MATH2021 NOTES

KYLE BRODER

1. Part 1 — What is a linear differential equation.

Before we introduce the Wronskian, let's recall what it means for a differential equation to be linear: For simplicity, we'll stick to ODEs of order (i.e., the number of derivatives in the equation) is at most 2. Consider the second-order ODE

$$y'' + p(x)y' + q(x)y^m = 0.$$

Then it common to define a such a differential equation to be linear if $m = 1$. That is, the equation

$$y'' + 2xy' + 3x^2y = 0$$

is linear, but

$$y'' + 2xy' + 3x^2y^3 = 0.$$

is not linear.

Now, while this is correct, it will be crucial for today's discussion to have a more meaningful expression of what it means of an ODE to be *linear*. Think about where you have studied linear things before, preferably prior to the study of ODEs. Of course, we study linear objects in linear algebra. What do we study in linear algebra, we study matrices, vectors, solutions of matrix equations, and so on, but primarily, we study vector spaces (which we will briefly review in a moment) and the maps between them, namely, linear maps.

Linear algebra is one of the most important subjects to have a thorough understanding of in all of mathematics and, in fact, one can make the argument that, really, the only thing we understand in mathematics is linear algebra. Think about the derivative of a function for example; the notion of a slope of a function is understood precisely by approximating by a linear object.

Definition. So to remind ourselves, let's recall that a **vector space** is a set V such that:

(i) if $\mathbf{u}, \mathbf{v} \in V$, then $\mathbf{u} + \mathbf{v} \in V$, this is commonly referred to as vector addition.

(ii) if $\mathbf{u} \in V$ and $\lambda \in \mathbb{R}$, then $\lambda\mathbf{u} \in V$, and this is commonly referred to as scalar multiplication.

Another way of formulating this first condition is saying that a vector space is a set which is closed under addition (of vectors), and the second condition states that a vector space is closed under scalar multiplication.

An important concept that was introduced in linear algebra was that of linear independence, expressing the "new linear information" (or lack thereof) that is obtained by introducing a new vector into a given set of vectors. For example, if we consider the vector $\mathbf{u} = (1, 0)$ and the vector $\mathbf{v} = (2, 0)$, then $\mathbf{v} = 2\mathbf{u}$, and so $\mathbf{v}$ gives no more "linear information" than what $\mathbf{u}$ gives, so these vectors, being scalar multiples of each other, are **linearly dependent.**

If we consider the vectors $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (0, 1)$, however, then these vectors represent very different information from the point of view of linear algebra. More precisely, if we looked at the equation $\lambda_1\mathbf{u} + \lambda_2\mathbf{v} = \mathbf{0}$, the fact that $\mathbf{u}$ and $\mathbf{v}$ carry different "linear information" is made precise by saying that this equation has only the trivial solution

$$\lambda_1 = \lambda_2 = 0.$$

Let's make this more explicit in the following definition:

Definition. A set of vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ in a vector space V is said to be **linear independent** if, for scalars $\lambda_1, ..., \lambda_n \in \mathbb{R}$,

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = 0 \implies \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

One of the key concepts that the notion of linear independence gives rise to is the notion of a basis: The notion of a basis allows us to express all the linear information of a vector space (commonly, an infinite-set) in terms of a finite set of vectors.

Definition. A set of vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ in a vector space V is said to be a basis for V if

(i) the vectors are linearly independent;
(ii) the vectors span V in the sense that any vector $\mathbf{u} \in V$ can be written as

$$\mathbf{u} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n.$$

Ok, so at this point you're probably thinking, I'm not supposed to be learning about linear algebra, I'm supposed to be learning about differential equations. Well, a differential equation is linear precisely when the set of solutions of the ODE form a vector space. That is, if y_1, y_2 are two solutions to a linear ODE, then $y_1 + y_2$ is a solution, and any scalar multiple λy_1 or λy_2 is also a solution.

Now, although this may seem like abstract unnecessary knowledge that gives supplemental context, this understanding of the underlying linear algebra will play a pivotal role in our discussion of the Wronskian. Before moving forward, however, let's test our understanding with an example:

Example. Consider the differential equation

$$y' + xy = 0.$$

We'll verify that this is a linear differential equation by showing that if y_1, y_2 are solutions, then $y_1 + y_2$ is a solution, and that λy_1 is a solution.

