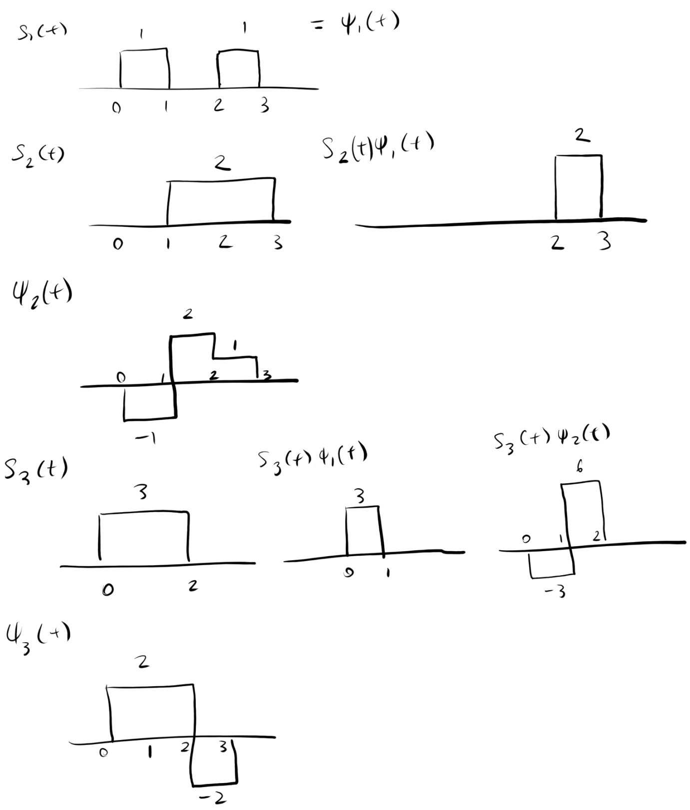
ECE211 – PSET2

Jonathan Lam

March 5, 2020

1. Define the inner product $\langle x, y \rangle \coloneqq \int_{-\infty}^{\infty} x(t)y(t)\,dt$ (area under the product of $x(t)$ and $y(t)$).

   (a) *Perform GSO on the three given signals.*
   Sketches: $\{s_i\}$ represent the original signals set; $\{\psi_i\}$ represent an orthogonal basis; and $\{\phi_i\}$ represents the orthonormal basis.



   Calculating orthogonal basis:

$$\psi_1 = s_1$$

$$||\psi_1||^2 = 1^2 + 1^2 = 2$$

$$\langle s_2, \psi_1 \rangle = 2 \cdot 1 = 2$$

$$\psi_2 = s_2 - \frac{\langle s_2, \psi_1 \rangle}{||\psi_1||^2}\psi_1 = s_2 - \frac{2}{2}\psi_1 = s_2 - \psi_1$$

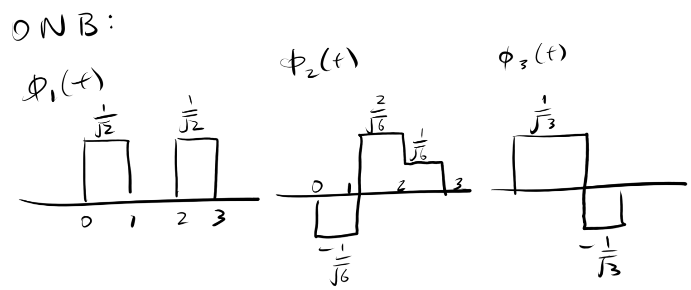
$$||\psi_2||^2 = (-1)^2 + 2^2 + 1^2 = 6$$

$$\langle s_3, \psi_1 \rangle = 3 \cdot 1 = 3$$

$$\langle s_3, \psi_2 \rangle = -3 \cdot 1 + 6 \cdot 1 = 3$$

$$\psi_3 = s_3 - \frac{\langle s_3, \psi_1 \rangle}{||\psi_1||^2}\psi_1 - \frac{\langle s_3, \psi_2 \rangle}{||\psi_2||^2} = s_3 - \frac{3}{2}\psi_1 - \frac{3}{6}\psi_2$$

Normalizing to get orthonormal basis:



$$||\psi_3||^2 = 2^2 + 2^2 + (-2)^2 = 12$$

$$\phi_1 = \frac{1}{||\psi_1||}\psi_1 = \frac{1}{\sqrt{2}}\psi_1$$

$$\phi_2 = \frac{1}{||\psi_2||}\psi_2 = \frac{1}{\sqrt{6}}\psi_2$$

$$\phi_3 = \frac{1}{||\psi_3||}\psi_3 = \frac{1}{\sqrt{12}}\psi_3$$

(b) **Express s_1, s_2, and s_3 w.r.t. the newly-calculated ONB.**

$$s_1 = \langle s_1, \phi_1 \rangle \phi_1 + \langle s_1, \phi_2 \rangle \phi_2 + \langle s_1, \phi_3 \rangle \phi_3 = \sqrt{2} \cdot \phi(1) + 0 \cdot \phi_2 + 0 \cdot \phi_3$$

$$s_2 = \langle s_2, \phi_1 \rangle \phi_1 + \langle s_2, \phi_2 \rangle \phi_2 + \langle s_2, \phi_3 \rangle \phi_3 = \sqrt{2} \cdot \phi(1) + \sqrt{6} \cdot \phi_2 + 0 \cdot \phi_3$$

$$s_3 = \langle s_3, \phi_1 \rangle \phi_1 + \langle s_3, \phi_2 \rangle \phi_2 + \langle s_3, \phi_3 \rangle \phi_3 = \frac{3}{\sqrt{2}} \cdot \phi(1) + \frac{3}{\sqrt{6}} \cdot \phi_2 + \frac{6}{\sqrt{3}} \cdot \phi_3$$

Alternatively, as tuples in $\mathbb{R}^3$:

$$\vec{s}_1 = (\sqrt{2}, 0, 0);\ \vec{s}_2 = (\sqrt{2}, \sqrt{6}, 0);\ \vec{s}_3 = \left(\frac{3}{\sqrt{2}}, \frac{3}{\sqrt{6}}, \frac{6}{\sqrt{3}}\right)$$

(c) *Express the projection of s_3 onto span$\{\phi_1, \phi_2\}$ as a vector in $\mathbb{R}^3$.*

$$\hat{s}_3 = \left(\frac{3}{\sqrt{2}}, \frac{3}{\sqrt{6}}, 0\right)$$

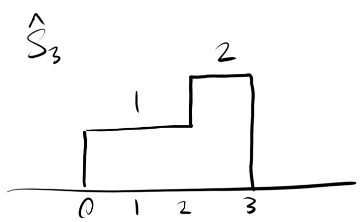
(d) *Compute $||s_3||$ using an integral and compare it to $||\vec{s}_3||$.*

$$||s_3|| = \sqrt{\int_{-\infty}^{\infty} s_3^2(t)\, dt} = \sqrt{3^2 + 3^2 + 0^2} = \sqrt{18}$$

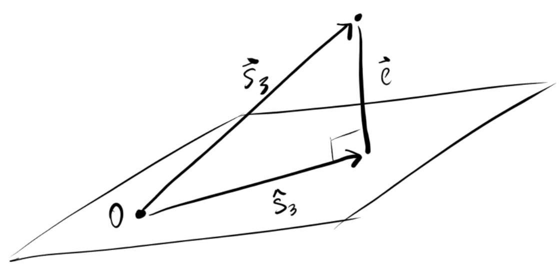
$$||\vec{s}_3|| = \sqrt{\left(\frac{3}{\sqrt{2}}\right)^2 + \left(\frac{3}{\sqrt{6}}\right)^2 + \left(\frac{6}{\sqrt{3}}\right)^2} = \sqrt{\frac{9}{2} + \frac{9}{6} + \frac{36}{3}} = \sqrt{18}$$

They match!

(e) *Sketch the approximate signal $\hat{s}_3$.*



(f) *Consider the error equation $e(t) = s_3(t) - \hat{s}_3(t)$. Write the equation that relates $||s_3||$, $||\hat{s}_3||$, and $||e||$ and use it to compute $||e||$.*
(Not really sketching e, just showing the general idea.)



Since $e(t) \perp \hat{s}_3(t)$, and $\hat{s}_3(t) + e(t) = s_3(t)$, the three magnitudes are related by the Pythagorean theorem:

$$||e|| = \sqrt{||s_3||^2 - ||\hat{s}_3||^2} = \sqrt{18 - \left(\left(\frac{3}{\sqrt{2}}\right)^2 + \left(\frac{3}{\sqrt{6}}\right)^2\right)} = \sqrt{18 - 6} = \sqrt{12}$$

2. A voltage amplifier provides a gain of 30dB.

(a) *Given $V_{in} = 1\,V$, find output amplitude.*

$$20 \log_{10} \frac{V_{out}}{1\text{V}} = 30$$

$$V_{out} = (1V)10^{30/20} = 31.6\text{V}$$

(b) *Given $P_{out} = 20dBm$, find input power.* Since the decibel scale is logarithmic, we may simply add the gain to an input voltage on an absolute decibel scale to get the output voltage on the same absolute decibel scale.

$$V_{in}(\text{dBm}) + V_{gain}(\text{dB}) = V_{out}(\text{dBm})$$

$$V_{in} = 20\text{dBm} - 30\text{dB} = -10\text{dBm}$$

(c) Express the input power from part (b) in Watts.

$$-10\text{dBm} = 10 \log_{10} \frac{P_{in}}{1 \times 10^{-3}\text{W}}$$

$$P_{in} = (1 \times 10^{-3}\text{W})10^{-10/10} = 1 \times 10^{-4}\text{W}$$

3. A real 40MHz sine wave is sampled at 100MHz.

(a) *List all frequencies up to 200MHz that would alias into 40MHz.*

40MHz, 60Mhz, 140MHz, 160MHz

(b) *Specify the appropriate cutoff frequency for the filter that should be employed by the 100MHz A/D converter.*

Cutoff frequency should be the Nyquist bandwidth: $f_s/2 = 100\text{MHz}/2 = 50\text{MHz}$ in order to avoid aliasing.

(c) *Specify the input signal frequencies in the different scales.*

analog radian frequency	$80\pi \times 10^6$ rad/s
normalized digital radian frequency	$4\pi/5$ rad
as a fraction of sampling rate	2/5
as a fraction of Nyquist bandwidth	4/5

(d) *The sampled data is passed to a D/A converter operated at a rate of 60kHz. List the six lowest frequencies that emerge at the output.*

24KHz, 36KHz, 84KHz, 96KHz, 144KHz, 156KHz

(e) *Specify the cutoff frequency of the anti-imaging filter that should be used in the D/A converter.*

Again, the cutoff frequency should be the Nyquist bandwidth: $f_{cutoff} = f_s/2 = 60\text{kHz}/2 = 30\text{kHz}$.

4. *Let $s(t)$ be a bandpass signal with reference carrier frequency f_0 and $s_{BB}(t)$. Derive the correct formula for the baseband equivalent for $s(t-t_0)$.*

$$s(t) = \Re\left(s_{BB}(t)e^{jw_0 t}\right)$$

$$s(t-t_0) = \Re\left(s_{BB}(t-t_0)e^{j2\pi f_0(t-t_0)}\right) = \Re\left(\left(e^{-j2\pi f_0 t_0}s_{BB}(t-t_0)\right)e^{2\pi f_0 t}\right)$$

Thus the baseband signal that corresponds with the shifted bandpass signal is multiplied by $\exp(-j\omega_0 t_0)$ (like shifting in the Laplace transform):

$$s'_{BB}(t) = e^{-j2\pi f_0 t_0}s_{BB}(t-t_0)$$

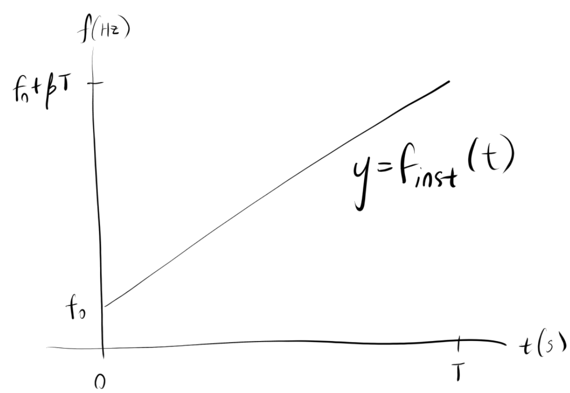
5. A chirp signal is given by

$$s(t) = A\cos(2\pi f_0 t + \pi\beta t^2 + \phi_0),\ 0 \le t \le T$$

(a) *Compute and sketch the instantaneous frequency $f_{inst}(t),\ 0 \le t \le T$.*

$$f_{inst}(t) = \frac{1}{2\pi}\frac{d\theta}{dt} = f_0 + \beta t$$

(i.e., frequency is f_0 at time $t = 0$, and increases as time increases with slope β.)



(b) *What would an appropriate unit for β be?*
Since β is the slope in the instantaneous-frequency/time plot, it should be Hz/s (rate of change of frequency).

(c) *Write the baseband equivalent for the chirp.*

$$s(t) = A\cos(2\pi f_0 t + \pi\beta t^2 + \phi_0) = \Re\left(Ae^{j(2\pi f_0 t + \pi\beta t^2 + \phi_0)}\right)$$

$$= \Re\left(Ae^{j(\pi\beta t^2 + \phi_0)}e^{j2\pi f_0 t}\right) \Rightarrow s_{BB}(t) = Ae^{j(\pi\beta t^2 + \phi_0)}$$

(d) *Suppose we receive a delayed version of the chirp: $r(t) = s(t - t_0)$. Write the baseband equivalent.*

$$r_{BB}(t) = e^{-j2\pi f_0 t_0} s_{BB}(t - t_0)$$

(e) *Given $r(t)$ at the receiver the signal is brought to baseband (i.e., $r_{BB}(t)$ is obtained) and then the following operation is performed at baseband:*

$$x_{BB}(t) = r_{BB}(t) s^*_{BB}(t)$$

Obtain the expression $x_{BB}(t)$. Which is the best way to recover the delay t_0?

$$x_{BB}(t) = \left(e^{-j2\pi f_0 t_0} \cdot Ae^{j(\pi\beta(t-t_0)^2+\phi_0)}\right)\left(Ae^{-j(\pi\beta t^2+\phi_0)}\right)$$

$$= A^2 e^{j(\pi\beta t_0^2 - 2\pi(f_0+t\beta)t_0)}$$

This is a complex sine wave with phase

$$\theta(t) = 2\pi(-\beta t_0)t + (\pi\beta t_0^2 - 2\pi f_0 t_0) = 2\pi f_0' t + \phi_0'$$

Thus the resulting wave is a sine wave with constant frequency f_0' and constant phase ϕ_0'. Since amplitude is nonchanging and there is no way to measure phase, we measure the frequency and solve for t_0:

$$-\beta t_0 = f_0' \Rightarrow t_0 = -\frac{f_0'}{\beta}$$

6. In communications applications, the target bandpass signal is called an RF signal, and uses some intermediate frequency signal IF. Let $RF = 1.0$GHz, $IF = 70$MHz.

(a) *What are the two possible choices for f_0?*
930.0MHz, 1070.0MHz

(b) *For each choice of f_0, what other frequencies will emerge when the 1.0GHz signal is mixed with f_0?*
For 930.0MHz: 1930MHz; for 1070.0MHz: 2070MHz

(c) *Find the image frequency of the lower choice of f_0.*
For 930.0MHz: 860MHz

(d) *Find the image frequency of the higher choice of f_0.*
For 1070.0MHz: 1140MHz

(e) *Does the (absolute value of the) difference between the RF and image frequencies depend on the choice of f_0 for a given f_{IR}? Does it depend on the value of f_{IF}? Can you surmise the formula for $|f_{RF} - f_{image}|$?*

i. No; the frequencies f_0 are $f_{RF} \pm f_{IF}$ and thus the image frequencies are $f_{RF} \pm 2f_{IF}$, both lying $2f_{IF}$ away from f_{RF}, and thus independent of the choice of f_0.

ii. Yes, as per the explanation in (i).

iii. $|f_{RF} - f_{image}| = 2f_{IF}$, as per the explanation in (i).

ECE211 – PSET 3

Jonathan Lam

March 5, 2020

1. Convolution.

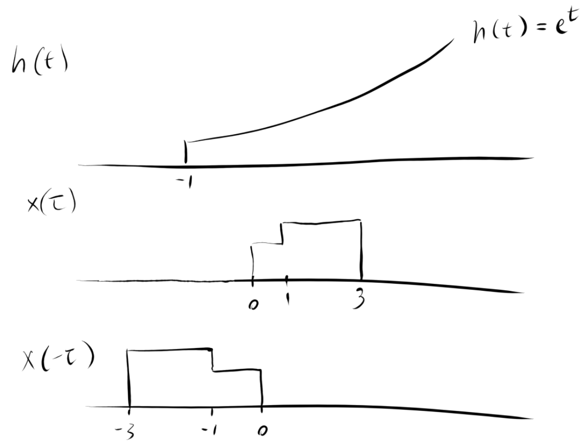
   Let

$$h(t) = e^t u(t+1)$$

$$x(t) = \begin{cases} 2 & 0 \le t \le 1 \\ 3 & 1 \le t \le 3 \\ 0 & \text{else} \end{cases}$$

   and

$$y = h * x$$



   There are four areas with different formulas here due to the piecewise nature of both functions

   (a) The first region is when $t \ge 2$, and the entire mirrored x region overlaps with the exponential where it is positive ($t \ge 1$). The integral with the proper bounds is

$$y(t) = \int_{t-3}^{t-1} 3e^{\tau}\,d\tau + \int_{t-1}^{t} 2e^{\tau}\,d\tau = 2e^t + e^{t-1} - 3e^{t-3},\ t \ge 2$$

(b) The second region is when $t \geq 2$, when the entire region where $x = 2$ overlaps with the positive exponential, but the overlap between the region where $x = 3$ and the exponential is only in the region $-1 \leq t \leq t - 1$. Thus the convolution here is

$$y(t) = \int_{-1}^{t-1} 3e^{\tau}\, d\tau + \int_{t-1}^{t} 2e^{\tau}\, d\tau = 2e^{t} + e^{t-1} - 3e^{-1},\ 0 \leq t \leq 2$$

(c) The third region is when none of the region where $x = 3$ overlaps with the positive exponential, and some of the region where $x = 2$ overlaps with the positive exponential.

$$y(t) = \int_{-1}^{t} 2e^{\tau}\, d\tau = 2(e^{t} - e^{-1}),\ -1 \leq t \leq 0$$

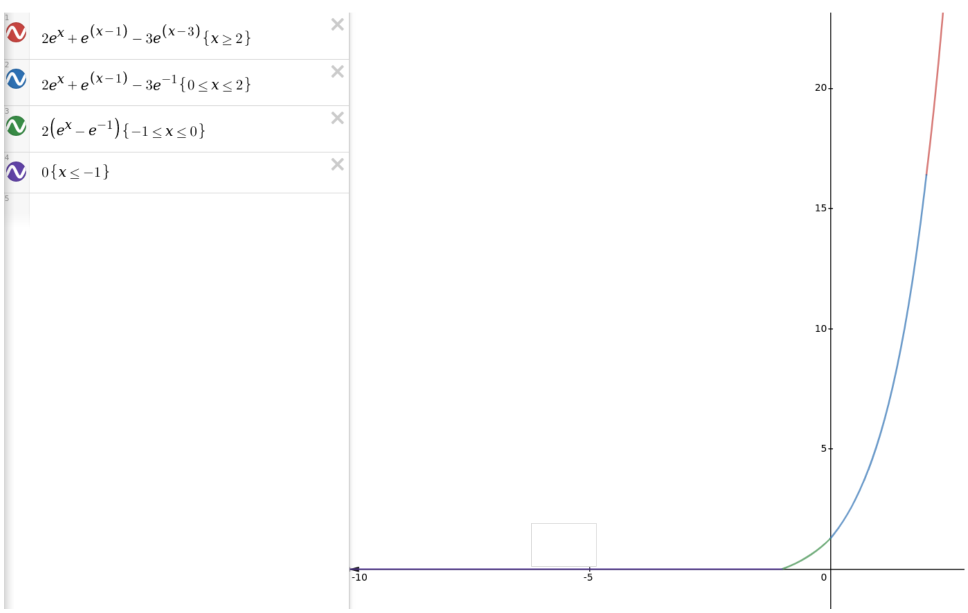
(d) The last region is where there is no overlap. Thus the convolution is zero.

$$y(t) = 0,\ t \leq -1$$

Then

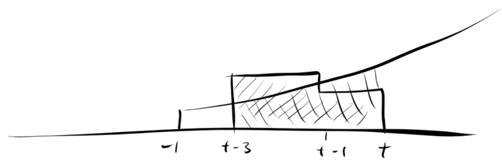
$$y(t) = \begin{cases} 2e^{t} + e^{t-1} - 3e^{t-3}, & t > 2 \\ 2e^{t} + e^{t-1} - 3e^{-1}, & 0 < t \leq 2 \\ 2(e^{t} - e^{-1}), & -1 < t \leq 0 \\ 0, & t \leq -1 \end{cases}$$

The bounds here are chosen arbitrarily, since $y(x)$ is a continuous function, as can be clearly seen in the following figure.



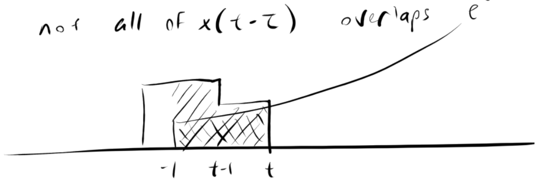
Case 1): $t \geq 2$

all of $x(t-\tau)$ overlaps with e^t



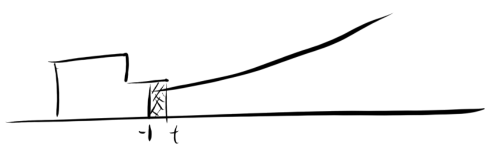
Case 2): $0 \leq t \leq 2$

not all of $x(t-\tau)$ overlaps e^t



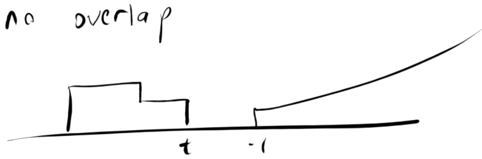
Case 3): $-1 \leq t \leq 0$

not all of $x(t-\tau)$ overlaps e^t



Case 4): $t \leq -1$

no overlap



2. Let $h = \{5, \underline{3}, 2, -2, -3\}$, $x = \{2, 3, \underline{4}, 5, 6\}$, $y = h * x$.

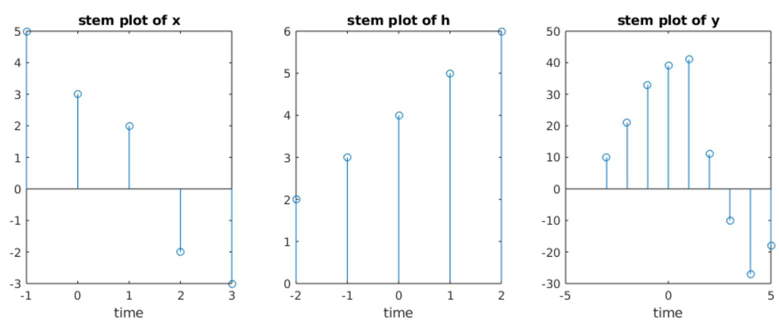
(a)

	Length	Support
x	5	$[-2, 2]$
h	5	$[-1, 3]$
y	9	$[-3, 5]$

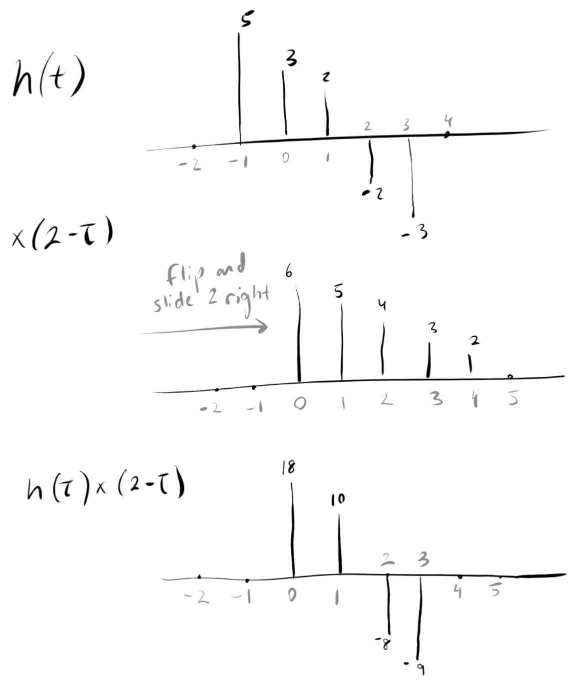
(b)

```
% generate signals
x = [5 3 2 -2 -3];
x_start = -1;
h = [2 3 4 5 6];
h_start = -2;
```

```
y = conv(h, x);
y_start = x_start + h_start;

% finite discrete stem
fds = @(s, s_start)
        stem((0:length(s)-1)+s_start, s);

% plot signals
subplot(1,3,1);
fds(x, x_start);
title('stem plot of x');
xlabel('time');

subplot(1,3,2);
fds(h, h_start);
title('stem plot of h');
xlabel('time');

subplot(1,3,3);
fds(y, y_start);
title('stem plot of y');
xlabel('time');
```

(c) 

(d) Flip 'n slide:



$$y[2] = \cdots + h[-3]x[5] + h[-2]x[4] + h[-1]x[3] + h[0]x[2] + h[1]x[1]$$

$$+h[2]x[0] + h[3]x[-1] + h[4]x[-2] + h[5]x[-3] \dots$$

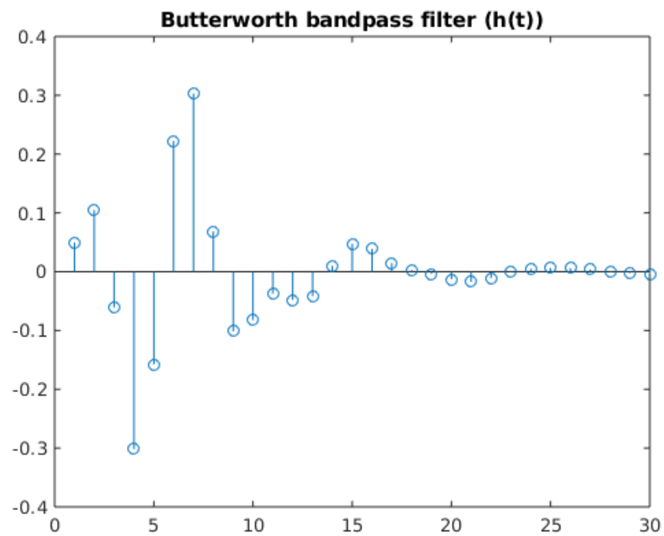
$$= \cdots + 0 \cdot 0 + 5 \cdot 0 + 3 \cdot 6 + 2 \cdot 5 + (-2) \cdot 4 + (-3) \cdot 3 + 2 \cdot 0 + 0 \cdot 0 + \cdots = 11$$

3. The Butterworth bandpass filter

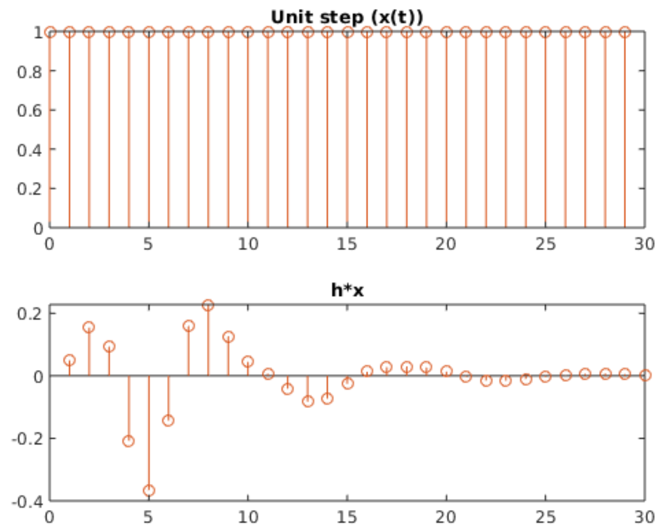
```
[b, a] = butter(3, [0.2 0.5]);
[h, t] = impz(b, a, 30);

figure;
stem(h);
title('Butterworth bandpass filter (h(t))');

figure;
x = ones(30);
subplot(2,1,1);
stem(t, x);
title('Unit step (x(t))');
subplot(2,1,2);
stem(filter(b, a, x));
title('h*x');

figure;
```

```
t = 0:30;
x = cos(0.4*pi*t);
subplot(2,1,1);
stem(t, x);
title('cos(4*pi*t) (x(t))');
subplot(2,1,2);
stem(filter(b, a, x));
title('h*x');
```

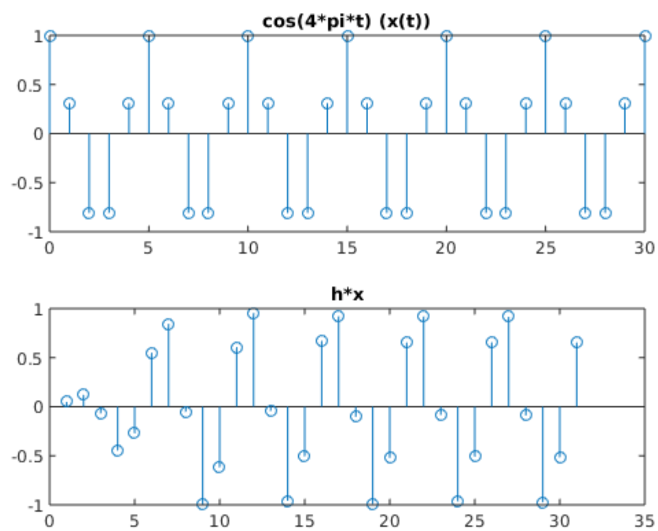
(a) Butterworth response with passband from $0.2 \leq \omega_{pass} \leq 0.5\,\text{rad}$



(b) Intially, the output looks like the impulse response $h(t)$, since giving it a starting value of 1 acts like an impulse. However, since $\omega = 0\text{rad}$ is out of the passband, the output quickly dies out to almost no amplitude.



(c) Since $\omega = 0.2\text{rad}$ is in the passband, the signal is not noticeably attenuated at all and matches the input at all of the sample points.



PSET 4: Basic z-Transform

Jonathan Lam

March 12, 2020

1. *Suppose $H(z)$ is such that it has a zero of multiplicity $m \geq 1$ at ∞. Then it has an inverse transform such that $h[N] \neq 0$ but $h[n] = 0\ \forall n \leq N-1$. Find N.*

   This means $H(z)$ has a single causal region. We focus on this causal region. Multiplying it by z will decrease the multiplicity of the zero at ∞ and shift the corresponding impulse response (its inverse transform) $h[n]$ left once. When the multiplicity of the zero at ∞ becomes negative (i.e., when there is a pole at ∞), then there will be no causal region of $H(z)$ anymore.

   The multiplicity of the zero at ∞ becomes zero when multiplying by z^m, which corresponds to $h[n+m]$. Since $h[n+m]$ is a causal signal and left-shifting it any more would make it non-causal by the above argument, this means that $h[m] \neq 0$ and $h[n] = 0\ \forall n \leq m-1$. Thus $N = m$.

   If we take $m = 0$, that means H has a finite nonzero limit at $z \to \infty$. In this case, what can we conclude about the causal inverse transform $h[n]$?

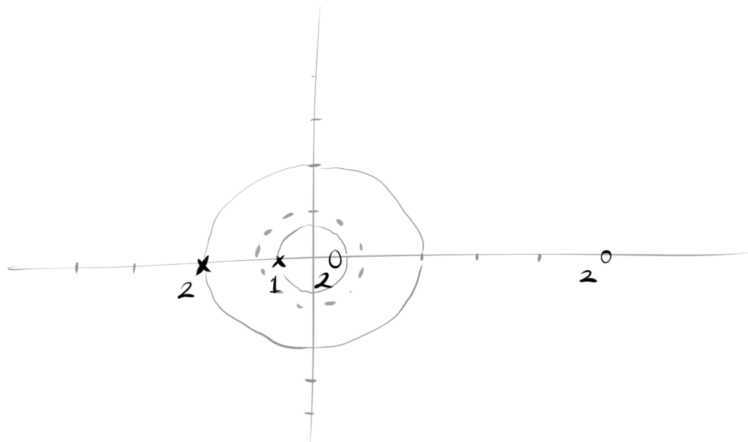
   With $N = m = 0$, $h[0] \neq 0$, by the above argument.

2. *Given:*

$$H(z) = \frac{3(2z-1)^2(z-5)^2}{(3z+2)(z+2)^2}$$

   (a) *List all poles and zeros, with multiplicities, and draw a pole-zero plot by hand.*
   Poles: -2/3 (m=1), -2 (m=2), ∞ (m=1)
   Zeros: 1/2 (m=2), 5 (m=2)

Figure 1: Pole-zero plot



(b) *Identify all possible ROCs for $H(z)$.*
$|z| < 2/3$, $2/3 < |z| < 2$, $2 < |z|$

(c) *Identify the ROC (if any) associated with:*

- *a causal system*: none; pole at ∞
- *a stable system*: $2/3 < |z| < 2$
- *a system with a well-defined frequency response*: $2/3 < |z| < 2$

(d) *We seek an integer L s.t. $z^L H(z)$ has a causal inverse transform such that $h[0] \neq 0$. Find all possible L.*

This uses much of the same logic as question (1). For there to be a causal inverse transform, there must not be a pole at ∞, so $L < 0$. The operation $z^L H(z)$ implies a left-shift of L on $h[n]$. Since $L < 0$ by the above statement, this means a right-shift by $-L$. Thus, specifying the condition $h[0] \neq 0$ puts a tight bound on L: we must shift $h[z]$ exactly enough units until it is causal, because only then will $h[0] \neq 0$ also be fulfilled.

By this reasoning, L must be -1, since then $z^{-1}H(z)$ is the smallest right-shift that can occur to create a causal region.

3. *A digital FIR filter with input x and output y is given by:*

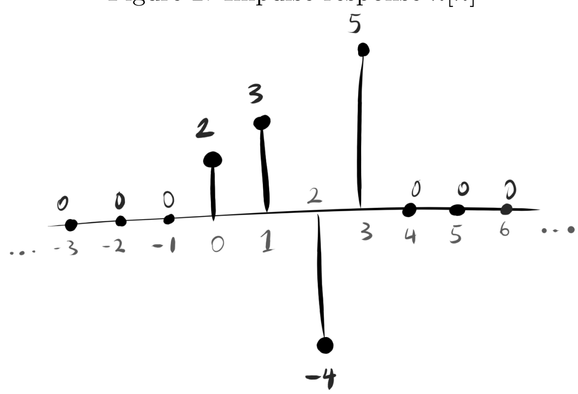
$$y[n] = 2x[n] + 3x[n-1] - 4x[n-2] + 5x[n-3]$$

(*a*) *Sketch the impulse response $h[n]$.* (See figure at top of next page.)

(*b*) *Express $h[n]$ as a superposition of impulses.*

$$h[n] = y[n](\delta) = 2\delta[n] + 3\delta[n-1] - 4\delta[n-2] + 5\delta[n-3]$$

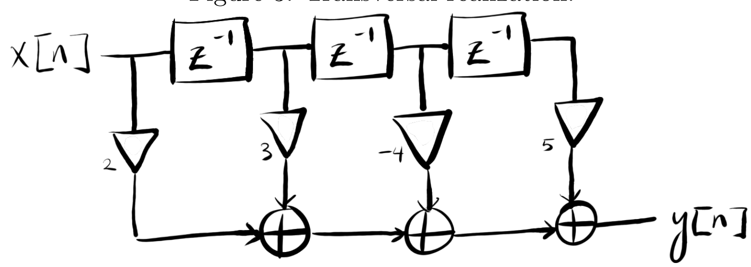
Figure 2: Impulse response $h[n]$



(c) *Specify the length and order of the filter.*
Length: 4 (number of terms)
Order: 3 (highest delay)

(d) *Sketch a transversal filter realization.*

Figure 3: Transversal realization.



4. *Given the following transfer function of a digital IIR filter:*

$$H(z) = \frac{3z^2 + 4z + 5}{10z^2 - z + 2}$$

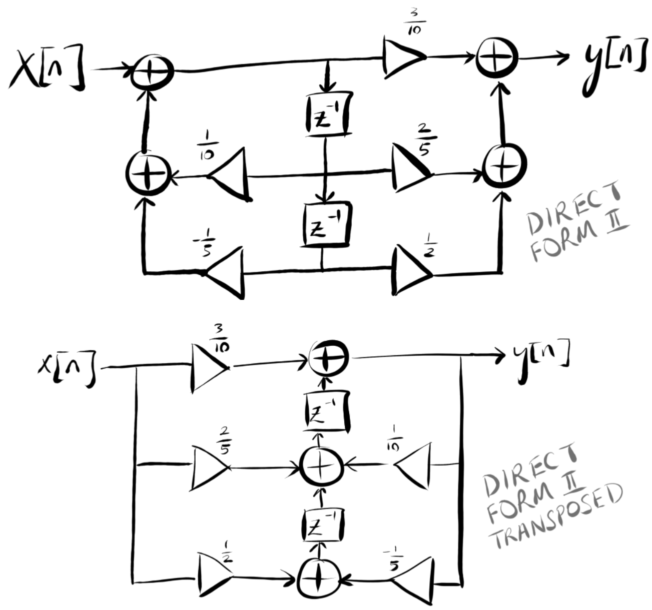
(a) *Write a difference equation with input x and output y.*
Rewriting in nicer form:

$$H(z) = \frac{\frac{3}{10}z^0 + \frac{2}{5}z^{-1} + \frac{1}{2}z^{-2}}{1 - \left(\frac{1}{10}z^{-1} - \frac{1}{5}z^{-2}\right)}$$

$$y[n] = \frac{3}{10}x[n] + \frac{2}{5}x[n-1] + \frac{1}{2}x[n-2] + \frac{1}{10}y[n-1] - \frac{1}{5}y[n-2]$$

(b) *Sketch direct form II and direct form II transposed realizations.*

Figure 4: Direct form II and direct form II transposed realizations.



5. *The following figure shows a block diagram. Find the overall transfer function by inspection.*

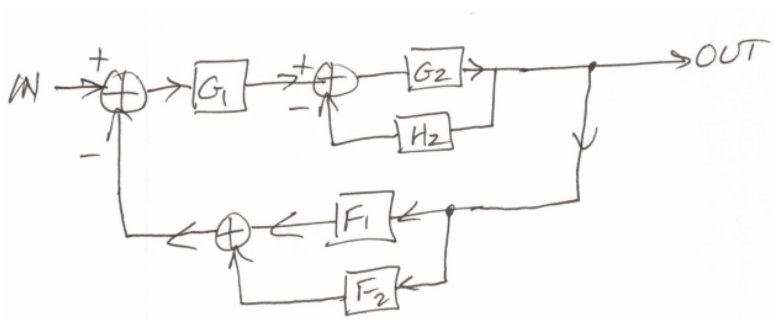


Figure 5: Block diagram of complex transfer function.

$$H = \frac{G_1\left(\frac{G_2}{1+G_2H_2}\right)}{1 + G_1\left(\frac{G_2}{1+G_1H_2}\right)(F_1+F_2)}$$

ECE211 – Pset 5

Jonathan Lam

March 31, 2020

1. *The spectrum $X(\omega)$ of a discrete time signal is shown below.*

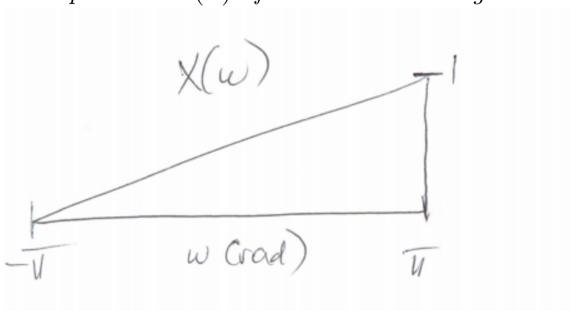


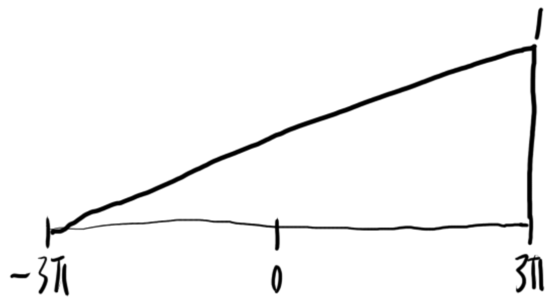
FIGURE 1: A DIGITAL SPECTRUM

(a) *Sketch $X(3\omega)$ on the range $-\pi < \omega < \pi$. Briefly state why this is a valid spectrum for some discrete-time signal.*

This spectrum is valid because it is contained within a 2π interval; that means that there are no "disagreements" for any frequencies that are $2\pi k$ apart. In other words, we are only looking at one branch of $e^{i\theta}$, so this is fine.

(b) *Sketch $X(4\omega)$ on the range $-\pi < \omega < \pi$.*

(c) *Sketch $X(\omega/3)$ on the range $-3\pi < \omega < 3\pi$. Briefly state why this is not a valid spectrum for some discrete-time signal.*



The spectrum of a signal must be 2π-periodic, since $e^{i\theta}$ is 2π-periodic; thus this spectrum cannot belong to any signal. (E.g., the frequencies at 0 and 2π should match if it were a valid spectrum.)

(d) *Let $Y(\omega) = X(\omega/3)$. Note that $X(\omega/3 - 2\pi k/3)$ ($k \in \mathbb{Z}$) is not obtained by shifting $Y(\omega)$ by $2\pi k/3$. How much is Y shifted by?*

By the definition of Y:

$$X\left(\frac{\omega - 2\pi k}{3}\right) = Y\left(3\left(\frac{\omega - 2\pi k}{3}\right)\right) = Y(\omega - 2\pi k)$$

Thus Y is shifted by $2\pi k$.

2. *Let $y[n] = x[Mn]$. This is called decimation by M. Here $M \geq 2 \in \mathbb{Z}$.. It turns out that*

$$Y(\omega) = \frac{1}{M}\sum_{k=0}^{M-1} X\left(\frac{\omega - 2\pi k}{M}\right)$$

Confirm that this produces a valid DTFT for some discrete-time signal.

If X is a valid spectrum, then it is 2π-periodic. Thus $X(\omega/M - 2\pi k/M)$ is $2\pi/M$-periodic, which is also 2π-periodic. Thus Y is the sum of M 2π-periodic spectra and scaling by $1/M$, and thus it should also be 2π-periodic and thus a valid spectrum for some signal.

3. *Let $x[n]$ have DTFT $X(\omega)$. Find the Fourier transform of $x^*[n_0 - n]$.*

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

$$\text{DTFT}\left[x^*[n_0 - n]\right] = \sum_{n=-\infty}^{\infty} x^*[n_0 - n] e^{-j\omega n} = \cdots$$

Let $n' = n_0 - n$ ($n = n_0 - n'$). We have a one-to-one correspondence between n and n', so we can make a change of indices.

$$\cdots = \sum_{n=n_0-n'=-\infty}^{\infty} x^*[n'] e^{-j\omega(n_0-n')} = e^{-j\omega n_0} \left[\sum_{n'=-\infty}^{\infty} x^*[n'] \left(e^{-j\omega n'}\right)^* \right]$$

$$= e^{-j\omega n_0} \left[\sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \right]^* = e^{-j\omega n_0} X^*(\omega)$$

(In other words, this is a reversal, shift, and conjugation in time, resulting in conjugation and multiplication by a constant in the transform domain.)

4. *Let $x(t)$ have CTFT $X(\omega)$. Find the inverse Fourier transform of $X^*(\omega_0 - \omega)$.*

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{j\omega t}\, d\omega$$

$$\text{CTFT}^{-1}\left[X^*(\omega_0 - \omega)\right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} X^*(\omega_0 - \omega) e^{j\omega t}\, d\omega = \cdots$$

Make the substitution $\omega' = \omega_0 - \omega$ ($\omega = \omega_0 - \omega'$, $d\omega = -d\omega'$). We have a one-to-one correspondence, so we can make a change of variable from ω to ω'.

$$\cdots = \frac{1}{2\pi} \int_{\infty}^{-\infty} X^*(\omega') e^{j(\omega_0-\omega')t} (-1) d\omega' = e^{j\omega_0 t} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} X^*(\omega') \left(e^{j\omega' t}\right)^* d\omega' \right]$$

$$= e^{j\omega_0 t} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{j\omega t}\, d\omega \right]^* = e^{j\omega_0 t} x^*(t)$$

(In other words, a shift, reversal, and conjugation in the transform domain will result in a conjugation and multiplication in the time domain.)

5. *Define the paraconjugate to be the following in the transform domain:*

$$\tilde{H}(z) = H^*(1/z^*)$$

$$\tilde{H}(s) = H^*(-s^*)$$

(a) *Check that each paraconjugate formula reduces to $H^*(\omega)$ in the frequency domain.*

Discrete-time case:

$$H^*(z) = \sum_{n=-\infty}^{\infty} h^*[n]z^n$$

$$H^*(1/z^*) = \sum_{n=-\infty}^{\infty} h^*[n]\left((z^*)^{-1}\right)^n = \sum_{n=-\infty}^{\infty} h^*[n]\,(z^*)^{-n}$$

The frequency domain is the circle $z = e^{j\omega}$.

$$H^*(1/z^*) = \sum_{n=-\infty}^{\infty} h^*[n]\left(e^{-j\omega}\right)^{-n} = \left(\sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n}\right)^* = H^*(\omega)$$

Continuous-time case:

$$H^*(-s^*) = \int_{-\infty}^{\infty} h^*(t)e^{(-s^*)t}\,dt$$

The frequency domain is the line $s = j\omega$.

$$H^*(-s^*) = \int_{-\infty}^{\infty} h^*(t)e^{(-j\omega)^*t}\,dt = \int_{-\infty}^{\infty} h^*(t)e^{j\omega t}\,dt$$

$$= \left(\int_{-\infty}^{\infty} h(t)e^{-j\omega t}\right)^* = H^*(\omega)$$

(b) *Let $H(s)$, $H(z)$ be second-order systems. Determine $\tilde{H}$ in each case.*

Discrete-time case:

$$H(z) = \frac{b_2z^2 + b_1z + b_0}{a_2z^2 + a_1z + a_0}$$

$$\tilde{H}(z) = H^*(1/z^*) = \left(\frac{b_2(1/z^*)^2 + b_1(1/z^*) + b_0}{a_2(1/z^*)^2 + a_1(1/z^*) + a_0} \times \left(\frac{z^*}{z^*}\right)^2\right)^*$$

$$= \frac{(b_2 + b_1(z^*) + b_0(z^*)^2)^*}{(a_2 + a_1(z^*) + a_0(z^*)^2)^*} = \frac{b_2^* + b_1^*z + b_0^*z^2}{a_2^* + a_1^*z + a_2^*z^2}$$

Continuous-time case:

$$H(s) = \frac{b_2s^2 + b_1s + b_0}{a_2s^2 + a_1s + a_0}$$

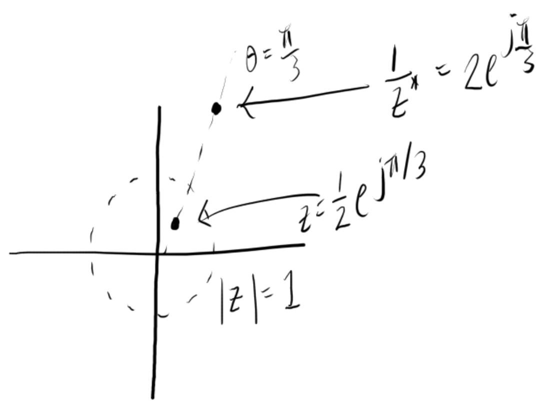
$$\tilde{H}(s) = H^*(-s^*) = \left(\frac{b_2(-s^*)^2 + b_1(-s^*) + b_0}{a_2(-s^*)^2 + a_1(-s^*) + a_0}\right)^*$$

$$= \frac{(b_2(s^*)^2 - b_1(s^*) + b_0)^*}{(a_2(s^*)^2 - a_1(s^*) + a_0)^*} = \frac{b_2^*s^2 - b_1^*s + b_0^*}{a_2^*s^2 - a_1^*s + a_0^*}$$

(c) *z and $1/z^*$ are called symmetric points w.r.t. the unit circle, and s and $-s^*$ are called symmetric points w.r.t. the $j\omega$-axis.*

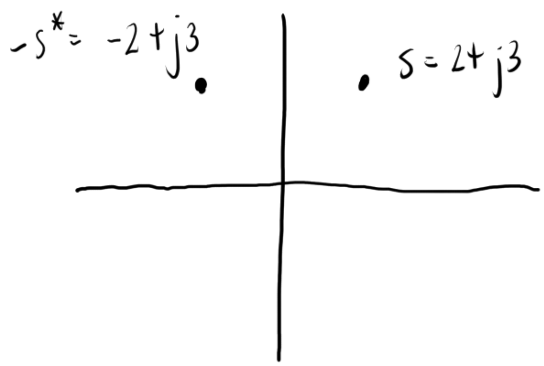
i. *If $z = r\exp(j\theta)$, write $1/z^*$ in polar form.*

$$\frac{1}{z^*} = \frac{1}{r\exp(-j\theta)} = \frac{1}{r}\exp(j\theta)$$



ii. *If $s = \sigma + j\omega$, write $-s^*$ in rectangular form.*

$$-s^* = -(\sigma - j\omega) = -\sigma + j\omega$$



(d) *An all-pass system has constant magnitude 1 at all frequencies. This can be expressed as $|A(\omega)| = 1$. We can also express this as $|A(\omega)|^2 = A(\omega)A^*(\omega) = 1$. In other words, an all-pass function has the following property in the transform domain:*

$$A\tilde{A} = 1$$

In general, the symmetric point to a pole will be a zero, and the most general form for a digital all-pass function is

$$A(z) = z^L \prod_{i=1}^{M} \frac{\alpha_i^* z - 1}{z - \alpha_i}$$

where $|\alpha_i| \neq 0$ *and* $L \in \mathbb{Z}$.

i. *Plug in* $z = \exp(j\omega)$ *and confirm* $A(z)$ *is an all-pass function.*

Since $|z| = 1$, $z = 1/z^*$. Thus $\tilde{A}(z) = A^*(1/z^*) = A^*(z)$, and $A\tilde{A} = AA^* = |A|^2$. Thus we have to show that $|A|^2 = 1$.

$$|A(z)| = \left| \left(e^{j\omega}\right)^L \prod_{i=1}^{M} \frac{\alpha_i^* e^{j\omega} - 1}{e^{j\omega} - \alpha_i} \right| = \left|e^{j\omega L}\right| \prod_{i=1}^{M} \left| \frac{\alpha_i^* e^{j\omega} - 1}{e^{j\omega} - \alpha_i} \right|$$

It is easy to see that $\left|e^{j\omega L}\right| = \sqrt{e^{j\omega L} e^{-j\omega L}} = 1$. The multiplicand (the magnitude inside the product operator) also has value 1:

$$\left| \frac{\alpha_i^* e^{j\omega} - 1}{e^{j\omega} - \alpha_i} \right| = \left| \frac{-e^{j\omega}\left(e^{-j\omega} - \alpha_i^*\right)}{e^{j\omega} - \alpha_i} \right| = \left|-e^{j\omega}\right| \left| \frac{\left(e^{j\omega} - \alpha_i\right)^*}{e^{j\omega} - \alpha_i} \right|$$

Since $|(-)e^{j\omega}| = 1$ and $|\beta^*/\beta| = 1 \, \forall \beta \in \mathbb{C}$, this magnitude has value 1. Thus:

$$|A(z)| = (1) \prod_{i=1}^{M} (1) = 1 \Rightarrow |A|^2 = A\tilde{A} = 1$$

ii. *If* $L \neq 0$, *the factor* z^L *introduces a set of pole-zero pairs. Assume* $L > 0$. *What are the poles and zeros for this factor?*

This introduces a set of poles at infinity and zeros at zero, with multiplicity L. This obeys the symmetry of the poles, since $|0| = |1/\infty|$.

6. *Use the method of partial fractions to find the (causal) inverse Laplace transform for each of the following.*

(a) *Basic case:*

$$H(s) = \frac{3s^3 + 2}{(s+2)(s+3)}$$

$$H(s) = As + B + \frac{C}{s+2} + \frac{D}{s+3}$$

$$\begin{aligned} As + B & &&= 3s - 15 \\ C &= \frac{3(-2)^3 + 2}{-2+3} &&= -22 \\ D &= \frac{3(-3)^3 + 2}{-3+2} &&= 79 \end{aligned}$$

$$H(s) = 3s - 15 - \frac{22}{s+2} + \frac{79}{s+3}$$

$$\mathcal{L}^{-1}\left\{H(s)\right\} = h(t) = 3\delta'(t) - 15\delta(t) - 22e^{-2t}u(t) + 79e^{-3t}u(t)$$

(b) *Pole with multiplicity:*

$$H(s) = \frac{3s^3+2}{(s+2)(s+3)^2}$$

$$H(s) = A + \frac{B}{s+2} + \frac{C}{s+3} + \frac{D}{(s+3)^2}$$

$$\begin{aligned}
A & & &= 3 \\
B &= \frac{3(-2)^3+2}{(-2+3)^2} & &= -22 \\
C &= \frac{d}{ds}\left(\frac{3s^3+2}{s+2}\right)\Bigg|_{s=-3} = \frac{6(-3)^3+18(-3)^2-2}{(-3+2)^2} & &= -2 \\
D &= \frac{3(-3)^3+2}{(-3+2)} & &= 79
\end{aligned}$$

$$H(s) = 3 - \frac{22}{s+2} - \frac{2}{s+3} + \frac{79}{(s+3)^2}$$

$$\mathcal{L}^{-1}\left\{H(s)\right\} = h(t) = 3\delta(t) - 22e^{-2t}u(t) - 2e^{-3t}u(t) + 79te^{-3t}u(t)$$

(c) *Complex poles:*

$$H(s) = \frac{3s^3+2}{(s+3)(s^2+4s+8)}$$

$$H(s) = A + \frac{B}{s+3} + \frac{C}{s-(-2-2j)} + \frac{D}{s-(-2+2j)}$$

$$\begin{aligned}
A & & &= 3 \\
B &= \frac{3(-3)^3+2}{(-3)^2+4(-3)+8} & &= \frac{-79}{5} \\
C &= \frac{3\left((2)^{3/2}\,e^{-j3\pi/4}\right)^3+2}{((-2-2j)+3)((-2-2j)+2-2j)} & & \\
&= \frac{3(16-16j)+2}{(1-2j)(-4j)} = \frac{25-24j}{-4-2j} & &= -\frac{13}{5} + \frac{73}{10}j \\
D &= \frac{3\left((2)^{3/2}\,e^{j3\pi/4}\right)^3+2}{((-2+2j)+3)((-2+2j)+2+2j)} & & \\
&= \frac{3(16+16j)+2}{(1+2j)(4j)} = \frac{25+24j}{-4+2j} & &= -\frac{13}{5} - \frac{73}{10}j
\end{aligned}$$

$$H(s) = 3 - \frac{79/5}{s+3} + \frac{-13/5+73/10j}{s-(-2-2j)} + \frac{-13/5-73/10j}{s-(-2+2j)}$$

$$\mathcal{L}^{-1}\{H(s)\} = h(t)$$

$$= 3\delta(t) - \frac{79}{5}e^{-3t}u(t) + \left(-\frac{13}{5} + \frac{73}{10}j\right)e^{(-2-2j)t}u(t) + \left(-\frac{13}{5} - \frac{73}{10}j\right)e^{(-2+2j)t}u(t)$$

$$= 3\delta(t) - \frac{79}{5}e^{-3t}u(t) + e^{-2t}\left[\left(-\frac{13}{5} + \frac{73}{10}j\right)e^{-2jt} + \left(-\frac{13}{5} - \frac{73}{10}j\right)e^{2jt}\right]u(t)$$

Let

$$z = \left(-\frac{13}{5} + \frac{73}{10}j\right)e^{-2jt} = \left(-\frac{13}{5} + \frac{73}{10}j\right)(\cos(2t) - i\sin(2t))$$

Then the part in the square brackets is

$$z + z^* = 2\,\mathrm{Re}(z) = 2\left(-\frac{13}{5}\cos(2t) - \frac{73}{10}(-\sin(2t))\right)$$

and

$$h(t) = 3\delta(t) - \frac{79}{5}e^{-3t}u(t) + e^{-2t}\left(-\frac{26}{5}\cos(2t) + \frac{73}{5}\sin(2t)\right)u(t)$$

7. *Compute the integral:*

$$\int_0^5 7e^{-2t}[\delta(t-1) + 2\delta(t-10) + 5\delta'(t-2)]\,dt$$

$$= 7\int_0^5 e^{-2t}\delta(t-1)\,dt + 14\int_0^5 e^{-2t}\delta(t-10)\,dt + 35\int_0^5 e^{-2t}\delta'(t-2)\,dt$$

$$= 7\left. e^{-2t}\right|_{t=1} + 14(0) + 35\left.\left[(-1)^1\frac{d}{dt}e^{-2t}\right]\right|_{t=2}$$

$$= 7e^{-2} - 35(-2)e^{-2(2)} = 7e^{-2} + 70e^{-4}$$

8. *Windowing*

```
clc; clear all; clear screen;

% Pset 5 Question 8
% Jonathan Lam

% 8c
N = 15;
r = 30;
beta = 3.05;

hr = rectwin(N);            % rectangular window
hc = chebwin(N, r);         % Chebyshev window
hk = kaiser(N, beta);       % Kaiser window
```

```

w = linspace(0, pi, 1000);

% normalization
hr = hr/sum(hr);
hc = hc/sum(hc);
hk = hk/sum(hk);

H = freqz(hr, 1, w);
plot(w, 20*log10(abs(H)));
hold on;

H = freqz(hc, 1, w);
plot(w, 20*log10(abs(H)));
hold on;

H = freqz(hk, 1, w);
plot(w, 20*log10(abs(H)));
ylim([-50 0]);
xlim([0 pi]);
ylabel(["Magnitude" "(dB)"]);
xticks(0:pi/4:pi);
xticklabels(["0","\pi/4","\pi/2","3\pi/4","\pi"]);
xlabel(["Digital frequency" "(rad)"]);
legend(["Rectangular" "Chebyshev", "Kaiser"]);
title("Magnitude response of window functions");

% 8c
figure;

subplot(1, 3, 1);
stem(hr);
ylim([0 0.12]);
xlim([0 16]);
ylabel("h[n]");
xlabel("n");
title("Rectangular window");

subplot(1, 3, 2);
stem(hc);
ylim([0 0.12])
xlim([0 16]);
ylabel("h[n]");
xlabel("n");
title("Chebyshev window");

```

```
subplot(1, 3, 3);
stem(hk);
ylim([0 0.12]);
xlim([0 16]);
ylabel("h[n]");
xlabel("n");
title("Kaiser window");

% 8d
figure;

H = freqz(hr, 1, w);
plot(w, unwrap(180/pi*angle(H)));
ylabel(["Phase" "(deg)"]);
xlabel(["Digital frequency" "(rad)"]);
xticks(0:pi/4:pi);
xticklabels(["0","\pi/4","\pi/2","3\pi/4","\pi"]);
title("Phase response of rectangular window");
```

(a) *Why does $h = h/sum(h)$ yield $H(0) = 1$? (This assumes $sum(h) \neq 0$; if it is zero, what is $H(0)$?)*

Normalizing h makes the sum of the coefficients equal to one:

$$\sum_{-\infty}^{\infty} h[n] = 1$$

Since

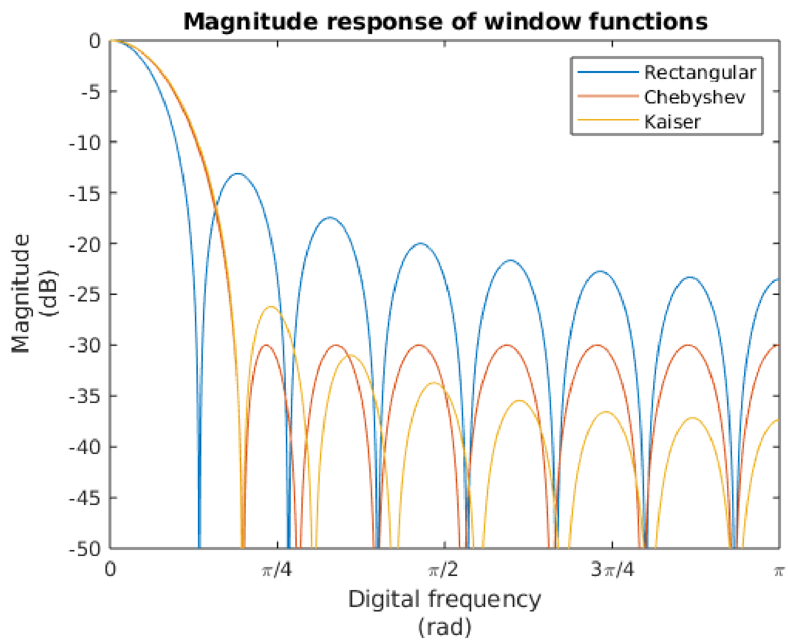
$$H(\omega) = \sum_{-\infty}^{\infty} h[n]e^{-j\omega n}$$

then

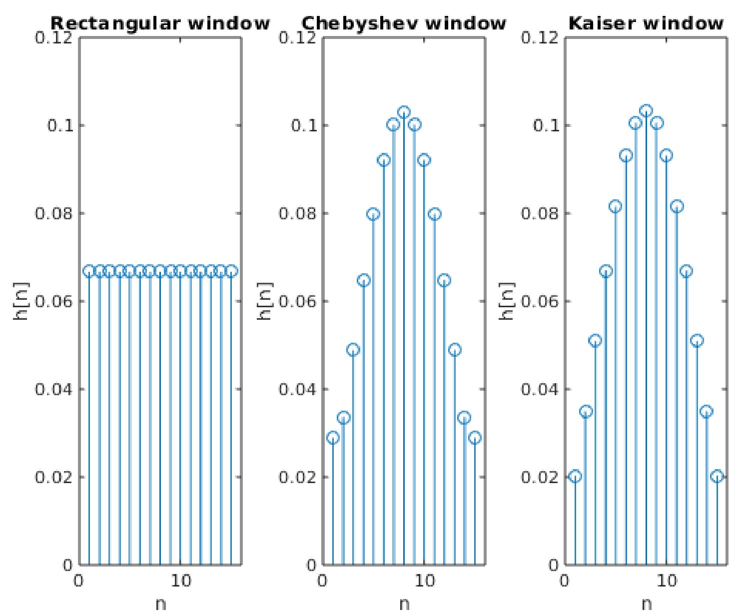
$$H(0) = \sum_{-\infty}^{\infty} h[n]e^{0} = \sum_{-\infty}^{\infty} h[n] = 1$$

If $\text{sum}(h) = 0$, then we wouldn't be able to normalize it in the first place. But if we wanted to calculate $H(0)$, it would be equal to $\text{sum}(h) = 0$ by the same reasoning as above.

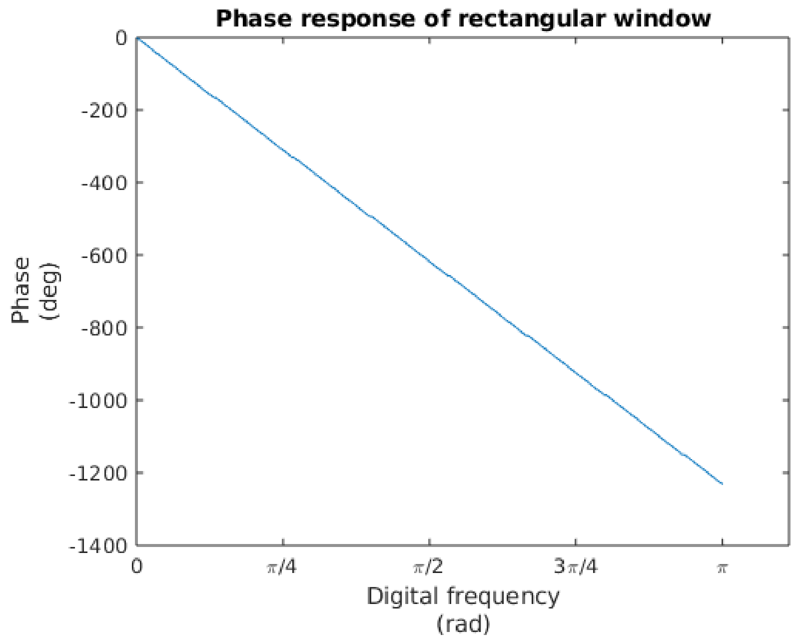
(b) Figure:



(c) Figure:



(d) Figure:

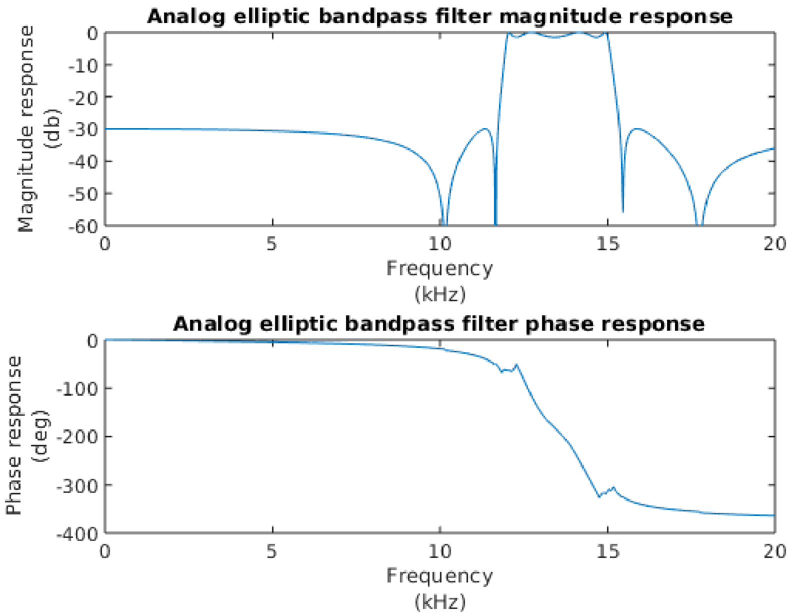


9. *Analog elliptic bandpass filter*

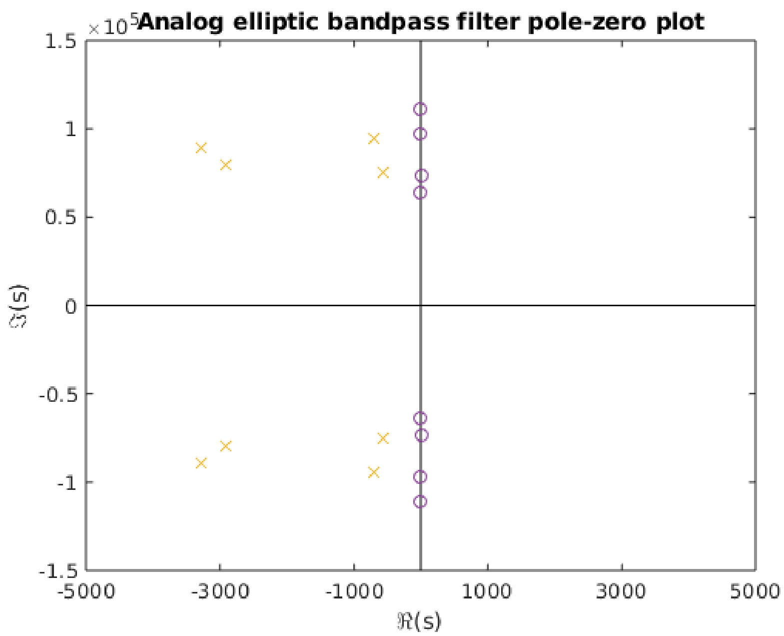
```
% Pset 5 Question 9
% Jonathan Lam
clc; clear all;

% design a bandpass analog elliptic filter
fp = [12e3 15e3];
fs = [10e3 16e3];
rp = 1.5;
rs = 30;
[n, wn] = ellipord(2*pi*fp, 2*pi*fs, rp, rs, 's');
[z, p, k] = ellip(n, rp, rs, wn, 's');
[b, a] = zp2tf(z, p, k);

% 9a
w = linspace(0, 2*pi*20e3, 1e3);
H = freqs(b, a, w);

% 9b
figure;
subplot(2, 1, 1);
plot(w, 20*log10(abs(H)));
ylim([-60 0]);
xlim([0 2*pi*20e3]);
```

```
yticks(-60:10:0);
xticks(0:2*pi*5e3:2*pi*20e3);
xticklabels(["0" "5" "10" "15" "20"]);
ylabel(["Magnitude response" "(db)"]);
xlabel(["Frequency" "(kHz)"]);
title("Analog elliptic bandpass filter magnitude response");

subplot(2, 1, 2);
plot(w, unwrap(180/pi*angle(H)));
xlim([0 2*pi*20e3]);
xticks(0:2*pi*5e3:2*pi*20e3);
xticklabels(["0" "5" "10" "15" "20"]);
ylabel(["Phase response" "(deg)"]);
xlabel(["Frequency" "(kHz)"]);
title("Analog elliptic bandpass filter phase response");

% 9c
figure;
plot([-1e7i 1e7i], 'k');     % plot imaginary axis
hold on;
plot([-1e4 1e4+0.1i], 'k'); % plot real axis
hold on;
plot(p, 'x');               % plot poles
hold on;
plot(z, 'o');               % plot zeroes
ylim([-1.5e5 1.5e5]);
xlim([-5e3 5e3]);
yticks(-1.5e5:0.5e5:1.5e5);
xticks(-5e3:2e3:5e3);
ylabel('\Im(s)');
xlabel('\Re(s)');
title("Analog elliptic bandpass filter pole-zero plot");
```

(a) Figure:



(b) Figure:



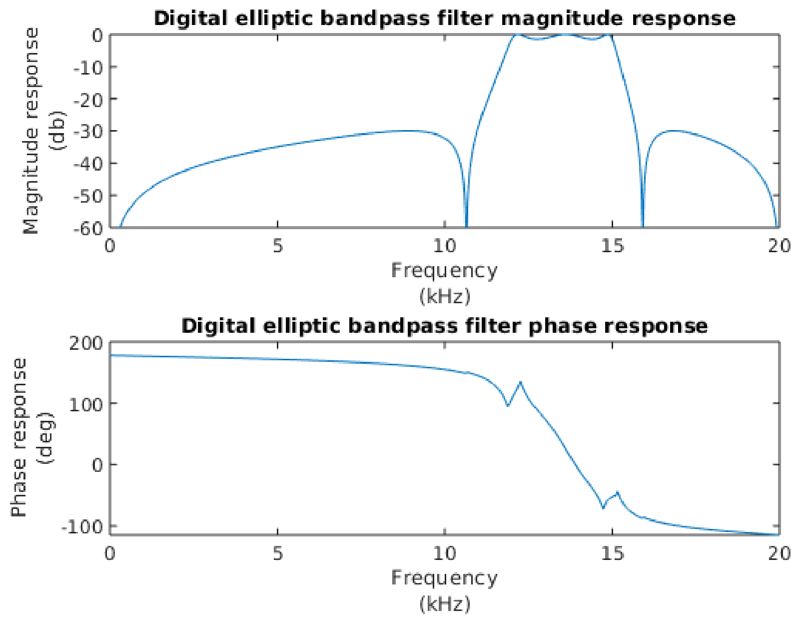
10. *Digital elliptic bandpass filter*

```
clc; clear all;
```

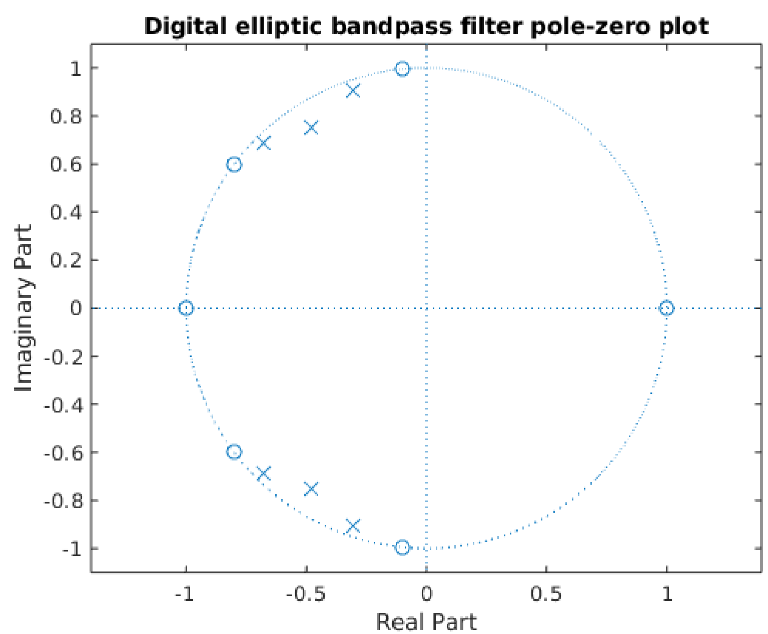
```
% Pset 5 Question 10
% Jonathan Lam

% design digital filter with same specs as q9
fp = [12e3 15e3];
fs = [10e3 16e3];
rp = 1.5;
rs = 30;
fsamp = 40e3;
Bnyq = fsamp/2;
fpd = fp/Bnyq;
fsd = fs/Bnyq;
[nd, fnd] = ellipord(fpd, fsd, rp, rs); % freq specs normalized to Nyq BW
[zd, pd, kd] = ellip(nd, rp, rs, fnd);
[bd, ad] = zp2tf(zd, pd, kd);

% compute frequency response
w = linspace(0, 2*pi*20e3, 1e3);
wd = w/fsamp;
H = freqz(bd, ad, wd);

% plot magnitude and phase response
figure;
subplot(2, 1, 1);
plot(wd, 20*log10(abs(H)));
ylim([-60 0]);
xlim([0 pi]);
yticks(-60:10:0);
xticks(0:pi/4:pi);
xticklabels((0:pi/4:pi)/2/pi*fsamp/1e3);
ylabel(["Magnitude response" "(db)"]);
xlabel(["Frequency" "(kHz)"]);
title("Digital elliptic bandpass filter magnitude response");

subplot(2, 1, 2);
plot(wd, unwrap(180/pi*angle(H)));
xlim([0 pi]);
xticks(0:pi/4:pi);
xticklabels((0:pi/4:pi)/2/pi*fsamp/1e3);
ylabel(["Phase response" "(deg)"]);
xlabel(["Frequency" "(kHz)"]);
title("Digital elliptic bandpass filter phase response");

% plot pole-zero plot
figure;
zplane(zd, pd);
```

```
title("Digital elliptic bandpass filter pole-zero plot");
```

(a) Figure:



(b) Figure:



ECE211 – Pset 6

Jonathan Lam

April 16, 2020

1. *Use the method of partial fractions (residue approach) to find the (right-sided) inverse z-transform of*

$$H(z) = \frac{3z^3+2}{(3z-1)(4z+3)}$$

$$z^{-1}H(z) = \frac{3z^3+2}{12z(z-1/3)(z+3/4)} = A + \frac{B}{z} + \frac{C}{z-1/3} + \frac{D}{z+3/4}$$

$$\begin{aligned}
A & &= \frac{1}{4} \\
B &= \frac{2}{12(-1/3)(3/4)} &= -\frac{2}{3} \\
C &= \frac{3(1/3)^3+2}{12(1/3)(1/3+3/4)} &= \frac{19}{39} \\
D &= \frac{3(-3/4)^3+2}{12(-3/4)(-3/4-1/3)} &= \frac{47}{624}
\end{aligned}$$

$$H(z) = \frac{z}{4} - \frac{2}{3} + \frac{19}{39}\frac{z}{z-1/3} + \frac{47}{624}\frac{z}{z+3/4}$$

$$h[n] = \frac{1}{4}\delta[n+1] - \frac{2}{3}\delta[n] + \frac{19}{39}\left(\frac{1}{3}\right)^n u[n] + \frac{47}{624}\left(-\frac{3}{4}\right)^n u[n]$$

2. *A causal discrete-time signal* $x[n]$ *has the following poles:* 0.4 *with multiplicity 3, and* $0.3\exp \pm j4\pi/5$*, each with multiplicity 2. Assume the underlying sampling rate is* 10kHz*.*

   (a) *Write a general expression for* $x[n]$*.*

$$\begin{aligned}
x[n] =& (a_1 n^2 + a_2 n + a_3)(0.4)^n u[n] \\
&+ (a_4 n + a_5)(0.3)^n e^{j4\pi n/5} u[n] + (a_6 n + a_7)(0.3)^n e^{-j4\pi n/5} u[n] \\
=& (a_1 n^2 + a_2 n + a_3)(0.4)^n u[n] + (a_8 n + a_9)(0.3)^n \cos\left(\frac{4\pi n}{5} + \theta\right) u[n]
\end{aligned}$$

(b) *For each mode, specify the time constant in seconds and, if there is an oscillation, the frequency of the oscillation in Hertz.*

For the pole at 0.4:

$$\tau = \frac{T}{|\ln|p||} = \frac{10000^{-1}\text{s}}{|\ln 0.4|} = 1.09 \times 10^{-4}s$$

For the pole at $0.3\exp(4\pi/5)$:

$$\tau = \frac{10000^{-1}\text{s}}{|\ln 0.3|} = 8.31 \times 10^{-5}s$$

And this pole oscillates at:

$$f = \frac{\omega_d}{2\pi} \cdot f_s = \frac{4\pi}{5}\frac{1}{2\pi}(10\text{kHz}) = 4\text{kHz}$$

3. *The input to a digital filter $H(z)$ is:*

$$x[n] = 3u[n] + 4(-0.5)^n u[n]$$

The output is

$$y[n] = 6u[n]+5(-0.5)^n u[n]+(6n+3)(0.5)^n u[n]+0.4(0.2)^n \cos(n\pi/3+\pi/4)$$

(a) *Identify the natural and forced responses in y.*

Natural:

$$(6n+3)(0.5)^n u[n] + 0.4(0.2)^n \cos(n\pi/3 + \pi/4)$$

Forced:

$$6u[n] + 5(-0.5)^n u[n]$$

(b) *Specify the system poles, with multiplicity, assuming no pole-zero cancellation with x.*

- 0.5 (double pole)
- $0.2\exp(\pm j\pi/3)$ (simple poles)

4. *An analog signal is given by:*

$$x(t) = 3\exp(-2t)u(t) + e^{-4t}(t\cos(3t) - \sin(3t))u(t)$$

Specify the poles of x with multiplicity. Also, the above contains two modes; for each mode, specify the time constant in seconds, and if there is oscillation the frequency in Hertz.

Pole at -2 have multiplicity 1
Poles at $-4 \pm j3$: have multiplicity 2 (because of t factor)

Mode for pole at -2: $\tau = 0.5$s, no oscillation
Mode for poles at $-4 \pm j3$: $\tau = 0.25$s, oscillating with frequency $\frac{3}{2\pi}$Hz

5. *The response of an analog filter $H(s)$ with input x is y. x has poles at $s = 0, -1, -4$, all simple. y has poles at $s = 0, -1, -2 \pm j3$ (simple), and $s = -3$ (double).*

   (a) *The fact that the output has no pole at -4 tells you something about $H(s)$. What?*

   There is a pole-zero cancellation with the filter. In particular, $H(s)$ has a zero at 4.

   (b) *Of the output poles, which must be the poles of $H(s)$ (and with what multiplicity)?*

   These would be all of the poles that don't come from the input. In particular: $s = -2 \pm j3$ (single) and $s = -3$ (double). With pole-zero cancellations, it's possible that these multiplicities are higher.

   (c) *At first glance the system appears to be stable – why? However, it turns out we are missing some information, that makes it possible for the system to be unstable! Explain what might be happening that would "disguise" the fact that the system is unstable?*

   It would appear stable because all of the poles of the system (from the previous question) are in the LHP. However, it is possible that there is a pole-zero cancellation (with the zero from the input and the pole from the output) that is hiding a pole of the transfer function in the RHP (or on the $j\omega$ axis).

ECE211 – Pset 7

Jonathan Lam

April 28, 2020

1 Discrete time

Consider a discrete-time state-space realization $\{A, B, C, D\}$ and corresponding transfer function matrix $H(z)$. Let $x[n]$, $u[n]$, $y[n]$ denote the state vector, input vector, and output vector, respectively. Assume N state variables, m inputs, and n outputs.

1. *Write the state-space equations.*

$$x[n+1] = Ax[n] + Bu[n]$$

$$y[n] = Cx[n] + Du[n]$$

2. *Write the formula for $H(z)$ in terms of A, B, C, D.*

$$H(z) = C(zI - A)^{-1}B + D$$

3. *Give the dimensions of A, B, C, D, H.*

$$A : N \times N$$

$$B : N \times m$$

$$C : n \times N$$

$$D : n \times m$$

$$H : n \times m$$

4. *Under certain conditions, the state evolves in the form $x[n] = \Phi[n]x[0]$.*

   (a) *Specify those conditions.*
   The input is zero. (i.e., $u[n] = 0$), $n > 0$, and the system is LTI.

   (b) *State the name of $\Phi[n]$.*
   It is called the state transition matrix.

   (c) *Express $\Phi[n]$ in terms of A, B, C, D.*

$$\Phi[n] = A^n$$

(d) *State the one-sided z-transform of* $\Phi[n]$, *in terms of* A, B, C, D.

$$z(zI - A)^{-1}$$

(e) *Give precise conditions on* A, B, C, D *for which* $\Phi[n] \Rightarrow 0$ *as* $n \Rightarrow \infty$.
All eigenvalues of A are in the stability region ($|\lambda| < 1$).

2 Continuous time

1. *Write the state-space equations*

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

2. *Write the formula for* $H(s)$.

$$H(s) = C(sI - A)^{-1}B + D$$

3. *Give the dimensions of the matrices.*

   Same as for the discrete-time case.

4. *Under certain conditions, the state evolves in the form* $x(n) = \Phi(n)x(0)$.

   (a) *Specify those conditions.*
   The input is zero, (i.e., $u(n) = 0$), $t > 0$, and the system is LTI.

   (b) *State the name of* $\Phi(n)$.
   It is called the state transition matrix.

   (c) *Express* $\Phi(n)$ *in terms of* A, B, C, D.

$$\Phi(n) = e^{At}$$

   (d) *State the one-sided Laplace transform of* $\Phi(n)$, *in terms of* A, B, C, D.

$$(sI - A)^{-1}$$

   (e) *Give precise conditions on* A, B, C, D *for which* $\Phi(n) \Rightarrow 0$ *as* $n \Rightarrow \infty$.
   All eigenvalues of A are in the stability region ($\Re[\lambda] < 0$).

3 Example transfer function

Given the following transfer function matrix of a continuous-time system:

$$H(s) = \begin{bmatrix} \frac{1}{s+2} & \frac{1}{(s+2)^2(s+3)} & 0 \\ \frac{1}{s+3} & \frac{s+4}{(s+2)(s+3)} & \frac{1}{s+4} \end{bmatrix}$$

Although not specified, assume underlying state-space realization $\{A, B, C, D\}$.

1. *Specify the transfer function from the first input to the second output.*

$$H(s) = \frac{1}{s+3}$$

2. *List the system poles with multiplicity*

$$p = -2, m = 2$$

$$p = -3, m = 1$$

$$p = -4, m = 1$$

3. *What information is known about the eigenvalues of A?*

   A must have the eigenvalue -2 (with multiplicity ≥ 1), eigenvalue -3 (with multiplicity ≥ 2), and eigenvalue -4 (with multiplicity ≥ 1). Their multiplicities may be higher because of pole-zero cancellation.

4. *What is the minimum size of A?*

   4×4

5. *Suppose there is a hidden pole at* 4. *How does that change the answers to parts (3) and (4), if all?*

   It means that A must also have an eigenvalue at 4 (with multiplicity ≥ 1), and that the minimum size of A is 5×5.

6. *Again assuming there is a hidden pole at* 4. *Is the system internally stable, and is the system externally stable?*

   Internally unstable (because of pole in A in the RHP), but externally stable (because no pole in the RHP in $H(z)$, input cancels out pole of internal state).

7. *Assume H is as given, there is a hidden pole at* 4, *and there are no other hidden poles (including no higher multiplicities other than as system poles). We are now interested in* e^{At}.

   (a) *What is the size of A?*

   5×5

(b) *We can compute e^{At} by evaluating a polynomial of the form $r_0(t)I + r_1(t)A + \cdots + r_M(t)A^M$. What is M?*

By the Cayley-Hamilton theorem, we can evaluate a polynomial of degree $M = 5 - 1 = 4$ to evaluate e^{At}.

(c) *Write the general formula for an entry of e^{At}.*

Let $\{\lambda_k\} = \{-2, -3, -4, 4\}$ denote the eigenvalues of A, $\{m_k\} = \{2, 1, 1, 1\}$ denote the multiplicities of the eigenvalues, and $L = 4$ denote distinct eigenvalue count. Then

$$\left(e^{At}\right)_{ij} = \sum_{k=1}^{L} p_k(t)e^{\lambda_i t} = (a_0 + a_1 t)e^{-2t} + a_2 e^{-3t} + a_4 e^{-4t} + a_5 e^{4t}$$

where $p_k(t)$ is a polynomial of degree at most $m_k - 1$.

ECE211 – Pset 8

Jonathan Lam

May 9, 2020

A real WSS signal $x[n]$ is modeled as:

$$x[n] = v[n] + 0.5v[n-1] + 0.4v[n-2] - 0.2x[n-1] - 0.8x[n-2]$$

where $v[n]$ is the 0-mean white noise with $\sigma_v^2 = 4$.

