Overview/Summary
MATH 6701
Mathematical Methods of Applied Sciences

John McCuan

November 19, 2020

This course had three components: Complex Analysis, Linear Algebra, and Ordinary Differential Equations. The main prerequisite, and perhaps the main thing to be learned in the course, might be called "function theory" or knowing what a function is and what it means to understand functions of various sorts. To understand any function requires some understanding of sets. Every function has a domain and a codomain. Hopefully, every student has had an opportunity to understand what is intended by the symbols $f : A \to B$ and $b = f(a)$. Every individual function $f : A \to B$ can be understood in terms of another set of ordered pairs

$$\mathcal{G} = \{(a, f(a)) \in A \times B : a \in A\}$$

called the graph of the function—if indeed that set can be understood. Again, hopefully every student has had the opportunity to appreciate the set notation,

$$\{x \in \mathcal{U} : P(x)\}$$

where P is some proposition specifying x, just used to express the graph.

Other sets of frequent interest in the consideration of a function $f : A \to B$ are the images of sets $S \subset A$ under f:

$$f(S) = \{f(a) \in B : a \in S\},$$

and the preimages of sets $T \subset B$ under f:

$$f^{-1}(T) = \{a \in A : f(a) \in T\}.$$

At length, one considers sets of functions, which we often refer to as collections of functions. These are often delineated by using (or assuming) certain structures on the domain and/or codomain. A domain X with structure is often called a space. Certainly a prototypical space for mathematical methods in applied sciences is the set of real numbers $X = \mathbb{R}$. The real numbers, in particular, provide an example of a set/space with many algebraic, geometric, and set theoretic properties. In fact, $\mathbb{R}$ is a very rich and complicated space. Abstracting various of the properties of $\mathbb{R}$ leads to the notions of other important spaces. Among these are two that have been particularly emphasized in this course:

1. fields
2. vector spaces over fields.

Also mentioned somewhat superficially, but in practice of rather great importance in applied science, were measure spaces. Hopefully every student had the opportunity to understand the detailed definitions of fields and vector spaces. Foundational to the notion of a field (and of course to a vector space as well) is that of a group. A group is a set with an operation satisfying three properties. To be precise, a set G is a group if there is an operation (i.e., a certain kind of function) $* : G \times G \to G$ with values often written as $*(a, b) = a * b$ and satisfying

1. The operation is associative:
$$a * (b * c) = (a * b) * c \qquad \text{for every } a, b, c \in G.$$
2. There exists an element $e \in G$ called the identity element defined by the property
$$e * g = g * e = g \qquad \text{for every } g \in G.$$
3. For every element $g \in G$ there exists an inverse element $h \in G$ defined by the property
$$g * h = h * g = e.$$

In various contexts the operation may be denoted by "+" (and called addition), "·" or adjacency "ab" (and called multiplication), "∘" (and called composition), or other things. It is usual for the inverse of an element a with respect to "addition" to be denoted by $-a$, while the inverse of an element a with respect to multiplication (or composition, either one) is usually denoted by a^{-1}. If the operation is an "addition," then the identity element is a "zero" and is denoted by something like 0 or $\mathbf{0}$. If the operation is a "multiplication," then the identity element is often called 1. For "composition" rules in a group G the symbol id or id_G is often used.

Additional structures may be placed upon that of a vector space. Two notable ones we have had occasion to consider, at least superficially, are

1. vector spaces with norms, i.e., normed vector spaces, and
2. vector spaces with inner products, or inner product spaces.

