave complete analysis of the first order ODE for $f(x) \neq 0$. In this section we study the remaining case: what can be said about the flow near an equilibrium? We show that in a neighborhood of a typical (in a precise sense) equilibrium the equation is equivalent to $\dot{x} = x$ or to $\dot{x} = -x$ (this is the one-dimensional version of the Hartman-Grobman theorem).

Without loss of generality, let the equilibrium be at the origin: $f(0) = 0$. We assume that $f'(0) \neq 0$; this is the precise meaning of "typical". The following important theorem states that $\dot{x} = f(x)$ is equivalent to $\dot{y} = y$ or to $\dot{y} = -y$, depending on the sign of $f'(0)$.

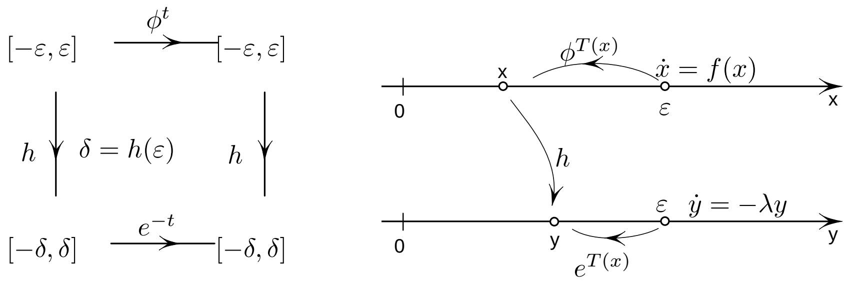
Theorem 15. *If f is continuously differentiable for $|x| < \varepsilon$ for some ε, and if $f(0) = 0$, $f'(0) < 0$, then there exists a continuous one-to-one mapping $h : [-\varepsilon, \varepsilon] \mapsto [-\delta, \delta]$ (for some $\delta > 0$) such that $h(\phi^t x) = e^{-t} h(x)$ for all $t > 0$ and for all $|x| \leq \varepsilon$. In other words, the flow $\phi^t x$ is conjugate to the linear flow $e^{-t}x$ via the conjugation h, i.e. the diagram in Figure* **??** *commutes.*

Think of looking at the x–axis through a nonlinear lens. The lens sends the point x on the line into the location $y = h(x)$ on the "retina". The normal form theorem says that there exists a "lens" h such that the retinal image of the nonlinear flow $\dot{x} = f(x)$ is linear!

$\diamondsuit$ By the assumption of f, there is a neighborhood $N = [-\varepsilon, \varepsilon]$ such that every solution $\phi^t x$ with $x \in N$ approaches 0 monotonically. Consider now two flows: $\dot{x} = f(x)$ and $\dot{y} = -y$, see figure **??** . We construct $h(x)$ as follows: given $0 < |x| < \varepsilon$, let $T = T(x) > 0$ be the time it takes to reach x from the boundary of the neighborhood: $\phi^T(\pm\varepsilon) = x$ (the sign in front of ε is meant to coincide with the sign of x). We then define the h–image of x by going back with the nonlinear flow to the boundary $\varepsilon = \pm\varepsilon$ and then forward for the same time with the linear flow $\dot{y} = -y$:

$$h(x) = e^{-T(x)t} \phi^{-T(x)} x, \ x \neq 0 \tag{3.16}$$

and $h(0) = 0$. With this definition, h is a homeomorphism (one-to-one and continuous), and it satisfies $h \circ \phi^t = e^{-t} h(x)$, as claimed. $\diamondsuit$



$h \circ \phi^t = e^{-t} \circ h$ Construction of h

Figure 3.4: The linearizing map h.

3.11 Bifurcations

3.12 Some paradoxes

The surprising fact regarding even the simplest IVP, such as $\dot{x} = \sqrt{x}$, $x(0) = 0$ is that the solution is not unique, as Figure illustrates. Indeed, $x \equiv 0$ is a solution, and so is $t^2/4$. In fact, any function vanishing on the interval $t \in [0, c]$ and equal $(t - c)^2/4$ for $t \geq c$ is a solution.

I want to make two remarks on this: first, what is the intuitive reason for such a strange behavior? - the answer has to do with exponential growth, and second, I want to point out that this example can be used to show that Newtonian mechanics is not always deterministic!