Questions

1. It has both AR ($x[n-N]$ terms) and MA ($v[n-N]$ terms) components, so it is ARMA.
2. Since it outputs x, it is the innovations filter (not the whitening filter).
3. Rearranging to solve for the inverse filter $v[n]$:

$$v[n] = x[n] + 0.2x[n-1] + 0.8x[n-2] - 0.5v[n-1] - 0.4v[n-2]$$

4. Transfer function:

$$H(z) = \frac{z^2 + 0.5z + 0.4}{z^2 + 0.2z + 0.8}$$

5. PSD:

$$S_x(\omega) = \sigma_v^2 \frac{|B(\omega)|^2}{|A(\omega))|^2} = 4\frac{\left|1 + 0.5e^{-j\omega} + 0.4e^{-2j\omega}\right|^2}{\left|1 + 0.2e^{-j\omega} + 0.8e^{-2j\omega}\right|^2}$$

PSD peak

(Part 4b-c) The peak of the pwelch-estimated PSD ($\omega_0 = 1.6322$), the PSD calculated from the transfer function ($\omega_0 = 1.6690$), and the angle of the pole ($\arg p = \pm 1.6828$) are very close (within a range of 0.05 for this specific x). This makes sense, as the pole makes the signal's magnitude (and thus power) blow up near the pole. Also note that the exact value of the calculated PSD peak and the argument of the pole do not have to be exactly the same; the behavior of the signal may make the exact position of the peak vary, but it should be fairly close to $\arg p$.

```
% 2
H = tf([1 0.5 0.4], [1 0.2 0.8])
```

```
H =

  s^2 + 0.5 s + 0.4
  -----------------
  s^2 + 0.2 s + 0.8

Continuous-time transfer function.
```

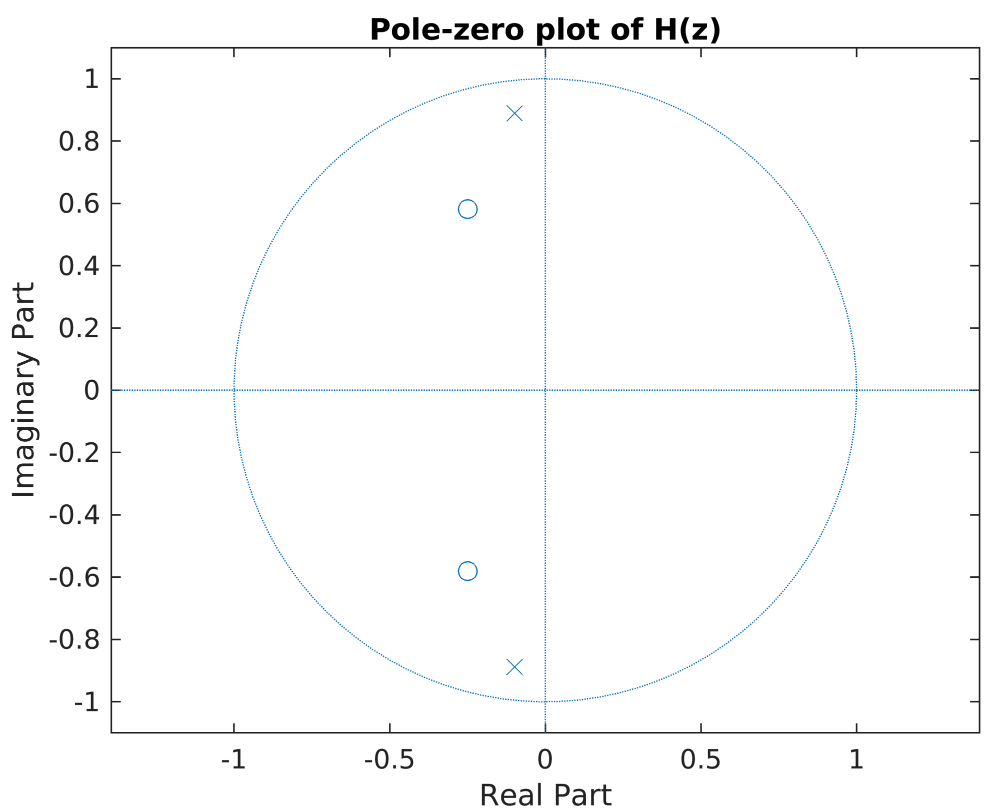
```
p = pole(H)
```

```
p = 2×1 complex
  -0.1000 + 0.8888i
  -0.1000 - 0.8888i
```

```
z = zero(H)
```

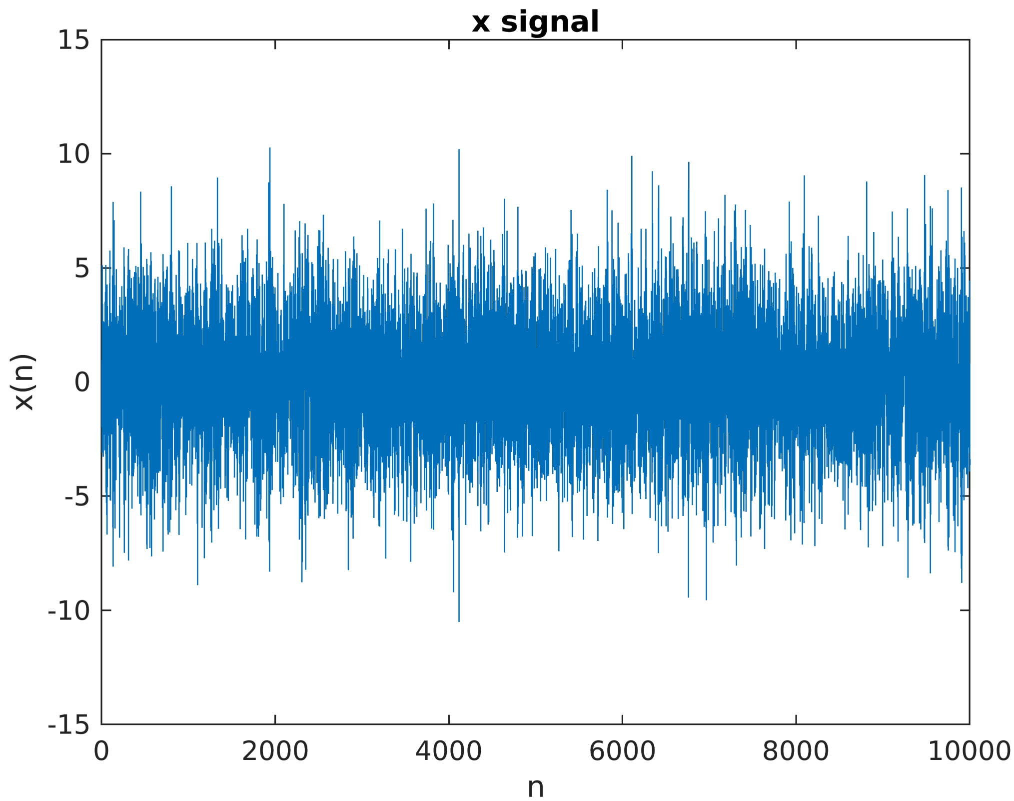
```
z = 2×1 complex
  -0.2500 + 0.5809i
  -0.2500 - 0.5809i
```

```
zplane(z, p)
title('Pole-zero plot of H(z)')
```

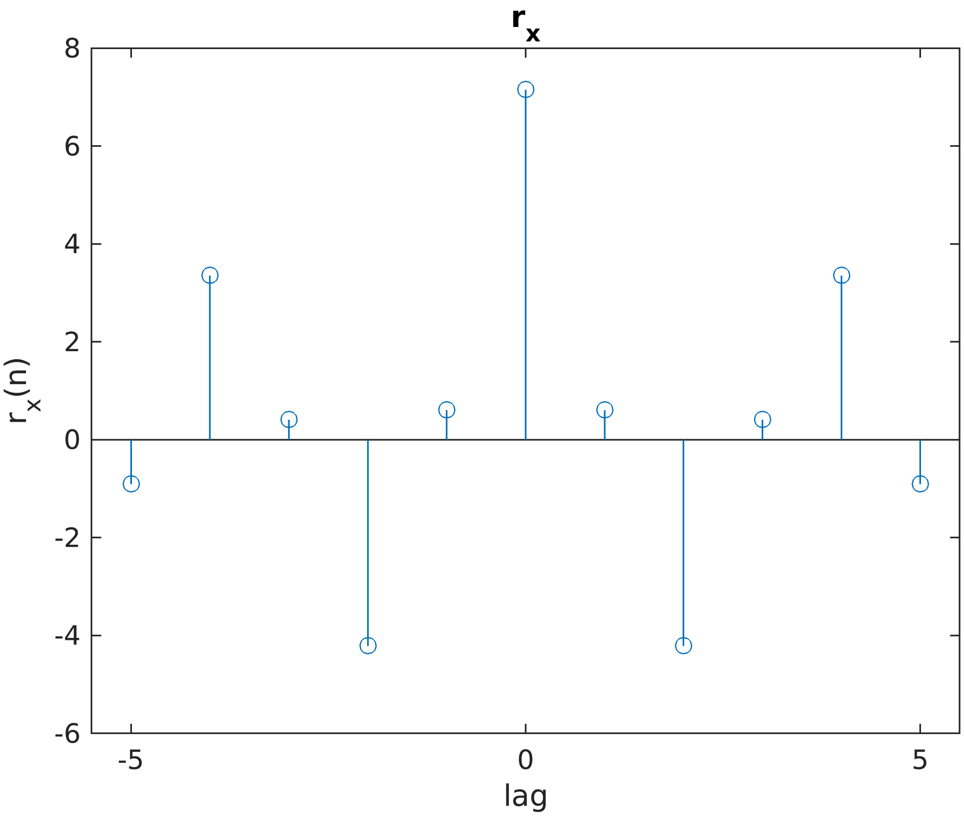



```
% 3a
N = 10000;
v = 2*randn(N, 1);
x = filter([1 0.5 0.4], [1 0.2 0.8], v);
```

```
plot(x);
title('x signal');
xlabel('n');
ylabel('x(n)');
```




```
% 3b) equivalent to doing:
% [xc, lags] = xcorr(x)
% stem(-5:5, xc(lags>=-5 & lags<=5))
lags = -5:5;
xcor = zeros([11 1]);
for lag=0:5
    xcor(lags==lag|lags==-lag) = dot(x(1:N-lag),x(lag+1:N))/(N-lag);
end
% 3c
stem(lags, xcor)
title('r_x')
xlabel('lag')
ylabel('r_x(n)')
xlim([-5.5 5.5])
```




```
% 3d
% M=6 here
tm = toeplitz(xcor(lags>=0))
```

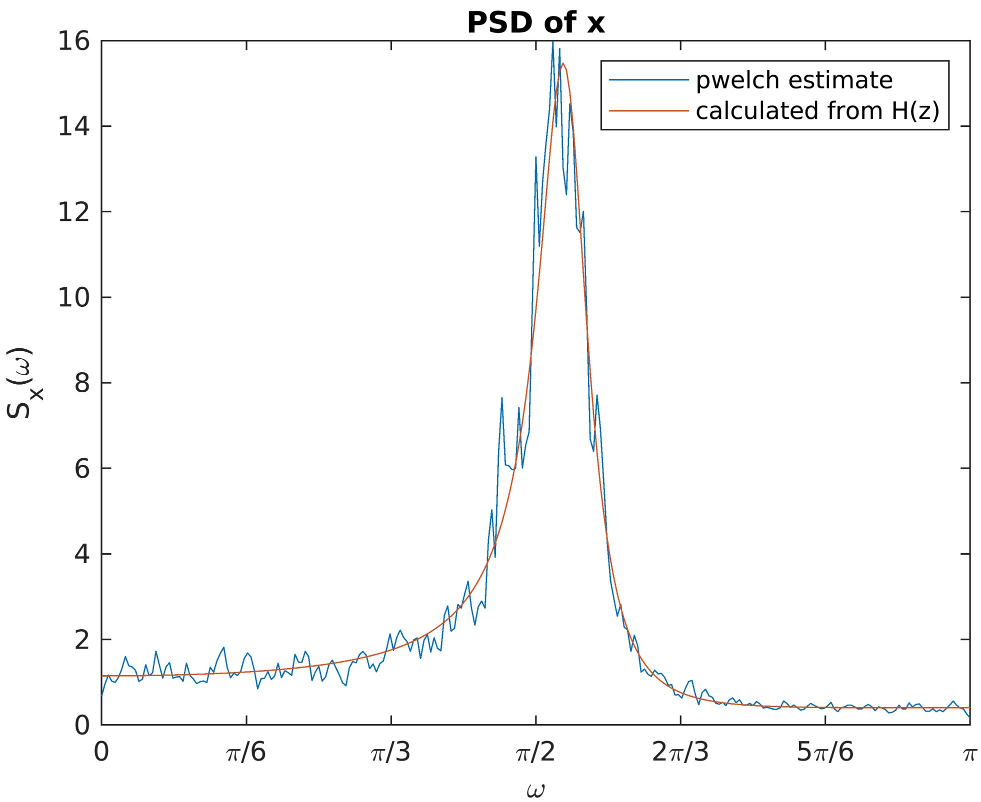
```
tm = 6×6
    7.1510    0.6043   -4.2092    0.4110    3.3563   -0.9066
    0.6043    7.1510    0.6043   -4.2092    0.4110    3.3563
   -4.2092    0.6043    7.1510    0.6043   -4.2092    0.4110
    0.4110   -4.2092    0.6043    7.1510    0.6043   -4.2092
    3.3563    0.4110   -4.2092    0.6043    7.1510    0.6043
   -0.9066    3.3563    0.4110   -4.2092    0.6043    7.1510
```

```
% 3e
ev = eig(tm)
```

```
ev = 6×1
    1.3598
    2.7045
    4.0701
    4.7076
   14.7870
   15.2768
```

```
% 4a
[s_est, w] = pwelch(x, hamming(512), 256, 512);
% sanity check; calculate from H(z) and normalize
s_x = 4*(abs(1+0.5*exp(-1j*w)+0.4*exp(-2j*w)) ...
    ./abs(1+0.2*exp(-1j*w)+0.8*exp(-2j*w))).^2/pi;
plot(w, s_est)
hold on;
```

```
plot(w, s_x)
hold off;
xlim([0 pi])
title('PSD of x')
xlabel('\omega')
xticks(0:pi/6:pi)
xticklabels({'0','\pi/6','\pi/3','\pi/2','2\pi/3','5\pi/6','\pi'})
ylabel('S_x(\omega)')
legend({'pwelch estimate', 'calculated from H(z)'})
```




```
% 4b
w_0_est = w(s_est==max(s_est))
```

```
w_0_est = 1.6322
```

```
w_0_calc = w(s_x==max(s_x))
```

```
w_0_calc = 1.6690
```

```
% 4c
angle(p)
```

```
ans = 2×1
    1.6828
   -1.6828
```

Blue Book

NAME Jonathan Lam

SUBJECT Signals & Systems

INSTRUCTOR Fontaine

EXAM SEAT NO. SECTION

DATE 3/5/20 GRADE

10 7/8 X 8 1/4

50-12 PAGE

1	9
2	6
3	4
4	3
5	4
6	15
7	4
8	10
9	6
10	~~8~~ 10
11	8
12	6
13	15

100

$h * x =$

1 a) $\int_{-\infty}^{\infty} h(\tau)\, x(t-\tau)\, d\tau = y(t)$

b)

$y[n] = h * x = \sum_{m=-\infty}^{\infty} h[m]\, x[n-m]$

9

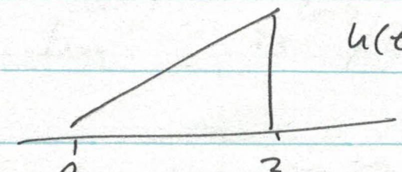
c) $\int_{-\infty}^{\infty} |h(t)|\, dt < \infty$. (i.e., $h \in L^1$).

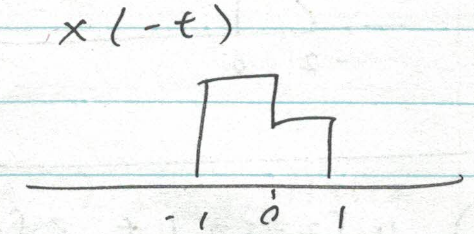
2 a) support: $-1 \le t \le 3$

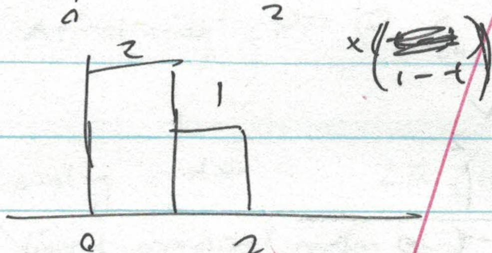
duration: 4.

6

b)



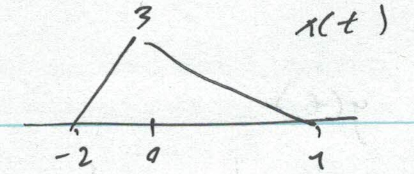




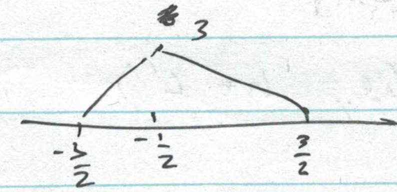
c) $\int_0^1 2t\, dt + \int_1^2 t\, dt$

$= t^2 \Big|_0^1 + \frac{t^2}{2}\Big|_1^2 = 1 + \left(2 - \frac{1}{2}\right) = \frac{5}{2}.$

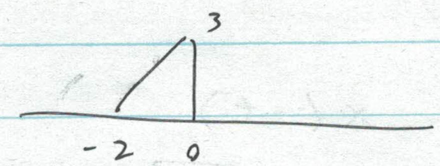
3).



a) $x(2t+1) = x(2(t+\frac{1}{2}))$

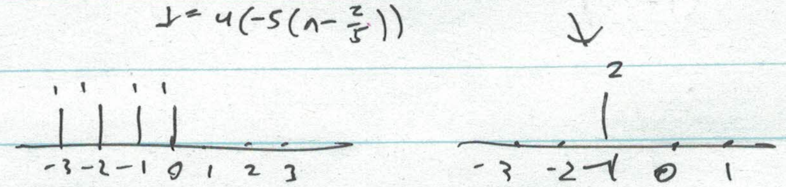


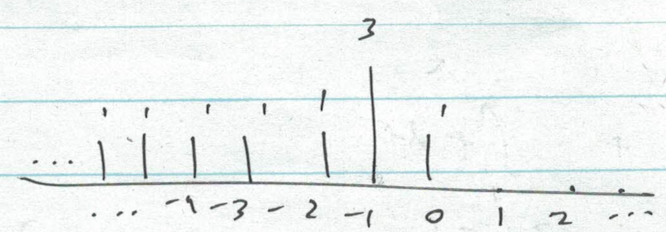
b) $x(t) \cdot u(-3t)$.



4.) $u[2-5n] + 2\delta[n+1]$

$= u(-5(n-\frac{2}{5}))$





5a) length h: 7

length x: 4

support y: $-2 \leq n \leq 7$

length y: 10

3b) $y[2] = \sum_{m=-\infty}^{\infty} h[m]\, x[n-m] = h[0]x[2] + h[1]x[1] + h[2]x[0] + h[3]x[-1]$

5c) 1) yes (because impulse response is causal)

2) yes (because FIR $\Rightarrow h(x) \in L^1$)

6.) a) $f_s^{(in)} = \cancel{}$ 22kHz

b) 32kHz, 34kHz

c) $f_{nyq}^{(in)} = f_s^{(in)}/2 = 11\,kHz$.

d) anti-aliasing LPF @ $f_{cutoff} = f_{nyq}^{(in)} = \cancel{}\ 11\,kHz$.

15

e) analog radian: $2\pi(10kHz) = 20000\pi$ rad/s.

digital normalized radian freq: $= \omega T = 2\pi\left(\frac{10kHz}{22kHz}\right) = \frac{10\pi}{11}$.

as frac of sampling rate: $\frac{10kHz}{22kHz} = \frac{5}{11}$

f). $f_s^{(out)} = 2f_s^{(in)} = 44kHz$.

cutoff freq $= f_{nyq}^{(out)} = 22kHz$.

anti-imaging filter.

g). 20kHz, 24kHz, 64kHz, 68kHz.

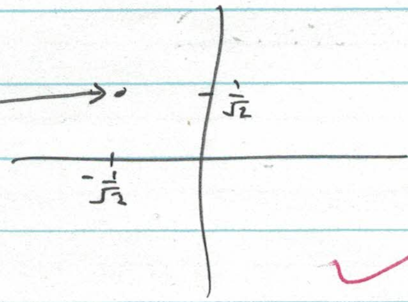
7.) $e^{jk\frac{\pi}{4}} = e^{j\left(\frac{19\pi}{4}\right)} = e^{j\left(4\pi + \frac{3\pi}{4}\right)}$

$\Rightarrow k = 3.$

4

$e^{j\frac{3\pi}{4}}$

$= \frac{1}{\sqrt{2}}(-1+j).$



8.) $\sqrt{2}\cos\left(\omega_0 t - \frac{\pi}{4}\right) + 5\cos\omega_0 t \quad - 4\sin\omega_0 t$

10 $= \sqrt{2}\, e^{j\left(-\frac{\pi}{4}\right)} + (5+4j)\, e^{0j}$

$= \sqrt{2}\, e^{-j\frac{\pi}{4}} + (5+4j)$

9.) $x(t) = A\cos\left(2\pi f_0 t + \alpha \sin 2\pi f_m t + \phi_0\right)$

a) $f_{inst}(t) = \frac{1}{2\pi}\frac{d\phi}{dt} = \frac{1}{\cancel{2\pi}}\left(2\pi f_0 + \alpha\, 2\pi f_m \cos 2\pi f_m t\right)$

$= f_0 + \alpha f_m \cos(2\pi f_m t)$

6 b) $f_{min} = f_0 - \alpha f_m.$

$f_{max} = f_0 + \alpha f_m$

For f_{min} to be > 0, then $f_0 > \alpha f_m$.

10.) a) $x(t) = \sum_{n=-\infty}^{\infty} c_n e^{j\omega_0 n t}$

write! $\omega_0 = \frac{2\pi}{T}$

$$c_n = \frac{1}{T}\int_0^T x(t) e^{-j\omega_0 n t} \, dt$$

b) $c_m = c_{-m}^*$

c) fundamental freq $= f_0 = \frac{1}{T}$ (Hz)

3rd harmonic $= 3f_0 = \frac{3}{T}$ (Hz)

d) DC power $= |c_0|^2$

10

e) at third harmonic, power is $|c_3|^2 + |c_{-3}|^2$.

11.) a) $x_m(t) = A\cos(2\pi(f_0 + m\Delta f)t)$, $0 \le t \le T$.

$$= Re\left(x_{BB_m}(t) e^{j2\pi f_0 t}\right).$$

$$= Re\left(Ae^{j(2\pi(f_0 + m\Delta f)t)}\right) = Re\left(Ae^{j(2\pi m\Delta f t)} e^{j2\pi f_0 t}\right)$$

8

$$\Rightarrow x_{BB,m}(t) = Ae^{j(2\pi m \Delta f t)}.$$

JUSTIFY THIS

b) $x_{BB,m}(t)$, $x_{BB,n}(t)$, $m \ne n$ orth. if.

$m\Delta f - n\Delta f = k\left(\frac{1}{T}\right)$.

Technically I didn't ask so OK

~~$\Rightarrow$ Let $\Delta f = \frac{1}{MT}$ for any particular M.~~

~~If Δf was to be the same for all $x_{BB,m}$, $0 \le m < M$~~,

$\Delta f = \frac{1}{T} \Rightarrow x_{BB,M}(t) = Ae^{j(2\pi(\frac{M}{T})t)}$, mutually orthogonal

12.) $P_{in} = 0.01W$, $P_{out} = 50\,dBm$.

a) $P_{in}\,(dBm) = 10\log_{10}\left(\frac{0.01W}{0.001W}\right)$

$= 10\log_{10}(10) = 10(1) = 10\,dBm.$

b). dB

c) $\overset{gain}{g}(dB) = P_{out} - P_{in} = \underbrace{50\,dBm - 10\,dBm}_{\text{Valid because of log magic.}} = 40\,dB$

d). $40\,dB = 20\log_{10}\left(\frac{V_{out}}{1V}\right)$

6

$2 = \log_{10}(V_{out})$

$V_{out} = 10^2\,V = 100\,V.$

13.) $\langle x, y\rangle = \int_{-\infty}^{\infty} x(t)\,y(t)\,dt.$

a) $\|\phi_1(t)\| = \int_{-\infty}^{\infty} |\phi_1^2(t)|\,dt = A_1^2(2) = 1$

$\Rightarrow \frac{1}{\sqrt{2}} = A_1$

15

$\|\phi_2(t)\| = \int_{-\infty}^{\infty} |\phi_2^2(t)|\,dt = A_2^2(1) + (-A_2)^2(1)$

$= 2A_2^2$

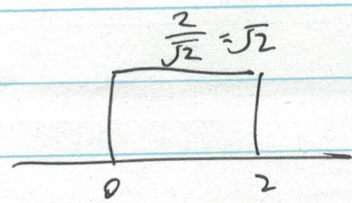
$\Rightarrow \frac{1}{\sqrt{2}} = A_2.$

b.) $\psi_3(t) = s_3 - \frac{\langle s_3, \phi_1 \rangle}{\cancel{\|\phi_1\|^2}} \phi_1 - \frac{\langle s_3, \phi_2 \rangle}{\cancel{\|\phi_2\|^2}} \phi_2.$

$\|\phi_1\| = \|\phi_2\| = 1$

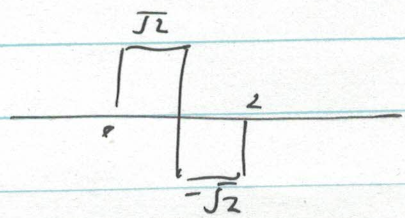
$= s_3 - \langle s_3, \phi_1 \rangle \phi_1 - \langle s_3, \phi_2 \rangle \phi_2.$

$s_3 \phi_1 =$



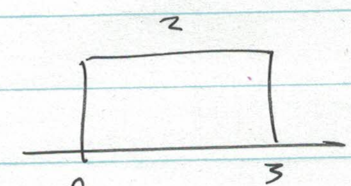
$\Rightarrow \langle s_3, \phi_1 \rangle = 2\sqrt{2}.$

$s_3 \phi_2 =$

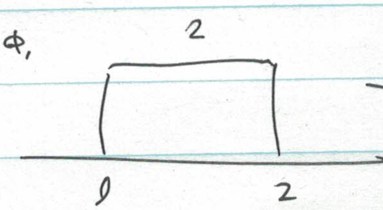


$\Rightarrow \langle s_3, \phi_2 \rangle = 0.$

$\Rightarrow$ s_3

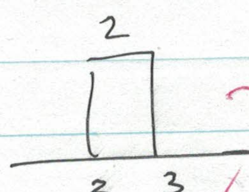


$-$ $\langle s_3, \phi_1 \rangle \phi_1$

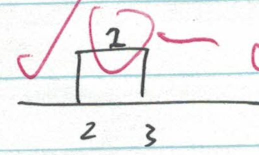


$-$ $\langle s_3, \phi_2 \rangle \phi$

$\Rightarrow \psi_3 =$



$\phi_3 = \frac{\psi_3}{\|\psi_3\|} = \frac{\psi_3}{2} =$



— ah ok, your 1 looks like 2

c). $\vec{s}_3 = \begin{pmatrix} 2\sqrt{2} \\ 0 \\ 2 \end{pmatrix}$

d) $\hat{s}_3 = \begin{pmatrix} 2\sqrt{2} \\ 0 \\ 0 \end{pmatrix}$ $\quad \vec{e} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$

e) GRAM SCHMIDT ORTHONORMALIZATION.

Jonathan Lam, Min Cheng, Joshua Yoon
Prof. Shlayan
ECE 241
Electronics I
5 / 11 / 20

Electric Car Charging Subsystem Project

Electronics Final Project

Introduction

The goal of this project was to design the charging circuitry for an electric vehicle modeled after the Gem Polaris e2. This charging subsystem would consist of a Li-ion battery, photovoltaic solar cells, and charging circuits to charge the battery from the solar cells or from a standard 120VAC 60Hz wall outlet. A circuit using basic power electronic concepts was built and components were chosen to fit design specifications outlined by the Gem Polaris characteristics, the solar cell chosen, and the wall outlets. Research was performed in order to better grasp the concepts behind the solar cell and the regulator technology, to gauge realistic and desirable specifications, and to properly simulate the different components of the circuit. The circuit was simulated in LTspice to provide insight into the battery charging rates and how the solar cell and the wall outlet representations interact with the voltage regulators.

Design and Analysis

Defining specifications

The following are tables of design specifications considered while designing the circuit:

Provided design requirements

Real-life model	Polaris Gem e2 electric vehicle
Motor	48V AC induction motor
Motor maximum power output	5.0kW (6.7hp)

Desired specifications

Maximum vehicle weight	907kg (2000lb) (1200lb dry weight + 800lb payload)
Maximum speed	11.2m/s (25mph)

Interface specifications

Solar panel	Count	4
	Maximum power rating	250W
	Maximum current rating	8.33A
	Open circuit voltage	37.4V
	Power efficiency	15.1%
	Dimensions	1652mm x 1000mm x 45mm
Battery	Voltage	48V
	Storage capacity	8.9kWHr
	Type	lithium-ion
Wall charger	Voltage	120V 60Hz AC
	Maximum power[1]	1.4kW

Circuit Design and Analysis

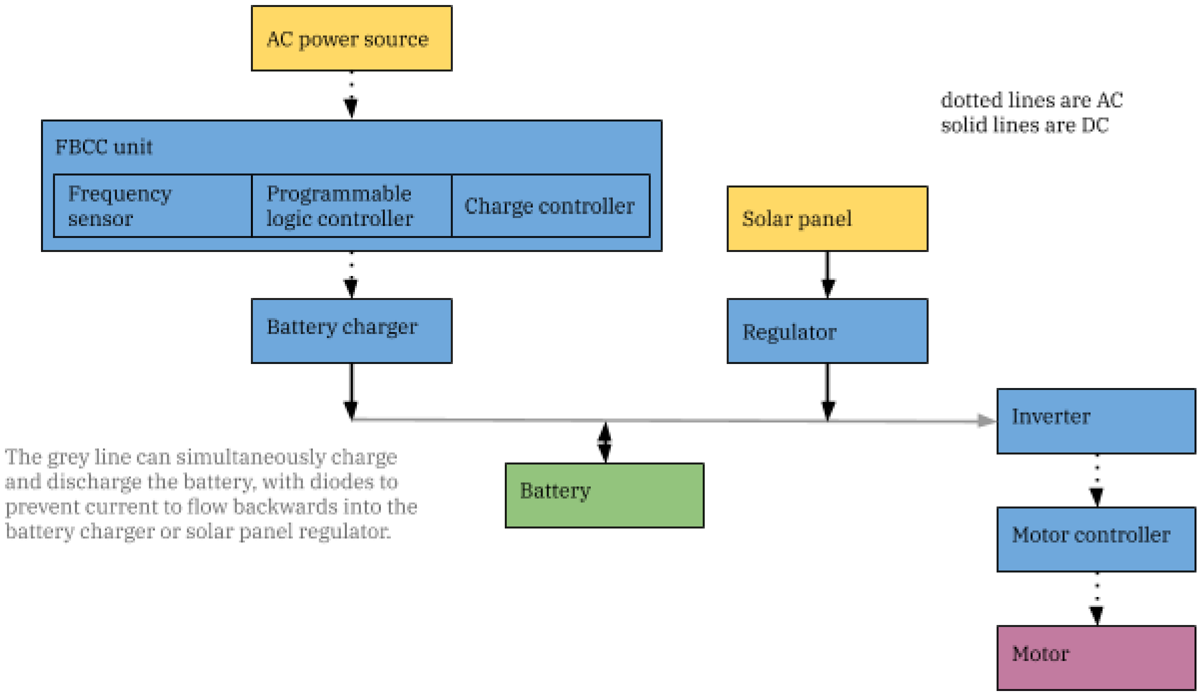
We aim to design the power system of an electric car similar to that of the Polaris Gem e2. Our design consists of three major components: a solar panel, an interface to connect with a standard 60Hz, 120V AC wall charger and the battery. There are various voltage regulators, transformers, diodes, and other devices to connect different parts of the circuit.

The below block diagram outlines a basic flow of our system. Each block will be further discussed. The above circuit outlines the basic circuitry of our system; note that the ATmega328-PU chip was used to facilitate the pins of the Arduino Uno.

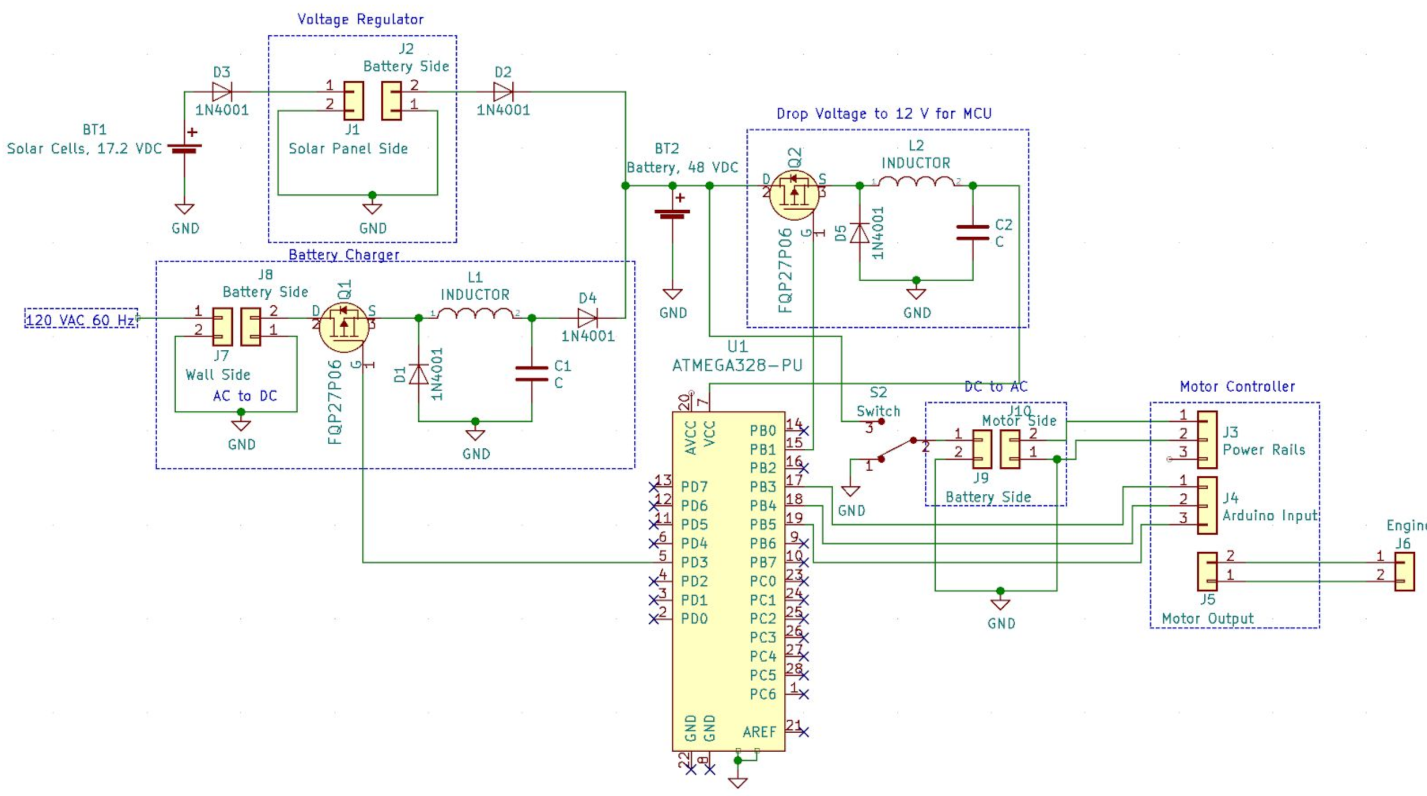
The flow of the system is as follows: the voltage from the solar panels will be fed into a voltage regulator to output a clean 50 V output, with diodes in place to protect the regulator and solar panel. Voltage from the wall will be fed into a transformer to step the AC voltage down to around 50 V, and then fed into a voltage regulator to produce a clean 50 V output. These voltage values were chosen specifically to be above the voltage value of the battery in order to facilitate charging. Also note that linear voltage regulators were chosen over switching regulators in order to better absorb changes to the input; solar panels produce notoriously noisy signals, and thus if switching regulators are used, while more power efficient with less I^2R loss, the circuit will require some kind of feedback loop to maintain a constant voltage value.

These two outputs will be coupled together, with diodes to prevent current from flowing in the incorrect direction. This output will charge a battery, which will be connected via an ignition switch to a DC to AC converter, which will be fed to the motor of the car. A microcontroller, powered by the battery via a buck converter that steps the voltage down to 12 V, will control the various buck converters and feedback loops throughout the circuit. The microcontroller and motor portion of the circuit will not be simulated.

Block Diagram



Overall Circuit (KiCAD)



Functional Design

Solar panel

Solar panel converts sunlight into DC power with solar cells to charge the battery. We use four 100W monocrystalline flexible solar panels. They can maximize the area for solar energy absorption because of their bending property. They are lighter and cheaper than the rigid solar panels, and the monocrystalline design makes them more efficient than polycrystalline ones. Every solar panel's dimension is 40.9" x 26.8" x 0.1", based on the WindyNation 100W flexible panels.

Regulator

The regulator ensures the battery is charged properly; it is a 48 V 60 A 3000W DC Step Up Converter Voltage Regulator[2]. The regulator has an input voltage range of 18 - 35 VDC, which encompasses the voltage range of the solar panel. This portion of the circuit also includes a diode to ensure no current flows in the wrong direction. The Arduino Uno will monitor the battery to ensure that when it is filled, no more power is fed to the battery; one way to do this would be to step down the battery voltage with a buck converter so that it can be read from the Arduino Analog pins. As stated before, however, this part of the circuit will not be simulated.

Battery Charger

The battery charger is used to charge the battery by converting the power of AC grid electricity into DC power that goes into the battery. The battery charger we ended up using simply consisted of a transformer, AC to DC converter and a linear voltage regulator.

Battery

We use a lithium-ion battery for our solar car. We assume our battery is a simple two-terminal battery. We can provide simultaneous charging and discharging[3] by connecting the charging inputs (the battery charger for AC outlet charging and the regulator for solar panel charging, with diodes to prevent current backflow), the load (the inverter for the motor), and the battery in parallel. We chose a capacity of 8.9kWHr, one of the available choices for the Gem and which allows reasonable drive and charging times. Our battery would be a four-cell, 48V battery (12 4.2V Li-Ion cells in series), which is similar to the e2's battery pack. This would match the e2's 48V motor, so that we wouldn't need to have voltage conversion between the battery and the motor.

Charging the Battery While in Use

The battery will be charged by the solar panels from the solar cell system when the car runs. The driver will be able to monitor charge levels. The solar panels, battery, and engine will be in parallel so you can charge and discharge the battery at the same time. The car can also be charged from the wall outlet when it is not running.

Battery charging considerations:

- The charging current and voltage of Li-Ion batteries should be monitored to be close to the desired wall charger specifications.
- Li-Ion batteries should not be charged at excessively low or high temperatures (lower than 0°C or higher than 45°C).
- Li-Ion batteries should not be overcharged (over 4.2V per cell) or charged with the wrong polarity.
- Li-Ion batteries are usually charged in two stages: at constant voltage until it is nearly full, and then at constant voltage near saturation. For the sake of finishing this project, we are only going to focus on the constant-voltage charging.

Inverter

The inverter is used to convert DC power into 48V AC electricity. We assume our inverter outputs a pure sine wave, which ensures maximum conversion efficiency. This will be used to convert the battery DC voltage to AC voltage that will power the engine/motor. The design simply consists of four diodes connected in a way to create the inverter (more details will be shown in the simulation portion of this report).

Performance Metrics

Power efficiency

We have power transfer in the circuit charging the battery from the solar panels, the circuit charging the battery from the charging station, the DC/AC conversion circuit, and any voltage converter circuits. We may evaluate power efficiency by finding, for each one of these circuits, the ratio of output power to input power. A “good” power efficiency rating for charging would be over 80%, and an “excellent” power efficiency would be over 90%. (For reference, Tesla chargers have a maximum power efficiency of 92%)[5]. We might also expect similar losses in energy from the discharging battery to the engine.

Speed of charging

A typical 120V 60Hz household outlet (not specifically designed for an EV) can deliver roughly 1.4kW max[4]. With an 80% efficiency, we hope to fully charge our 8.9kWHr battery in roughly 8 hours (without any contribution from solar panels, e.g., if parked in a garage). With an “excellent” 90% efficiency, this would be roughly 7 hours.

Similarly, if roughly 80-90% of the battery's power is sent to the engine, then the car may drive for roughly 1.5 hours on a single charge (without any charging from the solar panels). Thus we would have roughly 5 hours of charging time per hour of driving time, which should be reasonable for the average short-distance commuter.

Regulator Calculations

A very simple calculation to determine what kind of voltage regulator to use is as follows: 4 solar panels x 100 W = 400 W, so we should have a voltage regulator that handles this kind of power[3].

Summary of Research

Solar Cell Physics

Solar panels are made from n and p type silicon, i.e. pn junctions[14]. A depletion layer forms between the two types of doped silicon. Photons from the sun excite electrons in the n-doped region and cause more electron-hole pairs to be formed, widening the depletion region and increasing the builtin voltage; the electrons are driven to the top of the n type silicon and holes are driven to the bottom of the p type silicon by the electric field generated by the depletion layer, creating a potential difference. If a load is connected across the solar cell, the potential difference drives electrons from the n type silicon, through the load, and back into the p type silicon, thus delivering electrical power to the load.

If a photovoltaic cell is not connected to a load, the absorption of photons creates more free (excited) electrons and holes (akin to increasing doped-ness of both sides), which causes the widening of the depletion region, and higher drift and diffusion currents (still in equilibrium) for both holes and electrons. If there is an external path connecting the n-doped and p-doped sides with some load (resistance), then the electrons can become un-excited as they travel down that path, and will recombine with holes on the p-doped side. In other words, when an external (Ohmic) path is connected to the photovoltaic cell, it acts as a voltage source with hole-current traveling from the p-doped side to the n-doped side, until all excited electrons return to their ground state (i.e., the photogenerated charge carriers generate a current along the external path).

The Shockley diode equation[15] describes the IV characteristic for a diode and can be used to model the current (I_D) across the photovoltaic cell (which is essentially a pn-junction, or diode):

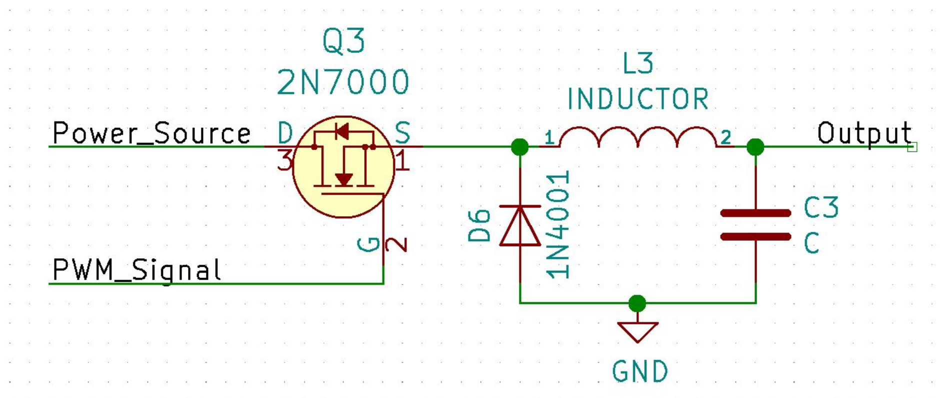
$$I_D = I_S \left(e^{V_D/V_T} - 1 \right)$$

where I_s is the saturation current, V_D is the voltage across the diode, and V_T is the thermal voltage (and all assuming an ideal diode).

Power Electronics

Some basic important power electronics devices used in the circuit are buck, boost, and buck/boost converters. Transformers were also used in the circuit.

The buck converter is used to step down a DC voltage to a lower DC voltage. Below is a diagram of a standard buck converter.

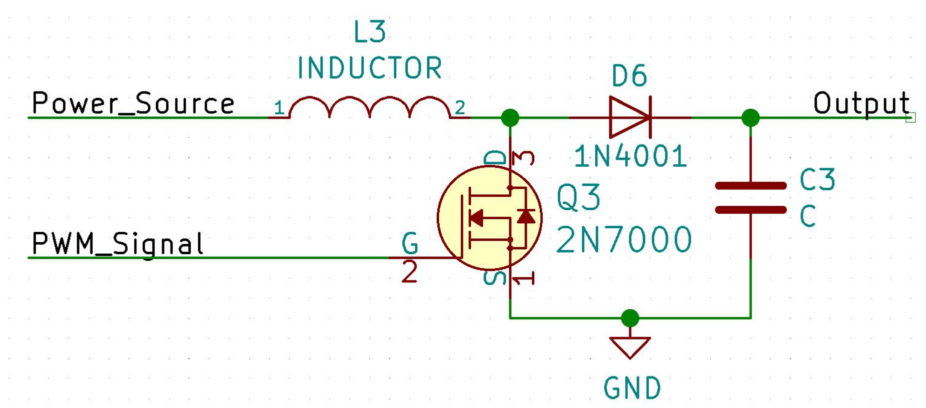


The output DC voltage value depends on the duty cycle of the signal that turns the transistor on and off, and is thus governed by the following equation:

$$V_{OUT} = V_{IN}\frac{t_{ON}}{T}$$

The buck converter is often used in favor of linear voltage regulators as it has less I2R loss. Note that the output does not depend on the values of the components in the circuit, but rather on the input signal. Because the diode needs to handle small voltages and large currents, a schottky diode is often favored for its smaller forward voltage drop[16].

The boost converter is used to step up a DC voltage to a higher DC voltage. Below is a diagram of a standard boost converter:

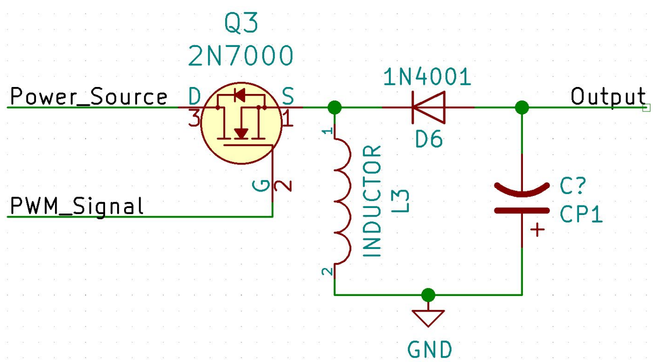


The output DC voltage, similarly to the buck converter, also depends on the duty cycle of the signal that turns the transistor on and off, and is governed by the following equation:

$$V_{OUT} = \frac{V_{IN}}{1 - D}, D = \frac{t_{ON}}{T}$$

The boost converter output voltage also does not depend on the component values. An example application of a boost converter is to maintain a certain voltage for a discharging battery. A boost converter is often controlled by a PWM feedback loop to maintain a target voltage[17]. Note that power is conversed, and thus less current can be sourced from the output.

The buck/boost converter combines the functionalities of the buck and boost converter and allows for fine voltage control. The following is a circuit diagram of a typical buck/boost converter (note that the capacitor is reverse polarized with respect to the input voltage!):



The output DC voltage depends on the duty cycle of the signal that turns the transistor on and off, and is governed by the following equation:

$$V_{OUT} = V_{IN}\frac{-D}{1 - D}, D = \frac{t_{ON}}{T}$$

When the transistor is on, the inductor charges and current is blocked from flowing into the capacitor by the diode; when the transistor is off, the inductor charges the capacitor, which discharges its load to the output. Like the boost and buck converters, this device has a high efficiency and is often used in PWM feedback loops to maintain a constant voltage. An example application would be if you had a battery at 15 V, and needed to power a sensitive circuit that only accepts 12 V. Initially, the voltage needs to be stepped down to 12 V, but once the battery is sufficiently discharged and its voltage falls below 12 V, the voltage must be stepped up. A buck/boost converter would be the ideal device to use due to its flexibility[18].

Batteries

There are numerous types of batteries used in industry; some examples include Lithium-ion, Lead-acid, and Nickel metal hydride batteries. The lithium-ion battery will be discussed in more detail, as it is the battery of choice for this project. The lithium-ion battery consists of an anode and cathode which store the lithium, a separator which allows flow of lithium ions but blocks electrons, electrolyte which acts as the medium through which the ions flow, and two current collectors (positive and negative)[6]. During a discharge cycle, lithium ions flow from the anode to the cathode, and the separated electrons flow through the load and recombine with the lithium ions arriving at the cathode. The ions and electrons both facilitate current, thus creating an electric circuit between the battery and load[7]. When an external power source is applied to the battery, it charges, reversing the previous process. After many charge and discharge cycles, imperfections in the metal that make up the anode and cathode lessen the capacity of the battery[8].

Table of Comparisons Between types of Batteries

Type	Pros	Cons
Lithium-Ion	• High energy density (thus useful in lightweight, small, and high-capacity applications). • Low maintenance (i.e. replenishing acid, periodic discharge, etc). • High cell voltage, requiring fewer cells per battery. • Good load characteristics provide relatively constant 3.6 V per cell until charge falls off[9].	• Requires protection circuitry to prevent being charged or discharged too far. • These batteries age, thus decreasing the number of charge-discharge cycles available. • Not safe for airline transportation. • Relatively expensive[9].
Lead-Acid	• Relatively cheap. • Easily rechargeable. • High power output capability[10]. Can withstand harsh conditions with little to no maintenance[11].	• Very heavy. • Batteries tend to be huge as power density is very low[10]. • Requires maintenance to remove lead sulfate buildup. • Have a limited lifespan. • Some types have risk of fire if shorted[11].

Nickel Metal Hydride	• High power density. • Often come in “standard” sizes. • Easy to recharge[12].	• Self-discharges quickly. • Relatively expensive[12]. • Cannot deliver high load. • Poor low temperature performance[13].

Circuit Implementation

Part Considerations

Overall Considerations

The starting cost of the Gem e2 vehicle is listed as $10,299 on their website. We aimed for an upper limit of roughly $15,000 for our cost, on the basis that we were buying parts individually (as opposed to in bulk).

Desired Battery Properties

There are a number of technical and non-technical aspects to consider when choosing a battery.

Cost	<$8000
Ease of implementation / simplicity of design	made for electric vehicles
Reliability	made for electric vehicles
Capacity	>8kWh
Voltage & current (both charging and discharging) specs	48V DC, 100A discharge current (to saturate motor maximum power output)
Efficiency	>95%
Warranty	A few years, if possible

Some of the choices we found when choosing the battery were:

Cost	Main features	Benefits	Order Link
$8200	8.9 kWh, Li-ion, 44 V	Use a proven electric car battery. Meets specifications we need	https://gem.polaris.com/en-us/shop/accessories/batteries-and-charging/4017366/
$7352	10 kWh, Li-ion (LiFePO4), 48V, 2500 charge cycles, 2 year	Cheaper and higher capacity than e2 8.9kWh battery pack,	https://www.amazon.com/4

	warranty	meets specifications, warranty, made for electric vehicles, internal BMS	8-VDC-Kwh-Battery-Pack/dp/B079348MM1
$3800	10 kWh, Li-ion, 51.2V, 25.6"x18.9"x7.5", 216lb, 4000 charge cycles, 10 year factory warranty, small self-discharge, high temperature range, charge voltage 58.4v, max discharge current 100A, max charge current 50A, LiPO4 cell technology 98% efficiency, operation temperature 32-122 deg. F, storage temperature -4-140 deg F	Cheap, high capacity and charge/discharge speeds, lists golf carts as one of its potential uses, long warranty, internal BMS (battery management system), much more detailed specs	https://www.electriccarpartscompany.com/10KWh-48V-200Ah-LiFePO4-Lithium-Battery-Solar-Energy-Storage-System

Of these three, we decided to go with the latter. Not only was it the cheapest, but it had many of the desired requirements, had the most useful information, and lists electric vehicles as one of its potential uses. We were unsure why this one, which seems so much cheaper than the others, is this cheap when it boasts many similar characteristics, but were unable to find any reasons for this.

Desired Solar Panel Properties

We changed and specified some of our requirements for the solar cars' roof-top panels. Instead of using rigid solar panels, we will use flexible solar panels. They can maximize the area for solar energy absorption because of its flexible bending property. They are lighter and cheaper than the rigid solar panels. The disadvantages are that they are less efficient than the rigid ones, and most of them only have about 100W output. We pick monocrystalline solar panels instead of polycrystalline ones because they are more efficient.

The number of solar panels are determined by the size of the solar panels and the roof area of our 2-passengers eVehicle. We plan to embed the solar panels in the windshield and the roof of the car to maximize the area of radiation absorption, so we can fit approximately four solar panels on the roof based on a 126" x 40" roof and windshield size of a golf car.

We will pick up solar panels based on the following criterias: work output, cost, solar cell efficiency, size, power ratings, power tolerance, and temperature coefficients (in the sequence from the most important to the least important). Solar cell efficiency decides how much electricity can be generated in a certain size of the solar panel. It is one of the most important criteria when we pick up solar panels. The efficiency for monocrystalline is normally between 18% and 25%. Power rating is the amount of power in DC form that can be produced under ideal lab conditions. Power tolerance, which is the difference between the power produced by the solar panels in reality and its nameplate rating. For example, a 250-watt panel with a +/= 5% power tolerance has the power output range from 237.5W to 262.5W. The more concentrate of the range, the better because it means more certainty for the output. Temperature coefficient quantifies how a panel's power capacity decreases as the temperature is higher than the standard temperature (77°F) the tests are performed. Panels with less temperature coefficients perform better in the long run.

Cost	<$500
Solar cell efficiency	between 18% and 25% (but higher is better)
Power tolerance	+-5%

The choices we encountered are shown below. We ended up choosing the WindyNation 100W flexible panels. We would order four of them, for a total of $434.

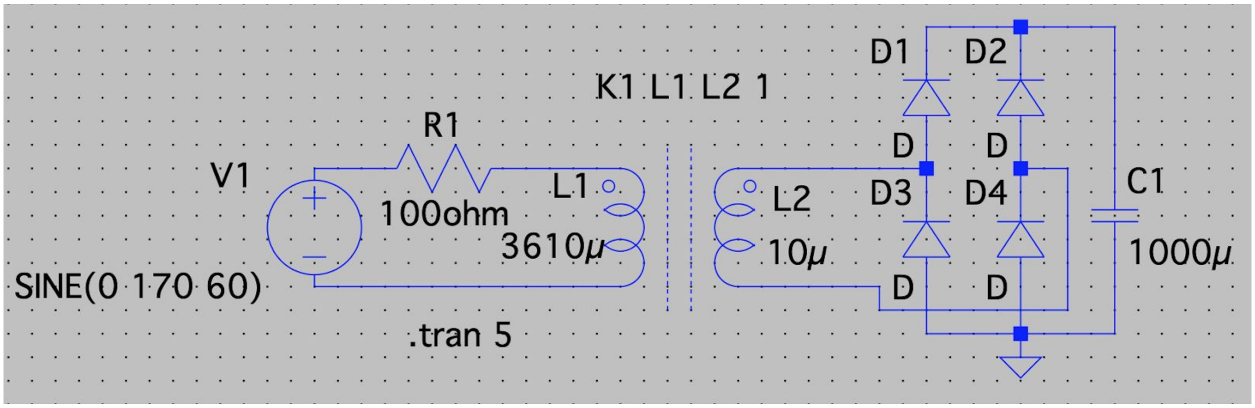
Solar Panels	Cost/panel	Reasons for choosing/ not choosing
WindyNation 100W flexible panels	$108.5	It's cheaper than other 100W flexible panels, and it is weatherproof
Giaride 100W flexible panels	$123	More expensive than the WindyNation 100W one
RichSolar 80W flexible panels	$106	Although its price is lower than the WindyNation one, it can produce less work output, which may drag the charging rate down too much.

Battery charger (Voltage rectifier)

We decided to design a simple smoothing voltage rectifier as part of the battery charger circuit. This features a transformer (AC voltage step-down), a bridge rectifier, and a smoothing capacitor to transform the 120VAC wall plug current to roughly 50VDC current to charge the battery. As mentioned before, we hoped to be able to harness the full 1.4kW maximum power output of a standard wall plug, so we had to look for high-power components when looking for components, which was slightly problematic sometimes (it was difficult to find such high power components, a problem that would come up again when designing the circuit on LTSpice).

Battery charger schematic

The schematic did not change from the first part of the project.



Battery charger component list

(Everything in the following table has quantity 1.)

Component	Cost	Specifications	Justification	Order Link
Step-down transformer	$310	115VAC to 56VAC (tapped for additional voltages as well), 1400VA (kW) power rating, 38 lb, 162x181mm	High power rating allows us to use full wall outlet wattage (1.4kW)	https://www.digikey.com/product-detail/en/signal-transformer/56-25/56-25-ND/1984771
Bridge rectifier	$42	400V, 60A, 4-pin	High power ratings and current ratings, should be safe (PIV should be accounted for)	https://www.sager.com/m5060sb400-4157601.html
Smoothing capacitor	$1.52	1000uF, 160VDC, aluminum electrolytic	Common capacity for smoothing capacitors, is rated for battery's voltage	https://www.mouser.com/ProductDetail/EPCOS-TDK/B43254D1108M000?qs=sGAEpiMZZMsh%252B1woXyUXj4oSl%252BFvVmKmmKDMFjfMzUE%3D

Other components

(Everything in the following table has quantity 1.)

Component	Price	Purpose	Features	Justification	Order Link
Voltage Regulator	$500	Regulate voltage from Solar Panel to power the battery	Input range from 18 to 35 V, rated for 0 - 60A, 3000W rating, efficiency up to 95%, over current and voltage protections	Need this part to regulate unstable output from solar panels, choose over buck converter to avoid a feedback loop, this part has the specs we need.	https://www.daygreen.com/products/24v-36v-to-48v-60a-3000w-dc-dc-step-up-converter-voltage-regulator
Arduino nano	$22	Control buck converters and motor controller via PWM, create feedback loops to properly maintain constant voltage	32 KB Flash, 16 Mhz, 8 Analog IN, 22 Digital IO, 6 PWM pins, input 7-12 V	Easy to use and integrate into design, need microcontroller to control voltage levels of buck converter output and for feedback control. Also use for motor controller.	https://www.digikey.com/products/en?mpart=A000005&v=1050
Voltage regulator	$34	Regulate voltage from smoothed rectifier circuit	Input 10-60VDC, output 12-90V (adjustable), 1500W maximum power rating, 130x52x85mm, conversion efficiency 92-97%, over current protection, low voltage protection, input	High power rating, voltage adjustable and in range, high efficiency, low cost, good built-in protection	https://www.amazon.com/Voltage-Converter-DROK-Regulator-Transformer/dp/B076TTBKFG

			reverse polarity detection		
DC-DC Converter	$12	Step down voltage of the battery to power the arduino	Voltage input range of 43.2 V - 52.8 V, Voltage output of 9 V, 2 W power capability, 82% efficiency	Need to step down voltage to Arduino for a reasonable price and reasonable efficiency, the Arduino won't need much power so this simple DC to DC converter will suffice. Meets specifications necessary.	https://www.digikey.com/product-detail/en/xp-power/IL4809S/1470-2442-5-ND/4785548

Overall design cost

Component(s)	Cost
Battery	$3800
Solar panels	$434
Voltage regulator	$354
Other components	$545
Total	$5163

This is much lower than the $15,000 original upper limit on cost, but there is still a large possible error in this estimate, given that we have no way to test these components to see if they are a good fit. For example, this cheap battery may have some unknown disadvantage as compared to the OEM e2 battery (which cost $4400 more) that we cannot know without further testing or inquiry. However, this cost estimate is reasonable given that the total e2's base configuration's cost is $10,299 and the charging subsystem is only one of various systems in the car (the engine and chassis are likely to contribute to another large portion of the car's cost).

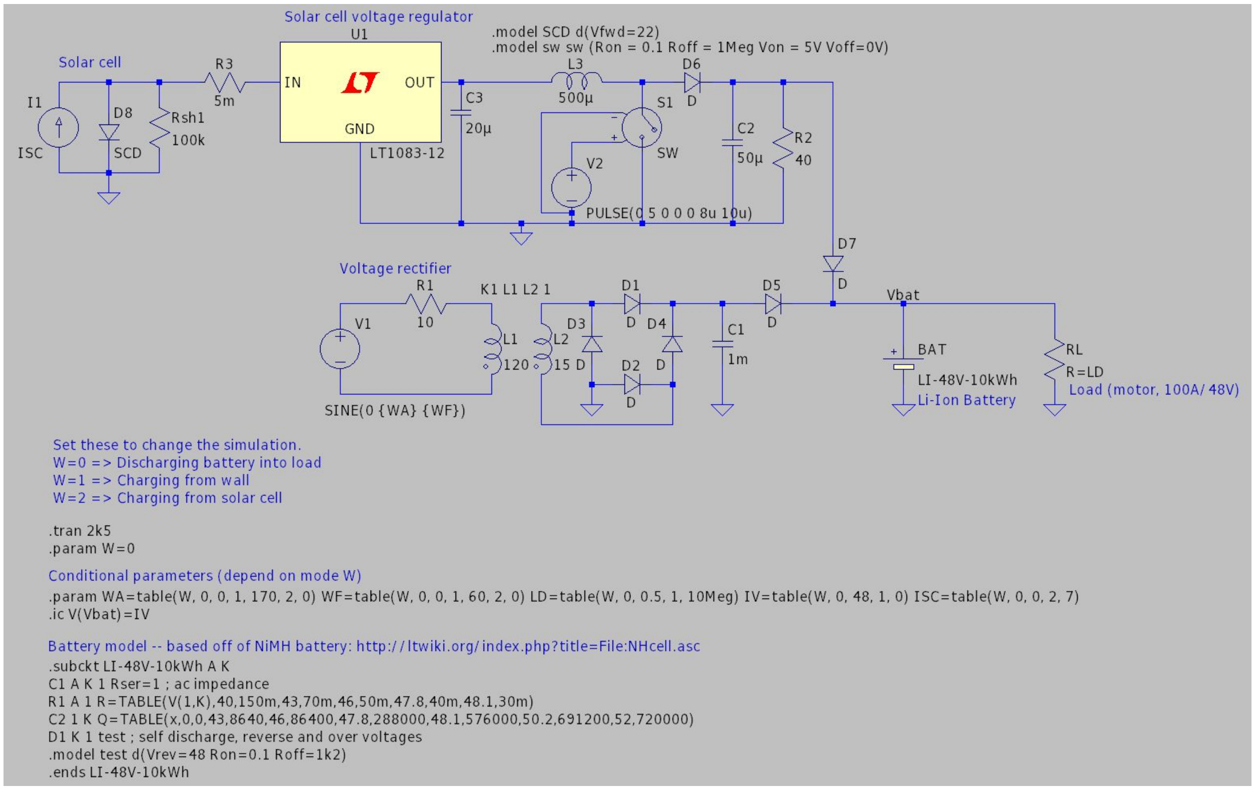
Simulations

We simulated the following blocks:

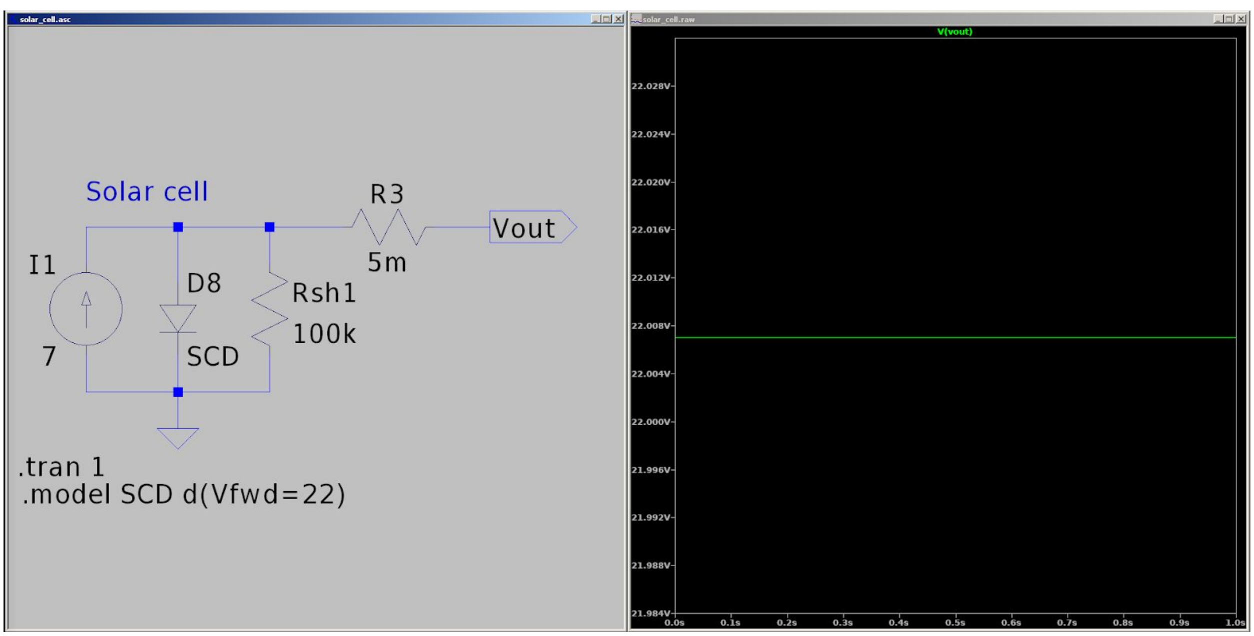
- The solar cell
- A 12V solar cell voltage regulator (IC model from LTSpice)
- A 12V to 48V boost converter from the solar cell voltage regulator
- The AC-DC voltage step-down/rectifier circuit
- A 48V 10kWh Li-Ion battery

We put everything into one schematic and tried to simulate it, but it turned out terribly slow to simulate (<1/1,000,000 simulation/real time ratio). The overall schematic is shown first, followed by the individual components.

Overall Schematic

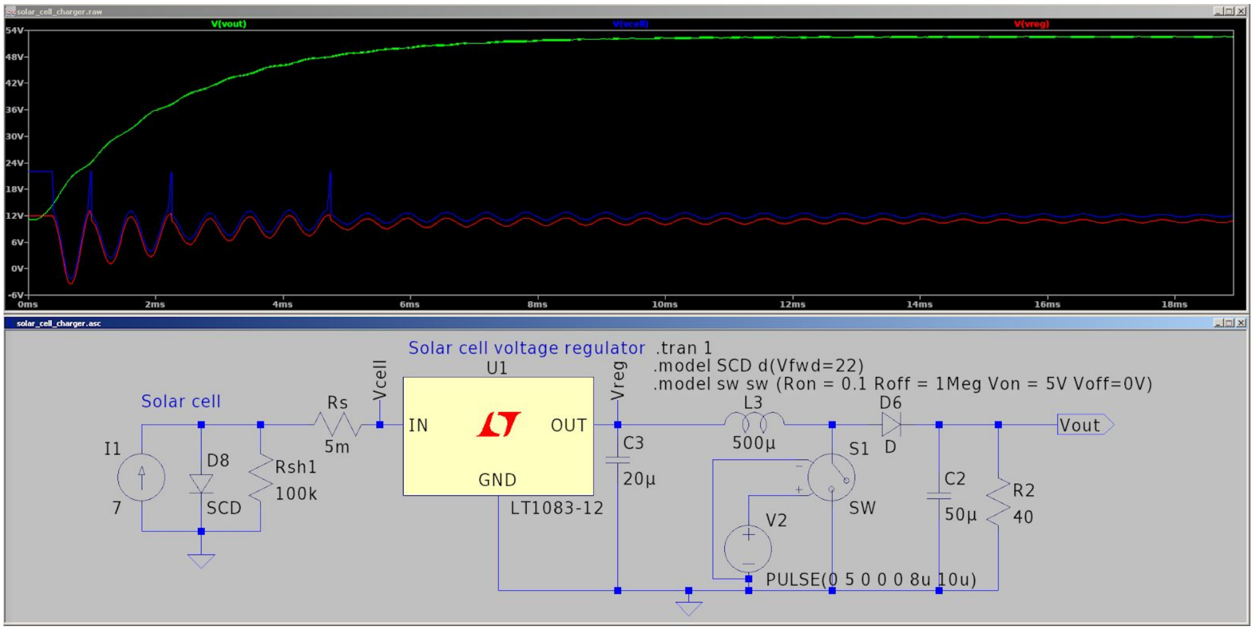


Solar Cell



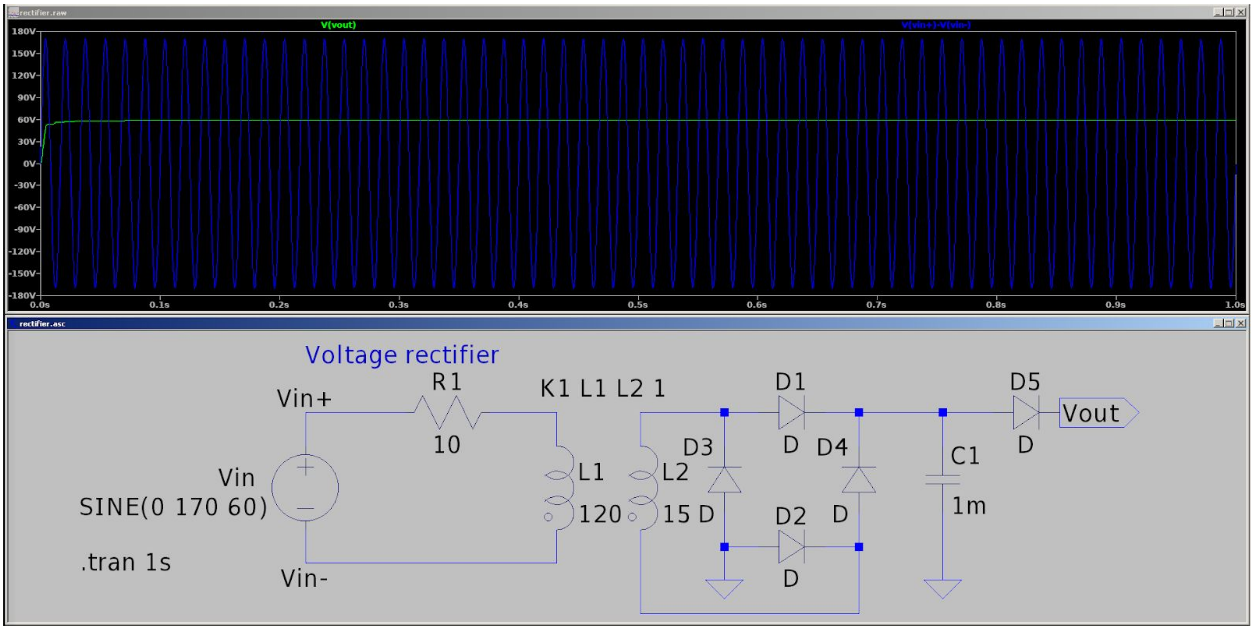
This solar cell was supposed to roughly match a 6V, 22Voc, 12Vout, 100W solar cell. While many sources online suggested this model of a solar cell (and this was the same solar cell schematic we found in the research stage of this project), we were unable to find good sources on how to calculate reasonable values for the resistors and characteristics of the diode for this circuit, so this is largely based on trial and error. By making the forward voltage of the diode 22V, this made the open-circuit voltage Voc=22V.

Solar Cell Battery Charger



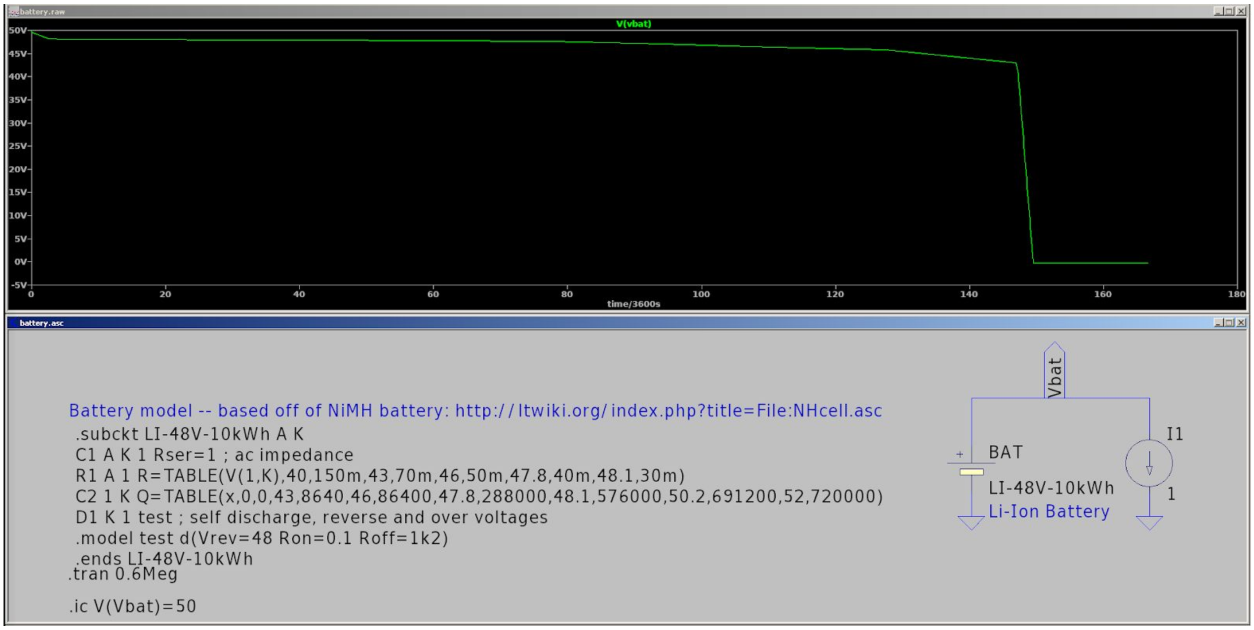
The voltage regulator after the solar panel is a 12 V linear voltage regulator and a 1:4 boost converter (75% duty cycle) to boost the voltage up to 48 V. We looked for spice models that fit the specifications of the model we found (10 - 60 V input range, 1500 W power rating) but were unsuccessful, so we had to use this combination of voltage regulator and boost converter. You can see the voltage regulator and boost converter cause the voltage at Vcell and Vreg to oscillate, but the oscillations decrease over time and the output voltage steadies around 50V, as desired.

Wall Charger Voltage Stepdown/Rectifier

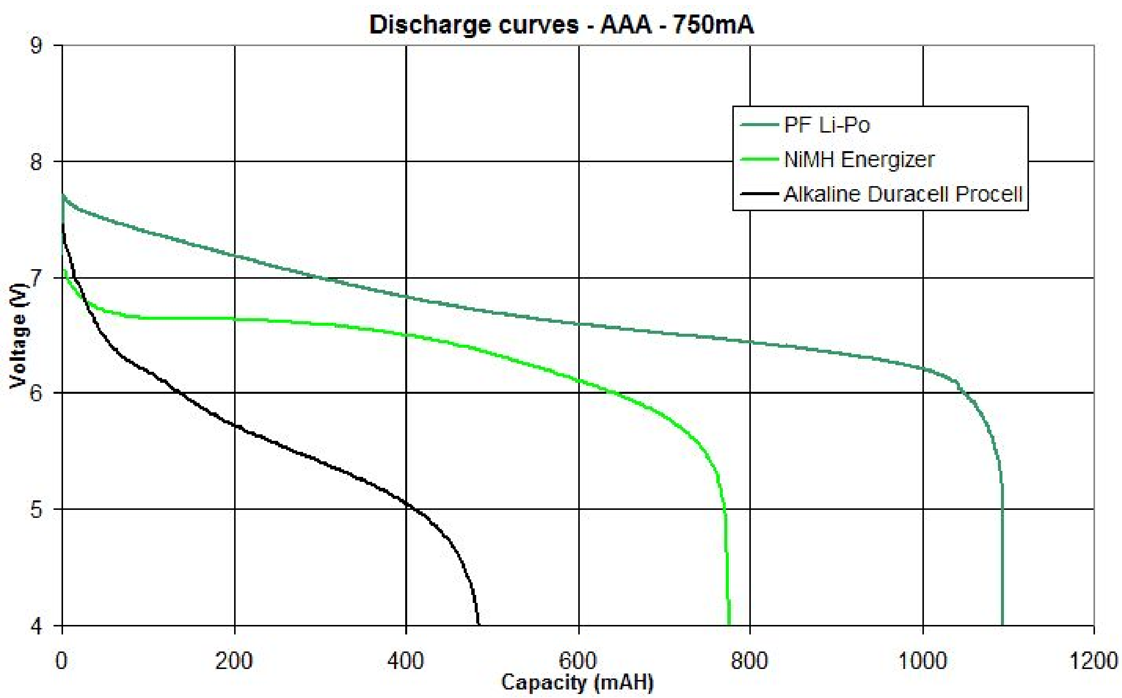


The principles behind this circuit are very similar to the bridge rectifier from class, and it works nicely to step down and rectify the wall current to the correct voltage. The diodes are left to the default model, and the capacitor and the “transformer” (modeled using inductors) have values chosen somewhat by trials and errors to get the expected result.

Li-Ion Battery

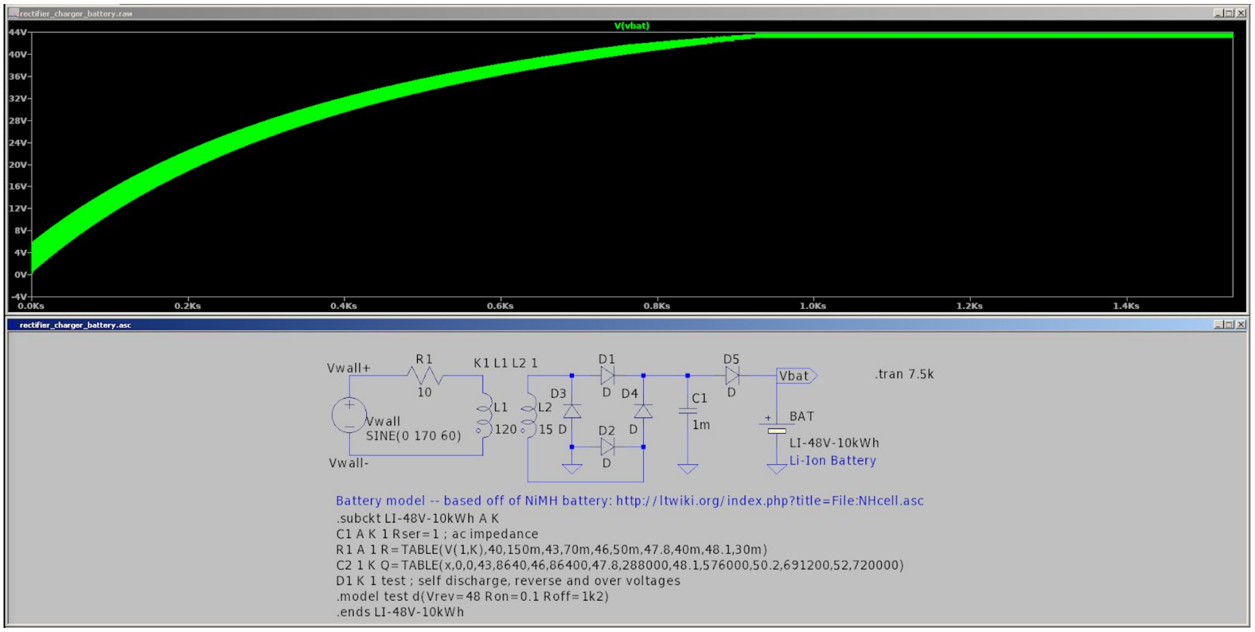


This was adapted from a 1.5V, 2.5Ah NiMH (nickel-metal hydride) battery model (from http://ltwiki.org/index.php?title=File:NHcell.asc). The model essentially hardcodes the charge based on the cell voltage. Looking at the voltage vs. charge graph for li-ion batteries vs. NiMH batteries, we see that the curve is not too different.

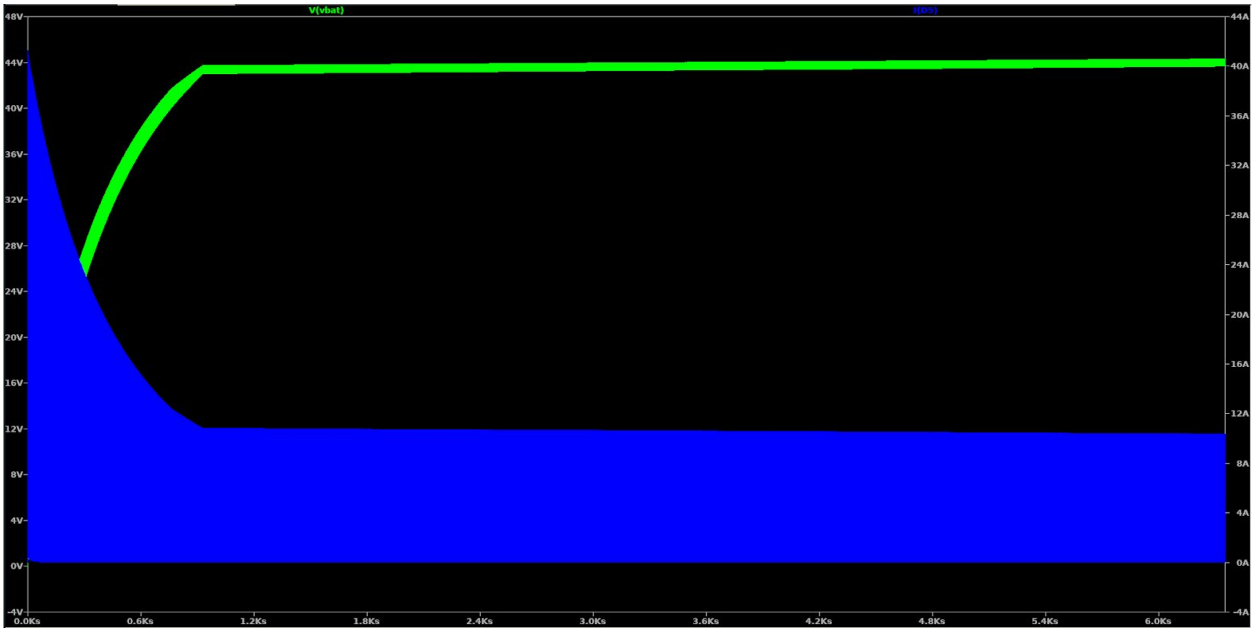


By tweaking the discharge rate (changing the resistor value) and diode voltage, we were able to make this act like a 10kWh (or roughly 200Ah), 48V battery. (This was verified by integrating the area under the current vs. time curve.) This was very much done empirically, and comments on how to improve this model would be welcome. The simulation shows that this battery, discharged at a rate of 1A, keeps its charge for roughly 150 hours, so it is actually around 150Ah (instead of the desired 200Ah), so this should require some tweaking to fix. Also, when charging (next section), it doesn't seem to reach exactly 48V, but is a little lower (~44V).

Charging the Battery with the Wall Charger



Once the battery is connected to the charger, there is a lot of oscillation in the output (battery voltage). You can clearly see the battery charging up slowly (as the battery's voltage is monotonic increasing w.r.t. its charge), and it reaches roughly its intended voltage around 900s. (This is the "plateau" of the battery's voltage-charge curve, so it still has a long way to charge.) Simulating for a longer time and also viewing the current through D5 (in blue):



You can see that there is a huge oscillation in the current through D5 (from 0-10A), which is probably not very efficient. Also, if we say that the average charging current after the voltage reaches the plateau is 5V, then this is only charging at roughly 44V*5A=220W, which wouldn't take into account the full 1400W of the wall voltage. We might need to add an additional regulator or find some other way to reduce oscillations.

Because of the higher modeling complexity of the solar cell charger (with the voltage regulator IC and boost converter), we were unable to get any useful simulation because the simulation was too slow to run.

Conclusions

Overall, the project was a useful exercise in thinking about what desired specifications (technical and nontechnical) should be met to fulfill a project, how to pick the components intended to fit the said specifications (as best as possible), how to design the basic circuitry for a photovoltaic system, and how to run proper simulation for various parts of the circuit. Finding the components that fit our specifications was aided by the filters available on different sites, but a deeper analysis into the datasheet information should be done before ordering parts. In addition, a more thorough cost analysis should be done before finalizing the bill of materials, as sometimes our group simply chose the first component that fit the set specifications. LTspice was a useful tool to use, but some problems hindered our progress, such as its slow simulation rates and its inability to simulate ICs. Finding the exact models for the components chosen in the LTspice libraries or some other online sources was almost impossible, and we were forced to create the components from the existing models or to tweak some properties from the online examples. The circuitry also required high voltage and high current rated products, and this hindered our search for the proper models to simulate.

Future Work

One aspect of this project that was frustrating was the lack of real simulations. While it would be infeasible to simulate power specifications on the power levels as those for a real electric car, even a small-scale simulation might dramatically help our understanding of the matter, even if not entirely accurate. This would aid mostly with simulations (as real life doesn't lag like LTSpice does in a simulation) and learning how some of the theorized parts might be better or worse than expected. Additionally, this could help us test how well our simplistic solar cell model works by comparing its output to that of a real solar cell.

Another aspect of this project that was left mostly unfinished was efficiency calculations. Because of the inability to satisfactorily simulate the entire circuitry in LTSpice, we could not obtain much useful information from the simulation. Nor were we able to calculate efficiency of the system manually due to the complexity of the system and the unknown properties of the battery and other blackbox components used such as voltage regulators.

If this project were to be continued in the future, a more robust circuit design would have to be created; one that optimized cost and included specifications for connections and properly accounted for input fluctuations. Designs for a potential custom printed circuit board or wiring diagram would be required to properly interface the different off the shelf components together. In addition, a more thorough size analysis would be required to gauge how the circuit would be mounted on the electric car.

Another way to expand the project would be to interface the battery charging circuit with the motor and microcontroller part of the car, as was originally intended for this project. Extensive testing on each component ordered and the circuitry built would be required to ensure safety and to better gauge how the different components interact with each other. Of course, funding would also have to be secured to purchase the off the shelf components and to prototype the circuitry and mechanical elements of the design.

Personal Takeaways

Josh - I enjoyed working with a team to do research and find components for a project that deals with a growing and important technology. I appreciated the chance to delve deeper into LTspice simulations and the intricacies of solar cells and power electronics, and feel that I can better search for parts and design circuitry. I hope to apply these skills to future projects, even if they don't involve solar cells.

Jon - Throughout most of this project I felt really lost. I don't have much experience with hardware or choosing components or knowing what to look for in components, and diving into it headfirst with a team and guidance from the TAs is very valuable. I remember last year in Motorsports when we were supposed to look for components and I was completely lost — now I feel (slightly) more confident in knowing what to look for.

Min - I feel a lot motivated working in a team, and I really enjoy our teamwork. Although this is just a simulation project, I figured out some research methods and testing steps that I was confused about before in some other projects. I didn't really understand the reason for spending time comparing different versions of components to finish the budget list previously, but after this project I found it critically related to the designing and testing standards we set and the realization of our design. I also learnt a lot from my teammates in time management and the habit of always doing an objective final check. I think I will approach all my future projects with all these skills confidently.

