Notes for MAT 308, Spring 2025

Joseph Helfer

Last updated: February 10, 2025

Contents

1 Jan 27, Intro and calculus review

1.1 Introductory remarks

- Differential equations are the central formalism in all of the most important physical theories, and are of great importance in many other branches of science and engineering, as well as in pure mathematics.
  - For example: "Newton's Second Law", "Maxwell's Equations" from electromagnetism, "Einstein's Field Equations" from general relativity, "Schrödinger's equation" from quantum mechanics, are all systems of differential equations.
- Roughly speaking a differential equation is an equation containing (ordinary or partial) derivatives.
- Example: Newton's Second Law $F = ma$ is really $F = m\frac{\mathrm{d}^2u}{\mathrm{d}t^2}$, where $u(t)$ is the position of a given object with mass m.
  - F typically depends on u and maybe on $\frac{\mathrm{d}u}{\mathrm{d}t}$.
  - For example, for a massive body in a central gravitation field, we have $F = -G\frac{u}{|u|^3}$ for some constant G, so the equation becomes $-G\frac{u}{|u|^3} = m\frac{\mathrm{d}^2u}{\mathrm{d}t^2}$.
  - The "unknown" in this equation, for which one might try to solve, is u. It appears twice, with different amounts of derivatives. A solution would not be a *number* as in an ordinary equation but a *function* of time: the trajectory of the body in question.
- In Newton's equation, the variable u is actually a *vector*-valued – i.e., $\mathbb{R}^3$-valued – function.
  - Thus, the equation is actually a system of *three* equations.
  - As you learned in linear algebra, a system of linear equations is most fruitfully understood as a single matrix equation $Av = b$.
  - The same kind of thing is true in the theory of differential equations, and moreover there are some interesting and important further ideas in linear algebra which come in.
  - This is why linear algebra is also part of this course.
- Newton's equation only contains *ordinary* derivatives and is thus a so-called **Ordinary Differential Equation** (or **ODE**).
  - By contrast, all of the other equations mentioned above are **Partial Differential Equations** (or *PDE*s).
  - We will be dealing almost exclusively with ODEs in this class.
  - Just to give you an idea, a simple example of a PDE is the *heat equation*, which looks like this: $\frac{\partial u}{\partial t}(t, x) = -\frac{\partial^2 u}{\partial x^2}(t, x)$.

1.2 Calculus review

1.2.1 On engineers and mathematicians

- There are two ways of approaching the subject of differential equations.
  - The "mathematician's way" gives precise definitions of all the objects under consideration, and rigorous proofs on the basis of those definitions.

- The "engineer's (or physicist's) way" rests on a basic intuitive understanding of all of the basic objects (real numbers, continuous functions, limits, etc.) and their properties, and this intuition is based on the physical objects and phenomena that these mathematical objects represent.
- Feynman ("The character of physical law"): *There are two kinds of ways of looking at mathematics, which [...] I [...] call the Babylonian tradition and the Greek tradition. In Babylonian schools in mathematics the student would learn something by doing a large number of examples until he caught on to the general rule. [...] Also he would know a large amount of geometry, a lot of the properties of circles, the theorem of Pythagoras, formulae for the areas of cubes and triangles; in addition, some degree of argument was available to go from one thing to another. [...] But Euclid discovered that there was a way in which all of the theorems of geometry could be ordered from a set of axioms that were particularly simple. The Babylonian attitude – or what I call Babylonian mathematics – is that you know all of the various theorems and many of the connections in between, but you have never fully realized that it could all come up from a bunch of axioms.*

- **Both of these ways are valuable and important**.
  - And in fact, the "ideal mathematician" or "ideal engineer" don't exist; everyone is somewhere in between the two.
  - As this is a *mathematics* course (MAT 308), we must to some extent emphasize the "mathematician's way": in principle, you should know the precise definitions of all the objects we consider, be able to give precise statements of all the theorems we consider, and be able to follow the proofs, and reproduce the simpler ones. (In principle, this also means being familiar with the axioms of set theory – which, though a good thing, we will not insist on.)
  - Thus, we begin with a quick review of the basic elements of set theory and calculus which we will be using, to make sure we are all on the same page.

1.2.2 Sets

- We take for granted all the notions of set theory, and the basic properties of and operations with sets, which you know well by now.
- The most important sets are $\mathbb{N} = \mathbb{Z}_{\geq 0}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ for the naturals, integers, rationals, reals, and complex numbers, to all of which we shall return shortly.
  - We write $\mathbb{N}_+$ or $\mathbb{Z}_{>0}$ or something like that for the set of positive integers.
- The product of two sets is the set of all ordered pairs $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$.
- We write $A \subset B$ or $A \subseteq B$ for "A is a subset of B".
  - If we want to express that A is a *proper* subset of B, we write $A \subsetneq B$
- A <u>relation</u> R between sets X and Y is a subset of the product $R \subset X \times Y$.
  - We write "xRy" for $(x, y) \in R$.
- A <u>function</u> $f\colon X \to Y$ is a relation $f \subset X \times Y$ such that for every $x \in X$, there is a unique $y \in Y$ with xfy., i.e.
$$\forall x \in X, \exists! y \in Y, xfy.$$
  - We write $f(x)$ for the unique $y \in Y$ such that xfy.
- A(n infinite) <u>sequence</u> in a set X is a function $\mathbb{N} \to X$ (or $\mathbb{Z}_{\geq n} \to X$ for any $n \in \mathbb{Z}$).

- As usual, we write $(x_i)_{i=0}^{\infty}$ for the sequence $\mathbb{N} \to X$ with $i \mapsto x_i$ for $i \in \mathbb{N}$.