Proof. Proof: If y_1, y_2 are solutions, then $y_1' + xy_1 = 0$ and $y_2' + xy_2 = 0$. We want to show that

$$(y_1 + y_2)' + x(y_1 + y_2) = 0.$$

To this end, we simply observe that

$$(y_1 + y_2)' + x(y_1 + y_2) = (y_1' + xy_1) + (y_2' + xy_2) = 0.$$

Now, if y is a solution, and $\lambda \in \mathbb{R}$ is a scalar, we want to show that λy is a solution, i.e., we want to show that $(\lambda y)' + x(\lambda y) = 0$. But this is clear, since

$$(\lambda y)' + x(\lambda y) = \lambda y' + \lambda xy = \lambda(y' + xy) = 0.$$

□

Now that we understand what it means for a differential equation to be linear, we want to understand how we can obtain all the solutions. This is one of the key advantages to a thorough understanding of linear algebra. Indeed, suppose we have a second-order linear differential equation, say,

$$y'' - y = 0.$$

Two solutions to this equation are given by $y_1(x) = e^x$ and $y_2(x) = e^{-x}$. It's natural to ask, are these all the solutions to our differential equation? Well, since the equation is linear, the set of solutions forms a vector space, whose dimension is at most the order of the differential equation, which in this case is two. It therefore suffices to verify that $y_1(x)$ and $y_2(x)$ form a basis, since then every solution to our ODE will just be given by

$$y(x) = \lambda_1 e^x + \lambda_2 e^{-x},$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ are scalars.

Moreover, since a set of n vectors in an n—dimensional vector space is a basis if and only if they are linearly independent, it suffices to check whether e^x and e^{-x} are linearly independent. Ok, so there's a lot of content here, so let's take a moment to recap on what we've said so far:

Reminder.

(†) An ODE is linear if the set of solutions to the ODE form a vector space (i.e., if we add two solutions, we get another solution, and if we scale a solution, we get another solution).

(ii) A set of vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ is a basis for V if and only if it is linearly independent and spans V. Note that if V is n—dimensional, $\mathbf{v}_1, ..., \mathbf{v}_n$ is a basis for V if and only if $\mathbf{v}_1, ..., \mathbf{v}_n$ are linearly independent.

So from this discussion so far, we have reduced the problem of finding all solutions, given that we know some number of solutions, to the problem of determining whether these solutions are linearly independent. This is where the Wronskian comes in.

2. Part 2 – The Wronskian

Definition. Suppose that y_1, y_2 are solutions to a linear differential equation. We define the **Wronskian** of this set $\{y_1, y_2\}$, by

$$\mathcal{W}(y_1, y_2) := \det \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} = y_1 y_2' - y_1' y_2.$$

Of course, we can define the Wronskian for a set of n functions $y_1, ..., y_n$which are (at least) $(n-1)$—times differentiable by setting

$$\mathcal{W}(y_1, ..., y_n) := \det \begin{bmatrix} y_1 & \cdots & y_n \\ y_1' & \cdots & y_n' \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix},$$

but for simplicity, we'll stick to the case where $n = 2$.

The main important result concerning the Wronskian is the following:

Key fact. A set of solutions $y_1, ..., y_n$ is linearly independent if $\mathcal{W}(y_1, ..., y_n) \neq 0$.

That is, a set of solutions to a linear ODE is linearly independent if the Wronskian is never zero. Let's look at an example:

Example. Compute the Wronskian of

$$y_1(x) = e^x, \qquad y_2(x) = e^{-x},$$

which are solutions of the second-order linear ODE $y'' - y = 0$.
Indeed, we see that $y_1'(x) = e^x$ and $y_2'(x) = -e^{-x}$. Therefore,

$$\mathcal{W}(y_1, y_2) = \det \begin{bmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{bmatrix} = e^x(-e^{-x}) - e^x(e^{-x}) = -e^{x-x} - e^{x-x} = -2 \neq 0,$$

so the solutions $y_1(x) = e^x$ and $y_2(x) = e^{-x}$ are linearly independent. Hence, all solutions of $y'' - y = 0$ are given by $y(x) = \lambda_1 e^x + \lambda_2 e^{-x}$, for scalars λ_1, λ_2.