Works Cited

1. https://www.clippercreek.com/wp-content/uploads/2017/12/SMUD_Charge-Times-Chart-20171208_Final_Low-Res.pdf
2. https://www.daygreen.com/products/24v-36v-to-48v-60a-3000w-dc-dc-step-up-converter-voltage-regulator
3. https://electronics.stackexchange.com/a/79108/195122
4. https://www.clippercreek.com/wp-content/uploads/2017/12/SMUD_Charge-Times-Chart-20171208_Final_Low-Res.pdf
5. https://forums.tesla.com/forum/forums/charging-efficiency-0
6. https://www.energy.gov/eere/articles/how-does-lithium-ion-battery-work
7. https://www.cei.washington.edu/education/science-of-solar/battery-technology/
8. https://www.youtube.com/watch?v=9OVtk6G2TnQ
9. https://www.electronics-notes.com/articles/electronic_components/battery-technology/li-ion-lithium-ion-advantages-disadvantages.php
10. https://learn.adafruit.com/all-about-batteries/lead-acid-batteries
11. https://www.envyride.com/lithium-ion-vs-lead-acid/
12. https://learn.adafruit.com/all-about-batteries/ni-mh-batteries-nickel-metal-hydride
13. http://www.urbanfox.tv/articles/batteries/b1batteries.html
14. https://www.popularmechanics.com/technology/infrastructure/a28186403/how-solar-panels-work/
15. https://en.wikipedia.org/wiki/Shockley_diode_equation
16. https://components101.com/articles/buck-converter-basics-working-design-and-operation
17. http://www.ti.com/lit/an/sloa127/sloa127.pdf
18. https://www.youtube.com/watch?v=ZiD_X-uo_TQ

ECE241 – Electronics I

Jonathan Lam

March 1, 2020

Contents

1 Introduction to Electronics

1.1 Basic Concepts

distortion when a signal's output is not a linear function of input (i.e., non-linear transfer function, signal output shape different from input shape)

analog signals difficult to process b/c of noise, distortion, difficult to store because require an analog storage method (e.g., capacitors)

many digital signals still have to be processed as analog signals (e.g., when retrieved from storage) before becoming a robust digital signal, as analog signals are the raw form and digital signals are an abstraction

analog functions include amplification (characterized by gain, speed, power distribution, (practical) bandwidth), filtering

KCL related to conservation of charge; KVL related to conservation of energy

2 Semiconductor Physics

2.1 Basic Concepts

band gap an energy range in which electrons cannot exist (i.e., between valence shell and conduction layer)

band gap energy the difference between top and bottom of band gap energy; i.e., the energy required to move an electron into the conduction layer from the valence shell

2.2 Transport of Carriers

drift movement of charge carriers due to an E-field; since E and q are both signed, both equations (for holes and electrons) are positive; drift current proportional to E and Anq (charge density per unit cross-sectional area) with proportionality constant μ_n (mobility)

diffusion movement of charge carriers due to an uneven charge distribution; for positive charges/current, moves in opposite direction of gradient, so sign is negative (positive for electrons b/c of sign of q); proportional to q, $\frac{dn}{dx}$, with proportionality constant D_n (diffusion constant)

3 pn Junction

3.1 pn Junction in Equilibrium

diffusion current diffusion current from p to n side

depletion region ions without free charges build up in the center region

E-field E-field from n to p side

equilibrium constraints on currents $|I_{dr,p}| = |I_{di,p}|$, $|I_{dr,n}| = |I_{di,n}|$

4 Carrier modeling in equilibrium

band gap The difference in energy between the conduction band and the valence band. In metals, it may be zero; in insulators, it is usually high; in semiconductors, it is small.

5 Symbols, Equations, and Units

electron volt (eV) 1eV $= 1.6 \times 10^{-19}$J (i.e., since $V = \frac{J}{C}$, $1V = \frac{1eV}{1e}$, and $e = 1.6 \times 10^{-19} C$) Brief review of physics because dumb:

$$\vec{F} = q\vec{E}$$

$$\int \vec{F} \cdot d\vec{s} = W,\ \vec{F} = -\nabla U$$
$$\int \vec{E} \cdot d\vec{s} = V,\ \vec{E} = -\nabla V$$
$$-U = W = -qV$$

(Voltage is EPE per unit charge, electric field is electric force per unit charge. Integrate force to get energy, or field to get potential. The electric potential is the negative of the work done by the field; electrical potential energy and electric potential have the same sign.)

band gap energy (E_g) for Si @ room temp., $E_g = 1.12\text{eV}$

mass action law $np = n_i^2$

intrinsic electron density (n_i) for Si, $n_i = 5.2 \times 10^{15} \exp \frac{-E_g}{2kT} \frac{e^-}{\text{cm}^3}$; in intrinsic material, $n \approx n_i \approx p$

negative charge carrier density (n) n-type: $n \approx N_D$; p-type: $n \approx \frac{n_i^2}{N_A}$

positive charge carrier density (p) p-type: $p \approx N_A$; n-type: $n \approx \frac{n_i^2}{N_D}$

current density $I = JA$

drift current density $J_{dr} = q(\mu_n n + \mu_p p)E$

mobility saturation $\mu(E) = \frac{\mu_0}{1+bE} \Rightarrow v \approx \mu_0 E$ at low E, $v \approx \frac{\mu_0}{b}$ at high E

diffusion current density $J_{di} = q\left(D_n \frac{dn}{dx} - D_p \frac{dp}{dx}\right)$

Einstein relation $\frac{D}{\mu} = \frac{kT}{q}$; at 300K, $\frac{kT}{q} \approx 26mV$

anode where current flows into device; ACID (Anode Current Into Device); for batteries negative terminal; for passive devices positive terminal

built-in potential of a diode $V_0 = \frac{D_p}{\mu_p} \ln \frac{p_p}{p_n} = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2}$

effective mass $\vec{F} = -q\vec{E} = m_0 \frac{dv}{dt}$ (but in a whole semiconductor crystal, use define masses m_n^* and m_p^*); allows treatment of carriers as quasi-classical particles

state densities just above the conduction band, the density of electrons is

$$g_c(E) = \frac{m_n^* \sqrt{2m_n^*(E - E_c)}}{\pi^2 \hbar^3}$$

and similarly below the valence band for the density of holes:

$$g_v(E) = \frac{m_p^* \sqrt{2m_p^*(E_v - E)}}{\pi^2 \hbar^3}$$

These are symmetric, albeit with a different constant.

Fermi function The Fermi function specifies how many of the existing states at the energy E will be filled with an electron:

$$f(E) = \frac{1}{1 + \exp \frac{E-E_F}{kT}}$$

$\lim_{T\to 0} f(E)$ is a the step function $1 - H(E_F)$. If T is a positive finite number, then it approaches a smooth curve with $f(F_E) = 1/2$. Increasing temperature makes the slope shallower (i.e., widening out the probability curve). The function is rotationally symmetric about the point $(F_E, 1/2)$; thus if the probability that a conduction state is full is equal to the probability that a valence state is empty, then $E_F = \frac{E_C+E_V}{2} = E_I$ (the midgap). We can make the approximations:

- If $E \geq E_F + 3kT$ (if E sufficiently higher than the Fermi level), then $f(E) \approx \exp -\frac{E-E_F}{kT}$ (the 1 becomes insignificant, and it acts like a negative exponential)
- If $E \leq E_F - 3kT$, then $f(E) \approx 1 - \exp \frac{E-E_F}{kT}$ (thus, as E decreases, then $1 - f(E) = \exp \frac{E-E_F}{kT}$ drops off exponentially, acting like a negative exponential)
- At room temperature, $3kT = 0.0777\text{eV}$ is quite small, so this approximation can often be used.

To summarize, approximations:

$$f(E) \approx \exp -\frac{E-E_F}{kT}, \;\; E \geq E_F + 3kT$$

$$1 - f(E) \approx \exp \frac{E-E_F}{kT}, \;\; E \leq E_F - 3kT$$

Note that the former models electron density in the conduction band, and the latter models hole density in the valence band. Note that, visually, the number of the negative carriers will be greater than the number of positive carriers if $E_C - E_F \leq E_F - E_V$, and vice versa; i.e., the band the Fermi level is closer to will have a higher density of charge carriers.

carrier densities The product of the state densities and Fermi probability function at any energy level gives the number of carriers at that energy level. Integrating, we get

$$n = \int_{E_C}^{E_{top}} g_c(E) f(E)\, dE$$

$$p = \int_{E_{bottom}}^{E_V} g_v(E)(1 - f(E))\, dE$$

To simplify this, we define the Fermi-Dirac integral of order 1/2, and other substitutions:

$$F_{1/2}(\eta_x) = \int_0^\infty \frac{\eta^{1/2}\, d\eta}{1 + e^{\eta-\eta_x}}$$

$$\eta = \frac{E - E_c}{kT}, \ \eta_c = \frac{E_F - E_c}{kT}, \ \eta_v = \frac{E_v - E_F}{kT}$$

$$N_C = 2\left(\frac{m_n^* kT}{2\pi\hbar^2}\right)^{3/2}, \ N_V = 2\left(\frac{m_p^* kT}{2\pi\hbar^2}\right)^{3/2}$$

then

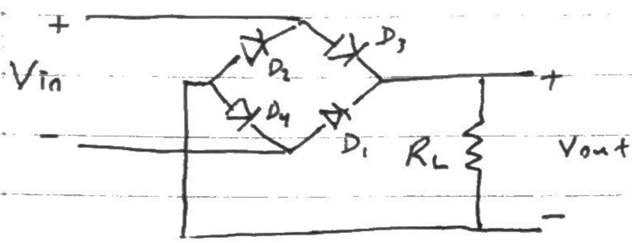
$$n = N_C \frac{2}{\sqrt{\pi}} F_{1/2}(\eta_c)$$

$$p = N_V \frac{2}{\sqrt{\pi}} F_{1/2}(\eta_v)$$

These are the general result (since this comes from an integral of the product of the unapproximated Fermi integral). Thus we use the same approximations when the semiconductor is nondegenerate (i.e., when $E_v + 3kT \leq E_F \leq E_c - 3kT$)

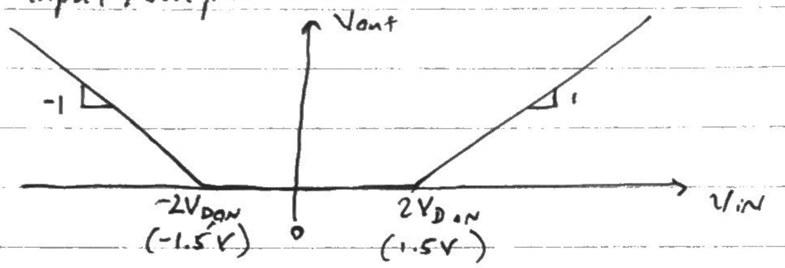
Jonathan Lam
Prof. Shlayan
ECE 241
Electronics I
4/20/20.

Quiz 2.

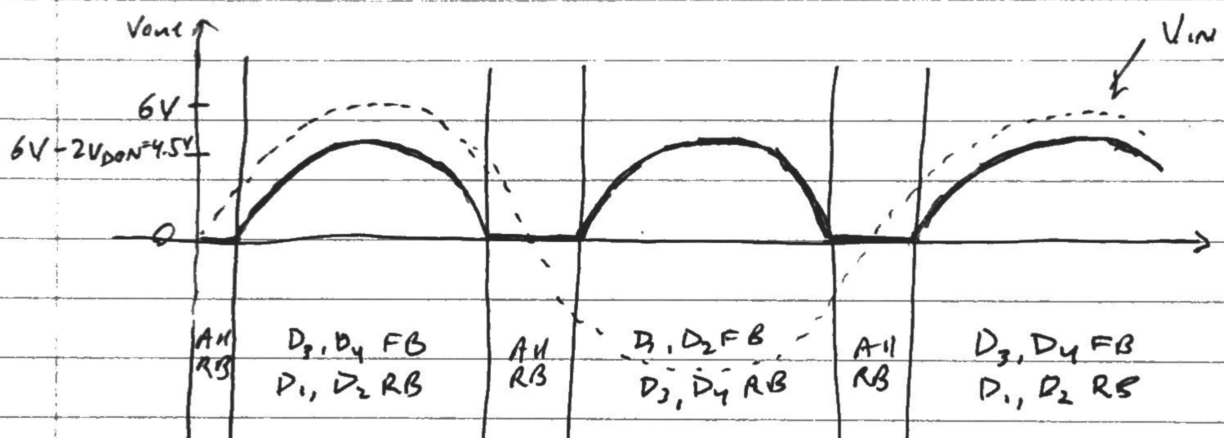


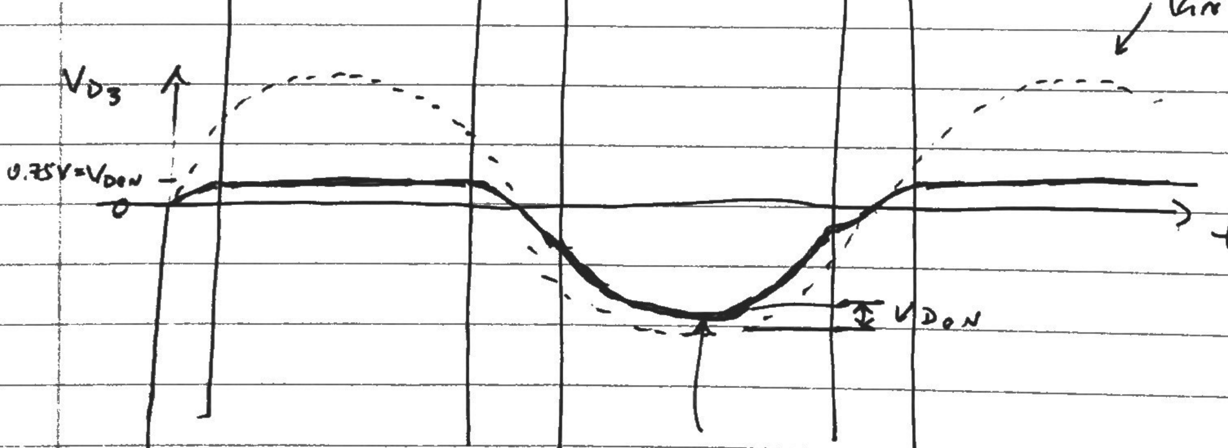
$V_{DON} = 750mV$
$V_{IN} = V_p \cos 2\pi f t$
$V_p = 6V$
$f = 60Hz.$

a) Input/output characteristics



Vout vs. time





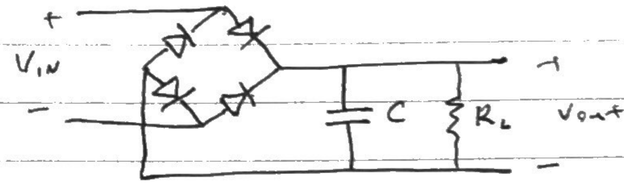
$PIV = -6V + V_{DON} = -5.25V.$

Jonathan Lam
Prof. Shlayan
ECE241
Electronics I
4/20/20.

Quiz 2.

b) $V_{r,max} = 0.4V$, $R_L = 50\Omega$, add smoothing cap. in parallel to load,

$$V_r = \frac{I_L}{Cf}$$

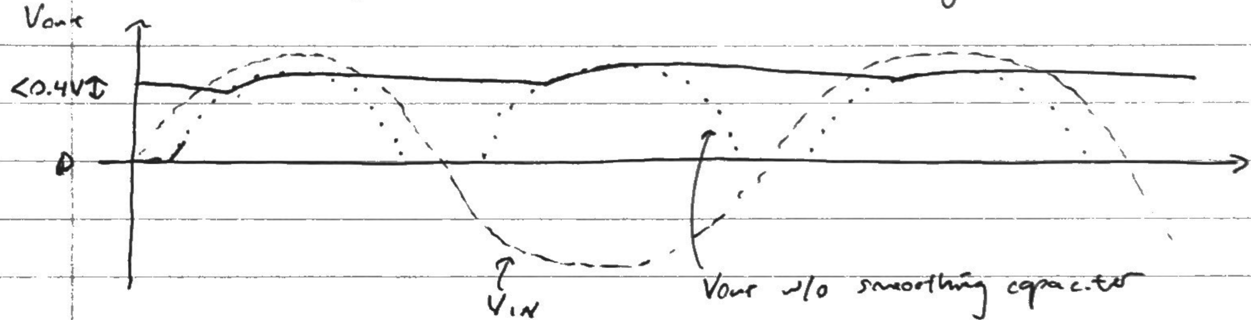


$f = 2 \times 60Hz = 120Hz$

full wave rectifier doubles ripple freq.

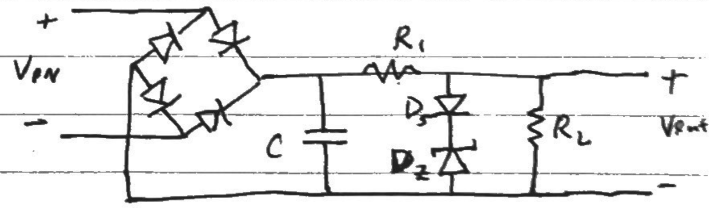
$$I_L = \frac{V_L}{R_L} = \frac{V_p - 2V_{D,ON}}{R_L} = \frac{4.5V}{50\Omega} = 90mA$$

$$C = \frac{I_L}{V_r f} = \frac{90mA}{(0.4V)(120Hz)} = 1.875mF \text{ (or greater)}$$



c) $R_1 = 100\Omega$, D_s : $V_{D,ON} = 800mV$, D_z : $V_{Dz} = 3.2V$, $r_{dz} = 5\Omega$.

Calculate line & load regulation.



Line regulation:

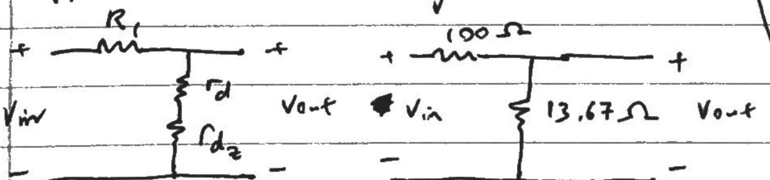
assume small signal model

(center of ripple)

$$I_D = \frac{4.3V - 4V}{100\Omega} = 3mA$$

$$r_d = \frac{V_T}{I_D} = \frac{0.026V}{0.003A} = 8.67\Omega$$

approximate w/ small signal model:



$$V_{out} = \frac{13.67\Omega}{(13.67+100)\Omega} V_{IN}$$

$$\Delta V_{out} = \frac{13.67}{113.67} \Delta V_{IN}$$

$$\text{Voltage line regulation} = \frac{\Delta V_{out}}{\Delta V_{in}} = \frac{13.67}{113.67} \times 100\% = 12.03\%$$

Load regulation: (assume 50Ω load)

$$I_{R_L} = -\Delta I_{D,loaded} = \frac{V_{regulated}}{R_L} = \frac{4V}{50\Omega} = 80mA$$

$$\Delta V_{loaded} = (\Delta I_{D,loaded})(r_d)$$

$$= (-80mA)(13.67\Omega)$$

$$= -1.09V$$

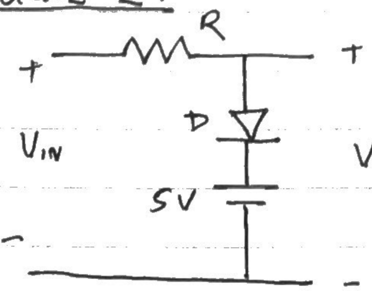
$$\text{load regulation} = \frac{\Delta V_{loaded}}{V_{regulated}}$$

$$= \frac{-1.09V}{4V} = -27.3\%$$

Jonathan Lam
Prof. Shlayan
ECE 241
Electronics 1
4/20/20.

Quiz 2.

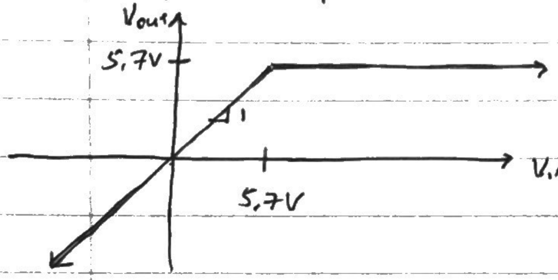
2a)



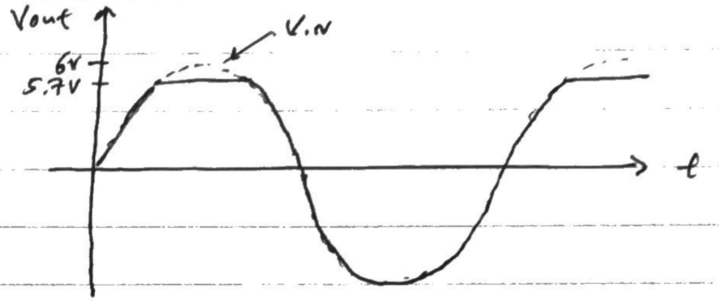
$V_{D,ON} = 0.7V$

$V_{IN} = V_p \cos \omega t$, $V_p = 6V$

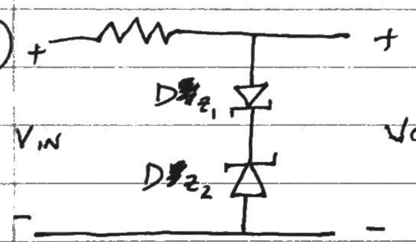
input/output characteristic:



Vout vs. time



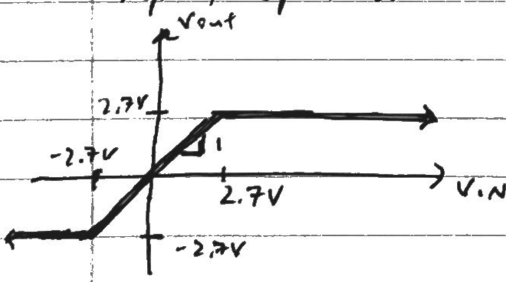
2b)



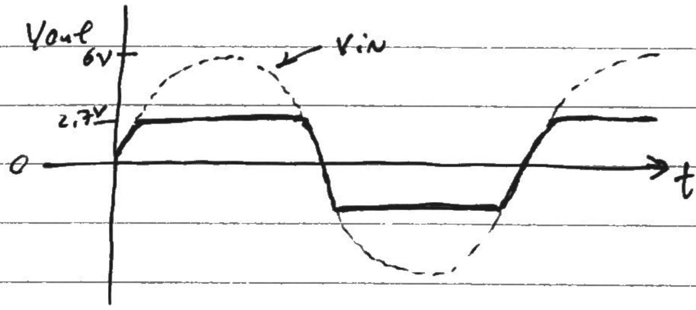
$V_Z = 2V$

V_{in} = (same as above)

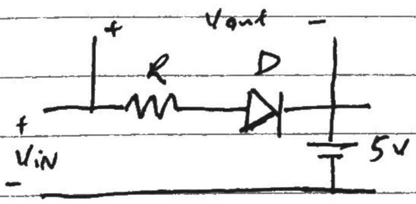
input/output characteristic:



Vout vs. time:

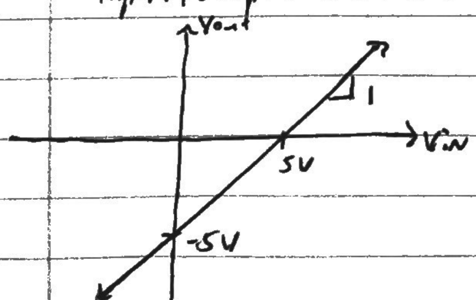


2c)

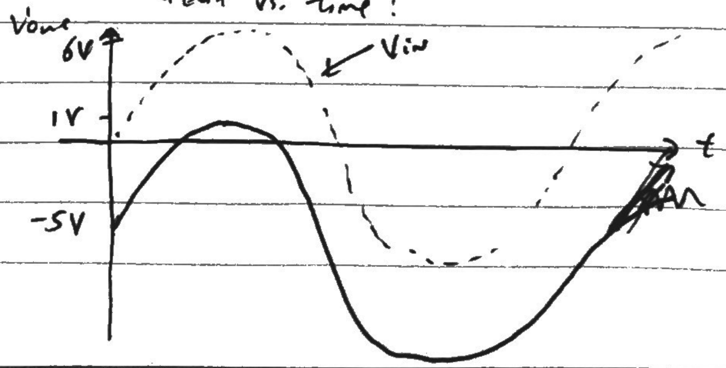


Vout always just $V_{in} - 5V$

input/output characteristic:



Vout vs. time:



ECE241 – Quiz 3

Jonathan Lam

May 15, 2020

BJT Voltage Amplifier

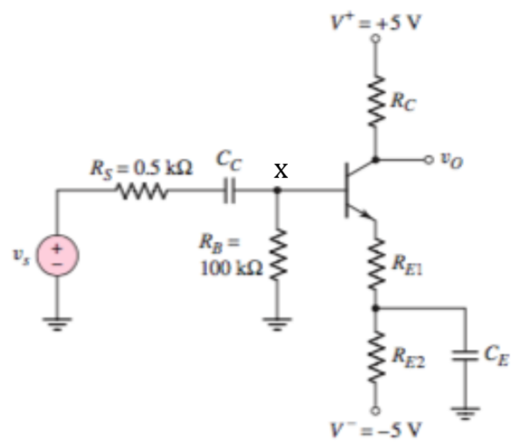


Figure 1: Common emitter stage

Assume that C_1, C_2 are large enough so that they act as open to DC and short to AC.

Common emitter stage voltage gain

Show that the voltage gain of the common emitter stage can be approximated as:

$$|A_v| = \frac{R_C}{\frac{1}{g_m} + R_{E_1}}$$

Consider only the common-emitter transistor implementation (and ignoring the input signal). Use the small signal model for the BJT, treating capacitor are

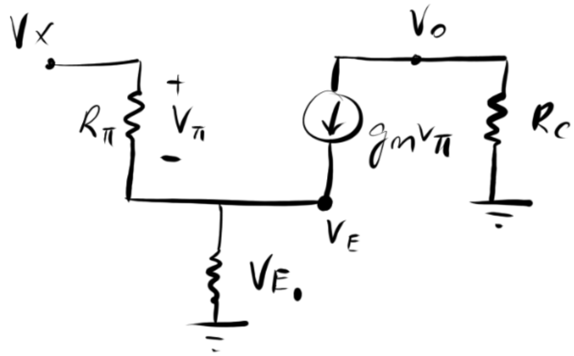


Figure 2: Small signal model of common emitter stage

short circuits. Perform KCL on the the two nodes.

$$\frac{V_x - V_E}{R_\pi} + \frac{0 - V_E}{R_{E_1}} + g_m V_\pi = 0$$
$$g_m V_\pi + \frac{V_o}{R_C} = 0$$

Since the base current is small relative to the amplified current, approximate that the first term of the first equation, $V_\pi / R_\pi \approx 0$. Then:

$$\begin{aligned}
V_x\left(g_m\right) &= V_E \frac{1}{R_{E_1}} + g_m \\
V_o &= g_m R_C V_\pi = g_m R_C \left(V_x - V_E\right) \\
&= g_m R_C V_x \left(1 - \frac{g_m}{\frac{1}{R_{E_1}} + g_m}\right) \\
&= g_m R_C V_x \left(\frac{\frac{1}{R_{E_1}} + g_m - g_m}{\frac{1}{R_{E_1}} + g_m}\right) \\
&= \frac{R_C V_x}{\frac{1}{g_m} + R_{E_1}} \\
\Rightarrow \left|A_v\right| = \frac{V_o}{V_x} &= \frac{R_C}{\frac{1}{g_m} + R_{E_1}}
\end{aligned}$$

which is the desired result. The attenuation factor is related to this voltage gain (the higher the gain, the lower the attenuation.) It can be seen that it is dependent on the current flowing the transistor, as $g_m \propto I_C$: a higher current causes a larger gain.

Circuit design to match specifications

Design the circuit such that $V_{CEQ} = 4\text{V}$ *and to amplify a* 12mV *sinusoidal signal from a microphone, having an output resistance of* $0.5\text{k}\Omega$*, to a* 0.4V *sinusoidal output signal. Approximate* $|A_v| \approx R_C/R_E$*. Assume that the transistor used in the design has nominal values of* $\beta = 100$*,* $V_{BE(ON)} = 0.7\text{V}$*, and* $I_s = 4 \times 10^{-16}\text{V}$*.*

Start with the large signal model to find R_C and R_E.

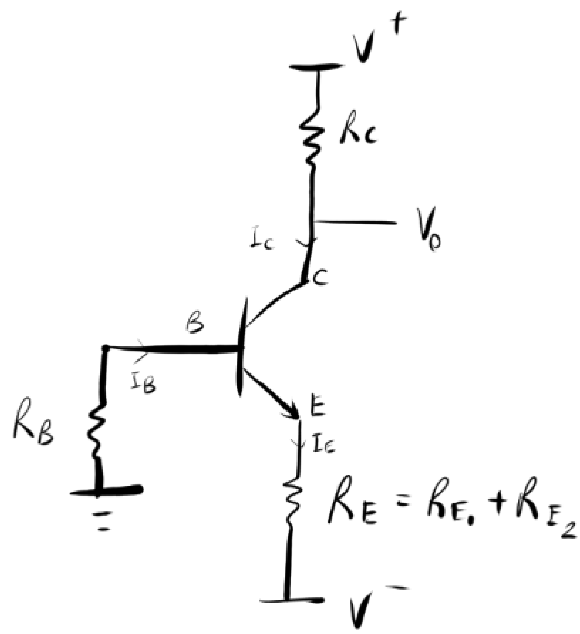


Figure 3: Large signal model

$$I_C = I_s \exp\left(\frac{V_{BE}}{V_T}\right) = 1.97 \times 10^{-4}\text{A}$$
$$I_B = \frac{I_C}{\beta} = 1.97 \times 10^{-6}\text{A}$$
$$I_E \approx I_C$$
$$V_B = -R_B I_B = 0.197\text{V}$$
$$V_E = V_B - V_{BE} = -0.897\text{V}$$
$$V_C = V_E + V_{CEQ} = 3.103\text{V}$$
$$R_C = \frac{V^+ - V_C}{I_C} = \frac{5\text{V} - 3.303\text{V}}{1.97 \times 10^{-4}\text{A}} = 9629\Omega$$
$$R_E = \frac{V_E - V^-}{I_E} = 20827\Omega = R_{E_1} + R_{E_2}$$

Now we need to find R_{E_1} and R_{E_2} (such that their sum equals R_E). We turn to the small signal model, where $g_m = \frac{I_C}{V_T}$ and $R_\pi = \frac{\beta}{g_m}$. By KCL:

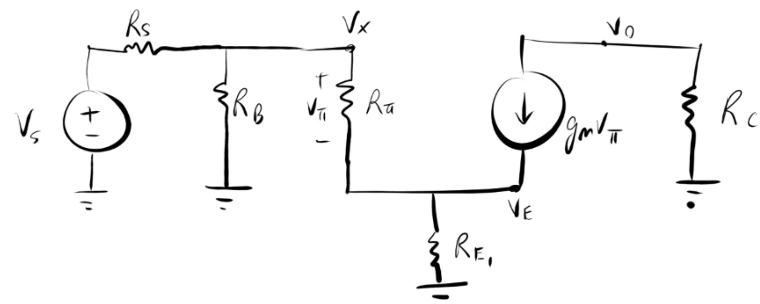


Figure 4: Small signal model

$$\frac{V_s - V_x}{R_s} + \frac{0 - V_x}{R_B} + \frac{V_E - V_x}{R_\pi} = 0$$
$$\frac{V_x - V_E}{R_\pi} + \frac{0 - V_E}{R_{E_1}} + g_m(V_x - V_E) = 0$$

Rearrange these equations to get V_s in terms of V_x.

$$V_s\left(\frac{1}{R_s}\right) = V_x\left(\frac{1}{R_s}+\frac{1}{R_B}+\frac{1}{R_\pi}\right) - V_E\left(\frac{1}{R_\pi}\right)$$

$$V_E\left(\frac{1}{R_\pi}+\frac{1}{R_{E_1}}+g_m\right) = V_x\left(\frac{1}{R_\pi}+g_m\right)$$

$$\frac{V_s}{V_x} = R_s\left(\frac{1}{R_s}+\frac{1}{R_B}+\frac{1}{R_\pi}-\frac{1}{R_\pi}\left(\frac{\frac{1}{R_\pi}+G_m}{\frac{1}{R_\pi}+\frac{1}{R_{E_1}+g_m}}\right)\right)$$

$$= 1+\frac{R_s}{R_\pi}+\frac{R_s}{R_B}-\frac{R_s}{R_\pi}\left(\frac{1+R_\pi g_m}{1+\frac{R_\pi}{R_{E_1}}+g_m R_\pi}\right)$$

Since we know that $|A_V| \approx \frac{R_C}{R_{E_1}}$ from the previous section, and since we know that the desired output $V_o = 0.4\text{V}$ when $V_s = 0.012\text{V}$, then we can solve for R_{E_1}:

$$A_V = \frac{R_C}{R_{E_1}} = \frac{V_o}{V_x} = \frac{V_s}{V_x}\frac{V_o}{V_s} = \left[1+\frac{R_s}{R_\pi}+\frac{R_s}{R_B}-\frac{R_s}{R_\pi}\left(\frac{1+R_\pi g_m}{1+\frac{R_\pi}{R_{E_1}}+g_m R_\pi}\right)\right]\frac{0.4\text{V}}{0.012\text{V}}$$

The only unknown here is R_{E_1}. Plugging into a graphing calculator (desmos.com) and visually solving the expression, we get $R_{E_1} = 286\Omega$. In summary:

$$R_C = 9629\Omega$$
$$R_{E_1} = 286\Omega$$
$$R_{E_2} = 20827\Omega - 286\Omega = 20541\Omega$$

Diode-connected transistor

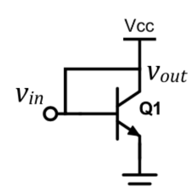


Figure 5: Diode-connected transistor

You may assume that the transistor, with common-emitter current gain β, is in forward active mode and has the following small signal parameters: r_π, g_m.

Small signal input impedance

Determine the small signal input impedance.

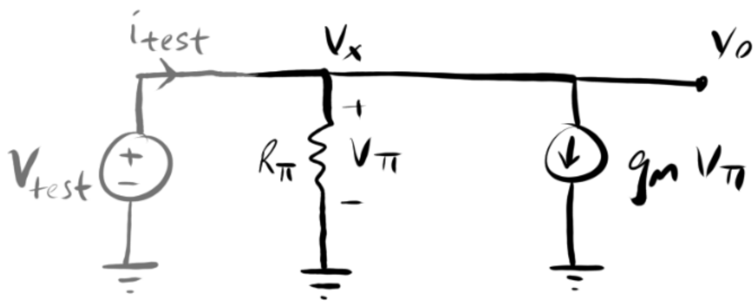


Figure 6: Small-signal model with test voltage

Use KCL at the only non-ground node (which is at voltage $V_\pi = V_o = V_{test}$):

$$I_{test} = \frac{V_\pi}{R_\pi} + g_m V_\pi = \frac{V_\pi}{\beta/g_m} + g_m V_\pi = g_m V_\pi \left(\frac{\beta+1}{\beta}\right) = V_{test} g_m \left(\frac{\beta+1}{\beta}\right)$$

$$R_{in} = \frac{V_{test}}{I_{test}} = \frac{\beta}{g_m(\beta+1)} \approx \frac{1}{g_m}$$

Since $g_m = I_C/V_T$, the input impedance is roughly inversely proportional to the current $I \approx I_C$ through the BJT, which means that it will maintain a roughly constant voltage across it ($V = IR_{in}$, and I and R_{in} are roughly proportional), similar to a diode.

Justify the name

In this circuit configuration, Q_1 is referred to as a diode-connected transistor. Justify!

Just by looking at the circuit, we can see that the collector and base are connected, so the voltage $V_{CB} = 0$ and there is no current flowing between those terminals. Thus, the "input" terminal is the base and the output is the emitter, which is just a PN-junction (diode).

ECE303 – Communication Networks

Jonathan Lam

February 9, 2020

Contents

1 Computer Networks and the Internet

1.1 What is the Internet?

We can think of the Internet:

- Physically, as a network of communication links connecting hosts to each other
- Functionally, as an infrastructure providing services to applications

1.2 Network edge

end systems/hosts devices connected to the internet (and not facilitating the Internet's connections); i.e., compute devices; i.e., they "host" applications; divide into client and server based on usage

communication links a network of communication links join the hosts

packet switch usually routers (usually in network core) and link-layer switches (access networks)

route/path the sequence of communication links and packet switches taken by a path from one host to another

Internet Service Provider (ISP) provide access to the internet; are a network of packet switches and communication links

Internet Engineering Task Force (IETF) manages internet standards and RFCs

Request For Comments (RFC) define protocols

distributed application applications that require connecting to another host; "distributed" because run on multiple computers; only run on edge devices (compute devices)

protocol defines the format and order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission and/or receipt of a message or other event

network access anything that takes an end device to the first router ("edge router")

1.3 Network access

DSL and cable two major types of broadband residential access; DSL (digital subscriber line) happens over a telephone network, so that the telephone company (telco) is the ISP; the signal gets sent to the central office (CO), which multiplexes the signal (DSLAM, DSL Access Multiplexer), which is connected to the network core and converts the analog frequencies into digital ones; cable uses the same infrastructure as cable television; CMTS (cable modem termination system) is similar to DSLAM; cable modem is a shared broadcast medium, so total bandwidth is limited

HFC hybrid fiber coax; a common implementation of cable systems

asymmetric transmission rates typically, download is allocated higher bandwidth than upload

FTTH Fiber To The Home; e.g., FIOS; fast, one-to-one w/ CO

satelliate, dial-up alternative methods for Internet access

physical data transmission media twisted-pair copper wire, coaxial cable, multimode fiber-optic cable, terrestrial/satellite radio spectrum

guided/unguided data transmission media wires: guided: WLAN/satellite: unguided

unshielded twisted pair (UTP) cheapest form of wire; commonly used in indoor wiring, but not as good as shielded

1.4 Network core

packet speed packets are transmitted over each link at its full transmission speed (i.e., they are atomic, not broken up or mixed with other packets)

store-and-forward packets are sent one at a time, only after fully finished receiving (b/c may need to process) (i.e., it builds entire packet in some buffer)

queueing delays packets must wait in output buffer/queue until communication link is done transmitting other packets

packet loss output buffer full

forwarding table in a router; maps (parts of) destination IPs to outbound links; get automatically set by routing protocols

circuit a communication link on the path between two hosts which maintain their state for a connection; a.o.t. packet switching; has a guaranteed constant transmission rate

frequency/time-division multiplexing (FDM/TDM) different frequency bands are reserved (full time, partial bandwidth) for a circuit or time is divided into frames with dedicated time slots (partial time, full bandwidth)

circuit switching benefits/downfalls over packet switching it has "silent periods" because of reserved pieces not being used; also some time to set up the circuit; it has the ability for real-time services; it is more costly to implement

ISP tiers tier-1: globally-spanning, doesn't have to pay anyone; local: connect to tier-1 or larger local ISPs

customer/provider ISPs customers pay providers to connect to them; show hierarchy of size

Points of Presence (PoP) a group of routers in the same location in the provider's network where customer ISPs can connect into the provider ISP

multi-homing when an ISP connects to multiple provider ISPs

peering when two ISPs (usually of the same level) exchange traffic without either side paying

Internet Exchange Point (IXP) a place where multiple ISPs can peer together

Content Provider Networks (CDNs) corporations with large internal networks largely separate from the rest of the Internet; allow greater control over their data and services

delays processing (examining headers); queueing (other items in queue); transmission delay (how long it takes to get all of packet onto link); propagation delay (based on physical medium's properties)

latency and throughput latency is how long it takes for a packet to reach its destination; throughput is the rate at which packets goes through (are received by the receiver) and is bounded above by the lowest bandwidth of any links along the route; access network (the "last mile") is usually the bottleneck for performance in today's Internet

1.5 Protocol layers and their service models

1.5.1 TCP/IP stack

A user application places some data into a message in some application-specific protocol. The OS manages TCP protocol and sets up the IP protocol. When traveling between hosts, while on the physical medium, physical-level protocols govern bit movement at the most basic level, and link-level protocols provide some abstraction that governs reliability and makes some decisions between any nodes in the network. At a node, the network-level protocols manage which next link to take. When the destination is reached, the transport- and application-level protocols are again managed by the host.

E.g., a NIC operates at levels 1 (sending out physical bits) and level 2 (making decisions based on MAC address); a switch also operates at levels 1 and 2; a router operates at 1-3.

application where network applications and their protocols live (e.g., HTTP, SMTP, FTP, DNS); packet of information in application layer is referred to as a "message"

transport TCP, UDP; supports application layer with connection and some reliability (TCP only); transport layer packet is a "segment"

network (IP layer) responsible for moving network-layer packets ("datagrams") between hosts, e.g., through the IP protocol and other routing protocols

link responsible for reliable movement of link-level packets ("frames") between network nodes, e.g., WiFi, Ethernet

physical protocols for moving bits along the specific physical medium

1.5.2 ISO OSI model

Open Systems Interconnection model, by the International Organization for Standardization

application

presentation provide services that allow communicating applications to interpret the meaning of data exchanged, e.g., data compression and encryption (e.g., SSL/TLS)

session for delimiting and synchronization of data exchange, such as the means to build a checkpointing and recovery scheme

transport

network

link

physical

Important idea of **data encapsulation**; at any node, any headers/packets at a higher level in the stack than what the node is implemented at is treated as data. Only the header at the level of implementation is interpreted. I.e., the **payload** of any packet comprises of the packet from the layer above (unless it is the application-layer, in which the payload is the user's data).

1.6 Networks under attack

botnet a network of malware-compromised devices that attackers use to perform coordinated attacks (e.g., DDoS)

self-replicating many malware are self-replicating

viruses vs. worms viruses require explicit user input; worms do not

DoS attacks usually by vulnerability attacks (well-crafted messages sent to the server) or bandwidth/connection flooding

packet sniffing e.g., WireShark; data over shared media (e.g., WiFi) are especially vulnerable

IP spoofing this creates the need for end-point authentication

2 Application layer

2.1 Principles of network applications

P2P architectures interesting b/c of self-scalability – each user also acts as part of the service; however, is not ISP friendly (due to asymmetric upload/download rate of residential ISPs), may have security and incentive concerns

loss-tolerant applications e.g., video streaming – not of utmost important that every bit gets through

bandwidth-sensitive vs. elastic applications the former requires a certain amount of throughput to function (e.g., media); the latter just uses whatever is available (e.g., email)

services provided by transfer protocols reliable data transfer, security provided by TCP (the latter with SSL/TLS); throughput and timing are not guaranteed by current internet protocols

2.2 Overview of the WWW and HTTP

HTTP and the transfer layer HTTP runs over TCP (not over UDP)

stateless HTTP is stateless – doesn't maintain any information about previous information sent

round-trip time (RTT) latency from client to server back to client

web cache may be implemented locally, or through a local server

conditional GET may be used to keep cache up to date; cache asks server for updates, server sends updated file or text message if not updated

Jonathan Lam
Prof. Chen
ECE 303
Communication Networks
2 / 22 / 20

Project 2 - Transport

<u>Port Scanning</u>

1. What nmap command would perform a scan of the top *x* most common ports, specifically TCP only? Let the target host be “synprint.com” and *x*=10.
Answer: nmap --top-ports 10 synprint.com

Cut and paste the result of the nmap command below:

```
Nmap scan report for synprint.com (192.241.168.54)
Host is up (0.0057s latency).

PORT     STATE    SERVICE
21/tcp   filtered ftp
22/tcp   open     ssh
23/tcp   filtered telnet
25/tcp   filtered smtp
80/tcp   open     http
110/tcp  filtered pop3
139/tcp  filtered netbios-ssn
443/tcp  open     https
445/tcp  filtered microsoft-ds
3389/tcp filtered ms-wbt-server

Nmap done: 1 IP address (1 host up) scanned in 1.35 seconds
```

2. Capture the packets of your nmap run. Save the file as “<date>_nmap.pcap”.
Provide the command you used to collect the packets:
Answer: tcpdump -w 21200218_nmap.pcap

Dump the textual result of the pcap file below (tshark -r <pcap_file>):

```
    1   0.000000 192.168.1.187 → 192.168.1.1  DNS 72 Standard query 0x2c56 A synprint.com
    2   0.000022 192.168.1.187 → 192.168.1.1  DNS 72 Standard query 0x7354 AAAA synprint.com
    3   0.003001  192.168.1.1 → 192.168.1.187 DNS 88 Standard query response 0x2c56 A synprint.com A 192.241.168.54
    4   0.003374  192.168.1.1 → 192.168.1.187 DNS 140 Standard query response 0x7354 AAAA synprint.com SOA ns01.domaincontrol.com
    5   0.004065 192.168.1.187 → 192.241.168.54 TCP 74 40294 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012205
TSecr=0 WS=128
```

```
    6   0.004089 192.168.1.187 → 192.241.168.54 TCP 74 41356 → 443 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012205
TSecr=0 WS=128
    7   0.010411 192.241.168.54 → 192.168.1.187 TCP 74 80 → 40294 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1
TSval=1854221149 TSecr=3146012205 WS=64
    8   0.010425 192.168.1.187 → 192.241.168.54 TCP 66 40294 → 80 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012212 TSecr=1854221149
    9   0.010431 192.241.168.54 → 192.168.1.187 TCP 74 443 → 41356 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1
TSval=1854221149 TSecr=3146012205 WS=64
   10   0.010435 192.168.1.187 → 192.241.168.54 TCP 66 41356 → 443 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012212 TSecr=1854221149
   11   0.010467 192.168.1.187 → 192.241.168.54 TCP 66 40294 → 80 [RST, ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012212
TSecr=1854221149
   12   0.010480 192.168.1.187 → 192.241.168.54 TCP 66 41356 → 443 [RST, ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012212
TSecr=1854221149
   13   0.010596 192.168.1.187 → 192.168.1.1  DNS 87 Standard query 0xec3f PTR 54.168.241.192.in-addr.arpa
   14   0.013613  192.168.1.1 → 192.168.1.187 DNS 113 Standard query response 0xec3f PTR 54.168.241.192.in-addr.arpa PTR synprint.com
   15   0.013695 192.168.1.187 → 192.241.168.54 TCP 74 33208 → 445 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   16   0.013712 192.168.1.187 → 192.241.168.54 TCP 74 41360 → 443 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   17   0.013724 192.168.1.187 → 192.241.168.54 TCP 74 40302 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   18   0.013735 192.168.1.187 → 192.241.168.54 TCP 74 59336 → 25 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   19   0.013745 192.168.1.187 → 192.241.168.54 TCP 74 41576 → 139 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   20   0.013757 192.168.1.187 → 192.241.168.54 TCP 74 46986 → 110 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   21   0.013768 192.168.1.187 → 192.241.168.54 TCP 74 47924 → 21 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   22   0.013779 192.168.1.187 → 192.241.168.54 TCP 74 57352 → 22 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   23   0.013789 192.168.1.187 → 192.241.168.54 TCP 74 43522 → 3389 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   24   0.013799 192.168.1.187 → 192.241.168.54 TCP 74 54516 → 23 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146012215
TSecr=0 WS=128
   25   0.021402 192.241.168.54 → 192.168.1.187 TCP 74 22 → 57352 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1
TSval=1854221151 TSecr=3146012215 WS=64
   26   0.021415 192.168.1.187 → 192.241.168.54 TCP 66 57352 → 22 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223 TSecr=1854221151
   27   0.021420 192.241.168.54 → 192.168.1.187 TCP 74 80 → 40302 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1
TSval=1854221151 TSecr=3146012215 WS=64
   28   0.021425 192.168.1.187 → 192.241.168.54 TCP 66 40302 → 80 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223 TSecr=1854221151
   29   0.021428 192.241.168.54 → 192.168.1.187 TCP 74 443 → 41360 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1
TSval=1854221151 TSecr=3146012215 WS=64
   30   0.021431 192.168.1.187 → 192.241.168.54 TCP 66 41360 → 443 [ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223 TSecr=1854221151
   31   0.021468 192.168.1.187 → 192.241.168.54 TCP 66 41360 → 443 [RST, ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223
TSecr=1854221151
   32   0.021480 192.168.1.187 → 192.241.168.54 TCP 66 40302 → 80 [RST, ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223
TSecr=1854221151
   33   0.021487 192.168.1.187 → 192.241.168.54 TCP 66 57352 → 22 [RST, ACK] Seq=1 Ack=1 Win=64256 Len=0 TSval=3146012223
TSecr=1854221151
   34   0.607564 192.168.1.187 → 192.168.1.158 TCP 176 52230 → 8009 [PSH, ACK] Seq=1 Ack=1 Win=705 Len=110 TSval=1155283588 TSecr=7281774
[TCP segment of a reassembled PDU]
   35   0.611794 192.168.1.158 → 192.168.1.187 TCP 176 8009 → 52230 [PSH, ACK] Seq=1 Ack=111 Win=528 Len=110 TSval=7282275
TSecr=1155283588 [TCP segment of a reassembled PDU]
   36   0.611835 192.168.1.187 → 192.168.1.158 TCP 66 52230 → 8009 [ACK] Seq=111 Ack=111 Win=705 Len=0 TSval=1155283592 TSecr=7282275
   37   0.932089 192.168.1.187 → 192.168.1.169 TCP 176 33522 → 8009 [PSH, ACK] Seq=1 Ack=1 Win=614 Len=110 TSval=1818755475 TSecr=6863540
[TCP segment of a reassembled PDU]
   38   0.937463 192.168.1.169 → 192.168.1.187 TCP 176 8009 → 33522 [PSH, ACK] Seq=1 Ack=111 Win=721 Len=110 TSval=6864041
TSecr=1818755475 [TCP segment of a reassembled PDU]
   39   0.937540 192.168.1.187 → 192.168.1.169 TCP 66 33522 → 8009 [ACK] Seq=111 Ack=111 Win=614 Len=0 TSval=1818755481 TSecr=6864041
   40   1.019453 208.255.115.145 → 192.168.1.187 TLSv1.2 105 Application Data
   41   1.019589 192.168.1.187 → 208.255.115.145 TCP 66 57112 → 443 [ACK] Seq=1 Ack=40 Win=501 Len=0 TSval=33648639 TSecr=2800984375
   42   1.019654 208.255.115.145 → 192.168.1.187 TCP 66 443 → 57112 [FIN, ACK] Seq=40 Ack=1 Win=117 Len=0 TSval=2800984375 TSecr=33528618
   43   1.020431 192.168.1.187 → 208.255.115.145 TCP 66 57112 → 443 [FIN, ACK] Seq=1 Ack=41 Win=501 Len=0 TSval=33648640 TSecr=2800984375
   44   1.023556 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 54516 → 23 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   45   1.023583 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 43522 → 3389 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   46   1.023592 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 47924 → 21 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   47   1.023601 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 46986 → 110 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   48   1.023608 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 41576 → 139 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   49   1.023614 192.168.1.187 → 192.241.168.54 TCP 74 [TCP Retransmission] 59336 → 25 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1
TSval=3146013225 TSecr=0 WS=128
   50   1.028425 208.255.115.145 → 192.168.1.187 TCP 66 443 → 57112 [ACK] Seq=41 Ack=2 Win=117 Len=0 TSval=2800984385 TSecr=33648640
   51   1.115229 192.168.1.187 → 192.241.168.54 TCP 74 54518 → 23 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013316
```

```
TSecr=0 WS=128
   52   1.115637 192.168.1.187 → 192.241.168.54 TCP 74 43528 → 3389 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
   53   1.115753 192.168.1.187 → 192.241.168.54 TCP 74 47936 → 21 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
   54   1.115797 192.168.1.187 → 192.241.168.54 TCP 74 47002 → 110 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
   55   1.115833 192.168.1.187 → 192.241.168.54 TCP 74 41596 → 139 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
   56   1.115857 192.168.1.187 → 192.241.168.54 TCP 74 59360 → 25 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
   57   1.115881 192.168.1.187 → 192.241.168.54 TCP 74 33240 → 445 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 SACK_PERM=1 TSval=3146013317
TSecr=0 WS=128
```

3. Write a script/program in a language of your choice to parse the resulting pcap file, process the parsed results, and create the same output as what nmap generated. Make sure you can handle the at least 2 states: open and filtered. Your output should look something like the following:

PORT   STATE SERVICE
21/tcp   filtered ftp
22/tcp   open   ssh
23/tcp   closed telnet
…

Cut and paste the analysis program/script below:
nmap_analysis.sh

```
#!/bin/sh

ip_addr=$(ip address show | grep "192.168.[0-9]\+\.[0-9]\+" -o | head -n 1)
open=$(tshark -r $1 | grep "$2.\+$ip_addr.\+\[SYN, ACK\]" | awk '{print
$8}' | sort -n | uniq)
attempted=$(tshark -r $1 | grep "$ip_addr.\+$2.\+\[SYN\]" | grep -oP
'[0-9]+(?= \[SYN\]+)' | sort -n | uniq)

printf "%12s\t%8s\t%s\n" "PORT" "STATE" "SERVICE"
for port in $attempted
do
  if [ $(echo "$open" | grep -c "^$port$") -eq 0 ]; then isopen="filtered";
else isopen="open"; fi
  name=$(getent services | grep -oP ".+(?= +$port/tcp)" | head -n 1)
  printf "%8s/tcp\t%8s\t%s\n" $port $isopen $name
done
```

Usage: ./nmap_analysis.sh [pcap_in] [dst_ip]

```
$ ./nmap_analysis.sh 21200218_nmap.pcap 192.241.168.54
        PORT          STATE     SERVICE
```

```
     21/tcp      filtered    ftp
     22/tcp          open    ssh
     23/tcp      filtered    telnet
     25/tcp      filtered    smtp
     80/tcp          open    http
    110/tcp      filtered    pop3
    139/tcp      filtered    netbios-ssn
    443/tcp          open    https
    445/tcp      filtered    microsoft-ds
   3389/tcp      filtered    ms-wbt-server
```

4. Write a script/program in a language of your choice to perform a simple TCP connection to a series of ports, specified from command line as a comma delimited list.

Cut and paste the TCP connection program/script below:
tcp_ports.sh

```
#!/bin/sh

ipdst=$1
ports=$(echo $2 | tr , '\n')

for port in $ports; do
  nc $ipdst $port <<<'' >/dev/null -w 1
done
```

Usage:./tcp_ports.sh [dst_ip] [port1,port2,...]

5. Putting it all together, write a script that uses tshark (or equivalent) and your programs/scripts from Part 3 and 4 above to perform the function of nmap, without ever running nmap. Your script should accept 2 inputs: target and a comma delimited list of ports to scan.

Cut and paste the analysis program/script below:

```
#!/bin/sh

target="$(dig +short $1)"
ports="$2"

tcpdump -w tmp.pcap 2>/dev/null &
tdpid=$!
sleep 1 # sleep to allow tcpdump to initialize
```

```
        # see: https://askubuntu.com/a/746061/433872
./tcp_ports.sh $target $ports
kill $tdpid
wait $tdpid

./nmap_analysis.sh tmp.pcap $target

rm tmp.pcap
```

Run your script using target="github.com" and ports="21,22,23,25,80,443" and cut and paste the output below:

```
$ ./nmap_copy.sh github.com 21,22,23,25,80,443
        PORT         STATE    SERVICE
      21/tcp      filtered    ftp
      22/tcp          open    ssh
      23/tcp      filtered    telnet
      25/tcp      filtered    smtp
      80/tcp          open    http
     443/tcp          open    https
```

PROTOCOL DESIGN TEMPLATE

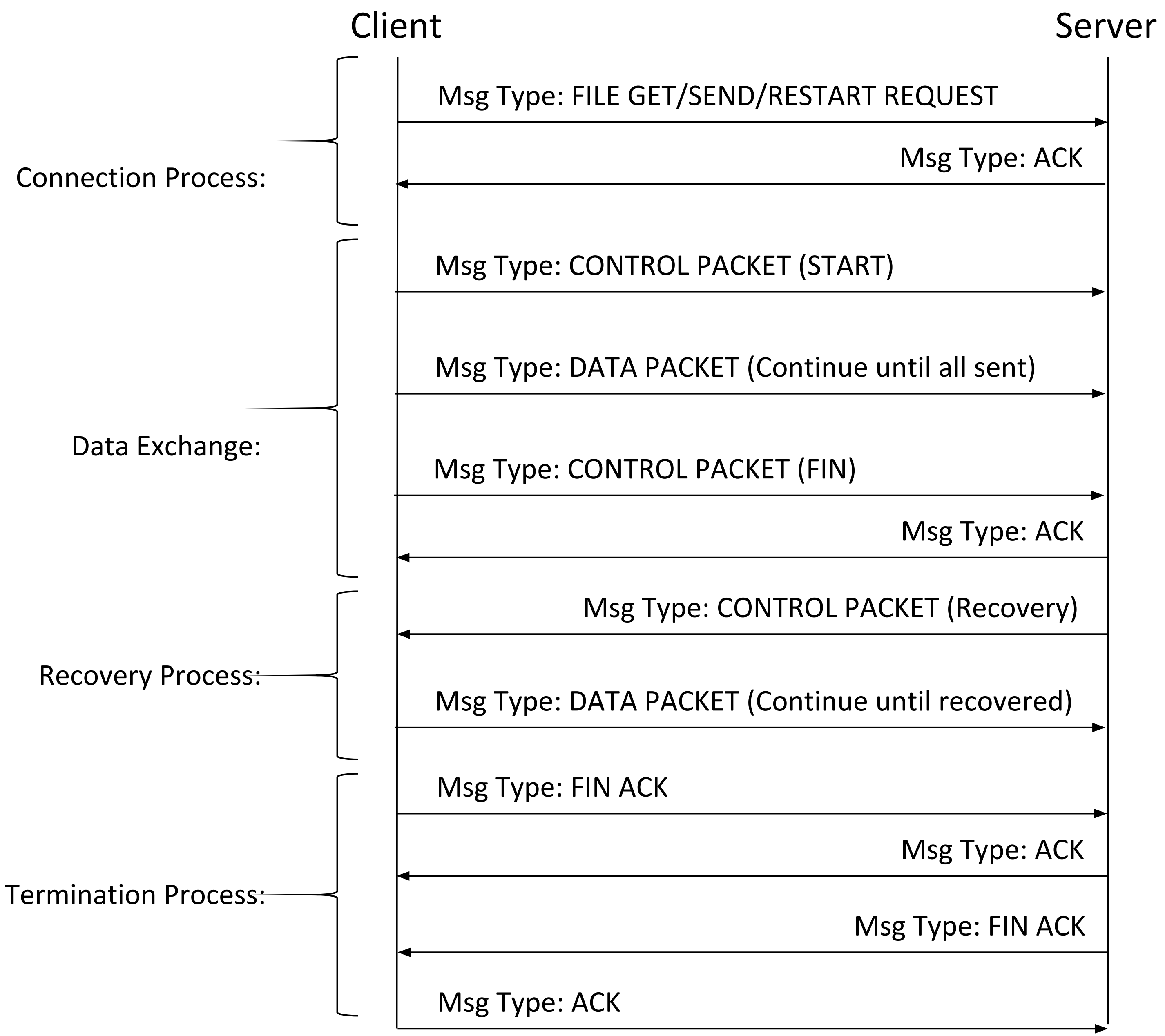
Dan and Jon's Wings (and Thodoris and Josh)

Dan & Jo~~h~~n's

WINGS

& Thodoris & Josh

Message Flow Template



Software Implementation Considerations

The software client must implement both the file sending and receiving parts. Both should listen on ports 10512 (control packets) and 10256 (data packets) by default (or can be configured otherwise).

Connection initiation

- Randomly generate a transaction ID, which will be sent on the initial file transfer request. The receiver of the file transfer request may indicate that this transaction ID is not unique, and this repeats until the transaction ID is indeed unique.

Software Implementation Considerations

ACKs:

- ACKs are sent on response of all control packets (which are not as numerous).
- ACKs are sent from file receiver every **n** data packets (**n** will be fixed based on recovery attempt number; i.e., initially **n**=2^8, then **n**=2^7, … **n**=2^1). Specific implementation of strict congestion control can be left up to the sender implementation (for example, compare time to send **n** number packets vs getting an ACK back from the receiver regarding those **n** packets, or slow down after a few lost ACKs).
- One way for the connection termination is after **z** lost consecutive ACKs (which would indicate that the receiver is not still connected).

Restart mechanism:

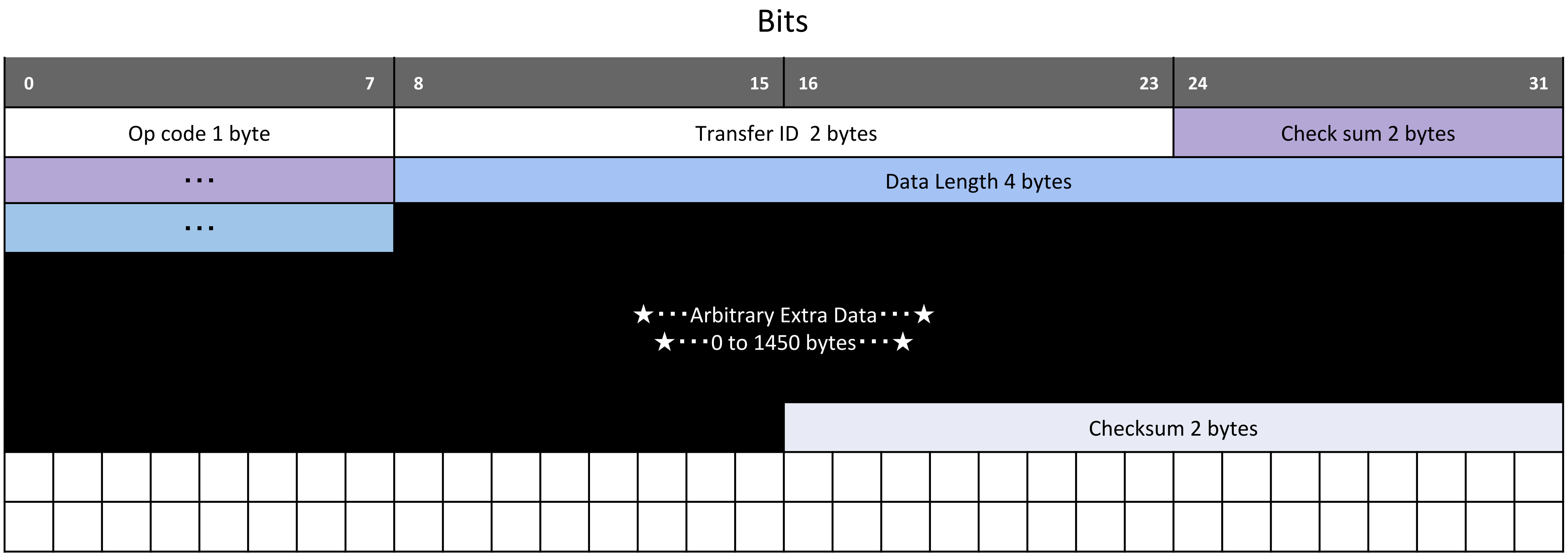
- Implemented purely software-side on file receiver side; if connection closed with some error, can retry seamlessly and just request the remaining packets with special restart opcode (basically starting file request and then immediately starting recovery with specified packets).

Recovery Process

- After all data packets sent, if any packets are lost, receiver sends a control packet specifying what packets were lost.
    - Flow loops back to sender sending data packets until all lost data packets sent.
    - If packets are still lost, flow loops back again to sending data packets.
    - This process loops **y** number of times, depending on the size of the file being sent.
    - If packets are still lost after **y** tries, terminate connection.

We can let **n** (number of packets between ACKs), **y** (number of recovery attempts), **z** (number of consecutive ACKs before connection is lost), be determined by the specific software implementation (or tweaked manually when starting the implementation).

Control Header Template ≪port:10512≫



Data Header Template 《port:10256》

Bits

<table>
<tr><th>0 … 7</th><th>8 … 15</th><th>16 … 23</th><th>24 … 31</th></tr>
<tr><td colspan="2">Session Transaction ID (2 bytes)</td><td colspan="2">Packet Size (2 bytes)</td></tr>
<tr><td colspan="4">Packet Sequence Id (8 bytes)</td></tr>
<tr><td colspan="4">Check sum (8 bytes)</td></tr>
<tr><td colspan="4">Data (variable length)</td></tr>
</table>

Control Header Values

Header	Values	Meaning/Derivation/Definition
TID	2 bytes	Transfer ID
Opcode	1 byte	Opcode
Data length	4 bytes	Length of data (in bytes)
Data	0-1450 bytes	Arbitrary length extra data (dependent on operation)
Checksum	2 bytes	Checksum

Control Header Opcodes

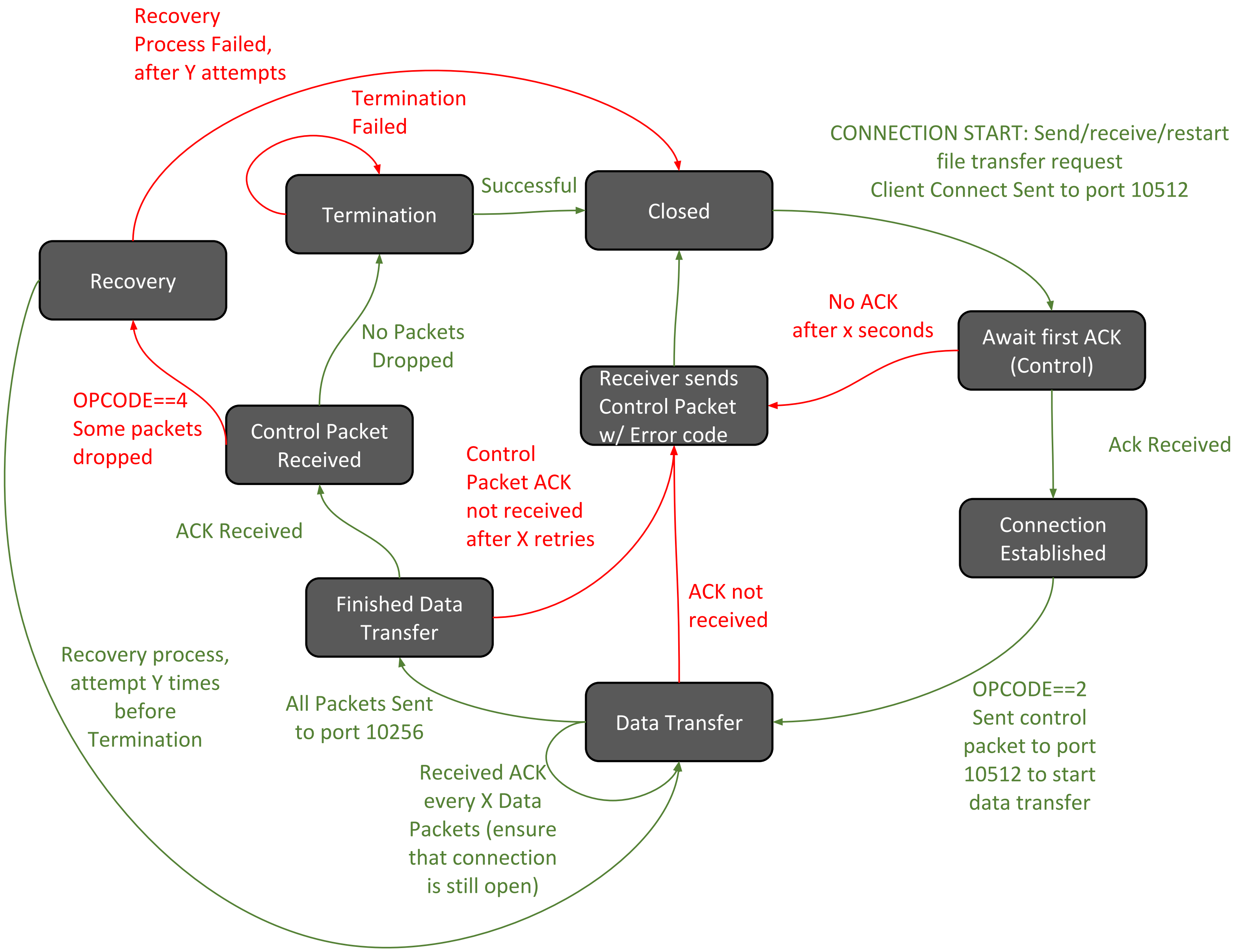
Opcode	Value	Extra data	Initiator	Meaning
Acknowledgement	0	Opcode of packet acking to	Either	Received packet, used in basic flow termination
Get file request	1	File path	Connection initiator	Request file to be sent over to initiator
Send file request	2	File path	Connection initiator	Request file to be sent from initiator
Finished sending	3	Retries	File sender	Current batch of packets sent
Bad packets	4	List of missing packets (sequence order; specific numbers or ranges)	File receiver	Not all packets received
Error	5	Error code	Either	Error in file transfer
Stop file sending	6		Either	Stop request
Restart request	7	List of missing packets (sequence order)		

Data Header Values

Header	Values	Meaning/Derivation/Definition
Packet Size	0-65535	Size of the Header Packet
Packet sequence index: 8 bytes	0-[2^{64}-1]	indexes each individual packet
Session/Transaction ID	0-65535	Unique identification for file transaction
Checksum	0-255	detects errors

State	Description
Closed	Client is not sending any packets
Await first ACK (control)	Wait for the first ACK after sending control packets
Connection Established	Send starting control packet to start data transfer
Data Transfer	Make sure ACKs are received after every x data packets, if not assume connection is dead.
Finished Data Transfer	After Data Transfer, determine if ACK was received
Control Packet Received	If no packets dropped, proceed to termination, else proceed to recovery
Recovery	Proceed with recovery for dropped packets, then proceed to termination
Termination	FIN ACK and ACK responses.
Send Control Packet w/ Error Code	Close connection with ERROR OPCODE; sender pings receiver for error Control Packet, restart connection w/ proper restart opcode, handled via software.

Client Behavior



Message Flow and State Diagram Questions

- How does the client initiate a connection? **Get/Send/Resend file request opcode.**
- Are you going to have authentication? **Nay**
- Are you going to have encryption? **Absolutely none**
- How will you confirm when a file has completed transfer? **Finish sending opcode**
- How will you ensure the integrity of the message in transit? **Checksums**
- How will you handle dropped packets? **After sender sends “finish sending” opcodes, receiver sends ack (on success) or sends “bad packets” opcode and lists sequence numbers of bad packets**
- Who initiates the closing of the connection? **Either with “stop file sending” opcode**
- Will you have error handling? **Basic — send error code and message in variable length**

Message Header and Value Questions

•How many header types do you need? Does it make sense to define one header and have an opcode to differentiate message types? Or does it make sense to have multiple header types? **Two header types.**

•How many bytes will the header(s) be? **Data header- fixed 20 bytes + DATA**

Control header- Fixed 11 bytes + Arbitrary Data

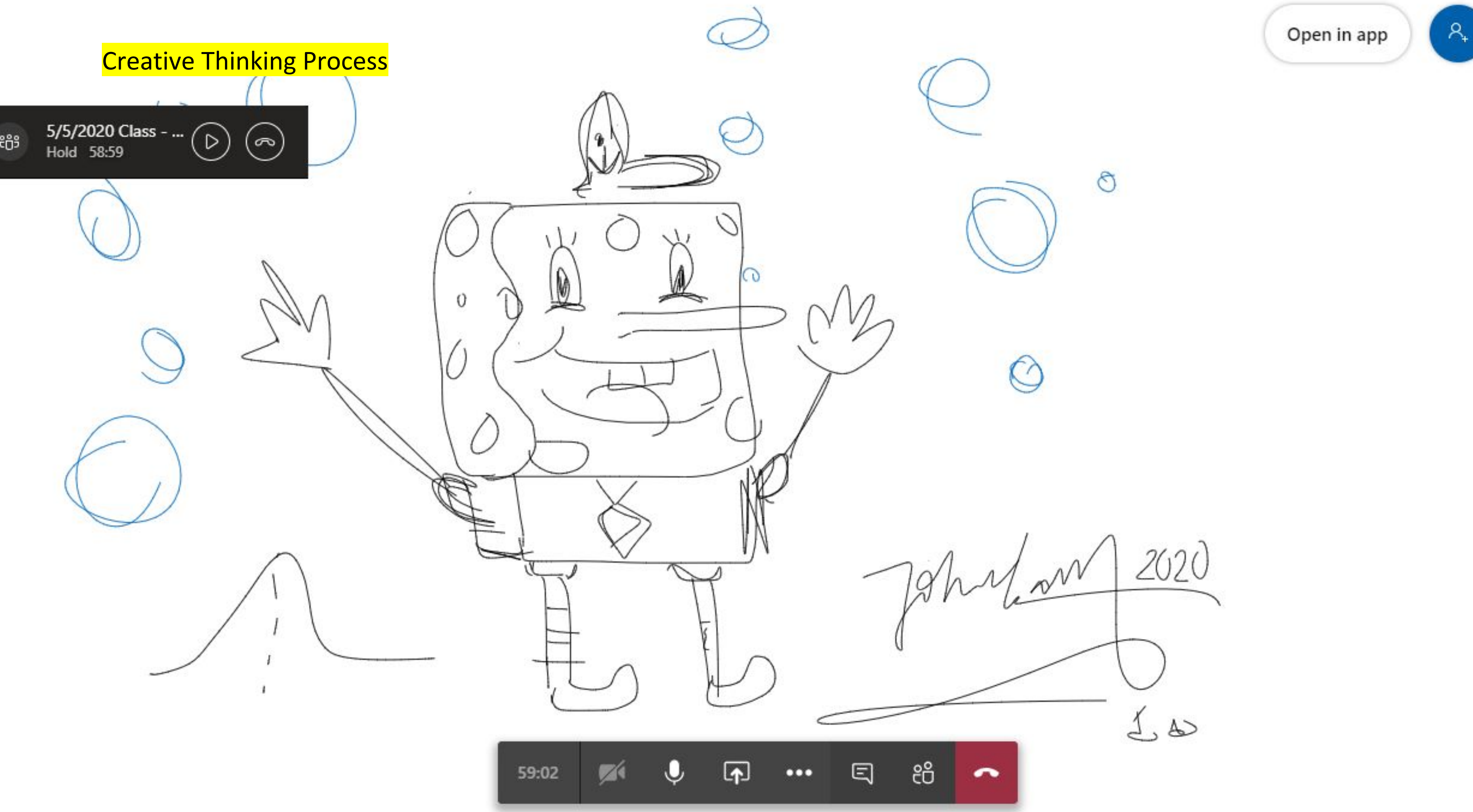
•Do you want the header to be a fixed size or do you want to enable variability? **We have variability in the packet size field**

•Do you want to make your header as small and efficient as possible, only catering to what is possible currently, or do you want to future-proof your header by making it larger than currently necessary? **Small headers, large variable for arbitrary data (control header)**

•Do you want to add options that allow you to introduce variability and future proof mechanisms?

Variable length extra data field and a whole byte for opcodes allows for expandability in control header.

Creative Thinking Process



ECE303 – Communication Networks

Jonathan Lam

February 9, 2020

Contents

1 Computer Networks and the Internet

1.1 What is the Internet?

We can think of the Internet:

- Physically, as a network of communication links connecting hosts to each other
- Functionally, as an infrastructure providing services to applications

1.2 Network edge

end systems/hosts devices connected to the internet (and not facilitating the Internet's connections); i.e., compute devices; i.e., they "host" applications; divide into client and server based on usage

communication links a network of communication links join the hosts

packet switch usually routers (usually in network core) and link-layer switches (access networks)

route/path the sequence of communication links and packet switches taken by a path from one host to another

Internet Service Provider (ISP) provide access to the internet; are a network of packet switches and communication links

Internet Engineering Task Force (IETF) manages internet standards and RFCs

Request For Comments (RFC) define protocols

distributed application applications that require connecting to another host; "distributed" because run on multiple computers; only run on edge devices (compute devices)

protocol defines the format and order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission and/or receipt of a message or other event

network access anything that takes an end device to the first router ("edge router")

1.3 Network access

DSL and cable two major types of broadband residential access; DSL (digital subscriber line) happens over a telephone network, so that the telephone company (telco) is the ISP; the signal gets sent to the central office (CO), which multiplexes the signal (DSLAM, DSL Access Multiplexer), which is connected to the network core and converts the analog frequencies into digital ones; cable uses the same infrastructure as cable television; CMTS (cable modem termination system) is similar to DSLAM; cable modem is a shared broadcast medium, so total bandwidth is limited

HFC hybrid fiber coax; a common implementation of cable systems

asymmetric transmission rates typically, download is allocated higher bandwidth than upload

FTTH Fiber To The Home; e.g., FIOS; fast, one-to-one w/ CO

satelliate, dial-up alternative methods for Internet access

physical data transmission media twisted-pair copper wire, coaxial cable, multimode fiber-optic cable, terrestrial/satellite radio spectrum

guided/unguided data transmission media wires: guided: WLAN/satellite: unguided

unshielded twisted pair (UTP) cheapest form of wire; commonly used in indoor wiring, but not as good as shielded

1.4 Network core

packet speed packets are transmitted over each link at its full transmission speed (i.e., they are atomic, not broken up or mixed with other packets)

store-and-forward packets are sent one at a time, only after fully finished receiving (b/c may need to process) (i.e., it builds entire packet in some buffer)

queueing delays packets must wait in output buffer/queue until communication link is done transmitting other packets

packet loss output buffer full

forwarding table in a router; maps (parts of) destination IPs to outbound links; get automatically set by routing protocols

circuit a communication link on the path between two hosts which maintain their state for a connection; a.o.t. packet switching; has a guaranteed constant transmission rate

frequency/time-division multiplexing (FDM/TDM) different frequency bands are reserved (full time, partial bandwidth) for a circuit or time is divided into frames with dedicated time slots (partial time, full bandwidth)

circuit switching benefits/downfalls over packet switching it has "silent periods" because of reserved pieces not being used; also some time to set up the circuit; it has the ability for real-time services; it is more costly to implement

ISP tiers tier-1: globally-spanning, doesn't have to pay anyone; local: connect to tier-1 or larger local ISPs

customer/provider ISPs customers pay providers to connect to them; show hierarchy of size

Points of Presence (PoP) a group of routers in the same location in the provider's network where customer ISPs can connect into the provider ISP

multi-homing when an ISP connects to multiple provider ISPs

peering when two ISPs (usually of the same level) exchange traffic without either side paying

Internet Exchange Point (IXP) a place where multiple ISPs can peer together

Content Provider Networks (CDNs) corporations with large internal networks largely separate from the rest of the Internet; allow greater control over their data and services

delays processing (examining headers); queueing (other items in queue); transmission delay (how long it takes to get all of packet onto link); propagation delay (based on physical medium's properties)

latency and throughput latency is how long it takes for a packet to reach its destination; throughput is the rate at which packets goes through (are received by the receiver) and is bounded above by the lowest bandwidth of any links along the route; access network (the "last mile") is usually the bottleneck for performance in today's Internet

1.5 Protocol layers and their service models

1.5.1 TCP/IP stack

A user application places some data into a message in some application-specific protocol. The OS manages TCP protocol and sets up the IP protocol. When traveling between hosts, while on the physical medium, physical-level protocols govern bit movement at the most basic level, and link-level protocols provide some abstraction that governs reliability and makes some decisions between any nodes in the network. At a node, the network-level protocols manage which next link to take. When the destination is reached, the transport- and application-level protocols are again managed by the host.

E.g., a NIC operates at levels 1 (sending out physical bits) and level 2 (making decisions based on MAC address); a switch also operates at levels 1 and 2; a router operates at 1-3.

application where network applications and their protocols live (e.g., HTTP, SMTP, FTP, DNS); packet of information in application layer is referred to as a "message"

transport TCP, UDP; supports application layer with connection and some reliability (TCP only); transport layer packet is a "segment"

network (IP layer) responsible for moving network-layer packets ("datagrams") between hosts, e.g., through the IP protocol and other routing protocols

link responsible for reliable movement of link-level packets ("frames") between network nodes, e.g., WiFi, Ethernet

physical protocols for moving bits along the specific physical medium

1.5.2 ISO OSI model

Open Systems Interconnection model, by the International Organization for Standardization

application

presentation provide services that allow communicating applications to interpret the meaning of data exchanged, e.g., data compression and encryption (e.g., SSL/TLS)

session for delimiting and synchronization of data exchange, such as the means to build a checkpointing and recovery scheme

transport

network

link

physical

Important idea of **data encapsulation**; at any node, any headers/packets at a higher level in the stack than what the node is implemented at is treated as data. Only the header at the level of implementation is interpreted. I.e., the **payload** of any packet comprises of the packet from the layer above (unless it is the application-layer, in which the payload is the user's data).

1.6 Networks under attack

botnet a network of malware-compromised devices that attackers use to perform coordinated attacks (e.g., DDoS)

self-replicating many malware are self-replicating

viruses vs. worms viruses require explicit user input; worms do not

DoS attacks usually by vulnerability attacks (well-crafted messages sent to the server) or bandwidth/connection flooding

packet sniffing e.g., WireShark; data over shared media (e.g., WiFi) are especially vulnerable

IP spoofing this creates the need for end-point authentication

2 Application layer

2.1 Principles of network applications

P2P architectures interesting b/c of self-scalability – each user also acts as part of the service; however, is not ISP friendly (due to asymmetric upload/download rate of residential ISPs), may have security and incentive concerns

loss-tolerant applications e.g., video streaming – not of utmost important that every bit gets through

bandwidth-sensitive vs. elastic applications the former requires a certain amount of throughput to function (e.g., media); the latter just uses whatever is available (e.g., email)

services provided by transfer protocols reliable data transfer, security provided by TCP (the latter with SSL/TLS); throughput and timing are not guaranteed by current internet protocols

2.2 Overview of the WWW and HTTP

HTTP and the transfer layer HTTP runs over TCP (not over UDP)

stateless HTTP is stateless – doesn't maintain any information about previous information sent

round-trip time (RTT) latency from client to server back to client

web cache may be implemented locally, or through a local server

conditional GET may be used to keep cache up to date; cache asks server for updates, server sends updated file or text message if not updated

Jonathan Lam, Richard Lee, Tiffany Yu, Victor Zhang
Prof. Lusterman
ECE366 — Software Engineering and Large System Design
5 / 13 / 20

Photocol: Final Report

Table of Contents:

Introduction

We present a novel method of handwriting recognition…

…hmmm. Wrong project. Let's start again.

Photocol is an intuitive website that allows users to upload and view photos, make collections of photos, and share them with other users of the site. The vision is for it to act as a fully-featured photo collection sharing site, with the ability to befriend other users, share and discover collections freely, and be able to view a list of public and discoverable collections personally selected to fit your interests. For this class, we implemented the ability to upload photos, make collections of photos and share collections with other users.

Nontechnical overview of project

(First of all, the name is a portmanteau of "photo" and "protocol": the former for the subject of the project, and the latter to indicate how we were learning to interface with different web protocols.)

When first arriving at the site, users are prompted to either make an account or login. Upon signing in, users can navigate to a number of pages, namely **Profile**, **Photos**, and **Collections**.

On the **Profile** page, a user can edit his/her profile picture, display name, as well as view Public and Discoverable Collections (see **Collections** below).

On the **Photos** page, a user can view all of his/her photos, neatly displayed in a horizontal masonry layout. The Photos page provides an Upload Photos button, a Select Photos Button, as well as a Delete Photos Button. A user can use the Select Photos Button to toggle Selection Mode. Selected photos can be subsequently deleted using the Delete Button.

Clicking on any photo leads a user to the **Photo** page, which displays an expanded version of that single image, followed by the original upload name, as well as a photo caption and some metadata, such as image dimensions.

On the **Collections** page, a user can create a collection as well as view any collections he/she has access to. Each collection is displayed in a box stating the collection title, its owner, as well as an optional cover image for the collection.

Clicking on any collection will lead a user to the page for that specific collection. The images are once again laid out in a masonry layout. The **Collection** page provides several buttons including, Edit Collection Details (shown only to the owner), Add Photos [to Collection], Select Multiple Photos [from photos currently in Collection], and Remove Photo [from Collection] (the latter three are only shown to the owner and editors). The Edit Collection Details button allows the owner to change the name and description of a collection, as well as add additional users to the collection as viewers, editors, or owners). The Add Photos button opens a modal of the Photos page, and provides the user with the ability to select photos to add to the collection, or upload photos straight from their computer. The Select Multiple Photos button allows a user to select multiple photos and subsequently use the Remove Photos button to remove the photos from the collection.

Collections can be set to Private, Discoverable, and Public. In private mode, only specific users in the Access Control List are allowed to view the collection details, as well as photos. In Discoverable mode, users not on the Access Control List for the collection can see the cover image, title, and description, but not any of the photos. Finally, in Public mode, any user or person who is not signed in can view the cover image, title, description, and collection photos. Discoverable and Public collections can be found on a user's Profile Page.

There is also a complex photo permissions system. As described earlier, collections can be set to public, discoverable, or private. A non-signed-in user can view a photo if it is in a public collection, if it is the cover photo for a discoverable collection, or if it is the profile photo for a user. A signed-in user can view the same photo if any of those conditions apply, if the user is in the access-control-list of a collection which contains the photo, or if the user owns the photo

Discoverable collections may seem somewhat artificial and unnecessary right now (as a user not in its ACL cannot take any action to join the collection), but this was because we hoped to make a larger system of collection discovery (of which a very basic version is implemented on the landing page), where users can send requests to join a collection (similar to how you can request access from a private Google Drive document or Facebook group that you have the link to).

Technical overview of project

Summary

We implemented the backend in three layers (handler, service, and store layers), driven by the Photocol class (initializing services) and the Endpoints class (directing HTTP requests to the handlers and handling authorization and CORS middleware). The handler layer manages basic input validation and packages the inputs into the internal types (User, PhotoCollection, and Photo). The service layer does most of the business logic (passing requests to the store layers as necessary). The store layer uses the DBConnectionManager utility class to make connections to the database (and does so by requesting a connection from the Hikari connection pool). Successful responses are sent back up the call chain to the handler, and serialized into JSON using GSON. Error conditions (including expected and unexpected conditions) are thrown as exceptions and transformed to a standardized JSON response using Java Spark framework's exception mapping, along with an appropriate HTTP status code and additional details, if applicable.

The users-collection and collection-photo relations are many-to-many, so these are stored in junction tables. (The photo-user relation is many-to-one, however, as each photo is directly linked with a particular owner) The database is set up in five tables: one for users, one for collections, one for photos, one serving as the junction table between users and collections (access-control list table), and another serving as the junction table between photos and collections. Foreign keys are implemented to make updates and deletion cascade easily. We use a lot of nested queries and joins to implement the many-to-one and many-to-many relationships.

On the front end, we used React as our main Javascript framework. We use react-redux to store global state (login information) and bind the login state to component props. For API requests, we created a utility class called ApiConnectionManager that manages fetch (AJAX) requests to the backend. Routing was managed using react-router-dom. (Since the backend is running on a different port than the website, proper CORS configuration is set up for all of the JSON endpoints on the backend.) We used Reactstrap (Bootstrap with React bindings) to make our website look more consistent, as well as the Font Awesome icon set. We used the react-photo-gallery library to create a PhotoSelectorList masonry-style photo list display component, which was embedded whenever we needed to display a list of photos (e.g., when listing a user's photos, for each collection, and when selecting a cover photo for a collection or the profile photo for a user).