- A finite sequence of length $n \in \mathbb{N}$ or n-tuple in a set X is a function $\{1, \ldots, n\} \to X$.
  - We write X^n for the set of finite sequences of length n.
  - As usual, we write $v = (v_1, \ldots, v_n)$ for the tuple $\{1, \ldots, n\} \to X$ given by $i \mapsto v_i$.
  - More generally, given sets $X_1, \ldots, X_n$, we write $X_1 \times \cdots \times X_n$ for the set of finite sequences $(x_1, \ldots, x_n)$ with $x_i \in X_i$ for each i.

- We write $X \cong Y$ if there exists a bijection between X and Y, and we write $f \colon X \xrightarrow{\sim} Y$ to indicate that f is a bijection.

- We write Y^X for the set of functions $X \to Y$ and $\mathcal{P}(X)$ for the power set of X, i.e., the set of subsets of X.
  - We have $2^X \cong \mathcal{P}(X)$ for any set X, where 2 is the set $\{0, 1\}$.

1.2.3 The real numbers

- We recall the main properties of the set of real numbers $\mathbb{R}$ with its operations $+ \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\cdot \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and its ordering "$\leq$" $\subset \mathbb{R} \times \mathbb{R}$:

- **Field axioms**
  - addition and multiplication are commutative and associative, and multiplication distributes over addition
  - 0 and 1 are identity elements for addition and multiplication, respectively (i.e., $0 + a = a$ and $1 \cdot a = a$ for all $a \in \mathbb{R}$).
  - (*Exercise*: 0 and 1 are each uniquely determined by this property!)
  - Every element $a \in \mathbb{R}$ has an additive inverse $(-a)$ (i.e., $a + (-a) = 0$).
  - Every element $a \in \mathbb{R} - \{0\}$ has a multiplicative inverse a^{-1}.

- **Ordering axioms**
  - $\leq$ is reflexive, meaning $a \leq a$ for all $a \in \mathbb{R}$.
  - $\leq$ is antisymmetric, meaning $a \leq b \wedge b \leq a \Rightarrow a = b$ for all $a, b \in \mathbb{R}$.
  - $\leq$ is transitive, meaning $a \leq b \wedge b \leq c \Rightarrow a = c$ for all $a, b, c \in \mathbb{R}$.
  - $\leq$ is total, meaning $a \leq b \vee b \leq a$ for all $a, b \in \mathbb{R}$.
  - If $a \leq b$ then $a + c \leq b + c$ for any $a, b, c \in K$.
  - If $a \leq b$ and $c \geq 0$, then $a \cdot c \leq b \cdot c$ for any $a, b, c \in K$.

- **Completeness axiom**
  - Given a set $S \subset \mathbb{R}$, we say that $a \in \mathbb{R}$ is an upper bound for S if $s \leq a$ for all $s \in \mathbb{R}$, and we say that S is bounded-above if it has some upper bound.
  - Completeness axiom: each bounded-above subset $S \subset \mathbb{R}$ has a least upper bound (or supremum) $\sup S$, meaning that $\sup S$ is an upper bound for S, and that $\sup S \leq a$ for every upper bound a for S.

- In short, one says that $(\mathbb{R}, +, \cdot, \leq)$ is a complete ordered field.

- On the basis of these axioms, one can prove all the familiar properties of the algebraic operations and the ordering relation (an activity you have hopefully done before, and should try your hand at if you haven't).
  - In fact, these axioms **completely determine** the real numbers in the following sense: given any system $(K, \tilde{+}, \tilde{\cdot}, \tilde{\leq})$ consisting of a set K, binary operations $\tilde{+}$ and $\tilde{\cdot}$ on K, and a binary relation $\tilde{\leq}$ on K satisfying the above axioms, there exists a unique bijection $F\colon \mathbb{R} \to K$ satisfying $F(a+b) = F(a)\tilde{+}F(b)$, $F(a \cdot b) = F(a)\tilde{\cdot}F(b)$, and $a \leq b \iff F(a)\tilde{\leq}F(b)$ for all $a, b \in K$.
    - (Such a bijection F is called an "isomorphism of ordered fields".)
    - (The aforementioned theorem establishes the *uniqueness* (up to isomorphism) of the real numbers. One may still reasonably inquire about their *existence*: why must there exist such a set satisfying these axioms at all? We simply take it for granted that it exists; if one wants to *construct* it, this must be on the basis of some more basic axioms: either the Peano axioms of arithmetic, or else the axioms of set theory.)
- The other number systems
  - We define the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q} \subset \mathbb{R}$ as follows.
  - $\mathbb{N} \subset \mathbb{R}$ is the smallest subset such that $0 \in \mathbb{N}$ and such that $n+1 \in \mathbb{N}$ whenever $n \in \mathbb{N}$.
    - Concretely, this means that $\mathbb{N}$ is the intersection $\bigcap_{S \subset \mathbb{R}} S$ of all subsets of $\mathbb{R}$ satisfying these two properties.
    - It follows immediately from this that $\mathbb{N}$ satisfies the all-important **principle of mathematical induction**: given any $S \subset \mathbb{N}$ such that $0 \in S$ and $n+1 \in S$ whenever $n \in S$, we have $S = \mathbb{N}$.
  - $\mathbb{Z} = \{a - b \mid a, b \in \mathbb{N}\} \subset \mathbb{R}$.
  - $\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z} \land b \neq 0\} \subset \mathbb{R}$.
  - Finally, $\mathbb{C}$ is simply defined to be $\mathbb{R}^2$.
    - Given $(a_1, b_1), (a_2, b_2) \in \mathbb{C}$, we set $(a_1, b_1) + (a_2, b_2) := (a_1 + b_1, a_2 + b_2)$, and $(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)$.
    - One may check that this makes $\mathbb{C}$ into a *field* (though of course, not an *ordered* field, since there is no natural ordering of points in the plane).
    - Given $a \in \mathbb{R}$, we write a as a shorthand for $(a, 0) \in \mathbb{C}$.
    - We also write i as a shorthand for $(0, 1) \in \mathbb{C}$; we have $i^2 = -1$.
    - Thus, we may write $(a, b) = a + bi$ for every $(a, b) \in \mathbb{C}$, and we then have $(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + b_1) + (a_2 + b_2)i$ and $(a_1 + b_1 i) \cdot (a_2 + b_2 i) = a_1 a_2 + b_1 b_2 i^2 + a_1 b_2 i + a_2 b_1 i$.