One technicality to point out is that there is a small subtlety concerning the above key fact, which is the following:

Refined Key Fact. A set of solutions $y_1, ..., y_n$ is linearly independent if $\mathcal{W}(y_1, ..., y_n) \neq 0$ everywhere, except possibly one point.

Example. Consider the functions $y_1(x) = e^{-x}$ and $y_2(x) = xe^{-x}$, which are solutions of $y'' + 2y' + y = 0$. The Wronskian is given by

$$\mathcal{W}(y_1, y_2) = \det \begin{bmatrix} e^{-x} & xe^{-x} \\ -e^{-x} & e^{-x}(1-x) \end{bmatrix} = e^{-2x}(1-x) + xe^{-2x} = e^{-2x}.$$

3. Part 3 – A neat application of Wronskians

Question. Let $f_1, ..., f_n$ be a set of smooth functions. How do we construct a linear ODE for which $f_1, ..., f_n$ are the solutions?

To really appreciate what we're about to do, pause the video and try to do this yourself...
OK, let's see how Wronskians can give an effective tool for the construction of such a differential equation. If we take the case $n = 1$, then we want to construct a linear differential equation such that $y = f_1$ is the solution. If we compute the Wronskian of $\{y, f_1\}$, we have

$$\mathcal{W}(y, f_1) = \det \begin{bmatrix} y & f_1 \\ y' & f_1' \end{bmatrix} = yf_1' - y'f_1.$$

Hence, we can take

$$f_1 y' - f_1' y = 0,$$

or equivalently,

$$y' - \frac{f_1'}{f_1} y = 0$$

as the differential equation.

Variation of Parameters

To understand the method of variation of parameters, let's first recall what we do when solving a first-order linear differential equation of the form

$$\frac{dy}{dx} + p(x)y = f(x).$$

The differential equation has the property that the solution is given by $y = y_{\text{hom}} + y_{\text{par}}$, where y_{hom} is the solution to the homogeneous equation

$$\frac{dy}{dx} + p(x)y = 0$$

and y_p is the partial solution – the solution to the in-homogeneous equation above. Indeed,

$$\frac{d}{dx}(y_{\text{hom}} + y_{\text{par}}) + p(x)(y_{\text{hom}} + y_{\text{par}}) = \underbrace{\left[\frac{dy_{\text{hom}}}{dx} + p(x)y_{\text{hom}}\right]}_{=0} + \underbrace{\left[\frac{dy_{\text{par}}}{dx} + p(x)y_{\text{par}}\right]}_{=f(x)} = f(x).$$

Homogeneous equation. Let us now observe that the homogeneous equation

$$\frac{dy}{dx} + p(x)y = 0$$

is a separable equation, which we can write as

$$\frac{dy}{y} + p(x)dx = 0.$$

Integrating, we get

$$\begin{aligned}
\frac{dy}{y} + p(x)dx = 0 \quad &\implies \quad \int \frac{dy}{y} + \int p(x)dx = 0 \\
&\implies \quad \ln|y| + \int p(x)dx = 0 \\
&\implies \quad y_{\text{hom}} = A\exp\left(-\int p(x)dx\right)
\end{aligned}$$

Remark. This is why we use an integrating factor when solving first-order linear homogeneous differential equations.