On deployment, we used docker-compose to set up a network for our containers. We hosted the static website (built using `npm run build` in the React app) on an nginx container, the database in a MariaDB container, and the server (which is packed into an uber-JAR (fatJAR) using maven) on an OpenJDK container. We are currently running this in an EC2 instance for the live demo.

Instructions to run the project are written into the READMEs of the photocol-website, photocol-server, and photocol-DB_SETUP project. The former two are to run the frontend and backend standalone, and the latter is to build and run their respective Docker images.

List of technologies used

Languages: HTML, JS, CSS, (My)SQL, Java
Backend

- Server
    - Java Spark Framework: main HTTP server, streamlines implementation of HTTP protocol, manages cookie-based sessions
    - sl4j: logging service used by the Java Spark Framework
    - GSON: Java implementation of a JSON parser/generator
    - Apache Commons 3: some utility classes
    - Amazon S3 SDK: Java SDK for the S3 service
    - Spring Security Core: used to implement bcrypt for password hashing
    - Metadata extractor: image metadata extractor library
- Server build tools
    - Maven: Java build automation tool and package manager
- Database
    - MariaDB: database technology (drop-in compatible with MySQL)
    - JDBC: Java driver for the database
    - Hikari Connection Pool: for maintaining a connection with the database

Frontend

- Main framework
    - React JS: fully-featured JS framework to handle reactivity and behavior
    - Redux: global state container (mostly used for login information)
    - react-router-dom: easy routing within the app
- Outside libraries
    - Bootstrap: really awesome HTML/CSS/JS library for consistency and responsiveness
    - Reactstrap: React bindings for Bootstrap
    - Font Awesome: really awesome icons
    - react-photo-gallery: nice masonry layout for a list of photos

Deployment

- Containerization:
    - Docker, docker-compose: main container technology
        - adminer: image for database web administration interface
        - mariadb:latest: image for mariadb package
        - nginx:mainline: image for serving the website
        - openjdk:15-jdk-alpine: image for running the packaged server
    - maven-shade plugin: uber-JAR build plugin
- (Ubuntu) Amazon EC2 instance: deployment of Docker images
- Amazon S3: cloud-based storage

Environment

- IntelliJ IDEA Ultimate: IDE
- GitHub (git): version control

Summary of first milestone

The first milestone was focused on implementing the main backend functionality. We began to organize a majority of the major backend endpoints (e.g., endpoints to create users, upload images, and create collections and add images to them). We tested this using a special CLI tool and with curl. Because the project had already reached a stage where persistent storage would be useful, we implemented a MariaDB database. Also, we implemented the S3 SDK. Basic error handling was implemented at this stage.

Summary of second milestone

The second milestone was mainly focused on implementing as much functionality as possible (without focusing too much on quality or design), mostly on the front end. We set up a React webapp and focused on hooking up the major endpoints to the website, and implementing missing endpoints as needed. At the end of this milestone, we had a mostly-working client-facing frontend. We also implemented middleware on the backend and more hierarchical routing endpoints to make routing more neat.

Summary of third milestone

This milestone was mostly about getting the product to be reliable, aesthetically-appealing, complete, and deployable. We implemented Font Awesome and Bootstrap, and spent much time on improving the general aesthetics of the website. There were many UX changes, such as improved error handling (through alerts and form error messages), semantic HTML (such as modals and title attributes), and little tweaks (such as editing the favicon). On the backend, some of the final endpoints were implemented, but most of the work was taken towards reliability (fixing some bugs and adding validation code for user input), extensibility (cleaning up and making code more consistently handle errors, and adding Javadoc-style comments), and security (adding proper salting/hashing for passwords). The database was expanded to include many more fields for users, collections (especially new permissions levels), and photos (especially photo metadata). Public and discoverable photo collections were implemented. The website photo caching policy was fixed so that only images viewable by the currently-logged-in user are accessible. We finished a Docker setup and deployed it to an EC2 instance to host an instance of the website online, and moved values that differed between the deployment and instance and the development instance to environment variables.

Future work

While we were able to accomplish most of our goals and the major functionality of this site, there are still many quality-of-life features (i.e., nice-to-haves) that could be implemented if we were given more time. For example, this could include adding the selected photos on the photos page to a certain collection; searching, filtering, sorting, and pagination features for long lists of collections or photos; more metadata extraction, showing which collections a photo is in; real-time updates when a collection or photo is updated (e.g., through server-sent events). On the technical side, using memcache to have logins persist server restart (right now, restarting the server invalidates all login sessions), finishing up the Docker implementation (including database backups, which are currently a work in progress), and implementing unit testing for the backend and/or the server would be desirable.

Key takeaways

From this project we learned a lot more about dockers as well as learning how to implement different features using react. Learning how to communicate well as a group and work together was another issue that we eventually improved. Creating a website takes a lot of foresight and attention to detail; there are a lot of little features that we consider basic and trivial that sometimes we overlook them are actually essential and make a huge difference. There's a lot that goes into making a website in both the frontend and the backend, which we were able to experience and learn for ourselves.

Team dynamics

After the second milestone was due, we all realized that we wanted the website to not only be aesthetic but also include many features. There were many late nights getting everything implemented, especially docker. Since it is finals week, we are all busy finishing up the semester and busy with final projects and finals. There were some issues when setting up the docker because the credentials kept accidentally being pushed, which compromised the security and caused an issue. For this milestone we held more meetings and worked together, which prevented a lot of conflicts. Additionally, there were some technical difficulties due to one member's computer "exploding," which hindered some progress but we were able to get things to work.

Resources

Github Release:
https://github.com/photocol/photocol-website/releases/tag/v0.3

Server source code:	https://github.com/photocol/photocol-server
Website source code:	https://github.com/photocol/photocol-website
Database setup and dockerfiles:	https://github.com/photocol/photocol-DB_SETUP
Live demo:	http://34.207.93.180/

The Photo Union
for the Advancement of Software and Photo Art

Photocol (backend)
Contributors: Jonathan Lam, Richard Lee, Tiffany Yu, Victor Zhang
GitHub: https://github.com/photocol/photocol-server

Macro-level Overview of the Project
The main server is stored in photocol-server.
DB setup/maintenance operations are stored in photocol-DB_SETUP.
The CLI is stored in photocol-cli.
Specifics about endpoints, error codes, and more are available in the server wiki.

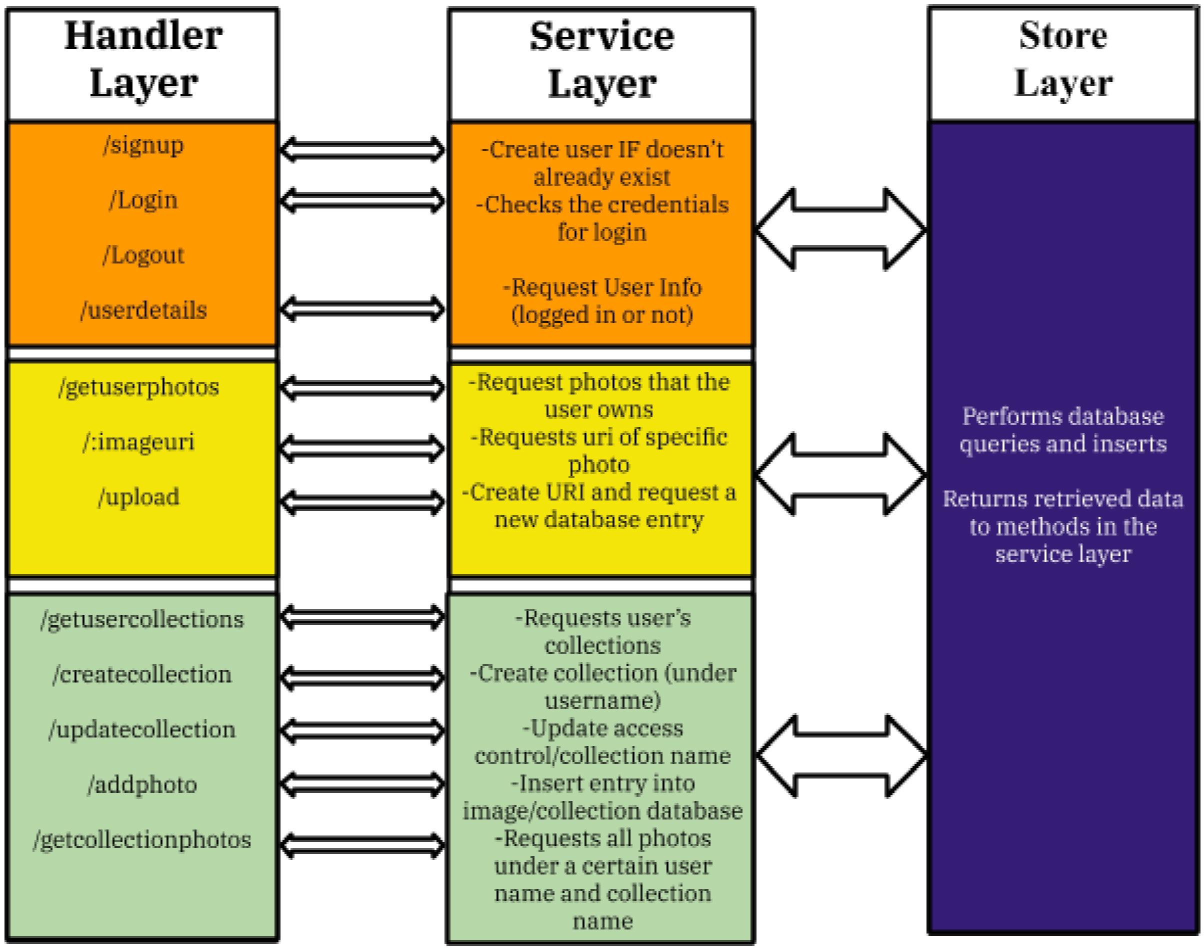
Dependencies: Spark Framework, Gson, SLF4J, Apache Commons 3, Amazon AWS SDK V2, MariaDB JDBC driver. Maven is used to manage dependencies.

User-oriented Overview of the System
The Photo Union is a photo management and collection system, where users can create accounts, upload photos and make collections to share with others.

At the current stage of the project, users can interact with the server via a custom CLI (photocol-cli), or via ordinary HTTP requests. Current features implemented include basic user creation and login session management, ability to create photo collections, ability to upload photos to account, ability to add photos to collections, and ability to get photos from the server, all with basic permissions checking.

Users have roles in collections: Owner, Editor, and Viewer as a form of access control. Images cannot be shared by themselves; rather, access permissions are set on entire collections. Image names (once uploaded to S3) are randomly-generated and unique site-wise; collections are accessible through a semantic URL namespaced by username (e.g., /someuser/theircollectionname, where someuser2 could also have a collection called theircollectionname). (Usernames are unique.)

Black-Box Diagram



High Level Overview of Codebase
Our system has three layers, the handler layer, the service layer, and the store layer. This type of abstraction allows for different team members to work on different aspects of the design without getting too hung up with all the aspects of the system.

The handler layer (package photocol.layer.handler) is responsible for responding to the HTTP requests from the end-user, deserializing and doing (currently minimal) validation on the user, and passing the deserialized objects down to the service layer. (Only the /logout endpoint doesn't pass down data to the next layer, since it only invalidates the http session).

The service layer (package photocol.layer.service) is responsible for the application logic. Some of these endpoints are a simple passthrough (e.g., PhotoService::getUserPhotos) to the store layer, but many of these perform multiple store-level logic (i.e., checking and validation).

The store layer (package photocol.layer.store) manages the database. This layer provides methods to interface with retrieve and insert data into the databases. See comments below about why we decided to use a database.

In addition, we have some utility classes (package photocol.util), which contains a MariaDB and AWS S3 connection manager.

This layered structure also enables us to have 'hot swappable' layers - any changes to the one layer are opaque to another layer. This means that upgrades to each layer are possible without breaking the entire system. The trade-off is that we have to be careful when making changes to our method, that is keep the return types and return values in sync.

Design Decisions (a slightly-lower-level overview)
While databases were out of scope, we decided to implement a MySQL (actually: MariaDB) DB scheme, because we felt that that would be the most logical way to implement all of the access control schemes, especially with joins and built in concurrency management. In particular, this means that in addition to basic tables storing user, photo, and collection data, in order to implement the many-to-one relationship between images to users, we use foreign keys in the photo table (currently not implemented as foreign keys, but that will be changed later). Similarly, for the many-to-many implementation between users and collections (i.e., access control permissions), and the many-to-many implementation between collections and images, we create two additional junction tables just storing lists of relations (pairs of foreign keys), which will be used in joins to check user permissions to images or collections.

Also, we decided to implement S3 capability for uploading/downloading, since this was an integral part of our system and provided a scalable interface for large volumes of uploads and downloads (as opposed to uploading directly onto our server, for which we have not received the specifications of and may not be suited to large network traffic). S3 has the nice-to-have's of huge scalability, high reliability, version control (e.g., ETags for caching/change detection), and concurrency support built-in.

Data and Entity Schemas
Common Java entity classes and their public members are shown. Most of these Java types are only meant to manage the fields of an incoming or outgoing request, and meant to be (more or less) directly serialized/deserialized to/from JSON.

class photocol.definitions.response.StatusResponse<T>: a wrapper for returning status codes with optional data.

enum Status: making application-specific response codes semantic

int status: hold the Status enum response code
T payload: hold the (optional) data payload of the request

class photocol.definitions.User: hold credentials on signup/login requests, or when retrieving user data

String email
String username
String passwordHash (currently-unhashed)

class photocol.definitions.PhotoCollection: hold information about a collection when performing some operation about it

boolean isPublic
String name
List<ACLEntry> aclList: list of access control entries, only for endpoints that request to change/view the access control list to the collection
String uri: this is automatically generated from name to allow collections to be accessible at /collections/:collectionuri

class photocol.definitions.Photo: hold information about an image when updating or retrieving image information

String uri
String description
String uploadDate

The schema for the database tables is shown below. These definitions can be found in the [photocol-DB_setup]() repo.

table photocol.user:

uid int not null auto_increment
email varchar(255) not null unique
username varchar(255) not null unique
password varchar(255) not null
primary key(uid)

table photocol.collection

cid int not null auto_increment
pub tinyint(1) not null
name varchar(255) not null

```
    uri varchar(255) not null
    primary key(cid)

table photocol.photo
    pid int not null auto_increment
    uri varchar(255)
    upload_date date
    description varchar(255)
    uid int not null
    primary key(pid)

table photocol.acl (Access Control Lists)
    cid int not null
    uid int not null
    role int not null
    primary key(cid, uid)

table photocol.icj (Image-Collection Junction table)
    pid int not null
    cid int not null
    primary key(pid, cid)
```

Team Dynamic
The project is definitely moving along rather well - a lot of functionality is coming together. We were/are having some version control/git syncing issues at points - it seems that IntelliJ doesn't automagically update files when we git pull (we have to click Refresh from Disk often). Also, since we are running 4 different OSes across the 4 of us, we do have some package version differences. MariaDB on one of our machines has a 767 max characters per row, which differs from the other three machines, and over the duration of this stage of the project, we were unable to resolve it. We may just have to recompile and reinstall a newer MariaDB.

In terms of team dynamic it seems we are *very* busy people with rather conflicting schedules so meeting up for during the week is a rather difficult task to accomplish. This means that we did have points in our project where two people wrote code for the same functionality because of improper communication. In the future we do need a better way of managing who does what and when in order to not waste anybody's time.

Jonathan Lam
Prof. Abdelwahed
HSS 4
4 / 10 / 20

The CARES Act: Subsidized Normalcy in an Intractable Era

In response to the demand and supply shocks caused by the coronavirus (COVID-19) crisis in the U.S., the two trillion dollar CARES (Coronavirus Aid, Relief, and Economic Security) Act was approved at the end of March to provide emergency relief to many institutions and individuals. Some of the largest components of the act included cash payments to most Americans, increased unemployment benefits, funding for both large and small businesses, and funding for medical institutions and research. While critics argue that it doesn't address some of the underlying problems in society or comes too late, the scale of the emergency has caused this bill and similar ones in many other countries to gain overwhelming support. The CARES Act cannot be compared to previous recessions because of the forced circumstances; we cannot even provide a "stimulus" bill at this time because of mandatory isolation and the resulting inability to buy goods, so the best that can be done is financial "lifelining." Given the current overall state of the nation and our prior economic situation, the CARES Act is a very necessary fiscal policy to fight the recession: it offers relief to those in the most dire situations, and it offers incentives to keep unemployment low (which will help recovery).

In providing a temporary universal basic income for low-income Americans, the CARES Act maintains a basic standard of living for low-income Americans, which prevents escalating the disaster further. It's no secret that income inequality in the U.S. has been growing in recent years, but this virus exacerbates the problem. Those in the lowest income bracket tend to have jobs in the service industry (increasing their likelihood for catching the disease), have less access to health insurance (increasing mortality rate), and have a smaller unemployment wealth buffer (forcing them to stay on their job, increasing the likelihood for spreading the disease). In an article by the New York Times, Fisher summarizes these effects, concluding that "inequality itself may be acting as a multiplier on the coronavirus's spread and deadliness," and claiming that other historical crises also tended to increase inequality (Fisher, 2020). The CARES Act has money allocated to providing almost all adults with a small income and increased unemployment insurance, which will help mitigate two of these effects and thus decreasing the inequality "multiplier." While the lowest income bracket is most in need of this basic income, roughly half of all Americans live paycheck-to-paycheck (Heflin, 2020), so this should help a great number of people be able to continue living pseudo-normally under isolation conditions, either with the government-subsidized incomes or unemployment insurance. With assurance of an income, improved insurance, and tax cuts, the CARES Act mitigates the spread of economic inequality and curbs panic, either of which would worsen the current situation. While the Act certainly doesn't solve the problems of high inequality (rather, it exposes many areas of improvement for low-income people, such as poor access to healthcare and high student loans), maintaining the economic status quo may be the best we can do in an emergency situation.

In addition to providing immediate relief to low-income Americans, another important factor of the CARES Act is its measures to keep many businesses afloat to allow for recovery after the recession. Similar to low-income individuals, many small businesses don't have funds to last very long without revenue; a recent report found that small businesses on average only have enough funds available for twelve days without revenue (Farrell et al., 2018). Mandated isolation periods and the closure of "non-essential businesses" have clearly exceeded that interval, even far before the release of the CARES Act, which forced many small businesses to close permanently or indefinitely. Even if the tsunami of unemployed workers could be supported by this temporarily bolstered unemployment insurance, it is an unsustainable pattern. Economics professor Ricardo Reis notes that the "re-matching of markets, companies and workers often makes recessions prolonged and painful. It is better to stop businesses failing in the first place, if the lockdown is going to be short-lived" (Reis, 2020). Fortunately, almost half of the CARES Act is allocated towards businesses toward fulfilling Reis's philosophy, with small businesses receiving a sizable amount: three-quarters as much as large businesses (Routley, 2020). While the methods and terms by which the government subsidizes companies differs, a similar policy is in place throughout modern countries facing the pandemic. Denmark, for example, subsidizes up to three-quarters of a worker's salary, under the condition that the workers stay employed and that they stay isolated; in France, the government will subsidize the entire worker's wages if the company cannot provide it (Synder, 2020). In the U.S., there is less of a guarantee of direct government subsidy of wages, but small businesses are incentivized with loans to continue to keep their workers employed and to rehire recently laid-off employees. While the U.S. doesn't force its workers to stay at home in order to receive government subsidy (like Denmark) or offer seamless handoff of paying wages to the government (like France), it does encourage businesses to keep employment and offers some flexibility on what to spend on.

Facing an unprecedented pandemic with the potential for a massive economic downturn, whatever that can be done to help citizens maintain their economic status and keep businesses open until the crisis is over is an effective measure. The CARES Act is a lifeline bill tackling both of these major problems, and thus is effectively helping us mitigate the economic disaster as we wait for the health disaster to pass us by.

Works Cited

Farrell, Diana, et al. “Growth, Vitality, and Cash Flows.” *J.P. Morgan Chase and Co. Institute*, J.P. Morgan Chase and Co., July 2018, https://institute.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/institute/pdf/institute-growth-vitality-cash-flows.pdf.

Fisher, Max, and Emma Bubola. “As Coronavirus Deepens Inequality, Inequality Worsens Its Spread.” *The New York Times*, The New York Times, 15 Mar. 2020, www.nytimes.com/2020/03/15/world/europe/coronavirus-inequality.html.

Heflin, Jay. “Nearly Half of Americans Live Paycheck to Paycheck: Bank Survey.” *Washington Examiner*, Washington Examiner, 20 Feb. 2020, www.washingtonexaminer.com/news/nearly-half-americans-live-paycheck-to-paycheck-bank-survey.

Reis, Ricardo. “How Do Countries Differ in Their Response to the Coronavirus Economic Crisis?” *The Guardian*, Guardian News and Media, 3 Apr. 2020, www.theguardian.com/commentisfree/2020/apr/03/coronavirus-economic-crisis-financial-impact.

Routley, Nick. “The Anatomy of the $2 Trillion COVID-19 Stimulus Bill.” *Visual Capitalist*, Visual Capitalist, 30 Mar. 2020, www.visualcapitalist.com/the-anatomy-of-the-2-trillion-covid-19-stimulus-bill/.

Snyder, Stephen. “How the US Coronavirus Stimulus Package Compares to Those of Europe.” *Public Radio International*, PRI Public Radio International, 3 Apr. 2020, www.pri.org/stories/2020-04-03/how-us-coronavirus-stimulus-package-compares-those-europe.

Jonathan Lam
Prof. Abdelwahed
HSS4 — Keynesian Economics
02/07/20

The Rise of a Dynamic Monetary Policy
Response to Keynes' Essay "The End of the Gold Standard"

Since ancient times, gold and money have been virtually synonymous. However, in 1931, Britain broke free of the gold standard that implemented the correspondence between money and gold, and John Maynard Keynes argues in his essay, "The End of the Gold Standard," that the gold standard only created "an intolerable strain" on national economies, motivated only by a historical "obligation of *honour*" to couple currencies with gold. He welcomes the change, citing its use in combating economic stall and the problematic concentration of gold in trade-surplus nations.

Keynes believes the gold standard is only "quixotry" — a romantic, old-fashioned want to keep money at its original worth; thus he cites contemporary and logical evidence to debunk its importance, such as the revitalization of the Stock Exchange in the weeks following the separation from the gold standard. He also reasons that the resulting decrease in national purchasing power is not detrimental to trade: a currency depreciation of 25% offers protectionist restrictions on imports, *and* a greater return on exports (i.e., a "bounty of the same 25 per cent"); benefits the economy bidirectionally and bilaterally; contrastly, tariffs are unidirectional. Keynes believes that becoming "the cheapest producer in the world in terms of gold" is of no consequence: maintaining an equivalent value in gold is not as important as keeping the economy flowing.

Then what are the fiat standard's downsides? For the countries following this new standard, Keynes presents no disadvantage, as each country may adjust their exchange rate to match their needs. However, the two major proponents of the gold standard — namely, France and the U.S., which had amassed much of the world's gold supply from trade or war reparations, and whose hoarding of gold led to its global deficit — are disadvantaged as their high-valued gold-backed currencies discourage trade.

Despite Keynes' optimism, the gold standard protected against hyperinflation, which plagued the Weimar Republic in 1921 after it left the gold standard in 1914; a prior and potential mishap Keynes fails to mention. Secondly, the fairness of currency manipulation is an activate debate, especially between inflated currencies in large national producers such as China, leading to the two-way benefits Keynes describes; this led to IMF regulation of currency exchange rates which

> "ensure that such countries regularly and promptly adjust the rate of exchange between their currencies and the United States dollar to permit effective balance of payments adjustments and to eliminate the unfair advantage" (United States, Dept. of Treasury, 1988).

One of the factors of the "trade war" between the U.S. and China, currently the two largest national economies, is that the IMF does not consider China as a currency manipulator while President Trump's administration does (United States, Dept. of Treasury, 2019). Now that the success of currency fluctuations as government policy has established it as an international norm, the legacy of the separation from the gold standard is how to determine when these measures go too far, and how to prevent against the possibility of the massive inflation that it promotes.

Works Cited

Keynes, John Maynard. "The End of the Gold Standard." *Essays in Persuasion*, 1931, https://gutenberg.ca/ebooks/keynes-essaysinpersuasion/keynes-essaysinpersuasion-00-h.html.

United States, Congress, "Omnibus Trade and Competitiveness Act of 1988 (H.R. 3)." *United States Department of the Treasury*, 1988.

United States, Congress, "Treasury Designates China as a Currency Manipulator." *United States Department of the Treasury*, 5 Aug. 2019, https://home.treasury.gov/news/press-releases/sm751.

7. The End of the Gold Standard
(Sept. 27, 1931)[29]
[29][On Sept. 21, 1931, the Gold Standard in Great Britain was suspended.]

There are few Englishmen who do not rejoice at the breaking of our gold fetters. We feel that we have at last a free hand to do what is sensible. The romantic phase is over, and we can begin to discuss realistically what policy is for the best.

It may seem surprising that a move which had been represented as a disastrous catastrophe should have been received with so much enthusiasm. But the great advantages to British trade and industry of our ceasing artificial efforts to maintain our currency above its real value were quickly realised.

The division of inside opinion was largely on a different point. The difficult question to decide was one of honour. The City of London considered that it was under an obligation of honour to make every possible effort to maintain the value of money in terms of which it had accepted large deposits from foreigners, even though the result of this was to place an intolerable strain on British industry. At what point—that was the difficult problem—were we justified in putting our own interests first?

As events have turned out, we have got the relief we needed, and, at the same time, the claims of honour have been, in the judgement of the whole world, satisfied to the utmost. For the step was not taken until it was unavoidable. In the course of a few weeks the Bank of England paid out £200,000,000 in gold or its equivalent, which was about half the total claims of foreigners on London, and did this at a time when the sums which London had re-lent abroad were largely frozen. No banker could do more. Out of the ashes the City of London will rise with undiminished honour. For she has played the game up to the limits of quixotry, even at the risk of driving British trade almost to a standstill.

No wonder, then, that we feel some exuberance at the release, that Stock Exchange prices soar, and that the dry bones of industry are stirred. For if the sterling exchange is depreciated by, say, 25 per cent, this does as much to restrict our imports as a tariff of that amount; but whereas a tariff could not help our exports, and might hurt them, the depreciation of sterling affords them a bounty of the same 25 per cent by which it aids the home producer against imports.

In many lines of trade the British manufacturer to-day must be the cheapest producer in the world in terms of gold. We gain these advantages without a cut of wages and without industrial strife. We gain them in a way which is strictly fair to every section of the community, without any serious effects on the cost of living. For less than a quarter of our total consumption is represented by imports; so that sterling would have to depreciate by much more than 25 per cent before I should expect the cost of living to rise by as much as 10 per cent. This would cause serious hardship to no one, for it would only put things back where they were two years ago. Meanwhile there will be a great stimulus to employment.

I make no forecast as to the figure to which sterling may fall in the next few days, except that it will have to fall for a time appreciably below the figure which cool calculators believe to represent the equilibrium. There will then be speculation and profit-taking in favour of sterling to balance speculation and panic selling on the other side. Our authorities made a great mistake in allowing sterling to open so high, because the inevitable gradual fall towards a truer level must sap confidence and produce on the ignorant the impression of a slide which cannot be stayed. Those who were guilty of undue optimism will quite likely succumb to undue pessimism. But the pessimism will be as unfounded as the optimism was. The equilibrium value of sterling is the same as it was a month ago. There are tremendous forces to support sterling when it begins to fall too far. There is no risk, in my judgement, of a catastrophic fall.

These, in brief, are the consequences in Great Britain. How will the rest of the world be influenced? Not in a uniform way. Let us take first the debtor countries to whom Great Britain has in the past lent large sums in sterling, and from whom interest is due in sterling, such as Australia, Argentina, and India. To these countries the depreciation of sterling represents a great concession. A smaller quantity of their goods will be sufficient to meet their sterling liabilities. The interest due to Great Britain from abroad, which is fixed in sterling, amounts to about £100,000,000 a year. In respect of this sum Great Britain now plays the part of a reasonable creditor who moderates his claim in view of so great a change in the situation as the recent catastrophic fall in commodity prices.

When we try to calculate the effect on other manufacturing countries, whose competition we are now in a better position to meet, the effect is more complex. A large part of the world will, I expect, follow Great Britain in reducing the former gold value of their money. There are already signs in many countries that no great effort will be made to maintain the gold parity. In the last few days Canada, Italy, Scandinavia have moved in our direction. India and the Crown Colonies, including the Straits Settlements, have automatically followed sterling. Australia and the whole of South America had already abandoned the effort to maintain exchange parity. I shall be astonished if Germany delays long before following our example. Will Holland deal final ruin to the rubber and sugar industries of the Dutch Indies by keeping them tied to gold? There will be strong motives driving a large part of the world our way. After all, Great Britain's plight, as the result of the deflation of prices, is far less serious than that of most countries.

Now, in so far as this is the case, we and all the countries following our example will gain the benefits of higher prices. But none of us will secure a competitive advantage at the expense of the others. Thus the competitive disadvantage will be concentrated on those few countries which remain on the gold standard. On these will fall the curse of Midas. As a result of their unwillingness to exchange their exports except for gold their export trade will dry up and disappear until they no longer have either a favourable trade balance or foreign deposits to repatriate. This means in the main France and the United States. Their loss of export trade will be an inevitable, a predictable, outcome of their own action. These countries, largely for reasons resulting from the war and the war settlements, are owed much money by the rest of the world. They erect tariff barriers which prevent the payment of these sums in goods. They are unwilling to lend it. They have already taken nearly all the available surplus gold in the whole world. There remained, in logic, only one way by which the rest of the world could maintain its solvency and self-respect; namely, to cease purchasing these countries' exports. So long as the gold standard is preserved—which means that the prices of international commodities must be much the same everywhere—this involved a competitive campaign of deflation, each of us trying to get our prices down faster than the others, a campaign which had intensified unemployment and business losses to an unendurable pitch.

But as soon as the gold exchange is ruptured the problem is solved. For the appreciation of French and American money in terms of the money of other countries makes it impossible for French and American exporters to sell their goods. The recent policy of these countries could not, if it was persistently pursued, end in any other way. They have willed the destruction of their own export industries, and only they can take the steps necessary to restore them. The appreciation of their currencies must also embarrass gravely their banking systems. The United States had, in effect, set the rest of us the problem of finding some way to do without her wheat, her copper, her cotton, and her motor-cars. She set the problem, and, as it had only one solution, that solution we have been compelled to find.

Yet this is quite the opposite of the note on which I wish to end. The solution to which we have been driven, though it gives immediate relief to us and transfers the strain to others, is in truth a solution unsatisfactory for every one. The world will never be prosperous without a trade recovery in the United States. Peace and confidence and a harmonious economic equilibrium for all the closely interrelated countries of the globe is the only goal worth aiming at.

I believe that the great events of the last week may open a new chapter in the world's monetary history. I have a hope that they may break down barriers which have seemed impassable. We need now to take intimate and candid conference together as to the better ordering of our affairs for the future. The President of the United States turned in his sleep last June. Great issues deserve his attention. Yet the magic spell of immobility which has been cast over the White House seems still unbroken. Are the solutions offered us always to be too late? Shall we in Great Britain invite three-quarters of the world, including the whole of our Empire, to join with us in evolving a new currency system which shall be stable in terms of commodities? Or would the gold standard countries be interested to learn the terms, which must needs be strict, on which we should be prepared to re-enter the system of a drastically reformed gold standard?

Jonathan Lam
Prof. Abdelwahed
HSS4 — Keynesian Economics
1/31/20

A Single Statistic in a Big Data World

A nation's GDP is a monetary measure of a nation's material production, arising in the booming WWII mass-material-production era. It was not designed to be *the* measure of the strength of a nation's economy, but it has established itself as a catch-all metric when discussing large-scale economic strength and changes. Newer competing proposals, such as Stiglitz's "dashboard" or Sen's "capabilities" approach, illustrate that the GDP too rigid to adapt to changes in the economic value—evolving ideas of the worth of inequality and of certain products— causing it to fail in any qualitative or long-term analysis.

In his podcast with WNYC, Stiglitz emphasizes the disparity between changes in GDP and the economy following the Great Recession. There is a "selective" growth in GDP, but the employment rate is historically low, and the majority population feels a great deal of anxiety (Lehrer & Stiglitz, 2019). According to Stiglitz, 91% of the GDP's growth between 2009 and 2012 went to the wealthiest 1%, which shouldn't measure economic success: "the top 1% was feeling a recovery," but the vast majority gain no economic power (disregarding the robotic "potential Pareto optimality" argument (Stanton, 2007)). Lack of economic insight in netting income inspires Stiglitz and Sen's newer "capabilities" approach to measuring economic growth, which combines the analysis of multiple factors indicative of a population's overall economic health, such as freedoms and health (Coyle, 2017). This is more powerful in that it is designed to measure economic *behavior*, albeit compromising objectivity.

But even the GDP's claim to objectivity is faltering, as the boundary between "final market" and "intermediate" product and economic value becomes more ambiguous. The technology market is a contributor and motivator towards a different value-based approach: for example, the BEA doesn't consider advertising a final product, so "a pay-per-view business model to an ad-supported model reduces GDP" (Varian, 2016) — this arguably discounts billions of dollars worth of productivity in the advertising industry. Personal and business computing power are becoming cheaper and more valuable by the day, skewing the productivity-price correspondence. Lastly, the rise of data collection and analysis firms allow for more multifaceted measures than the GDP (such as the Sen-supported HDI) to be developed and globally disemminated.

We're far past a war economy focused on material goods; the changing demographic and technology younger than the GDP and the effects they have on the average consumer's economic behavior form a set of statistics that can more holistically describe economic health and growth.

Works Cited

Coyle, Diane. “Rethinking GDP.” *Finance and Development*, Mar. 2017, pp. 17-19.

Lehrer, Brian, and Joseph Stiglitz. “WNYC.” *WNYC*, New York Public Radio, 30 Dec. 2019, https://www.wnyc.org/story/beyond-gdp-stiglitz/.

Stanton, Elizabeth A. “The Human Development Index: A History.” *University of Massachusetts Amherst*, Political Economy Research Institute, 2007, pp. 1–5.

Varian, Hal. “Intelligent Technology.” *Finance and Development*, Sept. 2016, pp. 6–9.

Jonathan Lam

Prof. Abdelwahed

HSS4 — Keynesian Economics

Original: 1/31/20; Revised: 3/27/20

The GDP: A Narrow Statistic in a Big Data World

A nation's GDP is the monetary measure of a nation's material production. Its usefulness as a measure of economic strength was clear when it arose in the booming WWII era of the mass production of material goods. However, while it was not designed to be *the* measure of the strength of a nation's economy, it has established itself as a catch-all metric when discussing large-scale economic strength and changes. There are two major flaws with the GDP in describing economic strength: it is a single number that lacks insight on many important economic indicators such as employment or financial certainty; and despite historical claims, it cannot be reliably measured in some modern industries. Together, these flaws suggest that when discussing economic strength, it may be more useful to consider a more holistic approach consisting of more specific statistics.

The largest flaw of the GDP is that it doesn't offer a holistic view of a nation's economy. Older schools of thought robotically proposed that growth in GDP would correspond with a stronger economy by the idea of "potential Pareto optimality" (Stanton, 2007): the idea that an increase in wealth would increase the capacity of the economy, which would eventually reach all members of the economy. However, this is often false; in his podcast with WNYC, Stiglitz emphasizes that there are seemingly opposite trends in GDP and the overall economy following the Great Recession. How can this be the case? Stiglitz describes the growth in GDP as "selective": 91% of the GDP's growth between 2009 and 2012 went to the wealthiest 1% (Lehrer, 2019), a fact that doesn't indicate *overall* economic growth. In other words, "the top 1% was feeling a recovery," but the vast majority of the population gained no economic power. Stiglitz also cited that the employment rate is historically low, and the majority population feels a great deal of anxiety (Lehrer, 2019), which are additional indicators that the economy is not faring well, despite the growth in GDP. In short, the GDP attempts to measure a complex and multifaceted entity into a single number, but this number may hide phenomena such as unemployment and individual economic stability. In response to these concerns, Stiglitz and Sen propose a newer approach involving measuring multiple human "capabilities" rather than production. This analysis combines multiple factors indicative of a population's overall economic health, such as freedoms and health (Coyle, 2017). While it is less objective and more difficult to measure than the GDP, it is more useful in that it attempts to measure how effectively the money can be used (how capable citizens are of using that money), rather than providing the sum total money produced (which, if not distributed wisely, may not indicate much about an economy).

Historically, a major justification for the use of GDP as an economic indicator was that it is an objective, practical measure of a nation's economy (as opposed to more subjective measures such as the "capabilities" approach), but this cannot be justified in the modern economy. Firstly, a clear distinction between "final market" and "intermediate" products is important to the definition of the GDP — the GDP doesn't include intermediate products, only final market items — but this distinction is ambiguous with modern goods. For example, the BEA doesn't consider advertising a final product, so "a pay-per-view business model to an ad-supported model reduces GDP" (Varian, 2016) — this arguably disregards from the GDP billions of dollars worth of productivity in the advertising industry. Secondly, our perception of and value of goods rapidly changes as digital technology advances at an exponential rate. Personal and business computing power are becoming both cheaper and more powerful by the day, skewing our idea of the monetary value of goods, i.e., the relation between productivity and price. The GDP was first proposed and widely used in the WWII era, where an important measure of the value of an economy was its production of wartime supplies, which could be easily measured and where the value of goods was not too volatile; however, we cannot apply that to the highly turbulent modern consumer industry.

We're far past a war economy focused on material goods; the rapidly-changing demographic and technology younger than the GDP makes it impossible to holistically summarize a national economy in a single statistic. To effectively describe the overall health of a modern national economy, more advanced and multifaceted measures have to be developed.

Notes on revisions (3/27/20):

Like in the midterm final draft, I tried my best to address your comments by making the overall format of this essay more clear, with clearer topic sentences and thesis statement, clearer and shorter sentences, and more logical paragraph content.

Works Cited

Coyle, Diane. “Rethinking GDP.” *Finance and Development*, Mar. 2017, pp. 17-19.

Lehrer, Brian. “WNYC.” *WNYC*, New York Public Radio, 30 Dec. 2019, https://www.wnyc.org/story/beyond-gdp-stiglitz/.

Stanton, Elizabeth A. “The Human Development Index: A History.” *University of Massachusetts Amherst*, Political Economy Research Institute, 2007, pp. 1–5.

Varian, Hal. “Intelligent Technology.” *Finance and Development*, Sept. 2016, pp. 6–9.

Jonathan Lam

Prof. Abdelwahed

HSS4

3/13/20

Anti-Immigrant Sentiments: Failed Attempts at Remedying the Great Depression Crisis

The prevalence of nativist sentiment dramatically reinforced in the U.S. with the Immigration Act of 1924, which issued quotas on different ethnic groups' entries into the U.S. Its implementation in 1929 coincided with the onset of the Great Depression and introduced a massive stop onto the flow of immigrants; U.S. census data from the 1820 to 2018 show that the number of immigrants gaining resident status per year, a number which had not dropped below 100,000 since the Civil War and peaked to over 1.2 million in the early 20th century, dropped from almost 300,000 in 1929 to 23,000 in 1933 (U.S. Department of Homeland Security, 2016). In effect, we observe a race-targeted movement that favored certain "preferred" immigrant groups and spurned certain others. While the original rationale behind many of these acts was initially based on eugenics (by Americans trying to keep its racial makeup more homogeneous), the nature of the arguments became more economic as the Great Depression progressed (Hoffman, 1974). Despite justification by the U.S. government, most of these anti-immigrant events of the early 1930's, such as the Mexican Repatriation and the quotas of the Immigration Act of 1924, exacerbated the economic crisis of the Great Depression. These acts hurt a historically hardworking population while idolizing other populations without sound economic reason, wrongfully attacked American citizens solely on account of race (especially in the case of the Mexican Repatriation), and led to no improvements in the economy for non-immigrant U.S. citizens and perhaps even exacerbating the financial situation.

While the late 19th century and early 20th century involved several pieces of legislation against immigrants, the Immigration Act of 1924 was perhaps the most extensive. In 1882, the

Chinese Exclusion Act was passed after the immigration of hundreds of thousands of Chinese into the U.S. in the prior few decades (Ngai, 1999). The so-called “Gentlemen's agreement” restricted Japanese immigration to the U.S. in 1907. The Immigration Act of 1917 imposed English literacy tests for immigration, making it difficult for many non-Anglo immigrants (especially Eastern Europeans and Asians) to enter. The Emergency Quota Act was passed in 1921, setting quotas on immigrants based on their percentage of the U.S. population in 1910, aimed to increase ethnic homogeneity amongst Americans, and thus restricting minorities proportionally to their rarity in the U.S. population. This was followed by the Immigration Act of 1924, which lowered the quotas further and prevented any immigration from Asia. As we can see in Figure 1, the largest drops in immigrants occurred around 1930 when the Immigration act of 1924 was put in place.

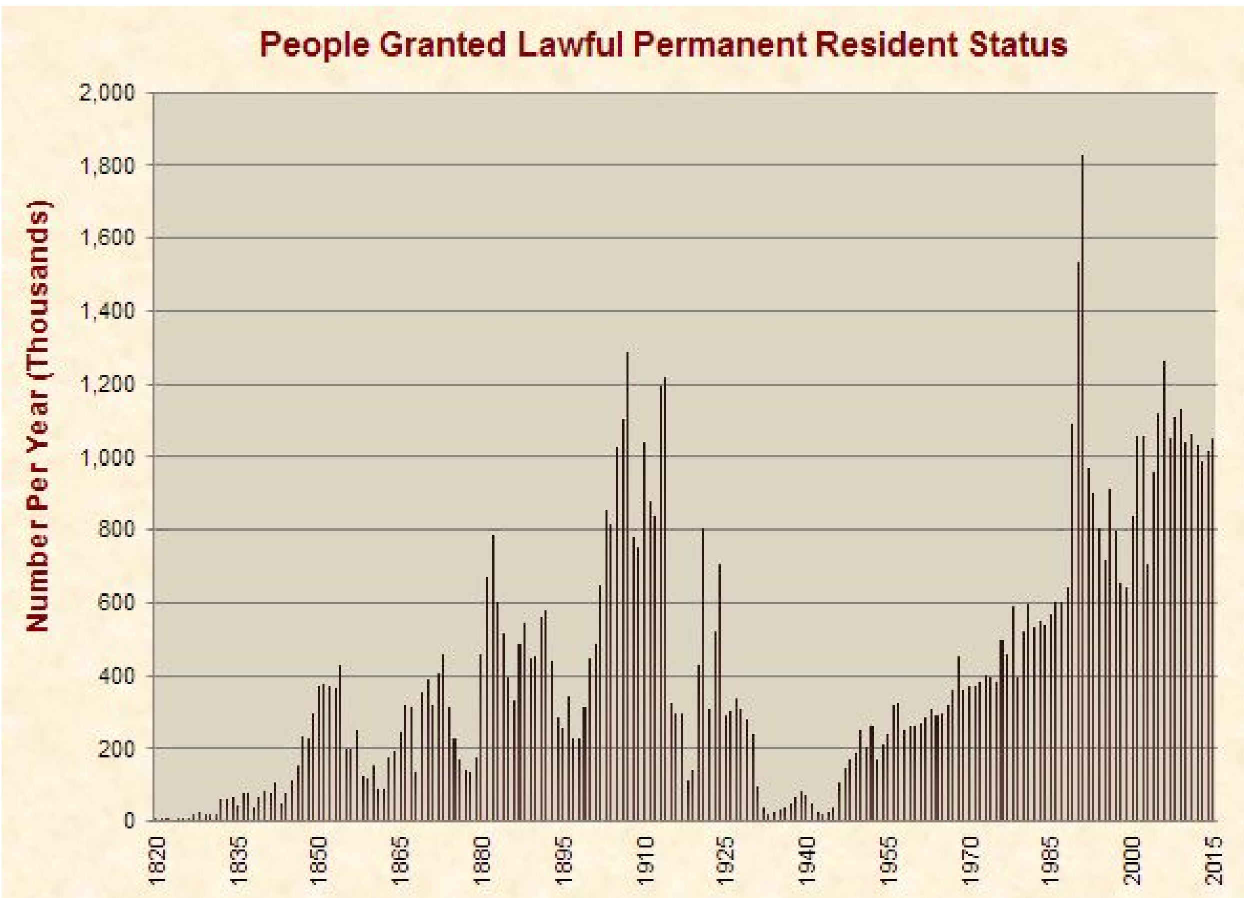


Figure 1. People Granted Lawful Permanent Resident Status, by Year.

Figure published at https://www.justfacts.com/immigration.asp, visualizing data from the "2015 Yearbook of Immigration Statistics," U.S. Department of Homeland Security.

While the Depression may have greatly compounded the drop in immigration, it doesn't discount the fact that this was the greatest change to the trend of immigration in the history of the U.S.. Even if economic factors were a greater contributor than sociopolitical factors, immigration reached all-time lows in the 1930s, at the same time that the Immigration Act of 1924, which limited immigration counts to 20% of their pre-WWI counts, was implemented (Murrin, 2015). As this act was passed before the Depression, the primary argument for limiting immigrant usage was for non-economic reasons: Senator Reed stated that "disregards entirely those of us who are interested in keeping American stock up to the highest standard—that is, the people who were born here" (Stephenson, 1926). In other words, eugenics was the initial main driving force for limiting immigration.

Firstly, we find that quality of life diminished for the average, non-English speaking immigrant. We turn to a study on what Inwood calls "reverse assimilation" for immigrant groups in Canada in the 1930s (Inwood et al., 2014). The term "reverse immigration" indicates that not only are new immigrants discouraged, but even the long-established non-Anglo immigrants were being financially uprooted, causing older immigrants to de-assimilate. Canada is relevant because it passed immigration acts roughly matching those of the U.S., including immigration quotas by national origin; Canada also implemented a better system of tracking citizens' earnings. Using earning reports from 1901-1931, Inwood found that there was a significant gap in earnings between "free" or "preferred" U.S. or British immigrants, and the non-preferred ones (Inwood et al., 2014). Inwood further finds that older immigrants appear to fare even worse than newer ones, as their decreased ability to invest in "language human capital" (i.e., to learn English) greatly impaired their employment opportunities. Thus this was a positive-feedback spiral that led to an increasingly large employment gap: the minority immigrant groups in Canada tended to have fewer individuals proficient in English, which limited their ability to immigrant, which limited their representation out of their population, which decreased their quota numbers; and vice a versa for larger immigrant groups of Anglophone nations. In addition, this led to the widening social disdain of minority groups and the belief that they were often public charges (i.e., likely to become dependent on the government). In other words, the disadvantages of not knowing English were compounded by these new acts and made the life of immigrants dramatically more difficult; this in turn hurt their financial stability, which made them appear as second-class citizens to non-immigrants. To reiterate, due to the vast similarities between Canadian demographics and immigration policy to the U.S., the general principle can be applied to the U.S. as well.

As the economic situation of the Great Depression worsened, the arguments against immigrants shifted more toward claims that they were less financially reliable and more prone to be dependent on the government. However, this argument is caused by the non-immigrant majority leveraging a popular argument in order to continue driving their eugenic goal forward, and does not have much grounding in fact. More respectable immigrant classes with larger quotas, such as Irish Americans, were just as prone to becoming dependent on the government; indeed, this particular group was hit especially hard by the Stock Market Crash of 1929, and in the years following many had “carried the derogatory label of ‘Returned Yank’” or “often cut off contact with family members back in Ireland because they were too embarrassed or guilty about their inability to meet their familial responsibilities” (O’Brien, 2002). Thus, even though this group was welcomed into America with open arms, they were no more able than other immigrant families (and perhaps more panicky and less prudent, as O’Brien’s depiction may show). Except perhaps the related work by Inwood and the inability of immigrants to find work, there is no strong evidence that non-preferred immigrant groups were less capable or hardworking than non-immigrant or preferred-immigrant groups; rather, the Asians’ efficiency and risk-tolerance in the dangerous and difficult jobs such as building the Transcontinental Railroad was a major reason it could be finished so quickly, and this helped foster jealousy and hatred amongst non-immigrants who didn’t take those jobs. This discredits the common contemporary sentiment that “preferred” immigrant groups were more financially capable and recovered better than “non-preferred” groups.

While many immigrant populations were blocked from entering the U.S. and were hurt by a poor English ability, one ethnic group was hit especially hard during the Great Depression: Mexican Americans. Like Filipino Americans, Mexicans were exempt from the quotas of the Immigration Act of 1924, and large numbers emigrated to the U.S. in the early 1900's, bolstering the farm economy of the southwestern states; they played an important part of the labor force like Asian Americans of the late 1800's in the Gold Rush and the railroad industry. With stronger proponents of eugenics and the worsening economic situation came the Mexican Repatriation, a policy beginning in 1929 supported by both the American and Mexican governments that involved "repatriating" many Mexican migrant workers in the U.S. to Mexico. Historian Abraham Hoffman describes the American supporters of the Repatriation to belong to several groups: "small farmers, progressives, labor unions, eugenicists, and racists" (Hoffman, 1974). The former three camps claimed that their presence occupied American jobs, and many Americans were on relief or were public charges; eugenics and racists found the Mexican barrios and the difference in physical appearance of Mexicans easy to use as a scapegoat for economic troubles. In the next few paragraphs, I find that both of these arguments are unsubstantiated.

First of all, the legitimacy of many of the "repatriations" is questionable; the action involved the deportation of more American citizens than the repatriation of citizens of Mexico. Estimates for the repatriation are difficult to pin down — there was no mechanism in place to accurately measure this kind of trend, much of it was illegitimate (deporting U.S. citizens on basis of appearance), the efforts were decentralized, and contemporary authors were not always consistent with their claims. A report by Hoffman designated specifically to the statistics of the Mexican repatriation claims that commonly-cited contemporary author Carey McWilliams often was inconsistent with his estimates in different articles (Hoffman, 1972); his own estimate was that of approximately 400,000 (Hoffman, 1974), although other estimates have ranged to almost two million (Ray, 2005); these and other reviews also conclude that the majority of the people (60%) forcibly-moved were U.S. citizens. The simplest conclusion to draw from this lack of clarity is how hasty and otherwise poorly-implemented this policy was put in place. The only result that this creates is the reinforcement of Mexican Americans (migrant workers or valid U.S. citizens), are second-class citizens who don't even have the basic right of protection from discrimination-based, unlawful deportation. The forced deportation of many American citizens also demonstrates the highly racialized aspect of the policy. Indeed, from a purely economic perspective, Mexicans migrant workers in the U.S. had been beneficial to the U.S. economy as mentioned above; however, the social (i.e., ethnic) aspects cause problems when it comes to the U.S. deciding on immigration policy. Aguila succinctly summarizes the ongoing dilemma for U.S. policymakers:

> "A significant cause of this quandary results from the obvious, but problematic, fact that each nation benefits from the existence of Mexicans (documented and undocumented) in the United States. The remittances from workers are an obvious subsidy for the Mexican economy, which today totals nearly ten billion dollars per year. However, the gains for U.S. society are highly controversial despite major sectors of the economy's dependence on this labor" (Aguila, 2007).

The fact that "the gains for U.S. society are highly controversial" emphasize the complexity of U.S.-Mexican relations past the facade of economic gain for Americans. Moreover, one of the major reasons the Mexican government supported the repatriation was that it would reintroduce many skilled workers from the U.S. back into their society — conversely, this implies that the U.S. was losing much skilled labor (that was willing to work at a lower pay rate), which would undoubtedly hurt the U.S. economy in the long term, even if there was a current economic downturn. Thus the Mexican American community was greatly economically maligned due to a largely social cause.

But did the Mexican Repatriation help with non-immigrant American employment as it was intended to do? Even if it did negatively affect the Mexican Americans, did it benefit non-immigrants? To assume that the repatriation of Mexicans would benefit Americans financially makes the assumption that there were enough Mexican migrants such that replacing their collective roles with a fewer number of American workers would be able to improve wages and fulfill their roles. In reality, since the pay rate of Mexican workers was much lower, this assumption was unrealistic. A statistical study with controls on the employment levels using available census data found that, in addition to decreasing employment for migrants, the repatriation was associated with a nonpositive net effect for American native citizens, leading to a slight increase in American unemployment (Lee et al., 2017). Thus the Mexican Repatriation not only harmed many Mexican Americans, but it was a situation of Pareto worsening — i.e., nobody (except perhaps the eugenicists from a social viewpoint) benefitted economically from this emergency measure.

One other major non-preferred immigrant group was heavily targeted during the Great Depression: Filipino Americans. Since the Philippines were a U.S. territory at the time, Filipinos were legal American citizens and thus did not have quotas placed on them; like with Mexico, Filipino immigration boomed in the 1920's. Especially with Chinese immigration curtailed in the 1880's with the Chinese Exclusion Act and Japanese immigration slowed in the 1910's following the Gentleman's Agreement, Filipinos played an important role in filling the role of the Chinese and Japanese (Flores, 2004). Despite their important role in the economy, they were also subjected to harsh racism like the other Asian immigrant groups, as well as the growing economic concern of job availability. This led to numerous conflicts, the most notorious being the Watsonville Riots in California in 1930, a deadly attack of local farmers on a Filipino institution mainly caused by them worrying about Filipino workers taking the scarce jobs. This set a precedent for further attacks and anti-Asian sentiment amongst non-immigrants, which eventually became a primary cause for the U.S. to lessen their occupation of the Phillipines, which in turn allowed them to restrict quotas on Filipinos with the Tydings Mcduffie Act (Sobredo, 2018).

The question of immigration policy cannot be ever dismissed as trivial, and this case study of various immigration events related to the Great Depression era demonstrates the sort of racial scapegoating that can easily arise as a result of economic panic. The anti-immigrant sentiment was originally a matter of eugenics, but as the Depression worsened and the immigrants were increasingly maligned by the immigration law, native U.S. citizens turned their excuse into an economic claim, which is largely unjustified, as shown by Lee et al. Furthermore, we see the general financial stability of immigrants decline, and there was an increase in racism as a means of venting economic frustration, but none of it benefitted any party.

Despite this insight into the Great Depression, we still see parallels to nativist socioeconomic concerns even today; for instance, one year before the 2008 Great Recession, Arnold warns of the “growing anxiety about Mexican immigrants ‘stealing jobs’ and usurping welfare, education and healthcare benefits is being tied to sovereign concerns” (Arnold, 2007); in recent memory, this “stealing jobs” rhetoric has been beaten to death by U.S. political leaders, likely in the same way that it was employed eighty years prior. Hopefully, policymakers of the future will learn from the immigration laws of the 1930’s and use methods other than attempting to effect economic or racial changes by immigration restrictions.

Works Cited

Aguila, Jaime R. "Mexican/US immigration policy prior to the Great Depression." *Diplomatic History* 31.2 (2007): 207-225.

Arnold, Kathleen. "Enemy invaders! Mexican immigrants and US wars against them." *Borderlands* 6.3 (2007).

“Figure: Figure 1. People Granted Lawful Permanent Resident Status, by Year..” Just Facts, 2015, https://www.justfacts.com/immigration.asp.

Flores, Melissa G. "Images from the Past: Stereotyping Filipino Immigrants in California." *Historical Perspectives: Santa Clara University Undergraduate Journal of History*, *Series II* 9.1 (2004): 8.

Hoffman, Abraham. "Mexican Repatriation Statistics: Some Suggested Alternatives to Carey McWilliams." *The Western Historical Quarterly* 3.4 (1972): 391-404.

Hoffman, Abraham. *Unwanted Mexican Americans in the Great Depression: Repatriation Pressures*, 1929-1939. Vnr Ag, 1974.

Inwood, Kris, Chris Minns, and Fraser Summerfield. "Reverse assimilation? Immigrants in the Canadian labour market during the Great Depression." *European Review of Economic History* 20.3 (2016): 299-321.

Lee, Jongkwan, Giovanni Peri, and Vasil Yasenov. *The Employment Effects of Mexican Repatriations: Evidence from the 1930's*. No. w23885. National Bureau of Economic Research, 2017.

Murrin, John M., et al. *Liberty, Equality, Power: A History of the American People*, Volume 2: Since 1863. Cengage Learning, 2015.

Ngai, Mae M. "The architecture of race in American immigration law: A reexamination of the Immigration Act of 1924." *The Journal of American History* 86.1 (1999): 67-92.

O'Brien, Matthew J. "Transatlantic connections and the sharp edge of the great depression." *Eire-Ireland* 37.1 (2002): 38-57.

Ray, Eric L. "Mexican Repatriation and the Possibility for a Federal Cause of Action: A Comparative Analysis on Reparations." *The University of Miami Inter-American Law Review* 37.1 (2005): 171-196.

Sobredo, James. "The 1934 Tydings-McDuffie Act and Filipino Exclusion: Social, Political and Economic Context Revisited." *Studies in Pacific History.* Routledge, 2018. 155-169.

Stephenson, George Malcolm. *A History of American Immigration, 1820-1924*. Ginn, 1926.

Yearbook of Immigration Statistics. Department of Homeland Security, 6 Jan. 2018, www.dhs.gov/immigration-statistics/yearbook.

Jonathan Lam

Prof. Abdelwahed

HSS4 — Keynesian Economics

4 / 17 / 20

IS-LM Model Summary

- A way to visualize equilibrium in Keynes' General Theory, created by John Hicks.
- Stands for "investment-savings" and "liquidity preference-money supply": these are the two curves on the model.
- The graph plots GDP on the horizontal axis and interest rates on the vertical axis.
- IS curve: the relationship between GDP and interest rates when investment equals savings (I=S)
    - As interest rates go down, then savings (and thus investments if I=S) go up, so overal output and GDP go up. (The independent variable is interest rates, and the dependent variable is income level and investment.) Thus there is a negative slope in the IS curve.
- LM curve: the relationship between GDP and interest rates when liquidity preference is equal to money supply (L=M)
    - Liquidity preference is how much cash people want to have at hand, money supply is how much is available in banks.
    - As GDP goes up, people's income tends to go up, and generally they want more free money (higher liquidity) to be able to invest more or save for use in emergencies. To have the money supply match this liquidity, higher interest rates are needed. (The independent variable is income level, and the dependent variable is interest rates.) Thus there is a positive slope in the LM curve.
- Intersection of IS and LM curves are an equilibrium point between the real market (real production) and money market. (i.e., won't decrease interest rates further because of high money market demand, won't increase interest rates further because of negative impact to investment)
- It was later regarded to be not very realistic because it doesn't incorporate many factors, such as taxes, inflation, international markets, among other factors. John Hicks admitted that it was more of a "classroom gadget" representation.
- Also considered unrealistic because it treats the two curves as independent, whereas changes in one should cause changes in the other.

Migration during the Great Depression

Jonathan Lam
Prof. Abdelwahed
HSS4
2/21/20

Some Numbers

- **Immigration** in the U.S. dropped from **279,678 in 1929 to 23,068 in 1933** (U.S. Census Bureau, 2006).

Some Numbers

- **Immigration** in the U.S. dropped from **279,678 in 1929 to 23,068 in 1933** (U.S. Census Bureau, 2006).

- An estimated **500,000 to 2,000,000 Mexicans** in the U.S. (many U.S. citizens) were **deported** to Mexico during the Great Depression (Ray, 2005).

Some Numbers

- **Immigration** in the U.S. dropped from **279,678 in 1929 to 23,068 in 1933** (U.S. Census Bureau, 2006).
- An estimated **500,000 to 2,000,000 Mexicans** in the U.S. (many U.S. citizens) were **deported** to Mexico during the Great Depression (Ray, 2005).
- An estimated **500,000** Americans migrated from the dry Midwest to California in the "**Okie Migration**" (Gutmann et al., 2016).

What I am studying

I am studying **migration patterns during the Great Depression**

because I want to find out **what movement was caused or impaired by the conditions in the 1930s**

in order to help my reader understand **how crisis drives mass mobility, and vice versa.**

format courtesy of *The Craft of Research, 4th Edition* by Wayne C. Booth

Outline

Tentative claim: During the Great Depression, attempts both locally and federally to change normal migration patterns were sometimes only effective at self-preservation in the short-term (e.g., in the case of the Dust Bowl) and otherwise not at all effective in curbing any of the crises' effects.

Outline

Tentative claim: During the Great Depression, attempts both locally and federally to change normal migration patterns were sometimes only effective at self-preservation in the short-term (e.g., in the case of the Dust Bowl) and otherwise not at all effective in curbing any of the crises' effects.

- **Subclaim**: Environmental factors
    - Evidence: Dust bowl (increased migration)

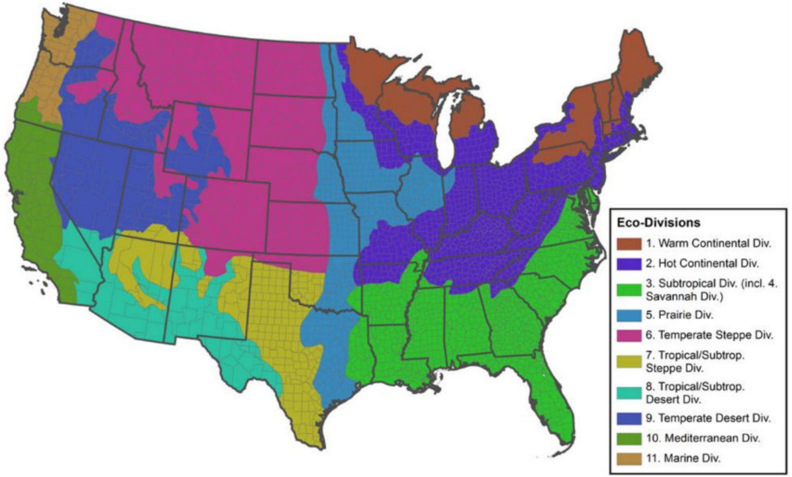
- **Subclaim**: Reduction of immigration (forced and voluntary)
    - Evidence: Mexican repatriation (forced migration)
    - Evidence: Irish immigration (limiting mobility)

Annotated Bibliography [1]

Article: “Migration in the 1930s: Beyond the Dust Bowl”

This summarizes the results of statistical research that aims to find statistically-significant correlation between domestic migration events in the U.S. using 1940 census data (including 1935 whereabouts) and agricultural data from throughout the 1930s. The paper looks at attributes of (especially agricultural) and net population flux per county, and ecodivisions. The results confirm other research that the majority of the migration is tightly coupled to ecological factors, but also acknowledge that some regional patterns cannot be explained by their model, and a study on a per-individual basis may be more informative.

Gutmann et al., 2016



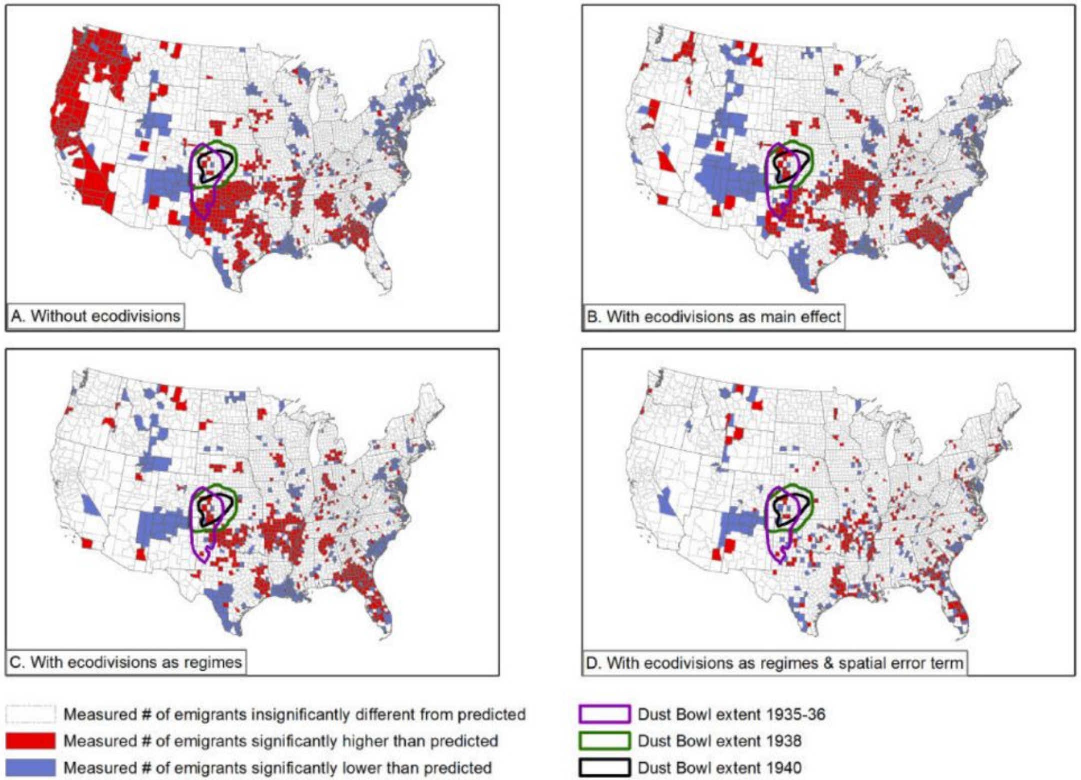


Figure 10.
Mapped residuals of four different implementations of the crop failure model.

Gutmann et al., 2016

Annotated Bibliography [2]

Article: “Transatlantic Connections and the Sharp Edge of the Great Depression”

This article focuses on how Irish emigration to the U.S., which had played a significant economic role in the former since the Irish famine, was severely disrupted by the Great Depression. Immigration was not impacted much by American immigration laws during the early 20th century, but many Irish breadwinners in America lost the earnings of the last decade in the stock market. Some others were too embarrassed to not be able to support their families and cut off contact; others who returned to Ireland were shunned from society; thus migration to the United States caused major economic and social strife in Ireland, even if the stock market and ecological effects didn’t affect Ireland directly.

O’ Brien, 2002

Annotated Bibliography [3]

Article: “Mexican Repatriation and the Possibility for a Federal Cause of Action: A Comparative Analysis on Reparations”

This report discusses the Mexican Repatriation Act, a policy instituted by President Hoover to free up jobs for Americans from 1929 to 1944 to deport an estimated 500,000 to 2,000,000 Mexicans (mostly U.S. citizens) to Mexico. Many Mexican workers had come to the U.S. in the early 20th century, but they were always “second-class citizens”: low-paid and an easy scapegoat during the Great Depression. More specifically, it aims to address the question of its justice and if a federal bill is necessary to aid in reparation for damages caused by it (after California Governor Schwarzenegger vetoed such a bill).

Ray, 2005

References

Gutmann, Myron P., et al. “Migration in the 1930s: Beyond the Dust Bowl.” *Social Science History*, vol. 40, no. 4, 2016, pp. 707–740., doi:10.1017/ssh.2016.28.

O’Brien, Matthew J. "Transatlantic Connections and the Sharp Edge of the Great Depression." *Éire-Ireland*, vol. 37 no. 1, 2002, p. 38-57. *Project MUSE*, doi:10.1353/eir.2002.0003.

Ray, Eric L. “Mexican Repatriation and the Possibility for a Federal Cause of Action: A Comparative Analysis on Reparations.” *The University of Miami Inter-American Law Review*, vol. 37, no. 1, 2005, pp. 171–196. *JSTOR*, www.jstor.org/stable/40176606. Accessed 21 Feb. 2020.

Jonathan Lam
Prof. Abdelwahed
HSS4
5 / 15 / 20

Misled Perceptions on Immigration during the Great Recession

Introduction
The Great Recession of 2008-2009 was the largest economic downturn in the United States since the Great Depression of the 1930's, and until the current Great Lockdown in 2020. Large uncertainty in the economic sector sometimes sparked panic and blame towards a common scapegoat: foreigners and immigrants. On April 20, 2020, a few months into the Great Lockdown, President Donald Trump tweeted the following message expressing this sentiment:

> "In light of the attack from the Invisible Enemy, as well as the need to protect the jobs of our GREAT American Citizens, I will be signing an Executive Order to temporarily suspend immigration into the United States!" @realDonaldTrump.

This tweet is hardly surprising or unprecedented in the modern context — a pamphlet prepared by UNESCO claimed that "all previous crises of the 1900s, including the Great Depression … affected migration in different ways and spurred resentment of foreigners and xenophobic actions" (Global Migration Group and UNESCO, 2009). What makes the Great Recession (when compared to immigration policies and attitudes from other crises, such as the Great Depression) especially interesting to study is its recentness and the amount of data gathered on perceptions of native U.S. citizens towards immigrants.

During the Great Recession, while immigrants were often blamed for exacerbating the economic crisis by native citizens, this was usually caused by misguided psychological perspectives. Immigrants in fact tended to aid the economy before and after the Recession by increasing diversification. As UNESCO noted, these are important themes not specific to the Great Recession, but strongly observed in the Great Recession of the 1930's and the current Great Lockdown of 2020. More widespread knowledge of the mental biases motivating this discrimination is an important step in preventing similar prejudice from affecting policy during financial crises in the future.

Summary of immigrant-related policies passed slightly prior to and during the Recession
It is important to provide an overview of the situation and policies relating to immigrants passed in the era of the Great Recession and Great Depression, and how other literature attempts to rationalize these policies. Having a firm comprehension of the laws enacted during these crises will help portray the sort of blatant prejudices present in society, something that can only be viewed with such a summarial hindsight summary. For example, in the early twenty-first century, there was the rise of stricter immigration laws and higher rates of deportation, especially among the Latino population, and this racial bias is often attributed to economic factors. The Latino and Asian populations grew by 43% between 2000 and 2010, constituting 23.2% of the U.S. population in 2010 (Humes et al., 2011). In response to the high influx, state legislatures passed 156 immigration-related laws in 2012 (National Conference of State Legislatures, 2013). These new policies have ranged from protectionist policies (encouraging immigrants to leave and protecting native-born workers) to more forceful policies, such as a greater emphasis on the removal of undocumented immigrants (i.e., leading to higher deportation rates) (Ybarra et al., 2016).

One particular policy that was especially harsh towards immigrants was deportation, and this was disproportionately biased towards specific racial groups. Between 1993 and 2011, the Department of Homeland Security (DHS) reported a ten-fold increase in Mexican deportees and a twelve-fold increase in Central American deportees (in contrast with relatively smaller changes in deportees from Europe and Asia) (Golash-Boza, 2013); by 2011, 97% of the reported deportees came from Latin America. He finds that the high joblessness due to the Great Recession, a changing labor market relying more on automation and less on construction and manufacturing jobs (of which industries Latino-Americans were a large constituent), and the historically-targeted Latino demographic made them easy targets for deportation. Additionally, Golash-Boza finds that the majority of the deportees had “strong ties” to the U.S. (most living in the U.S. for 5-10 years prior to deportation) and that a large proportion were legal immigrants — for example, only three-quarters of Dominican deportees are undocumented (Siulc, 2009). This kind of “messy” deportation, similar to that during the Mexican Repatriation of the 1930’s, hints strongly at racial (as opposed to economic) motivations. The chart shown in Fig. 1 demonstrates a similar effect; when the hysteria was highest in 2007-2009, the number of non-criminal deportations, and the ratio of non-criminal to criminal deportations, reached their highest values (McCarthy, 2014).

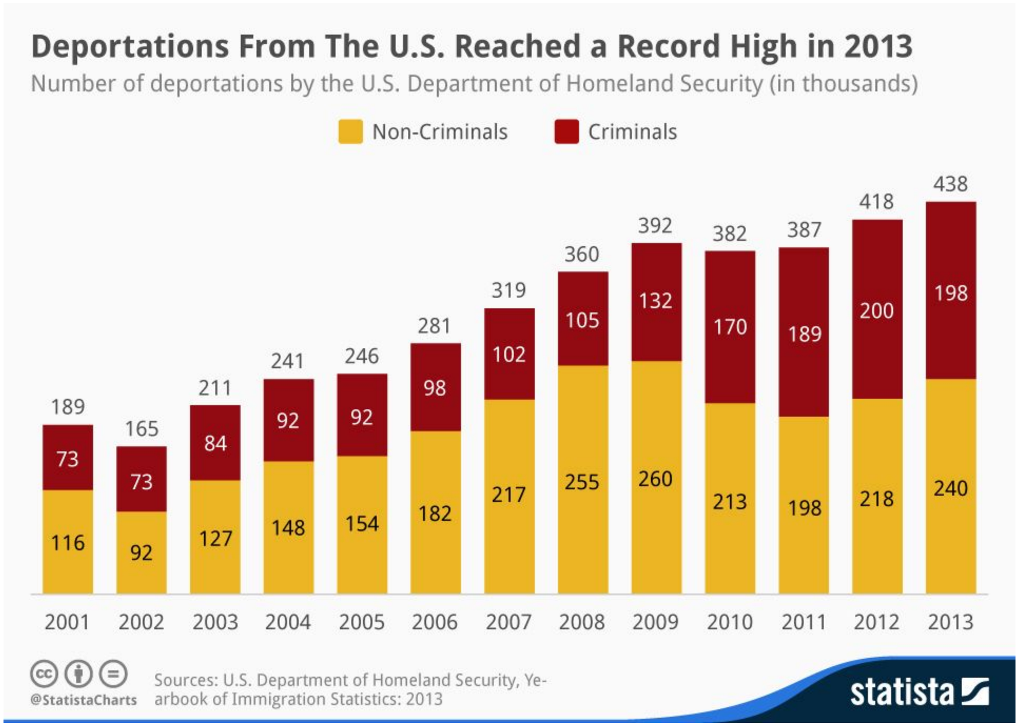


Fig. 1. Deportation count in the United States, 2001-2013. (McCarthy, 2014)

There were also many strict domestic policies passed during the Great Recession regarding immigrant rights. One example of a policy passed in the not-yet-recovered American economy in 2011, Alabama’s House Bill 56, is a good example of the anti-immigrant bias that built up during the Recession. This bill requires police officers in Alabama to check immigration status of all stopped and detained people, requires public schools to check immigration status of all enrolled students, prohibits undocumented immigrants from attending higher education, and prohibits public state benefits for undocumented immigrants (Ybarra et al., 2016). This act was primarily in response to a high Hispanic population influx in the years leading up to the Recession, and it caused fear of punishment among the Hispanic population, who were indeed the most adversely affected by Alabama HB 56 (Conley, 2015).

The U.S. observed similar kinds of policies during the Great Depression eight decades prior. Slightly before the Depression, a series of increasingly strict immigration laws were passed up to and including the strict Immigration Act of 1924, due mostly to eugenics (Hoffman, 1974). Economic rationale (the protection of native jobs) justified the Mexican Repatriation, a massive deportation of recent Mexican immigrants, but the selective deportation was carried out largely based on appearance rather than actual immigrant status. The supposed economic motive was not proven to be effective, and a study by Lee actually suggests that deportation of Mexican workers exacerbated the economic crisis (Lee et al., 2017). Like the Great Recession, there were massive deportations of Mexican Americans due to an economic argument, and while not explicitly stated, there was definitely a racialized aspect of the stricter immigrant policies (with harsher policies towards most non-European immigrants). The parallels are strong enough with the Great Depression that some of the analyses used here can likely be applied postmortem to the Great Recession in order to rationalize some of the anti-immigrant events that occurred then.

Psychological appeals of anti-immigrant bias

While many Americans felt threatened by the presence of immigrant groups, it turns out this feeling was more associated with media pressure against immigrants and a sense of personal financial stability, rather than with being a reflection of one's real financial situation. The literature often points to a psychological basis justifying most of the discrimination laws. In particular, Ybarra claims that there is a strong relationship between the American public's perception of immigrants and the enactment of stricter laws against immigrants. He cites a survey by Kohut, which finds that most Americans in 2011 believe that immigrants are a "burden" to the country (Kohut et al., 2011), and papers from Humes and Perez reporting a relationship between negative attitudes towards immigrants and restrictive immigration policy (Humes et al., 2011 and Perez, 2010). The problem is that many of the claims of immigrants as a financial "burden" are unfounded, and the fact that immigrants in the U.S. have never led the U.S. into a depression nor is there strong evidence that the immigrant workforce had a negative effect on native employment make the economic arguments for anti-immigration seem to be a political sham. The question arises of why exactly Americans believed in and passed the policies.

The first point to address is how any of these anti-immigrant policies, which are almost blatantly racist, can justify their discriminatory nature. There must be some critical defense in the anti-immigration policymakers' against the claim that the selective immigration summarized earlier is not racist, and understanding this argument will be important for countering it. While not explicitly targeting certain immigrant groups, there is no denial that these policies affect some immigrant groups more than others: almost all deportees come from Latin America, and several domestic laws were effectively targeted towards Latin Americans, such as Alabama HB 56. This is similar to the case in the Great Depression, in which there were more explicitly labeled "preferred" (primarily European) and "non-preferred" (primarily Latino and Asian) immigrants. The key to mass approval of prejudiced policies lies in the wording— even without sufficient factual evidence supporting the economic claims, a "threat perception" is a powerful motivator. Even when modern democratic societies culturally tend to favor egalitarian rules, any threat to the society takes on a higher priority than maintaining non-discriminatory policies (Ramos et al., 2016). Ramos describes how it is not contradictory for an educated, ideological citizen of a modern democracy to both accept discriminatory measures such as HB 56 and believe that they themselves are unprejudiced. By reframing the issue by worrying about the effect that immigrants might have on culture and the economy, politicians can portray immigrants as a threat, which is a legitimate cause in and of itself.

Even if the reasons for being a threat (i.e., the cultural and economic harm it might inflict on American society) are not entirely sound, the potential of damage caused by a threat is enough to rally supporters around. The same logic worked during the Great Depression, in which legitimate data was scarcer about the effect of immigrants on the economy: viewing Mexicans as a threat to the already-scarce job market allowed the federal and state governments to perform deportations en masse (even messily deporting many U.S. citizens) without any opposition.

What is the primary argument for such a strong rejection of immigrants in the U.S. during economic recessions? In other words, why do Americans believe that immigrants will hurt the economy? From a certain objective point of view, it may seem obvious that free immigration might help the economy. Goldstein concisely summarizes a such an argument as a counterargument to those who view immigration as financially damaging:

> "Immigration policy remains one of the most divisive issues on the American political landscape. While policymakers have legislated a policy that reflects the benefits of open borders to goods and services, they have forestalled a similar policy toward the movement of peoples. This divergence, that is, open borders for goods but closed borders for people, is a puzzle from a purely materialist perspective. Economic logic suggests that trade and immigration policy are tightly connected and have similar material effects; thus, opinions on one flow should be similar to opinions on the other" (Goldstein et al., 2014).

While free trade is outside the scope of this paper and introduces many extra complexities, it does poke a hole of doubt in the anti-immigrant rhetoric. Free trade is loudly promoted by large capitalist nations (which tend to be the same as those democratic nations who perceive immigration as a threat), but the material benefit of the people exchange is not as often mentioned in economic discussions.

To understand why Americans don't support immigration, Goldstein analyzes the results of a survey with questions about perceptions on immigrants and about personal financial status. The survey was taken six times throughout the Recession to gauge Americans' beliefs over time. The most interesting result of the survey is that high-skilled immigrants tended to be favored over low-skilled immigrants in general, but during the crisis were less-favored by native high-skilled citizens. Low-skilled natives didn't change their opinion on high-skilled immigrants during the recession, because they didn't feel that their job security was threatened by this certain demographic. Incorporating data from the survey participants' financial situation (housing and employment situations, etc.) and self-reported (perceived) anxiety levels, the self-perceived anxiety during the Great Recession was a much stronger indicator of opinion on immigration than real financial situation. Likewise, real financial stimulus spending and real immigration numbers had no effect on opinion on immigration. Goldstein's report shows the important result that the real effect (i.e., the supposed "threat") of immigrants really is not correlated with the anti-immigrant effect, weakening the justification for the "threat" of anti-immigrant sentiment.

A separate paper illustrates that the perceived immigrant threat is highest where economic instability is high, such as "where there is economic stagnation, where labor unions are growing weaker, where the minimum wage is low, and where corporate restructuring has taken place" (Wallace and Figueroa, 2012); note that this list doesn't include immigrant presence or economic effect. This study similarly illustrates that it is more of a personal sense of financial instability that is more afraid of an immigrant "threat" than any real effect caused by immigrant populations. Together, Goldstein and Wallace's reports illustrate that even in the absence of a real economic threat by immigrants, the anxiety about one's own financial situation can meddle with opinions about immigrants. In other words, there's a high chance that the immigrant threat to the job market can be imaginary but hypersensitized by the stress of a recession.

The real effects of immigrants on economics before the recession and during recovery
Contrary to the belief that immigrants hurt the economy during a Recession (sensitivized by financial insecurities as discussed above), immigrants tend to benefit the U.S. economy, before, during, and after recessions.

Zaretsky provides a detailed view of the impact of immigrants on the U.S. economy during ordinary times (i.e., not during a recession). He acknowledges the common argument that introducing more labor into the market means lower wages for all, but states that this simplistically assumes that “all workers are the same. But all workers are not the same” (Zaretsky, 1997). His research finds that there is a net increase in wages across all native worker groups (except for unskilled domestic workers) caused by increased immigration. Additionally, Zaretsky’s research finds that immigrants are also unable to receive government benefits such as Social Security, Medicare, and Medicaid, and typically immigrate into wealthier states with higher tax rates, which means that they are not a fiscal burden on government social funding. All of these statements are evidence that immigrants certainly do not hurt the economy, and benefit almost all worker groups’ wages. However, Zaretsky didn’t examine the effect of immigrants on the economy during Recessions, which is when the anti-immigrant sentiment is strongest.

During Recessions, a higher population of immigrants improves overall economic diversification, allowing for shallower recessions and quicker recovery than a more homogenous economy. Lester and Mai find that a greater integration of immigrants in a local economy is correlated with a greater resistance to economic shocks, where resilience is coarsely measured in changes in real wages and unemployment (Lester and Mai, 2016). After analyzing the factors that influence immigrant integration into the U.S. economy in the years leading up to the Recession, they find that U.S. immigration seems to favor high- and low-skilled migrant workers, and state that their results advise that a more widely-accepting immigration policy is economically favorable.

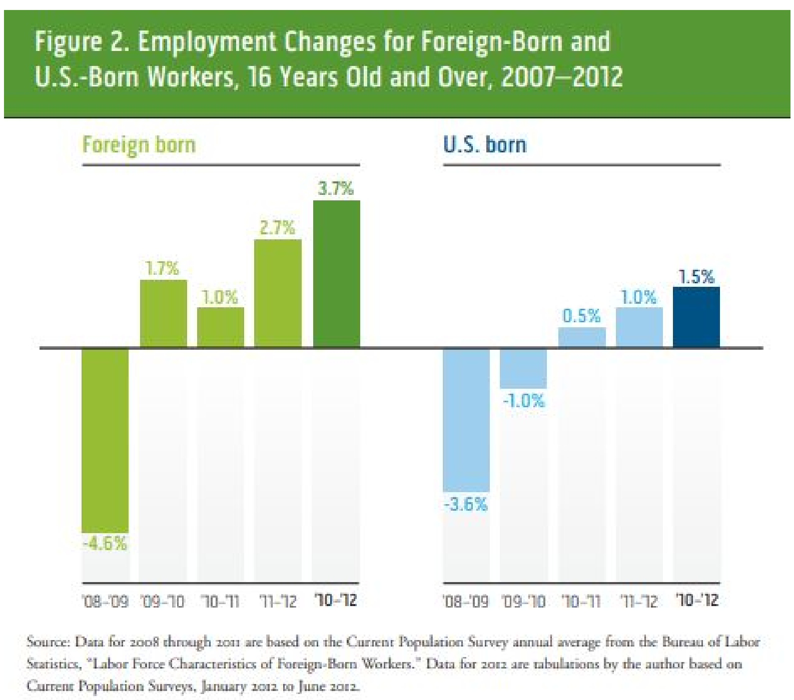


Fig. 2. Immigrant economic crisis and recovery, 2008-2012 (Enchautegui, 2012)

A report analyzing data from the Bureau of Labor Statistics in the years during and immediately after the Great Recession in the U.S. finds that immigrants were hit harder in terms of relative unemployment than non-immigrant groups, but their growth in employment was also greater in the years following the Recession (Enchautegui, 2012). In particular, Enchautegui finds that out of the working class without a bachelor's degree, immigrants had greater rates of employment growth in the years 2009 to 2012 and were able to surpass their pre-recession employment rates, while the native workforce were unable to (see Fig. 2). While Enchautegui doesn't measure the same statistics as Lester and Mai, their general result is the same: higher long-term resilience of a local economy is correlated with higher immigrant integration. Roughly speaking, this indicates that a higher percentage of immigrants in the workforce would indicate an overall speedier recovery of the economy.

Conclusion and present implications

Distanced from the contemporary panic that came with the Great Depression and Great Recession, it is clear that some of the arguments used against immigrants were actually unsound. There was the perception that immigrants could threaten job security for natives, but higher integration of immigrants actually aided the recovery of the economy. The idea of the "threat" used to allow discriminatory measures is perpetuated by the fear induced by an economic recession and only hurts immigrants. Changes to the education system, in which the benefits of free immigration are discussed alongside free trade, could be a potential way to help combat the strong innate nativist sentiment many Americans harbor. As Americans now enter into a recession even worse than the Great Recession a decade ago, it's important that we understand what are facts and what is subjective, so that recovery is as speedy as possible and unnecessary. Public messages from authorities such as the recent tweet from President Trump are clearly baseless but blindly followed and agreed with in times of financial sensitivity, and as a result can easily hurt the economy and immigrant livelihoods. There are no numbers or statistics behind President Trump's words, just the depiction of an attack on Americans (by the virus) and a threat to American jobs.

Zaretsky suggested that "those who compete with—are substitutes for—immigrants will receive a lower wage than they would without immigration, while those who complement immigrants will receive a higher wage" (Zretsky, 1997). This wording emphasizes that immigrant labor is a normal and established part of the economy, and that even natives can "substitute... for" or "complement immigrants," just as they can supplant or complement native jobs. This view is very progressive and should be professed more commonly in order to (subconsciously) acknowledge the importance and deep integration of the immigrant economy with the native one. If we, as a nation, can learn to express ourselves in a similar way, putting immigrants at the same level as native citizens and objectively observing facts, then we have a chance at providing more fair treatment to groups during economic recessions in the U.S.. Discriminatory anti-immigrant laws should no longer be justifiable in any modern context without proper evidence, even during stressful times.

Works Cited

Conley, Meghan. "In times of uncertainty: The Great Recession, immigration enforcement, and Latino immigrants in Alabama." *Immigrant Vulnerability and Resilience*. Springer, Cham, 2015. 147-162.

Dancygier, Rafaela, and Michael Donnelly. "Attitudes Toward Immigration in Good Times and Bad." *Mass Politics in Tough Times: Opinions, Votes and Protest in the Great Recession*. : Oxford University Press, April 01, 2014. Oxford Scholarship Online.

Enchautegui, Maria E. "Hit hard but bouncing back: The employment of immigrants during the Great Recession and recovery." *Washington, DC: Urban Institute* (2012).