2 Jan 29, More calculus review and 10.1: Direction fields

2.1 Calculus review continued

2.1.1 Calculus

- We will be reviewing various topics from calculus (and linear algebra) as they come up; we will just recall a couple of the most important ones here
- We consider functions $f\colon I \to \mathbb{R}$ defined on some domain $I \subset \mathbb{R}$.
  - Typically, $I \subset \mathbb{R}$ will be an interval: a set of the form (a,b), $[a,b)$, $(a,b]$, or $[a,b]$, where $a \in \mathbb{R} \cup \{-\infty\}$ and $b \in \mathbb{R} \cup \{\infty\}$ and $a \leq b$.
- Limits:
  - Given a function $f\colon I \to \mathbb{R}$ and $x \in I$, we have the limit $\lim_{x\to a} f(x)$ and the one-sided limits $\lim_{x\nearrow a} f(x)$ and $\lim_{x\searrow a} f(x)$ (each of which may or may not exist).
  - Ideally, you should know the **definition** of the limit: $\lim_{x\to a} f(x) = b$ means $\forall \epsilon > 0, \exists \delta > 0, \forall x \in I, (0 < |x-a| < \delta \Rightarrow |f(x) - b| < \varepsilon)$.
  - But the most important thing is that you know the basic **rules** for limits: $\lim_{x\to a} f(x) + g(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$ (assuming the right-hand sides exists), and so on.
- A function $f\colon I \to \mathbb{R}$ is continuous at $a \in I$ if $\lim_{x\to a} f(x) = f(a)$, and is continuous on I if it is continuous at each $a \in I$.
  - Continuous functions satisfy the **intermediate value theorem**.
- Derivatives:
  - $f\colon I \to \mathbb{R}$ is differentiable at $a \in I$ if $\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$ exists, and if so, this limit is called $f'(a)$.
  - f is differentiable on I if it is differentiable at each $a \in I$; in this case, we have the derivative $f'\colon I \to \mathbb{R}$.
    - **Theorem:** Differentiability implies continuity.
  - We sometimes write $y = f(x)$ and then denote the derivative by $f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x}$.
    - Here, we are thinking of x and y as "variables", with y varying "dependently" on x; then $\mathrm{d}x$ represents an "infinitesimal variation" of x, $\mathrm{d}y$ represents the resulting infinitesimal variation of y, and the derivative is their quotient.
    - This is the classical way of thinking about functions, which has been preserved by the physicists and engineers, but among mathematicians has been replaced by the set-theoretic perspective.
    - With some care, this notation can be used in a consistent and rigorous way (in fact, doing so leads to some very interesting mathematics), but in general, it is best to follow the physicists and engineers, and freely use our intuition about infinitesimals without worrying about the formal definition; and then if we want a rigorous proof, we can always revert to the formal, set-theoretic definitions.
    - In any case, we often simply write $\frac{\mathrm{d}f}{\mathrm{d}x}$ for f'.
  - We write f'' or $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$ for the derivative of f' (if it exists), and f''' and $f^{(4)}$ and $f^{(5)}$ and so on.
  - Again the most important thing is to know all the **rules** for derivatives: the sum rule, product rule, quotient rule, **chain rule**, the rule for x^n, and the derivatives of everyone's favourite functions: sin, cos, e^x, $\ln x$, and so on.

- We denote by $\mathcal{C}^k(I)$ the set of functions $f\colon I \to \mathbb{R}$ such that the derivatives $f', f'', \ldots, f^{(k)}$ exist and are continuous; $\mathcal{C}^0(I)$ is simply the set of continuous functions on I, and $\mathcal{C}^\infty(I)$ is the set of infinitely differentiable (or smooth) functions on I. All of the most common functions are smooth.

- Partial derivatives
  - We consider functions $f\colon U \to \mathbb{R}$ defined on a domain $U \subset \mathbb{R}^n$.
  - Often, U will be all of $\mathbb{R}^n$, or maybe some product of intervals $U = I_1 \times \cdots \times I_n$.
  - The definition of **limit** and **continuity** for such functions is exactly the same as in the single-variable case,
    * except that now, the absolute values $|x - a|$ and so on appearing in the definition are now considered norms.
    * We recall the norm of $v \in \mathbb{R}^n$ is given as in the Pythagorean theorem: $|v| = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$.
  - Given $f\colon U \to \mathbb{R}$ and $a \in U$, the i-th partial derivative of f at a, if it exists, is the "derivative of f at a in the i-th coordinate direction, holding all other coordinates constant".
  - For example, if $n = 2$, the two partial derivatives at $(a, b) \in U$ are

$$\lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} \qquad \text{and} \qquad \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

  - In general, the i-th partial derivative at $a \in U$ is

$$\lim_{h\to 0} \frac{f(a_1,\ldots,a_{i-1},a_i+h,a_{i+1},\ldots,a_n) - f(a_1,\ldots,a_n)}{h}.$$

  - There are many notations for the i-th partial derivative; the simplest are $\partial_i f$ or f_i; often we write $y = f(x_1, \ldots, x_n)$, and then write $\frac{\partial y}{\partial x_i}$ or $\frac{\partial f}{\partial x_i}$ or $\partial_{x_i} f$ or f_{x_i}.
  - In practice, one evaluates the partial derivative with respect to x_i by "pretending all the other variables are constant and taking the ordinary derivative with respect to x_i".

