Math 218 — Study guide for the final

Alex Cowan
alex.cowan@uwaterloo.ca
https://www.math.uwaterloo.ca/~a2cowan/

- WW: Wainwright–West
- BdP: Boyce–DiPrima 11^{th} edition
- A: Assignment and its solutions
- EP: Extra problems
- M: Midterm

1 First order differential equations

WW §1.

1.1 Separable equations

WW 1.2.3; A1Q1; EP1 1.2, 1.3 b)d), 1.10; M1Q1.

1.2 Exact equations

Linear first order ODEs. Integrating factors. Non-linear first order ODEs which can be recognized as the derivative of a product.

WW 1,2,4; BdP 2.1; A1Q2, A8Q2–3; EP1 1.1, 1.3 a)c), 1.10, 1.11.

1.3 Qualitative understanding

Limiting behaviour: EP1 1.5–1.7; A8Q4.6.

Direction fields: WW 1.2.5; BdP 1.1; A1Q3; EP1 1.4, 1.8, 3.8.

2 Initial conditions

First order: WW 1.2.8.

Second order: WW 2.1.4; A2Q2; EP1 2.4.

Systems: WW 4.1.3; A6Q2.

3 Constants of integration

Basic understanding of the following: Usually the number of free parameters is the order of the differential equation minus the number of conditions. When multiple variables are involved (e.g. $x' = Ax$), you get one free parameter per variable per order.

If there are discontinuities in the solution, the preceding statements are true for every continuous chunk of solution. Free parameters (and other choices like branch of square root) on different chunks are independent.

Ensuring that choice of constants of integration lead to solutions being defined on the required intervals $\longrightarrow$ sometimes they're not just arbitrary real numbers.

A1Q1 solution, A2Q2, A8Q1–2; EP1 1.13.

4 Decomposition into homogeneous and particular solutions

Second order: WW 2.2.1–2.2.2; A2Q3b); EP1 2.6, 2.8.

Systems: WW 4.1.3–4.1.4; BdP 7.4.

5 Second order linear differential equations with constant coefficients

$y'' + 2\lambda y' + \omega_0^2 y = 0$ etc. WW §2; BdP §3.

5.1 General solutions for homogeneous equations

Including repeated roots. See also Section 7.6.

WW 2.2.3; A2Q1; EP1 2.1, 2.3; M1Q2.2

5.2 Particular solutions in the case of exponential driving functions

WW 2.2.4; A2Q3, A3Q1; EP1 2.5.

Repeated roots: WW 2.2.4; BdP 3.4; EP1 2.7.

5.3 Damped simple harmonic oscillators

Homogeneous part: Undamped/Underdamped/Overdamped/Critically damped behaviour:
WW 2.3.2; BdP 3.7; A3Q2; EP1 2.2, 2.3, 2.9; M1Q2.1.

Particular part: Resonance frequency & amplitude:
WW 2.3.3; BdP 3.8; A3Q1, A4Q2; EP 2.9–2.13; M1Q2.3.

6 Euler's method

Conceptual understanding. Relationship to linear approximation. Code/pseudocode. Step-size tension between speed and precision. $< 10^6$ steps $\rightarrow$ Instantaneous, 10^6 – 10^9 steps $\rightarrow$ Noticeable but not problematic, 10^9 – 10^{12} steps $\rightarrow$ Realistic but needs consideration, $> 10^{12}$ steps $\rightarrow$ Not realistic in the context of this course. Checking $\frac{1}{2}y''(t)(\Delta t)^2$ term in series expansion to guess precision/step-size needed.

BdP 2.7, 8.1; A4, A5Q2.1–2.2, A9Q2; EP1 §3; Example code from lectures 2024/09/30 and 2024/10/02, Example code from tutorial 2024/10/02; M1Q3. Appendix B

7 Systems of first order linear differential equations with constant coefficients

$x' = Ax$ etc. WW §4; BdP §7.

7.1 Diagonalization

Computing eigenvalues and eigenvectors by hand given A. Writing $A = PDP^{-1}$ when possible. Interpretation of P and D. View as a change of coordinates to a basis on which A acts diagonally. $x(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots$ solves $x' = Ax$.