Underlying geometric considerations, in a manner somewhat analogous to the manner in which a group underlies the algebraic notions of field and vector space, is the notion of a topological space. A topological space is one with a specified collection of subsets called the open sets. The collection $\mathcal{T}$ of open subsets of a topological space X is called the topology on X and is defined by the following three properties:

1. The empty set ϕ and the entire space X are both open, i.e., $\phi, X \in \mathcal{T}$.
2. $\mathcal{T}$ is closed under unions, that is given any collection of open sets

$$\mathcal{U} = \{U \in \mathcal{T} : P(U)\} \subset \mathcal{T},$$

the set

$$\bigcup_{U \in \mathcal{U}} U \in \mathcal{T}.$$

3. $\mathcal{T}$ is closed under finite intersections: If $\{U_1, U_2, \ldots, U_k\} \subset \mathcal{T}$ is a finite collection of open sets, then

$$\bigcap_{j=1}^{k} U_j \text{ is open.}$$

In the applied sciences most topoligical spaces X of interest have the additional property that given two elements $a, b \in X$ with $a \neq b$, there exist disjoint open sets U and V with $a \in U$ and $b \in V$. A topological space with this property is said to have the Hausdorff property to be be a Hausdorff (topological) space. A Hausdorff space, in the most abstract sense, is one in which it is possible to measure the "closeness" of elements perhaps without having an actual distance as produced by a metric, a norm, or an inner product (concepts which will be covered in more detail next semester). I give you an example of a non-Hausdorff topological space (which may be considered to be of interest in applied science) in an exercise below.

A structure with somewhat the same set theoretic flavor as that of a topology is the notion of a measure space. Sometimes measure spaces are considered to have, in fact, a geometric structure, or put another way, a measure space is sometimes considered a geometric space, the defining characteristic being the ability to "measure" sets and, in particular, the closeness of sets much in the manner that a topology or metric measures the closeness of elements. One simple definition is the following: A measure space X is a set with a specified collection $\mathcal{M}$ of subsets called the measurable sets and a function $\mu : \mathcal{M} \to [0, \infty]$ called a measure defined by the following properties:

1. $X \in \mathcal{M}$,

2. $\mathcal{M}$ is closed under complments, i.e., if $A \in \mathcal{M}$ is measurable, then $A^c = X \backslash A = \{x \in X : x \notin A\}$ is also measurable.

3. $\mathcal{M}$ is closed under countable unions: If $\{A_1, A_2, A_3, \ldots\} = \{A_j\}_{j=1}^{\infty} \subset \mathcal{M}$, then
$$\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}.$$

4. If $\{A_j\}_{j=1}^{\infty} \subset \mathcal{M}$ is a sequence of measureable sets that are pairwise disjoint, then
$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

5. $\mu(\phi) = 0$, i.e., the empty set has zero measure.

The first three properties are said to make the measurable sets $\mathcal{M}$ a sigma algebra of sets. The fourth property is called countable additivity. The last property is really included to avoid one particular degenerate case which is of no particular interest.

Exercise 1 *Show that if the properties of a measure space hold for X except, possibly, for the fifth and last property, and there exists a set $A \in \mathcal{M}$ with $\mu A < \infty$, then X is a measure space with measurable sets $\mathcal{M}$ and measure μ, i.e., you can prove the last property using the first four properties.*

Let $(X, \mathcal{M}, \mu)$ be a triple satisfying the following:

(i) *X is a set.*

(ii) *$\mathcal{M}$ is a collection of subsets of X.*

(iii) $\mu : \mathcal{M} \to [0, \infty]$ *is a function.*

(iv) *Properties 1-4 of a measure space hold but* $\mu(\phi) \neq 0$.

Describe the resulting "measure space" $(X, \mathcal{M}, \mu)$.

Given a vector space X, a function $\lfloor\ \rfloor : X \to [0, \infty]$ is a **seminorm** on X if

(i) $\lfloor af \rfloor = a\lfloor af \rfloor$ for every $a \in \mathbb{R}$ and $f \in X$.

(ii) $\lfloor af \rfloor = a\lfloor f + g \rfloor$ for $f, g \in X$.