- The chain rule
  - The sum, product, quotient rules for partial derivatives are just like those for ordinary derivatives.
  - The **chain rule** is more interesting: if $y_i = f_i(x_1, \ldots, x_m)$ for $i = 1, \ldots, n$ and $z = g(y_1, \ldots, y_n)$, so that

$$z = g\big(f_1(x_1,\ldots,x_m),\ldots,f_n(x_1,\ldots,x_m)\big),$$

    then

$$\frac{\partial z}{\partial x_i} = \sum_{j=1}^n \frac{\partial z}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_i}.$$

  - The chain rule is most elegantly expressed using the **Jacobian** (or **total derivative**) matrix. For this, we must recall a bit about vector-valued functions and matrix multiplication, and we will do this when it comes up.

- Sequences and series
  - The limit $\lim_{n\to\infty} x_n$ of a sequence $(x_n)_{n=0}^\infty$, if it exists, is defined as the unique $a \in \mathbb{R}$ such that $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - a| < \varepsilon$.
    * We also write $x_n \xrightarrow{n\to\infty} a$.

- Again, the most important thing is that you know the main **rules** for limits: the sum rule $\lim_{n\to\infty} a_n + b_n = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$ (assuming the right-hand side exists), the difference and quotient rules, and the rule that if f is a function which is continuous at a and $x_n \xrightarrow{n\to\infty} a$, then $f(x_n) \xrightarrow{n\to\infty} f(a)$.
- The sum of the infinite series $\sum_{n=0}^{\infty} a_n$, if it exists, is the limit $\lim_{N\to\infty} S_n$ of the partial sums $S_N = \sum_{n=0}^{N} a_n$.
- You should know the various **tests** for convergence: the comparison test, alternating series test, absolute convergence test, ratio test, root test, integral test.

• Integration

- If $I \subset \mathbb{R}$ contains the interval $[a, b]$ and $f\colon I \to \mathbb{R}$ is any function, we can talk about the integral $\int_a^b f(x)\,\mathrm{d}x$, which may or may not exist, but always *does* exist when f is continuous, or even continuous outside of finitely many points.
- Again, it is best – though perhaps not essential – to know the **definition** of the integral, in terms of Riemann sums: at least for continuous functions, the integral $\int_a^b f(x)\,\mathrm{d}x$ is given by the limit $\lim_{N\to\infty} \frac{b-a}{N} \sum_{i=1}^{N} f(a + i\frac{b-a}{N})$.
  * (In general (for not necessarily continuous f), the definition is a bit more complicated: one has to consider arbitrary partitions $a = x_0 \leq \cdots \leq x_N = b$, and not just the "regular partition" $x_i = a + i\frac{b-a}{N}$.)
  * The notation $\int_a^b f(x)\,\mathrm{d}x$ is meant to represent an "infinite sum of infinitesimals" (the integral sign is an "S" for "sum"). This is the classical notion, which again has been superseded by the modern set-theoretic definition. Again, it can be made rigorous with some effort, and again, we do well to follow the physicists and engineers in freely using this intuition when it is useful.
- Again, the most important thing to know are the various **rules**, among which the most important is the **fundamental theorem of calculus**:
  * The integral $\int_a^b f(x)\,\mathrm{d}x$ is equal to $F(b) - F(a)$ where $F\colon [a, b] \to \mathbb{R}$ is any <u>anti-derivative</u> of f, meaning a function with $F' = f$ (assuming that an anti-derivative exists!).
  * Moreover, if f is continuous, then an anti-derivative *does* exist.
  * When anti-derivative F of f exists is uniquely determined up to a constant $C \in \mathbb{R}$, and the "indefinite integral" is defined as $\int f(x)\,\mathrm{d}x = F(x) + C$.
- Other rules you should know include: the sum and difference rule, the substitution rule, and integration by parts, as well as the million little tricks you've learned to apply these in various special cases.
- We will review the important topics of multi-variable integrals and improper integrals if and when they come up.

2.2 10: First-order ODEs

• We now begin our study of differential equations.

- As a preliminary definition, a *differential equation of order* n is an equation whose unknown is a function $y(x)$, and which may involve the derivatives of y up to order n, that is $y, y', y'' \ldots, y^{(n)}$, as well as x itself.
- Examples of order 1 and 2 are

$$y'(x) + y(x) = x \qquad \text{and} \qquad y''(x) + y'(x) = 0.$$

which we usually just write as

$$y' + y = x \qquad \text{and} \qquad y'' + y' = 0.$$

- A solution to this equation on an interval $I \subset \mathbb{R}$ is then a function $y\colon I \to \mathbb{R}$ which satisfies this equation for all $x \in I$ (in particular, y must be n times differentiable).
  - ∗ In principle, one can consider differential equations on *any* domain $I \subset \mathbb{R}$, and not just on an interval I. But if I is not connected, say for example, $I = I_1 \cup I_2$, where I_1 and I_2 are disjoint intervals, then a solution to the differential equation on I just amounts to a solution on I_1 and a separate solution on I_2.
- For example, $y(x) = x - 1$ and $y(x) = e^{-x}$ are solutions to the above two equations.
- (As with ordinary equations, we are often not interested in finding *a* solution, but in finding *all* possible solutions.)

• The above definition will be perfectly satisfactory for our purposes, though it is also good to give a more precise, formal definition of what we mean by a "differential equation", which one can do as follows.