WW 4.2; BdP 7.5–7.6; A5, A6Q1-3; M2Q1.

7.2 The matrix exponential

Definition as a power series. $x' = Ax \iff x = \exp(tA)x(0)$. $\exp(tPDP^{-1}) = P\exp(tD)P^{-1}$. Proofs of these two facts. Conceptual understanding of the matrix exponential acting on initial conditions and evolving them through time. $\exp(tA)^{-1} = \exp(-tA)$. Understanding that determining $\exp(tA)$ tells you basically everything about the system.

WW 4.4.3–4.4.4; BdP 7.7; A5–6 solutions; EP2Q1–5; M2Q2;
3Blue1Brown https://www.youtube.com/watch?v=O85OWBJ2ayo especially 22:11 onwards.

7.3 Non-diagonalizable matrices

The Jordan normal form. Over $\mathbb{C}$, can't always write $A = PDP^{-1}$, but can always write $A = PJP^{-1}$. General solution for non-diagonalizable A: $x(t) = \exp(tA)x(0)$. Understanding that the Jordan normal form J generalizes diagonal matrices (i.e. diagonal matrices are a special case of matrices in Jordan normal form), as opposed to a different thing entirely.

A5Q3, A6Q1.3. (A more concrete method for solving such systems is in WW 4.2.4 and BdP 7.8. These could help you understand the material, but focus on conceptual understanding.)

7.4 General solution to homogeneous equations

WW 4.2; BdP 7.5–7.6; A5Q1, 2.3; Material from tutorial 2024/10/30; M2Q1.

7.5 Non-homogeneous equations

Variation of parameters. Derivation of the method of variation of parameters. (You are not responsible for methods relying on undetermined coefficients or Laplace transforms.)

WW 4.4.3; BdP 7.9; A6Q1, 2.2, 2.3; M2Q2; Appendix A.

7.6 Conversion to and from higher order linear differential equations with constant coefficients

WW 4.1.2; BdP 7.1.

7.7 Phase space

WW 4.3; BdP 7.5+; A6Q3, A8Q4; M2Q4.

8 The Laplace transform $\mathscr{L}$

WW §3; BdP §6.

8.1 Strategy for solving differential equations using the Laplace transform

WW "3.0", 3.2; BdP 6.2; A7Q2; M2Q3.

8.2 Definition and properties

Definition: WW 3.1.1–3.1.2; BdP 6.1; A6Q4.

Basic properties of the Laplace transform: WW 3.1; BdP 6.2+; A7Q1,3.

A sense of how to derive these basic properties, and how to compute Laplace transforms for given functions.

The convolution theorem: WW 3.3; BdP 6.6; A7Q2.4, A9Q3.3. (Not in great depth; basic awareness is sufficient.)

How to compute the Laplace transform of at least the exponential function (WW 3.1.1; BdP 6.1 Ex5; A6Q4) and the delta function (WW 3.5.2; BdP 6.5; A7Q2.3). A notion of how to determine the Laplace transform of related functions using the basic properties of the Laplace transform. (You will not need a table of Laplace transforms.)

8.3 Partial fraction decomposition

WW 3.1.4 and throughout; BdP 6.2 and throughout (ctrl+F "partial fraction"); A7Q2; M2Q3.

8.4 Discontinuous driving functions; the Heaviside function

WW 3.2.3; BdP 6.4; A7Q2.2.

8.5 The Dirac delta function δ

$\int_{-\infty}^{\infty} \delta(t-c)f(t)\,dt = f(c)$. Identity of the convolution operation. Perspective as a limit of $\frac{1}{2z}\mathbb{1}_{|x|<z}$ as $z \to 0^+$. Indefinite integral is the Heaviside function. "Lives under the integral sign".

WW 3.5.2, BdP 6.5, A7Q2.3.

9 Series solutions

Basic idea and simple examples only.

WW appendix A examples 1 and 2 only; BdP 5.2; A9Q1.