A heuristic explanation of solutions splitting. What property of f allows two solutions $x_1(t) = 0$ and $x_2(t) = t^2/4$ split apart? In other words, how can the point $[x_1(0), x_2(0)]$, split into an interval at $t > 0$? The length $L(t) = x_2(t) - x_1(t)$ satisfies $\dot{L} = \dot{x}_2 - \dot{x}_1 = f(x_2) - f(x_1)$ (we write $x_1 = x_1(t)$ for brevity), or

$$\dot{L} = a(t)L, \tag{3.17}$$

where $a(t) = f'(\overline{x}(t))$, where $\overline{x}(t)$ point in $[x_1, x_2]$. Now if f' is bounded, L grows no faster than an exponential, and thus cannot become nonzero if $L(0) = 0$. This explains why solutions are unique if f' is bounded. On the other hand, in our example, $a(0) = \infty$ in (3.17), suggesting (as the calculation shows) that L can become nonzero.

The answer will also tell us what will prevent such a split, i.e. what is a condition for uniqueness (the condition is called Osgood's criterion).

3.13 Problems

Problem 92. A particle's position x on the line evolves according to $\dot{x} = f(x)$, where x is a given function. Find the particle's acceleration as a function of its position.

Problem 93. A particle moves according to $\dot{x} = f(x)$, where $f(x) > 0$ for all $x \in \mathbb{R}$. Find the time of travel from a to $b > a$.

Problem 94. Show that no solution of $\dot{x} = f(t, x)$ blows up in finite time if f satisfies $|f(x)| \leq a(t)|x| + b(t)$ for all $t,\ x \in \mathbb{R}$, where $a, b \in C(\mathbb{R})$ (here $C(\mathbb{R})$ denotes the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$).

Hint. Use the comparison theorem.

Problem 95. (See Problem 100) Consider two differential equations $\ddot{x} + q(t)x = 0$ and $\ddot{y} + p(t)y = 0$ with $p(t) > q(t)$.

1. Write each equation as a system of two ODEs, and consider the angle formed with the positive x-axis by the solution vector of each system.

2. Show that these angles satisfy $\dot{\alpha} = f(t, \alpha)$ and $\dot{\beta} = g(t, \beta)$ with $f(t, \alpha) > g(t, \alpha)$. In other words, whenever the vectors $(y, \dot{y})$ and $(x, \dot{x})$ are aligned, one rotates faster than the other.

3. Let $x(t)$ and $y(t)$ be two solutions of the above equations. Show that between any two consecutive zeros of $x(t)$ there is a zero of $y(t)$ (a zero of $x(t)$ is, by the definition, a value of t for which $x(t) = 0$). Hint. Use the result of the preceding question.

Problem 96. Find all equilibria and classify their stability for the differential equation $\dot{x} = \sin x$. Sketch the phase portrait of the above differential equation and sketch enough solutions in the x, t–plane to fully understand the behavior of all solutions. Draw the bifurcation diagram for the above system.

Problem 97. The logistic equation. In this problem we consider a simple model of the propagation of a rumor, or a joke, or the flu in a population. Let $x(t)$ be the proportion of the people who have heard the joke. Assume the population to be so large that we can treat x as a continuous variable. Think of an enormous party where people mingle randomly, with the average person tells the joke to all the people (s)he meets (whether they heard it or not). Write a simple differential equation which would model the evolution of x in time.

Problem 98. Find the mistake in the following "solution" of the differential equation $\dot{x} = x$: integrating, we obtain $x = x^2/2 + c$ and solve this quadratic equation for x.

Problem 99. *Solve the following initial value problems. Also, describe the asymptotic behavior of* **all** *solutions. In other words, specify what is the limit of a solution depending on the initial condition.*

1. $\dot{x} = x,\ x(0) = 1$
2. $\dot{x} = x(1-x),\ x(0) = \frac{1}{2}$
3. $\dot{x} = \sin x,\ x(0) = \pi/2$
4. $\dot{x} = \sin^2 x,\ x(0) = \pi/2$

Problem 100. Consider the system

$$\begin{cases} \dot{x} = y \\ \dot{y} = -\omega^2 x \end{cases}, \tag{3.18}$$

where ω is a constant. Find the first order ODE for the angle $\theta = \tan^{-1}(y/x)$ formed by the solution vector (x, y) with the x-axis. What becomes of this ODE if $\omega = 1$? Note that $\dot{\theta}$ does not depend on the distance of the point (x, y) to the origin, but only on θ. What special property of the ODE is responsible for this fact?