- We define a(n ordinary) <u>differential equation of order n</u> to be an arbitrary function $F\colon \mathbb{R}^{n+1} \to \mathbb{R}$ (or more generally $U \to \mathbb{R}$ for some $U \subset \mathbb{R}^{n+2}$).
  - ∗ Intuitively, we think of F as representing the equation $y^{(n)} = F(x, y, y', \ldots, y^{(n-1)})$.
  - ∗ (One can more generally consider so-called "implicit" equations given by a function $E(x, y, y', \ldots, y^{(n)})$, which we think of as representing the equation $0 = F(x, y, y', \ldots, y^{(n)})$, but this is rarely done.)
- A <u>solution</u> to the differential equation F on $I \subset \mathbb{R}$ is then a n-times differentiable function $y\colon I \to \mathbb{R}$ such that $((x, y(x), y'(x), \ldots, y^{(n-1)}(x)) \in U$ and) $y^{(n)}(x) = F(x, y(x), y'(x), \ldots, y^{(n-1)}(x)) = 0$ for all $x \in I$.
- For example, in this formalism, the above two equations would be given by the functions $F(x, y_0) = x - y_0$ and $E(x, y_0, y_1) = -y_1$, respectively.
- An example where we have to take some $U \subsetneq \mathbb{R}^n$ is the equation $y'' = 1/y' + y$, where $F(x, y_0, y_1) = 1/y_1 + y_0$, and so the largest U we can take is $\{(x, y_0, y_1) \in \mathbb{R}^4 \mid y_1 \neq 0\}$.

• To begin with, we will be considering only **first-order** differential equations.

2.3 10.1: Direction fields

• We consider an arbitrary first-order equation

$$y' = F(x, y).$$

- Any solution $y(x)$ passing through (x_0, y_0) must have slope $F(x_0, y_0)$ at that point.
- Thus, we think of $F(x, y)$ as assigning a slope to each point (which we can represent as a short line segment through that point with that slope); this is called a **slope field** or **direction field**.
- This is not to be confused with a **vector field**.
- We can draw the direction field $F(x, y)$ by sampling several points and drawing line segments, or we can first find several solutions to the differential equation, and then simply plot various tangents along the resulting graphs.

2.3.1 Example 10.1.1

- Consider the equation $y' = -y/x$ for $x \neq 0$.
- We sample the slopes at a few points:

(x, y)	$y' = -y/x$
(1,1)	-1
(1,2)	-2
(2,1)	-1/2
(-1,2)	2
(-2,2)	1

- We obtain a picture as in Figure 10.2 (a)
- Let's solve the equation!
  - Let $I \subset \mathbb{R}$ be one of $(-\infty, 0)$ or $(0, \infty)$.
  - Suppose that $y \colon I \to \mathbb{R}$ is a solution to $y' = -y/x$.
  - Rearrange: $xy' + y = 0$ (equivalent by the assumption $x \neq 0$).
  - Product rule: $(xy)' = 0$.
  - Thus $x \cdot y(x)$ is some constant c.
  - So $y = c/x$.
- Thus, every solution is of the form $y = c/x$.
  - Conversely, since each step above was a **logical equivalence**, it follows that $y = c/x$ **is** always a solution (as one can also check directly).
  - We may say that the general solution to the equation is $y = c/x$ with $x \in \mathbb{R}$.
- We can plot the solutions and see they are tangent to our slope field.

3 Feb 3, 10.2: Separation of variables

3.1 10.1: Direction fields

3.1.1 Example 10.1.2

- If $F(x, y)$ is independent of y, the equation takes the form

$$y' = G(x);$$

  assume G is continuous.

- Then the solutions are simply the antiderivatives of G: $y = \int G(x)\,\mathrm{d}x + C$.
  - All of the solution curves will be parallel to each other, as one can also see from the direction field.

- Example: $y' = \cos x$.
  - Then the general solution is $y = \sin x + c$.
  - Note: when discussing the solutions to a differential equation, one should always first state *on which interval* I one is seeking solutions $y\colon I \to \mathbb{R}$, though this is often left implicit.
  - So here, we should say we are seeking solutions y defined on all of $\mathbb{R}$, and then the general such solution is $y = \sin x + c$.

3.1.2 Example 10.1.3

- We see that differential equations tend to have many solutions.
- We can add extra conditions to a differential equation to single out a particular solution.
- If we demand that y passes through a given point (x_0, y_0), i.e., that $y(x_0) = y_0$, this is called an **initial condition**. (More generally, we can demand that $y^{(i)}(x_i) = y_i$.)
- The problem of satisfying a differential equation with a given initial condition is called an **initial value problem** (IVP).
- The name comes from thinking of x as the "time parameter" and x_0 as the "initial time".
- We return to $y' = -y/x$, $x \neq 0$ with solutions $y = c/x$.
  - Now consider the initial value problem $(x_0, y_0) = (1/2, 2)$.
  - We thus have $2 = c/(1/2)$ and hence $c = 1$.
  - The (unique!) solution to the IVP is thus $y = 1/x$, $x > 0$.
- We can also solve the general IVP for a given (x_0, y_0).
  - We have $y_0 = c/x_0$, so $c = x_0 y_0$, and the solution is $y = x_0 y_0/x$, with $x \in (0, \pm\infty)$ depending on the sign of x_0.

3.2 10.2B Separation of variables

- Among the first-order equations $y' = F(x, y)$, we have seen that the simplest case is when $F(x, y)$ doesn't depend on y. We can then solve it simply by integration.
- We now study the more general case in which in which $F(x, y)$ is of the form $f(x) \cdot h(y)$.
  - Supposing h is non-zero and setting $g(y) = h^{-1}(x)$, these are thus the equations of the form $g(y) \cdot y' = f(x)$.
  - These can often also be solved simply by integration, using a method called "separation of variables".
  - Writing $g(y)\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$, we obtain "$g(y)\,\mathrm{d}y = f(x)\,\mathrm{d}x$" and hence

$$\int g(y)\,\mathrm{d}y = \int f(x)\,\mathrm{d}x.$$

  - Computing antiderivatives F and G of f and g, we thus have $G(y) = F(x) + C$, and can thus find y as long as we can solve for y in this (non-differential-)equation.
  - (We can do the above computation without using "differentials": we simply have $\frac{\mathrm{d}G(y(x))}{\mathrm{d}x} = y'(x) \cdot g(y(x)) = f(x)$ by the chain rule, and hence $G(y(x))$ is an antiderivative of $f(x)$.)