Golash-Boza, T., Hondagneu-Sotelo, P. Latino immigrant men and the deportation crisis: A gendered racial removal program. *Lat Stud* 11, 271–292 (2013). https://doi.org/10.1057/lst.2013.14.

Goldstein, Judith L., and Margaret E. Peters. "Nativism or economic threat: Attitudes toward immigrants during the Great Recession." *International Interactions* 40.3 (2014): 376-401.

Global Migration Group and UNESCO. *Fact-sheet on the impact of the economic crisis on discrimination and xenophobia*, 2009. https://www.gfmd.org/files/documents/gfmd_athens09_contr_unesco_factsheet_discrimination_and_xenophobia_en.pdf.

Hoffman, Abraham. Unwanted Mexican Americans in the Great Depression: Repatriation Pressures, 1929-1939. *Vnr Ag*, 1974.

Humes, Karen, Nicholas A. Jones, and Roberto R. Ramirez. "Overview of Race and Hispanic Origin, 2010." *US Department of Commerce, Economics and Statistics Administration*, US Census Bureau, 2011.

Kohut, Andrew, Caroli Doherty, Michael Dimock, and Scott Keeter. 2011. "Beyond Red vs. Blue: Political Typology." *The Pew Research Center for the People & the Press*.

Lester, William T. and Mai Thi Nguyen. The Economic Integration of Immigrants and Regional Resilience, *Journal of Urban Affairs*, 38:1, 42-60, DOI: 10.1111/juaf.12205, 2016.

McCarthy, Niall. "Deportations from the United States Reached a Record High in 2013." *Statista*, 8 Oct. 2014, https://www.statista.com/chart/2802/deportations-from-the-united-states-reached-a-record-high-in-2013/.

National Conference of State Legislatures. 2013. "2012 Immigration-Related Laws and Resolutions in the States (Jan. 1-Dec. 31, 2012)." http://www.ncsl.org/research/immigration/2012-immigration-related-laws-jan-december-2012.aspx.

Lee, Jongkwan, Giovanni Peri, and Vasil Yasenov. The Employment Effects of Mexican Repatriations: Evidence from the 1930's. *National Bureau of Economic Research*, 2017.

Lester, T. William, and Mai Thi Nguyen. "The economic integration of immigrants and regional resilience." *Journal of Urban Affairs* 38.1 (2016): 42-60.

Perez, Efren O. 2010. "Explicit Evidence on the Import of Implicit Attitudes: The IAT and Immigration Policy Judgments." *Political Behavior* 32 (4): 517-45.

Ramos, Alice, Cícero Roberto Pereira, and Jorge Vala. "Economic crisis, human values and attitudes towards immigrants." *Values, economic crisis and democracy*. Routledge, 2016. 130-163.

@realDonaldTrump. *Twitter*, 20 Apr. 2020, 10:06 p.m., https://twitter.com/realDonaldTrump/status/1252418369170501639.

Siulc, Nina. Unwelcome citizens, criminalized migrants, and the quest for freedom: Deportees in the Dominican Republic. New York University, 2009.

Wallace, Michael, and Rodrigo Figueroa. "Determinants of Perceived Immigrant Job Threat in the American States." *Sociological Perspectives*, vol. 55, no. 4, Dec. 2012, pp. 583–612, doi:10.1525/sop.2012.55.4.583.

Ybarra, Vickie D., Lisa M. Sanchez, and Gabriel R. Sanchez. "Anti-immigrant anxieties in state policy: The great recession and punitive immigration policy in the American states, 2005–2012." *State Politics & Policy Quarterly* 16.3 (2016): 313-339.

Zaretsky, Adam M. "A burden to America? Immigration and the economy." *The Regional Economist,* Oct (1997): 5-9.

HSS4 – Transcript for Final Paper Presentation

Jonathan Lam

05/01/2020

Note: This was a quick overview of points to mention and not a polished draft.

This is my presentation for the final paper, “Misleading perceptions on immigration during the Great Recession” I decided not to change my topic from the middle of the semester.

Let’s begin with a quote: In light of the attack from the Invisible Enemy, as well as the need to protect the jobs of our Great American citizens, I will be signing an executive order to temporarily suspend immigration into the United States

Guess who and when this quote was from? This was from ten days ago, by President Donald Trump, before passing an act that essentially brings immigration to a halt for the next two months. While much of my research and paper doesn’t focus too much on immigration legislature per se, this kind of attitude precisely fits what I was trying to discover.

So what was I hoping to find out? I believe I shared this earlier, but to reiterate: I am studying the economic effect of immigrants and attitudes towards them before, during and after the Great Recession because I want to find out how perception (notably, psychology) differed from actual economic effects in order to help my reader understand the actual vs. perceived economic effect of immigrants with respect to the Great Recession.

I want to start by bringing it back to the Great Depression. Back then, there was a large wave of nativism. There wa a huge influx of immigration in the mid- 1800s on the West Coast, and one of the first major immigration bans was the Chinese Exclusion Act in 1880. There was also a large immigration of Mexican migrant workers in the early 1900s. There were many immigration laws passed in this time, the most strict being the Immigration Act of 1924, which passed quotas based on country of origin, and thus was clearly racially biased towards European immigrants. The Mexican Repatriation was the joint efforts of the U.S. and Mexican governments to move hundreds of thousands of Mexican migrant workers back to Mexico. Mexico wanted to do it to reintroduce skilled workers into their economy, and the U.S. purportedly had a very familiar argument: to save American jobs for natives. Then, there was some violence against Filipinos in California for ”taking” their farmer jobs, such as in the Watsonville riots.

Turning to the Great Recession, it was also an unrestful time. Like the era before the Great Recession, the 1990s and early 2000s were a time of high immigration flux. Of course, things were more on edge after 9/11, and this combined with higher immigration caused a spike in deportations and funding to enforce these laws. During the Great Recession, there was a higher distrust of immigrant populations, especially Latinos, leading to stronger laws that were essentially racially discriminatory against Latinos, the most prominent being Alabama's House Bill 56, which enforced more frequent documentation checks and stronger restrictions against immigrants from receiving government aid or public education.

You can see in these charts here a little bit more. From the graph on the right, maybe it's not so clear because there is an underlying rising trend, but the number of noncriminal deportations peaked in 2008 and 2009, as the anti-immigrant histeria peaked. Similarly, there was a very strong bias towards deporting Latino immigrants. You can't see it from the graph, but it was disproportionally targeting Latino immigrants over other immigrant populations based on the size of the immigrant groups, and often legal immigrants were deported, just like during the Mexican Repatriation.

To summarize the arguments behind these anti-immigrant laws briefly, during the Great Depression it could be rationalized from a eugenics standpoint (i.e., to preserve the homogeneity of the American demographic and culture), but clearly this faded out after WWII. But in general the cultural argument still persists. Secondly, and a common one for us to hear, is the threat immigrants pose to American jobs, since they are known to work for lower wages. Similarly, that immigrant families tend to be low-income and depend more on government handouts, thus costing the government more.

It's pretty clear that there's some racial discrimination. We can classify this under the term "threat rhetoric," which has historically been very effective. Both for the Mexican Repatriation and the acts during the Great Depression, this has caused little backlash when passing strict anti-immigrant laws. The basic idea is to associate immigrants with some threat, such as taking jobs. Then it becomes the task of protecting Americans rather than discriminating against "other people."

Similarly, some surveys taken throughout the Recession era have shown that more misconceptions come about with selfishness on a personal level. Unsurprisingly, people favor higher-skilled immigrants. However, high-skilled natives tended to favor high-skilled immigrants less and less during the Recession because they feared that they would lose their jobs. However, these surveys, which also took data on actual financial standings of the respondents, found that there was not a high correlation between negative attitudes of natives towards immigrants and actual threats (such as rate of immigration, and local and personal economic situation) but more on perceived threat to one's own financial situation. In other words, people tended to hate immigrants more if they felt threatened financially, even if it had nothing to do with immigrants. This is clearly a logic flaw.

This brings us full circle. What effect do immigrants actually have on the economy? A study found that after roughly two generations, there isn't much difference between immigrant's and native's net effect on the economy. They still have to pay taxes, still receive Social Security the same way as regular citizens, and undocumented citizens neither contribute via taxes nor receive social security. All in all, it basically balances out, but the same study suggests that there is a small net positive benefit from immigrants in general due to some subtle factors.

Specifically, during the Great Recession, we have a few things going on: first, immigrants tended to increase diversification of the economy, which is a good thing. Secondly, they tended to get hit worst in the initial part of the Recession, with "less essential jobs" in a very similar experience to what we're seeing nowadays. Thus we can hardly say they "stole" American jobs, but perhaps provided some of a buffer to natives during the hardest part of the Recession. And lastly, they tended to help most with recovery, quickly gaining back reemployment in the years following the Recession, which refutes the idea that they are lazy and stay unemployed and dependent on government aid.

That's the whole story. Again, we have a history of the threat rhetoric, the conflation of personal stability with attitudes on immigrantion policy, and some studies on the actual effect of immigrants on the economy, which tend to show they benefit the economy in the time around a Recession.

I guess I'd finish on another quote, from UNICEF in a report written after the worst part of the Recession. As we are going into another recession, we have to be sure to consider the facts without bias.

Misled **Perceptions** on **Immigration** during the **Great Recession**

Jonathan Lam
Prof. Abdelwahed
HSS4K

“In light of the attack from the Invisible Enemy, as well as the need to protect the jobs of our GREAT American Citizens, I will be signing an Executive Order to temporarily suspend immigration into the United States!”

"In light of the attack from the Invisible Enemy, as well as the need to protect the jobs of our GREAT American Citizens, I will be signing an Executive Order to temporarily suspend immigration into the United States!"

Twitter,
@realDonaldTrump
20 April 2020

Topic

I am studying economic effect of and attitudes towards immigrants

because I want to find out how perception differs from reality

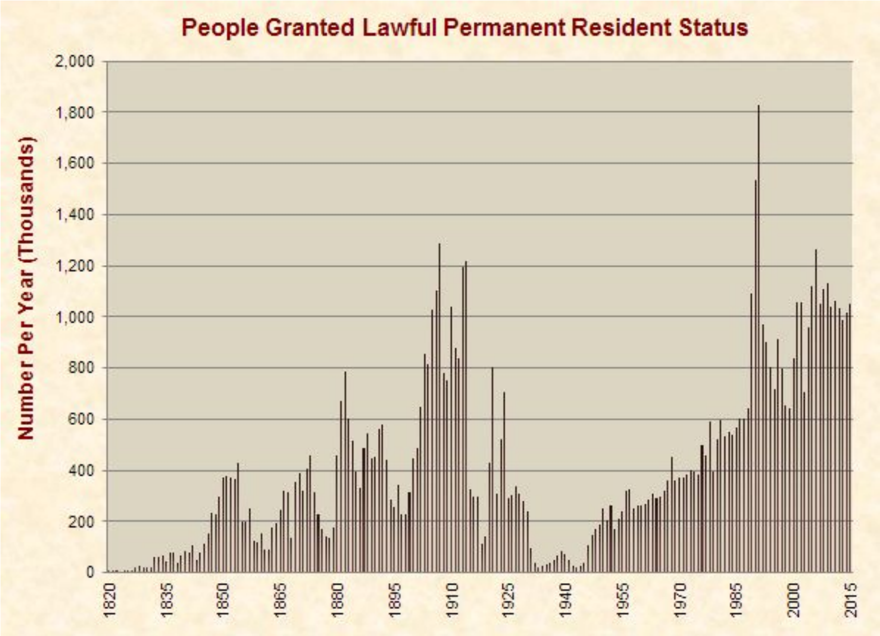
in order to help my reader understand the actual vs. perceived economic effect of immigrants

Significance

Great Depression during a time of nativism

- Large immigrant flux in mid-1800s, 1920s
- Immigration Act of 1924 (1929)
- Mexican Repatriation (1929-1936)
- Watsonville Riots (1930)

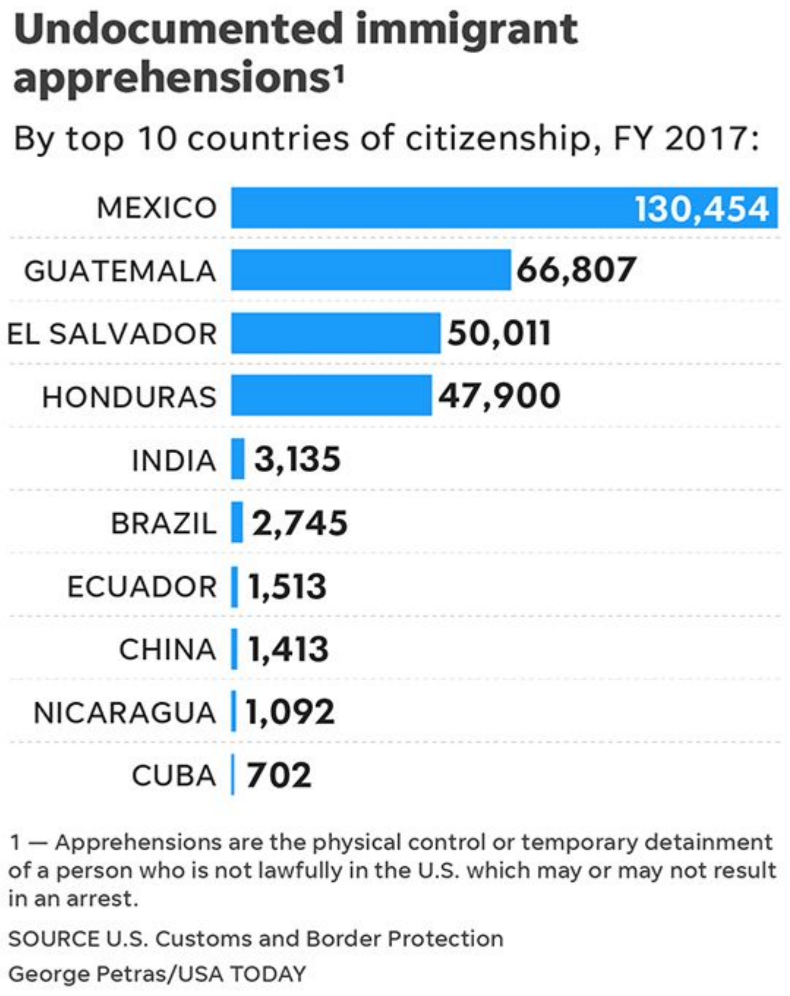
Chart source: https://www.justfacts.com/immigration.asp

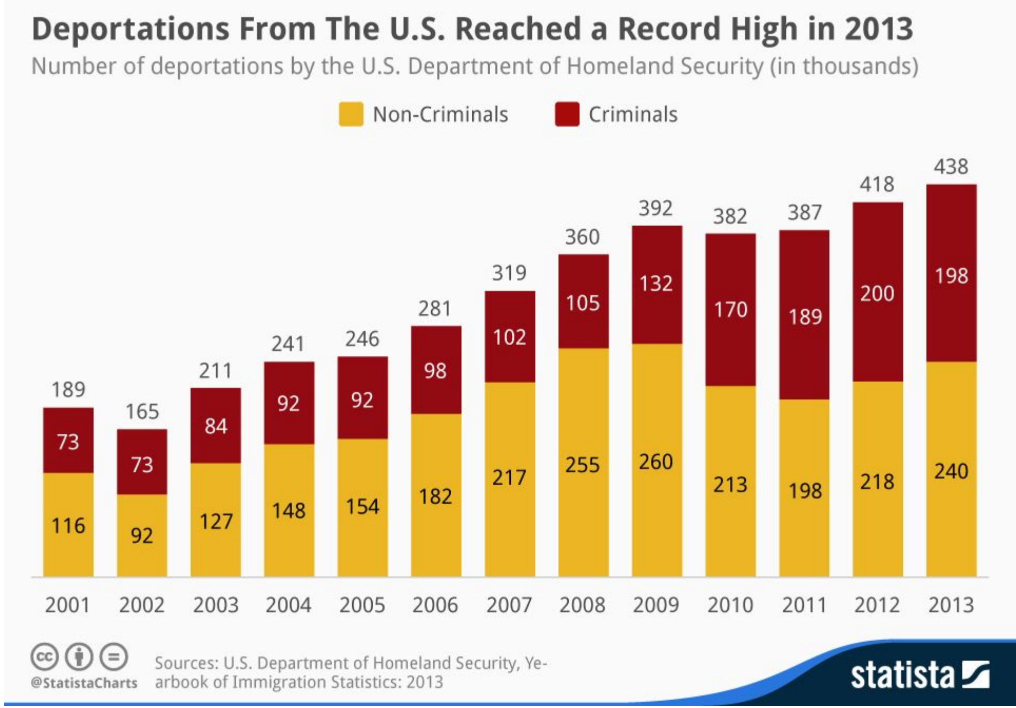


Significance

Great Recession during a time of unrest

- Large rise in immigrant population in 1990-2010
- War on Terror (2001-)
- Mass deportations (1993-2011)
- 156 immigration-related laws in 2012
- Anti-immigrant legislation biased towards Latino populations





Graphics sources:
(left) https://www.usatoday.com/pages/interactives/graphics/deportation-explainer/
(right) http://www.crfimmigrationed.org/featured-news

Arguments against Immigration

- Threat to American culture (social)
    - Nativism and eugenics (mostly Great Depression)

Arguments against Immigration

- Threat to American culture (social)
  - Nativism and eugenics (mostly Great Depression)

- Threat to American jobs (economic)
  - Widespread belief
  - Undocumented immigrants are larger strain on government aid

Threat Rhetoric

"In light of the attack from the <u>Invisible Enemy</u>, as well as the need to <u>protect the jobs</u> of our GREAT American Citizens, I will be signing an Executive Order to temporarily suspend immigration into the United States!"

Conflation of immigrants with threats to U.S. citizens ⇔ not a racial bias

No real evidence but high support

Personal Insecurities

- General fear of immigrants threatening low-skilled natives

Personal Insecurities

- General fear of immigrants threatening low-skilled natives
- Increased fear of similarly-skilled immigrant groups during Recession

Personal Insecurities

- General fear of immigrants threatening low-skilled natives
- Increased fear of similarly-skilled immigrant groups during Recession
- With respect to negative attitudes on immigration:
    - **low correlation** with **actual immigration numbers**
    - **low correlation** with **actual economic instability** in the local region
    - **low correlation** with **actual personal financial position**
    - **high correlation** with **perception of personal financial position**

$\Rightarrow$ Negative attitudes not rooted in fact

Actual effects of immigrants on economy

In general:

- Over time
- Costs of enforcement of immigration laws (ICE budget)
- No large changes on natives’ salary or unemployment levels

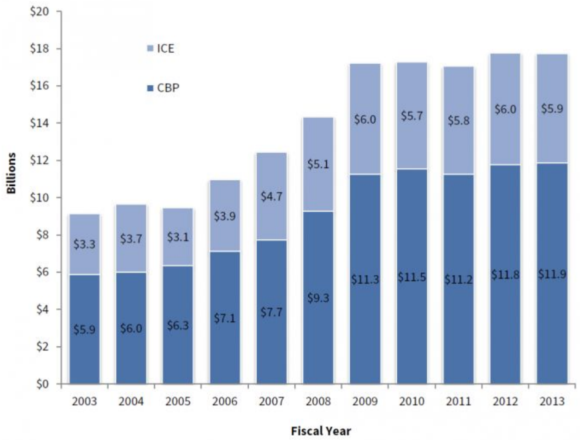


Image source: https://www.americanimmigrationcouncil.org/research/growth-us-deportation-machine

Actual effects of immigrants on economy

During the Great Recession:

- Better long-term economic resilience
  - Diversification (industry, geography)

Actual effects of immigrants on economy

During the Great Recession:

- Better long-term economic resilience
  - Diversification (industry, geography)
- Tended to be hit harder during Recession
  - Not “stealing jobs” during worst part

Actual effects of immigrants on economy

During the Great Recession:

- Better long-term economic resilience
  - Diversification (industry, geography)
- Tended to be hit harder during Recession
  - Not “stealing jobs” during worst part
- Tended to recover quicker after Recession
  - Not more dependent on government aid

Actual effects of immigrants on economy

During the Great Recession:

- Better long-term economic resilience
  - Diversification (industry, geography)
- Tended to be hit harder during Recession
  - Not “stealing jobs” during worst part
- Tended to recover quicker after Recession
  - Not more dependent on government aid

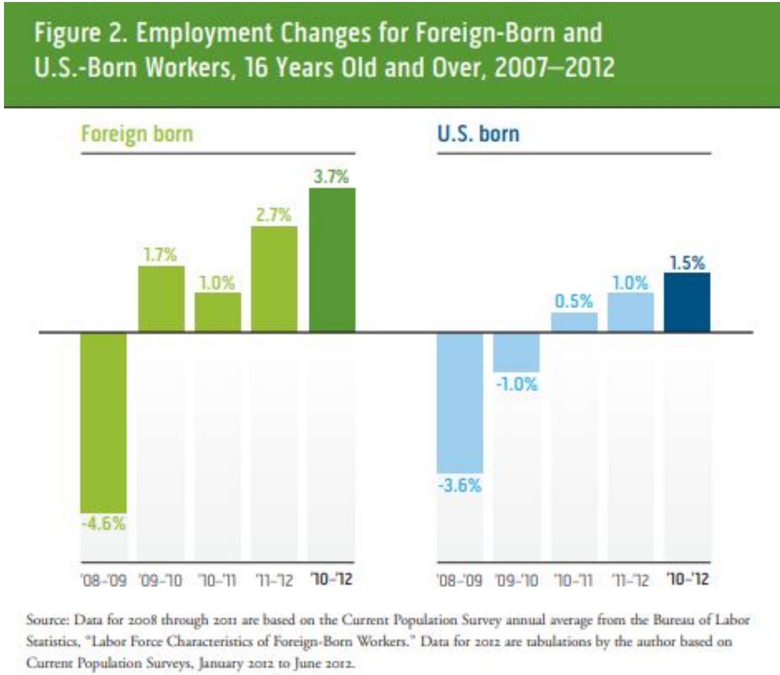


Image source:
https://www.urban.org/sites/default/files/publication/26181/412691-Hit-Hard-but-Bouncing-Back-The-Employment-of-Immigrants-During-the-Great-Recession-and-the-Recovery.PDF

Conclusions

1. History of popular anti-immigrant “threat rhetoric”
   a. Allowed politicians to easily enable discriminatory and sometimes illegal behavior.

Conclusions

1. History of popular anti-immigrant “threat rhetoric”
    a. Allowed politicians to easily enable discriminatory and sometimes illegal behavior.
2. Anti-immigrant sentiment not heavily rooted by fact

Conclusions

1. History of popular anti-immigrant “threat rhetoric”
    a. Allowed politicians to easily enable discriminatory and sometimes illegal behavior.
2. Anti-immigrant sentiment not heavily rooted by fact
3. Immigrants may be more sensitive to economic changes, but:
    a. They don’t contribute heavily to “taking American jobs,” instead they get hardest hit
    b. Recover quickly from the Recession
    c. Provide greater long-term stability

"all previous crises of the 1900s, including the Great Depression ... affected migration in different ways and spurred resentment of foreigners and xenophobic actions"

GMG and UNESCO, 2009

Works Cited

Conley, Meghan. "In times of uncertainty: The Great Recession, immigration enforcement, and Latino immigrants in Alabama." Immigrant Vulnerability and Resilience. Springer, Cham, 2015. 147-162.

Dancygier, Rafaela, and Michael Donnelly. "Attitudes Toward Immigration in Good Times and Bad." Mass Politics in Tough Times: Opinions, Votes and Protest in the Great Recession. : Oxford University Press, April 01, 2014. Oxford Scholarship Online.

Enchautegui, Maria E. "Hit hard but bouncing back: The employment of immigrants during the Great Recession and recovery." Washington, DC: Urban Institute (2012).

Golash-Boza, T., Hondagneu-Sotelo, P. Latino immigrant men and the deportation crisis: A gendered racial removal program. Lat Stud 11, 271–292 (2013). https://doi.org/10.1057/lst.2013.14.

Goldstein, Judith L., and Margaret E. Peters. "Nativism or economic threat: Attitudes toward immigrants during the Great Recession." International Interactions 40.3 (2014): 376-401.

Global Migration Group and UNESCO. Fact-sheet on the impact of the economic crisis on discrimination and xenophobia, 2009. https://www.gfmd.org/files/documents/gfmd_athens09_contr_unesco_factsheet_discrimination_and_xenophobia_en.pdf.

Hoffman, Abraham. Unwanted Mexican Americans in the Great Depression: Repatriation Pressures, 1929-1939. Vnr Ag, 1974.

Humes, Karen, Nicholas A. Jones, and Roberto R. Ramirez. "Overview of Race and Hispanic Origin, 2010." US Department of Commerce, Economics and Statistics Administration, US Census Bureau, 2011.

Kohut, Andrew, Caroli Doherty, Michael Dimock, and Scott Keeter. 2011. "Beyond Red vs. Blue: Political Typology." The Pew Research Center for the People & the Press.

Lester, William T. and Mai Thi Nguyen. The Economic Integration of Immigrants and Regional Resilience, Journal of Urban Affairs, 38:1, 42-60, DOI: 10.1111/juaf.12205, 2016.

National Conference of State Legislatures. 2013. "2012 Immigration-Related Laws and Resolutions in the States (Jan. 1-Dec. 31, 2012)." http://www.ncsl.org/research/immigration/2012-immigration-related-laws-jan-december-2012.aspx.

Lee, Jongkwan, Giovanni Peri, and Vasil Yasenov. The Employment Effects of Mexican Repatriations: Evidence from the 1930's. No. w23885. National Bureau of Economic Research, 2017.

Lester, T. William, and Mai Thi Nguyen. "The economic integration of immigrants and regional resilience." Journal of Urban Affairs 38.1 (2016): 42-60.

Perez, Efren O. 2010. "Explicit Evidence on the Import of Implicit Attitudes: The IAT and Immigration Policy Judgments." Political Behavior 32 (4): 517-45.

Ramos, Alice, Cícero Roberto Pereira, and Jorge Vala. "Economic crisis, human values and attitudes towards immigrants." Values, economic crisis and democracy. Routledge, 2016. 130-163.

@realDonaldTrump. “In light of the attack from the Invisible Enemy, as well as the need to protect the jobs of our GREAT American Citizens, I will be signing an Executive Order to temporarily suspend immigration into the United States!” Twitter, 20 Apr. 2020, 10:06 p.m., https://twitter.com/realDonaldTrump/status/1252418369170501639.

Siulc, Nina. Unwelcome citizens, criminalized migrants, and the quest for freedom: Deportees in the Dominican Republic. New York University, 2009.

Wallace, Michael, and Rodrigo Figueroa. “Determinants of Perceived Immigrant Job Threat in the American States.” Sociological Perspectives, vol. 55, no. 4, Dec. 2012, pp. 583–612, doi:10.1525/sop.2012.55.4.583.

Ybarra, Vickie D., Lisa M. Sanchez, and Gabriel R. Sanchez. "Anti-immigrant anxieties in state policy: The great recession and punitive immigration policy in the American states, 2005–2012." State Politics & Policy Quarterly 16.3 (2016): 313-339.

Zaretsky, Adam M. "A burden to America? immigration and the economy." The Regional Economist Oct (1997): 5-9.

TEST 2

Started : 3:54
Ended : 4:43

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis
4/6/20

1) $\log(1+i)$ $\quad 1+i = \sqrt{2}e^{i\frac{\pi}{4}}$

$= \ln|z| + i\arg z = \ln\sqrt{2} + i\left(\frac{\pi}{4} + 2\pi n\right),\ n \in \mathbb{Z}$

$= \frac{1}{2}\ln 2 + i\left(\frac{\pi}{4} + 2\pi n\right),\quad n \in \mathbb{Z}.$

2) $(-1+i)^{2i}$ (p.v.).

$= \exp(2i\,\mathrm{Log}(-1+i))$ $\quad -1+i = \sqrt{2}e^{\frac{3\pi}{4}i}$

$= \exp\left(2i\left(\ln\sqrt{2} + i\left(\frac{3\pi}{4}\right)\right)\right)$

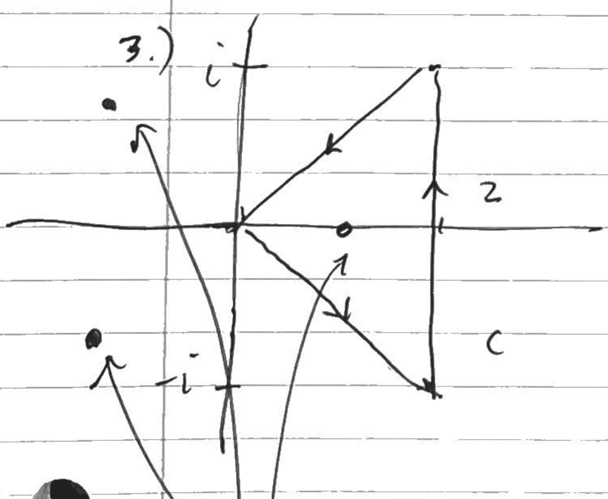
$= \exp\left(\cancel{2}i\left(\frac{1}{\cancel{2}}\right)\ln 2 + 2i^2\left(\frac{3\pi}{4}\right)\right)$

$= \exp\left(i\ln 2 - \frac{3\pi}{2}\right)$

$= \exp\left(-\frac{3\pi}{2}\right)\exp(i\ln 2)$

$= \exp\left(-\frac{3\pi}{2}\right)2^i = \exp\left(-\frac{3\pi}{2}\right)\left(\cos(\ln 2) + i\sin(\ln 2)\right)$

3.)



determine value of $\int_C \frac{z^2}{z^3-1}\,dz$

$z^3 - 1 = 0 \Rightarrow z = (1)^{\frac{1}{3}}$

$=$ (skipping algebra) : $\left\{e^{\pm\frac{2\pi}{3}i}, 1\right\}$

let $f(z) = \frac{z^2}{z^2+z+1}$

its singularities are at $e^{\pm\frac{2\pi}{3}i}$, so it is AOI C

singularities of integrand
(ones in LHP are singularities of f).

$$\int_C \frac{z^2}{z^3-1}\,dz =$$

f AOIC

CIF(0)

$$~~\text{[illegible]}~~ \qquad \int_C \frac{f(z)}{z-1}\,dz = 2\pi i\, f(1)$$

POSCC

1 is an interior point to C

$$= 2\pi i \left(\frac{1^2}{1^2+1+1}\right)$$

$$= 2\pi i \left(\frac{1}{3}\right)$$

$$= \frac{2}{3}\pi i.$$

4.) a) $C = |z| = 1$, PO.

a) Find $\int_C \bar{z} + z \, dz$

$$z = e^{i\theta}, \quad 0 \le \theta \le 2\pi$$

$$z' = ie^{i\theta}$$

$$= \int_0^{2\pi} \left(e^{-i\theta} + e^{i\theta}\right) ie^{i\theta}\, d\theta$$

$$= \int_0^{2\pi} i\left((1) + e^{2i\theta}\right) d\theta = i\int_0^{2\pi} d\theta + \int_0^{2\pi} ie^{2i\theta}\, d\theta$$

antiderivative is θ

antiderivative is $\frac{1}{2}e^{2i\theta}$

$$= i\theta \Big|_0^{2\pi} + \frac{1}{2}e^{2i\theta}\Big|_0^{2\pi}$$

$$= 2\pi i + \frac{1}{2}\left(e^{4i\pi} - e^{0}\right)$$

$$= 2\pi i + \cancel{\tfrac{1}{2}(1-1)} = 2\pi i$$

Jonathan Lau

4b). $\int_C e^{z^2} dz$.

composition of two entire functions (e^z and z^2), thus also entire.

thus $\int_C e^{z^2} dz = 0$.

SCC   AOIC   CG

5.) $C: |z| = 2$ (P.O.)

AOIC (entire, since composition of entire fns)

$$\int_C \frac{e^{5z}}{(z-1)^3} = \frac{2\pi i}{2!} \frac{d^2}{dz^2} e^{5z} \Big|_{z=1}$$

POSCC   1 is an interior point to C   CIF(2)

$$\frac{d}{dz} e^{5z} = 5e^{5z}$$

$$\frac{d^2}{dz^2} e^{5z} = 25 e^{5z}$$

$$= \frac{\cancel{2}\pi i}{\cancel{2}} \cdot 25 e^{5(1)} = 25\pi i e^5.$$

6.) $z(t) = 4e^{i2t} + e^{i11t}$, $0 \le t \le 2\pi$.

find $\displaystyle\int_C \frac{dz}{z^2}$

$z(0) = 4e^{i(0)} + e^{i(0)} = 4 + 1 = 5$,
$z(2\pi) = 4e^{i(4\pi)} + e^{i(22\pi)} = 4 + 1 = 5$

$\Downarrow$

final pt. = initial pt $\Rightarrow$ CC.

antiderivative of $\frac{1}{z^2} = \frac{-2}{z^3}$, antiderivative

exists everywhere.

antiderivative thm.

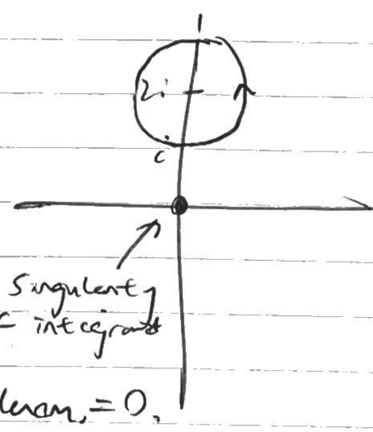
thus $\displaystyle\int_C \frac{dz}{z^2} = 0$

closed contour

antiderivative exists everywhere

7.) $C: |z - 2i| = 1$ (P.O.)

Find $\displaystyle\int_C \frac{e^z}{z}\, dz$



Let $f(z) = \frac{e^z}{z}$

f is quotient of two entire fns, so only singularities when denom. = 0.
Thus analytic in $\mathbb{C} \setminus \{0\}$.

AOIC

$\displaystyle\int_C f(z)\, dz = 0$

SCC

CG

Jonathan Lam

8.) $C_R: \ |z| = R$ (P.O.).

a) $\lim_{R\to\infty} \int_{C_R} \frac{1}{z^2+5z+i}\, dz$.

guess limit will be 0. use ML.

$|z^2+5z+i| \geq \big||z^2| - |5z+i|\big| = \big|R^2 - |5z+i|\big|$
(in limit as $R\to\infty$, $R^2 > 5z+i$, so we can remove outer absolute val. bars)
$= R^2 - |5z+i| \geq R^2 - \big||5z| - |i|\big| = R^2 - \big|5|z| - 1\big|$
$= R^2 - |5R-1|$
(same reasoning here: $R > 1$ in limit as $R \to \infty$)
$= R^2 - 5R - 1$.

thus $\left|\frac{1}{z^2+5z+i}\right| \leq \frac{1}{R^2-5R-1} = M$

length of arc $= 2\pi R$.

By ML, $\left|\int_{C_R} \frac{1}{z^2+5z+i}\, dz\right| \leq ML = \frac{2\pi R}{R^2-5R-1}$.

$\lim_{R\to\infty} \left|\int_{C_R} \frac{1}{z^2+5z+i}\, dz\right| = \lim_{R\to\infty} \frac{2\pi R}{R^2-5R-1} = 0$

(formal evaluation of limit not shown here)

$\Rightarrow \lim_{R\to\infty} \int_{C_R} \frac{1}{z^2+5z+i} = 0$.

(if limit of modulus $\to 0$, then limit of value $\to 0$).

8b) $\int_{C_6} \frac{1}{z^2+5z+i}\,dz$

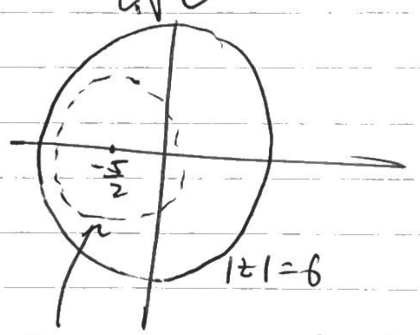
factor denom:

$$z = \frac{-5 \pm (5^2-4(1)(i))^{\frac{1}{2}}}{2(1)}$$

$$= \frac{-5 \pm (25-4i)^{\frac{1}{2}}}{2} = \pm z_s \text{ (singularities)}$$

$$\frac{\left(\sqrt{25^2+4^2}\right)^{\frac{1}{2}}}{2} \le 3.5$$

$$= \int_{|z|=6} \frac{1}{(z-z_s)(z+z_s)}\,dz$$



two singularities lie somewhere ~~on small circle~~ both within C. on this dotted circle.

(singularities lie on circle $|z+\frac{5}{2}| = \left|\frac{(25-4i)^{\frac{1}{2}}}{2}\right| \le 6$)

partial fractions:

$$\frac{1}{(z-z_s)(z+z_s)} = \frac{A}{z-z_s} + \frac{B}{z+z_s}$$

$A = \frac{1}{2}$

$B = -\frac{1}{2}$

Jonathan Lam

using partial fractions:

$$\int_C \frac{1}{(z-z_s)(z+z_s)}\,dz = \frac{1}{2}\left(\int_C \frac{1}{z-z_s}\,dz \cancel{+} - \int_C \frac{1}{z+z_s}\,dz\right)$$

numerator ADIC

interior point

posc

$$\overset{\text{CIF}(0)\ \text{twice}}{=} \quad \frac{1}{2}\left(2\pi i(1) - 2\pi i(1)\right) = 0.$$

Jonathan Lam
Prof. Smyth
MA345
Complex Analysis
3/31/20

PSET 5.

§53 #2, 4-7.

2. Let C_1 denote the POSCC of the square whose sides lie along the lines $x=\pm 2$, $y=\pm 1$, let C_2 be the POSCC $|z|=4$. Point out why $\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz$ when:

a) $f(z)=\frac{1}{3z^2+1}$

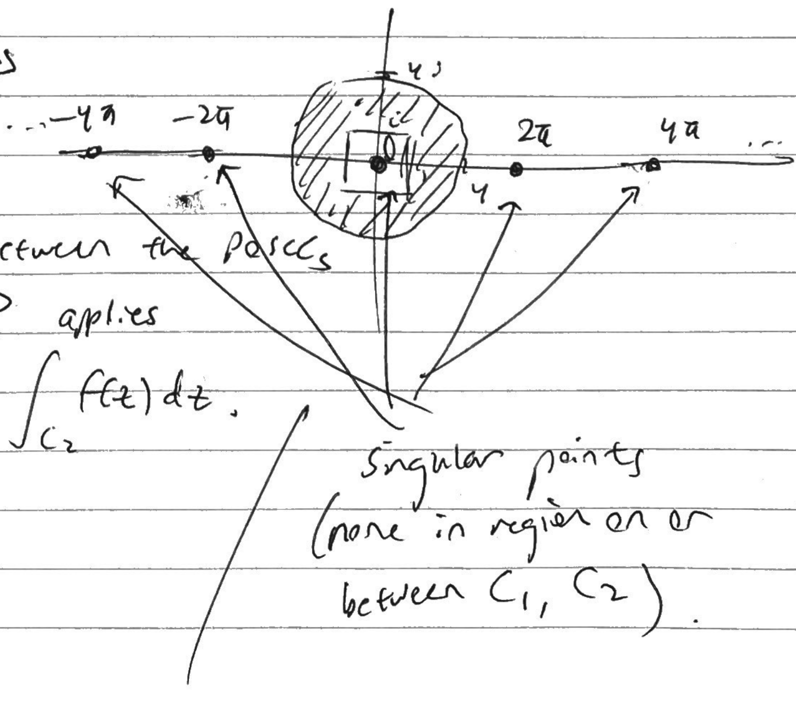
f is only non-analytic when $3z^2+1=0$, i.e., $z=\left(-\frac{1}{3}\right)^{\frac{1}{2}}$ at which points $|z|<1$. Thus f is analytic on the square C_1 and further from the origin, and thus in the region on and between C_1 and C_2. By the principle of deformation of paths ~~(PODOP)~~ (PDP), $\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz$.

b) $f(z) = \frac{z+2}{\sin\left(\frac{z}{2}\right)}$

f is analytic everywhere except where $\sin\left(\frac{z}{2}\right)=0$, since it is the ~~composition~~ quotient of entire functions. This occurs when $z=2\pi k$, $k\in\mathbb{Z}$, i.e.,

$$\sin\left(\frac{z}{2}\right) = \frac{e^{\frac{iz}{2}}-e^{-\frac{iz}{2}}}{2i}=0 \Rightarrow e^{\frac{iz}{2}}=e^{-\frac{iz}{2}} \Rightarrow e^{iz}=1 \Rightarrow z=2\pi k,\ k\in\mathbb{Z}$$

None of the singular points lies in the region on and between C_1 and C_2, i.e.,

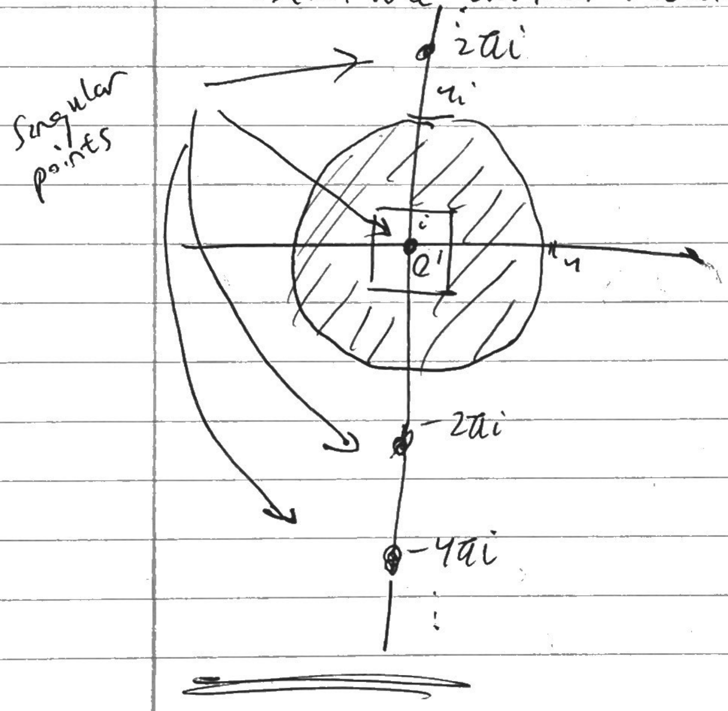


thus f is analytic everywhere in the region on and between the POSCCs C_1 and C_2, so the PDP applies to show that $\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz$.

3/31/20.

2c) $f(z) = \frac{z}{1-e^z}$.

Since this is the quotient of two analytic functions, it is analytic everywhere except where $1-e^z = 0$, i.e., when $e^z = 1$, i.e., where $z = 2\pi i k$, $k \in \mathbb{Z}$. As with part (b), none of these singular points lies in the region on and between the POSCCs C_1 and C_2, thus PDP applies to obtain the desired result: $\int_{C_1} f(z)dz = \int_{C_2} f(z)dz$.

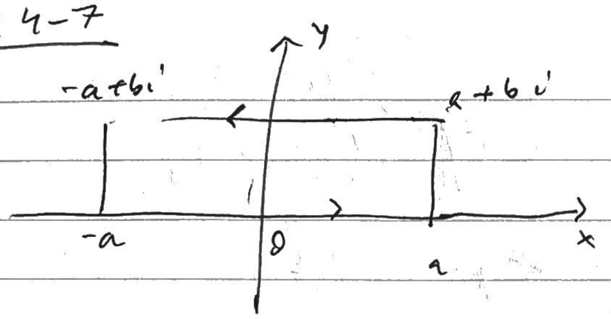


Jonathan Lam
Prof. Smyth
MA345
Complex Analysis
3/31/20

PSET 5

553 #4-7

4.)



Show that the sum of the integrals of e^{-z^2} along the lower and upper horizontal legs can be written

$$2\int_0^a e^{-x^2}dx - 2e^{b^2}\int_0^a e^{-x^2}\cos 2bx\,dx.$$

lower leg: $z(x)=x$, $z'(x)=dx$
$-a\le x\le a$

$$\int f(z)dz = \int_{-a}^{a} e^{-x^2}\cdot 1\,dx$$

e^{-x^2} is even, so this can be written:

$$= 2\int_a^a e^{-x^2}dx$$

upper leg: $z(x)=-x+bi$ $z'(x)=-dx$.
$-a\le x\le a$

$$\int F(z)dz = \int_{-a}^{a} e^{-(-x+bi)^2}\cdot(-1)\,dx = -\int_{-a}^{a} e^{-(x^2-b^2-2xbi)}dx$$

$$= -\int_{-a}^{a} e^{+b^2}e^{-x^2}e^{i(2xb)}dx = -e^{b^2}\int_{-a}^{a}e^{-x^2}(\cos(2xb)+i\sin(2xb))\,dx$$

$$= -e^{b^2}\int_{-a}^{a}\underbrace{\underbrace{e^{-x^2}}_{\text{even}}\underbrace{\cos 2bx}_{\text{even}}}_{\text{even}}\,dx - ie^{b^2}\int_{-a}^{a}\underbrace{\underbrace{e^{-x^2}}_{\text{even}}\underbrace{\sin 2bx}_{\text{odd}}}_{\text{odd}}\,dx$$

$$= -2e^{b^2}\int_0^a e^{-x^2}\cos 2bx\,dx \cancel{- ie^{b^2}(0)}$$

Thus the sum of the integrals on the horizontal legs (in the correct orientations) is:

$$2\int_0^a e^{-x^2}dx - 2e^{b^2}\int_0^a e^{-x^2}\cos 2bx\,dx.$$

3/31/20

Show that the sum of the integrals along the vertical legs can be written

$$ie^{-a^2}\int_0^b e^{y^2}e^{-i2ay}dy - ie^{-a^2}\int_0^b e^{y^2}e^{i2ay}dy.$$

Left leg: $z(y) = -a + yi,\quad 0 \le y \le b$ ← this is in the reverse direction, so negate the value later.

$z'(y) = i\,dy.$

$$-\int f(z)\,dz = -\int_0^b e^{-(-a+yi)^2}\cdot i\,dy$$

$$= -i\int_0^b e^{-(a^2-y^2-2ayi)}dy = -ie^{-a^2}\int_0^b e^{y^2}e^{i2ay}dy.$$

Right leg: $z(y) = a + yi \quad 0 \le y \le b.$

$z'(y) = i\,dy$

$$\int f(z)\,dz = \int_0^b e^{-(a+yi)^2}\,i\,dy = i\int_0^b e^{-(a^2-y^2+2ayi)}dy$$

$$= ie^{-a^2}\int_0^b e^{y^2}e^{-i2ay}dy.$$

Sum of integrals on left and right legs:

$$ie^{-a^2}\int_0^b e^{y^2}e^{-i2ay}dy - ie^{-a^2}\int_0^b e^{y^2}e^{i2ay}dy$$

$f(z) = e^{-z^2}$ is the composition of two entire functions, so it is analytic everywhere. Since the rectangle (C) is a SCC and f is AOIC, then $\int_C f(z)\,dz = 0$.

Thus:

$$\int_C f(z)\,dz = 2\int_0^a e^{-x^2}dx - 2e^{b^2}\int_0^a e^{-x^2}\cos 2bx\,dx + ie^{-a^2}\int_0^b e^{y^2}e^{-i2ay}dy - ie^{-a^2}\int_0^b e^{y^2}e^{i2ay}dy = 0.$$

$$\Rightarrow \underbrace{2e^{b^2}\int_0^a e^{-x^2}\cos 2bx\,dx}_{2e^{b^2}} = \underbrace{2\int_0^a e^{-x^2}dx + ie^{-a^2}\int_0^b e^{y^2}\left(e^{-2ayi} - e^{2ayi}\right)dy}_{2e^{b^2}}$$

$$\int_0^a e^{-x^2}\cos 2bx\,dx = e^{-b^2}\int_0^a e^{-x^2}dx + e^{-a^2}e^{-b^2}\int_0^b e^{y^2}\left(\frac{e^{(2ay)i} - e^{(-2ay)i}}{2i}\right)dy$$

$$= e^{-b^2}\int_0^a e^{-x^2}dx + e^{-(a^2+b^2)}\int_0^b e^{y^2}\sin 2ay\,dy.$$

Jonathan Lam
Prof. Smyth
MA345
Complex Analysis
3/31/20.

PSET 5

§53 #4b, -7.

4b.) Accept that $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$ (**), $\left|\int_0^b e^{y^2}\sin 2ay\,dy\right| \le \int_0^b e^{y^2}dy$ (*).

Take the limit of the expression from 4a as $a \to \infty$ to get the desired integration formula.

$$\int_0^\infty e^{-x^2}\cos 2bx\,dx = \lim_{a\to\infty}\int_0^a e^{-x^2}\cos 2bx\,dx = \lim_{a\to\infty}\left[e^{-b^2}\int_0^a e^{-x^2}dx + e^{-(a^2+b^2)}\int_0^b e^{y^2}\sin 2ay\,dy\right]$$

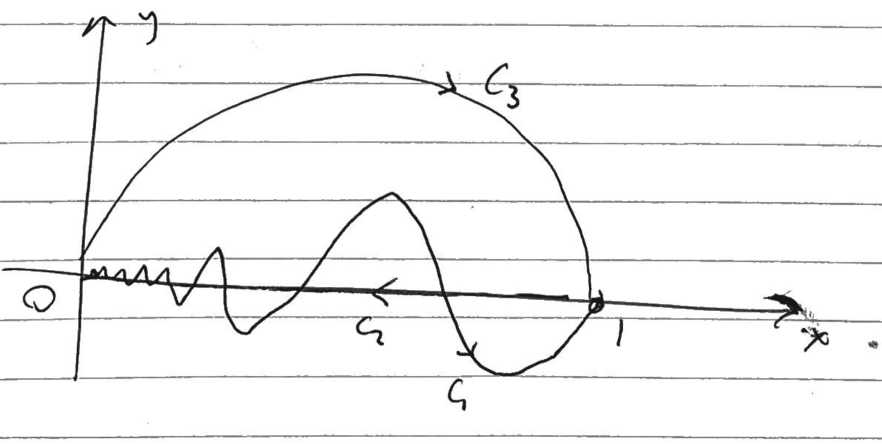
$$= \underbrace{e^{-b^2}\int_0^\infty e^{-x^2}dx}_{=\frac{\sqrt{\pi}}{2}\ (**)} + \underbrace{\underbrace{\lim_{a\to\infty}\left(e^{-a^2}\right)}_{=0} e^{-b^2}\underbrace{\int_0^b e^{y^2}\sin 2ay\,dy}_{\text{bounded by }(*)}}_{\text{limit of } 0 \text{ \& (some bounded finite value)} = 0}$$

$$= \frac{\sqrt{\pi}}{2}e^{-b^2} + 0 = \frac{\sqrt{\pi}}{2}e^{-b^2}.$$

5.) It was established in a previous exercise that the arc C_1 along the graph of the function

$$y(x) = \begin{cases} x^3\sin\left(\frac{\pi}{x}\right), & 0 < x \le 1 \\ 0 & x = 0 \end{cases}$$

is a smooth arc. Let C_2 denote the line segment along the real axis from $z=1$ to the origin, and let C_3 denote any smooth arc from $z=1$ back to the origin not intersecting C_1 or C_2.

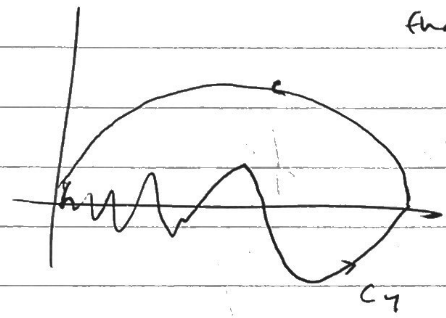


3/31/20.

Show that if f is entire, then

$$\int_{C_1} f(z)\,dz = \int_{C_3} f(z)\,dz, \qquad \int_{C_2} f(z)\,dz = -\int_{C_3} f(z)\,dz.$$

Let $C_4 = C_1 - C_3$. Then C_4 is a SCC (it is already established that C_1 is smooth, so this is a contour).

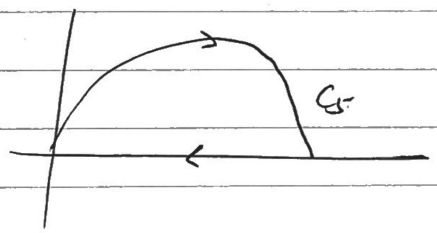


Since f is entire, then it is AOIC,

and $\int_{C_4} f(z)\,dz = \int_{C_1} f(z)\,dz - \int_{C_3} f(z)\,dz = 0$ ← C-G

$$\Rightarrow \int_{C_1} f(z)\,dz = \int_{C_3} f(z)\,dz. \quad (*)$$

Similarly, let $C_5 = C_2 + C_3$. Then C_5 is a (NO) SCC, and by the same reasoning:



$$\int_{C_5} f(z)\,dz = \int_{C_2} f(z)\,dz + \int_{C_3} f(z)\,dz = 0$$

SCC, AOIC, CG

$$\Rightarrow \int_{C_2} f(z)\,dz = -\int_{C_3} f(z)\,dz. \quad (**)$$

Thus, we can conclude that

$$\int_{C_1+C_2} f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz$$

$$= \underbrace{\int_{C_3} f(z)\,dz}_{(*)} - \underbrace{\int_{C_3} f(z)\,dz}_{(**)} = 0$$

even though the curve $C = C_1 + C_2$ self-intersects infinitely many times.

Jonathan Lam
Prof. Smyth
MA 345
Complex Analysis
3/31/20

PSET 5

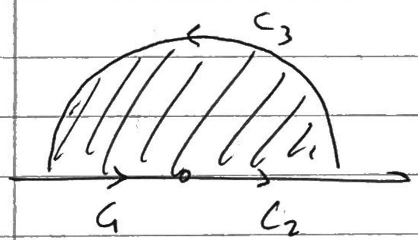
S53 # 6, 7

6.) Let C denote the POSCC of the half-disk $0 \le r \le 1$, $0 \le \theta \le \pi$, and let $f(z)$ be a cont. fn. defined on that half disk by writing $f(0)=0$ and using the branch

$$f(z) = \sqrt{r}\, e^{i\frac{\theta}{2}} \qquad \left(r>0,\ -\tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2}\right)$$

of the multiple-valued fn. $z^{\frac{1}{2}}$. Show that $\int_C f(z)\,dz = 0$ by evaluating separately the integrals of $f(z)$ over the semicircle and the two radii that make up C. Why does C-G not apply here?

C-G doesn't apply since f is not analytic at $z=0$, which lies on the curve.



in the negative direction

$z(r) = re^{i\pi}$, $0 \le r \le 1$, $z' = -1\,dr$

$$C_1:\ -\int_0^1 \sqrt{r}\, e^{i\left(\frac{\pi}{2}\right)} (-1)\, dr$$

$$= \int_0^1 \sqrt{r} \cdot i\, dr = \frac{2i}{3} r^{\frac{3}{2}}\Big|_0^1 = \frac{2i}{3}.$$

C_2: $z(r) = re^{i0}$, $0 \le r \le 1$, $z' = 1$

$$\int_0^1 \sqrt{r}\, e^{i\left(\frac{0}{2}\right)} \cdot 1\, dr = \int_0^1 \sqrt{r}\, dr = \frac{2}{3} r^{\frac{3}{2}}\Big|_0^1 = \frac{2}{3}.$$

C_3: $z(\theta) = e^{i\theta}$, $z' = ie^{i\theta}\, d\theta$, $0 \le \theta \le \pi$

$$\int_0^\pi \sqrt{1}\, e^{i\left(\frac{\theta}{2}\right)} \cdot ie^{i\theta}\, d\theta = \int_0^\pi i e^{i\frac{3\theta}{2}}\, d\theta = \frac{2}{3} e^{i\frac{3\theta}{2}}\Big|_0^\pi$$

use antiderivative

$$= \frac{2}{3}\left(e^{-\frac{3\pi}{2}} - e^{0}\right) = \frac{2}{3}(-i-1)$$

$$\int_{C_1+C_2+C_3} f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz + \int_{C_3} f(z)\,dz$$

$$= \frac{2i}{\cancel{3}} + \frac{2}{\cancel{3}} - \frac{2i}{\cancel{3}} - \frac{2}{\cancel{3}} = 0.$$

3/31/20

7.) Show that if C is a POSCC, then the area of the region enclosed by C can be written $\frac{1}{2i}\int_C \bar{z}\,dz$.

In Sec. 50, we derived expression (4), which can be applied to any SCC on any complex-valued function:

$$\int_C f(z)\,dz = \iint_R (-v_x - u_y)\,dA + i\iint_R (u_x - v_y)\,dA.$$

Since this only used the basic concepts of parametric integration and Green's theorem, and doesn't require analyticity of f, it may be used with $f(z) = \bar{z}$. Begin with the proposed formula:

$$\frac{1}{2i}\int_C \bar{z}\,dz = \frac{1}{2i}\left(\iint_R (-v_x - u_y)\,dA + i\iint_R (u_x - v_y)\,dA\right)$$

$$\left(\begin{array}{llll} f(z) = x - iy & u = x, & u_x = 1, & u_y = 0, \\ & v = -y, & v_y = -1, & v_x = 0. \end{array}\right)$$

$$= \frac{1}{2i}\left(\cancel{\iint_R (-0-0)\,dA} + i\iint_R (1-(-1))\,dA\right)$$

$$= \frac{1}{2i}\, i\iint_R 2\,dA = \frac{2i}{2i}\iint_R dA = A.$$ ✓

This confirms the hypothesis.

Jonathan Lam
Prof. Smyth
MA345
Complex Analysis
3/31/20.

PSET 5.

S57 # 1, 2a, 3.

1.) Let C denote the POSCC of the square whose sides lie on the lines $x = \pm 2$, $y = \pm 2$. Evaluate the integrals:

a) $$\int_C \frac{e^{-z}\,dz}{z - \left(\frac{\pi i}{2}\right)} = 2\pi i\left(e^{-\frac{\pi i}{2}}\right) = 2\pi i(-i) = -2\pi(-1) = 2\pi$$

e^{-z} ← AOIC; $\int_C$ ← POSCC; $\frac{\pi i}{2}$ ← interior point; $=$ ← CIF

b) $$\int_C \frac{\cos z}{z(z^2+8)}\,dz = \int_C \frac{\left(\frac{\cos z}{z^2+8}\right)}{(z-0)}\,dz = 2\pi i\left(\frac{\cos 0}{0^2+8}\right) = \frac{\pi i}{4}$$

AOIC (poles lie outside enclosed region); $\int_C$ ← POSCC; 0 ← interior point; $=$ ← CIF

c) $$\int_C \frac{z\,dz}{2z+1} = \int_C \frac{\frac{z}{2}}{z+\frac{1}{2}} = 2\pi i\left(\frac{-\frac{1}{2}}{2}\right) = -\frac{\pi i}{2}$$

$\frac{z}{2}$ ← AOIC; $=$ ← CIF; $\int_C$ ← POSCC; $-\frac{1}{2}$ is an interior point

d) $$\int_C \frac{\cosh z}{z^4}\,dz = \int_C \frac{\cosh z}{(z-0)^4}\,dz = \frac{2\pi i}{3!}\left(\frac{d^3}{dz^3}\cosh z\right)\Big|_{z=0}$$

$\cosh z$ ← AOIC; $\int_C$ ← POSCC; 0 ← interior point; $=$ ← (extended) CIF

$$\left(\frac{d^3}{dz^3}\cosh z = \sinh z\right)$$

$$= \frac{2\pi i}{6}\sinh(0) = \frac{2\pi i}{6}(0) = 0.$$

e) $$\int_C \frac{\tan\left(\frac{z}{2}\right)}{(z-x_0)^2}\,dz \qquad (-2 < x_0 < 2)$$

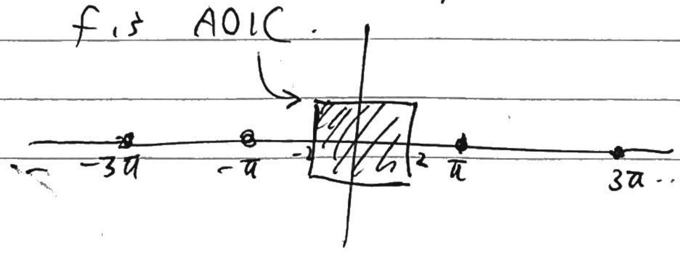
$\tan\left(\frac{z}{2}\right)$ ← AOIC; $\int_C$ ← POSCC; x_0 ← interior point

$$\underset{\text{CIF}}{=} \frac{2\pi i}{1!}\frac{d}{dz}\left(\tan\left(\frac{z}{2}\right)\right)\Big|_{x_0}$$

$$= \cancel{2}\pi i\left(\sec^2\left(\frac{z}{2}\right)\cdot\frac{1}{\cancel{2}}\right)\Big|_{x_0}$$

$$= \pi i \sec^2\left(\frac{x_0}{2}\right).$$

~~tan is analytic when cos~~
$\tan = \frac{\sin}{\cos}$ is analytic when $\cos \neq 0$, i.e., when $\frac{z}{2} \neq (2n+1)\frac{\pi}{2}$ i.e., when $z \neq (2n+1)\pi$. None of these singularities lie on C or in the enclosed region, thus f is AOIC.

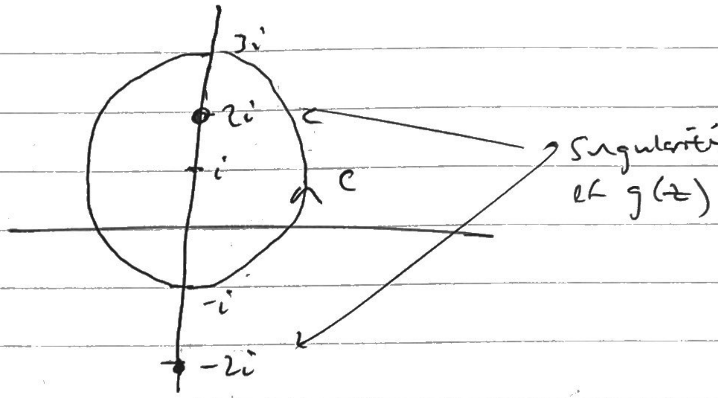


3/31/20.

2a) Find the value of the integral of $g(z)$ around the circle $|z-i| = 2$ in the positive sense when

a) $g(z) = \frac{1}{z^2+4}$

Roots of $z = (-4)^{\frac{1}{2}}$
$= \pm 2i$



let $h(z) = \frac{1}{z+2i}$

Then h is AOIC

$$\int_C \frac{1}{z^2+4}\,dz = \int_C \frac{\left(\frac{1}{z+2i}\right)}{z-2i}\,dz = 2\pi i\left(\frac{1}{2i+2i}\right) = \frac{8\pi i}{24i} = \frac{\pi}{2}.$$

posCC — interior point — CIF

3.) Let C be the circle $|z| = 3$. in the positive orientation. Show that if $g(z) = \int_C \frac{2s^2-s-2}{s-z}\,ds \quad (|z| \neq 3)$

then $g(2) = 8\pi i$. What is $g(z)$ when $|z| > 3$.

$$g(2) = \int_C \frac{(2s^2-s-2)}{s-2}\,ds = 2\pi i(2(2^2)-2-2) = 2\pi i(4) = 8\pi i.$$

AOIC — CIF — posCC — interior to $|z| < 3$

When $|z| > 3$, then the integrand becomes AOIC (ratio of polynomials, denominator $\neq 0$ since z is exterior to the circle). Thus, by C-G, the integral is zero.

Jonathan Lam
Prof. Smyth
MA 345
Complex Analysis
3/31/20.

PSET 5.

S57 # 4, 5, 7, 10

4) Let C be any POSCC, and write $g(z) = \int_C \frac{s^3+2s}{(s-z)^3}\,ds$.

Show that $g(z) = 6\pi i z$ when z is inside C and $g(z)=0$ when z is outside.

AOIC (polynomial)

when z is inside C: $$g(z) = \int_C \frac{s^3+2s}{(s-z)^3} \underset{\text{CIF}}{=} \frac{2\pi i}{2!}\left(\frac{d^2}{ds^2} s^3+2s\right)\Big|_z$$

POSCC — interior to C by hyp.

$$= \pi i (6s)\big|_z = 6\pi i z.$$

When z is outside C, then the integrand becomes AOIC (ratio of polynomials, denominator $\neq 0$), thus by C-G the integral evaluates to 0.

5.) Show that if f is AOIC, where C is a SCC and z_0 is not on C, then: $$\int_C \frac{f'(z)\,dz}{z-z_0} = \int_C \frac{f(z)\,dz}{(z-z_0)^2}$$

Let $g = f'$. If f AOIC, then g AOIC (since g is analytic where f is). Assume z_0 is within C and C is a POSCC.

Case 1.

AOIC

$$\int_C \frac{g(z)}{z-z_0}\,dz \underset{\text{CIF}}{=} 2\pi i\, g(z_0) = \frac{2\pi i}{1!} f'(z) \underset{\text{CIF}}{=} \int_C \frac{f(z)\,dz}{(z-z_0)^2}$$

POSCC — interior point — AOIC — POSCC — interior point

Now, assume z_0 is within C and C is a NOSCC,

Case 2.

Then $$\int_C \frac{g(z)}{z-z_0}\,dz = -2\pi i\, g(z_0) = \frac{-2\pi i}{1!} f'(z_0) = \int_C \frac{f(z)}{(z-z_0)^2}\,dz$$

for much of the same reasoning as case 1, except that the orientation is reversed.

Case 3.

Assume z_0 is not an interior point of the region enclosed by C. Then the integrand becomes AOIC and by C-G the integral evaluates to 0 (for both integrals).

3/31/20.

7.) Let C be the unit circle $z=e^{i\theta}$ $(-\pi \le \theta \le \pi)$.
First, show that for any real constant a;

$$\int_C \frac{e^{az}}{z}\,dz = 2\pi i.$$

$$\underbrace{\int_C}_{\text{Pos cc}} \frac{e^{az} \leftarrow \text{Analytic}}{\underbrace{z-0}_{\text{interior point}}}\,dz \underset{\text{CIF}}{=} 2\pi i\left(e^{a(0)}\right) = 2\pi i\left(e^{0}\right) = 2\pi i. \quad (*)$$

Then write this integral in terms of θ to derive the integration formula $\int_0^{\pi} e^{a\cos\theta}\cos(a\sin\theta)\,d\theta = \pi$.

$z = e^{i\theta}$, $-\pi \le \theta \le \pi$

$z' = ie^{i\theta}\,d\theta$.

$$\int_{-\pi}^{\pi} \frac{e^{a(\exp i\theta)}}{\cancel{e^{i\theta}}}\, i\cancel{e^{i\theta}}\,d\theta = i\int_{-\pi}^{\pi} e^{a(\cos\theta + i\sin\theta)}\,d\theta$$

$$= i\int_{-\pi}^{\pi} e^{a\cos\theta} e^{ia\sin\theta}\,d\theta$$

$$= i\int_{-\pi}^{\pi} e^{a\cos\theta}\left(\cos(a\sin\theta) + i\sin(a\sin\theta)\right)d\theta$$

$$= i\int_{-\pi}^{\pi} \underbrace{\underbrace{e^{a\cos\theta}}_{\text{even}}\underbrace{\cos(a\sin\theta)}_{\text{even}}}_{\text{even}}\,d\theta + i^2\int_{-\pi}^{\pi} \underbrace{\underbrace{e^{a\cos\theta}}_{\text{even}}\underbrace{\sin(a\sin\theta)}_{\text{odd}}}_{\text{odd}}\,d\theta$$

$$= \frac{\cancel{2i}\int_0^{\pi} e^{a\cos\theta}\cos(a\sin\theta)\,d\theta}{\cancel{2i}} + \cancel{i^2(0)} \overset{(\text{by } (*))}{=} \frac{2\pi i}{2i}$$

$$\Rightarrow \int_0^{\pi} e^{a\cos\theta}\cos(a\sin\theta)\,d\theta = \pi.$$

Jonathan Lam
Prof. Smyth
MA345
Complex Analysis
3/31/20

PSET 5.

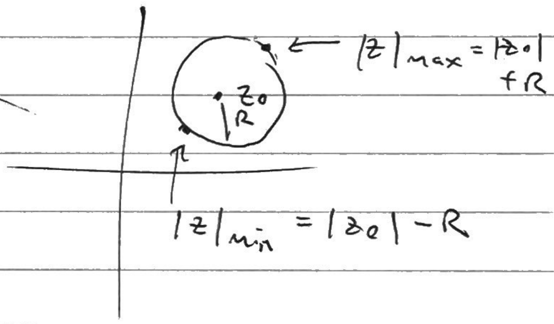
S57 #10.

10.) Let f be an entire function s.t. $|f(z)| \le A|z| \ \forall z \in \mathbb{C}$, where A is a fixed positive number. Show that $f(z) = a_1 z$, where a_1 is a complex constant.

Fix some point $z_0 \in \mathbb{C}$ and some $R \in \mathbb{R}^+$. Suppose C is a POSCC circle centered at z_0 w/ radius R.

Let M_R denote the maximum value of $|f(z)|$ on the circle C_R.

On C_R, $|z_0| - R \le |z| \le |z_0| + R$



$$\Rightarrow A(|z_0| - R) \le |f(z)| = A|z| \le A(|z_0| + R)$$

Thus $M_R \le A(|z_0| + R)$.

Since f is entire, then f is ADIC, and we may apply Cauchy's inequality:

$$|f^{(n)}(z_0)| \le \frac{n! M_R}{R^n} = \frac{n! A(|z_0| + R)}{R^n}$$

Note that this result must hold true for any choice of n, R, and z_0. Thus we can use this result to find $|f^{(n)}(z_0)|$ using a circle of any radius centered at z_0. Let $n=2$, and let R approach ∞:

$$\lim_{R\to\infty} |f''(z_0)| \left(= |f''(z_0)|\right) = \lim_{R\to\infty} \frac{(2!) A (|z_0| + R)}{R^2}$$

we can rewrite this infinite limit in another way:

$$|f''(z_0)| = \lim_{R\to 0} \left(\frac{2A\left(|z_0| + \frac{1}{R}\right)}{\left(\frac{1}{R}\right)^2} \right) \cdot \frac{R^2}{R^2} = \lim_{R\to 0} \frac{2A(|z_0|R^2 + R)}{1}$$

$$= 2A\left(\cancel{|z_0|(0)^2} + \cancel{0}\right) = 0 \Rightarrow |f''(z_0)| = 0 \Rightarrow f''(z_0) = 0$$
$$\forall z_0 \in \mathbb{C}.$$

Thus f must be a zeroth- or first-order polynomial of z, or else its second derivative would be nonzero. Thus $f(z) = a_1 z$, where $a_1 \in \mathbb{C}$.

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis
4/13/20

PSET 5.

§59 # 1, 2, 4

1.) Suppose that $f(z)$ is entire and that the harmonic function $u(x,y) = \mathrm{Re}[f(z)]$ has an upper bound u_0. Show that $u(x,y)$ must be constant through the plane.

If f is entire, then $g(z) = e^{f(z)}$ is entire (composition of entire functions is entire). If $v = \mathrm{Im}[f(z)]$, then

$g(z) = e^{u(x,y)} e^{iv(x,y)}$

and $|g(z)| = e^{u(x,y)} \le e^{u_0}$

(e^x is monotonically increasing and x has an upper bound)

By Liouville's thm., since $g(z)$ is analytic everywhere and $|g(z)|$ is bounded, then $g(z)$ is constant

$\Rightarrow |g(z)|$ is constant

$\Rightarrow e^{u(x,y)}$ is constant

$\Rightarrow u(x,y)$ is constant (since e^x is 1-to-1).

2.) Let a function f be continuous on a closed bounded region R, and let it be analytic and not constant through the interior of R. Assuming that $f(z) \neq 0$ anywhere in R, prove that $|f(z)|$ has a minimum value m in R which occurs on the boundary of R and never in the interior.

Let $g(z) = \frac{1}{f(z)}$ in the domain $D = R \setminus \partial R$ (interior of R). Since $f \neq 0$ in D and is analytic, g is analytic everywhere in D. Since f is not constant in D, g is not constant in D, and by Liouville's Thm, g has no maximum value in D.

PSET 5

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis
4/13/20

S59 # 2,4.

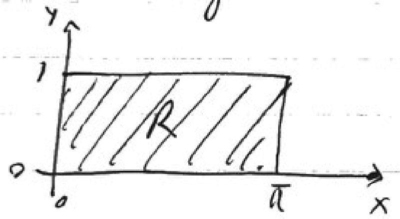
2, cont'd) This means that $\nexists\, z_0 \in D$ s.t. $|g(z)| \le |g(z_0)|$ $\forall z \in D$

$$\Rightarrow \nexists\, z_0 \in D \text{ s.t. } \left|\frac{1}{f(z)}\right| = \frac{1}{|f(z)|} \le \left|\frac{1}{f(z_0)}\right| = \frac{1}{|f(z_0)|} \quad \forall z \in D$$

$$\Rightarrow \nexists\, z_0 \in D \text{ s.t. } |f(z)| \ge |f(z_0)| \quad \forall z \in D$$

i.e., f has no minimum value in D. Since f is continuous on R, and R is closed and bounded, and doesn't achieve a minimum in its interior, it must achieve a minimum on the boundary of R.

4.) Let R be the region $0 \le x \le \pi$, $0 \le y \le 1$. Show that the modulus of the entire function $f(z) = \sin z$ has a maximum value in R at the boundary point $z = \frac{\pi}{2} + i$.



From S37, we know that:

$$|f(z)|^2 = |\sin z|^2 = \sin^2 x + \sinh^2 y \quad (\text{if } z = x+iy).$$

$|f(z)|$ should achieve a maximum when $|f(z)|^2$ achieves a maximum, and $|f(z)|^2$ should achieve a maximum value when $\sin^2 x$ and $\sinh^2 y$ achieve their maximums within R (since they are both positive values).

$\sin^2 x$ achieves a maximum when $|\sin x| = 1$, i.e., when $x = \frac{\pi}{2}(2n+1)$, $n \in \mathbb{Z}$. In R, it only achieves this once, at $n = \frac{1}{2}$ $(x = \frac{\pi}{2})$.

$\sinh^2 y$ achieves its maximum in R when $y = 1$, since $\sin^2 y$ is monotonically increasing for $y > 0$.

Thus $|f(z)|$ reaches its maximum @ $x = \frac{\pi}{2}$, $y = 1$ $(z = \frac{\pi}{2} + i)$. This agrees w/ the result given by the Liouville thm.

PSET 5

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
4/13/20

S65 # 2, 3, 11

2) Obtain the Taylor Series:
$$e^z = e\sum_{n=0}^{\infty}\frac{(z-1)^n}{n!} \qquad (|z-1|<\infty)$$

for $f(z)=e^z$ by:

a) Using $f^{(n)}(1)$, $n=0,1,2,\ldots$

This is a taylor series centered @ $z=1$.

$f^{(n)}(z) = e^z$, $n=0,1,2,\ldots$

$f^{(n)}(1) = e$, $n=0,1,2,\ldots$

using Taylor series formula:
$$f(z) = \sum_{n=0}^{\infty}\frac{f^{(n)}(z_0)}{n!}(z-z_0)^n \qquad \left(\begin{matrix} z_0=1, \\ |z-z_0|<\infty\end{matrix}\right)$$
since f is entire.
$$= \sum_{n=0}^{\infty}\frac{e}{n!}(z-1)^n = e\sum_{n=0}^{\infty}\frac{(z-1)^n}{n!}$$

b) writing $e^z = e^{z-1}e$.

From S64, we know that:
$$e^z = \sum_{n=0}^{\infty}\frac{z^n}{n!} \qquad (|z|<\infty)$$

make the substitution $z = z-1$

e^{z-1} analytic everywhere
$$e^{z-1} = \sum_{n=0}^{\infty}\frac{(z-1)^n}{n!} \qquad (|z-1|<\infty).$$

Thus
$$f(z) = ee^{z-1} = e\sum_{n=0}^{\infty}\frac{(z-1)^n}{n!} \qquad (|z-1|<\infty).$$

3.) Find the Maclaurin Series expansion of $f(z)=\frac{z}{z^4+4} = \frac{z}{4}\cdot\frac{1}{1+\left(\frac{z^4}{4}\right)}$.

From S64, we know that: $\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n, \quad (|z|<1)$.

Make the substitution: $z = -\frac{z^4}{4}$.
$$\frac{1}{1-\left(-\frac{z^4}{4}\right)} = \sum_{n=0}^{\infty}\left(-\frac{z^4}{4}\right)^n, \quad \left(\left|-\frac{z^4}{4}\right|<1\right) \Rightarrow \left(\frac{|z|^4}{4}<4 \Rightarrow |z|<\sqrt{2}\right)$$
$$\Rightarrow f(z) = \frac{z}{4}\cdot\frac{1}{1-\left(-\frac{z^4}{4}\right)} = \frac{z}{4}\sum_{n=0}^{\infty}\frac{(-1)^n z^{4n}}{4^n} = \sum_{n=0}^{\infty}\frac{(-1)^n z^{4n+1}}{4^{n+1}} \qquad (|z|<\sqrt{2}).$$

4/13/20.

11). Show that when $0<|z|<4$,

$$\frac{1}{4z-z^2} = \frac{1}{4z} + \sum_{n=0}^{\infty} \frac{z^n}{4^{n+2}}$$

$$f(z) = \frac{1}{4z-z^2} = \frac{1}{4z} \cdot \frac{1}{1-\frac{1}{4}z}$$

From S64, $\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$, $(|z|<1)$

make substitution : $z = \frac{1}{4}z$

$$\frac{1}{1-\frac{1}{4}z} = \sum_{n=0}^{\infty} \left(\frac{1}{4}z\right)^n , \quad \left(\left|\frac{z}{4}\right|<1\right)$$

$$= \sum_{n=0}^{\infty} \frac{z^n}{4^n} , \quad (|z|<4).$$

$$f(z) = \frac{1}{4z} \cdot \sum_{n=0}^{\infty} \frac{z^n}{4^n} \qquad (0<|z|<4)$$

$z \neq 0$ b/c of $\frac{1}{4z}$ term.

$$= \sum_{n=0}^{\infty} \frac{z^{n-1}}{4^{n+1}}$$

make substitution $n' = n-1$, $n = n'+1$

$$= \sum_{n'=-1}^{\infty} \frac{z^{n'}}{4^{n'+2}} = \frac{z^{-1}}{4^{-1+2}} + \sum_{n=0}^{\infty} \frac{z^n}{4^{n+2}} = \frac{1}{4z} + \sum_{n=0}^{\infty} \frac{z^n}{4^{n+2}}$$

rename back to n

PSET 5.

Jonathan Lam
Prof. Smyth
MA 345
Cplx-Analysis
4/13/20.

S68 # 1, 4, 6, 7, 8, 10.

1.) Find the Laurent series that represents the function,

$f(z) = z^2 \sin\left(\frac{1}{z^2}\right)$. in the domain $0<|z|<\infty$.

From S64: $\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$ $\quad (|z|<\infty)$.

Substituting: $z = \frac{1}{z^2}$:

$$\sin\frac{1}{z^2} = \sum_{n=0}^{\infty} (-1)^n \frac{\left(\frac{1}{z^2}\right)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{z^{-4n-2}}{(2n+1)!} \quad \left(z \neq 0 \text{ and } \left|\frac{1}{z^2}\right| < \infty\right)$$

thus $f(z) = z^2 \sin\left(\frac{1}{z^2}\right) = \cancel{z^2} \sum_{n=0}^{\infty} \frac{(-1)^n z^{-4n} \cancel{z^{-2}}}{(2n+1)!}$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n z^{-4n}}{(2n+1)!} \quad (0<|z|<\infty)$$

4.) Give two Laurent series expansions in powers of z (i.e., centered @ 0) for the function $f(z) = \frac{1}{z^2(1-z)}$ and specify the regions in which those expansions are valid.

a) for ROC $|z|<1$, from S64:

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$$

$$\Rightarrow f(z) = \frac{1}{z^2}\cdot\frac{1}{1-z} = \underset{\substack{\uparrow \\ z\neq 0}}{\frac{1}{z^2}} \sum_{n=0}^{\infty} z^n = \sum_{n=0}^{\infty} z^{n-2} \quad (0<|z|<1)$$

b) $\frac{1}{z^2(1-z)} = \frac{-1}{z^3}\cdot\frac{1}{1-\frac{1}{z}}$ ← make ~~sub~~ substitution $z = \frac{1}{z}$ in expansion of $\frac{1}{1-z}$.

$$= -\underset{\substack{\uparrow \\ z\neq 0}}{\frac{1}{z^3}} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n, \quad \underbrace{\left(|z|\neq 0 \text{ and } \left|\frac{1}{z}\right|<1\right)}_{\Downarrow \; \infty>|z|>1}$$

$$= -\sum_{n=0}^{\infty} z^{-n-3}, \quad (1<|z|<\infty).$$

make substitution $n' = n+3$

$$= -\sum_{n'=3}^{\infty} z^{-n'}$$

4/13/20

6.) Show that when $0 < |z-1| < 2$,

$$\frac{z}{(z-1)(z-3)} = -3\sum_{n=0}^{\infty} \frac{(z-1)^n}{2^{n+2}} - \frac{1}{2(z-1)}.$$

(i.e., express in powers of $(z-1)$, i.e., Laurent series centered at 1).

$$\frac{z}{(z-1)(z-3)} = \frac{A}{z-1} + \frac{B}{z-3}.$$

Use heaviside cover-up: $A = -\frac{1}{2}$, $B = \frac{3}{2}$.

$$= -\frac{1}{2(z-1)} + \frac{3}{2(z-3)}$$

$$= -\frac{3}{4}\left(\frac{1}{\frac{3}{2} - \frac{z}{2}}\right) - \frac{1}{2(z-1)}$$

$$= -\frac{3}{4}\left(\underbrace{\frac{1}{1-\left(\frac{z-1}{2}\right)}}\right) - \frac{1}{2(z-1)}$$

use (Taylor) expansion for $\frac{1}{1-z}$, substituting $z = \frac{z-1}{2}$.

$$= -\frac{3}{4}\sum_{n=0}^{\infty}\left(\frac{z-1}{2}\right)^n - \frac{1}{2(z-1)} \qquad \left(\left|\frac{z-1}{2}\right| < 1,\ |z-1| \neq 0\right)$$

$$= -\frac{3}{2^2}\sum_{n=0}^{\infty}\frac{(z-1)^n}{2^n} - \frac{1}{2(z-1)} \qquad (0 < |z-1| < 2)$$

$$= -3\sum_{n=0}^{\infty}\frac{(z-1)^n}{2^{n+2}} - \frac{1}{2(z-1)} \qquad (0 < |z-1| < 2)$$

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
4/23/20.

PSET 5

S68 # 7, 8, 10.

7.) Let a denote a real number, where $-1 < a < 1$, and derive the Laurent series representation:

$$\frac{a}{z-a} = \sum_{n=1}^{\infty} \frac{a^n}{z^n} \qquad (|a| < |z| < \infty).$$

Case 1: $a = 0$

$$\frac{0}{z-a} = \sum_{n=1}^{\infty} \frac{0^n}{z^n} \Rightarrow 0 = 0 \checkmark \qquad (0 < |z| < \infty)$$

doesn't work at $z=0$ b/c results in division by 0.

Case 2: $a \neq 0$:

$$\frac{a}{z-a} = \frac{a}{z} \cdot \underbrace{\frac{1}{1-\frac{a}{z}}}$$

↳ use Taylor series representation from S64 for $\frac{1}{1-z}$, substituting $z = \frac{a}{z}$

$$= \frac{a}{z} \cdot \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n \qquad \left(\left|\frac{a}{z}\right| < 1\right)$$

$$= \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^{n+1} \qquad (|a| < |z| < \infty)$$

Substitute $n' = n+1$

$$= \sum_{n'=1}^{\infty} \left(\frac{a}{z}\right)^{n'} \qquad (|a| < |z| < \infty)$$

4/13/20.

8.) Suppose that a series $X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n} = \sum_{n=-\infty}^{\infty} c_{-n} z^{-n}$

converges to an analytic function $X(z)$ in some annulus, $R_1 < |z| < R_2$. That sum $X(z)$ is called the z-transform of $x[n]$ $(n \in \mathbb{Z})$. Show that if the annulus contains the unit circle $|z| = 1$, then the inverse z-transform of $X(z)$ can be written:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\theta}) e^{in\theta} \, d\theta , \quad n \in \mathbb{Z}.$$

In general, an analytic fn. in some annulus can be expressed as $f(z) = \sum_{n=-\infty}^{\infty} c_n z^n$, where $c_n = \frac{1}{2\pi i} \int_C \frac{f(z)\,dz}{(z-z_0)^{n+1}}$

In this case, since $x[n]$ is the coefficient of a negative power of z, $x[n] = c_{-n}$.

By hypothesis, the domain is an annulus centered at the origin containing the unit circle. Thus let C: $z(\theta) = e^{i\theta}$ $(-\pi \le \theta \le \pi)$,

$z'(\theta) = ie^{i\theta}$,

$z_0 = 0$.

Thus $x[n] = c_{-n} = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{X(e^{i\theta})}{(e^{i\theta} - 0)^{-n+1}} ie^{i\theta} \, d\theta$

$= \frac{\cancel{i}}{2\pi \cancel{i}} \int_{-\pi}^{\pi} \frac{X(e^{i\theta})}{e^{-ni\theta}} \cdot \frac{\cancel{e^{i\theta}}}{\cancel{e^{i\theta}}} \, d\theta$

$= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{i\theta}) e^{ni\theta} \, d\theta .$

PSET 5.

Jonathan Lam
Prof. Smyth
MA 345
Cplx. Analysis
4/13/20.

568 #10.

(10) a) Let $f(z)$ denote a function which is analytic in some annular domain about the origin that includes the unit circle $z = e^{i\phi}$ $(-\pi \le \phi \le \pi)$. Show that

$$f(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\,d\phi + \frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} f(e^{i\phi})\left[\left(\frac{z}{e^{i\phi}}\right)^n + \left(\frac{e^{i\phi}}{z}\right)^n\right]d\phi$$

where z is any point in the annular domain.