The following exercise describes two non-Hausdorff spaces which may be of interest in the applied sciences:

Exercise 2 *Consider* $\mathfrak{L}^1(0, 1)$ *the set of all (Lebesgue) integrable functions* $u : (0, 1) \to \mathbb{R}$ *and*

$$\mathcal{W} = \{f \in X : f \text{ has a weak derivative in } \mathfrak{L}^1(0, 1)\}.$$

(a) *Show* $\lfloor\ \rfloor_{\mathcal{L}^1} : \mathcal{L}^1(0, 1) \to \mathbb{R}$ *by*

$$\lfloor f \rfloor_{\mathcal{L}^1} = \int_{(0,1)} |f|$$

is a seminorm on $\mathcal{L}^1(0, 1)$

(b) *Show* $\lfloor\ \rfloor_{\mathcal{W}} : \mathcal{W} \to \mathbb{R}$ *by*

$$\lfloor f \rfloor_{\mathcal{W}} = \int_{(0,1)} |g|$$

where g is the weak derivative of f is a seminorm on $\mathcal{W}$.

(c) *Let X be a vector space with a seminorm. Show that the collection* $\mathcal{T}$ *of all subsets of X satisfying the following property is a topology on X:* $U \in \mathcal{T}$ *if and only if*

For each $x_0 \in U$*, there is some* $r > 0$ *such that*

$$\{x \in X : \lfloor x - x_0 \rfloor < r\} \subset U.$$

(d) *Show that the seminorm topoligies on* $\mathfrak{L}^1(0, 1)$ *and* $\mathcal{W}$ *are non-Hausdorff topologies.*

1 Complex Analysis

Our discussion of complex analysis has been framed around the consequences of the existence of a field extension of $\mathbb{R}$ having a solution to the quadratic polynomial equation $x^2 + 1 = 0$. Thus one obtains the complex numbers

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}$$

which also constitutes a field.

The field of complex numbers may be viewed as no more than a different notation for capturing certain properties and structures associated with the Euclidean plane $\mathbb{R}^2$ under the identification $(x, y) \longleftrightarrow x + iy$. However, the economy and complexity of the information captured using the complex product in conjunction with the plane is both surprising and beautiful.

In particular, one obtains the collection of (complex) differentiable, or analytic, functions $f : \mathcal{U} \to \mathbb{C}$ where $\mathcal{U}$ is an open subset of $\mathbb{C}$. These turn out to be the functions which are locally approximated (up to first order) at each point by a translation, a rotation, and an isotropic scaling.

Exercise 3 *Give expressions for the local rotation and scaling factor of an analytic function $f : \mathcal{U} \to \mathbb{C}$ at the point $z = z_0 \in \mathcal{U} \subset \mathbb{C}$.*

The scaling factor is also called the conformal factor at $z = z_0$ and complex analytic functions are called conformal maps. The analytic functions are a relatively small collection of functions which is still quite rich and has a number of important applications.

Exercise 4 *The characterization of analytic functions in terms of conformality given above is not quite correct. What is the exceptional case in which an analytic function may not be approximated by a rigid motion at a point? What else can happen in this exceptional case?*

Functions $L : \mathbb{C} \to \mathbb{C}$ which are linear over $\mathbb{C}^1$ considered as a vector space over the field $\mathbb{C}$, satisfying

$$L(az + bw) = aLz + bLw \qquad \text{for } a, b \in \mathbb{C} \text{ and } z, w \in \mathbb{C}^1$$

provide a special case of analytic functions in which

1. $L(0 + 0i) = 0 + 0i$.

2. The conformal factor and the rotation are global, i.e., the same for all points.

Exercise 5 *Express a linear function* $L : \mathbb{C}^1 \to \mathbb{C}^1$ *in terms of matrix multiplication using the identification* $(x, y) \longleftrightarrow x + iy$*. What are the rotation and conformal factor in terms of this matrix?*

Exercise 6 *Given a linear function* $L : \mathbb{C}^1 \to \mathbb{C}^1$*, is it possible to have*

$$\{z \in \mathbb{C}^1 : Lz = 0 + 0i\}$$

a straight line?

Exercise 7 *What does Liouville's theorem say about a linear function* $L : \mathbb{C}^1 \to \mathbb{C}^1$*?*

2 Linear Algebra

General linear functions $L : V \to W$ where V and W are any vector spaces, can display considerably more variety than observed in linear functions $L : \mathbb{C}^1 \to \mathbb{C}^1$. We approached this subject by trying to understand some of the simplest linear functions. Perhaps it is in order now to say what we mean by "understanding" such a function. Broadly speaking, this means we can answer any (reasonable) question one might be inclined to ask about such a function. Among the simplest questions one might ask about a linear function $L : V \to W$ are the following:

1. What are the images of various subsets of V?
2. What are the preimages of various subsets of W?
3. How do the properties of these subsets transform under L?