10 Translating word problems into differential equations

Cursory knowledge is sufficient, e.g. at the level of A8Q4, or the baking example with Newton's law of cooling and the Laplace transform done in lecture.

WW 1.1, 1.3.1, 1.3.6, 4.1.1; A8Q4.

11 Soft skills

Focusing on what you do know, not on what you don't. Trying to be distracted as little as possible by the sea of darkness representing all the things you don't know. The bright spot in the sea of darkness. Focusing on what's important. Being able to agree that a certain statement is true and continue on, even if you don't yet know why it's being brought up. Being literal and staying focused regarding questions, definitions, etc. Guess-and-check. Ansatz. Getting your hands dirty. Avoiding being standoffish and looking for a square peg to fit into a square hole. Being confident in your conclusions. Being okay with not having explicit expressions for functions. Piecing together information from different perspectives. Making use of other perspectives. Certain things must be true because of conceptual reasons, even when you don't have a calculation demonstrating that they're true. Not relying solely on algorithms nor endeavouring to be the world's worst computer. Using a conceptual understanding to guide you through long calculations. Formulas are where ideas go to die — Jim Loveys. Viewing math as a web affording many different approaches, as opposed to a pile of computational techniques. Having an overall story for the subject, and a sense of what is key and what are details. The distinction between "I don't understand" and "I don't know" regarding information you have never been told or have not been given the opportunity to process. Being transparent and unashamed when you don't know or don't understand something. Being able to function as much as possible whilst being terrified.

See e.g. A3Q2d), A4Q2, A7Q3, A9Q3; EP1 1.9–1.11, 2.1i), 2.6e), 3.8–3.9; EP2 §§2,3; M1Q4; M2Q2, 4. However these ideas are ubiquitous and impossible to cram, so this is more of a reminder than anything else.

A Variation of parameters derivation

I was asked a couple times for a derivation of the method of variation of parameters, so here is a writeup.

Let's say you have

$$x' = Ax + f(t). \tag{1}$$

The general solution to (1) will be the sum of the general homogeneous solution (i.e. the general solution to $x' = Ax$) and a particular solution. Finding the general homogeneous solution to (1) isn't what this exposition is about; I'll talk about finding a particular solution to (1).

Key idea
The key idea for finding a particular solution is to write

$$x(t) =: \exp(tA)w(t). \tag{2}$$

Here I'm writing $=:$ to mean that the equation (2) is what defines $w(t)$. This does not make any assumptions, because $\exp(tA)^{-1} = \exp(-tA)$ — i.e. the matrix inverse of $\exp(tA)$ is given by changing the t to $-t$ — and therefore the inverse of $\exp(tA)$ always exists, so this is just giving the name $w(t)$ to the quantity $\exp(-tA)x(t)$.

Substitution
Using (2), we have

$$x' = A\exp(tA)w + \exp(tA)w' \tag{3}$$

by the product rule. Referencing (2), we can recognize the term $A\exp(tA)w$ on the right hand side of (3) as Ax. Rewriting (3) with this observation that $A\exp(tA)w = Ax$ turns (3) into

$$x' = Ax + \exp(tA)w'. \tag{4}$$

Isolating particular solution
Now we compare (4) to (1):

$$x' = Ax + \exp(tA)w',$$
$$x' = Ax + f(t).$$

Thus, we'll get a particular solution when

$$\exp(tA)w' = f(t). \tag{5}$$

Solving for $w(t)$
We can solve for w' (I will use the fact mentioned above that $\exp(tA)^{-1} = \exp(-tA)$):

$$w' = \exp(-tA)f(t) \tag{6}$$

To get w from (6) we can integrate (which means integrating entry-by-entry). This can be an indefinite integral if you like, and the constant of integration doesn't matter, just like it didn't matter when doing integrating factors. The way I present things is instead to solve for w as

$$w(t) = \int_0^t \exp(-uA)f(u)\,du. \tag{7}$$

Here I'm taking the definite integral from 0 to t. Any lower endpoint would work. Picking one is like picking a constant of integration. I change the variable in the integral from t to u because something like $\int_0^t g(t)\,dt$, with a t both in the bounds of integration and as the variable being integrated over, doesn't make sense syntactically.