By hyp, the annular domain (centered @ the origin) contains the unit circle, so we can use this as our path of integration.

$$C: z(\phi) = e^{i\phi} \qquad -\pi \le \phi \le \pi$$
$$z'(\phi) = ie^{i\phi}$$
$$z_0 = 0.$$

Using the formulation for a Laurent series:

$$a_n = \frac{1}{2\pi \cancel{i}}\int_{-\pi}^{\pi} \frac{f(e^{i\phi})}{(e^{i\phi})^{n\cancel{+1}}}\,\cancel{ie^{i\phi}}\,d\phi = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(e^{i\phi})}{e^{in\phi}}\,d\phi,$$

$$b_n = \frac{1}{2\pi \cancel{i}}\int_{-\pi}^{\pi} \frac{f(e^{i\phi})}{(e^{i\phi})^{-n\cancel{+1}}}\,\cancel{ie^{i\phi}}\,d\phi = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\cdot e^{in\phi}\,d\phi$$

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{f(e^{i\phi})}{e^{in\phi}}\right)(z-0)^n\,d\phi + \sum_{n=1}^{\infty} \frac{f(e^{i\phi})\,e^{in\phi}}{(z-0)^n}\,d\phi$$

take out 0^{th} term

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\cancel{\left(\frac{z}{e^{i\phi}}\right)^0}\,d\phi + \sum_{n=1}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\left(\frac{z}{e^{i\phi}}\right)^n d\phi$$
$$+ \sum_{n=1}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\left(\frac{e^{i\phi}}{z}\right)^n d\phi$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\,d\phi + \frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} f(e^{i\phi})\left[\left(\frac{z}{e^{i\phi}}\right)^n + \left(\frac{e^{i\phi}}{z}\right)^n\right]d\phi.$$

4/13/20

10b). Write $u(\theta) = Re[f(e^{i\theta})]$ and show that

$$u(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(\phi)\,d\phi + \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} u(\phi)\cos(n(\theta-\phi))\,d\phi.$$

(thus ~~show~~ deriving one form of the Fourier series expansion of a real-valued fn. $u(\theta)$ on the interval $-\pi \le \theta \le \pi$.)

$$u(\theta) = Re[f(e^{i\theta})]$$

$$= Re\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi})\,d\phi\right] + Re\left[\frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} f(e^{i\phi})\left[\left(\frac{e^{i\theta}}{e^{i\phi}}\right)^n + \left(\frac{e^{i\phi}}{e^{i\theta}}\right)^n\right]d\phi\right]$$

From S42, we can bring Re inside of integral (i.e., $Re\left[\int w(t)\,dt\right] = \int Re[w(t)]\,dt$.).

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} Re[f(e^{i\theta})]\,d\phi + \frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} Re\left[f(e^{i\phi})\underbrace{\left[e^{in(\theta-\phi)} + e^{in(\phi-\theta)}\right]}_{=2\cos(n(\theta-\phi))}\right]d\phi$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} u(\phi)\,d\phi + \frac{2}{2\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} Re\Big[\underbrace{f(e^{i\phi})}_{\text{complex}}\underbrace{\cos(n(\theta-\phi))}_{\text{real}}\Big]\,d\phi$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} u(\phi)\,d\phi + \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} u(\phi)\cos(n(\theta-\phi))\,d\phi.$$

Jonathan Lam
Prof. Smyth
MA 325
Cplx Analysis
4/27/20.

§72 # 3, 6, 7

3.) Find the Taylor series representation for:

$$\frac{1}{z} = \frac{1}{2+(z-2)} = \frac{1}{2} \cdot \frac{1}{1+\frac{z-2}{2}}$$

about $z_0 = 2$. Then, by differentiating termwise, show that:

$$\frac{1}{z^2} = \frac{1}{4}\sum_{n=0}^{\infty} (-1)^n (n+1)\left(\frac{z-2}{2}\right)^n, \quad (|z-2| < 2)$$

$$\frac{1}{2} \cdot \frac{1}{1+\frac{z-2}{2}} = \frac{1}{2} \cdot \sum_{n=0}^{\infty} (-1)^n \left(\frac{z-2}{2}\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n (z-2)^n}{2^{n+1}}$$

$$\left(\left|-\left(\frac{z-2}{2}\right)\right| \le 1 \Rightarrow |z-2| < 2\right).$$

differentiating:

$$\frac{d}{dz}\frac{1}{z} = -\frac{1}{z^2} = \frac{d}{dz}\sum_{n=0}^{\infty} \frac{(-1)^n}{2^{n+1}}(z-2)^n = \sum_{n=0}^{\infty} \frac{d}{dz}\frac{(-1)^n}{2^{n+1}}(z-2)^n$$

term by term differentiating allowed

$$\Rightarrow \frac{1}{z^2} = -\sum_{n=0}^{\infty} \frac{(-1)^n}{2^{n+1}} n (z-2)^{n-1}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^{n-1}}{2^{n+1}} n (z-2)^{n-1}$$

let $m = n-1$; $$= \sum_{m=-1}^{\infty} \frac{(-1)^m}{2^{m+2}}(m+1)(z-2)^m$$

($m = -1$ term $= 0$, so can start @ 1)

$$= \frac{1}{4}\sum_{m=0}^{\infty} (-1)^m (m+1)\left(\frac{z-2}{2}\right)^m$$

§72 #6, 7

6) In the w plane, integrate the Taylor series representation

$$\frac{1}{w} = \sum_{n=0}^{\infty} (-1)^n (w-1)^n \qquad (|w-1|<1)$$

along a contour interior to its circle of convergence from $w=1$ to $w=z$ to obtain the representation:

$$\mathrm{Log}\, z = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (z-1)^n .$$

allowed to integrate power series representation termwise and get the correct result (§71 result):

$$\int \frac{1}{w} = \sum_{n=0}^{\infty} \int (-1)^n (w-1)^n = \sum_{n=0}^{\infty} (-1)^n \frac{(w-1)^{n+1}}{n+1}$$

$$\text{let } m = n+1 \quad = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(w-1)^m}{m}.$$

$$\int_1^z \frac{1}{w}\,dw = \mathrm{Log}\, w \Big|_1^z = \underbrace{\mathrm{Log}\, z - \cancel{\mathrm{Log}\, 1} = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}}_{\text{desired result.}}$$

(Note that this only is true when entire contour lies in circle of convergence)

7). Show that if $f(z) = \frac{\mathrm{Log}\, z}{z-1}$ when $z \neq 1$ and $f(1) = 1$, then f is analytic throughout the domain $0<|z|<\infty$, $-\pi < \mathrm{Arg}\, z < \pi$.

We already know that this function is analytic throughout the specified domain $\setminus \{1\}$, since the numerator is analytic throughout this domain and the denominator is entire and has a zero at $z=1$. We can use the Taylor series expansion for $\mathrm{Log}\, z$ in a deleted nbd of $z=1$ to make an analytic continuation of f @ 1:

$$f(z) = \underbrace{\frac{\sum_{n=1}^{\infty} (-1)^{n+1} \frac{(z-1)^n}{n}}{z-1}}_{\text{expression}} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(z-1)^{n-1}}{n} = \sum_{n=0}^{\infty} (-1)^n \frac{(z-1)^n}{n+1}$$

$$(0<|z-1|<1)$$

Since this ~~function~~ expression is clearly analytic (only positive powers of $(z-1)$) @ $z=1$ and matches the value of f in a deleted nbd of 1, if we assign $f(1) = \sum_{n=0}^{\infty} (-1)^n \frac{(1-1)^n}{n+1} = 1$, it also becomes analytic @ 1.

§§73 # 1, 5.

1.) <u>Show that</u> $\dfrac{e^z}{z(z^2+1)} = \dfrac{1}{z} + 1 - \dfrac{1}{2}z - \dfrac{5}{6}z^2 + \dots \quad (0<|z|<1)$

$$f(z) = \frac{1}{z} \cdot e^z \cdot \frac{1}{1-(-z^2)} = \frac{1}{z}\left(1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \dots\right)\left(1 - z^2 + z^4 - z^6 + \dots\right)$$

$$= \frac{1}{z}\left(\begin{array}{l} 1 \quad -z^2 \quad +z^4 \quad -z^6 + \dots \\ \quad +z \quad -z^3 \quad +z^5 \quad -z^7 + \dots \\ \quad + \frac{z^2}{2} \quad - \frac{z^4}{2} \quad + \frac{z^6}{2} \quad + \dots \\ \quad + \frac{z^3}{6} \quad - \frac{z^5}{6} \quad + \dots \end{array}\right)$$

$$= \frac{1}{z}\left(1 + z - \frac{z^2}{2} - \frac{5z^3}{6} + \dots\right) = \frac{1}{z} + 1 - \frac{z}{2} - \frac{5z^2}{6} + \dots$$

5.) Note how the expansion:

$$\frac{1}{z^2 \sinh z} = \frac{1}{z^3} - \frac{1}{6} \cdot \frac{1}{z} + \frac{7}{360} z + \dots \quad (0<|z|<\pi)$$

follows from the example in §73. Show that:

$$\int_C \frac{dz}{z^2 \sinh z} = -\frac{\pi i}{3}, \quad C \text{ is the P.O. circle } |z|=1.$$

z^2 has a zero @ $z=0$,

$\sinh z$ has zeros @ $z = n\pi i$, $n \in \mathbb{Z}$ (§38).

Thus f is analytic in the ~~annulus~~ punctured disk $(0<|z-0|<\pi)$,

So there is a (unique) Laurent series representation of $f(z) = \sum_{n=-\infty}^{\infty} c_n (z-0)^n$, where C is a ~~[illegible]~~ basic circle encircling the origin and completely lying in the punctured disk (which is true by hypothesis), and $c_n = \dfrac{1}{2\pi i}\displaystyle\int_C \frac{f(z)}{(z-0)^{n+1}}\,dz$.

Thus $\displaystyle\int_C \frac{dz}{z^2 \sinh z} = \int_C \frac{1}{(z^2 \sinh z)\, z^{-1+1}}\,dz = \int_C \frac{f(z)}{z^{-1+1}}\,dz$

$$= 2\pi i\, c_{-1} = 2\pi i \left(-\frac{1}{6}\right) = -\frac{\pi i}{3}.$$

coefficient of the $\frac{1}{z}$ term

§77 # 1a,b, 2a,c, 5.

1) Find the residue @ $z=0$ of:

a) $\frac{1}{z+z^2} = \frac{1}{z} \cdot \frac{1}{1+z} = \frac{1}{z}\left(\frac{1}{1-(-z)}\right)$

punctured disk centered @0

$(0<|z-0|<1)$

$= \frac{1}{z}(1 - z + z^2 - z^3 + \dots) = \frac{1}{z} - 1 + z - \dots$

$\operatorname{Res}_{z=0} f(z) = 1$

b) $z\cos\left(\frac{1}{z}\right) = z\left(1 - \frac{\left(\frac{1}{z}\right)^2}{2!} + \frac{\left(\frac{1}{z}\right)^4}{4!} - \dots\right)$ $(0<|z-0|<\infty)$

in punctured disk centered at 0.

$= z - \frac{1}{2z} + \frac{1}{24z} - \dots$

$\operatorname{Res}_{z=0} f(z) = -\frac{1}{2}$

2a) Evaluate the integral of each of these fns. around the P.O. circle $|z|=3$ using Residue thm.

$f(z) = \frac{\exp(-z)}{z^2}$ one singularity of f, at $z=0$, enclosed in C.

$\int_C f(z)\,dz = 2\pi i \sum_{\substack{\text{enclosed}\\ \text{sing's}}} \operatorname{Res} f(z) = 2\pi i \operatorname{Res}_{z=0} f(z)$ (*)

posCC

AOIC except at $z=0$

residue thm

$(0<|z-0|<\infty)$

$f(z) = \frac{1}{z^2}\left(1 - z + \frac{z^2}{2} - \frac{z^3}{6} + \dots\right) = \frac{1}{z^2} - \frac{1}{z} + \frac{1}{2} - \dots$

$\operatorname{Res}_{z=0} f(z) = -1$

plugging this into (*), we get:

$\int_C f(z)\,dz = -2\pi i.$

§77 #2c, 5.

2c). $f(z) = z^2 \exp\left(\frac{1}{z}\right)$, C = (same curve as before).

only one singularity of f (enclosed within C) @ z = 0.

$f(z) = z^2\left(1 + \frac{1}{z} + \frac{\left(\frac{1}{z}\right)^2}{2} + \frac{\left(\frac{1}{z}\right)^3}{6} + \cdots\right)$ $(0<|z-0|<\infty)$

$= z^2 + z + \frac{1}{2} + \frac{1}{6z} + \cdots$

$\operatorname*{Res}_{z=0} f(z) = \frac{1}{6}$

residue thm

$\int_C f(z)\,dz = 2\pi i \sum_{\text{sing's}} \operatorname{Res} f(z) = 2\pi i\left(\frac{1}{6}\right) = \frac{\pi i}{3}$.

POSCC

AOIC except @ one interior singularity

5.) Let C denote the P.O. circle $|z| = 1$. Show that $\int_C \exp\left(z + \frac{1}{z}\right) dz = 2\pi i \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!}$.

a) expand using the Maclaurin series for $\exp z$

$\int_C \exp\left(z+\frac{1}{z}\right) dz = \int_C \exp z \exp\frac{1}{z}\,dz = \int_C \sum_{n=0}^{\infty} \frac{z^n}{n!} \exp\frac{1}{z}\,dz$

$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_C z^n \exp\left(\frac{1}{z}\right) dz.$ ⊛

integration can be taken termwise, §71

b) use residue thm. Let $f(z) = z^n \exp\left(\frac{1}{z}\right)$

$\int_C \underbrace{z^n \exp\left(\frac{1}{z}\right)}_{\text{AOIC except of a single singular point, } z=0} dz = 2\pi i \sum_{\text{enclosed sing's}} \operatorname{Res} f(z) = 2\pi i \operatorname*{Res}_{z=0} f(z).$

POSCC

$f(z) = z^n \sum_{m=0}^{\infty} \left(\frac{1}{z}\right)^m \cdot \frac{1}{m!} = \sum_{m=0}^{\infty} \frac{z^{n-m}}{m!}$. $\operatorname*{Res}_{z=0}$ = coefficient of z^{-1}

term, i.e. when $n = m-1$ = $\frac{1}{(n+1)!}$ $\Rightarrow \int_C f(z)\,dz = \frac{2\pi i}{(n+1)!}$

Plugging into ⊛, $\int_C \exp\left(z+\frac{1}{z}\right) dz = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{2\pi i}{(n+1)!}\right) = 2\pi i \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!}$

§79 #1a, b, 2, 4.

1) Write the principal part of the function at its isolated singular point and determine its type:

a) $f(z) = z\exp\left(\frac{1}{z}\right)$ isolated singular pt @ $z=0$.

$$= z\left(\sum_{n=0}^{\infty}\left(\frac{1}{z}\right)^n \cdot \frac{1}{n!}\right) = z\left(1+\frac{1}{z}+\frac{1}{2z^2}+\dots\right) \quad (0<|z-0|<\infty)$$

$$= z+1+\underbrace{\frac{1}{2z}+\frac{1}{6z^2}+\frac{1}{24z^3}+\dots}_{\text{principal part}}$$

principal part has infinitely-many terms, so essential singularity.

b). $f(z) = \frac{z^2}{1+z}$ isolated singular pt. @ $z=-1$. $(0<|z+1|<\infty)$

$$= \frac{((z+1)-1)^2}{z+1} = \left((z+1)^2 - 2(z+1)+1\right)\cdot\frac{1}{z+1}$$

$$= (z+1) - 2 + \underbrace{\frac{1}{z+1}}_{\text{principal part}} \Rightarrow \text{simple pole @ } z=-1.$$

2.) Show that the singular part of each of the following fns is a pole. Determine the order m of that pole and corresponding residue B,

a) $f(z) = \frac{1-\cosh z}{z^3}$ isolated sing. @ $z=0$.

$$= \frac{1}{z^3}\left(\cancel{1}-\left(\cancel{1}+\frac{z^2}{2!}+\frac{z^4}{4!}+\frac{z^6}{6!}+\dots\right)\right) \quad (0<|z-0|<\infty)$$

$$= \underbrace{-\frac{1}{2z}}_{\text{P.P.}} - \frac{z}{4!} - \frac{z^3}{6!} - \frac{z^5}{8!} - \dots$$

$B = \operatorname{Res}_{z=0} f(z) = -\frac{1}{2}$, order $= 1$.

§79 # 2, 4.

2b). $f(z) = \dfrac{1-\exp(2z)}{z^4}$ isolated sing. pt. @ $z=0$.

$= \frac{1}{z^4}\left(\cancel{1} - \left(\cancel{1} + (2z) + \frac{(2z)^2}{2!} + \frac{(2z)^3}{3!} + \cdots\right)\right)$ $(0<|z-0|<\infty)$

$= \underbrace{-\frac{2}{z^3} - \frac{4}{2z^2} - \frac{8}{6z}}_{P.P.} - \frac{16}{24} - \cdots$

$m = 3$ $B = \operatorname{Res}_{z=0} f(z) = -\frac{4}{3}$

2c). $f(z) = \dfrac{\exp(2z)}{(z-1)^2}$ isolated sing. pt. @ $z=1$.

$= \frac{1}{(z-1)^2}\exp(2(z-1)+2)$ $(0<|z-1|<\infty)$

$= e^2\left(\frac{1}{(z-1)^2}\right)\left(1 + 2(z-1) + \frac{2^2(z-1)^2}{2!} + \frac{2^3(z-1)^3}{3!} + \cdots\right)$

$= \underbrace{\frac{e^2}{(z-1)^2} + \frac{2e^2}{z-1}}_{P.P.} + 2e^2 + \cdots$

$B = \operatorname{Res}_{z=1} f(z) = 2e^2$, $m = 2$.

4.) write the function $f(z) = \dfrac{8a^3z^2}{(z^2+a^2)^3}$ $(a>0)$

as $f(z) = \dfrac{\phi(z)}{(z-ai)^3}$ where $\phi(z) = \dfrac{8a^3z^2}{(z+ai)^3}$

Point out why $\phi(z)$ has a Taylor series representation about $z=ai$, then use it to show that the principal part of f at that point is:

$$\frac{\phi''(ai)/2}{z-ai} + \frac{\phi'(ai)}{(z-ai)^2} + \frac{\phi(ai)}{(z-ai)^3} = \frac{-i/2}{z-ai} - \frac{a/2}{(z-ai)^2} - \frac{a^2i}{(z-ai)^3}$$

§79 #4.

4, cont'd.) $\phi(z) = \frac{8a^3z^2}{(z+ai)^3}$ clearly has its only (isolated) singular point @ $z=-ai$, so it is analytic in some nbd of $z=ai$, so it must have a Taylor series expansion at $z=ai$:

$$\phi(z) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}(ai)}{n!}(z-ai)^n$$

Thus $f = (z-ai)^{-3} \sum_{n=0}^{\infty} \frac{\phi^{(n)}(ai)}{n!}(z-ai)^n$

$$= \sum_{n=0}^{\infty} \frac{\phi^{(n)}(ai)}{n!}(z-ai)^{n-3}$$

$$\circledast \quad = \underbrace{\frac{\phi(ai)}{(z-ai)^3} + \frac{\phi'(ai)}{(z-ai)^2} + \frac{\phi''(ai)}{2(z-ai)}}_{\text{p.p. of } f \text{ @ } z=ai} + \frac{\phi'''(ai)}{6\cancel{\ldots}}$$

---

to show that this equals the other side of the inequality, just have to calculate $\phi(ai)$, $\phi'(ai)$, $\phi''(ai)$.

$$\phi(ai) = \frac{8a^3(ai)^2}{(2ai)^3} = \frac{\cancel{8}a^{\cancel{5}}\cancel{i^2}}{\cancel{2^3}\cancel{a^3}i^{\cancel{3}}} = -a^2 i$$

$$\phi'(z) = 8a^3\left(\frac{(z+ai)^3(2z) - z^2(3(z+ai)^2)}{(z+ai)^6}\right)$$

$$= 8a^3\left(\frac{2z(z+ai) - 3z^2}{(z+ai)^{4}}\right)$$

$$\phi'(ai) = 8a^3\left(\frac{2ai(2ai) - 3(ai)^2}{(2ai)^5}\right) = \cancel{8}a^3\left(\frac{4(ai)^2 - 3(ai)^2}{\cancel{32}(ai)^{5}}\right) = -\frac{a}{2}$$

$$\phi''(z) = 8a^3\left(\frac{(z+ai)^4(\cancel{4z} + 2ai - 6\cancel{z}z) - (2z(z+ai) - 3z^2)(4(z+ai)^3)}{(z+ai)^8}\right)$$

$$\phi''(ai) = 8a^3\left(\frac{(2ai)^4\overbrace{\cancel{(4ai+2ai-6ai)}}^{=0} - \overbrace{\cancel{2ai(2ai) - 3(ai)^2}}^{(ai)^2}(4(2ai)^3)}{(2ai)^8}\right)$$

$$= \frac{-\cancel{2^5}\cancel{a^5}i^5}{\cancel{2^8}\cancel{a^8}i^8} = -i^{-3} = -i$$

Substituting into $\circledast$, we get,

$$\text{P.P. } f \text{ @ } (z=2) = \frac{-a^2 i}{(z-ai)^3} - \frac{a/2}{(z-ai)^2} - \frac{i/2}{z-ai}$$

§81 # 1b-d, 2b, c, 4, 6.

Henceforth call thm. in §80 the pole-residue thm.

1) Show that any singular point of each fn. below is a pole. Determine the order of each pole, and find the corresponding residue B.

b.) $f(z) = \dfrac{z^2+2}{z-1}$ (only) singular point: $z=1$.

$f(z) = \dfrac{\phi(z)}{(z-1)^1}$, $\phi(z) = z^2+2$

$\phi(1) = 1+2 = 3 \neq 0$

$\Downarrow$ ϕ AANZ @ $z=1$.

By pole-residue thm, f has simple pole @ $z=1$, and $B = \operatorname{Res}_{z=1} f(z) = \phi(1) = 3$.

c) $f(z) = \left(\dfrac{z}{2z+1}\right)^3$ only singular point: $z = -\frac{1}{2}$

$f(z) = \dfrac{\phi(z)}{(z+\frac{1}{2})^3}$, $\phi(z) = \left(\dfrac{z}{2}\right)^3$

$\phi(\frac{1}{2}) = \left(\dfrac{-1}{2}\right)^3 \cdot \frac{1}{2} = \frac{-1}{16}$

$\Downarrow$ ϕ AANZ @ $z = \frac{-1}{2}$

By pole-residue thm, f has triple pole @ $z = \frac{-1}{2}$, and $\operatorname{Res}_{z=-\frac{1}{2}} f(z) = \dfrac{\phi''(z_0)}{2!} = \dfrac{\frac{3}{4}z_0}{2} = \dfrac{-3}{16}$

d) $f(z) = \dfrac{e^z}{z^2+\pi^2}$. f has isolated singularities @ $z = \pm\pi i$.

$\phi_1 = \dfrac{e^z}{z-\pi i}$, $\phi_2 = \dfrac{e^z}{z+\pi i}$.

$f(z) = \dfrac{\phi_1(z)}{z+\pi i} = \dfrac{\phi_2(z)}{z-\pi i}$.

$\begin{cases} \phi_1(-\pi i) = \dfrac{e^{-\pi i}}{-2\pi i} = \dfrac{1}{2\pi i} \\ \phi_2(\pi i) = \dfrac{e^{\pi i}}{2\pi i} = -\dfrac{1}{2\pi i} \end{cases}$

ϕ_1 AANZ @ $z = -\pi i$,
ϕ_2 AANZ @ $z = \pi i$

Thus f has a pole of order 1 at both $z = \pm \pi i$,

and $\operatorname{Res}_{z=\pi i} f(z) = \phi_2(\pi i) = -\dfrac{1}{2\pi i}$

and $\operatorname{Res}_{z=-\pi i} f(z) = \phi_1(-\pi i) = \dfrac{1}{2\pi i}$

§81 #2b, c, 4, 6,

2). Show that:

b) $\underset{z=i}{\text{Res}} \dfrac{\text{Log } z}{(z^2+1)^2} = \dfrac{\pi+2i}{8}$

$f(z) = \dfrac{\text{Log } z}{(z^2+1)^2} = \dfrac{\frac{\text{Log } z}{(z+i)^2}}{(z-i)^2} = \dfrac{\phi(z)}{(z-i)^2}$, $\phi(z) = \dfrac{\text{Log } z}{(z+i)^2}$

$\phi(i) = \dfrac{\cancel{\ln(1)+} i(\frac{\pi}{2})}{(2i)^2} \neq 0 \Rightarrow \phi$ AANZ @ $z=i$

$\phi'(z) = \dfrac{(z+i)^2(\frac{1}{z}) - \text{Log } z\,(2(z+i))}{(z+i)^4}$

$\phi'(i) = \dfrac{(2i)^2(\frac{1}{i}) - (\cancel{\ln(1)} + i(\frac{\pi}{2}))(2(2i))}{(2i)^4}$

$= \dfrac{4i - i(\frac{\pi}{2})(4i)}{2^4} = \dfrac{\pi+2i}{8}.$

By pole-residue thm, f has double pole @ $z=i$,

and $\underset{z=i}{\text{Res}} f(z) = \dfrac{\phi'(i)}{1!} = \dfrac{\pi+2i}{8}$

c) $\underset{z=i}{\text{Res}} \dfrac{z^{\frac{1}{2}}}{(z^2+1)^2} = \dfrac{1-i}{8\sqrt{2}}$ $(|z|>0,\ 0<\arg z<2\pi)$

$f(z) = \dfrac{z^{\frac{1}{2}}}{(z^2+1)^2} = \dfrac{\frac{z^{\frac{1}{2}}}{(z+i)^2}}{(z-i)^2} = \dfrac{\phi(z)}{(z-i)^2}$, $\phi(z) = \dfrac{z^{\frac{1}{2}}}{(z+i)^2}$

$\phi(i) = \dfrac{i^{\frac{1}{2}}}{(2i)^2} = \dfrac{e^{(\frac{1}{2}\log i)}}{4i^2} = \dfrac{e^{\frac{1}{2}(\frac{\pi i}{2})}}{-4} = -\dfrac{e^{\frac{i\pi}{4}}}{4} \neq 0.$

$\Downarrow$ ϕ AANZ @ $z=i$.

By pole-residue thm, f has double pole @ $z=i$, and

$\underset{z=i}{\text{Res}} f(z) = \dfrac{\phi'(i)}{1!}$.

$\phi'(z) = \dfrac{(2i)^2(\frac{1}{2}z^{-\frac{1}{2}}) - z^{\frac{1}{2}}(2(z+i))}{(z+i)^4}$

$\phi'(i) = \dfrac{(2i)^2(\frac{1}{2}i^{-\frac{1}{2}}) - i^{\frac{1}{2}}(2(2i))}{(2i)^4}$

81 § 2c, 4, 6

2c, cont'd.

$$\left(\begin{array}{l} i^{-\frac{1}{2}} = e^{-\frac{1}{2}\log i} = e^{-\frac{1}{2}\left(i\frac{\pi}{2}\right)} = e^{-i\frac{\pi}{4}} \\ i^{\frac{1}{2}} = e^{i\frac{\pi}{4}} \end{array} \right)$$

$$\phi'(i) = \frac{(-4)\left(\frac{1}{2}\right)e^{-\frac{\pi}{4}i} - 4ie^{\frac{\pi}{4}i}}{16}$$

$$= \frac{-2e^{-\frac{\pi}{4}i} - 4e^{\frac{3\pi}{4}i}}{16} = \frac{-2e^{-\frac{\pi}{4}i} - 4\left(-e^{-\frac{\pi}{4}i}\right)}{16}$$

$$= \frac{e^{-\frac{\pi}{4}i}}{8} = \frac{1-i}{8\sqrt{2}} = \operatorname*{Res}_{z=i} f(z).$$

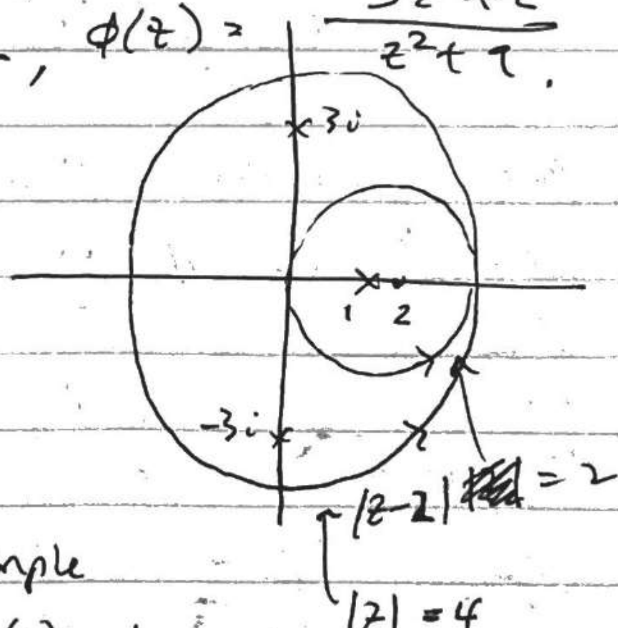
4.) Find the value of the integral

$$\int_C \frac{3z^3+2}{(z-1)(z^2+9)}\,dz$$

a) where C = p.o. circle $|z-2| = 2$.

$$f(z) = \frac{3z^3+2}{(z-1)(z^2+9)} = \frac{\phi(z)}{z-1}, \quad \phi(z) = \frac{3z^3+2}{z^2+9}.$$

only singularity enclosed within C is at $z=1$



$$\phi(1) = \frac{3+2}{1+9} = \frac{1}{2} \neq 0$$

$\Downarrow$

ϕ AANZ @ $z=1$.

By pole-residue thm, f has simple pole @ $z=1$, and $\operatorname*{Res}_{z=1} f(z) = \phi(1) = \frac{1}{2}$.

$$\text{Thus } \underbrace{\int_C}_{\text{posCC}} \underbrace{f(z)}_{\text{AOLC, except at simple pole } z=1}\,dz \overset{\text{residue thm}}{=} 2\pi i\left(\frac{1}{2}\right) = \pi i.$$

§81 # 4b, 6.

4b). $|z| = 4$, P.O.

$$f(z) = \frac{3z^3+2}{(z-1)(z^2+9)}.$$ 3 enclosed isolated singularities @ $z = 1, \pm 3i$.

$$f(z) = \frac{\frac{3z^3+2}{(z-1)(z+3i)}}{z-3i} = \frac{\phi_1(z)}{z-3i}. \longrightarrow \phi_1 \text{ AANZ } @ z = 3i$$

$$= \frac{\frac{3z^3+2}{(z-1)(z-3i)}}{z+3i} = \frac{\phi_2(z)}{z+3i}. \longrightarrow \phi_2 \text{ AANZ } @ z = -3i.$$

By pole-residue thm, f has a simple pole @ $z = \pm 3i$,

and $\displaystyle\operatorname*{Res}_{z=3i} f(z) = \phi_1(3i) = \frac{3(3i)^3+2}{(3i-1)(3i+3i)} = \frac{3\cdot 27(-i)+2}{(3i-1)(6i)}$

$$= \frac{2-81i}{-18-6i}$$

and $\displaystyle\operatorname*{Res}_{z=-3i} f(z) = \phi_2(-3i) = \frac{3(-3i)^3+2}{(-3i-1)(-3i-3i)} = \frac{81i+2}{-18+6i}.$

From part a, $\displaystyle\operatorname*{Res}_{z=1} f(z) = \frac{1}{2}.$

$$\int_C f(z)\,dz \overset{\text{residue thm}}{=} 2\pi i\left(\frac{1}{2} + \frac{2-81i}{-18-6i} + \frac{2+81i}{-18+6i}\right)$$

POSCC   AOLC except @ 3 isolated sing's

$$= 2\pi i\left(\frac{1}{2} + \frac{(2-81i)(-18+6i)+(2+81i)(-18-6i)}{18^2+6^2}\right)$$

$$= 2\pi i\left(\frac{1}{2} + \frac{-36+12i+1458i+486-36-12i-1458i+486}{324+36}\right)$$

$$= 2\pi i\left(\frac{1}{2} + \frac{900}{360}\right) = 2\pi i\left(\frac{1}{2}+\frac{5}{2}\right) = 6\pi i.$$

§81 #6

6). Evaluate the integral $\int_C \frac{\cosh \pi z}{z(z^2+1)}\,dz$ where C is the P.O. circle $|z|=2$.

let $f(z)=$ integrand. There are singularities @ $z=0, \pm i$. all lie within C.

Singularity @ $z=0$:

$f(z)=\frac{\phi(z)}{z}$, $\phi(z)=\frac{\cosh \pi z}{z^2+1}$, $\phi(0)=\frac{\cosh 0}{1}=1\neq 0$
$\Downarrow$
ϕ AANZ @ $z=0$.

By pole-residue thm, simple pole @ $z=0$,
$\operatorname{Res}_{z=0} f(z)=\phi(0)=1$.

Singularity @ $z=i$. $f(z)=\frac{\phi(z)}{z-i}$, $\phi(z)=\frac{\cosh \pi z}{z(z+i)}$.

$\phi(i)=\frac{\cosh \pi i}{i(2i)}$
$=\frac{e^{\pi i}+e^{-\pi i}}{2(-2)}=\frac{-1-1}{-4}=\frac{1}{2}\neq 0 \Rightarrow \phi$ AANZ @ $z=i$.

By pole-residue thm, $\operatorname{Res}_{z=i} f(z)=\frac{1}{2}$.

Singularity @ $z=-i$. $f(z)=\frac{\phi(z)}{z+i}$, $\phi(z)=\frac{\cosh \pi z}{z(z-i)}$.

$\phi(-i)=\frac{\cosh(-\pi i)}{-i(-2i)}=\frac{\cosh \pi i}{-2i}=\frac{1}{2}=0$ (same calculation as above)
$\Downarrow$
ϕ AANZ @ $z=-i$.

By pole-residue thm, $\operatorname{Res}_{z=-i} f(z)=\phi(-i)=\frac{1}{2}$.

$\int_C f(z)\,dz = 2\pi i\left(1+\frac{1}{2}+\frac{1}{2}\right)=4\pi i$.

POSCC

AOIC except at finitely many isolated sing's

residue thm

§83 # 1, 3, 5, 6, 8.

Henceforth call the thm 2 in §83 the simple pole-residue thm.

1). Show that the point $z=0$ is a simple pole of the fn:
$f(z) = \csc z = \frac{1}{\sin z}$ and that the residue is 1.

$f(z) = \frac{p(z)}{q(z)}$, $p(z)=1$, $q(z) = \sin z$, p, q entire
$p(0)=1$, $q(0) = \sin(0) = 0$,
$q'(0) = \cos(0) = 1$

By the simple pole-residue thm, f has a simple pole @ $z=0$ and $\operatorname{Res}_{z=0} f(z) = \frac{p(0)}{q'(0)} = \frac{1}{1} = 1$.

3.) Show that

a) $\operatorname{Res}_{z=\frac{i\pi}{2}} \frac{\sinh z}{z^2\cosh z} = -\frac{4}{\pi^2}$.

$f(z) = \frac{p(z)}{q(z)}$, $p(z) = \sinh z$, $q(z) = z^2\cosh z$ } Both entire fns.

$p\left(\frac{i\pi}{2}\right) = \frac{e^{\frac{i\pi}{2}} - e^{-\frac{i\pi}{2}}}{2} = \frac{2i}{2} = i \neq 0$.

$q\left(\frac{i\pi}{2}\right) = \left(\frac{i\pi}{2}\right)^2 \frac{e^{\frac{i\pi}{2}} + e^{-\frac{i\pi}{2}}}{2} = \left(\frac{i\pi}{2}\right)^2\left(\frac{i-i}{2}\right) = 0$

$q'\left(\frac{i\pi}{2}\right) = 2\left(\frac{i\pi}{2}\right)\cosh\left(\frac{i\pi}{2}\right) + \left(\frac{i\pi}{2}\right)^2 \sinh\left(\frac{i\pi}{2}\right)$

$= \cancel{2\left(\frac{i\pi}{2}\right)(0)} + \left(\frac{i\pi}{2}\right)^2 (i) = -\frac{i\pi^2}{4}$.

By simple pole-residue thm, f has simple pole @ $z = \frac{i\pi}{2}$,
and $\operatorname{Res}_{z=\frac{i\pi}{2}} f(z) = \frac{p\left(\frac{i\pi}{2}\right)}{q'\left(\frac{i\pi}{2}\right)} = \frac{i}{-\frac{i\pi^2}{4}} = -\frac{4}{\pi^2}$.

b) $\operatorname{Res}_{z=\pi i} \frac{\exp(zt)}{\sinh z} + \operatorname{Res}_{z=-\pi i} \frac{\exp(zt)}{\sinh z} = -2\cos(\pi t)$.

$p(z) = \exp(zt)$, $q(z) = \sinh z$, p, z are entire,
$z_{0_1} = \pi i$, $z_{0_2} = -\pi i$, $q'(z) = \cosh z$.

$p(z_{0_1}) = \exp(\pi i t) \neq 0$.
$q(z_{0_1}) = \sinh(\pi i) = \frac{e^{\pi i} - e^{-\pi i}}{2} = \frac{-1-(-1)}{2} = 0$.
$q'(z_{0_1}) = \cosh(\pi i) = \frac{e^{\pi i} + e^{-\pi i}}{2} = \frac{-1-1}{2} = -1$.

§83 #3b, 5, 6, 8.

3b, cont'd.) $p(z_{0_2}) = \exp(-\pi i t) \neq 0$.

$q(z_{0_2}) = \sinh(-\pi i) = -\sinh(\pi i) = 0$.

$q'(z_{0_2}) = \cosh(-\pi i) = \cosh(\pi i) = -1$.

by simple pole residue thm, f has simple poles @ $\pm \pi i$,

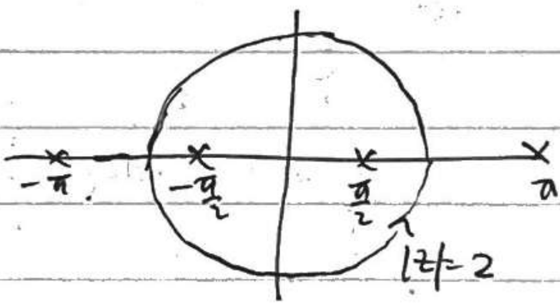
and $\operatorname{Res}_{z=\pi i} f(z) = \dfrac{p(\pi i)}{q'(\pi i)} = \dfrac{e^{\pi i t}}{-1}$

and $\operatorname{Res}_{z=-\pi i} f(z) = \dfrac{p(-\pi i)}{q(-\pi i)} = \dfrac{e^{-\pi i t}}{-1}$

so $\operatorname{Res}_{z=\pi i} f(z) + \operatorname{Res}_{z=-\pi i} f(z) = -2\left(\dfrac{e^{\pi i t} + e^{-\pi i t}}{2}\right) = -2\cos \pi t$

5.) Let C denote the p.o. circle $|z|=2$, and evaluate.

a) $\int_C \tan z \, dz$.



$\cos z$ has zeros @ (only)

$z = (2n+1)\frac{\pi}{2}$, $n \in \mathbb{Z}$

(from §38), thus, only two

enclosed singularities @ $z = \pm\frac{\pi}{2}$.

$f(z) = \dfrac{p(z)}{q(z)} = \dfrac{\sin z}{\cos z}$, ~~...~~

$p = \sin z$, $q = \cos z$, both entire

$p(\frac{\pi}{2}) = \sin(\frac{\pi}{2}) = 1$

$q(\frac{\pi}{2}) = \cos(\frac{\pi}{2}) = 0$. $q'(\frac{\pi}{2}) = -\sin(\frac{\pi}{2}) = -1$.

$p(-\frac{\pi}{2}) = \sin(-\frac{\pi}{2}) = -1$.

$q(-\frac{\pi}{2}) = \cos(-\frac{\pi}{2}) = 0$. $q'(-\frac{\pi}{2}) = -\sin(-\frac{\pi}{2}) = 1$.

by simple pole residue thm, f has simple poles @ $z = \pm\frac{\pi}{2}$,

and $\operatorname{Res}_{z=\frac{\pi}{2}} f(z) = \dfrac{p(\frac{\pi}{2})}{q'(\frac{\pi}{2})} = -1$,

$\operatorname{Res}_{z=-\frac{\pi}{2}} f(z) = \dfrac{p(-\frac{\pi}{2})}{q'(-\frac{\pi}{2})} = -1$,

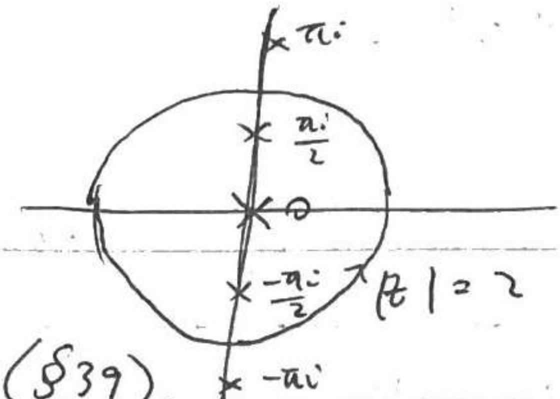
$\int_C f(z)\,dz = 2\pi i(-1 + (-1)) = -4\pi i$

pos cc

Aolc except two interior points

residue thm

§83 #5b, 6, 8.



5b) $\int_C \frac{dz}{\sinh 2z}$

let $f(z)$ be integrand.

$\sinh z$ only has zeros @ $z = n\pi i$ (§39),
so $f(z)$ has ~~poles~~ sing's @ $z = \frac{n\pi i}{2}$, only enclosed singularities are @ $z = 0, \pm\frac{\pi i}{2}$.

$f(z) = \frac{p(z)}{q(z)}$, $p(z) = 1$, $q(z) = \sinh 2z$, p, q entire,
$z_{0_1} = 0$, $z_{0_2} = \frac{\pi i}{2}$, $z_{0_3} = -\frac{\pi i}{2}$, $q'(z) = 2\cosh 2z$.

@ z_{0_1}: $p(z_{0_1}) = 1 \neq 0$,
$q(z_{0_1}) = \sinh 0 = 0$, $q'(z_{0_1}) = 2\cosh(0) = 2$.

@ z_{0_2}: $p(z_{0_2}) = 1 \neq 0$
$q(z_{0_2}) = \sinh(\pi i) = 0$, $q'(z_{0_2}) = 2\cosh \pi i = 2(-1) = -2$

@ z_{0_3}: $p(z_{0_3}) = 1 \neq 0$
$q(z_{0_3}) = \sinh(-\pi i) = 0$, $q'(z_{0_3}) = 2\cosh(-\pi i) = -2$.

by simple pole residue thm, f has simple poles @ these singularities, and: $\operatorname{Res}_{z=0} f(z) = \frac{p(0)}{q'(0)} = \frac{1}{2}$

$\operatorname{Res}_{z=\frac{\pi i}{2}} f(z) = \frac{p(\frac{\pi i}{2})}{q'(\frac{\pi i}{2})} = -\frac{1}{2}$, $\operatorname{Res}_{z=-\frac{\pi i}{2}} = \frac{p(-\frac{\pi i}{2})}{q'(-\frac{\pi i}{2})} = -\frac{1}{2}$.

$$\int_C \underbrace{f(z)}_{\text{AOLC except at 3 enclosed points}} dz \overset{\text{residue thm}}{=} 2\pi i \left(\cancel{\tfrac{1}{2} - \tfrac{1}{2}} - \tfrac{1}{2}\right) = -\pi i.$$

C: POSCC

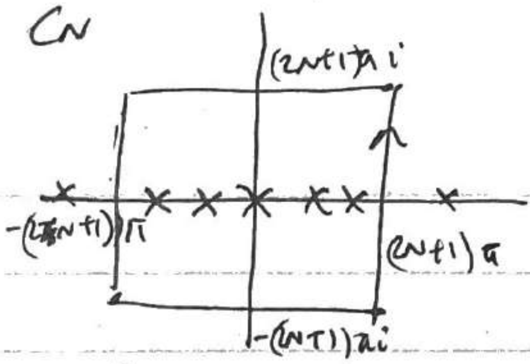
6.) Let C_N denote the P.O. boundary of the square whose edges lie along the lines $x = \pm\left(N + \frac{1}{2}\right)\pi$, $y = \pm\left(N + \frac{1}{2}\right)\pi$,

where N is a positive integer. Show that

$$\int_{C_N} \frac{dz}{z^2 \sin z} = 2\pi i \left[\frac{1}{6} + 2\sum_{n=1}^{N} \frac{(-1)^n}{n^2\pi^2}\right]$$

and then show that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}$ in the limit as $N \to \infty$.

§83 #6, 8.



case when $N=2$.

6, cont'd. let integrand $= f(z)$.

denominator $= z^2 \sin z$.

z^2 has a zero @ $z=0$

$\sin z$ has zeros @ $z = n\pi$ (§38).

Thus f has $(2n+1)$ enclosed zeros

at $-N\pi, -(N-1)\pi, \ldots, (N-1)\pi, N\pi$.

$$f(z) = \frac{p(z)}{q(z)}, \quad p(z) = 1, \; q(z) = z^2 \sin z, \; p, q \text{ entire}.$$

$$p(n\pi) = 1 \neq 0, \quad q(n\pi) = (n\pi)^2 \sin(n\pi) = (n\pi)^2(0) = 0,$$

$$q'(n\pi) = 2(n\pi)\sin(n\pi) + (n\pi)^2 \cos n\pi$$
$$= \cancel{2(n\pi)(0)} + (n\pi)^2(-1)^n \neq 0 \text{ if } n \neq 0.$$

simple pole residue thm applies for nonzero singularities, at which

$$\operatorname*{Res}_{\substack{z=n\pi,\\ n\neq 0}} f(z) = \frac{p(n\pi)}{q'(n\pi)} = \frac{1}{(n\pi)^2(-1)^n}.$$

for sing @ $z=0$: use long div.

$$f(z) = \frac{1}{z^2 \sin z} = \frac{1}{z^2\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots\right)}$$

$$= \frac{1}{z^3 - \frac{z^5}{3!} + \frac{z^7}{5!} - \cdots}$$

$$\begin{array}{l} z^{-3} + \frac{1}{3!} z^{-1} + \cdots \\ \hline 1 \\ 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \cdots \\ \hline \frac{z^2}{3!} - \frac{z^4}{5!} - \cdots \\ \frac{z^2}{3!} - \cdots \\ \hline \end{array}$$

$$\operatorname*{Res}_{z=0} f(z) = b_1 = \frac{1}{6}.$$

$$\int_C \underbrace{f(z)}_{\text{ADIC except @ finitely many interior pts}} dz \overset{\text{residue thm}}{=} 2\pi i\left[\frac{1}{6} + \sum_{\substack{n=-N\\ n\neq 0}}^{N} \frac{1}{n^2\pi^2(-1)^n}\right]$$

POSCC

$$= 2\pi i\left[\frac{1}{6} + 2\sum_{n=1}^{N} \frac{(-1)^n}{n^2\pi^2}\right].$$

§83 #6, 8.

6, cont'd.) From §47 #8, $\left|\int_{C_N} \frac{dz}{z^2 \sin z}\right| \le \frac{16}{(2N+1)\pi A}$,

so $\lim_{N\to\infty} \left|\int_{C_N} \frac{dz}{z^2 \sin z}\right| \le \lim_{N\to\infty} \frac{16}{(2N+1)\pi A} = 0$

$\Rightarrow \lim_{N\to\infty} \int_{C_N} \frac{dz}{z^2 \sin z} = 0$.

Thus $\lim_{N\to\infty} \left[\int_{C_N} f(z)\,dz\right] = 0 = \lim_{N\to\infty}\left[2\pi i\left(\frac{1}{6} + 2\sum_{n=1}^{N} \frac{(-1)^n}{n^2\pi^2}\right)\right]$

$\Rightarrow -\frac{1}{6} = \frac{2}{\pi^2}\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \Rightarrow \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12}$.

8.) Consider the function $f(z) = \frac{1}{(q(z))^2}$ where q is analytic at z_0, and $q(z_0) = 0$, $q'(z_0) \neq 0$. Show that z_0 is a pole of order $m=2$ of f, with residue $B_0 = -\frac{q''(z_0)}{(q'(z_0))^3}$

By definition, z_0 is a pole of order 1 of q, so
$q = (z-z_0)g(z)$, where g is AANZ @ z_0.
Thus $f(z) = \frac{\phi(z)}{(z-z_0)^2}$, where $\phi(z) = \frac{1}{(g(z))^2}$, and

$\phi(z)$ is clearly AANZ @ z_0. (since g is AANZ @ z_0).
Thus by the pole-residue thm, f has a pole of order 2 @ z_0, and $\operatorname{Res}_{z=z_0} f(z) = \frac{\phi'(z_0)}{1!} = \frac{-2g'(z_0)}{(g(z_0))^3}$. (*)

$q'(z) = \frac{d}{dz}((z-z_0)g(z)) = g(z) + (z-z_0)g'(z)$.

(**) $q'(z_0) = g(z_0) + \cancel{(z_0-z_0)g'(z_0)}$

$q''(z) = \frac{d}{dz}\left(g(z) + (z-z_0)g'(z)\right) = g'(z) + g'(z) + (z-z_0)g''(z)$

(***) $q''(z_0) = 2g'(z_0) + \cancel{(z_0-z_0)g''(z_0)}$

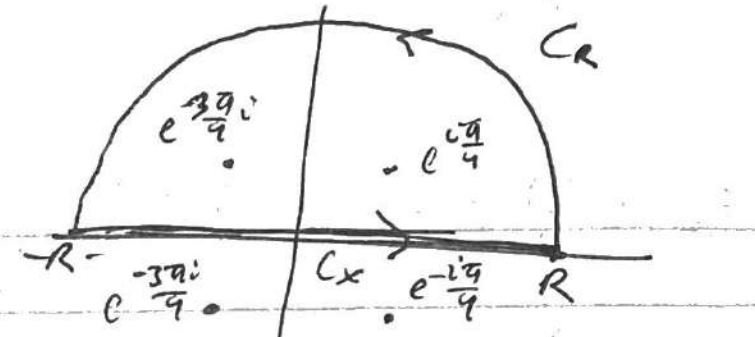
Substituting (**) and (***) into (*), we get the desired result.

$B_0 = \frac{-2g'(z_0)}{(g(z_0))^3} = \frac{-q''(z_0)}{(q(z_0))^3}$.

§86 #3, 6, 9.

Show:

3). $\displaystyle\int_0^{\infty} \frac{dx}{x^4+1} = \frac{\pi}{2\sqrt{2}}$.



let $f(z) = \dfrac{1}{z^4+1}$.

f has four singularities @ $z = (-1)^{\frac{1}{4}}$, two lie above the x-axis - enclose them with a semicircular path such as that on the right.

let C_x be the flat part of the contour, C_R be the curved part, $C = C_x + C_R$ is the SCC of integration, $(R > 1)$.

$$\underset{\substack{\uparrow \\ \text{SCC}}}{\int_C} \underbrace{f(z)}_{\substack{\uparrow \\ \text{ACIC except for} \\ \text{two internal points}}} dz \underset{\substack{\uparrow \\ \text{residue} \\ \text{thm}}}{=} 2\pi i \sum_{\substack{\text{enclosed} \\ \text{sing's}}} \operatorname{Res} f(z) = \int_{-R}^{R} \frac{dx}{x^4+1} + \int_{C_R} f(z)\,dz \quad (*)$$

$f(z) = \dfrac{p(z)}{q(z)}$, $p = 1$, $q(z) = z^4+1$, p, q entire.
let $z_{01} = e^{\frac{i\pi}{4}}$, $z_{02} = e^{\frac{3i\pi}{4}}$, $q'(z) = 4z^3$.

$p(z_{01}) = p(z_{02}) = 1$.
$q(z_{01}) = q(z_{02}) = 0$.
$q'(z_{01}) = 4z_{01}^3$, $q'(z_{02}) = 4z_{02}^3$.

By simple pole-residue thm, f has simple poles @ $z = z_{01}, z_{02}$,
and $\displaystyle\operatorname*{Res}_{z=z_{01}} f(z) = \frac{p(z_{01})}{q'(z_{01})} = \frac{1}{4z_{01}^3} = \frac{z_{01}}{4z_{01}^4} = -\frac{1}{4} z_{01}$

Similarly, $\displaystyle\operatorname*{Res}_{z=z_{02}} f(z) = -\frac{1}{4} z_{02}$.

and $\displaystyle\sum_{\substack{\text{enclosed} \\ \text{sing's}}} \operatorname{Res} f(z) = -\frac{1}{4}(z_{01}+z_{02}) = -\frac{1}{4}\left(\cancel{\tfrac{1}{\sqrt{2}}} + \tfrac{1}{\sqrt{2}} i + \cancel{\left(-\tfrac{1}{\sqrt{2}}\right)} + \tfrac{1}{\sqrt{2}} i\right)$

$(**)$ $= -\dfrac{1}{4}\left(\dfrac{2}{\sqrt{2}} i\right) = -\dfrac{i}{2\sqrt{2}}$

For any point on C_R, $|f(z)| = \left|\dfrac{1}{z^4+1}\right| = \dfrac{1}{|z^4+1|} \underset{\substack{\uparrow \\ \Delta 2.}}{\leq} \dfrac{1}{|z^4| - |1|}$

$= \dfrac{1}{|z|^4 - 1} = \dfrac{1}{R^4-1} = M_R$.

§8.6 #3, 6, 9.

3, cont'd). length of $C_R = \pi R, = L_R$

By ML-inequality, $\left| \int_{C_R} f(z)\,dz \right| \leq M_R L_R = \pi R \cdot \frac{1}{R^4 - 1}$

$$\lim_{R\to\infty} \left| \int_{C_R} f(z)\,dz \right| \leq \lim_{R\to\infty} \frac{\pi R}{R^4 - 1} = 0.$$

$$\Rightarrow \lim_{R\to\infty} \int_{C_R} f(z)\,dz = 0. \quad ***$$

Taking the limit of * as $R \to \infty$, and substituting in ** and ***

$$\lim_{R\to\infty} \int_{-R}^{R} \frac{dx}{x^4+1} = 2\pi i \left(-\frac{i}{2\sqrt{2}} \right) - \cancel{\lim_{R\to\infty} \int_{C_R} f(z)\,dz}^{\,0}$$

$$\int_{-\infty}^{\infty} \frac{dx}{x^4+1} = \frac{\pi}{\sqrt{2}}.$$

Since $\frac{1}{x^4+1}$ is even (quotient of two even polynomials),

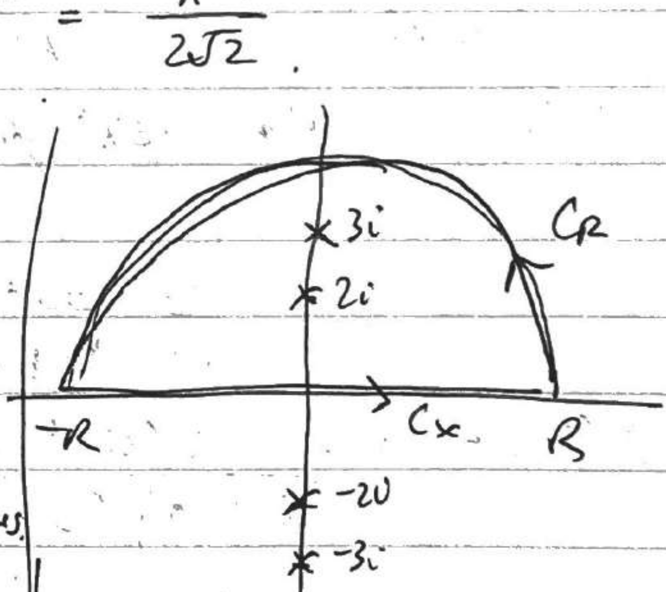
$$\int_0^{\infty} \frac{dx}{x^4+1} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{dx}{x^4+1} = \frac{\pi}{2\sqrt{2}}.$$

6). $\displaystyle \int_0^{\infty} \frac{x^2\,dx}{(x^2+9)(x^2+4)^2} = \frac{\pi}{200},$

$$f(z) = \frac{z^2}{(z^2+9)(z^2+4)^2}.$$

f has singularities @ $z = \pm 3i, \pm 2i$.

C encloses only the positive-i values.



$$\underbrace{\int_C}_{\text{PosCC}} \underbrace{f(z)\,dz}_{\text{AOIC except @ 2 enclosed singular pts}} = \underbrace{2\pi i}_{\text{residue thm}} \underbrace{\sum}_{\text{enclosed sing's.}} \operatorname{Res} f(z)$$

again, $C = C_x + C_R$, $R > 3$

$$= \int_{-R}^{R} \frac{x^2\,dx}{(x^2+9)(x^2+4)^2} + \int_{C_R} f(z)\,dz. \quad *$$

§8.6 #6, 9.

6, cont'd.) Singularity @ $z = 3i$:

$$f = \frac{\dfrac{z^2}{(z+3i)(z^2+4)^2}}{z-3i} = \frac{\phi(z)}{z-3i}, \quad \phi(z) = \frac{z^2}{(z+3i)(z^2+4)^2}.$$

$$\phi(3i) = \frac{(3i)^2}{(3i+3i)(3i^2+4)^2} = \frac{3i}{2\cancel{(3i)}(-9+4)^2} = \frac{3i}{50} \neq 0.$$

$\Downarrow$

ϕ AANZ @ $z = 3i$.

By pole-residue thm, f has simple pole @ $z = 3i$,

and $\operatorname{Res}_{z=3i} f(z) = \phi(3i) = \frac{3i}{50}$.

Singularity @ $z = 2i$.

$$f = \frac{\dfrac{z^2}{(z^2+9)(z+2i)^2}}{(z-2i)^2} = \frac{\phi(z)}{(z-2i)^2}, \quad \phi(z) = \frac{z^2}{z^4+4iz^3+5z^2+36iz-36}$$

$\downarrow$

AANZ @ $z = 2i$.

By pole-residue thm, f has double pole @ $z = 2i$,

and $\operatorname{Res}_{z=2i} f(z) = \frac{\phi'(2i)}{1!} = \phi'(2i)$.

$$\phi'(z) = \frac{(z^4+4iz^3+5z^2+36iz-36)(2z) - z^2(4z^3+12iz+10z+36i)}{\left((z^2+9)(z+2i)^2\right)^2}$$

$$= \frac{-2z^5-4iz^4+36iz^2-72z}{\left((z^2+9)(z+2i)^2\right)^2}$$

$$\phi'(2i) = \frac{-2(2i)^5-4i(2i)^4+36i(2i)^2-72(2i)}{\left(((2i)^2+9)(2i+2i)^2\right)^2}$$

$= \ldots$ ~~algebra~~ arithmetic $\ldots = \frac{-13i}{200} = \operatorname{Res}_{z=2i} f(z)$.

$$\sum_{\text{sing's}} \operatorname{Res} f(z) = \frac{3i}{50} - \frac{13i}{200} = \frac{-i}{200} \quad (**).$$

On C_R: $|f(z)| = \left|\frac{z^2}{(z^2+9)(z^2+4)^2}\right| = \frac{|z|^2}{|z^2+9|\,|z^2+4|^2}$

$$\leq \frac{R^2}{(R^2-9)(R^2-4)^2} = M.$$

$\uparrow$ Δ_2 twice

length of $C_R = \pi R = L$.

§86 # 6, 9

6, cont'd.) By ML inequality: $\left|\int_{C_R} f(z)\,dz\right| \le ML = \dfrac{\pi R^3}{(R^2-9)(R^2-4)^2}$

$$\lim_{R\to\infty}\left|\int_{C_R} f(z)\,dz\right| = \lim_{R\to\infty}\frac{\pi R^3}{(R^2-9)(R^2-4)^2} = 0.$$

$$\Rightarrow \lim_{R\to\infty}\int_{C_R} f(z)\,dz = 0. \quad (\ast\ast\ast).$$

Take limit of $(\ast)$ as $R\to\infty$ and substitute in $(\ast\ast)$ and $(\ast\ast\ast)$.

$$\lim_{R\to\infty}\int_{-R}^{R}\frac{x^2\,dx}{(x^2+9)(x^2+4)^2} = 2\pi i\left(\frac{-i}{200}\right) - \cancel{\lim_{R\to\infty}\int_{C_R} f(z)\,dz}^{\,0}$$

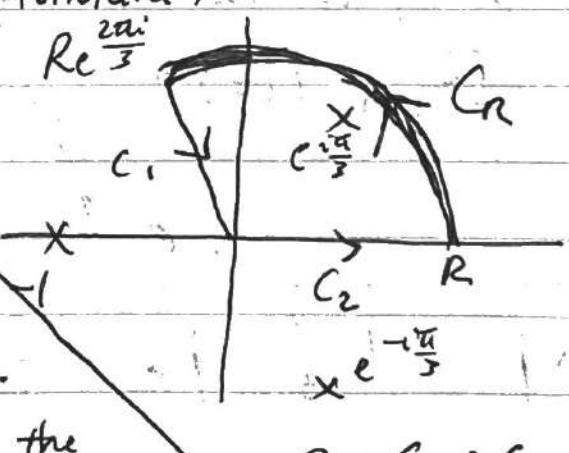
$$\int_{-\infty}^{\infty}\frac{x^2\,dx}{(x^2+9)(x^2+4)^2} = \frac{\pi}{100}.$$

Since integrand is even (quotient of two even polynomials):

$$\int_0^{\infty}\frac{x^2\,dx}{(x^2+9)(x^2+4)^2} = \frac{1}{2}\int_{-\infty}^{\infty}\frac{x^2\,dx}{(x^2+9)(x^2+4)^2} = \frac{\pi}{200}.$$

9) Use a residue and the contour below, where $R>1$, to establish the integration formula:

$$\int_0^{\infty}\frac{dx}{x^3+1} = \frac{2\pi}{3\sqrt{3}}.$$



$$f(z) = \frac{1}{z^3+1}$$

has 3 singularities: $z = (-1)^{\frac{1}{3}}$.

Only one is enclosed within the loop: $z = e^{\frac{i\pi}{3}}$. As before:

$C = C_1 + C_2 + C_R$, $R > 1$

$$\underset{\substack{\uparrow\\ \text{POSCC}}}{\int_{C}} \underbrace{f(z)}_{\substack{\uparrow\\ \text{AOIC except}\\ \text{at one singularity}\\ \text{inside } C}}\,dz \underset{\substack{\uparrow\\ \text{residue}\\ \text{thm}}}{=} 2\pi i \sum_{\substack{\text{enclosed}\\ \text{sing.}}} \operatorname{Res} f(z)$$

$$= \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz + \int_{C_R} f(z)\,dz. \quad (\ast)$$

§86 #9.

9, cont'd.) $C_1: z(r) = re^{\frac{2\pi i}{3}}, \quad 0 \le r \le R$

$z'(r) = e^{\frac{2\pi i}{3}}\,dr.$

$$\int_{C_1} f(z)\,dz = \int_R^0 \frac{1}{(re^{\frac{2\pi i}{3}})^3+1}\, e^{\frac{2\pi i}{3}}\,dr$$

$$= -e^{\frac{2\pi i}{3}}\int_0^R \frac{1}{r^3\underbrace{(e^{2\pi i})}_{=1}+1}\,dr = -e^{\frac{2\pi i}{3}}\int_0^R \frac{dx}{x^3+1}.$$

Also, it is clear that:

$$\int_{C_2} f(z)\,dz = \int_0^R \frac{dx}{x^3+1}, \quad \text{so:}$$

$$\int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz = \left(1-e^{\frac{2\pi i}{3}}\right)\int_0^R \frac{dx}{x^3+1}. \quad (**)$$

We have to find the residue @ $z = e^{\frac{i\pi}{3}}$.

$f(z) = \frac{p(z)}{q(z)}$, $p(z) = 1$, $q(z) = z^3+1$, p, z ative,

$p(e^{\frac{i\pi}{3}}) = 1$, $q(e^{\frac{i\pi}{3}}) = -1+1 = 0$, $q'(e^{\frac{i\pi}{3}}) = 3(e^{\frac{i\pi}{3}})^2$

By simple-pole-residue thm, f has a simple pole @ $z = e^{\frac{i\pi}{3}}$,

and $$\operatorname*{Res}_{z=e^{\frac{i\pi}{3}}} f(z) = \frac{p(e^{\frac{i\pi}{3}})}{q'(e^{\frac{i\pi}{3}})} = \frac{1}{3(e^{\frac{i\pi}{3}})^2} = \frac{e^{\frac{i\pi}{3}}}{3(e^{\frac{i\pi}{3}})^3}$$

$$= -\frac{1}{3}e^{\frac{i\pi}{3}}. \quad (***)$$

on C_R:

$$|f(z)| = \left|\frac{1}{z^3+1}\right| = \frac{1}{|z^3+1|} \underset{\Delta_2}{\le} \frac{1}{|z|^3-1} = \frac{1}{R^3-1} = M$$

length of $C_R = \frac{2}{3}\pi R$. By ML-inequality:

$$\left|\int_{C_R} f(z)\,dz\right| \le ML = \frac{\frac{2}{3}\pi R}{R^3-1}$$

$$\lim_{R\to\infty}\left|\int_{C_R} f(z)\,dz\right| \le \lim_{R\to\infty}\frac{\frac{2}{3}\pi R}{R^3-1} = 0.$$

$$\Rightarrow \lim_{R\to\infty}\int_{C_R} f(z)\,dz = 0. \quad (****)$$

§86 #9.

9, cont'd.)

Combining ⊛, ⊛⊛, ⊛⊛⊛, and ⊛⊛⊛⊛ and taking the limit as $R \to \infty$:

$$\lim_{R\to\infty} \frac{\left(1-e^{\frac{2\pi i}{3}}\right)\int_0^R \frac{dx}{x^3+1}}{1-e^{\frac{2\pi i}{3}}} = \frac{2\pi i\left(-\frac{1}{3}e^{\frac{i\pi}{3}}\right)}{1-e^{\frac{2\pi i}{3}}} - \cancel{\lim_{R\to\infty}\int_{C_R} f(z)\,dz} \to 0$$

$$\int_0^\infty \frac{dx}{x^3+1} = \frac{-\frac{2}{3}\pi i\, e^{\frac{i\pi}{3}}}{1-e^{\frac{2\pi i}{3}}} \cdot \left(\frac{1-e^{-\frac{2\pi i}{3}}}{1-e^{-\frac{2\pi i}{3}}}\right)$$

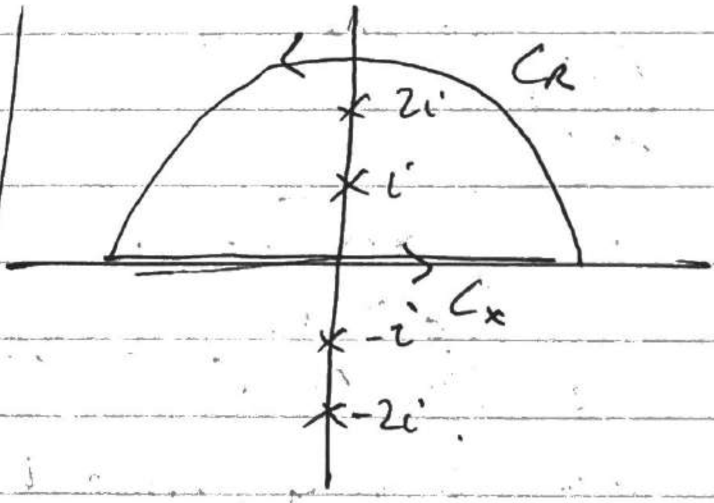
$$= \frac{-\frac{2}{3}\pi i\left(e^{\frac{i\pi}{3}} - e^{-\frac{i\pi}{3}}\right)}{\left(1-\left(-\frac{1}{2}\right)\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2}$$

$$= \frac{\frac{4\pi}{3}\left(\frac{e^{\frac{i\pi}{3}} - e^{-\frac{i\pi}{3}}}{2i}\right)}{3} = \frac{4\pi}{9}\sin\frac{\pi}{3} = \frac{\cancel{4}^{2}\pi}{\cancel{9}_{3\sqrt{3}}}\frac{\cancel{\sqrt{3}}}{\cancel{2}} = \frac{2\pi}{3\sqrt{3}}$$

§88 #6. $\displaystyle\int_{-\infty}^{\infty} \frac{x\sin x\,dx}{(x^2+1)(x^2+4)}$.

let $f(z) = \dfrac{ze^{iz}}{(z^2+1)(z^2+4)}$

two enclosed singularities: @ $z = i, 2i$.



$C = C_x + C_R$

$R > 2$.

As before:

$$\int_C \underbrace{f(z)}_{\text{ADIC except two interior singular pts}}\,dz \overset{\text{residue thm}}{=} 2\pi i \sum_{\text{sing's}} \operatorname{Res} f(z)$$

$\uparrow$ POSCCC

$$= \int_{-R}^{R} \frac{xe^{ix}\,dx}{(x^2+1)(x^2+4)} + \int_{C_R} f(z)\,dz. \quad ⊛$$

§88 #6.

6, cont'd.)

$f(z) = \dfrac{p(z)}{q(z)}$, $\quad p(z) = ze^{iz}$

$q(z) = z^4 + 5z^2 + 4. = (z^2+1)(z^2+4)$.

p, q entire ; $q'(z) = 4z^3 + 10z$.

$z_{0_1} = i, \quad z_{0_2} = 2i$.

$p(i) = ie^{i^2} = ie^{-1} \neq 0$.

$p(2i) = 2ie^{2(i)^2} = 2ie^{-2} \neq 0$.

$q(i) = 0, \quad q(2i) = 0$.

$q'(i) = 6i, \quad q'(2i) = -12i$.

By simple-pole-residue thm, f has simple poles @ $z = i, 2i$,

and $\operatorname*{Res}_{z=i} f(z) = \dfrac{p(i)}{q'(i)} = \dfrac{\cancel{i}e^{-1}}{6\cancel{i}} = \dfrac{1}{6e}$,

$\operatorname*{Res}_{z=2i} f(z) = \dfrac{p(2i)}{q'(2i)} = \dfrac{2\cancel{i}e^{-2}}{-12\cancel{i}} = \dfrac{-1}{6e^2}$

$\sum_{\substack{\text{enclosed}\\ \text{sing's}}} \operatorname{Res} f(z) = \dfrac{1}{6e} - \dfrac{1}{6e^2} = \dfrac{e-1}{6e^2}$ (**).

Plugging (**) into (*) ;

$$\int_{-R}^{R} \frac{xe^{ix}\,dx}{(x^2+1)(x^2+4)} = 2\pi i\left(\frac{e-1}{6e^2}\right) - \int_{C_R} f(z)\,dz$$

taking imaginary parts of both sides :

$$\int_{-R}^{R} \frac{x\sin x\,dx}{(x^2+1)(x^2+4)} = \frac{2\pi}{6e^2}(e-1) - \operatorname{Im}\int_{C_R} f(z)\,dz. \quad (***)$$

$g(z) = \dfrac{z^2}{(z^2+1)(z^2+4)}$ is analytic in the upper half

plane, exterior to the circle $|z| = 2$.

$|g(z)| = \left|\dfrac{z^2}{(z^2+1)(z^2+4)}\right| = \dfrac{|z^2|}{|z^2+1||z^2+4|} \underset{\Delta_2 \text{ and some simplification.}}{\leq} \dfrac{R^2}{(R^2-1)(R^2-4)} = M_R$

§88 #6.

6, cont'd.

$$\lim_{R\to\infty} M_R = \lim_{R\to\infty} \frac{R^2}{(R^2-1)(R^2-4)} = 0.$$

By Jordan's thm: $\lim_{R\to\infty}\int_{C_R} g(z)e^{iz}dz = \lim_{R\to\infty}\int_{C_R} f(z)\,dz = 0.$

$$\Rightarrow \lim_{R\to\infty}\left(\mathrm{Im}\int_{C_R} f(z)\,dz\right) \le \lim_{R\to\infty}\int_{C_R} f(z)\,dz = 0$$

$$\Rightarrow \lim_{R\to\infty}\left[\mathrm{Im}\int_{C_R} f(z)\,dz\right] = 0. \quad ****$$

Plugging in **** into *** and taking the limit as $R\to\infty$:

$$\lim_{R\to\infty}\int_{-R}^{R}\frac{x\sin x\,dx}{(x^2+1)(x^2+4)} = \frac{2\pi}{6e^2}(e-1) - \cancel{\lim_{R\to\infty}\left[\mathrm{Im}\int_{C_R} f(z)\,dz\right]}$$

$$\int_{-\infty}^{\infty}\frac{x\sin x\,dx}{(x^2+1)(x^2+4)} = \frac{\pi(e-1)}{3e^2}.$$

§92 #1.

1). Show that: $\displaystyle\int_0^{2\pi}\frac{d\theta}{5+4\sin\theta} = \frac{2\pi}{3}.$

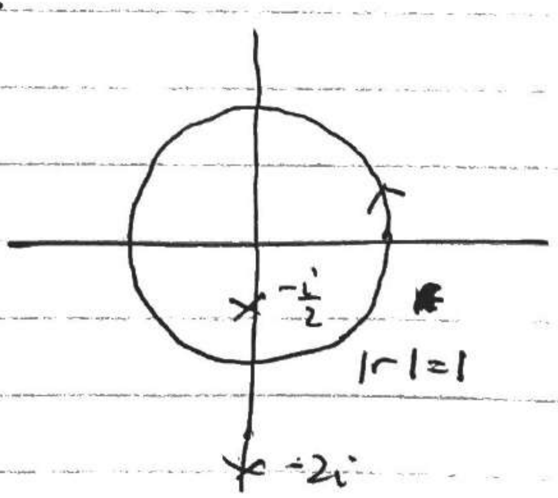
let $\sin\theta = \dfrac{z-z^{-1}}{2i}$, $z = e^{i\theta}$, $d\theta = \dfrac{dz}{iz}$, $0\le\theta<2\pi$

C = unit circle, P.O.

$$\int_0^{2\pi}\frac{d\theta}{5+4\sin\theta} = \int_C \frac{dz}{iz\left(5+4\left(\frac{z-z^{-1}}{2i}\right)\right)}$$

$$= \int_C \frac{dz}{5iz+2z^2-2}$$

$$= \int_C \frac{2\,dz}{\left(z+\frac{i}{2}\right)(z+2i)}$$



let $f(z)$ be the integrand.

f has only one singularity enclosed within C, $z = -\frac{i}{2}$.

§92 #1

1, cont'd

$$f(z) = \frac{p(z)}{q(z)}, \quad \left.\begin{aligned} p(z) &= 1, \\ q(z) &= 2z^2 + 5iz - 2 \end{aligned}\right\} \text{entire functions}$$

$p\left(-\frac{i}{2}\right) = 1 \neq 0$

$q\left(-\frac{i}{2}\right) = 0$

$q'\left(-\frac{i}{2}\right) = 3i.$

$q'(z) = 4z + 5i$

By simple pole residue thm, $\operatorname*{Res}\limits_{z=-\frac{i}{2}} f(z) = \dfrac{p\left(-\frac{i}{2}\right)}{q'\left(-\frac{i}{2}\right)} = \dfrac{1}{3i}$

residue thm

$$\underset{\text{pos cc}}{\int_C} \underbrace{f(z)}_{\text{AOLC except at one interior singular pt}} dz = 2\pi i \sum_{\text{sing's}} \operatorname*{Res}\limits_{z_k} f(z) = 2\pi \cancel{i}\left(\frac{1}{3\cancel{i}}\right) = \frac{2\pi}{3}.$$

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20.

§94 #5, 7a.

5) Suppose that a function f is analytic inside and on a POPCC C and has no zeros on C. Show that if f has n zeros z_k $(k=1,2,\dots,n)$ inside C, where each z_k is of multiplicity m_k, then:

$$\int_C \frac{z f'(z)}{f(z)}\,dz = 2\pi i \sum_{k=1}^{n} m_k z_k$$

By the winding # thm, $\frac{1}{2\pi}\Delta_C \arg f(z) = Z - P$,

where Z is the number of enclosed zeros and P the number of enclosed poles (counted by multiplicity). Since f is AOIC, $P=0$ and in the context of this problem, $Z = \sum_{k=1}^{n} m_k$.

(by Thm 1, §82)

At $z=z_k$ $(k=1,2,\dots,n)$, $f(z) = (z-z_k)^{m_k} g(z)$,
where g is a function that is AANZ @ z_k.

Thus $f'(z) = m_k (z-z_k)^{m_k-1} g(z) + (z-z_k)^{m_k} g'(z)$.

Let h = integrand of contour integral in question. Then

$$h(z) = \frac{z\left(m_k(z-z_k)^{m_k-1}g(z) + (z-z_k)^{m_k} g'(z)\right)}{(z-z_k)^{m_k} g'(z)}$$

$$= \underbrace{\overbrace{z_k}^{\text{Residue}}\left(m_k(z-z_k)^{-1}\right.}_{\substack{\text{pseudo power series}\\ \text{centered at } z=z_k}} + \left.\underbrace{\frac{g'(z_k)}{g(z_k)}}_{\substack{\text{power series}\\ \text{centered @ } z=z_k}}\right)$$

This is clearly a Laurent series centered @ $z=z_k$, where the first term is the pseudo-power series and the second term is analytic @ $z=z_k$ (since g is AANZ @ z_k) and thus forms an ordinary power series @ z_k.

Thus $\underset{z=z_k}{\text{Res}}\ h(z) = b_1 = m_k z_k$.

Residue thm.

$$\int_C \underbrace{h(z)}\, dz \overset{\downarrow}{=} 2\pi i \sum_{k=1}^{n} \underset{z=z_k}{\text{Res}}\ h(z)$$

$$= 2\pi i \sum_{k=1}^{n} m_k z_k$$

POSCC

AOLC except for (finitely many) poles inside C when $f(z)=0$ (i.e., not analytic @ $z=z_k$, $k=1,2,\ldots n$)

7a). Determine the number of zeros, counting multiplicities, of $z^4 - 2z^3 + 9z^2 + z - 1$ inside the circle $|z|=2$.

Let $f(z) = 9z^2$,

$g(z) = z^4 - 2z^3 + z - 1$.

On C: ($|z|=2$):

$|f(z)| = |9z^2| = 9|z|^2 = 36$.

$|g(z)| = |z^4 - 2z^3 + z - 1| \leq |z^4| + |-2z^3| + |z| + |-1|$

extended Δ_1

$= |z|^4 + 2|z|^3 + |z| + 1$

$= 2^4 + 2(2)^3 + 2 + 1 = 16 + 16 + 2 + 1 = 35$

Since $|g(z)| \leq 35 < 36 = |f(z)|$ on C, and f, g are entire (and thus AOLC). Thus we can apply Rouche's Thm: f and $f+g$ have the same number of zeros inside C (counting multiplicity). Clearly f only has a zero of multiplicity 2 inside C, so $f+g$ (the function of interest) must also have 2 zeros inside C (counted by multiplicity).

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20.

§98 #3.

3) Show that the image of the half-plane $y > c_2$ under the transformation $w = \frac{1}{z}$ is the interior of a circle when $c_2 > 0$. Find the image when $c_2 < 0$ and when $c_2 = 0$.

Use the result from §98.

For a line or circle in the xy plane,

$A(x^2+y^2) + Bx + Cy + D = 0 \quad (B^2+C^2 > 4AD)$

the transformation $w = \frac{1}{z}$ maps this line or circle into the line or circle $D(u^2+v^2) + Bu - Cv + A = 0$ in the uv plane.

Consider the images of horizontal lines in the xy plane, $y = c_0$.

Case I : $c_0 \neq 0$.

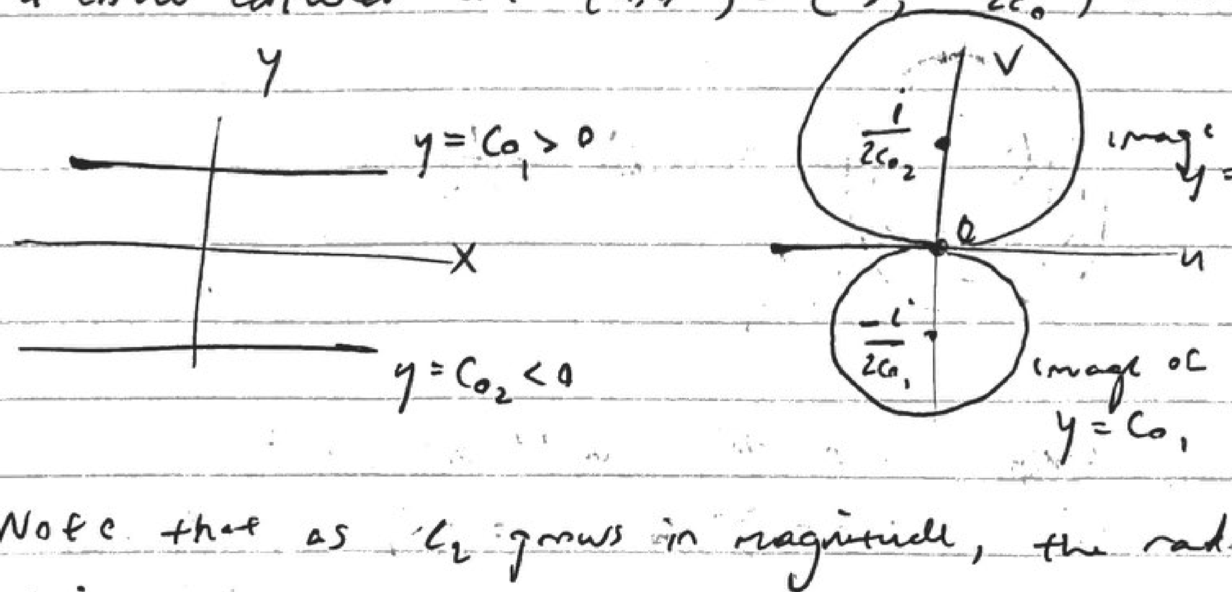
$\underbrace{A = 0,\ B = 0,\ C = 1,\ D = -c_0}_{\text{mapped to}}, \ (B^2+C^2 = 1 > 0 = 4AD)$

$-c_0(u^2+v^2) - v = 0$

$u^2 + v^2 + \frac{v}{c} + \left(\frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2$

$u^2 + \left(v + \frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2.$

(a circle centered at $(u,v) = \left(0, -\frac{1}{2c_0}\right)$ with radius $\frac{1}{2c_0}$)



Note that as c_2 grows in magnitude, the radius shrinks.

Case 2: $c_0 = 0$.

$\underbrace{A = 0,\ B = 0,\ C = 1,\ D = 0}$ $\quad (B^2 + C^2 = 1 > 0 = 4AD)$

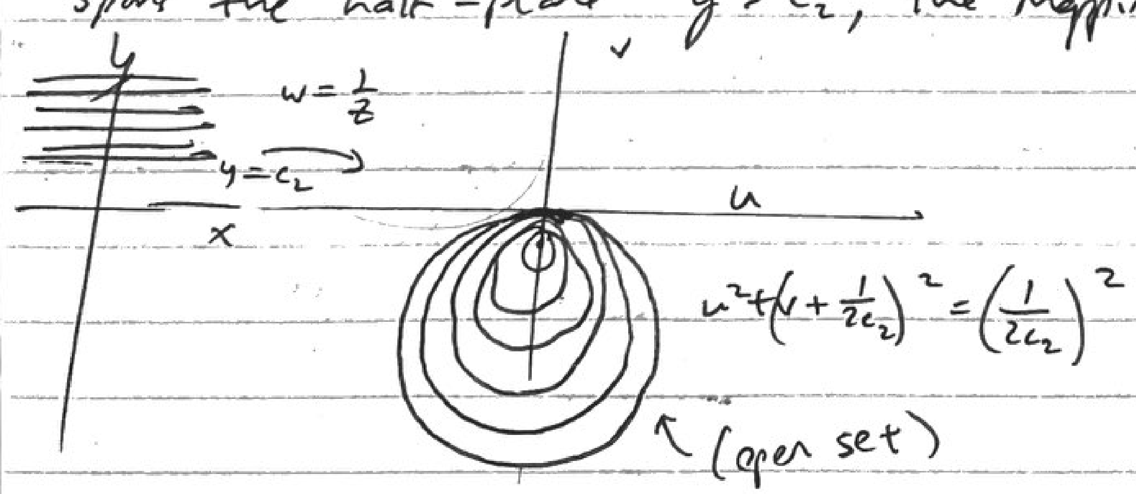
gets mapped to

$-v = 0$

Thus $y = 0$ gets mapped to $v = 0$ under this transformation.

$y = c_0 > c_2 > 0$ is the set of circles with maximum radius $\frac{1}{2c_2}$. As $c_2 \to \infty$, their centers move toward the origin but their radii get smaller so that each circle $u^2 + \left(v + \frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2$ is contained within the outermost circle, $u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$.

Also, this set of circles spans the outermost circle; for any point (u_0, v_0) in the outer circle $u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$, you can find a circle $u^2 + \left(v + \frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2$, $c_0 > c_2$ that passes through this point, so this ~~span~~ set of circles spans the disk $u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$. Since the preimage of this set is the ~~entire~~ set $y = c_0 > c_2$, which spans the half-plane $y > c_2$, the mapping is established.

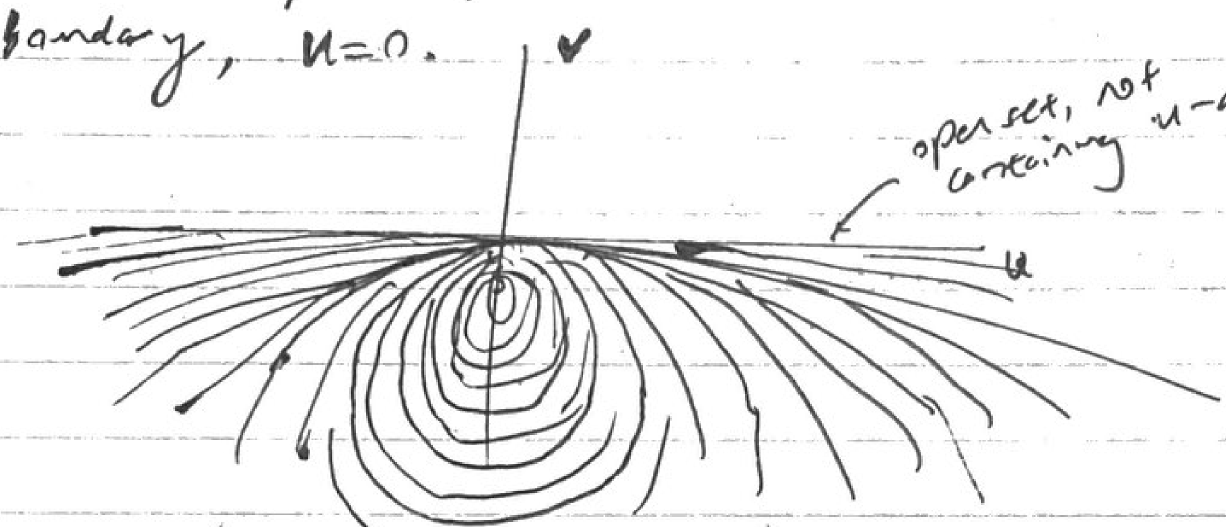


• If $c_2 = 0$, ~~then~~ then this would be the set of all circles $u^2 + \left(v + \frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2$, $c_0 > 0$. (the images of all lines $y = c_0 > 0$). This would be all circles whose top point (where "top" is the maximum v-coordinate point) is the origin, extending out to

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis 3
5/11/20

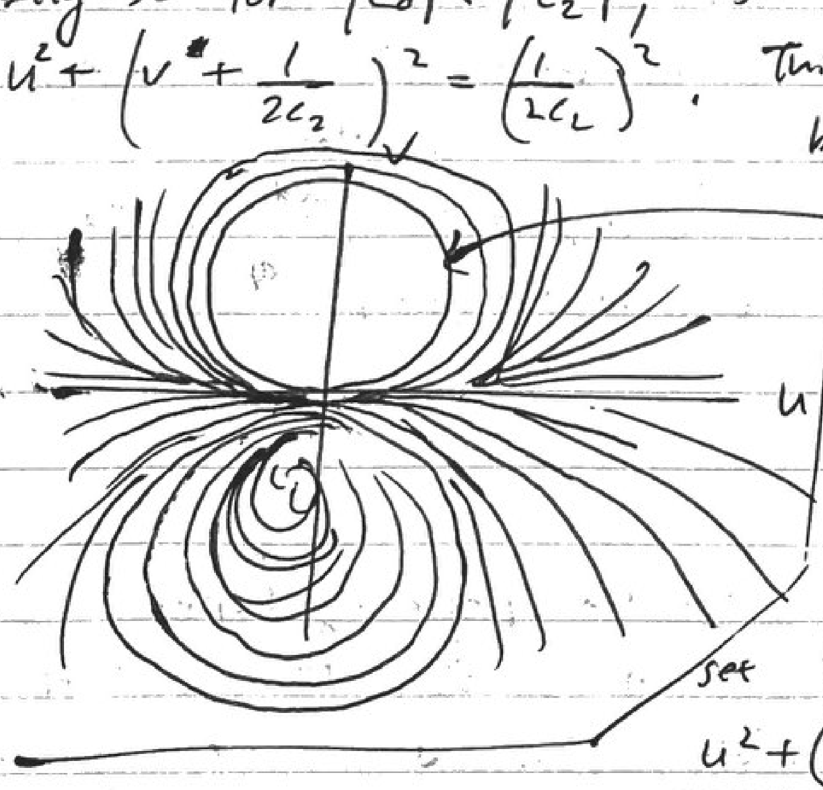
§98 #3, cont'd.