Our cursory attempt at understanding "all" linear functions on vector spaces included the following:

1. $L : \mathbb{R}^1 \to \mathbb{R}^n$.
2. $L : \mathbb{R}^n \to \mathbb{R}^1$.
3. $L : \mathbb{R}^2 \to \mathbb{R}^2$.

This, it was claimed, is a pretty good start. We did not obtain a complete understanding of the third case, as I will try to indicate briefly below.

Exercise 8 *Explain "everything" about linear functions* $L : \mathbb{R}^1 \to \mathbf{r}^n$.

Exercise 9 *Explain "everything" about linear functions* $L : \mathbb{R}^n \to \mathbf{r}^1$.

If one were to know "everything" about a linear function $L : \mathbb{R}^2 \to \mathbb{R}^2$, then one should be able to answer, for example, the following three[1] questions:

1. What is the image of an ellipse in standard position centered at the origin under L, i.e.,
$$\left\{ L \begin{pmatrix} x \\ y \end{pmatrix} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}?$$
2. What is the preimage of $\{\mathbf{0}\}$, i.e., $\ker(L)$?
3. If $S \subset \mathbb{R}^2$ has area A, then what is the area of the image $L(S) = \{L(\mathbf{x}) : \mathbf{x} \in S\}$?

Exercise 10 *Answer the last question about the transformation of area under a linear function. Hint: Consider the image of a small square* $[0, \epsilon] \times [0, \epsilon] = \{(x, y)^T : 0 \le x, y \le \epsilon\}$, *and then express the area of any set in terms of the areas of small squares.*

We did not fully answer the first question about ellipses. But we did answer it in some special cases.

Exercise 11 *Describe the image of an ellipse (and a rectangle* $\{(x, y)^T : 0 \le x \le a,\ 0 \le y \le b\}$*) in standard position in the following cases:*

(a) $L : \mathbb{R}^2 \to \mathbb{R}^2$ *by*
$$L\mathbf{x} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{x}.$$

(b) $L : \mathbb{R}^2 \to \mathbb{R}^2$ *by*
$$L\mathbf{x} = \begin{pmatrix} \lambda & -\mu \\ \mu & \lambda \end{pmatrix} \mathbf{x}.$$

(b) $L : \mathbb{R}^2 \to \mathbb{R}^2$ *by*
$$L\mathbf{x} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathbf{x}.$$

Hopefully, every student had the opportunity to discover that one can talk about linear functions without using matrices—and that one can also talk about linear functions $L : \mathbb{R}^n \to \mathbb{R}^m$ using matrices.

[1]These three questions, it may be noted, are representative of the three general questions above.

Exercise 12 *Given a linear function $L : \mathbb{R}^n \to \mathbb{R}^m$, explain how to find the matrix A such that*

$$L\mathbf{x} = A\mathbf{x}.$$

In particular, hopefully every student knows the difference between a linear function and a matrix.

Another question arose which we were able to answer for linear functions $L : \mathbb{R}^2 \to \mathbb{R}^2$:

> What linear function is represented by
>
> $$e^L = \sum_{j=0}^{\infty} \frac{1}{j!} L^j$$
>
> where L^j represents function composition (repeated j times)?

Exercise 13 *Describe the three canonical forms for linear functions $L : \mathbb{R}^2 \to \mathbb{R}^2$ and the meaning of the matrix exponential above.*

This last topic of matrix exponentiation turns out to be related to some of our work on the final topic of the course.

3 Ordinary Differential Equations (ODE)

This subject was presented in terms of finding a function whose (first) derivative (with respect to one real variable) is specified. This can be an interesting question for a single real valued function:

$$y' = f(y, t)$$

and there are various techniques for solving such equations.