Inhomogeneous equation. The particular solution is achieved by trying to find a function $u(x)$ such that $y_{\text{par}}(x) = u(x)y_{\text{hom}}(x)$.
Insert $y_{\text{par}} = u(x)y_{\text{hom}}$ into

$$\frac{dy}{dx} + p(x)y = f(x),$$

we get

$$\begin{aligned}
\frac{dy_{\text{par}}}{dx} + p(x)y_{\text{par}} = f(x) \quad &\Longrightarrow \quad \frac{du}{dx}y_{\text{hom}}(x) + u(x)\frac{dy_{\text{hom}}}{dx} + p(x)u(x)y_{\text{hom}} = f(x) \\
&\Longrightarrow \quad \frac{du}{dx}y_{\text{hom}}(x) + u(x)\underbrace{\left[\frac{dy_{\text{hom}}}{dx} + p(x)y_{\text{hom}}\right]}_{=0} = f(x) \\
&\Longrightarrow \quad \frac{du}{dx}y_{\text{hom}}(x) = f(x) \\
&\Longrightarrow \quad du = \frac{f(x)}{y_{\text{hom}}(x)}dx \\
&\Longrightarrow \quad u = \int \frac{f(x)}{y_{\text{hom}}(x)}dx
\end{aligned}$$

4. Variation of parameters – Second-Order Equations

We will now look at second-order inhomogeneous equations

$$y'' + p(x)y' + q(x)y = f(x).$$

The particular solution will be of the form

$$y_{\text{par}} = u_1(x)y_1(x) + u_2(x)y_2(x),$$

where y_1, y_2 are the fundamental solutions to the associated homogeneous equation. As before, we want to insert y_{par} into the ODE:

$$
\begin{aligned}
& y_{\text{par}}'' + p(x)y_{\text{par}}' + q(x)y_{\text{par}} = f(x) \\
\implies \quad & (u_1(x)y_1(x) + u_2(x)y_2(x))'' + p(x)(u_1(x)y_1(x) + u_2(x)y_2(x))' \\
& +q(x)(u_1(x)y_1(x) + u_2(x)y_2(x)) = f(x) \\
\implies \quad & (u_1'y_1 + u_1y_1' + u_2'y_2 + u_2y_2')' + p(x)(u_1'y_1 + u_1y_1' + u_2'y_2 + u_2y_2') \\
& +q(x)(u_1y_1 + u_2y_2) = f(x) \\
\implies \quad & u_1''y_1 + u_1'y_1' + u_1'y_1' + u_1y_1'' + u_2''y_2 + u_2'y_2' + u_2'y_2' + u_2y_2'' + p(x)(u_1'y_1 + u_1y_1' + u_2'y_2 + u_2y_2') \\
& +q(x)(u_1y_1 + u_2y_2) = f(x) \\
\implies \quad & u_1 \underbrace{\left[y_1'' + py_1' + qy_1\right]}_{=0} + u_2 \underbrace{\left[y_2'' + py_2' + qy_2\right]}_{=0} + y_1u_1'' + u_1'y_1' \\
& +y_2u_2'' + u_2'y_2' + p(y_1u_1' + y_2u_2') + y_1'u_1' + y_2'u_2' = f(x) \\
\implies \quad & \frac{d}{dx}(y_1u_1') + \frac{d}{dx}(y_2u_2') + p(y_1u_1' + y_2u_2') + y_1'u_1' + y_2'u_2' = f(x) \\
\implies \quad & \frac{d}{dx}(y_1u_1' + y_2u_2') + p(y_1u_1' + y_2u_2') + y_1'u_1' + y_2'u_2' = f(x).
\end{aligned}
$$

Because we seek to determine two unknown functions u_1 and u_2, we need two equations. We can obtain these equations by making the further assumption that the functions u_1 and u_2 satisfy $y_1u_1' + y_2u_2' = 0$. We make this assumption because it simplifies

$$\frac{d}{dx}(y_1u_1' + y_2u_2') + p(x)(y_1u_1' + y_2u_2') + y_1'u_1' + y_2'u_2' = f(x)$$

to

$$p(x)(y_1u_1' + y_2u_2') + y_1'u_1' + y_2'u_2' = f(x).$$

This gives two equations

$$\begin{aligned} y_1(x)u_1'(x) + y_2(x)u_2'(x) &= 0 \\ \text{and} \quad y_1'(x)u_1'(x) + y_2'(x)u_2'(x) &= f(x). \end{aligned}$$

By Cramer's rule (from linear algebra), we see that

$$u_1'(x) = -\frac{y_2(x)f(x)}{W(x)} \qquad \text{and} \qquad u_2'(x) = \frac{y_1(x)f(x)}{W(x)}.$$