Solution. We have:

$$\frac{d}{dt}\theta = \frac{d}{dt}\tan^{-1}(y/x) = \frac{\dot{y}x - \dot{x}y}{x^2 + y^2}$$

(a small algebraic step was skipped). Substituting $x = r\cos\theta,\ y = r\sin\theta$, where $r = \sqrt{x^2 + y^2}$, $\dot{x} = y$ and $\dot{y} = -\omega^2 x$, we get

$$\dot{\theta} = -\omega^2\cos^2\theta - \sin^2\theta. \tag{3.19}$$

If $\omega = 1$ the ODE for θ reduces to $\dot{\theta} = -1$. This fits with the geometric analysis we did directly the first day of class. What makes the right-hand side in Eq. (3.19) independent of r is the linearity of the given system of ODEs: the time t map is linear, as shown in the chapter on linear systems.

Problem 101. Consider the ODE $\dot{x} = f(x)$ where $-2x \le f(x) \le -x$ for all $x > 0$. Prove that the solution with $x(0) = 1$ satisfies $e^{-2t} \le x(t) \le e^{-t}$. In particular, although the solution decreases, we still have $x(t) > 0$ for all $t > 0$.

Solution. Use the comparison theorem, comparing with $\dot{x} = -x$ and $\dot{x} = -2x$.

In the preceding problem, the solution never becomes zero. The next problem looks at this question closer.

Problem 102. Can the solution of $\dot{x} = f(x)$ with $x(0) = 1$, where f is a continuous function with $f(0) = 0$ and $f(x) < 0$ for $x > 0$, reach zero in finite time?

Hint: If speed $f(x)$ is of the same order of magnitude as x, i.e. the speed towards the origin is on the order of the distance to the origin – think of $\dot{x} = -kx$, then x decreases no faster than some exponential and thus will never become zero. The only hope is to make the speed $f(x)$ be much larger than x.

Solution: Following the hint, take, for example, $f(x) = -x^{1/3}$. (Even very close to the origin – at, say, $x = 1/1000$, the speed is relatively high: $f(x) = 1/10$. While still approaching the origin, we do so with greater speed than if were approaching exponentially.)

Let us compute the time referred to in the problem and see whether it is finite. Let T be the time it takes to get from (say) $x = 1$ to $x = 0$:

$$T = \int \frac{d(\text{displacement})}{\text{speed}} = \int_{x_0}^{0} \frac{dx}{\dot{x}} = \int_{x_0}^{0} \frac{dx}{-x^{1/3}} = \frac{3}{2}$$

– we reach the origin in finite time! We thereby discovered that the solution $x(t) \equiv 0$ is not unique: another solution (the one with $x(0) = 1$) touches it (at time $T = \frac{3}{2}$), see Figure.

The following problem amounts to the well known **Osgood uniqueness theorem**; behind this theorem is a tautologically sounding idea: for two solutions to avoid a meeting the coalescence time must be infinite!

Problem 103 (Osgood uniqueness theorem)**.** Consider the Cauchy problem $\dot{x} = f(x)$, $x(0) = 0$, where $f(0) = 0$, and $f(x) \neq 0$ for $\neq 0$. Prove that if the improper integral $\int_0^x \frac{dy}{f(y)} = \infty$ diverges, then the solution $x \equiv 0$ is unique.

Remark 3. *If $|f'(0)| < \infty$, then the zero solution is unique (indeed, the above integral diverges). Thus finite divergence $f'(0)$ guarantees uniqueness; this makes intuitive sense. One can say, roughly speaking, that that non-uniqueness is caused by infinite divergence. For example, $f(x) = x^{\frac{1}{3}}$ has infinite divergence: $f'(0) = \infty$ and exhibits non-uniqueness.*

The following is a very interesting, but simple fact:

Problem 104. Show that uniqueness implies continuous dependence. In other words, show that if the solution of a Cauchy problem is unique for all initial data, then the solution depends on the initial data continuously.

Chapter 4

Dynamical Systems in the Plane

4.1 Classification, a bird's eye view

4.2 Linear Systems with constant coefficients

4.3 Linearization at an equilibrium point

4.4 The Poincaré index

4.5 Limit cycles

4.6 The Andronov–Hopf bifurcation

4.7 The Poincaré–Bendixson theory

4.8 The Bohl-Brouwer fixed point theorem

4.9 Hamiltonian Systems

4.10 Gradient Systems

4.11 Functions of Complex Variable and Hamiltonian Systems

4.12 Lyapunov's function, Energy

Chapter 5

Second Order Systems $\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$

5.1 Classification, a bird's eye view

5.2 Linear Vibrations

5.3 Lissajous Figures

5.4 Kepler's problem

Chapter 6

Chaos

6.1 A chaotic Pendulum

6.2 The Tent Map

6.3 The Horseshoe Map

6.4 Symbolic Dynamics