Exercise 14 *List all the single first order ODEs you know how to solve. What first order ODEs can you not solve?*

The subject, however, is really about systems of equations, or a single vector valued function of one real variable $\mathbf{x} : (a, b) \to \mathbb{R}^n$ or $\mathbf{z} : (a, b) \to \mathbb{C}^n$:

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}, t). \tag{1}$$

The collection of solutions $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ of ODEs having the form (1) includes all linear functions $L : \mathbb{R}^1 \to \mathbb{R}^n$ and a great many more interesting functions with numerous applications.

Central to all considerations of these problems are existence and uniqueness theorems. Hopefully, every student had an opportunity to appreciate the implications of the existence and uniqueness theory for ODEs (at least to some extent).

Exercise 15 *What conditions guarantee (local) existence of solutions to the initial value problem (IVP) associated with an ODE (1)?*

Exercise 16 *What conditions guarantee (local) uniqueness of solutions to the IVP associated with an ODE (1)?*

Exercise 17 *Solve the autonomous single equations $y' = y^2$ and $y' = \sqrt{|y|}$ and explain their significance with respect to the existence and uniqueness theorems.*

The first order ODE which you cannot solve can often be analyzed numerically using a "canned" solver like `NDSolve` in Mathematica or `ode45` in Matlab. The "45" in `ode45` stands for "adaptive fourth and fifth order Runge-Kutta method," and `NDSolve` uses essentially the same algorithm. We did not discuss these algorithms, but it is probably more useful for you to focus on the existence and uniqueness theorems that are often necessary to interpret the results of using these "canned" numerical solvers. Of course, it is well worth taking a few minutes to make sure you know how to implement some such numerical solver.

We also considered briefly linear ODE and the associated "linear theory."

Exercise 18 *What makes an ODE a linear ODE? What is the associated homogeneous equation? How do the requirements for existence and uniqueness change/improve for linear ODE?*

It should be noted here that any regular n-th order ODE

$$y^{(n)} = F(y^{(n-1)}, \dots, y', y, t) \tag{2}$$

is equivalent to a first order system of ODEs.

Exercise 19 *Explain the equivalence between an ODE of the form (2) and a first order system. Use matrix multiplication to express the equivalent system in the case where (2) is linear.*

We focused to some extent on autonomous equations:

$$y' = f(y) \qquad \text{and} \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x}).$$

These equations have solutions that are invariant under a shift in the independent variable. More precisely, if $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ is a solution of $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, then for every fixed $t_0 \in \mathbb{R}$, the function $\tilde{\mathbf{x}} : \mathbb{R} \to \mathbb{R}^n$ by

$$\tilde{\mathbf{x}}(t) = \mathbf{x}(t - t_0)$$

is also a solution. This property makes consideration of orbits in phase space the natural setting for studying autonomous ODE.

Exercise 20 *Explain the difference between solution space, phase space, and specification space. Define orbits and equilibrium points.*

As a special case, we explained that the solution of the IVP

$$\begin{cases} \mathbf{x}' = A\mathbf{x} \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases}$$

where A is an $n \times n$ matrix with constant entries is

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0.$$

An understanding of these solutions, which hopefully every student had a chance to review/learn in the case where A is a 2×2 matrix, can be used to analyze more interesting nonlinear autonomous systems

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}) \tag{3}$$

for $\mathbf{x} : (a, b) \to \mathbb{R}^n$. In particular, given an equilibrium point $\mathbf{x}_*$ in phase space, the linearization of (3) at $\mathbf{x}_*$ is the constant coefficient linear ODE

$$\mathbf{x}' = D\mathbf{F}(\mathbf{x}_*)\mathbf{x}.$$

Exercise 21 *Explain how to compute the matrix* $D\mathbf{F}(\mathbf{x}_*)$.

We also reviewed briefly some related concepts like nullclines and separatrices. These are important when the system (3) is for $\mathbf{x} : \mathbb{R} \to \mathbb{R}^2$ and one wishes to obtain a phase plane analysis.

And that's about where we left it.