3.2.1 Example 10.2.3

- We consider the differential equation $\frac{\mathrm{d}P}{\mathrm{d}t} = kP$, where $k > 0$ is a constant.
  - This describes the rate of growth of a population of bacteria of size $P(t)$ (say in grams), in which each bacterium reproduces at a rate of k bacteria per second.
  - As always in mathematical modelling, we are making simplifying assumptions in this equation, so we can only expect the solution of the equation to give an approximation to the population growth of the bacteria.
- We know from experience that the solution is $P(t) = Ke^{kt}$ for some constant K.
  - But let us see how we could have arrived at that using separation of variables.
  - Assuming P is always positive, we have $\int \frac{1}{P}\,\mathrm{d}P = \int k\,\mathrm{d}t$, hence $\ln P = kt + C$, hence $P(t) = e^{kt+C} = e^C \cdot e^{kt}$.
  - The solution to the IVP $P' = kP; P(0) = P_0$ is thus $P(t) = P_0 \cdot e^{kt}$.
  - (Assuming instead that P is always *negative* would given $\ln(-P) = kt + C$ and hence $P(t) = -e^C \cdot e^{kt}$ and thus again $P(t) = P_0e^{kt}$; assuming just that P is non-zero would also lead to the same conclusion, since being differentiable and hence continuous, it is then always positive or negative.)
  - **Be careful**: in this computation, we had to make the assumption that P is non-zero, so that we could divide by it. If this assumption wasn't valid, our solution may be incorrect. However, we can just check directly that our solution *is* correct.
- This obviously isn't a realistic solution for large t; one problem is that our model doesn't include the amount of *food* available to the bacteria.
- If we look carefully at what we did, we see that we *almost* showed that this is the unique solution to the IVP; the only problem is we had to assume P was non-zero; the possibility remains that there are other solutions which are zero somewhere (besides the obvious one $P(t) = 0$).

- The trick to show uniqueness without this assumption is to consider the product $P(t) \cdot e^{-kt}$.
- We then have $\frac{\mathrm{d}}{\mathrm{d}t}(Pe^{-kt}) = kPe^{-kt} - kPe^{-kt} = 0$. Hence $e^{-kt} \cdot P(t)$ is equal to a constant K and $P = Ke^{-t}$, as desired.
- We will return to this trick of "exponential multipliers" soon.
- (We also have uniqueness by the general uniqueness theorem we stated, but that is overkill.)

3.2.2 Example 10.2.6

- We consider a tank of a chemical solution, in which a particular chemical is flowing in and out at a particular rate.
- We let $S = S(t)$ be the amount of chemical at time t. Then

$$\frac{\mathrm{d}S}{\mathrm{d}t} = (\text{rate of inflow}) - (\text{rate of outflow}).$$

- Example: a 100-gallon tank contains 150 pounds of salt of salt in solution at time $t = 0$.
  - A salt solution with 2 pounds of salt per gallon is being added at a rate of 2 gallons per minute.
  - The solution, which we take to be homogeneous, is flowing out of the tank at a rate of 2 gallons per minute.
  - Thus salt is flowing in at 4 pounds per minute, and out at $2S(t)/100$ pounds per minute at time t.
- Thus S satisfies the IVP $\frac{\mathrm{d}S}{\mathrm{d}t} = 4 - \frac{2S}{100}; S(0) = 150$.
- We solve it using separation of variables
  - We have $\int \frac{\mathrm{d}S}{S-200} = -\int \frac{\mathrm{d}t}{50}$; we see that this will only be valid if $S - 200$ remains non-zero for all time. Since $S(0) - 200 = -50$, we suppose it remains negative for all t.
  - Hence $\ln\big(-(S-200)\big) = -t/50 + C$.
  - Hence $S = 200 - e^C e^{-t/50}$.
  - The initial condition gives $150 = S(0) = 200 - e^C$ and hence $e^C = 50$.
  - Hence $S = 200 - 50e^{-t/50}$; this is plotted in Figure 10.9; it has a positive slope and an asymptote $\lim_{t\to\infty} S = 200$.
  - Checking directly, we see that this is indeed a solution to the equation.
- (Again, we had to assume $S \neq 200$ here, so we haven't proven that this is the unique solution; again, we can circumvent this using an exponential multiplier.)
  - Note however that the uniqueness *is* important, from a scientific perspective: to draw the conclusion that the amount of chemical in the tank evolves in the way that we found, on the basis of the fact that it solves the given IVP, we have to know that it is the **only** solution to the IVP.

4 Feb 5, 10.3: More separation of variables; existence and uniqueness theorem

4.1 More of 10.2B Separation of variables

4.1.1 Example 10.2.4

- Consider $y' = y/x$, for $x \in \mathbb{R} - \{0\}$.
  - The direction field is shown in Figure 10.7.
- By separation of variables we have $y^{-1}\,\mathrm{d}y = x^{-1}\,\mathrm{d}x$, hence $\ln|y| = \ln|x| + C$, hence $|y| = e^C|x|$, hence $y = \pm e^C x$.
  - Thus, the solutions are $y = kx$ for $k \in \mathbb{R}$.
  - (We should really be more careful here and say that we are looking for solutions on $(-\infty, 0)$ or on $(0, \infty)$, since otherwise we could also have solutions like

$$y(x) = \begin{cases} x & x < 0 \\ 2x & x > 0. \end{cases}$$

  - (The same is true in general when we say $\int x^{-1}\,\mathrm{d}x = \ln|x| + C$: if we take the function x^{-1} to be defined on all of $\mathbb{R} - \{0\}$, then the anti-derivative is only determined up to *two* constants: one on $(-\infty, 0)$ and one on $(0, \infty)$.)
- To prove uniqueness without needing y to be non-zero, we can use a similar trick to last time: we have $0 = y' - y/x$ and hence $0 = y'/x - y/x^2 = \frac{\mathrm{d}(y/x)}{\mathrm{d}x}$. Hence y/x is a constant.
  - (This is secretly again a case of the method of "exponential multipliers".)