Obtaining $x_p(t)$
Once you've found $w(t)$ via (7), you can find your particular solution with (2):

$$x_p(t) = \exp(tA)\int_0^t \exp(-uA)f(u)\,du. \tag{8}$$

Simplification when $A = PJP^{-1}$
If you have a decomposition $A = PJP^{-1}$, then, by looking at the power series defining $\exp(tA)$, one finds that $\exp(tA) = P\exp(tJ)P^{-1}$. This is an important fact and calculation. In this case (8) becomes

$$x_p(t) = P\exp(tJ)P^{-1}\int_0^t P\exp(-uJ)P^{-1}f(u)\,du. \tag{9}$$

Integrating a vector or matrix means integrating entry by entry, which means that integration and multiplication by the constant matrix P commute, i.e.

$$\begin{aligned} &P\exp(tJ)P^{-1}\int_0^t P\exp(-uJ)P^{-1}f(u)\,du \\ &= P\exp(tJ)P^{-1}P\int_0^t \exp(-uJ)P^{-1}f(u)\,du \\ &= P\exp(tJ)\int_0^t \exp(-uJ)P^{-1}f(u)\,du. \end{aligned}$$

This makes the calculations easier. See assignment 6 solutions for some more detail here.

Summary
Summarizing, a particular solution to the differential equation

$$x' = Ax + f(t)$$

is given by

$$\begin{aligned} &\exp(tA) \int_0^t \exp(-uA) f(u)\, du \\ &= P \exp(tJ) \int_0^t \exp(-uJ) P^{-1} f(u)\, du. \end{aligned}$$

The key idea is writing

$$x(t) =: \exp(tA) w(t),$$

after which you can substitute into the differential equation and follow your nose.

B Euler's method step size

There is a central tension between speed and precision when approximating solutions to the differential equation $y' = F(t, y)$ numerically using Euler's method. The more steps you have to take, the longer it takes to compute a solution. Too many steps, and computing a solution is not realistic. The smaller your step size, the more steps you have to take. On the other side, if your step size is too big, then your approximation will be imprecise.

Speed
Generally speaking,

- Fewer than 10^6 steps is effectively instantaneous.
- Between 10^6 and 10^9 steps takes a noticeable but unproblematic length of time.
- Between 10^9 and 10^{12} steps is realistic, but takes long enough that it needs consideration.
- More than 10^{12} steps isn't realistic without access to serious computational resources.

```
sage: sumval = 0

sage: time for k in range(10**5): sumval += k
CPU times: user 35.6 ms, sys: 52 µs, total: 35.7 ms
Wall time: 34.9 ms

sage: time for k in range(10**6): sumval += k
CPU times: user 350 ms, sys: 556 µs, total: 351 ms
Wall time: 351 ms

sage: time for k in range(10**7): sumval += k
CPU times: user 2.61 s, sys: 702 µs, total: 2.61 s
Wall time: 2.62 s

sage: time for k in range(10**8): sumval += k
CPU times: user 22.5 s, sys: 8.37 ms, total: 22.5 s
Wall time: 22.5 s
```

It's very tempting to view 10^6, 10^7, 10^8, ... as more "arithmetic" (i.e. additively spaced, like 6-7-8) and less "geometric" (i.e. multiplicatively spaced, like 8-64-512) than they actually are. Count to 2. Now count to 20. These are very different! That's the difference between 10^7 steps and 10^8 steps in the example above. Be deliberate and mindful when deciding how many steps your code will run for.

Precision

Expanding y in a power series around t and evaluating at $t + \Delta t$ gives

$$y(t + \Delta t) = y(t) + y'(t)\Delta t + \tfrac{1}{2}y''(t)\Delta t^2 + \dots. \tag{10}$$

(Here I'm writing Δt^2 to mean $(\Delta t)^2$.) Euler's method truncates the RHS of (10) after the first two terms:

$$y(t + \Delta t) \approx y(t) + y'(t)\Delta t. \tag{11}$$

Let's imagine that the error incured by the approximation (11) is of size roughly equal to the first missing term, $\frac{1}{2}y''(t)\Delta t^2$. In some cases this is not true, but it usually is, and when it isn't the reasoning we'll explore can easily be modified.