3, cont'd) arbitrarily large radii. This mapping spans the entire LHP of the uv plane, as any point (u_0, v_0), $v_0 < 0$ has some circle of the form $u + \left(v + \frac{1}{2c_0}\right)^2 = \left(\frac{1}{2c_0}\right)^2$ through it. Again, this is an open set, because it doesn't include its boundary, $u = 0$.



(entire LHP of uv plane spanned by $y > 0$ mapped by $w = \frac{1}{z}$).

In the case that $c_2 < 0$, then we ~~include~~ add the v-axis (mapped to by $y = 0$) and some circles in the UHP of the uv plane. These circles would only be for $|c_0| < |c_2|$, so these are circles <u>outside of</u> $u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$. This is the open set illustrated below:



$u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$

By the same arguments as before, this spans the set outside of the circle $u^2 + \left(v + \frac{1}{2c_2}\right)^2 = \left(\frac{1}{2c_2}\right)^2$. Thus the whole image of the set $y > (c_2 < 0)$ is the set $u^2 + \left(v + \frac{1}{2c_2}\right)^2 > \left(\frac{1}{2c_2}\right)^2$.

§100 #2.

2). Find the linear fractional transformation that maps the points $z_1 = -i$, $z_2 = 0$, $z_3 = i$ onto the points $w_1 = -1$, $w_2 = i$, $w_3 = 1$. Into what curve is the imaginary axis $x=0$ transformed?

$w = \dfrac{az+b}{cz+d}$ $(ad-bc \neq 0)$

$\Rightarrow w(cz+d) = az+b.$

$-1\,(c(-i)+d) = a(-i)+b \Rightarrow ci - d = b - ai$ ①

$i\,(c(0)+d) = a(0)+b \Rightarrow id = b.$ ②

$1\,(ci+d) = ai+b \Rightarrow ci+d = b+ai$ ③

①+③ : $2ci = 2b \Rightarrow b = ci \underset{②}{=} id$

③−① : $2d = 2ai \Rightarrow d = ai$

$\Rightarrow$ let $M := c = d = ai = -b. \Rightarrow c = d = M,\ a = -iM,\ b = iM$

$w = \dfrac{az+b}{cz+d} = \dfrac{-Miz+Mi}{Mz+M} = \dfrac{M(-iz+i)}{M(z+1)} = \dfrac{-iz+i}{z+1}.$

Reiterating the argument in §99, we show that the linear fractional transformation of a line or circle in the xy plane must be a line or circle in the uv plane:

when $c \neq 0$ and $w = \dfrac{az+b}{cz+d} = \dfrac{a}{c} + \dfrac{bc-ad}{c}\cdot\dfrac{1}{cz+d}$,

then w is the composition $w_3 \circ w_2 \circ w_1$:

$w_1 = cz+d$

$w_2 = \dfrac{1}{z}$

$w_3 = \dfrac{a}{c} + \dfrac{bc-ad}{c}z$

w_1 and w_3 are linear and thus preserve general shape. w_2 is the inverse map, which was shown in §98 to map lines and circles only to lines and circles. Thus w must map lines and

Jonathan Lam
Prof. Smyth
MA 345
Complex Analysis
5/11/20

§100 #2, cont'd

circles only to lines and circles.

Let $z=iy$ (the line $x=0$).

Then $w = \dfrac{-i(iy)+i}{iy+1} = \dfrac{y+i}{iy+1}$

Thus $|w| = \dfrac{|y+i|}{|iy+1|} = \dfrac{\sqrt{y^2+1}}{\sqrt{y^2+1}} = 1$.

Since mappings of a line in the xy plane must map to a line or a circle, and ~~[illegible]~~ w always on the circle $|w|=1$, then the line $x=0$ must map onto the (entire) circle $|w|=1$.

§102 #1, 2

1). Recall from Ex 1 in §102 that the transformation

$$w = \frac{i-z}{i+z}$$

maps the half-plane $\text{Im } z > 0$ onto the disk $|w|<1$ and the boundary of the half-plane onto the boundary of the disk. Show that a point $z=x$ is mapped onto the point

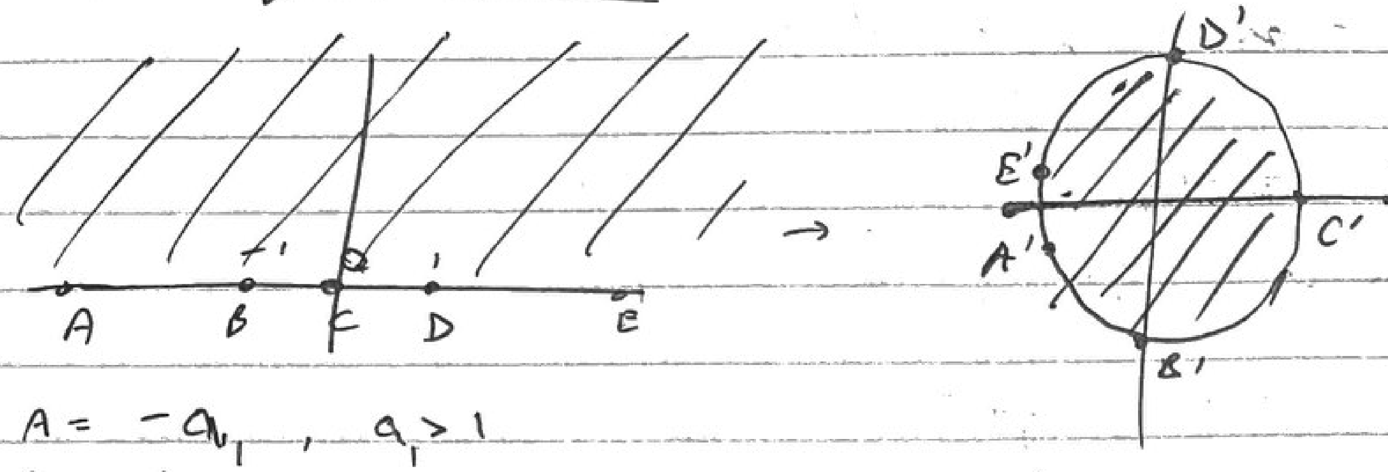
$$w = \frac{1-x^2}{1+x^2} + i\frac{2x}{1+x^2}$$

and then complete the verification of the mapping illustrated in Fig. 13, Appendix 2.

$$w = \frac{i-x}{i+x} = \frac{(i-x)(-i+x)}{(i+x)(-i+x)} = \frac{1+ix+ix-x^2}{x^2+1}$$

$$= \frac{1-x^2}{x^2+1} + i\frac{2x}{x^2+1}.$$

from fig 13, appendix 2:



$A = -a_1, \quad a_1 > 1$

$B = -1$

$C = 0$

$D = 1$

$E = a_2, \quad a_2 > 1$

$$A' = \frac{1-(-a_1)^2}{1+(-a_1)^2} + i\frac{2(-a_1)}{1+(-a_1)^2} = \frac{1-a_1^2}{1+a_1^2} - \frac{ia_1}{1+a_1^2}.$$

$$B' = \frac{1-(-1)^2}{1+(-1)^2} + \frac{i(-2)}{1+(-1)^2} = -i.$$

$$C' = \frac{1-0^2}{1+0^2} + \frac{i(2(0))}{1+0^2} = 1.$$

$$D' = \frac{1-1^2}{1+1^2} + \frac{i(2)(1)}{1+1^2} = i.$$

$$E' = \frac{1-a_2^2}{1+a_2^2} + \frac{i2a_2}{1+a_2^2}.$$

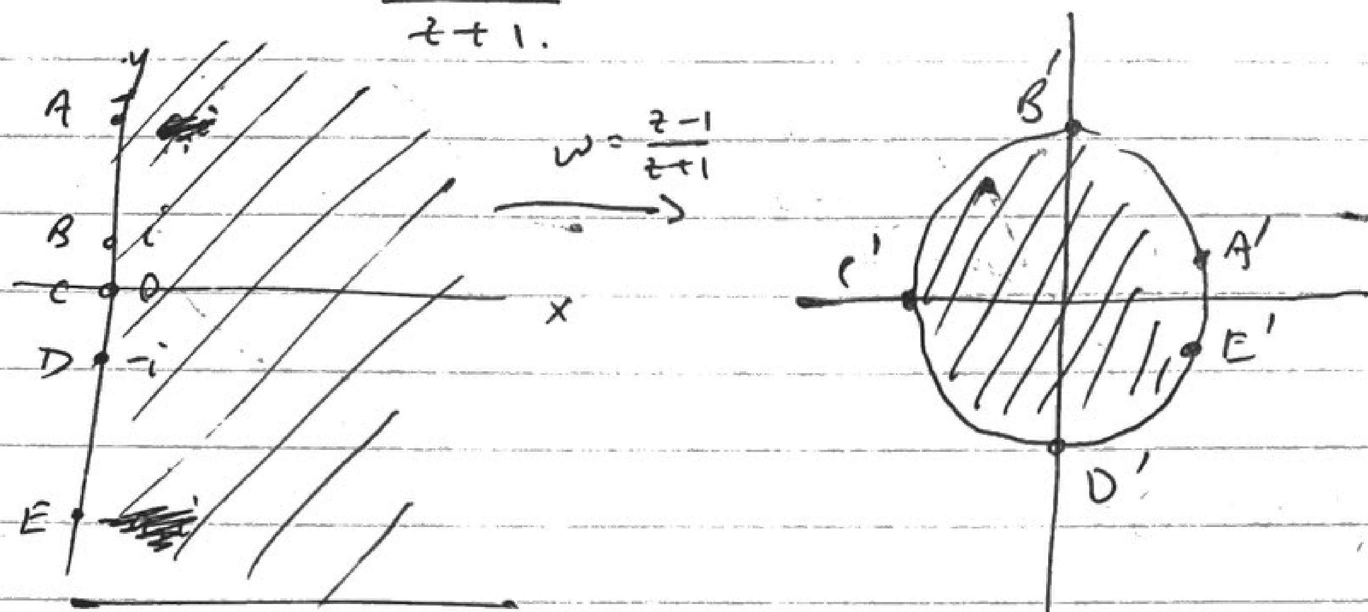
B', C', and D' clearly match the transformed points from the figure.

A' is in the 3rd quadrant, approaching -1 as $a_1 \to \infty$.
C' is in the 2nd quadrant, approaching -1 as $a_2 \to \infty$.
These values also match the figure.

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20.

§102 #2.

2.). Verify the mapping shown in Appendix 2, Fig 13., where $w = \dfrac{z-1}{z+1}$.



$$w = \frac{z-1}{z+1} = -\left(\frac{1-z}{z+1}\right) = -\left(\frac{i-iz}{i+iz}\right) = -\left(\frac{i-(iz)}{i+(iz)}\right)$$

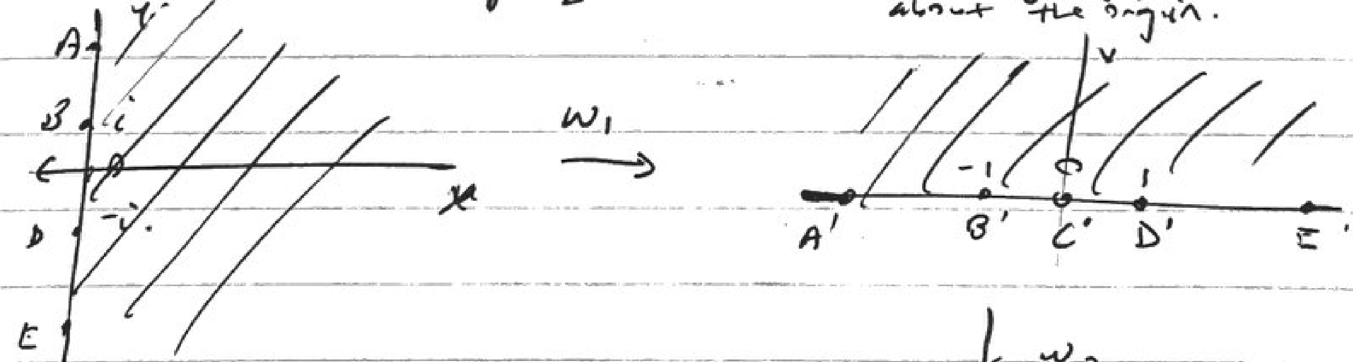
i.e., $w = w_3 \circ w_2 \circ w_1$, $\quad w_1 = iz,$

$$w_2 = \frac{i-z}{i+z},$$

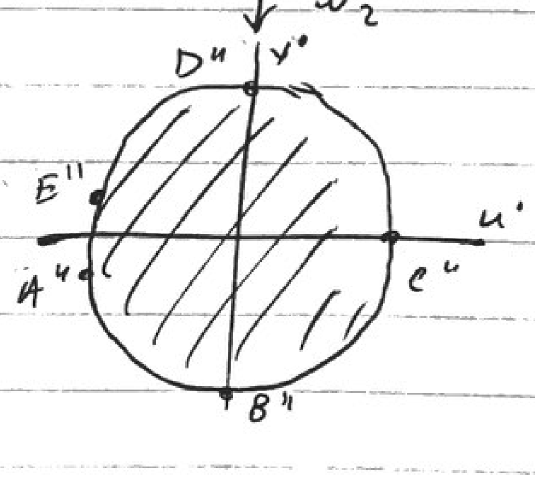
$$w_3 = -z.$$

---

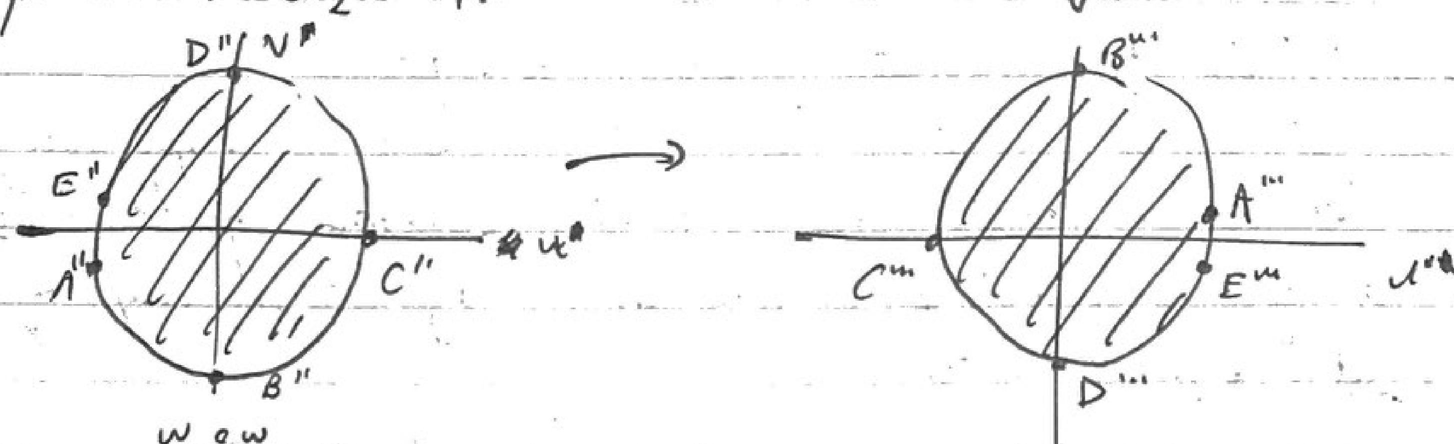
w_1 is a rotation by $\frac{\pi}{2}$ rad about the origin (and scaling by $|i| = 1$):



w_2 is the transformation from the previous exercise:



Lastly, applying $w_3 = -z$ flips the signs of the real and imaginary parts of $w_2 \circ w_1$.



$w_2 \circ w_1$

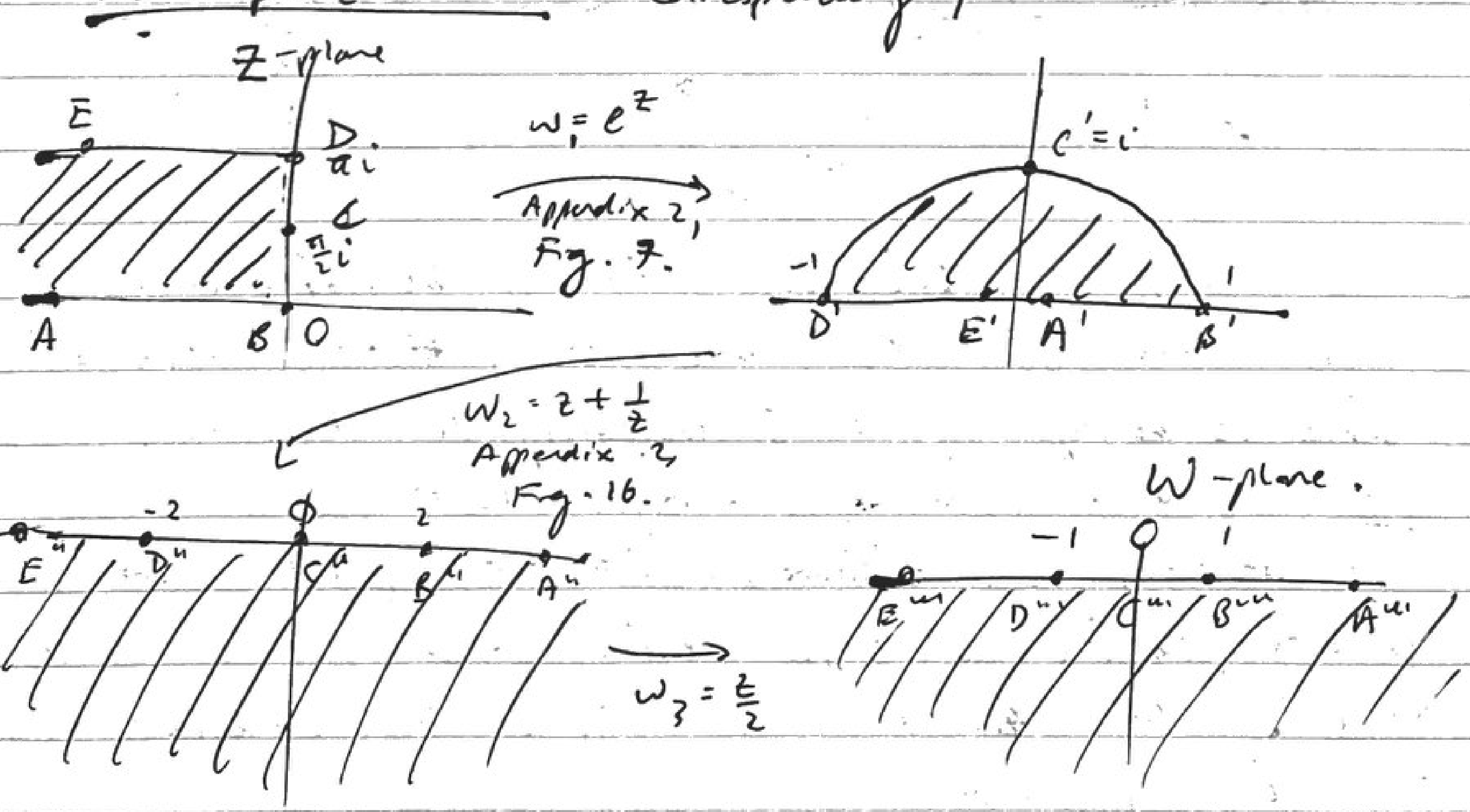
$W = w_3 \circ w_2 \circ w_1$.

§106 #10.

10). Observe that the transformation $w = \cosh z$ can be expressed as a composition of the mappings:

$$W = w_3 \circ w_2 \circ w_1, \quad w_1 = e^z, \quad w_2 = z + \frac{1}{z}, \quad w_3 = \frac{z}{2}.$$

By figures 7 and 16 in Appendix 2, show that when $w = \cosh z$, the semi-infinite strip $x \le 0$, $0 \le y \le \pi$ in the z-plane is mapped to the lower half $v \le 0$ of the w-plane. Indicate corresponding parts of boundaries.



Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20.

§114 # 3, 10.

3.) Show that under the transformation $w = \frac{1}{z}$, the images of the lines $y = x-1$ and $y = 0$ are the circle $u^2+v^2-u-v=0$ and the line $v=0$, respectively. Sketch all four curves, determine corresponding directions along them, and verify the conformality of the mapping at the point $z_0 = 1$.

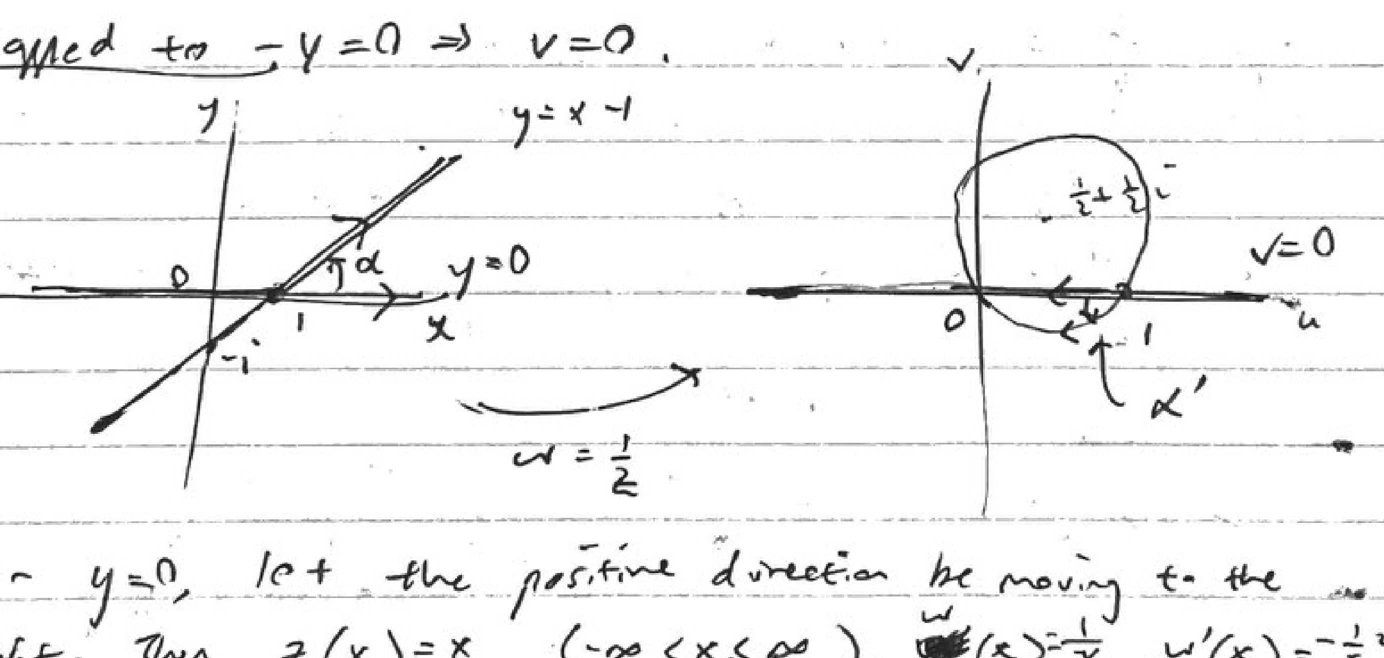
Recall from §98, that a line or circle of the form
$A(x^2+y^2) + Bx + Cy + D = 0$ gets mapped by $w = \frac{1}{z}$
to $D(u^2+v^2) + Bu - Cv + A = 0$.

$y = x-1 \Rightarrow A = 0, B = -1, C = 1, D = 1.$

mapped to $u^2+v^2-u-v=0 \Rightarrow (u-\frac{1}{2})^2+(v-\frac{1}{2})^2 = \frac{1}{2}$

$y = 0 \Rightarrow A=0, B=0, C=1, D=0$

mapped to $-v=0 \Rightarrow v=0$.

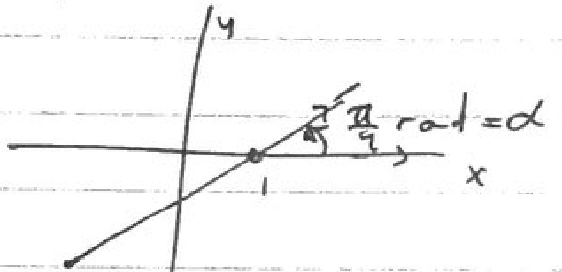


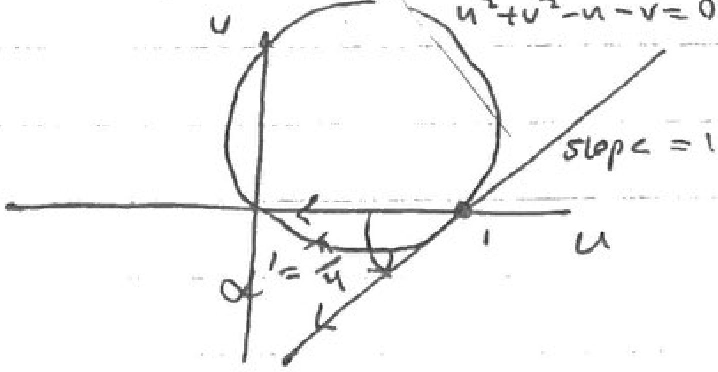
For $y=0$, let the positive direction be moving to the right. Then $z(x) = x$ $(-\infty \leq x \leq \infty)$, $w(x) = \frac{1}{x}$, $w'(x) = -\frac{1}{x^2}dx$ which is always negative. So the corresponding direction for $v=0$ in the uv plane is from right to left.

For $y = x-1$, let the positive direction be up and to the right. In this direction, arg z is increasing, so arg $w = -\arg z$ is decreasing (since $w = \frac{1}{z} = \frac{1}{r}e^{-i\theta}$), so this corresponds to CCW rotation around the circle.

Let α be the angle from $y=0$ to $y=x-1$ in the CCW direction, and α' be the angle between their respective images in the same direction, @ $z=1$ and $w=1$ ($z=1$ maps to $w=1$).

By inspection, the angle between $y=x-1$ and $y=0$ is $\frac{\pi}{4}$ rad.





The slope of the circle $u^2+v^2-u-v=0$ at $w=1$ is (by a Calc II thm for slope of implicit fn):

$$\frac{dv}{du} = -\frac{F_v}{F_u} = -\frac{\frac{\partial}{\partial v}(u^2+v^2-u-v)}{\frac{\partial}{\partial u}(u^2+v^2-u-v)} = -\frac{2v-1}{2u-1}$$

at $(u,v)=(1,0)$, $\frac{dv}{du} = \frac{-2(0)-1}{2(1)-1} = 1.$

Using the orientations previously established, it is clear that $\alpha' = \frac{\pi}{4}$ rad $= \alpha$, thus confirming conformality.

10.) Suppose that a function f is analytic @ z_0 and has a zero of order m there ($m \geq 1$). Also write $w_0 = f(z_0)$

a) Use the Taylor series for f about z_0 to show that there is a nbd of z_0 in which the difference $f(z)-w_0$ can be written. $f(z)-w_0 = (z-z_0)^m \frac{f^{(m)}(z_0)}{m!}[1+g(z)]$

where $g(z)$ is continuous and zero valued @ z_0.

Taylor series centered at $z=z_0$:

$$f(z) = \underbrace{\sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n}$$

zero coefs because of zero of order m.

$$= \underset{\text{nonzero}}{w_0} + \cancel{\frac{f'(z_0)}{1!}}(z-z_0) + \cancel{\frac{f''(z_0)}{2!}}(z-z_0)^2 + \dots +$$

$$\underbrace{\cancel{\frac{f^{(m-1)}(z_0)}{(m-1)!}}(z-z_0)^{m-1}}_{\text{zero coef}} + \underbrace{\frac{f^{(m)}(z_0)}{m!}(z-z_0)^m + \dots}_{\text{nonzero coef}}$$

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis
5/11/20.

§114 #10, cont'd.

$$f(z) = w_0 + \frac{f^{(m)}(z_0)}{m!}(z-z_0)^m + \frac{f^{(m+1)}(z_0)}{(m+1)!}(z-z_0)^{m+1} + \dots$$

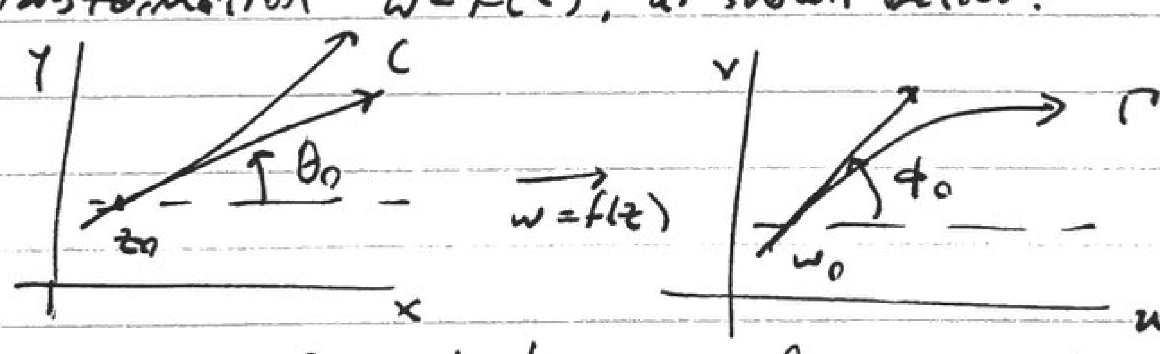
$$f(z) - w_0 = \frac{f^{(m)}(z_0)}{m!}(z-z_0)^m \left(1 + \underbrace{\frac{f^{(m+1)}(z_0)}{f^{(m)}(z_0)} \frac{m!}{(m+1)!}(z-z_0) + \frac{f^{(m+2)}(z_0)}{f^{(m)}(z_0)} \frac{m!}{(m+2)!}(z-z_0)^2 + \dots}_{\text{let this } = g(z).} \right)$$

$$= \frac{f^{(m)}(z_0)}{m!}(z-z_0)^m \left[1 + g(z) \right]$$

$g(z)$ is clearly an ordinary power series and thus analytic (and continuous) @ z_0, and $g(z_0) = 0$ since every term has a (positive) factor of $(z-z_0)$ in it.

---

b). Let Γ be the image of a smooth arc C under the transformation $w = f(z)$, as shown below.



Note that θ_0 and ϕ_0 are $\lim_{z\to z_0} \arg(z-z_0)$ and $\lim_{z \to z_0} \arg(f(z) - w_0)$, respectively. Use the result from the previous section to show that θ_0 and ϕ_0 are related by: $\phi_0 = m\theta_0 + \arg(f^{(m)}(z_0))$.

Recall that arg of a product is the sum of the args of the factors.

$$\arg(f(z)-w_0) = \arg\left((z-z_0)^m \frac{f^{(m)}(z_0)}{m!}(1+g(z))\right)$$

$$= \underbrace{\arg\left((z-z_0)^m\right)}_{\substack{= m\arg(z-z_0) \\ \text{by splitting product into} \\ \text{sum of factors by} \\ \text{above remark}}} + \underbrace{\arg\left(\frac{f^{(m)}(z_0)}{m!}\right)}_{\substack{\Downarrow \\ = \arg f^{(m)}(z_0) \\ \text{(scalar multiplication} \\ \text{doesn't change arg)}}} + \underbrace{\arg(1+g(z))}_{\substack{\searrow \\ g(z)=0 \\ @\, z=z_0}}$$

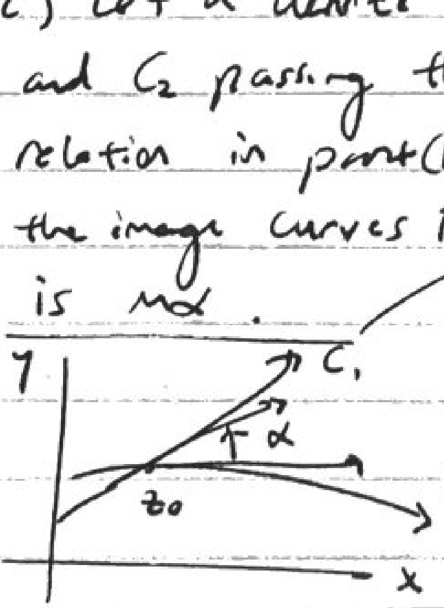
$\phi_0 =$

$$\lim_{z\to z_0} \arg(f(z)-w_0) = \lim_{z\to z_0}\left[m\arg(z-z_0)\right] + \lim_{z\to z_0}\left[\arg\left(f^{(m)}(z_0)\right)\right] + \lim_{z\to z_0}\left[\arg(1+g(z))\right]$$

$$= \underbrace{m\lim_{z\to z_0}\arg(z-z_0)}_{=\theta_0} + \arg\left(f^{(m)}(z_0)\right) + \underbrace{\cancel{\arg(1+\underbrace{g(z_0)}_{0})}}_{\arg(1)=0}$$

$$= m\theta_0 + \arg\left(f^{(m)}(z_0)\right).$$

c) Let α denote the angle between two smooth arcs C_1 and C_2 passing through z_0. Show how it follows from the relation in part (b) that the corresponding angle between the image curves Γ_1 and Γ at the point $w_0 = f(z_0)$ is $m\alpha$.



angle between $C_1, C_2 = \theta_1 - \theta_2 = \alpha$.

angle between $\Gamma_1, \Gamma_2 = \phi_1 - \phi_2$

$$= \left(m\theta_1 + \cancel{\arg(f^{(m)}(z_0))}\right) - \left(m\theta_2 + \cancel{\arg f^{(m)}(z_0)}\right)$$

$$= m(\theta_1 - \theta_2) = m\alpha.$$

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20

§115 #2b.

2). Show that the function $u(x,y)$ is harmonic throughout the xy plane. Then, find its harmonic conjugate. Also, write the corresponding fn. $f(z) = u(x,y) + iv(x,y)$ in terms of z.

b) $u(x,y) = y^3 - 3x^2y$.

Show harmonicity:

$u_x = -6xy$, $u_{xx} = -6y$, $u_y = 3y^2 - 3x^2$, $u_{yy} = 6y$

$u_{xx} + u_{yy} = -6y + 6y = 0$ ✓

Finding harmonic conjugate:

$$v(x,y) = \int_{(x_0,y_0)}^{(x,y)} -u_t(s,t)\,ds + u_s(s,t)\,dt.$$

$$= \int_{(0,0)}^{(x,y)} -\frac{\partial}{\partial t}(t^3 - 3s^2t)\,ds + \frac{\partial}{\partial s}(t^3 - 3s^2t)\,dt$$

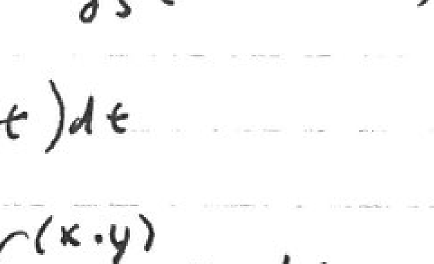
$$= \int_{(0,0)}^{(x,y)} -(3t^2 - 3s^2)\,ds - (6st)\,dt$$

$$= 3\int_{(0,0)}^{(x,y)} s^2 - t^2\,ds - 6\int_{(0,0)}^{(x,y)} st\,dt$$

$$= 3\int_0^x s^2 - \cancel{0^2}\,ds$$

$$-6\int_0^y xt\,dt$$

use this diagram and conservativity of this line integral



$$= 3\frac{s^3}{3}\Big|_0^x - 6x\frac{t^2}{2}\Big|_0^y$$

$$= x^3 - 3xy^2$$

$f(z) = u(x,y) + iv(x,y) = y^3 - 3x^2y + ix^3 - 3ixy^2$

$= i(-iy^3 - 3xy^2 + 3x^2yi + x^3) = i((iy)^3 + 3x(iy)^2 + 3x^2(iy) + x^3)$

$= i(x+iy)^3 = iz^3.$

§117 #4.

Under the transformation $w = \exp z$, the image of the segment $0 \le y \le \pi$ of the y-axis is the semicircle $u^2+v^2=1$, $v \ge 0$. Also, the function $h(u,v) = Re\left(2 - w + \frac{1}{w}\right) = 2 - u + \frac{u}{u^2+v^2}$ is harmonic everywhere in the w plane except the origin, and it assumes the value $h=2$ on the semicircle. Write an explicit expression for the function $H(x,y)$ in the form of §117. Then illustrate the thm. by directly showing that $H=2$ along the segment $0 \le y \le \pi$ of the y axis.

$$H = h(u(x,y),\ v(x,y))$$

$$= h(e^x\cos y,\ e^x \sin y)$$

$$= 2 - e^x\cos y + \frac{e^x \cos y}{(e^x\cos y)^2 + (e^x \sin y)^2}$$

$$= 2 - e^x\cos y + \frac{e^x\cos y}{e^{2x}(\cos^2 y + \sin^2 y)}$$

$$= 2 - e^x \cos y + e^{-x}\cos y = 2 - 2\cos y \sinh x$$

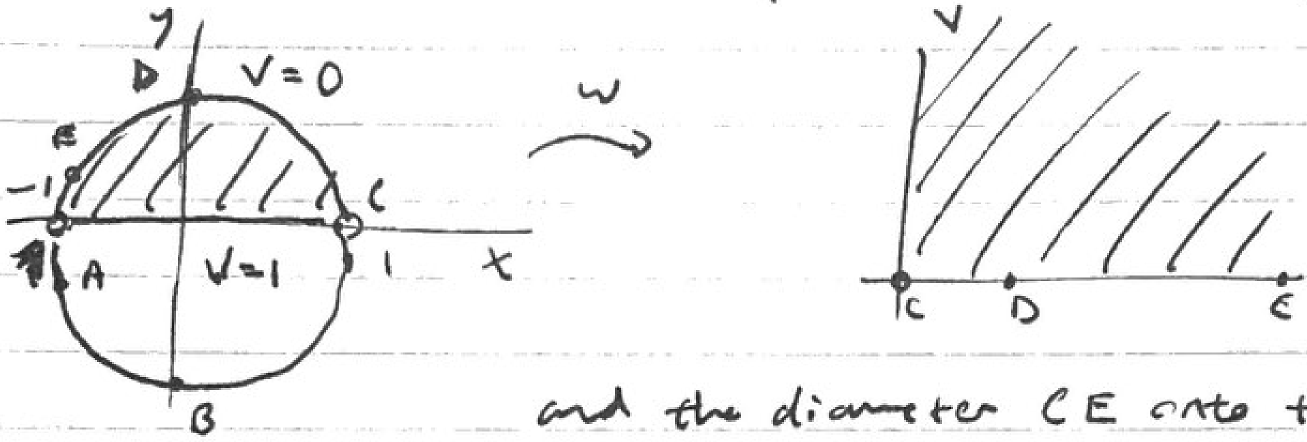
On the line segment $0 \le y \le \pi$ of the y-axis, $x=0$, so $\sinh x = 0$, and

$$H(0,y) = 2 - 2\cos y \sinh(0) = 2 - \cancel{2\cos y (0)} = 2 \checkmark.$$

Jonathan Lam
Prof. Smyth
MA 345
Cplx Analysis
5/11/20.

§123 #2, 10.

2). Show that the transformation $w = i\left(\frac{1-z}{1+z}\right)$ of §123 maps the upper half of the circular region shown below onto the first quadrant of the w plane



and the diameter CE onto the positive v-axis. Then find the electrostatic potential in the space enclosed by the half-cylinder $x^2+y^2=1$, $y \geq 0$ and the plane $y=0$ when $V=0$ on the cylindrical surface and $V=1$ on the planar surface.

we already know that the arc CDE gets mapped to the positive real axis by $w = i\left(\frac{1-z}{1+z}\right)$ from Fig. 13 in Appendix 2 and the example in § 123. Thus, we only need to show that CE maps to the positive v-axis.

EC: $z = x, \quad -1 \leq x \leq 1$

$w = i\left(\frac{1-x}{1+x}\right)$ is clearly pure imaginary, continuous, and its imaginary component is monotonic on $x \in (-1, 1)$.

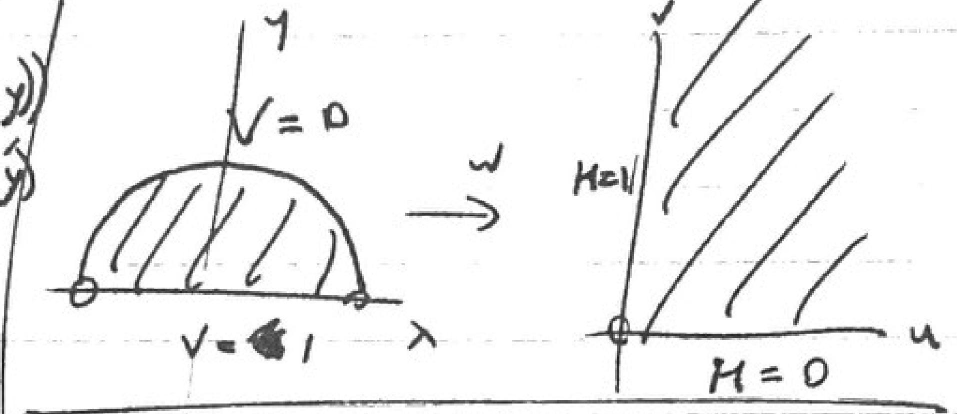
$\lim_{x \to -1} w = +\infty, \quad w = 0 \ @ \ x = 1.$

Thus CE must map to the positive v-axis, and the semicircular region must map to the first quadrant in the w-plane.

BVP setup:
$$\begin{cases} V_{xx}(x,y) + V_{yy}(x,y) = 0 & \begin{pmatrix} x^2+y^2<1, \\ y>0 \end{pmatrix} \\ V(x,y) = 1 & \begin{pmatrix} x^2+y^2=1, \\ y>0 \end{pmatrix} \\ V(x,0) = 0 & (-1 \le x < 1) \end{cases}$$

By the Thm in §116 and §117 we can map this problem to a similar one in the w plane, preserving harmonicity and maintaining the boundary condition values on the mapped boundaries:

Let $H = V(u(x,y), v(x,y))$
where $w = u(x,y) + iv(x,y)$



This translates to the problem:
$$\begin{cases} H_{uu}(u,v) + H_{vv}(u,v) = 0 & (u,v>0) \\ H(0,v) = 1 & (v>0) \\ H(u,0) = 0 & (u>0) \end{cases}$$

This can be solved by inspection. The imaginary part of the analytic function

$$\frac{2}{\pi}\operatorname{Log} w = \frac{2}{\pi}\ln|w| + i\underbrace{\frac{2}{\pi}\arg w}_{H(u,v)}$$
(analytic on all but negative real axis and origin, which includes desired domain)

is harmonic and matches these boundary conditions.

Here, (in the first quadrant) the arctangent function matches the value of arg,

so $H = \frac{2}{\pi}\arctan\left(\frac{v}{u}\right) = V(u,v)$ $\left(0 < \arctan\theta < \frac{\pi}{2}\right)$

(Need to find v, u in terms of x, y)

$$w = i\left(\frac{1-z}{1+z}\right) = i\left(\frac{(1-z)(1+z^*)}{|1+z|^2}\right) = \frac{i}{(1+x)^2+y^2}(1+z^*-z-zz^*)$$

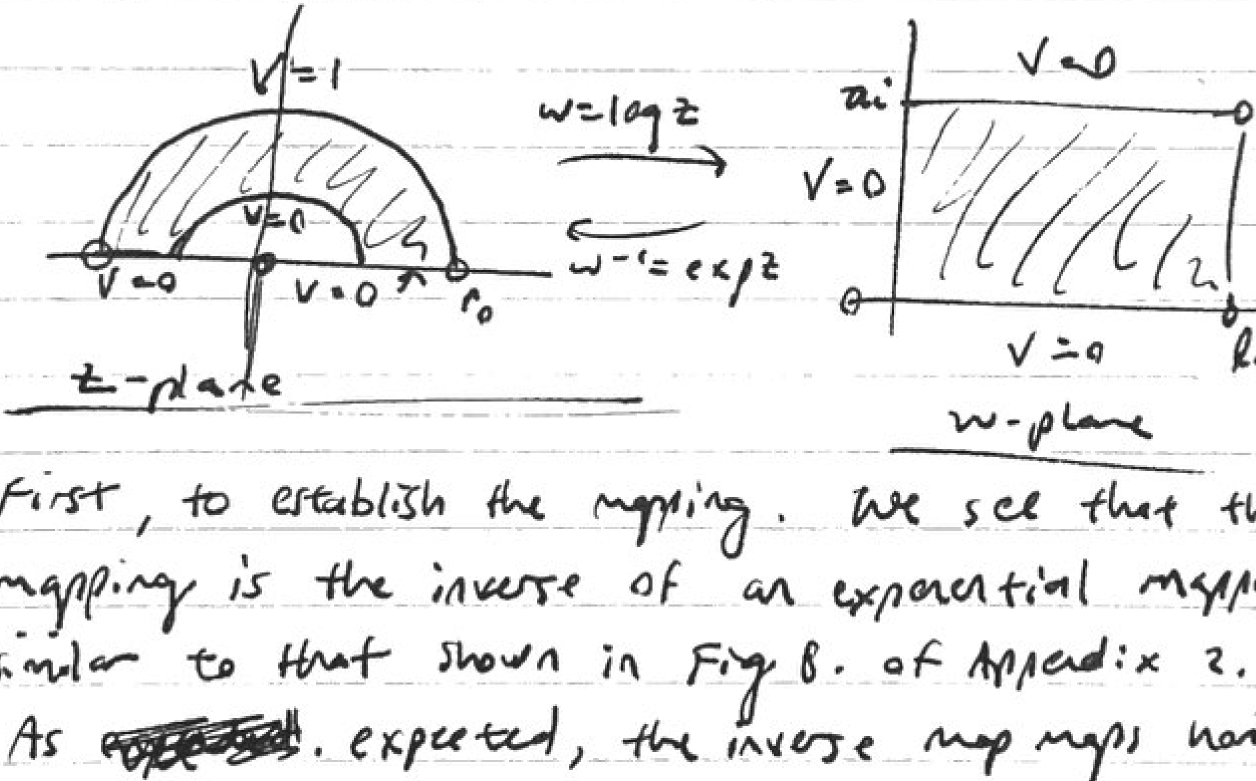
$$= \frac{i}{(1+x)^2+y^2}(1-x^2-y^2-2yi) = \underbrace{\frac{2y}{(1+x)^2+y^2}}_{u(x,y)} + i\underbrace{\frac{1-x^2-y^2}{(1+x)^2+y^2}}_{v(x,y)}$$

So $\underline{V(x,y) = \frac{2}{\pi}\arctan\left(\frac{1-x^2-y^2}{2y}\right)}$

Jonathan Lam
Prof. Smyth
MA345
Cplx Analysis
5/11/20.

§123 # 10.

10) The soln. to the Dirichlet problem on the right of the below figure is
$$V = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sinh mu}{m \sinh(m \ln r_0)} \sin mv$$
where $m = 2n-1$. By using the branch
$$\log z = \ln r + i\theta \quad \left(r > 0, -\frac{\pi}{2} < \theta < \frac{3\pi}{2}\right)$$
of the logarithmic function, derive the following soln of the Dirichlet problem on the bottom left:
$$V(r,\theta) = \frac{4}{\pi} \sum_{n=1}^{\infty} \left(\frac{r^m - r^{-m}}{r_0^m - r_0^{-m}}\right) \frac{\sin m\theta}{m}.$$
where $m = 2n-1$.



First, to establish the mapping. We see that this mapping is the inverse of an exponential mapping similar to that shown in Fig 8. of Appendix 2. As expected, the inverse map maps horizontal line segments to segments of radially-oriented lines, and vertical line segments are mapped to circular arcs centered at the origin. In particular:

$$\left.\begin{array}{l}\text{top edge of rectangle}\\ \text{left edge of rectangle}\\ \text{bottom edge of rectangle}\\ \text{right edge of rectangle}\end{array}\right\} \xrightarrow[z=\exp w]{w^{-1}} \left\{\begin{array}{l}\text{bottom left segment on x-axis}\\ \text{bottom arc, } r=1\\ \text{bottom right segment on x-axis}\\ \text{top arc, } r=r_0\end{array}\right.$$

Since $\exp z = w^{-1}$, we may use the branch of the log:

$$w = \log z = \ln r + i\theta \quad (r > 0, -\tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2})$$

to map the z-plane to the w plane. This branch of log is singular only on the negative imaginary axis, so it is analytic on the desired domain. Thus a conformal mapping is established.

Since we have a conformal mapping, and a harmonic function, by the thm in §116, $V(u,v)$ is harmonic in the mapped region, and by the thm in §117 this harmonic function maintains the same values for the boundary conditions in the z-plane as the mapped conditions in the w-plane.

Note that the mapped boundary conditions already match, i.e.:

$$\left.\begin{array}{c}\text{lower edges in } z\text{-plane,} \\ V = 0\end{array}\right\} \xrightarrow{w} \left\{\begin{array}{c}\text{left, top, bottom edges} \\ \text{in } w\text{-plane, } V = 0\end{array}\right.$$

$$\left.\begin{array}{c}\text{top arc in } z \text{ plane,} \\ V = 1\end{array}\right\} \xrightarrow{w} \left\{\begin{array}{c}\text{right edge in } w\text{-plane,} \\ V = 1\end{array}\right.$$

and we are already provided with a solution $V(u,v)$. Thus, to obtain the solution for the preimage, all we must do is express (u,v) in terms of z-plane coordinates (r,θ).

$$u = \ln r, \quad v = \theta, \quad \left(-\tfrac{\pi}{2} < \theta < \tfrac{3\pi}{2}\right)$$

$$V(\ln r, \theta) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sinh(m \ln r)}{m \sinh(m \ln r_0)} \sin(m\theta)$$

$$= \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cancel{\tfrac{1}{2}}(e^{m \ln r} - e^{-m \ln r}) \sin(m\theta)}{m \cdot \cancel{\tfrac{1}{2}}(e^{m \ln r_0} - e^{-m \ln r_0})}$$

$$= \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{r^m - r^{-m}}{r_0^{\,m} - r_0^{\,-m}} \frac{\sin(m\theta)}{m}$$

(where $m = 2n - 1$)

MA345 Test 1 Notes

Jonathan Lam

February 20, 2020

Contents

1 Basic algebraic properties and representations

$$z = x + iy = re^{i\theta} = r(\cos\theta + i\sin\theta) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$$

$$(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

$$(x + iy)^{-1} = \frac{1}{x^2 + y^2}(x - iy) = \frac{\bar{z}}{|z|^2},\ z \neq 0$$

$$\Re(z_1 + z_2) = \Re(z_1) + \Re(z_2),\ \Im(z_1 + z_2) = \Im(z_1) + \Im(z_2)$$

$$|z|^2 = \Re(z)^2 + \Im(z)^2$$

$$\Re(z) \leq |\Re(z)| \leq |z|,\ \Im(z) \leq |\Im(z)| \leq |z|$$

$$r = |z| = \sqrt{x^2 + y^2},\ \theta = \operatorname{atan2}(y, x)$$

$$|z_1 - z_2| = \mathrm{dist}(z_1, z_2)$$

$$\left|\prod_i z_i\right| = \prod_i |z_i|$$

$$\bar{\bar{z}} = z,\ |\bar{z}| = |z|$$

$$\overline{\prod_i z} = \prod_i \bar{z},\ \overline{\sum_i z} = \sum_i \bar{z}$$

(Products also work with division, sums also work with subtraction.)

$$\Re(z) = \frac{z + \bar{z}}{2},\ \Im(z) = \frac{z - \bar{z}}{2i}$$

$$z \in \mathbb{R} \iff z = \bar{z},\ z \in \mathbb{R} \cup \{z : z = ni,\ n \in \mathbb{R}\} \iff z^2 = \bar{z}^2$$

$$z^2 + \bar{z}^2 = c \text{ is a hyperbola}$$

$$\arg z = \mathrm{Arg} z + 2\pi n,\ n \in \mathbb{Z}$$

$$-\pi < \mathrm{Arg} z \leq \pi$$

$$\arg z_1 z_2 = \arg z_1 + \arg z_2$$

$$\arg z^{-1} = -\arg z$$

$$\overline{e^{i\theta}} = e^{-i\theta}$$

By De Moivre's Theorem:

$$c_k = z^{1/n} = \sqrt[k]{r}\left(\cos\frac{\theta_0 + 2\pi k}{k} + i\sin\frac{\theta_0 + 2\pi k}{k}\right)$$

If $\theta_0 = \mathrm{Arg}\, z$, c_0 is called the principal k-th root of z. For square roots, note that

$$\cos\frac{\theta}{2} = \sqrt{\frac{1 + \cos\theta}{2}}$$

$$(z^{1/m})^{-1} = (z^{-1})^{1/m}$$

1.1 Other notes

- Complex numbers are not ordered; comparisons may only happen between two reals, since $\mathbb{R}$ is a total order.

2 Neighborhoods and regions

- An ε neighborhood of z_0 is the domain $|z - z_0| < \varepsilon$.
- A deleted ε neighborhood of z_0 is the domain $0 < |z - z_0| < \varepsilon$.
- An interior point z_0 of S is one such that there exists some neighborhood of z_0 such that it only contains points in S. E.g., a finite set of points cannot have any interior points.
- An exterior point z_0 of S is one such that there exists no neighborhood of z_0 that only contains points in S. In other words, there exists a neighborhood of z_0 that contains no points in S. In other words, it is a point not in S and not a boundary point in S.
- A boundary point z_0 is one such that for every neighborhood of z_0, there exists points in S and points not in S. E.g., a single isolated point is a boundary point.
- An open set contains none of its boundary points (thus contains only interior points). A closed set contains all of its boundary points. A set's closure is the union of a set and its boundary (thus any closure is closed). If some boundary points contained and some not, then neither open nor closed. If no boundary points, then both open and closed (e.g., $\mathbb{C}$).
- A set is connected if each pair of points may be joined by a polygonal line (a finite collection of line segments joined end-to-end).
- A nonempty open connected set is called a domain.
- A domain with none or any of its boundary points is called a region. (Note: a region cannot be a single point, since there is no empty domain.)
- A set is bounded if there exists $R < \infty$ s.t. every point in S lies in the region $|z| < R$.
- A point is an accumulation point (limit point) if each deleted neighborhood of z_0 contains at least one point in S. S is closed iff it contains all of its accumulation points. E.g., an isolated point is not an accumulation point. E.g., any boundary point that is not a isolated point is a boundary point. E.g., any interior point is an accumulation point. E.g., a finite set of points cannot have any accumulation points.

3 The squaring function

This maps hyperbolas centered at the origin to straight lines. E.g., if $x^2 - y^2 = c_1$, then this gets mapped to the line $u = c_1$. Mappings are not necessarily 1-1.

4 Limits, derivatives, and continuity

$$\lim_{z \to z_0} f(z) = w_0 \iff \exists \delta : |f(z) - w_0| < \varepsilon \text{ whenever } 0 < |z - z_0| < \delta$$

i.e., The limit exists if there exists some deleted δ-neighborhood of the approach point s.t. the image of the entire δ-neighborhood is contained within the ε-neighborhood of the limit. If δ has been found, then any smaller δ may also be used. This definition of a limit guarantees uniqueness and is only applicable to interior points (which is fine because we mostly deal with open sets).

Limits of a complex variable iff limits of real and imaginary parts. (same with continuity)

Continuity works the same as limits, and composition of continuous functions is continuous where they are defined.

Can use different values along two paths for showing limit doesn't exist.

5 Limits at infinity

The ε-neighborhood of infinity is $|z - \infty| < \varepsilon$. $|\frac{1}{z} - 0| < \varepsilon$. Thus:

$$\lim_{z \to z_0} f(z) = \infty \iff \lim_{z \to z_0} \frac{1}{f(z)} = 0$$

$$\lim_{z \to \infty} f(z) = w_0 \iff \lim_{z \to 0} f\left(\frac{1}{z}\right) = w_0$$

$$\lim_{z \to \infty} f(z) = \infty \iff \lim_{z \to 0} \frac{1}{f(1/z)} = 0$$

A set is unbounded iff every neighborhood of the point at infinity contains at least one point in S.

6 Cauchy-Riemann equations

If a function is differentiable at a point z_0, then the C-R equations are satisfied at that point:

$$u_x = v_y, \ u_y = -v_x$$

and $f'(x) = u_x + iv_x$. The converse is also true if u_x, u_y, v_x, v_y exist throughout a neighborhood of z_0 and are continuous at z_0 (this gives a sufficient condition for differentiability).

In polar coordinates, the equivalent expressions are $ru_r = v_\theta$, $-rv_\theta = u_\theta$, and $f'(z) = e^{-i\theta}(u_r + iv_r)$, with the same sufficient conditions.

In general,

$$\frac{\partial F}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial F}{\partial x} + i\frac{\partial F}{\partial y}\right)$$

By applying the C-R equations, for a differentiable function, this simplifies to

$$\frac{\partial F}{\partial \bar{z}} = 0$$

i.e., a differentiable function is independent of $\bar{z}$.

7 Analytic functions

A function f of the complex variable z is analytic (also: regular, holomorphic) in an open set S if it has a derivative everywhere in the set. It is analytic at a point if it is analytic in some neighborhood of z_0. (Thus analyticity never occurs if analytic only in some isolated point, i.e., not at an accumulation point). If a function is differentiable everywhere in an open set, it is analytic on that set.

If a function is analytic at point z_0, it must also be analytic at each point in a neighborhood of z_0. If we say a $f(z)$ is analytic in a non-open set S, we mean it is analytic in some open set containing S.

Rules of differentiation (e.g., sums, differences, products, quotients, composition, etc.) for differentiable functions on open sets carry over to analytic functions.

If a function is not analytic at a point z_0 but is analytic at some point of every neighborhood of z_0, z_0 is called a singularity.

If a function $f(z)$ and its conjugate $\overline{f(z)}$ are both analytic over a domain D, then $f(z)$ is constant over D (uses C-R to show that $f'(z) = 0$ throughout D); thus, if $f(z)$ is real-valued throughout domain D, then $f(z)$ is constant. Similarly, if $f(z)$ is analytic over D and its modulus is constant over D, then $f(z)$ is constant over D.

8 Harmonic functions

Harmonic functions satisfy Laplace's equation, i.e., $H_{xx}+H_{yy} = 0$. The real and imaginary components of a function over a domain which it is analytic satisfy Laplace's equation.

The families of level curves of the component functions of an analytic function are orthogonal if $f'(z_0) \neq 0$.

The polar form of Laplace's equation holds for the component functions when given an input w.r.t. r, θ.

$$r^2 u_{rr} + r u_r r, \theta + u_{\theta\theta} = 0$$

9 Misc. Theorems

Theorem 1 (Triangle inequality)

$$||z_1| - |z_2|| \leq |z_1 + z_2| \leq |z_1| + |z_2|$$

(This can be extended to higher orders by induction.)

Theorem 2 *For any polynomial, $\exists R$ s.t.*

$$\left|\frac{1}{P(z)}\right| \leq \frac{2}{|a_n|R^n}$$

whenever $z > R$.

Theorem 3 (De Moivre's Theorem)

$$z^n = r^n(\cos n\theta + i \sin n\theta)$$

Theorem 4

$$|z_1| = |z_2| \iff \exists c_1, c_2 \in \mathbb{C} : z_1 = c_1 c_2,\ z_2 = c_1 \bar{c_2}$$

(For the forward direction, let $c_1 = \sqrt{r}\exp(\frac{i}{2}(\theta_1+\theta_2))$, $c_2 = \sqrt{r}\exp(\frac{i}{2}(\theta_1-\theta_2))$).

Theorem 5

$$1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}$$

Theorem 6 (Lagrange's trigonometric identity)

$$1 + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{1}{2} + \frac{\sin\frac{(2n+1)\theta}{2}}{2\sin\frac{\theta}{2}}$$

MA345 – Test 2 Review

Jonathan Lam

April 5, 2020

This test covers Chapter 2: Elementary Functions, and Chapter 3: Integrals of *Complex Variables and Applications, 9th edition* by Churchill and Brown.

Contents

1 The exponential and log functions

1.1 Definitions

$$e^z := e^x e^{iy}$$

Here, if x is a root $1/n$, then the *positive* nth root is used, i.e., $e^{x=1/n} = \sqrt[n]{e}$. If $w = \rho e^{i\phi} = e^z$, then $\log w = z$. We define the inverse to be the log function:

$$\log w = \log(\rho e^{i\phi}) := \ln|z| + i\arg z = \ln\rho + i(\phi + 2n\pi),\ n \in \mathbb{Z}$$

1.2 Notes

- For calculating the exponential, we begin in rectangular form and end in polar form; the reverse is true for the logarithm.
- e^z is many-to-one and $\log z$ is one-to-many. Thus $e^{\log z} = z$ is a single value and $\log e^z = z + 2n\pi,\ n \in \mathbb{Z}$ is multi-valued. We may choose a single value of the log function by selecting a particular branch of the arg function, in particular, the principal logarithm $\operatorname{Log} z$ is defined as:

  $$\operatorname{Log} z := \ln|z| + i\operatorname{Arg} z$$

  and this reduces to the real case when z is a positive real.
- The complex exponential is defined over the entire complex plane and is entire; the logarithm is defined over the punctured complex plane. Since we need to choose a branch of the logarithm to get a singly-valued function, the logarithm is not analytic anywhere on its branch cut or at the origin (the branch point); however, it is analytic everywhere else.
- Because of branches and the "wrapping" nature of the exponential and logarithm, some properties of the exponential and log may not hold true everywhere, e.g., $\log z^c = c\log z$ is not true for all values of z, c and all branch cuts.

1.3 Properties

$$|e^z| = e^x$$

thus $|e^z| \neq 0$.

$$\arg e^z = y + 2n\pi,\ n \in \mathbb{Z}$$

$$e^{z_1}e^{z_2} = e^{z_1+z_2}$$

$$\frac{e^{z_1}}{e^{z_2}} = e^{z_1-z_2}$$

$$\frac{1}{e^z} = e^{-z}$$

$$\frac{d}{dz}e^z = e^z$$

$$e^{z+2\pi in} = e^z,\ n \in \mathbb{Z}$$

thus e^z is $2\pi i$-periodic and is many-to-one.

$$\frac{d}{dz}\mathrm{Log}\, z = \frac{1}{z}\ (z \neq 0, \mathrm{Arg}\, z \neq \pi)$$

This uses the polar form of the C-R equations (since it is a pain to express the angle in terms of rectangular coordinates). The same is true for any branch, applying the appropriate condition for its branch cut.

$$\log(z_1 z_2) = \log z_1 + \log z_2$$

$$\log \frac{z_1}{z_2} = \log z_1 - \log z_2$$

These properties are not always true for the principal branch.

2 The power function

2.1 Definitions

$$z^c := e^{c \log z}$$

where $c \in \mathbb{C}$ is some constant. In general, z is multiple-valued unless $c \in \{0, 1, 2, \dots\}$. The principal value of z^c is as expected:

$$\mathrm{P.V.}\, z^c := c^{\mathrm{Log}\, z}$$

We can use this to define the exponential function with base $c \in \mathbb{C}$:

$$c^z := e^{z \log c}$$

and has a principal value as expected.

2.2 Notes

- $z^n,\ n \in \mathbb{Z}$ and $z^{1/n},\ n \neq 0 \in \mathbb{Z}$ agrees with the previous definition (i.e., repeated multiplication and square roots). The former is singly-valued.
- Since this definition of the exponential function with arbitrary base suggests that exponential functions are multiply-valued, this suggests that the fundamental exponential is also multiply-valued. Thus, our usual interpretation of e^z is the principal exponential function with base e.
- Some real properties may not work here because of wrapping, e.g., $(z_1 z_2)^c = z_1^c z_2^c$ is not a valid identity for all z_1, z_2, and c.

2.3 Properties

$$\frac{1}{z^c} = z^{-c}$$

Again, we may have branches of the power function determined by the branches of $\log z$. To find the derivative, we must choose a particular branch of $\log z$, and then by the Chain rule:

$$\frac{d}{dz} z^c = cz^{c-1}$$

If we choose a specific value of $\arg c$, then c^z is entire, and its derivative is

$$\frac{d}{dz} c^z = c^z \log c$$

3 The trigonometric and hyperbolic functions

3.1 Definitions

The real sine and cosine functions can be defined using the exponential function and Euler's formula alone. We can define their complex analogues by replacing the real parameters with complex ones:

$$\sin z := \frac{e^{iz} - e^{-iz}}{2i}$$

$$\cos z := \frac{e^{iz} + e^{-iz}}{2}$$

The other trigonometric functions are defined in terms of these two functions in the same way. Their derivatives are all in the same form as their real counterparts.

We define the hyperbolic functions as their analogues are defined:

$$\sinh z = \frac{e^z - e^{-z}}{2}$$

$$\cosh z = \frac{e^z + e^{-z}}{2}$$

Likewise, we define the rest of the hyperbolic functions in the same way, and get results of the same form as their real counterparts.

We can obtain the inverses of the trigonometric and hyperbolic functions by solving a quadratic expression, where the unknown variable is e^{iz} or e^z. We get the following results:

$$\sin^{-1} z = -i \log \left[iz + (1 - z^2)^{1/2} \right]$$

$$\cos^{-1} z = -i \log \left[z + i(1-z^2)^{1/2} \right]$$

$$\tan^{-1} z = \frac{i}{2} \log \frac{i+z}{i-z}$$

$$\sinh^{-1} z = \log \left[z + (z^2+1)^{1/2} \right]$$

$$\cosh^{-1} z = \log \left[z + (z^2-1)^{1/2} \right]$$

$$\tanh^{-1} z = \frac{1}{2} \log \frac{1+z}{1-z}$$

3.2 Notes

- As expected, the trigonometric and hyperbolic functions reduce to their real counterparts when $z \in \mathbb{R}$.
- While these functions are periodic and bounded in the real direction for a particular y, they are unbounded in the imaginary direction (and grow exponentially in magnitude).
- The inverse functions are multiply-valued, unless a single branch of the logarithm and square root are used.

3.3 Properties

$$\frac{d}{dz} \sin z = \cos z, \ \frac{d}{dz} \cos z = -\sin z$$

$$\sin -z = -\sin z, \ \cos -z = \cos z$$

$$\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2$$

$$\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$$

$$\sin\left(z + \frac{\pi}{2}\right) = \cos z, \ \sin\left(z - \frac{\pi}{2}\right) = -\cos z$$

$$\sin^2 z + \cos^2 z = 1$$

$$\sin(z+2\pi) = \sin z, \ \sin(z+\pi) = -\sin z, \ \cos(z+2\pi) = \cos z, \ \cos(z+\pi) = -\cos z$$

We can express the real and imaginary parts w.r.t. the trigonometric and hyperbolic functions of real variables:

$$\sin z = \sin x \cosh y + i \cos x \sinh y$$

$$\cos z = \cos x \cosh y - i \sin x \sinh y$$

Note that to derive the above expressions, we use the sum identity on $z_1 = x$, $z_2 = iy$ and then differentiate the sin decomposition to achieve the cos decomposition. These identities can also be used to show:

$$|\sin z|^2 = \sin^2 x + \sinh^2 y$$

$$|\cos z|^2 = \cos^2 x + \sinh^2 y$$
$$\sin z = 0 \iff z = n\pi \, n \in \mathbb{Z}$$
$$\cos z = 0 \iff z = (2n+1)\frac{\pi}{2}, \; n \in \mathbb{Z}$$

i.e., the zeros of the complex sine and cosine are the same as those of their real analogues. (And the zeros of the hyperbolic functions are of the same magnitude, but on the imaginary axis.)

$$\sinh iz = i \sin z, \; \cosh iz = \cos z$$

$$\sin iz = i \sinh z, \; \cos iz = \cosh z$$
$$\frac{d}{dz} \sin^{-1} z = \frac{1}{(1-z^2)^{1/2}}$$
$$\frac{d}{dz} \cos^{-1} z = -\frac{1}{(1-z^2)^{1/2}} \; \left[= -\frac{d}{dz} \sin^{-1} z\right]$$
$$\frac{d}{dz} \tan^{-1} z = \frac{1}{1+z^2}$$

4 Derivatives and definite integrals of complex-valued functions of a real variable

4.1 Prelimary results

Let

$$w(t) = u(t) + iv(t)$$

where $t \in \mathbb{R}$, and u and v are real-valued. Then, when u' and v' exist, then the derivative $w'(t)$ is

$$w'(t) = u'(t) + iv'(t)$$

In other words, this shows us how to differentiate a function if we parameterize it w.r.t. a real variable. Since u and v are real-valued, we may integrate over them. Thus, the definite integral of w is something we already know from regular calculus:

$$\int_a^b w(t)\, dt = \int_a^b u(t)\, dt + i \int_a^b v(t)\, dt$$

This is well-defined whenever u and v are piecewise-continuous, as we know from integrals of real functions. FTC also applies; if the complex-valued function of a real variable W is found s.t.

$$U'(t) = u(t), \; V'(t) = v(t)$$

(i.e., if W is the antiderivative of w), then

$$\int_a^b w(t)\, dt = W(b) - W(a)$$

4.2 Notes

- The mean value theorem for derivatives or integrals doesn't always apply, due to the wrapping (periodicity) of the complex plane.

5 Contours, contour integrals, and the ML-inequality

5.1 Definitions

An arc is the set of points generated by a continuous parameterization of x and y on the complex plane over the same interval. In other words, $z = z(t)$ $(a \leq t \leq b)$ is an arc if

$$x = x(t),\ \ y = y(t)$$

are piecewise continuous over the interval $[a, b]$. An arc is simple if it does not self-intersect. An arc is called a simple closed curve if it is simple except for the fact that its end and starts points are the same. It is positively oriented if the interior of the loop is always on the left (i.e., traveling CCW). A differentiable arc is one in which $z'(t)$ is continuous, where

$$z'(t) = \sqrt{(x'(t))^2 + (y'(t))^2}$$

An arc is smooth if its derivative is continuous and its value is nonzero in the open interval (a, b).

A contour is a piecewise smooth arc (i.e., it is composed of a finite number of smooth arcs joined end to end). A simple closed countour (SCC) is a contour where only the final and initial points are the same. Like in the real case, the contour (line) integral is defined as follows:

$$\int_C f(z)\,dz = \int_a^b f[z(t)]z'(t)\,dt$$

If $|f(z)| \leq M\ \forall z \in C$, then the upper bound on the modulus of the integral is

$$\left|\int_C f(z)\,dz\right| \leq ML$$

where L is the arc length (shown below). A slightly more general case is that

$$\left|\int_a^b w(t)\,dt\right| \leq \int_a^b |w(t)|\,dt$$

5.2 Notes

- The same arc can be represented by multiple parameterizations, and there may be different arcs with the same set of points (i.e., if an arc overlaps itself)

- If we change the parameterization of an arc, we still get the same value when integrating some function over it. I.e., no matter the parameterization of an arc,

$$L = \int_a^b |z'(t)|\, dt$$

is the arc length and is invariant of the parameterization of the arc.

- The contour integral may begin or end on a branch cut of the integrand.

6 Major integral theorems

6.1 Definitions

A simply connected domain is one s.t. every SCC completely inside the domain only encloses points in the domain. (Anything else is a multiply-connected domain.)

6.2 Antiderivative theorem

Suppose that a function $f(z)$ is continuous in a domain D. TFAE:

1. $f(z)$ has an antiderivative $F(z)$ throughout D.
2. Integrals of $f(z)$ along contours lying entirely in D have the same value (path independence):

$$\int_{z_1}^{z_2} f(z)\, dz = F(z_2) - F(z_1)$$

3. For a CC lying completely in D,

$$\oint_C f(z)\, dz = 0$$

Note that the antiderivative of a function is unique (if it exists) and is necessarily differentiable.

6.3 Cauchy-Goursat theorem (CG)

If a function f is analytic at all points interior to and on a simple closed contour C, then

$$\oint_C f(z)\, dz = 0$$

6.4 Principle of deformation of paths (PDP)

Let C_1 and C_2 denote POSCCs s.t. C_1 is interior to C_2. If a function f is analytic on and between them, then

$$\oint_{C_1} f(z)\, dz = \oint_{C_2} f(z)\, dz$$

6.5 Cauchy integral formula (CIF)

Let f be analytic AOIC a SCC C. If z_0 is a point interior to C, then:

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz$$

We denote the use of this formula with argument n to be CIF(n).

6.6 Summary of major conditions and results

Conditions:

1. Known antiderivative throughout some domain. (Doesn't have to be simply connected.)

2. A closed curve in that domain. (Doesn't have to be simple.)

then integrals around a closed loop evaluate to zero, and two integrals to same place evaluate to the same value.

Conditions:

1. SCC C

2. f AOIC C

then the integral around the curve evaluates to zero.

Conditions:

1. POSCC C_2

2. POSCC C_1 completely contained within C_2

3. f analytic on and between C_1 and C_2

then the integrals of f around C_1 and C_2 are equal.

Conditions:

1. POSCC C

2. f AOIC C

3. Integral of form

$$\oint_C \frac{f(z)}{(z-z_0)^{n+1}}\,dz$$

where z_0 lies in the interior of C, and $n \geq 0 \in \mathbb{Z}$

then the integral evaluates to

$$\oint_C \frac{f(z)}{(z-z_0)^{n+1}}\,dz = \frac{2\pi i}{n!} f^{(n)}(z_0)$$

7 Other theorems (not for calculating integrals)

7.1 Theorems for simply-connected domains

- If a function is analytic throughout a domain D, then

$$\oint_C f(z)\,dz = 0$$

  for any CC C contained completely within D.

- A function that is analytic throughout a domain D has an antiderivative throughout D.

- Entire functions always possess antiderivatives.

7.2 Consequences of CIF

- If f is analytic at a point, then its derivatives of all orders are analytic there too.

- Let f be continuous on a domain D. If

$$\oint_C f(z)\,dz = 0$$

  for every CC C in D, then f is analytic throughout D. (This is the converse to a theorem in (Sec. 6.4.).)

- (Cauchy's inequality) Assume f is AOIC a PO circle C_R, centered at z_0 and with radius R. If M_R denotes the maximum value of $|f(z)|$ on C_R, then

$$\left|f^{(n)}(z_0)\right| \leq \frac{n! M_R}{R^n}$$

7.3 Maximum modulus principle

A secondary result is Gauss's mean value theorem: if f is analytic within and on a given circle centered at z_0 and with radius ρ, then

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + \rho e^{i\theta})\, d\theta$$

i.e., the value of f at the center is the mean of the values of f along the edge of the circle. The main result is the maximum modulus principle: If a function f is analytic and not constant in a given domain D, then $|f(z)|$ has no maximum value in D.

8 Uncovered results

8.1 Liouville's theorem

If a function f is entire and bounded in the complex plane, then f is constant throughout the plane.

8.2 The fundamental theorem of algebra

Any n-th order polynomial

$$P(z) = \sum_{i=0}^{n} a_i z^i$$

has at least one zero. By applying this theorem repeatedly, then an n-th order polynomial has exactly n zeros (not necessarily all distinct). In other words, we can factor any n-th order polynomial $P(z)$ into

$$P(z) = \prod_{i=1}^{n} (z - z_i)$$

where $\{z_i\}$ are the zeros of P, not necessarily all distinct.

Arithmetical functions

Jonathan Lam

February 14, 2020

This is a more explicit understanding of Chapter 2: Arithmetical functions, from Alan Baker's *A concise introduction to the theory of numbers.*

Important results about the floor function

Basic properties of the floor function: $[x+y] \geq [x]+[y]$ (floor function "triangle inequality"); if $n \in \mathbb{Z}$, then $\left[\frac{x}{n}\right] = \left[\frac{[x]}{n}\right]$ and $[x+n] = [x]+n$.

Theorem 1. *Let p be prime. Let $l(n,p)$ be the largest integer such that p^l divides $n!$. Then the formula for $l(n,p)$ is*

$$l(n,p) = \sum_{j=1}^{\infty} \left[\frac{n}{p^j}\right]$$

Proof. We count.

$$l(n,p) = \sum_{m=1}^{n} \sum_{\substack{j=1 \\ p^j \mid m}}^{\infty} 1 = \sum_{j=1}^{\infty} \sum_{\substack{m=1 \\ p^j \mid m}}^{n} 1 = \sum_{j=1}^{\infty} \left[\frac{n}{p^j}\right]$$

If this is confusing, here is a brief explanation of the middle two expressions:

1. The inner summand is the highest power α_i of p that is a factor of m_i. Sum this over all m_is in $1 \ldots n$ to get the highest power of p that is a factor of $n!$. I.e., if p^{α_i} is the highest power factor of p in m_i, then $p^{\alpha_1} p^{\alpha_2} \ldots p^{\alpha_n} = p^{\sum_{m=1}^{n} \alpha_m} \Rightarrow l = \sum_{m=1} \alpha_m$.

2. This assumes a different interpretation of the problem. Instead of summing over all the powers of p within each m, sum over all of the ms divisible by p_j, and then sum over all of the js. This should give you the same result. The benefit is that the inner sum is easily representable with the floor function.

$\square$

Corollary 1.1. $l(n,p) \leq \left[\frac{n}{p-1}\right]$

Corollary 1.2. *For $m, n \in \mathbb{Z}$, $m \geq n \geq 0$, $\binom{m}{n}$ is an integer. Moreover, if $n_1 + n_2 + \cdots + n_k = m$, then the multinomial coefficient $\binom{m}{n_1, n_2, \ldots, n_k}$ is an integer.*

Proof. Express $m!$, $n!$, and $(m-n)!$ in their prime-factor representations over the same set of primes $\{p_i\}$.

$$m! = \prod_i p_i^{\alpha_i},\ n! = \prod_i p_i^{\beta_i},\ (m-n)! = \prod_i p_i^{\gamma_i},\ n!(m-n)! = \prod_i p_i^{\beta_i+\gamma_i}$$

From the "triangle inequality" for the floor function, observe that

$$\left[\frac{m}{p^j}\right] \geq \left[\frac{n}{p^j}\right] + \left[\frac{m-n}{p^j}\right]$$

By (Theorem 1), for every prime p_i of $\{p_i\}$, $\alpha_i = l(m, p_i)$, $\beta_i = l(n, p_i)$, and $\gamma_i = l(m-n, p_i)$. By the above inequality and the formula for l, $\alpha_i \geq \beta_i + \gamma + i$; i.e., the power α_i of each prime in the prime factorization of $m!$ is at least the power $\beta_i + \gamma_i$ of the same prime in the factorization of $n!(m-n)!$. Thus $n!(m-n)!|m!$. $\square$

Multiplicative functions

Definition 2. *A real function f defined over the positive integers is said to be multiplicative if $f(m)f(n) = f(mn)$ $\forall m, n$ s.t. $(m,n) = 0$.*

Theorem 3. *If f is multiplicative, either f is identically zero or $f(1) = 1$.*

Note that it is often useful to illustrate properties of multiplicative functions by using the fundamental theorem of arithmetic to factor any positive integer into mutually coprime factors.

Notes on divisors of products of coprime integers: Let $m, n \in \mathbb{Z}_+$. Then it is easy to show

$$D = \{d : d|mn\} = \{xy : (x,y) \in \{x : x|m\} \times \{y : y|n\}\}$$

is the set of divisors of mn. If $(m,n) = 1$, then $|D| = |\{x : x|m\}| \times |\{y : y|n\}|$; i.e., every factor of mn is *uniquely* factorable into the product of one divisor of m and one divisor of n. This allows the rewriting of the operation

$$\sum_{d|mn} f(d) = \sum\sum_{(x,y)\in\{x:x|m\}\times\{y:y|n\}} f(xy) = \sum_{x:x|m}\sum_{y:y|n} f(xy)$$

where the summation may be replaced by any other aggregate operation over a set, and f is a generic function defined over positive integers.

Here is the proof of uniqueness of factorization. By the fundamental theorem of arithmetic,

$$m = \prod_{i=1}^{k} p_i^{\alpha_k},\ n = \prod_{j=1}^{l} q_j^{\beta_j}$$

where p_i, q_j are prime for all i, j. Since $(m, n) = 1$, $\{p_i\} \cap \{q_j\} = \varnothing$. Let $a|m$, $b|n$; then

$$a = \prod_{i=1}^{k'} p_i^{\alpha_i},\ b = \prod_{j=1}^{l'} q_j^{\beta_j},\ ab = \prod_{i=1}^{k'} p_i^{\alpha_i} \prod_{j=1}^{l'} q_j^{\beta_j}$$

Since this factorization of ab is unique and $\{p_i\}, \{q_i\}$ are disjoint, it is clear that there is no way to factor this into the product of one divisor of m (the product of some subset of $\{p_i\}$) and one divisor of n (the product of some subset of $\{q_i\}$) except by $a \cdot b$.

Theorem 4. *Let f be a multiplicative function, and define g to be*

$$g(n) = \sum_{d|n} f(d)$$

Then g is multiplicative.

Proof. Let $(m, n) = 1$. Then we can split the single sum over the factors of mn into a double sum over the factors of m and the factors of n (proved above).

$$g(mn) = \sum_{d|mn} f(d) = \sum_{x|m} \sum_{y|n} f(mn) = \sum_{x|m} f(m) \sum_{y|n} f(n) = g(m)g(n)$$

□

Euler's totient function (Sylvester)

Theorem 5 (Euler's totient function)**.**

$$\phi(n) = n \prod_{p_i|k} \left(1 - \frac{1}{p_i}\right)$$

where p_i represents the set of unique prime factors of n.

Proof. Based on an argument provided by Sylvester, and not requiring showing that the totient is multiplicative beforehand. We work backwards, using the known result and showing that it is correct. By expanding the product, we see that this is equivalent to

$$(\phi(n)) = n - \sum_{p_r|n} \frac{n}{p_r} + \sum_{\substack{p_r p_s|n,\\ r<s}} \frac{n}{p_r p_s} - \cdots$$

Note that $\frac{n}{p_r}$ denotes the number of numbers in $1 \dots n$ are divisible by p_r, $\frac{n}{p_r p_s}$ is the number of numbers in that range divisible by $p_r p_s$, and so on. We may reformulate this counting into instead summing over each number m in that range $1 \dots n$, and counting the number of primes or products of primes that divide m. Let $l(m) = |\{p : p|m,\, p|n\}|$; i.e., $l(m)$ denotes the number of common prime factors of m and n and $l(m) = 0 \iff m, n$ are coprime. Thus the above expression is equivalent to

$$(\phi(n)) = \sum_{m=1}^{n} \left(1 - \sum_{\substack{r \\ p_r|m}} 1 + \sum_{\substack{r>s \\ p_r p_s|m}} 1 - \cdots \right)$$

The inner sums are equivalent to combinations over common prime factors of m and n, thus

$$(\phi(n)) = \sum_{m=1}^{n} \left(1 - \binom{l(m)}{1} + \binom{l(m)}{2} - \cdots \right) = \sum_{m=1}^{n} \sum_{r=1}^{l(m)} (-1)^r \binom{l(m)}{r}$$

Note that the inner summation is of the form of a binomial power expansion, i.e.,[1]

$$\sum_{b=0}^{a} (1)^{a-b}(-1)^b \binom{a}{r} = (1-1)^a = \begin{cases} 1 & \text{if } a = 0 \\ 0 & \text{else} \end{cases}$$

Thus

$$c(m) = \sum_{r=1}^{l(m)} (-1)^r \binom{l(m)}{r} = \begin{cases} 1 & \text{if } m, n \text{ coprime} \\ 0 & \text{else} \end{cases}$$

and $c(m)$ is a simple indicator of whether m, n are coprime. This finally simplifies the totient formula down to a clearly correct statement: $\phi(n)$ counts the integers in $1 \dots n$ coprime to n, thus concluding the proof.

$$(\phi(n)) = \sum_{m=1}^{n} c(m)$$

□

Note that, while this proof of the formula for the totient function does not require its multiplicativity, the formula itself may be used to prove its own multiplicativity.

Theorem 6. *Define $g(n)$ as follows:*

$$g(n) = \sum_{d|n} \phi(n)$$

Then $g(n) = n$.

[1] The following formulation includes first term 1 into the summation, but this introduces the somewhat-iffy 0^0 case. You can see that the value of this summation is clearly 1 when $l(m)$ is zero, as you have no primes to choose from in the pre-simplified version. So this is absolutely correct (and often is, working outside of an analysis context).

Proof. By (Theorem 4), the summation on the left is a multiplicative function. Thus it may be broken up over the prime factorization of n as follows:

$$g(n) = \sum_{d|n} \left((n =) \prod_{p_i|d} p_i^{\alpha_i} \right) = \prod_{p_i|d} \left(\sum_{d|p_i^{\alpha_i}} \phi(p_i^{\alpha_i}) \right)$$

The inner summation can be computed using (Lemma 7):

$$\sum_{d|p_i^{\alpha_i}} \phi(p_i^{\alpha_i}) = 1 + (p-1) + (p^2 - p) + \cdots + (p^{\alpha_i} - p^{\alpha_i - 1}) = p^{\alpha_i}$$

Thus the product in $g(n)$ turns back into the prime factorization for n. Thus

$$g(n) = \prod_{p_i|d} p_i^{\alpha_i} = n$$

□

Euler's totient function (based on multiplicativity)

Lemma 7 (Totient function on prime powers)**.** *Let p be prime. Then*

$$\phi(p^n) = p^n - p^{n-1}$$

Proof. Since p^n has only one unique prime factor p, p must divide any number in $1 \ldots p^n$ that shares a common prime factor with p^n. There are $p^n/p = p^{n-1}$ such numbers. Thus the number of numbers in $1 \ldots p^n$ that are coprime with p^n is $p^n - p^{n-1}$. □

Corollary 7.1. *Let p be prime. Then $\phi(p) = p - 1$.*

For now, assume that the totient function is multiplicative (which must be shown later). This, along with (Lemma 7), makes the proof for the totient function very simple.