4.2 End of 10.1: existence and uniqueness

4.2.1 Example 10.1.4

- In the previous examples, there was a unique solution to each IVP $y' = F(x, y); y(x_0) = y_0$, and this solution extended over the whole given domain; both of these features can fail.
- Consider

$$y' = \begin{cases} \sqrt{y}, & y \geq 0 \\ 0, & y < 0. \end{cases}$$

  - The direction field for this is shown in Figure 10.3 (a).
  - It has two solutions $y\colon \mathbb{R} \to \mathbb{R}$ passing through $(x_0, y_0) = (0, 0)$: $y(x) = 0$ and

$$y(x) = \begin{cases} 0, & x < 0 \\ x^2/4, & x \geq 0. \end{cases}$$

    - To conclude that y is differentiable and compute $y'(0)$, we use the **Theorem** from calculus that if a function f has a "left" derivative $\lim_{h \nearrow 0}(f(a+h) - f(a))/h$ and a "right" derivative $\lim_{h \searrow 0}(f(a+h) - f(a))/h$ at a given point a, and they agree, then f is differentiable at a.
  - In fact, there are infinitely many solutions to this IVP.
- As we will see, there is a simple condition on $F(x, y)$ which guarantees that this cannot not happen.

4.2.2 Example 10.1.5

- Now consider $y' = 1 + y^2$.
  - The direction field is shown in Figure 10.3 (b).
  - It has the solution $y = \tan x$ passing through $(x_0, y_0) = (0, 0)$.
  - This can only be extended to $x \in (-\pi/2, \pi/2)$ since it tends to $\pm\infty$, despite $F(x, y)$ being well-behaved everywhere.
  - In particular, had we stated the problem in the form "find a solution to this IVP which is defined on all of $\mathbb{R}$", the conclusion would have been that *this problem does not have a solution.*
- Again, this can be avoided by putting a simple condition on F.

4.2.3 Existence and uniqueness theorem

- We now state a general theorem which guarantees that a unique solution to a given IVP exists, if we assume that F satisfies certain conditions which in particular rule out the above pathologies.
  - The existence result should be cared to analogous phenomena for polynomial equations.
  - For degree 2 polynomial equations, we have an explicit method of solution, and even an explicit formula; this is analogous to the explicit methods we have introduced and will introduce to solve certain classes of differential equations.
  - On the other hand, we also know the general result (whose proof is easy): *every odd degree polynomial equation has a solution.* In this case, we are guaranteed on general grounds that a given equation has a solution, but we are not given any means to find it. The following existence theorem is of a similar nature.
- **Theorem**: Suppose $F\colon U \to \mathbb{R}$ is continuous, where $U = I_1 \times I_2$ for some intervals $I_1, I_2 \subset \mathbb{R}$, and that $F_y\colon U \to \mathbb{R}$ exists and is continuous.
  - Then for any $(x_0, y_0) \in U$, the IVP $y' = F(x, y); y(x_0) = y_0$ has a solution defined on some interval $I \subset I_1$, and this solution is unique in the sense that for any two solutions $y_1\colon I \to \mathbb{R}$ and $y_2\colon I' \to \mathbb{R}$, $y_1(x) = y_2(x)$ for $x \in I \cap I'$.
  - If moreover $I_2 = \mathbb{R}$ and there exists $B \in \mathbb{R}$ with $F_y(x, y) < B$ for all $(x, y) \in U$, then the solution will exist on the entire interval I_1.
- The first condition fails for Example 10.1.4, since $y \mapsto \sqrt{y}$ is not differentiable at $y = 0$.
  - The second condition fails for Example 10.1.5, since $1 + y^2$ is not bounded as $y \to \pm\infty$.

5 Feb 10, Proof of existence and uniqueness; 10.1: Numerical methods; and 10.3: Linear equations

5.1 End of 10.1: existence and uniqueness

- We will sketch the proof of the existence and uniqueness statement from last time.
- The proof of the second part is fairly easy: the point is that if the slope of $y(x)$ is bounded, then it can only increase by a finite amount in finite time by the **mean value theorem**.
  - The general fact is that if a function $y\colon I \to \mathbb{R}$ satisfies $|y'(x)| < B$ for all $x \in I$, then for any $a < b$ in I, we have $|y(b) - y(a)| < B(b-a)$.
  - Indeed, if we had $|y(b) - y(a)|/(b-a) \geq B$, then the mean value theorem would given $y'(\xi) = |y(b) - y(a)|/(b-a) \geq B$ for some $\xi \in (a, b)$, contradicting the assumption that $y'(x) < B$ for all $x \in I$.
  - One then has to show that the only way that a solution to the IVP can *fail* to extend over the entire interval I is if it diverges to $\pm\infty$.
- The proof of the first part uses an ingenious trick called **Picard iteration**.
  - One first converts the equation to the equivalent **integral equation** $y(x) = y_0 + \int_{x_0}^{x} F(x, y(x))\,\mathrm{d}x$; this integral exists since F is continuous.
  - One then finds a sequence of better and better **approximations** to the solution:
  - The first is simply $f_1(x) = y_0$.
  - Then we inductively define $f_{n+1}(x) = y_0 + \int_{x_0}^{x} F(x, f_n(x))\,\mathrm{d}x$.
  - Finally, we define $y(x) = \lim_{n\to\infty} f_n(x)$.
  - We then have