Exercise 22 *Explain why there are no separatrices in phase space for* $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ *for* $\mathbf{x} : \mathbb{R} \to \mathbb{R}^3$.

4 Epilogue

The course has also been presented with the assumption that the best review of mathematical topics is attained by using those topics to learn something new, or at least to view those topics from a different perspective or point of view, so it seems like you are learning something new and can consolidate and expand the material/framework of what you have learned before. Of course, if you didn't really learn anything about linear algebra or ordinary differential equations before this course, then it is difficult to use what you learned or build anything on it. The material on complex analysis was probably quite new to most of you, but the principle and approach are the same: Start with something you know, and find a good question upon which to focus; attempt to understand that question and try to answer that question. If the question you've chosen is too difficult, then someone who understands the answer already may be able to suggest a different question or point you in a different direction. The best questions, it seems, have the quality that one can say about them:

> This is something I really should know or be able to figure out...

but I don't know it. If you have no question(s), then you really have no place to start, and reviewing (or learning) is going to be difficult.

With this in mind, I mentioned one "building question," which was the following:

> The standard existence and uniqueness theorems for ODEs require the specification of the derivative to be continuous (and to be continuously differentiable in some variables for uniqueness). Under what circumstances can one relax these assumptions, and how would one do that?

Transferred to the simplest ODE $y' = f(x)$ which one encounters in calculus, and particularly in the fundamental theorem of calculus, one finds the the framework of $C^0[a,b]$ specification and $C^1[a,b]$ solutions is very clean and convenient, but does not include all situations of interest. In particular, one discovers that there are functions with derivative existing that are not in $C^1[a,b]$, i.e., the derivative is not continuous.

Exercise 23 *Find/remember a differentiable function whose derivative is not continuous.*

Natural questions are, perhaps, the following:

1. Under what conditions can one define a derivative of a function? When does the limit of a difference quotient exist? And (not so obvious) are there alternatives to defining a derivative which are different from using the limit of a difference quotient?

2. Under what conditions is the indefinite integral

$$F(x) = \int_a^x f(t)\,dt$$

differentiable? The fundamental theorem of calculus gives one, namely that f be continuous (so that f is Riemann integrable). But there are other situations we know about, using for example, the heaviside function in the integrand, where it seems reasonable to consider some kind of derivative of the indefinite integral. What is a natural framework which is more general than the elelgant $C^0(a,b)$-$C^1(a,b)$ framework?

3. Given that the indefinite integral F from the previous problem is differentiable, when is it true that $F'(x) = f(x)$?

There are various answers to the questions above. One could consider, for example, Carathéodory's strong solutions of the ODE $y' = f(x)$. A more popular approach these days is through the notion of weak solutions and Lebesgue integrable functions, so I attempted a cursory introduction to at least some topics related to this approach.

At the very least, I hope every student had the opportunity to appreciate the heirarchy of regularity associated with the inclusions

$$C^\omega(a,b) \subset C^\infty(a,b) \subset \cdots \subset C^k(a,b) \subset \cdots \\ \subset C^1(a,b) \subset C^0(a,b) \subset \text{Riem}_{loc}(a,b) \subset \mathfrak{L}^1_{loc}(a,b)$$

and the associated inclusions

$$C^1(a,b) \subset W^1(a,b) \subset \mathfrak{L}^1_{loc}(a,b).$$

Weak solutions are very natural if on believes, as most people who study differential equations do, in integration by parts. I hope it may seem natural, to at least some students, at this point to ask about weak solutions for other ordinary differential equations beyond the FTC equations. For example, one might consider the next simplest ODE to be the linear first order equation $y' + ay = f(x)$ where a is a constant.

Exercise 24 *Formulate what it means for u to be a weak solution of the linear first order ODE $y' + ay = f(x)$. What can you say about the "linear theory" of weak solutions for this ODE? Can you make sense of a linear operator corresponding to $Ly = y' + ay$? (You must be careful because you can't take a classical derivative.) What about the kernel? Is it still a subspace of some appropriate vector space of functions? Is it still one dimensional?*