How small must we take Δt to get a good approximation to $y(t)$ on an interval $t_0 < t < t_f$? To guess, I recommend the following rules of thumb:

1. $y''(t)\Delta t^2 \ll y'(t)\Delta t$ for all $t_0 < t < t_f$,
2. $y'(t)\Delta t \ll y(t)$ for all $t_0 < t < t_f$,
3. $\frac{t_f - t_0}{\Delta t} y''(t)\Delta t^2 \ll y(t_f)$ for all $t_0 < t < t_f$.

The symbol $\ll$ means "much less than", so like 100 times smaller or something. Absolute values are implied.

Rule 1

In Euler's method, we store a variable y, and update it by an amount $y'(t)\Delta t = F(t, y)\Delta t$ at each step. This is an application of the linear approximation (11). We've assumed that the first term missing from this approximation, $\frac{1}{2}y''(t)\Delta t^2$, is of roughly the same size as the total error incurred by this approximation during this update.

The first rule of thumb above, $y''(t)\Delta t^2 \ll y'(t)\Delta t$ for all $t_0 < t < t_f$, says that, for each step,

> The amount by which y is updated, $y'(t)\Delta t$, should be much larger than the error, $\frac{1}{2}y''(t)\Delta t^2$.

You can imagine that if the approximation (11) used in Euler's method leads you to add 0.001 to y, but an exact calculation would have you add 0.005, that's no good.

Rule 2

Colouring (10):

$$y(t + \Delta t) = y(t) + y'(t)\Delta t + \tfrac{1}{2}y''(t)\Delta t^2 + ...,$$

the first rule of thumb above says that the blue term should be much smaller than the green term. The second rule of thumb is saying that the green term should, in turn, be much smaller than the red term, i.e.

> The amount by which y is updated, $y'(t)\Delta t$, should be much smaller than y itself.

You can imagine that if y is 1, and then you take some small step and suddenly y is 5, that's no good.

Rule 3

The third rule of thumb above pertains to the accumulation of errors over the course of the computation. To get from t_0 to t_f by taking steps of size Δt requires $\frac{t_f - t_0}{\Delta t}$ steps in total. If each individual step incurs an error of size approximately $y''(t)\Delta t^2$, then the sum of these errors could be as large as

$$\sum_t y''(t)\Delta t^2 \asymp \text{Number of steps} \cdot \text{Error from each step}$$

$$\asymp \frac{t_f - t_0}{\Delta t} y''(t)\Delta t^2.$$

This is an error accumulated over the course of the entire computation, so it should be compared to the end result of the computation, $y(t_f)$. In other words, this rule is saying that

The sum of the errors from the individual steps should be small compared to your final answer $y(t_f)$.

A conservative ways to estimate the accumulated error is to take the largest single-step error and multiply by the number of steps you have. In other words, see how big the quantity $y''(t)\Delta t^2$ can get over the entire interval $t_0 < t < t_f$, and assume all step sizes will have errors that big, and with the same sign.

This is of course quite pessimistic, and it shows in practice; usually this rule of thumb will indicate that you should take a value of Δt much smaller than what the other two rules of thumb suggest. The pessimism manifests in two ways: assuming that all step sizes will have an error as big as the biggest, and assuming that all step sizes will have the same sign. The first of these is actually quite often pretty reasonable, but the second is usually pretty far off; generally speaking, if you add up N random numbers of size about 1 and with random signs, your sum will be of size roughly $\sqrt{N}$, not N. This is a consequence of the central limit theorem.

Rule 3 is one that you can skirt a bit in practice if you're finding that it's too uncompromising with the speed side of the fundamental tension (but don't neglect it when there's no reason to). In contrast, skirting rules 1 or 2 is playing with fire.