    $$y(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_{n+1}(x) = \lim_{n\to\infty} y_0 + \int_{x_0}^{x} F(x, f_n(x))\,\mathrm{d}x = y_0 + \int_{x_0}^{x} F(x, y(x))\,\mathrm{d}x,$$

    as desired.
  - The tricky part is to show that the limit defining y **actually converges** (and that the exchanging of limit and integral in the last equation is legitimate); this is where the assumption is used that F_y exists and is continuous.
- (This is an instance of the general technique of *finding fixed points using iteration*: given a domain U and a continuous function $G\colon U \to U$, if we want to find a *fixed point* of G, i.e., a point $x \in U$ with $G(x) = x$, we can choose some arbitrary $x_0 \in U$, and iteratively define $x_{n+1} = G(x_n)$, and set $x = \lim_{n\to\infty} x_n$ *if this limit exists.* Using the continuity of G, we then have $G(x) = G(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} G(x_n) = \lim_{n\to\infty} x_{n+1} = x$, as desired.)

5.2 10.1B: Numerical Methods

- In scientific applications, it is often important not to explicitly *solve* a given IVP $y' = F(x, y); y(x_0) = y_0$ (which often cannot be done anyway), but to compute numerically an approximation to its solution.
  - This amounts to running a simulation of a quantity y whose rate of change at each time x is given by $F(x, y)$.
  - Geometrically, it amounts to tracing out a curve tangent to a given direction field.

- A straightforward way to do this **Euler's method**:
  - Fix some step size $h > 0$.
  - We are given x_0 and y_0.
  - Now define $x_n = x_0 + n \cdot h$ for $n > 0$, and recursively define $y_{n+1} = y_n + F(x_n, y_n) \cdot h$.
  - It is easy to implement this in any programming language; an example is given in the book.
- There are many ways to improve this algorithm, and there is a whole field dedicated to studying such things.
  - A central problem that arises is the accumulation of **rounding errors**, for example if one tries to make the above step-size parameter h too small.
  - One such improved algorithm is discussed in the book.

5.3 10.3: Linear equations

- A first-order equation $y' = F(x, y)$ is called linear if F is of the form $F(x, y) = -g(x)y + f(x)$.
  - The equation can then be written as $y' + g(x)y = f(x)$; this is called its **normalized form**.
  - The name *linear* corresponds to the fact that $L(y) = y' + g(x)y$ is a *linear* function of y: we have $L(y_1 + y_2) = L(y_1) + L(y_2)$ and $L(a \cdot y) = a \cdot L(y)$ for $a \in \mathbb{R}$ (we will return to this when we do some linear algebra review).
  - Thus, a linear equation has the form $L(y) = f(x)$ for some linear operator L.
  - Just like in linear algebra, we call the equation homogeneous if $f(x) = 0$ and inhomogeneous otherwise.

5.3.1 Example 10.3.1

- In the homogeneous case, if we assume y is never 0, we obtain $\frac{y'}{y} = -g(x)$.
- Integrating, this gives $\ln y = -G(X) + C$ for some $C \in \mathbb{R}$ if $y > 0$ and $\ln(-y) = -G(x)\,\mathrm{d}x + C$ if $y < 0$, where G is some anti-derivative of g, i.e., $\int g(x)\,\mathrm{d}x = G(x) + C$.
- Hence $y = \pm e^C e^{G(x)}$ and so in either case, $y = Ke^{-G(x)} = Ke^{-\int g(x)\,\mathrm{d}x}$ for some $K \neq 0$.
  - As usual, we can now check directly that this is indeed a solution.
  - (And we notice that $K = 0$ also yields a solution.)

5.4 10.3A: Exponential integrating factors

- In order to show that the solution $y = Ke^{-G(x)}$ is the *general* solution (which we don't quite know since we had to assume $y \neq 0$), we use a trick we have used before, and divide the original equation by the known solution $y = e^{-G(x)}$.
  - We obtain $0 = e^{G(x)}y' + e^{G(X)}g(x) = \frac{\mathrm{d}}{\mathrm{d}x}(e^{G(x)}y)$.
  - Note that this is equivalent to the original equation, since we have multiplied by both sides by a non-vanishing quantity.
  - Thus, we see that for any solution y to this equation, $e^{G(x)y}$ must be constant, and so $y = Ke^{-G(x)}$ for some $K \in \mathbb{R}$, as desired.
- This is the trick of "exponential integrating factors" and it also allows us to solve linear equations in the *inhomogeneous case*.

- For the equation $y' + g(x)y = f(x)$ in normalized form, the exponential integrating factor is $M(x) = e^{\int g(x)\,\mathrm{d}x}$.
- We multiply the equation by M and obtain

$$e^{\int g(x)\,\mathrm{d}x} y' + g(x) e^{\int g(x)\,\mathrm{d}x} = f(x) e^{\int g(x)\,\mathrm{d}x}.$$

- Using the product rule, the left-hand side is $\frac{\mathrm{d}}{\mathrm{d}x}(y e^{\int g(x)\,\mathrm{d}x})$, and the equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}x}(y e^{\int g(x)\,\mathrm{d}x}) = f(x) e^{\int g(x)\,\mathrm{d}x}.$$

- Now we can just integrate the right-hand side and solve for y.
  - Again, the original equation is equivalent to the equation after multiplying by M since M is nowhere zero.

5.4.1 Example 10.3.2

- Let's solve $y' = xy + x$.
- We rewrite it in normalized form $y' - xy = x$.
- Multiply by the integrating factor $y' e^{-x^2/2} - xy e^{-x^2/2} = x e^{-x^2/2}$.
- Now integrate $y e^{-x^2/2} = -e^{-x^2/2} + C$.
- Thus $y = -1 + C e^{-x^2/2}$.
- Since each step was an equivalence, this is the *general* solution to the equation (with domain $\mathbb{R}$).