Theorem 8 (Euler's totient function (based on multiplicativity))**.** *Since the totient function is multiplicative and is defined on powers of primes (in a relation given by (Lemma 7)),*

$$\phi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

Proof. By the fundamental theorem of arithmetic, n is uniquely factorable into a product of its primes, i.e.,

$$n = \prod_{p_r|n} p_r^{\alpha_r}$$

$$\phi(n) = \phi\left(\prod_{p_r|n} p_r^{\alpha_r}\right) = \prod_{p_r|n} \phi(p_r^{\alpha_r}) = \prod_{p_r|n} p_r^{\alpha_r}\left(1 - \frac{1}{p_r}\right)$$
$$= \left(\prod_{p_r|n} p_r^{\alpha_r}\right)\left(\prod_{p_r|n}\left(1 - \frac{1}{p_r}\right)\right) = n\prod_{p|n}\left(1 - \frac{1}{p}\right)$$

□

This proof is clearly much simpler and straightforward than that in the previous section, and doesn't require the clever interpretation and manipulation of counting principles. However, we still need to show multiplicativity of the totient function to finish this proof. This can be done using the formula of the totient function (if proved already by some other manner, such as in the previous section) or using the Chinese remainder theorem (this will be proved later).

The Möbius function $\mu(n)$

The following result is not shown in the book and may be obvious, but it was not immediately apparent to me, so here it is.

Lemma 9. *Let f be an arithmetic function. Then*

$$\sum_{d|n} f(d) = \sum_{d|n} f\left(\frac{n}{d}\right)$$

Proof. The condition for d is symmetric to a condition for $\frac{n}{d}$; i.e., $n|d \iff \frac{n}{d}|d$. The result follows by substituting $d = \frac{n}{d'}$:

$$\sum_{d|n} f(d) = \sum_{\frac{n}{d'}|n} f\left(\frac{n}{d'}\right) = \sum_{d'|n} f\left(\frac{n}{d'}\right)$$

. □

Definition 10. *Let $n \in \mathbb{Z}^+$ have k distinct prime factors. Define the Möbius function $\mu(x)$ as follows:*

$$\mu(x) = \begin{cases} 1 & \textit{if } n = 1 \\ 0 & \textit{if } n \textit{ has any repeated prime factors} \\ (-1)^k & \textit{else} \end{cases}$$

μ is multiplicative: if $(m, n) = 1$, where m has k_1 distinct prime factors and n has k_2 distinct prime factors, then if m (or n) has a repeated prime factor, then the product will also have a repeated prime factor and $0 \cdot \mu(n) = \mu(m)\mu(n) = \mu(mn) = 0$; otherwise, all of the prime factors will be unique and $(-1)^{k_1}(-1)^{k_n} = \mu(m)\mu(n) = \mu(mn) = (-1)^{k_1+k_2}$. The convention that $\mu(1) = 1$ further makes the Möbius function fit well with multiplicative properties.

Lemma 11. *Define $v(n)$ as follows:*

$$v(n) = \sum_{d|n} \mu(x)$$

Then

$$v(n) = \begin{cases} 1 & \textit{if } n = 0 \\ 0 & \textit{if } n > 0 \end{cases}$$

The proof for this lemma is immediate.

Theorem 12 (Möbius inversion formula)**.** *Let $g(x) = \sum_{d|n} f(d)$, where f is an arithmetic function. Then*

$$f(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right)$$

(i.e., f is the Dirichlet convolution of its sum-function g and μ) and vice versa.

Proof. ($\Rightarrow$) We use (Lemma 9) and manipulate equivalent indexing to achieve the result:

$$\sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \sum_{d'|d} \mu\left(\frac{n}{d}\right) f(d') = \sum_{d'|d|n} \mu\left(\frac{n}{d}\right) f(d')$$

$$= \sum_{d'|n} f(d') \sum_{d'|d|n} \mu\left(\frac{n}{d}\right)$$

Note that $d'|d|n \iff \frac{n}{d}|\frac{n}{d'}$, so this becomes

$$= \sum_{d'|n} f(d') \sum_{\frac{n}{d}|\frac{n}{d'}} \mu\left(\frac{n}{d}\right) = \sum_{d'|n} f(d') v\left(\frac{n}{d'}\right)$$

Since $v(n/d') = 1 \iff n = d'$ by (Lemma 11), this simplifies to

$$f(n) \cdot 1 + \sum_{\substack{d'|n \\ d' \neq n}} f(d') \cdot 0 = f(n)$$

Proof. ($\Leftarrow$) Let f be defined in the following form:

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d)$$

Then

$$\sum_{d|n} f(d) = \sum_{d|n} \sum_{d'|d} \mu\left(\frac{d}{d'}\right) g(d') = \sum_{d'|d} g(d') \sum_{d'|d|n} \mu\left(\frac{d}{d'}\right) = \sum_{d'|n} g(d') \sum_{\frac{d}{d'}|\frac{n}{d'}} \mu\left(\frac{n}{d'}\right)$$

$$= \sum_{d'|n} g(d') v\left(\frac{n}{d'}\right) = g(n)$$

(This proof is very similar to that in the forward direction, so some intermediate steps and explanation are not shown here.) □

Some of the reindexing from the book was difficult for me to follow and understand. Online resources such as `https://math.berkeley.edu/~stankova/MathCircle/Multiplicative.pdf` and `https://math.stackexchange.com/a/1757370/96244` were helpful to me.

Theorem 13 (Totient-Möbius relationship)**.**

$$\phi(n) = n \sum_{d|n} \frac{\mu(d)}{d}$$

Proof. (Using totient formula) Expanding the totient formula in the same way as in the proof of (Theorem 5), we get

$$\phi(n) = n\left(1 - \sum_{p_r|n} \frac{1}{p_r} + \sum_{\substack{p_r p_s|n, \\ r<s}} \frac{1}{p_r p_s} - \cdots\right)$$

$$= n\left(\frac{\mu(1)(=(-1)^0)}{1} + \sum_{p_r|n} \frac{\mu(p_r)(=(-1)^1)}{p_r} + \sum_{\substack{p_r p_s|n \\ r<s}} \frac{\mu(p_r p_s)(=(-1)^2)}{p_r p_s} + \cdots\right)$$

$$= n \sum_{d|n} \frac{\mu(d)}{d}$$

In the original formulation of the totient formula, we are only counting factors of p that have only unique prime factors – this is accommodated nicely for because $\mu(n) = 0$ for all of the other factors (those with repeated prime factors), and thus are implicitly accounted for in this summation.

Proof. (Using Möbius inversion formula) The sum-function $g(n) = \sum_{d|n} \phi(n)$ is equal to n by (Theorem 6). Apply the Möbius inversion formula:

$$\phi(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \frac{\mu(d)}{d} = n \sum_{d|n} \frac{\mu(d)}{d}$$

□

Divisors and the Euclidean Algorithm

Jonathan Lam

February 1, 2020

DEF Let $a, b \in \mathbb{Z}, a \neq 0$. Then a **divides** b (den. $a|b$) if $\exists q \in \mathbb{Z} : aq = b$

(Note that $\forall a \neq 0, a|0$).

THM (Division algorithm) $\forall a, d \in \mathbb{Z}, d \neq 0 \; \exists! \; q, r \in \mathbb{Z} : a = dq + r, \; 0 \leq r < d$.

THM (Well-ordering principle) For any nonempty set $S \subseteq \mathbb{Z}_+$, S has a (unique) minimum element.

DEF (gcd) $\forall a, b \in \mathbb{Z}_+, d = \gcd(a, b)$ if $d|a$, $d|b$, and $d'|a \wedge d'|b \Rightarrow d'|d \; \forall d' \in \mathbb{Z}_+$.

(I.e., d divides both a and b, and all divisors of both a and b also divide d. Clearly, d is the maximum of all divisors of both a and b.)

THM (Alternate definition of the gcd) Define

$$S := \left\{ ax + by \; (\in \mathbb{Z}_+) : \begin{array}{r} a, b \in \mathbb{Z}_+ \\ x, y \in \mathbb{Z} \\ ax + by > 0 \end{array} \right\}$$

Then, for some $a, b \in \mathbb{Z}_+$, define

$$d := \min S$$

Then d exists, and

$$\gcd(a, b) = d$$

PF Three things need to be proved: (1): existence of d; (2): $d|a$ and $d|b$; (3): if $d'|a$ and $d'|b$, then $d'|d$. (2) and (3) are the hypotheses for the definition of gcd.

1. **Existence of** d
   S is a nonempty subset of $\mathbb{Z}$; by the well-ordering principle, it has a minimum element. Thus d exists (and is unique).

2. d **divides** a **and** b

   By the division algorithm, $a = dq + r,\ q, r \in \mathbb{Z},\ r < d$. To show that $d|a$, we have to show that $r = 0$. We prove this by contradiction: assume $r > 0$.

   $$d \in S \Rightarrow \exists x, y \in \mathbb{Z} : ax + by = d$$

   $$d > r = a - dq = a - (ax + by)q = a(q - xq) + b(yq) \in S \geq d \Rightarrow\perp$$

   By the contradiction, $r = 0$. The same logic applies to show that $d|b$.

3. **Any divisor** d' **of both** a **and** b **also divides** d

   By hypothesis, $\exists h, k \in \mathbb{Z} : d'h = a, d'k = b$.

   $$d \in S \Rightarrow \exists x, y \in \mathbb{Z} : ax + by = d$$

   $$d = (d'h)x + (d'h)y = d'(hx + ky) \Rightarrow d'|d \blacksquare$$

THM Let $a, b \in \mathbb{Z}_+,\ b \neq 0$. Then $\gcd(a, b) = \gcd(b, r)$, where r is obtained by the division algorithm applied on a and b.

PF By division algorithm, $a = bq + r,\ q, r \in \mathbb{Z}$. Define D to be the set of integers that divide both a and b, and define D' to be the set of integers that divide both b and r. Suppose $c \in D$; i.e., c divides both a and b Then

$$c|bq \Rightarrow c|a - bq \Rightarrow c|r$$

Thus c is divides both b and r, and thus

$$c \in D' \Rightarrow D' \subseteq D$$

Conversely, suppose $c \in D'$, i.e., c divides both b and r. Then

$$c|bq \Rightarrow c|bq + r \Rightarrow c|a$$

Thus c divides both a and r, and thus

$$c \in D \Rightarrow D \subseteq D'$$

Thus $D = D'$. In particular, $\gcd(D) = \max D = \max D' = \gcd(D')$. $\blacksquare$

(Note that this theorem doesn't guarantee termination of the Euclidean algorithm; for this to be true, $a \geq b$ is a necessary condition).

ALG (Euclidean Algorithm) Define the following algorithm:

```
gcd(a, b) {
    if(a < b)
        swap(a, b)
    while(b != 0) {
        r = a mod b
        a = b
        b = r
    }
    return a
}
```

Alternatively, recursively:

```
// the driver assures that a<=b
gcd(a, b) {
    if(b = 0)
        return a
    return gcd(b, a mod b)
}
gcd_drv(a, b) {
    return (a < b) ? gcd(b, a) : (a, b)
}
```

Intuitive PF This is applying the above theorem to a and b, recursively. Since $b \leq a$ (in the algorithm after the appropriate swapping, not necessarily for the initial invocation), $r < \min\{a, b\}$, the size of the inputs (from (a, b) to (b, r)) are (strictly) decreasing with subsequent invocations of the function. This means it will eventually reduce to the base case $(c, 0)$; since $gcd(c, 0) = c = gcd(a, b)$ (by the above theorem), this algorithm is correct.

THM The equation $ax + by = m\, a, b, m \in \mathbb{Z}$, a and b not both 0, has a solution $(x, y) \in \mathbb{Z}^2$ iff $\gcd(a, b)|m$.

PF (⇐) Since $d = \gcd(a, b) \in S$ (S defined in earlier theorem), $\exists x', y' \in \mathbb{Z} : d = ax' + by'$. Since $d|m$, $kd = m \Rightarrow k(ax' + by') = m \Rightarrow a(kx') + b(ky') = m \Rightarrow (kx', ky')$ is a solution to the equation.

PF (⇒) Let $ax + by = m$, $dh = a$, $dk = b$. Then $(dh)x + (dk)y = d(hx + ky)m \Rightarrow d|m$. ∎

ALG Algorithm to find a solution to $sa + tb = \gcd(a, b)$. (When the right side of the equation is $m \gcd(a, b)$, then multiply the determined coefficients by m).

```
gcd_st(a, b) {
    if(a < b)
        swap(a, b)
    if(b = 0)
        return (gcd=a, s=1, t=0)
    (gcd, s', t') = gcd_st(b, a mod b)
    return (gcd=gcd, s=t', t=s'-t'*(a/b))
}
```

(Here, the division a/b represents the integer division quotient.) This algorithm works due to the recurrence relation where $\gcd(a, b) = sa + tb = (t')a + (s' - t'(a/b))b$, where s' and t' are the integers such that $\gcd(b, r) = s'b + t'r$. The base case for this recursive relation is the case $\gcd(d, 0)$, which can be expressed as $(1)d + (0)0 = d$.

Primes stuff Fundamental Theorem of Arithmetic, Infinite Primes, GCD/LCM in terms of Prime Factorization, LCM in terms of GCD

MA352 Midterm Study Guide

Jonathan Lam

March 18, 2020

Contents

1 Sets

This can be seen as a boolean algebra $(S, \cup, \cap, \bar{\ }, \varnothing, U)$.

1.1 Properties

- Involution: $\bar{\bar{A}} = A$
- Absorption: $A \cup (A \cup B) = A \cup B$, $A \cap (A \cap B) = A \cap B$
- Bound: $A \cup U = U$, $A \cap \varnothing = \varnothing$
- Idempotent: $A \cup A = A$, $A \cap A = A$
- Complement: $A \cup \bar{A} = U$, $A \cap \bar{A} = \varnothing$

- Identity: $A \cup \varnothing = A$, $A \cap U = A$
- (Distributive law both ways)
- 0/1: $\bar{\varnothing} = U$, $\bar{U} = \varnothing$
- De Morgan's: $\bar{A} \cup \bar{B} = \overline{A \cap B}$, $\bar{A} \cap \bar{B} = \overline{A \cup B}$

A collection S of nonempty subsets (i.e., S is a set of sets) of X is said to be a partition of X if every element in X belongs to exactly 1 member o S (i.e., S pairwise disjoint, but $\cup_i S = X$).

If X and Y are sets, define the Cartesian product $X \times Y = \{(x, y) : x \in X, y \in Y\}$.

2 Relations

Define $\varnothing = R \subseteq X \times Y$ to be a relation from X to Y (any nonempty subset of the Cartesian product). If from a set to itself, call it a relation on X. Dom$(R)) \subseteq X$ set of elements $x \in X$ s.t. $(x, y) \in R$; the analogous definition also exists for Range(R). The inverse is the set $R^{-1} = \{(y, x) : (x, y) \in R\} \subseteq Y \times X$.

Can define function composition, $R_2 \circ R_1$ is R_2 composed with R_1. A relation is a function if Dom$(R) = X$, and if it is injective.

Injective, surjective, bijective.

Binary and unary operators.

Denote operators on elements of sets with (S, op). (E.g., $(\mathbb{R}, +)$).

Can represent a relation on a set with a digraph.

2.1 Classifications of relations

These all apply to relations on a set, not from one set to another.

- Reflexive: $(x, x) \in R \ \forall x \in X$
- Symmetric: $(x, y) \in R \Rightarrow (y, x) \in R$
- Antisymmetric: $x \neq y, (x, y) \in R \Rightarrow (y, x) \notin R$ (in other words, $(x, y), (y, x) \in R \Rightarrow x = y$)
- Transitive: $(x, y), (y, z) \in R \Rightarrow (x, z) \in R$
- Partial ordering: reflexive, antisymmetric, and transitive
- Total ordering: partial ordering, and every pair of elements is comparable
- Equivalence: reflexive, symmetric, and transitive

2.2 Equivalence relations

Let R be an equivalence relation on X. Then the equivalence class of $a \in X$ is $[a] = \{b : (a, b) \in R\}$.

The set of equivalence classes of X is a partition of X.

2.3 Matrices of relations

The matrix of relation is a transformation matrix. Each column is the transformation of an input element to the output elements, where 1 denotes a relation and 0 doesn't. I.e., line up the input set along the horizontal direction, output direction along the vertical direction, draw 1's where relations happen. Composition is multiplication, just like a transformation matrix, and multiplication by an element of the input set gives you the elements it is related to.

3 Mathematical induction

1. Basis step: Prove that $S(1)$ is true.
2. Inductive step: Prove that $S(k) \to S(k+1)$ is true.

Sometimes you can avoid mathematical induction; not the only way to prove things like these (e.g., geometric series, or using a similar, known mathematical rule and applying some (usually linear) rule like differentiation.)

4 Groups

... TODO ...

5 Modular arithmetic and elementary number theory

5.1 Basic rules of divisibility and modular arithmetic

Division algorithm:

$$\forall a, b \in \mathbb{Z},\ b > 0\ \exists! q, r \in \mathbb{Z} : a = bq + r,\ 0 \leq r < b$$

$$a|b \Rightarrow a|nb$$
$$a|b, c|d \Rightarrow (a+c)|d$$

$$a_1 \equiv a_2 \mod b \Leftrightarrow r_1 = r_2$$

where r_1 and r_2 come from the division algorithm of a_1 and a_2 with b.

$$a \equiv b \mod n, c \equiv d \mod n \Rightarrow a + b \equiv (b+c) \mod n$$

5.2 GCD and LCM

$$a = \prod_i^k p_i^{\alpha_i},\ b = \prod_i^k p_i^{\beta_i}$$

where $\{p_i\}$ denotes the set of distinct prime factors that divide a or b. Then:

$$\gcd(a, b) = \prod_i^k p_i^{\min \alpha_i, \beta_i}$$

$$\text{lcm}(a, b) = \prod_i^k p_i^{\max \alpha_i, \beta_i}$$

$$\gcd(a, b)\text{lcm}(a, b) = ab$$

Another way to characterize gcd is: for any integer d s.t. $d|a$ and $d|b$, then $d|\gcd(a, b)$.

5.3 Euclid's algorithm

To find $\gcd(a, b)$: let $m = \max(a, b)$, $n = \min(a, b)$. If $n = 0$, then $\gcd(a, b) = \gcd(m, n) = \gcd(m, 0) = m$. If not, then $\gcd(a, b) = \gcd(n, m \mod n)$.

5.3.1 Linear diophantine equations

Given $a, b \in \mathbb{Z}$ linear Diophantine equations $ax + by = c,\ x, y \in \mathbb{Z}$ always has a solution x, y if $\gcd(a, b)|c$, and Euclid's algorithm can help discover it. Namely, this involves the recurrence relation $x = y'$, $y = x' - y'q$ at any given step; at the basis step, $x = 1$, $y = 0$ (since $1(m) + 0(0) = \gcd(a, b)$ at the last step of Euclid's algorithm).

It's easier to understand by going the full depth of Euclid's algorithm, and then expressing $\gcd(a)b$ as a linear combination of the m and n from that step; again, on the bottom-most step, $1(m) + 0(n) = m = \gcd(a, b)$; work your way up from here. However, this can often be found by inspection.

This gives you a particular solution of x, y. To find the general solution, find the associated homogeneous solution and go from there. I.e., solve $ax = -by$. Thus $x = \frac{-b}{\gcd(a,b)}t$, $y = \frac{a}{\gcd(a,b)}t$. Thus the general solution is

$$x = x_0 - \frac{b}{\gcd(a, b)}t,\ y = y_0 + \frac{a}{\gcd(a, b)}t$$

5.3.2 Computing inverse modulo n

In general, $ax \equiv 1 \mod p$ has $\gcd(a, p)$ solutions. I.e., a has a unique solution to this equation $x = a^{-1}$ (inverse) iff $\gcd(a, p) = 1$.

MA352 – Pset 3

Jonathan Lam

April 16, 2020

Thief gang *A gang of 19 thieves has a pile of coins containing fewer than 8000 coins. They have to divide the pile evenly but there are 9 coins left over. As a result, a fight breaks out and one of the thieves is killed. They try to divide the pile again, and now they have 8 coins left over. Again, they fight, and again, one of the thieves dies and once more, they try to divide the pile but now they have 3 coins left.*

1. *How many coins are in the pile?*

$$x \equiv 9 \pmod{19}$$

$$x \equiv 8 \pmod{18}$$

$$x \equiv 3 \pmod{17}$$

(Let the i-th congruence relation is of the form $x \equiv a_i \pmod{n_i}$.)

$$N = \prod_{i=1}^{3} n_i = 5814$$

The solution to this congruence relation $(\bmod N)$ is:

$$x = \left[\sum_{i=1}^{3} a_i y_i z_i\right] \bmod N$$

where $y_i = N/n_i$ and $z_i = y_i^{-1} \bmod n_i$. Solving for y_i values by division and z_i by the extended Euclidean algorithm:

$$y_1 = 18 \times 17 = 306$$

$$z_1 = 306^{-1} \bmod 19$$

(Extended Euclidean algorithm:)

$$306 z_1 \equiv 1 \bmod 19$$

$$306 z_1 + 19w = 1$$

$$306 = 19 \times 16 + 2, \qquad 2 = 306 - 19 \times 16$$
$$19 = 2 \times 9 + 1, \qquad 1 = 19 - 2 \times 9$$

$$1 = 19 \times 1 - (306 - 19 \times 16) \times 9 = 19 \times 145 + 309 \times (-9)$$
$$z_1 \equiv -9 \equiv 10 \pmod{19}$$

Use the same method to solve for y_2, z_2, y_3, z_3 (not shown here):

$$y_2 = 323, \qquad z_2 = 17$$
$$y_3 = 342, \qquad z_3 = 9$$

$$x = \sum_{i=1}^{3} a_i y_i z_i = 9 \times 306 \times 10 + 8 \times 323 \times 17 + 3 \times 342 \times 9 = 80702$$
$$x \mod 5814 = 5120$$

Since the next value of x that would solve this system of congruences would be $2 \times 5120 = 10240 > 8000$, this is the unique solution of x.

2. *If they continue this process of fighting, losing one thief and redividing, how man thieves will be left when the pile is finally divided evenly with no remainder?*
   Since $16|5120$, 16 thieves will be left.

Conductor .

1. *Find* $\gcd(5, 8)$. Using (extended) Euclidean algorithm:

$$8 = 5 \times 1 + 3, \qquad 3 = 8 - 5 \times 1$$
$$5 = 3 \times 1 + 2, \qquad 2 = 5 - 3 \times 1$$
$$3 = 2 \times 1 + 1, \qquad 1 = 3 - 2 \times 1$$
$$2 = 1 \times 2 + 0$$

   Thus

$$\gcd(5, 8) = 1$$

2. *Find* x, y *s.t.* $5x + 8y = \gcd(5, 8)$. Using equations from the previous question:

$$\begin{aligned} 1 &= 3 \times 1 - 2 \times 1 \\ &= 3 \times 1 - (5 - 3 \times 1) \times 1 \\ &= 3 \times 2 + 5 \times (-1) \\ &= (8 - 5 \times 1) \times 2 + 5 \times (-1) \\ &= 8 \times 2 + 5 \times (-3) \end{aligned}$$

   Thus a particular solution is:

$$(x, y) = (-3, 2)$$

   and the general solution is

$$(x, y) = (-3 - 8t, 2 + 5t), \ t \in \mathbb{Z}$$

3. *It is a fact that given two nonnegative integers a, b where $\gcd(a,b) = 1$, there is always a point beyond which every integer is representable as $ax + by$ where x and y are both nonnegative integers. The least such result is denoted $\text{cond}(a,b)$. Find $\text{cond}(5,8)$.*

   Since $\gcd(5,8) = 1$, we can use the result derived in the next section.

$$\text{cond}(5,8) = (5-1)(8-1) = 28$$

4. *Find the general formula for $\text{cond}(a,b)$ where $\gcd(a,b) = 1$.*

   Since $\gcd(a,b) = 1$, we know that there are infinitely many solutions to the linear Diophantine equation:

$$ax + by = c, \ c \in \mathbb{Z}$$

   If $(x,y) = (x_0,y_0)$ is a particular solution, then we know that other particular solutions are of the form $(x,y) = (x_0 + bt, y_0 - at), \ t \in \mathbb{Z}$. Thus, a number cannot be represent by a positive tuple (x,y) if $x_0 < 0$ and $y_0 < a$ (or, alternatively, if $y_0 < 0$ and $x < b$). The largest of such would be when $x = -1$, $y = a - 1$ (or, alternatively, when $x = b - 1$, $y = -1$; both give the same answer). If we plug this into the Diophatine equation, this gives us the largest integer *not* representable by a positive tuple, so $\text{cond}(a,b)$ is this number plus one:

$$\text{cond}(a,b) = (a(-1) + b(a-1)) + 1 = -a + ab - b + 1 = (a-1)(b-1)$$

5. *In the United Kingdom, chicken nuggets are sold in packs of 9 or 20. What is the largest number of chicken nuggets that you cannot buy?*

   The answer is $\text{cond}(9,20) - 1$. Since $\gcd(9,20) = 1$, we can use the result from the previous section.

$$\text{cond}(9,20) - 1 = (9-1)(20-1) - 1 = 151$$

PH214: Optics and Modern Physics

Jonathan Lam

March 22, 2020

Commentary on Prof. Yecko's and Prof. Debroy's wonderful lecture notes. No guarantees on correctness or completeness. Not meant to be standalone, but to supplement the lecture notes and their diagrams.

Contents

1 Maxwell's equations

1.1 General form

Illustrated previously in E&M. In order: Gauss's flux laws (electric and magnetic); Faraday's induction law; Ampere's current law. Shown in integral and differential forms.

$$\oiint_{\partial V} \vec{E} \cdot d\vec{s} = \frac{1}{\varepsilon} \iiint_V \rho \, dV \equiv \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon} \tag{1}$$

$$\oiint_{\partial V} \vec{B} \cdot d\vec{s} = 0 \equiv \nabla \cdot \vec{B} = 0 \tag{2}$$

$$\oint_C \vec{E} \cdot d\vec{l} = -\iint_A \frac{\partial \vec{B}}{\partial t} \cdot d\vec{s} \equiv \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{3}$$

$$\oint_C \vec{B} \cdot d\vec{l} = \mu \iint_A \left(\vec{j} + \varepsilon \frac{\partial \vec{E}}{\partial t} \right) \cdot d\vec{s} \equiv \nabla \times \vec{B} = \mu \left(\vec{j} + \varepsilon \frac{\partial \vec{E}}{\partial t} \right) \tag{4}$$

where ε is permittivity, μ is the permeability, and $\vec{j}$ is current density.

1.2 In a vacuum

In a vacuum, there is no charge density, and thus no current density. Thus Maxwell's equations in a vacuum (in differential form) simplify to

$$\nabla \cdot \vec{E} = 0 \tag{5}$$

$$\nabla \cdot \vec{B} = 0 \tag{6}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{7}$$

$$\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{8}$$

where ε_0 is the permittivity of free space and μ_0 is the permeability of free space.

1.3 Review of differential operators

1.3.1 The del operator ∇

Note that the nabla/del operator in the below forms are principally defined for use with Cartesian coordinates. In polar (cylindrical or spherical) coordinates, these operations have to be converted from these Cartesian equivalents. It can be treated as a "vector":

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \tag{9}$$

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \tag{10}$$

$$\nabla \cdot \vec{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{11}$$

$$\nabla \times \vec{A} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \tag{12}$$

The Laplacian is a second-order differential operator, and can be treated as a "scalar" operator obtained from the dot product of nabla with itself:

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{13}$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{14}$$

$$\nabla^2 \vec{A} = \left(\nabla^2 A_x, \nabla^2 A_y, \nabla^2 A_z \right) \tag{15}$$

1.4 Other useful identities

$$\vec{A} \cdot (\vec{A} \times \vec{B}) = 0 \tag{16}$$

$$\vec{A} \cdot (\vec{A} \times \vec{C}) = (\vec{A} \times \vec{B}) \cdot \vec{C} \tag{17}$$

$$\nabla \times \nabla f = 0 \text{ (curl of gradient field is zero)} \tag{18}$$

$$\nabla \cdot (\nabla \times \vec{A}) = 0 \text{ (divergence of curl field is zero)} \tag{19}$$

$$\nabla \times (\nabla \times \vec{A}) = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A} \tag{20}$$

$$\nabla \times (\vec{A} \cdot \vec{B}) = (\nabla \times \vec{A}) \cdot \vec{B} \tag{21}$$

1.5 Proving the wave equation from Maxwell's equations

The canonical form for a wave (what we desire to achieve):

$$\frac{\partial^2 \psi}{\partial t^2} = v^2 \frac{\partial^2 \psi}{\partial x^2} \tag{22}$$

Remembering the side of v makes sense since it is time over distance, canceling out the units. Assuming a vacuum, we start with (Eq. 7) and apply identity (Eq. 20):

$$\nabla \times \left(\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \right)$$

$$\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{\partial}{\partial t}(\nabla \times \vec{B})$$

$$\nabla^2 \vec{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \tag{23}$$

This is in the form of (Eq. 22) in the 3D case, with $c^2 = \frac{1}{\mu_0 \varepsilon_0}$. The same can be shown for $\vec{B}$.

Note that since we know that sinusoids are solutions to the wave equation, then any periodic $\vec{E}$ function is a solution for this $\vec{E}$ wave equation (since it can be decomposed into a linear combination of sinusoids using Fourier analysis); this would then force constraints on $\vec{B}$ (and vice versa for arbitrary $\vec{B}$ waves). If there is a pair of matching $\vec{E}$ and $\vec{B}$ oscillations in space that match both wave equations simultaneously and satisfies the Maxwell equations, then it is an EM wave.

1.6 Polarization of an EM wave

For now, assume that $\vec{E}$ (and $\vec{B}$) waves are transverse (i.e., a transverse electromagnetic wave (TEM); not proven yet; see (Eq: 35)); i.e., $\vec{E} \cdot \hat{k} = 0$. Assume the wave is traveling in the z direction. Then

$$\vec{E} = E_x(z,t)\hat{\imath} + E_y(z,t)\hat{\jmath} \tag{24}$$

Since $E_x\hat{\imath}$ and $E_y\hat{\jmath}$ are themselves waves, then E_x and E_y both fit the form:

$$f(z,t) = A\cos(k(z-ct)+\delta) \tag{25}$$

If $\delta_x = \delta_y$, then $\vec{E}$ is linearly polarized. If $\delta_x \neq \delta_y$, then $\vec{E}$ is elliptically polarized. If $\delta_x \neq \delta_y$ and $E_x = E_y$, then $\vec{E}$ is circularly polarized. We can also write $\vec{E}$ as:

$$\vec{E} = E\cos\theta\hat{\imath} + E\sin\theta\hat{\jmath} \tag{26}$$

and we call θ the polarization angle.

Polarization comes from the direction of acceleration of a charge, which will be seen in the scattering section.

1.7 Plane waves

We can express a simple (real) cosine wave as its phasor equivalent:

$$\vec{E} = E_0\cos(k(z-ct)+\delta)\hat{n} = \Re\left(E_0 e^{i(k(z-ct))}\right)\hat{n} \tag{27}$$

where E_0 is a real constant in the cosine form, and a complex constant (including the phase) in the phasor form, and $\hat{n}$ is the unit vector in the direction of the polarization. k is the wave number, and c is the speed of light.

Alternatively, we can express the phase as $i(kz-\omega t)$, recognizing that $\omega = kc$. (Also, $c = \omega/k$, and $k = \omega/c$.)

More generally, if a wave is propagating in direction $\hat{k}$ with wave number k, we define the wave vector $\vec{k}$ and express the wave as

$$\vec{E} = \Re\left(E_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)}\right)\hat{n} \tag{28}$$

1.7.1 Plane wave derivatives

Let $\vec{E}$ be a plane wave. Then:

$$\frac{\partial \vec{E}}{\partial t} = -i\omega\vec{E} \tag{29}$$

$$\nabla\cdot\vec{E} = i\vec{k}\cdot\vec{E} \tag{30}$$

$$\nabla\times\vec{E} = i\vec{k}\times\vec{E} \tag{31}$$

In other words, we have correspondences between the differential operators and multiplication (for plane waves). In particular:

$$\frac{\partial}{\partial t} \leftrightarrow -i\omega \tag{32}$$

$$\nabla \leftrightarrow i\vec{k} \tag{33}$$

$$\nabla^2 = \nabla\cdot\nabla \leftrightarrow i\vec{k}\cdot i\vec{k} = -k^2 \tag{34}$$

Thus, to prove that plane waves are transverse:

$$\nabla\cdot\vec{E} = i\vec{k}\cdot\vec{E} = 0 \Rightarrow \vec{k}\perp\vec{E} \tag{35}$$

(The same result is true for $\vec{B}$.) To find the relative magnitudes of $\vec{E}$ and $\vec{B}$:

$$\nabla\times\vec{E} = i\vec{k}\cdot\vec{E} = -\frac{\partial\vec{B}}{\partial t} = i\omega\vec{B} \Rightarrow \vec{k}\times\vec{E} = \omega\vec{B} \tag{36}$$

$$\nabla\times\vec{B} = i\vec{k}\cdot\vec{B} = \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t} = -\frac{i\omega}{c^2}\vec{E} \Rightarrow \vec{k}\times\vec{B} = -\frac{\omega}{c^2}\vec{E} \tag{37}$$

Note that (Eq. 36) shows that $\vec{B}$ is also transverse since it results from a cross product with $\vec{E}$. The negative sign arises from a right-handed coordinate system. Since $\vec{k}$, $\vec{B}$, $\vec{E}$ are mutually orthogonal, the cross products become the products of the magnitudes, i.e.:

$$kE = \omega B \Rightarrow B = \frac{k}{\omega}E = \frac{1}{c}E \tag{38}$$

Note that this factor of $\frac{1}{c}$ is meaningless; it is more or less a consequence of our unit systems of $\vec{E}$ and $\vec{B}$ fields.

---

2 Poynting vector

2.1 Energy density of E and B fields

Denote energy density with u. These equations were derived in E&M.

$$u_E = \frac{1}{2}\varepsilon_0 \vec{E}\cdot\vec{E} = \frac{1}{2}\varepsilon_0 |\vec{E}|^2 \tag{39}$$

$$u_B = \frac{1}{2\mu_0}\vec{B}\cdot\vec{B} = \frac{1}{2\mu_0}|\vec{B}|^2 \tag{40}$$

For an E&M wave, at any point in space and moment in time, $u_E = u_B$, so total energy density is $u = u_E + u_B = 2u_E = 2u_B$.

2.2 Poynting vector

The Poynting vector points in the direction of (outward) energy flux. It can be thought of as the energy density multiplied by a velocity.

$$\vec{S} = \frac{1}{\mu_0}\vec{E}\times\vec{B} \tag{41}$$

To derive this, calculate the rate of change in energy density.

$$\frac{du}{dt} = \frac{\partial}{\partial t}\left(\frac{1}{\mu_0}\vec{B}\cdot\vec{B}\right) = \frac{1}{\mu_0}\frac{\partial\vec{B}}{\partial t}\cdot\vec{B} = \frac{1}{\mu_0}(-\nabla\times\vec{E})\cdot\vec{B} = -\nabla\cdot\left(\frac{1}{\mu_0}\vec{E}\times\vec{B}\right) = -\nabla\cdot\vec{S} \tag{42}$$

This makes sense from a conservation-of-energy perspective (this is called a continuity statement):

$$\frac{du}{dt} + \nabla\cdot\vec{S} = 0 \tag{43}$$

This statement can be read as: the rate of change of (internal) energy is equal to the energy flux (or divergence, or rate of loss of energy).
Time-averaging the Poynting vector:

$$\begin{aligned}\langle\vec{S}\rangle = \frac{1}{T}\int_0^T \frac{1}{\mu_0}\vec{E}\times\vec{B}\,dt = \frac{1}{T\mu_0}E_0B_0\int_0^T sin^2[\phi(t)]\,dt = \frac{1}{2\mu_0}E_0B_0 \\ = \frac{1}{2}c\varepsilon_0 E_0^2\hat{k}\end{aligned} \tag{44}$$

where $\hat{k}$ is the unit vector in the direction of the wave's propagation. Since $\langle u\rangle = \frac{1}{2}\varepsilon_0 E_0^2$, we obtain the "Suck" equation:

$$\langle\vec{S}\rangle = \langle u\rangle c\hat{k} \tag{45}$$

This is as we interpreted at the beginning of this section: the energy flux is an energy density being carried at the wave's propagation speed. The units for this are W/m^2; the rate of energy flux penetrating some patch of closed loop in space.

Lastly, we define intensity as the magnitude of $\langle\bar{S}\rangle$, i.e.,

$$I = |\langle\bar{S}\rangle| = \langle u\rangle c \tag{46}$$

2.3 Energy density as pressure

It makes sense that packing more energy in a unit volume may seem to give it more "pressure." Working dimensionally, this also works:

$$P = \frac{\mathrm{N}}{\mathrm{m}^2} = \frac{\mathrm{N} \cdot \mathrm{m}}{\mathrm{m}^2 \cdot \mathrm{m}} = \frac{\mathrm{J}}{\mathrm{V}} = u \tag{47}$$

This "light pressure" is theoretically able to be able to exert forces on objects, but its applicability is only theoretical right now (see "solar sails").

---

3 Radiation and scattering

After receiving an impulse from an EM wave, an electron radiates energy. We only study Thomson (free electron) and Rayleigh (electron bound to an atom) scattering, both of which are elastic scattering models.

3.1 Radiation

First, we talk about radiation due to an atom due to some acceleration. When radiation comes into contact with matter, it may be reflected, refracted, absorbed, or scattered. In this section we will discuss scattering models only.

3.1.1 Setup and fundamental results

An electron is at rest at the origin. It is briefly accelerated (wlog., in the positive x-direction) for a short interval of time τ, after which it moves at constant velocity $v \ll c$. We create a model to examine its $\vec{E}$ field at time $T \gg \tau$. Since EM waves travel at the speed of light, the EM wave updates in "rings" propagating outward from the particle at c.

The first boundary is at radius $R_0 = cT$; outside of this circle, an observer doesn't know the particle has moved. Since the acceleration stops after τ, the second important boundary is at $R_I = c(T - \tau)$; inside this circle, an observer sees the charge moving at a uniform speed. We make the approximation that these two circles are centered at the origin, since $\tau \ll T$.

Now, picture an $\vec{E}$ wave leaving the electron at angle θ. It travels in a straight line until the first boundary. Outside of the second boundary, however, an observer doesn't know the electron has moved from the origin yet, however, so the ray outside of the second boundary would be a straight line towards the origin, also roughly at angle θ. Thus, in the region between these two boundaries (where the information about the electron's acceleration is currently reaching), there is a "kink" in the electric field waves; we approximate this with a straight line. We can also approximate the curvature of the two circles to be small in this region, and assume them to be straight lines in the tangential direction.

Now we have a "rectangle" with side lengths E_{kink_r} and E_{kink_θ} (since $\vec{E}_{kink}$ is the diagonal). From the geometry and approximations given (now, let $\vec{E} = \vec{E}_{kink}$):

$$E_\theta = vT \sin\theta$$

$$E_r = c\tau$$

$$\frac{E_\theta}{E_r} = \frac{vT \sin\theta}{c\tau}$$

The following substitutions are also appropriate given the setup:

$$R = cT$$

$$a = \frac{v}{\tau}$$

From Coulomb's/Gauss's law for a point charge, we know that the radial $\vec{E}$ field is also:

$$E_r = \frac{1}{4\pi\epsilon_0} \frac{q}{R^2} \tag{48}$$

So we can rewrite E_θ as (using the appropriate substitutions):

$$E_\theta = \frac{qa \sin\theta}{4\pi\epsilon_0 c^2 R} \tag{49}$$

Note that, as R increases, $\vec{E}_\theta$ dominates over $\vec{E}_r$ ($E_\theta \propto R^{-1}$, not R^{-2} like radial $\vec{E}$). Thus, for reasonable distances, is practically a transverse wave with magnitude E_θ caused by an acceleration of the charge at the retarded time $-R/c$. This $\vec{E}_\theta$ is thus the $\vec{E}$ field of the radiated EM wave; we are not concerned with the radial component. Using (Eq. 38), we have:

$$B = \frac{qa \sin\theta}{4\pi\epsilon_0 c^3 R} \tag{50}$$

We can also find the Poynting vector associated with this EM wave. Using (Eq. 41), we can find the magnitude of the Poynting flux:

$$S = \frac{q^2 a^2 \sin^2\theta}{16\pi^2 \epsilon_0 r^2 c^3} \tag{51}$$

To recap, these are the magnitudes of the $\vec{E}$, $\vec{B}$, and $\vec{S}$ fields for the EM wave radiated by a charge accelerated at some retarded time, where $\vec{E}$ and $\vec{B}$ are in the tangential directions and mutually perpendicular, and $\vec{S}$ is in the radial direction.

3.1.2 Geometry of the radiation

The $\vec{E}$, $\vec{B}$, and $\vec{S}$ fields are clearly anisotropic (are a function of angle). Looking at the $\vec{S}$ vector, we can see that the power radiated is proportional to $\sin^2\theta$, and thus is sort of a weird donut-y shape. Specifically, in the direction of acceleration, there is no radiation; in the perpendicular direction, there is the maximum ("beaming") with magnitude

$$S = \frac{q^2 a^2}{16\pi^2 \epsilon_0 r^2 c^3} \tag{52}$$

which can be clearly seen from (Eq. 51).

If the donut is sitting flat on a table, then the $\vec{E}$ vector is tangent to the surface and points up or down; the $\vec{B}$ is tangent to the surface and points around the donut. Since the orientation of the donut is dictated by the direction of acceleration, the directions of $\vec{E}$, $\vec{B}$, and $\vec{S}$ are determined by the direction of acceleration; this is where polarization comes from.

The region of no radiation will show up again later when encountering Brewster's angle.

3.1.3 Larmor power

To get the total radiated power (Larmor power), we integrate the Poynting vector flux through some closed surface containing the charge. For simplicity, choose a sphere (and integrate using cylindrical coordinates):

$$P = \int_0^\pi (S)(2\pi r)(r\sin\theta)d\theta = \frac{q^2a^2}{8\pi\epsilon_0 c^3}\int_0^\pi \sin^3\theta\, d\theta = \frac{q^2a^2}{6\pi\epsilon_0 c^3} \tag{53}$$

This is the power, or total energy loss by an accelerated charge (in Watts).

3.2 Thomson scattering

Thomson scattering is the model used for free electrons. Imagine an incident $\vec{E}$ wave approaching a free electron, and the wave is traveling in the $\hat{x}$ direction. Let:

$$E = E_x = E_0\cos(kz - \omega t) = E_0 \cos\omega t$$

Then:

$$a = a_x = \frac{qE_0}{m}\cos\omega t \tag{54}$$

Plugging into (Eq. 53), we get:

$$P(t) = \frac{q^2a^2(t)}{6\pi\epsilon_0 c^3} = \frac{q^4E_0^2}{6\pi\epsilon_0 m^2 c^3}\cos^2\omega t$$

Time averaging, we get:

$$\langle P \rangle_{scatt_{th}} = \frac{q^4 E_0^2}{12\pi\epsilon_0 m^2 c^3} \tag{55}$$

This is the Thomson scattering by a single electron due to an incident $\vec{E}$ wave with amplitude E_0. We may want to express this in terms of the incident power; from (Eq. 44), we can do precisely this (and some simplifying) to obtain:

$$\langle P \rangle_{scatt_{th}} = \frac{8\pi}{3} \left(\frac{q^2}{4\pi\epsilon_0 mc^2} \right)^2 \langle S \rangle_{inc} \tag{56}$$

Note that this equation works dimensionally: we have a power equal to some area multiplied by a power flux. In particular, the quantity inside the parens is a length and is called the classical electron radius, r_0:

$$r_0 = \frac{e^2}{4\pi\epsilon_0 mc^2} \tag{57}$$

and this area is called the Thomson scattering cross-section:

$$\sigma_{th} = \frac{8\pi}{3} r_0^2 \tag{58}$$

So we can abbreviate the Thomson scattering using this new constant:

$$\langle P \rangle_{scatt_{th}} = \sigma_{th} \langle S \rangle_{inc} \tag{59}$$

Note that the scattering coefficient is intrinsic of the material and not of the intensity of the light; that is, for a given scatterer, the scattered power is proportional to the incident power flux.

3.2.1 Application: transmission through the Sun's corona

We can model the Sun's corona as a bunch of free electrons and protons. The protons are much weaker scatterers (their σ_{th} value is much smaller due to their larger mass), so we'll ignore their contribution. The change in total power after passing through a length dz through a portion of the Sun's atmosphere is

$$dP = -A\, d\langle S \rangle = NA\, dz\, \sigma \langle S \rangle$$

where A is the cross-sectional area of interest (we can see this quickly cancels out), and N is the number density of scatterers. The right hand side of this equation comes directly from (Eq. 59) through $NA\,dz$ electrons. Rearranging, we get a simple ODE, to which the solution is

$$\langle S \rangle = \langle S_0 \rangle \exp(-N\sigma z) \tag{60}$$

For the solar corona, we get $\langle S \rangle \approx 0.99995 \langle S_0 \rangle$, so very little of the Sun's total radiative power is scattered by its corona (and that is why it is so hard to see).

3.3 Rayleigh scattering

Rayleigh scattering uses a mechanical "electron on a spring" model of the atom, and is more accurate for electrons in matter. For simplicity, we can treat it in the undamped case, and then move onto the damped case. We see that it is similar to Thompson scattering: the radiated power is also proportional to the incident power flux, except now the proportionality "constant" is frequency-dependent (and thus not really a "constant").

3.3.1 Setup and fundamental results

We treat the electron as an undamped spring with driving force $qE_0 \cos \omega t$, similar to the original setup for Thomson scattering. However, the acceleration is not as simple: we have the additional restoring force term. Luckily this is still just a linear ODE that we can solve normally just like we would for a spring mechanics problem:

$$m\frac{d^2x}{dt^2} + m\omega_0^2 x = qE_0 \cos \omega t \tag{61}$$

Solving:

$$x = \frac{qE_0}{m(\omega_0^2 - \omega^2)} \cos \omega t \tag{62}$$

Solving for acceleration, we can plug this back into (Eq. 53) to get the Larmor power from Rayleigh scattering:

$$\frac{d^2x}{dt^2} = a = \frac{\omega^2}{\omega^2 - \omega_0^2} \frac{qE_0}{m} \cos \omega t \tag{63}$$

Compare this to (Eq. 54); the only difference is the frequency term. Since $S \propto a^2$, it can be seen from analogy to the derivation of Larmor power of Thomson scattering that the Larmor power for Rayleigh scattering is

$$\langle P \rangle_{scatt_{ray}} = \frac{8\pi}{3} \left(\frac{q^2}{4\pi\epsilon_0 mc^2} \right)^2 \left(\frac{\omega^2}{\omega^2 - \omega_0^2} \right)^2 \langle S \rangle_{inc} \tag{64}$$

If we express the scattering terms as a single coefficient, we get

$$\sigma_{ray} = \left(\frac{\omega^2}{\omega^2 - \omega_0^2} \right)^2 \sigma_{th} \tag{65}$$

and then we can express Rayleigh scattering as the simple equation just like (Eq. 59):

$$\langle P \rangle_{scatt_{ray}} = \sigma_{ray} \langle S \rangle_{inc} \tag{66}$$

3.3.2 Properties and results from frequency dependence

As ω_0 goes to zero, then the restoring force becomes negligible, and thus it should behave like a Thompson scattering. Indeed,

$$\lim_{\omega_0 \to 0} \sigma_{ray} = \sigma_{th} \tag{67}$$

On the other hand, for gases, $\omega_0 \gg \omega$, and we can make the approximation:

$$\langle P \rangle_{scatt_{ray}} \approx \left(\frac{\omega}{\omega_0}\right)^4 \sigma_{th} \langle S \rangle_{inc} \tag{68}$$

This fourth-power means that higher frequencies are scattered much more strongly than lower ones, and causes the phenomenon that blue light is scattered much more strongly by the atmosphere, hence causing it to appear blue.

3.3.3 Resonance and damping

The undamped model doesn't make sense, because it would imply extremely unstable conditions when resonance occurs. Thus it makes sense to introduce a damping term, making the original differential equation into

$$m\frac{d^2x}{dt^2} + m\beta\frac{dx}{dt} + m\omega_0^2 x = qE_0 \cos \omega t \tag{69}$$

The general equation for this is:

$$\vec{x} = \frac{q}{m((\omega_0^2 - \omega^2) + i\gamma\omega)}\vec{E} \tag{70}$$

This equation shows the movement of an electron relative to an atom; in other words, it shows the displacement of charges, or the dipole moment. In a static $\vec{E}$ field, then the driving frequency $\omega = 0$, and there is an average dipole moment in the same direction with magnitude:

$$\langle \vec{p} \rangle = q\langle \vec{d} \rangle = q\langle \vec{x} \rangle = \frac{q^2}{m\omega_0^2}\langle \vec{E}_{ext} \rangle \tag{71}$$

but we will not directly reference this solution to the undamped electron-on-a-spring equation until (Sec. 7.1) on absorption, which will deal with electrons responding to an incident $\vec{E}$ wave.

---

4 Dipoles and Potentials

We already know from integrating the $\vec{E}$ field that the electric potential is a scalar field:

$$\phi_E = -\nabla \phi_E \tag{72}$$

We will obtain the same for the "magnetic field strength", the $\vec{H}$ (not yet revealed).

$$\phi_H = -\nabla \phi_M \tag{73}$$

For a point charge, ϕ_E is already known:

$$\phi_E = \frac{q}{4\pi\epsilon_0}\frac{1}{r} \tag{74}$$

4.1 Dipoles and beyond

Electric dipoles are of interest because they are created when molecules distort in an electric field. I.e., displacing a charge is equivalent to adding a dipole. Define $\vec{d}$ as the vector from the vector from the negative charge to the positive charge in a dipole. Also, define $\vec{p}$, the dipole moment, to be

$$\vec{p} = q\vec{d} \tag{75}$$

There are electric also monopoles, quadrupoles, octupoles, etc. They can be in either a 3-D general form or a linear (axi-symmetric) form, but both have the same potential magnitudes. For a monopole, the potential is proportional to r^{-1}; for a dipole, it is r^{-2}; and so on. We deal primarily with the axi-symmetric forms.

4.2 Potential of a dipole

We already know from E&M that to calculate the voltage at any point in space, we add up the contributions from all of the charges (or poles). For a dipole, in general, if we let $\vec{r}$ be the positional vector of the point of interest w.r.t. to the center of the dipole, and define:

$$\vec{r}_+ = \vec{r} - \frac{1}{2}\vec{d},\ \vec{r}_- = \vec{r} + \frac{1}{2}\vec{d}$$

(i.e., r_+ is the distance between the point of interest to the positive charge, and likewise for r_- and the negative charge) then:

$$\phi_{dip} = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r_+} - \frac{1}{r_-}\right) \tag{76}$$

From the definition of $\vec{r}_+$ and $\vec{r}_-$, we can represent r_+^{-1} and r_-^{-1} as $|\vec{r} - \vec{\delta}|^{-1}$, where $\vec{\delta} = \pm\frac{1}{2}\vec{d}$. We can rewrite this in three ways:

$$\frac{1}{|\vec{r} - \vec{\delta}|} = \begin{cases} \frac{1}{r} - \vec{\delta} \cdot \nabla\left(\frac{1}{r}\right) + \cdots \\ \frac{1}{r}\left(1 - 2\left(\hat{\delta} \cdot \hat{r}\frac{\delta}{r}\right) + \frac{\delta^2}{r^2}\right)^{-1/2} \\ \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{\delta}{r}\right)^l P_l(\hat{\delta} \cdot \hat{r}) \end{cases} \tag{77}$$

where P_l is the lth Legendre polynomial. Note:

1. The former is the 3-D Taylor expansion.
2. The second is the law of cosines in the denominator.
3. The latter is a series expansion of the second. (Note that Legendre polynomials result from performing Gram Schmidt orthonormalization on the polynomial standard ordered basis (i.e., an orthonormal set created from

$\{1, x, x^2, x^3, \ldots\}$), so you can write any function as an expansion (linear combination) of the Legendre polynomials by using the inner product of the function with a Legendre polynomial as the coefficicent.)

Using the first two terms of the first representation, we can see that the potential of a dipole is:

$$\begin{aligned}\phi_E &= \frac{q}{4\pi\epsilon_0}\left(\frac{1}{|\vec{r} - \frac{1}{2}\vec{d}|} - \frac{1}{|\vec{r} + \frac{1}{2}\vec{d}|}\right) \\ &\approx \frac{q}{4\pi\epsilon_0}\left(\left[\frac{1}{r} + \frac{1}{2}\vec{d}\cdot\nabla\left(\frac{1}{r}\right)\right] - \left[\frac{1}{r} - \frac{1}{2}\vec{d}\cdot\nabla\left(\frac{1}{r}\right)\right]\right) \\ &= \frac{q}{4\pi\epsilon_0}\left(\vec{d}\cdot\nabla\left(\frac{1}{r}\right)\right) = \frac{1}{4\pi\epsilon_0 r^2}\vec{p}\cdot\hat{r} \qquad (78)\end{aligned}$$

This is the single-term approximation for a dipole moment, and it is clearly proportional to r^{-2}. However, there are more terms in the expansion, but this approximation is considered good enough. This expression is not an approximation for a "point dipole"; i.e., if $d \to 0$ but $\vec{p}$ is finite (i.e., then $q \to \infty$).

4.3 Potentials inside/outside a distribution of charges

With the third representation from (Eq. 77), we may express potentials within a net-neutral distribution of charges. For a point outside of the charge distribution, then the potential can be expressed as:

$$\phi_{out} = \frac{1}{4\pi\epsilon_0}\sum_{l=0}^{\infty}\frac{A_l}{r^{l+1}}P_l(\cos\theta) \qquad (79)$$

and inside the charge distribution, the potential can be expressed as:

$$\phi_{in} = \frac{1}{4\pi\epsilon_0}\sum_{l=0}^{\infty}B_l r^l P_l(\cos\theta) \qquad (80)$$

(Not sure how to derive these specifically. You can see the general pattern, however, that outside the material, we involve positive coefficients of r^{-1}; inside the material we have nonpositive coefficients of r^{-1}.) Generally, we have to solve for the coefficients $\{A_l\}$ and $\{B_l\}$ given some other constraint.

---

5 EM Fields in dielectric materials

The PDF at `http://www.phys.nthu.edu.tw/~thschang/notes/EM04.pdf` was very helpful for this topic.

5.1 The polarization and displacement fields

5.1.1 The polarization field

If we apply an $\vec{E}$ field on a dielectric material, then dipoles form. The net of these molecular dipoles forms the polarization field, $\vec{P}$:

$$\vec{P} = N\langle\vec{p}\rangle \tag{81}$$

where N is the number density of the dielectric particles. This can be thought of as a "dipole moment density." This is also proportional to the applied $\vec{E}$ field:

$$\vec{P} = \epsilon_0 \chi_E \vec{E} \tag{82}$$

where χ_E is the dielectric susceptibility (can be thought of as how willing a substance is to polarizing; free space doesn't polarize, so $\chi_{E_0} = 0$). For this course, we concern ourselves with linear material; i.e., $\vec{D}$ is directly proportional to the applied $\vec{E}$, or χ_E is a constant proportionality factor. (Real materials are not always so simple.)

5.1.2 The displacement fields

The electric and polarization fields combine to form the displacement field $\vec{D}$, which is like the net dipole moment density:

$$\vec{D} = \epsilon_0 \vec{E} + \epsilon_0 \chi_E \vec{E} = \epsilon_0 (1 + \chi_E) \vec{E} \tag{83}$$

We define the relative permittivity κ as follows:

$$\kappa \coloneqq 1 + \chi_E \tag{84}$$

and electric permittivity (of material) to be:

$$\epsilon = \kappa \epsilon_0 \tag{85}$$

(This gives the name "relative permittivity" meaning, as $\kappa = \epsilon/\epsilon_0$.) Thus, with these new definitions,

$$\vec{D} = \epsilon \vec{E} \tag{86}$$

From an intuitive perspective, this means that the displacement field (think of it as some sort of net electric field) is the applied electric field along with some induced amount of electric field, the amount that is induced is dependent on the material.

Taken directly from the notes: "Another way to think of the $\vec{P}$ field is in terms of 'bound charge density,' ρ_b, where:

$$-\nabla \cdot \vec{P} = \rho_b \tag{87}$$

For example, in [a dipole distribution in which all $\vec{p}$ vectors point outwards from a single point], $\nabla \cdot \vec{P} \neq 0$." In other words, the negative of the divergence of the polarization field is the "bound charge density," the net charge density due to dipolar (thus bound) electrons. In this particular dipole distribution, there is a net divergence, and thus $\rho_{bound} \neq 0$; intuitively, this makes sense since if all of the dipole moments are pointing outwards, then there is a positive polarization field divergence; also, the dipoles are slightly stretched, and the net interior charge due to the dipoles is slightly negative (since the surface/boundary charge is slightly more positive, and the total charge must be zero).

5.2 The $\vec{H}$ and magnetization fields

5.2.1 The $\vec{H}$ field

We define the field $\vec{H}$ (sometimes called the magnetic field strength field) as a more primitive "magnetic" field, one that doesn't rely on matter and is the magnetic analog to the $\vec{E}$ field. This is because the $\vec{B}$ field that is often used is actually more of an empirical measure traditionally used because we typically see magnetic fields alongside matter, but this is analogous to the $\vec{D}$ field that is a combination of some "driving," more fundamental force ($\vec{E}$ field) and its response from matter (the $\vec{P}$ field).

5.2.2 The magnetization field

Similar to the polarization field, we have the magnetization field, $\vec{M}$. Like polarization, this is the net of the dipole moments of the individual particles; unlike polarization, there are two types of magnetic dipole moments caused by atoms.

First of all, we need to define the magnetic moment $\vec{m}$:

$$\vec{m} = I\vec{A} \tag{88}$$

where I is some quantity of current, and $\vec{A}$ is a vector with magnitude the area or the surface enclosed by the loop and direction normal to the surface enclosed by the loop (similar to angular momentum). This can be applied both to electron orbits (if we consider the electron to be traveling in the classical sense of "orbiting" the nucleus), or in the quantum-mechanical sense of electron "spin."

The magnetization field is the bulk of the average magnetic fields:

$$\vec{M} = N\langle\vec{n}\rangle \tag{89}$$

Like the polarization field, it should be proportional to the magnetic field strength (assuming linear material):

$$\vec{M} = \chi_M\vec{H} \tag{90}$$

where χ_M is the magnetic susceptibility (the same conclusions can be drawn as for the electric susceptibility). Note the slight asymmetry from the polarization field: there is no multiplication by μ_0 here.

5.2.3 Revisiting the magnetic field

We provide a more precise formulation of the $\vec{B}$ field analogous to that of the displacement field:

$$\vec{B} = \mu_0 \vec{H} + \mu_0 \vec{M} = \mu_0(1 + \chi_M)\vec{H} \tag{91}$$

If we define the magnetic permeability of dielectric material:

$$\mu = \mu_0(1 + \chi_M) \tag{92}$$

(Note that for many materials, $\mu \approx \mu_0$, so this is a valid approximation.) Then the equation for $\vec{B}$ simplifies to:

$$\vec{B} = \mu \vec{H} \tag{93}$$

fitting a simple linear relationship analogous to the displacement field (assuming linear material).

5.2.4 Categories of magnetic materials

We can generally classify materials into one of three categories:

Paramagnetic: $\mu > \mu_0$, $\chi_M > 0$. Thus the magnetization occurs in the same direction as the imposed $\vec{H}$ field. This is due to a dominating magnetic moment caused by electron spin (i.e., many half-full orbitals). E.g., O_2, Al.

Diamagnetic: $\mu < \mu_0$, $\chi_M < 0$. Thus the magnetization opposes the $\vec{H}$ field. This is due to small contribution to magnetic moment due to electron spin (i.e., mostly full orbitals), and thus the electron "orbits" around the nucleus dominate the magnetic moment. E.g., H_2O, N_2, Cu.

Ferromagnetic: $\mu \gg 1$ (not really a subset of paramagnetic because of its more extreme properties). This is due to more complex (nonlinear) material, i.e., $\mu = \mu(H)$, above some critical temperature (the "curie temperature" T_{curie}). These exhibit hysteresis (can be interpreted as "memory" or "delay" in their response to changes in the $\vec{H}$ field). E.g., Fe (hence the name), Ni, Fe_3O_4.

5.3 Maxwell equations in material

Rearranging (Eq. 1), which still holds in material, we get:

$$\nabla \cdot \epsilon_0 \vec{E} = \rho \tag{94}$$

Noting that $\epsilon - \epsilon_0 = \epsilon_0 \chi_E$, we get:

$$\nabla \cdot (\epsilon - (\epsilon - \epsilon_0))\vec{E} = \nabla \cdot (\epsilon \vec{E}) - \nabla \cdot (\epsilon_0 \chi_E \vec{E}) = \nabla \cdot \vec{D} - \nabla \cdot \vec{P} = \rho \tag{95}$$

Noting that ρ consists of "free charge" ρ_f (free electrons and those in conducting material, which conduct current when an electric field is applied) and "bound charge" ρ_b (from dielectric dipoles), then:

$$\rho = \rho_f + \rho_b \tag{96}$$

A hand-wavy proof of the Maxwell equation for the divergence of $\vec{D}$ in matter says that if the bound charge density comes from the divergence of the polarization field (dielectrics), then the free charge density is from the divergence of the displacement field (credit to `https://em.geosci.xyz/content/maxwell1\_fundamentals/formative\_laws/gauss\_electric.html`). (This isn't really a proof, but this result is true.) Then we get the following two equations:

$$\nabla \cdot \vec{D} = \rho_f \tag{97}$$

$$-\nabla \cdot \vec{P} = \rho_b \tag{98}$$

The former becomes our first Maxwell equation in material; the latter is another way to arrive at (Eq. 87).

The rest of the Maxwell equations are mostly the same, but can be rewritten using the constituent equations. We can assume that there are no free charges in dielectric material, so the Maxwell equations in dielectric material:

$$\nabla \cdot \vec{D} = \rho_f = 0$$

$$\nabla \cdot \vec{H} = \frac{1}{\mu}\nabla \cdot \vec{B} = \frac{1}{\mu}(0) = 0$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

The induction law is a little more complicated; see `http://www.oceanopticsbook.info/view/radiative_transfer_theory/level_2/maxwells_equations_in_matter` for a more in-depth explanation (it involves looking at the free, bound, and polarization currents); for our purposes, we can just think of it as an extension of the vacuum equation, replacing permittivity and permeability of free space with their material counterparts (this again assumes no free current):

$$\nabla \times \vec{B} = \mu \nabla \times \vec{H} = \mu\epsilon \frac{\partial \vec{E}}{\partial t} = \mu \frac{\partial}{\partial t}(\epsilon \vec{E}) \Rightarrow \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t}$$

Summary of equations and constituent relations (in linear dielectric material with no free charge):

$$\nabla \cdot \vec{D} = 0 \tag{99}$$

$$\nabla \cdot \vec{B} = 0 \tag{100}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{101}$$

$$\nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} \tag{102}$$

$$\vec{D} = \epsilon \vec{E} \tag{103}$$

$$\vec{B} = \mu \vec{H} \tag{104}$$

This set of equations is more symmetric between the electric and magnetic fields, with the introduction of the "underlying" $\vec{H}$ field and "overlying" $\vec{D}$ field. To complete the picture, these cause some "response" in dielectric matter (the size of the response depends on the material's susceptibility), which are the $\vec{M}$ and $\vec{P}$ fields; combining the fundamental and material response fields leads us to the $\vec{B}$ and $\vec{D}$ fields.

As expected, these simplify to the vacuum Maxwell equations when in a vacuum, since the "response" fields are zero since the susceptibility of free space is zero.

A straightforward application of these new relations is to show another representation of the Poynting vector, which also nicely more symmetric once we only involve the fundamental fields:

$$\vec{S} = \frac{1}{\mu} \vec{E} \times \vec{B} = \vec{E} \times \vec{H} \tag{105}$$

We will revisit the Maxwell relations in material, with the possibility of free charges and current, in the last section (Sec. 7.2).

5.4 Interface conditions

For these, we show the more general forms and the less general ones.

5.4.1 Pillbox approach for flux

We approach the first two Maxwell equations. First, don't assume no free charge. Imagine a small "pillbox" (or other prism) straddling the interface between two materials, the first with permittivity ϵ_1 and the second with permittivity ϵ_2, with its two flat ends (locally approximately) parallel to the interface, one flat end in each material, and the sides of pillbox of infinitesimal length. The total $\vec{E}$ flux through this surface is almost all from the two parallel interfaces, since the area of the sides is negligible, and all of it is normal to the interface.

For the displacement field, we have the general form (Eq. 97). Assuming $\hat{n}$ is the unit normal vector pointing from material 1 to material 2, the net flux through this is

$$\nabla \cdot \vec{D} = \hat{n} \cdot \vec{D}_2 + (-\hat{n}) \cdot \vec{D}_1 = \hat{n} \cdot (\vec{D}_2 - \vec{D}_1) = \rho_f \tag{106}$$

Since this pillbox is essentially on the boundary of the material, this free charge is not actually the free charge volume density, but rather the free charge surface density, σ_f, so this may rewrite this as:

$$\hat{n} \cdot (\vec{D}_2 - \vec{D}_1) = \sigma_f \tag{107}$$

Of course, assuming no free charge (as we do for this course), $\sigma_f = 0$, so:

$$\hat{n} \cdot (\vec{D}_2 - \vec{D}_1) = 0 \tag{108}$$

This means that the normal component of the $\vec{D}$ field is continuous across a boundary (when there is no free charge), or that:

$$\vec{D}_1^{\perp} = \vec{D}_2^{\perp} \tag{109}$$

$$\epsilon_1 \vec{E}_1^{\perp} = \epsilon_2 \vec{E}_1^{\perp} \tag{110}$$

The same pillbox method can be used to obtain analogous relations for $\vec{B}$ across an interface (but this holds whether there is free charge or not):

$$\vec{B}_1^{\perp} = \vec{B}_2^{\perp} \tag{111}$$

$$\mu_1 \vec{H}_1^{\perp} = \mu_2 \vec{H}_2^{\perp} \tag{112}$$

5.4.2 Loop approach for curl

We can create a similar setup to the pillbox method, but instead of a pillbox, we have a rectangular loop: two sides (locally approximately) parallel to the surface, and the two other sides perpendicular to the surface and infinitesimally long. In a similar way to the pillbox method, we find the curl around the loop, neglecting that caused by the normal components, and apply Stoke's Theorem on (Eq. 101):

$$\oiint_A (\nabla \times \vec{E}) \cdot d\vec{s} = \oint_C \vec{E} \cdot \, d\vec{l} = -\oiint_A \frac{\partial \vec{B}}{\partial t} \cdot d\vec{s} \tag{113}$$

This essentially transforms it back into the integral form. Since as we take the limit of the sides normal to the interface as their length goes to 0, the limit of the area goes to zero, so the enclosed $\vec{B}$ flux also goes to zero (the right side of the equation). The curl of $\vec{E}$ (the left side of the equation) is the sum of the $\vec{E}$ tangential to the interface in opposite directions; this magnitude of the tangential field can be expressed as $\hat{n} \times \vec{E}$. Thus:

$$\hat{n} \times \vec{E}_2 + (-\hat{n}) \times \vec{E}_1 = \hat{n} \times (\vec{E}_2 - \vec{E}_1) = 0 \tag{114}$$

This means that the tangential part of the $\vec{E}$ field is continuous. The same result can be shown for $\vec{H}$ fields using the fourth Maxwell equation; this gives us the following tangential interface conditions:

$$\vec{E}_1^{\parallel} = \vec{E}_2^{\parallel} \tag{115}$$

$$\frac{1}{\epsilon_1}\vec{D}_1^{\|} = \frac{1}{\epsilon_2}\vec{D}_2^{\|} \tag{116}$$

$$\vec{H}_1^{\|} = \vec{H}_2^{\|} \tag{117}$$

$$\frac{1}{\mu_1}\vec{B}_1^{\|} = \frac{1}{\mu_2}\vec{B}_2^{\|} \tag{118}$$

5.4.3 Example: refraction of a (steady) electric field

These boundary conditions describe how static fields act near boundaries; this describes refraction. If α describes angle from the normal of an $\vec{E}$ field vector as near a horizontal interface, then from the normal and tangential matching conditions, respectively, we have:

$$\epsilon_1 E_1 \cos\alpha_1 = \epsilon_2 E_2 \cos\alpha_2 \tag{119}$$

$$E_1 \sin\alpha_1 = E_2 \sin\alpha_2 \tag{120}$$

Combining these two:

$$\frac{\tan\alpha_1}{\epsilon_1} = \frac{\tan\alpha_2}{\epsilon_2} \tag{121}$$

---

6 Reflection and transmission

This section is an application of the boundary conditions derived in the previous section, but now we apply it to plane waves (as opposed to static fields). We deal first with the simpler case of normal incidence (when the wave vector is normal to the interface) and then extend the results to oblique incidence.

6.1 Index of refraction, propagation speed, and impedance

From the Maxwell equations in material, we get:

$$v = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon}} \tag{122}$$

Which is an extension of the vacuum case (clearly, this agrees with the vacuum case since $v = c$ if $\mu = \mu_0$ and $\epsilon = \epsilon_0$). The index of refraction is defined as the ratio of c to v:

$$n = \frac{c}{v} = \frac{\frac{1}{\sqrt{\mu_0\epsilon_0}}}{\frac{1}{\sqrt{\mu\epsilon}}} = \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}} \approx \sqrt{\frac{\epsilon}{\epsilon_0}} = \sqrt{\kappa} \tag{123}$$

Usually, $n > 1$, but it is possible to have n be less than unity: see `https://en.wikipedia.org/wiki/Refractive_index#Refractive_index_below_unity`. Now, define the impedance Z of a material to be:

$$Z = \sqrt{\frac{\mu}{\epsilon}} \tag{124}$$

Note that this has units of electrical resistance ("impedance"). Plugging this into (Eq. 38), we can rewrite the relationship between $\vec{B}$ and $\vec{E}$ as:

$$E = \frac{\omega}{k}(\mu H) = \frac{\mu}{\sqrt{\mu\epsilon}} H = ZH \tag{125}$$

and thus we can rewrite Z as:

$$Z = \frac{E}{H} \tag{126}$$

Also, note that Z is inversely proportional to n (and thus directly proportional to v). (It may be a little counterintuitive that a higher impedance corresponds to a higher wave speed; this can be thought of like in a string wave, in which a wave moves faster the tauter the string is (which can be thought of as having a higher "impedance")).

$$Z = \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{\mu}{\epsilon}\frac{\mu}{\mu}} = \sqrt{\frac{\mu^2}{\mu\epsilon}} = \mu v = \mu\frac{c}{n} \Rightarrow Z \propto n^{-1} \tag{127}$$

6.2 Fresnel equations

(For this section, refer to the notes for diagrams.)

There are two polarization cases: the first, s-polarization (a.k.a., TE, transverse electric), when $\vec{E}$ is parallel to the interface (and normal to the incidence plane); and p-polarization (a.k.a., TM, transverse magnetic), when $\vec{E}$ is in the incidence plane. This is the general oblique case, for which normal incidence can be derived when $\theta_i = 0$.

Note that while s- and p-polarization are specific polarization cases, they are orthogonal and span the polarization plane (which is two-dimensional); thus, any EM wave with any polarization can be represented as a linear combination of the two. Also, there are a few assumptions made about the directions of the reflected and transmitted waves that are not explained; these are due to solving all of the interface conditions simultaneously.

Setup: the interface is the (horizontal) plane $z = 0$, the incidence plane is the plane $x = 0$. The incident wave has a wave vector aimed toward the origin from the second quadrant of the y-z plane, i.e.,

$$\vec{k_i} = k_i \sin\theta_i \hat{y} - k_i \cos\theta_i \hat{z} \tag{128}$$

$$\vec{k_r} = k_r \sin\theta_r \hat{y} + k_r \cos\theta_r \hat{z} \tag{129}$$

$$\vec{k_t} = k_t \sin\theta_t \hat{y} - k_t \cos\theta_t \hat{z} \tag{130}$$

Note that all three vectors lie in the incidence plane; see (Sec. 6.2.2).

6.2.1 s-polarization case

In this case, $\vec{E}$ vectors are normal to the plane of incidence ("s" for "perpendicular" (in German)), and $\vec{B}$ vectors lie in the plane of incidence, with $\vec{B}_i$ pointing NE, $\vec{B}_r$ pointing SE, and $\vec{B}_t$ pointing NE. Assume $\vec{E}_i$ and $\vec{E}_t$ point into the sheet of paper, and $\vec{E}_r$ points out of the sheet of paper.

(N.B. These signage decisions are arbitrary but important to keep consistent through a proof; note that we keep right-handed triads of $\vec{E}$ and $\vec{B}$ consistent, and thus their magnitudes can go negative as long as the signs of the $\vec{E}$ and $\vec{B}$ vectors match: i.e., $E_x/|E_x| = B_x/|B_x|$. [In this orientation, as $\theta_i = 0$, the signs of $\vec{E}_i$ and $\vec{E}_r$ are opposite, which is consistent with the setup for the normal case; we could have changed the orientation so that $\vec{E}_i$ and $\vec{E}_r$ would be in the same direction if $\theta_i = 0$, but the magnitude would be inverted. Either way, it makes no difference to the reflection or transmission (power) coefficients, which are the square of the magnitudes of the reflected and transmitted waves, respectively.] However, note that we are asserting that these are the orientations of the reflected and transmitted $\vec{E}$ and $\vec{B}$ vectors (i.e., that the $\vec{E}$ waves stay parallel to the plane of incidence and that the $\vec{B}$ waves stay in the plane of incidence) without proof, and using these to solve for the magnitudes assuming they're in this orientation. I have no idea how to prove that these are the correct orientations (and can't seem to find it online), but assume it is some problem of simultaneously solving all of the boundary conditions at once (including the normal ones, which we don't deal with here), and/or some calculus/optimization problem.)

Then, on the boundary:

$$\vec{E}_i = \vec{E}_{i_0} e^{i(yk_i \sin\theta_i - \omega_i t)} \tag{131}$$

$$\vec{E}_r = \vec{E}_{r_0} e^{i(yk_r \sin\theta_r - \omega_r t)} \tag{132}$$

$$\vec{E}_t = \vec{E}_{t_0} e^{i(yk_t \sin\theta_t - \omega_t t)} \tag{133}$$

Note here that $\vec{k} \cdot \vec{r}$ simplifies to only having a $\hat{y}$ component, since $k_x = 0$ and $z = 0$ (on the boundary). Then, apply the tangential boundary conditions:

$$\vec{E}_i^{\parallel} + \vec{E}_r^{\parallel} = \vec{E}_t^{\parallel} \tag{134}$$

$$\vec{H}_i^{\parallel} + \vec{H}_r^{\parallel} = \vec{H}_t^{\parallel} \tag{135}$$

In the case of s-polarization, the $\vec{E}$ vectors are already parallel to the plane, but the $\vec{B}$ are not. Thus, for s-polarization, we can write these more concretely as:

$$-E_{i_0} e^{i(yk_i \sin\theta_i - \omega_i t)} - E_{r_0} e^{i(yk_r \sin\theta_r - \omega_r t)} = -E_{t_0} e^{i(yk_t \sin\theta_t - \omega_t t)} \tag{136}$$

$$H_{i_0} \cos\theta_i e^{i(yk_i \sin\theta_i - \omega_i t)} + H_{r_0} \cos\theta_i e^{i(yk_r \sin\theta_r - \omega_r t)} = H_{t_0} \cos\theta_i e^{i(yk_t \sin\theta_t - \omega_t t)} \tag{137}$$

6.2.2 The laws of reflection and refraction

We make the observation that this set of equations must be true for all moments in time and all points in space. Thus, at any given point in space, in order to satisfy both boundary equations, all of the time terms must factor out. In other words, two sinusoids cannot sum to another sinusoid unless their frequencies are equal, and the sum sinusoid must have the same frequency. Thus they must all be oscillating at the same rate in time:

$$e^{i(-\omega_i t)} = e^{i(-\omega_r t)} = e^{i(-\omega_t t)} \Rightarrow \omega_i = \omega_r = \omega_t \tag{138}$$

Note that since this is a simple result of sinusoid math and does not involve the boundary conditions, it is true not only for EM waves, but other waves at boundaries (e.g., waves on a string). By the same logic, we can also deduce that the oscillations in space must also be due to an equal space-phase:

$$e^{iyk_i \sin\theta_i} = e^{iyk_r \sin\theta_r} = e^{iyk_t \sin\theta_t} \Rightarrow k_i \sin\theta_i = k_r \sin\theta_r = k_t \sin\theta_t \tag{139}$$

Since $k = \omega/v$, and we know from (Eq. 138) that $\omega_i = \omega_r$, and $v_i = v_r = v_1$ since waves travel at one speed through the one medium dependent on its index of refraction, then $k_i = k_r = k_1$. (Thus, it makes sense to define $k_2 \coloneqq k_t$, and $\theta_2 \coloneqq \theta_t$.) This allows us to draw two conclusions (both of which are general to waves at an interface, again without the EM boundary conditions). Firstly, the law of reflection:

$$\theta_i = \theta_r (\coloneqq \theta_1) \tag{140}$$

and then the law of refraction (Snell's law)

$$k_1 \sin\theta_1 = k_2 \sin\theta_2 \tag{141}$$

6.2.3 Aside: $\vec{k}$ lying in the incidence plane

With these relations and this logic, we can now show that all of the $\vec{k}$ vectors lie in the incidence plane. We assumed that $k_{r_x} = k_{t_x} = 0$; for now, assume they are not. Then, assuming that $\vec{k}_r$ and $\vec{k}_t$ lie ψ_r and ψ_t off of the normal axis (and off the incidence plane) in the $\hat{x}$-direction, the space-phase component ϕ_s of the three vectors are:

$$\begin{aligned} \phi_{i_s} &= iyk_i \sin\theta_i \\ \phi_{r_s} &= ik_r(y\sin\theta_r + x\sin\psi_r) \\ \phi_{t_s} &= ik_t(y\sin\theta_t + x\sin\psi_t) \end{aligned}$$

By the same logic as above, these three space-phases must be equal at all points in the $\hat{x}$ and $\hat{y}$ directions at any point in time (the same logic applies for the periodicity of the $\hat{x}$ and $\hat{y}$ dimensions in a complex sinusoid as it does for the t dimension). Thus these must work when $x \neq 0$ and $y = 0$, in which case setting all of the equations equal yields:

$$0 = x\sin\psi_r = x\sin\psi_t \tag{142}$$

Thus forcing $\sin\psi = 0$, so $\vec{k}$ must lie on the incidence plane.

6.2.4 Results of the s-polarization case

With the law of reflection and refraction, our boundary conditions get simplified to:

$$-E_{i_0} + E_{r_0} = -E_{t_0} \tag{143}$$

$$H_{i_0} \cos\theta_1 + H_{r_0} \cos\theta_2 = H_{t_0} \cos\theta_2 \tag{144}$$

Some magic/algebra happens here... We obtain the s-polarization reflectivity and transmission (amplitude) coefficients:

$$r_s = \frac{E_{r_0}}{E_{i_0}} = -\frac{\cos\theta_i - \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}} \tag{145}$$

$$t_s = \frac{E_{t_0}}{E_{i_0}} = \frac{2\cos\theta_i}{\cos\theta_i + \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}} \tag{146}$$

6.2.5 p-polarization case setup and results

In this case, the $\vec{E}$ fields lie in the plane of incident ("p" for "parallel" (also in German)). Imagine that the $\vec{k}$ vectors are situated the same way, all three $\vec{B}$ vectors point straight out of the page toward you, and $\vec{E}_i$ points in the NE direction, $\vec{E}_r$ points in the NW direction, and $\vec{E}_t$ points in the NE direction. Keep in mind the N.B. from the s-polarization section. Again, we use the tangential boundary conditions; here, $\vec{B}$ is already tangential to the interface, but $\vec{E}$ is not. The results are very similar (including the law of reflection, law of refraction, and the wave vectors in the plane of incidence), so we will skip much of the algebra and skip to the results. We have the p-polarization reflectivity and transmission (amplitude) coefficients:

$$r_p = \frac{E_{r_0}}{E_{i_0}} = -\frac{\left(\frac{n_2}{n_1}\right)^2 \cos\theta_i - \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}}{\left(\frac{n_2}{n_1}\right)^2 \cos\theta_i + \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}} \tag{147}$$

$$t_p = \frac{E_{t_0}}{E_{i_0}} = \frac{2\frac{n_2}{n_1}\cos\theta_i}{\left(\frac{n_2}{n_1}\right)^2 \cos\theta_i + \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_i}} \tag{148}$$

6.3 Normal incidence Fresnel equations

We can derive these simply from the more complex oblique cases derived above. The orientations match the limit of the orientations for both the s- and p-polarization case; i.e., the $\vec{E}_r$ is antiparallel to $\vec{E}_i$ and $\vec{E}_t$; and all of the $\vec{B}$s are parallel and pointing in the same direction. With $\theta_i = 0$, the two polarization cases converge to the same result. We have the normal case reflectivity and transmission (amplitude) coefficients:

$$r_\perp = \frac{n_2 - n_1}{n_2 + n_1} = \frac{Z_1 - Z_2}{Z_1 + Z_2} \tag{149}$$

$$t_\perp = \frac{2n_1}{n_2 + n_1} = \frac{2Z_2}{Z_1 + Z_2} \tag{150}$$

The relation to Z is due to (Eq. 127). What this means is that if the index of refraction of material 2 is greater than that of material 1 (or, alternatively, if the impedance of material 1 is greater than that of material 2), then the reflected $\vec{E}$ will be antiparallel; and vice versa. If the materials have the same index of refraction, then there will be zero reflection and total transmission (this is called impedance matching). The transmission, however, will always be parallel to the incident wave.

6.4 Reflection and transmission power coefficients

We can define the reflection and transmission power coefficients:

$$R \coloneqq \frac{\langle \vec{S}_r \rangle \cdot \hat{n}}{\langle \vec{S}_i \rangle \cdot \hat{n}} \tag{151}$$

$$T \coloneqq \frac{\langle \vec{S}_t \rangle \cdot \hat{n}}{\langle \vec{S}_i \rangle \cdot \hat{n}} \tag{152}$$

In other words, R and T are the power flux coefficients. Extending (Eq. 44) to use ϵ and μ, we have:

$$\langle \vec{S} \rangle = \frac{1}{2Z} E_0^2 \hat{k} \tag{153}$$

and thus R and T can be simplified to (using Snell's law and the fact that $Z_i = Z_r \neq Z_t$):

$$R = \left| \frac{E_{r_0}}{E_{i_0}} \right|^2 \tag{154}$$

$$T = \left| \frac{E_{t_0}}{E_{i_0}} \right|^2 \frac{n_2 \cos\theta_2}{n_1 \cos\theta_1} \tag{155}$$

6.5 Evanescent waves

We define the critical angle to be:

$$\theta_{crit} \coloneqq \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{156}$$

When $\theta_i < \theta_{crit}$, then the transmitted amplitude coefficient t is real. When $\theta_i = \theta_{crit}$, then there is no transmission $t = 0$ and total internal reflection

($r = 1$). However, when $\theta_i > \theta_{crit}$, then $r \in \mathbb{C}$ (but $|r| = 1$ and $t = 0$, so it is still total internal reflection). This in and of itself is not very revealing, so just let the math speak for itself.

$$\vec{E}_t = \vec{E}_{t_0} e^{i(\omega_t - k_2(y \sin\theta_2 + z\cos\theta_2))} \tag{157}$$

Making the substitutions (cosine identity, Snell's law, definition of critical angle) and the fact that $\sin\theta_1 > \sin\theta_{crit}$:

$$\cos\theta_1 = \pm\sqrt{1 - \sin^2\theta_1} \tag{158}$$

$$\sin\theta_2 = \frac{n_1}{n_2}\sin\theta_1 \tag{159}$$

$$\frac{n_2}{n_1} = \sin\theta_{crit} \tag{160}$$

then we can rewrite the transmitted $\vec{E}$ wave as:

$$\vec{E}_t = \vec{E}_{t_0} e^{-k_2 z\sqrt{\left(\frac{\sin\theta_1}{\sin\theta_{crit}}\right)^2 - 1}} e^{i\omega t - ik_2 y \frac{n_2}{n_1}\sin\theta_1} \tag{161}$$

Note here that the wave number k (what is being dotted with the position vector) becomes complex; the complex part gets multiplied with i again in the exponent to become real. (This will come up again in the next two topics.) We can define the folding distance z_0 to be:

$$z_0 = \frac{1}{k_2\sqrt{\left(\frac{\sin\theta_1}{\sin\theta_{crit}}\right)^2 - 1}} \tag{162}$$

(in general, the folding distance is z_0 s.t. the solution has an $\exp(z/z_0)$ coefficient) and trivially solve for the velocity and wave number of this wave:

$$v = \frac{\omega}{k_2\left(\frac{n_1}{n_2}\right)\sin\theta_1} \tag{163}$$

$$k = k_2\left(\frac{n_1}{n_2}\right)\sin\theta_1 \tag{164}$$

to rewrite this wave as:

$$\vec{E}_t = \vec{E}_{t_0} e^{-\frac{z}{z_0}} e^{ik(y - vt)} \tag{165}$$

This is an electric wave traveling parallel to the interface with exponentially-decreasing magnitude the further you get away from the interface. This is called the "evanescent wave," and it is traveling in the denser material 2, which is now called the "forbidden region." If you get very close to the surface, then you can detect or interact with this wave; this is called "frustrated total internal reflection" (FTIR); otherwise, this energy just travels next to the surface and is not lost. In the case of FTIR, then the reflection is not total; even though the wave is hitting the interface at an angle greater than the critical angle, this is a way to still transmit (lose) energy across the interface.

It is also sometimes referred to as "evanescent wave coupling," as messing with the evanescent wave will change the totality of the reflection, and is analogous to a quantum-mechanical phenomenon (quantum tunneling of wave functions; according to Wikipedia).

6.6 Brewster's angle

Now that we've found the case for zero transmission ($\theta = \theta_{crit}$) and zero-ish transmission ($\theta > \theta_{crit}$), what about the case for zero reflection? $r = 0$ only occurs in the p-polarization case:

$$r_p = 0 \Rightarrow \left(\frac{n_2}{n_1}\right)^2 \cos\theta_1 = \sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_1} \tag{166}$$

Define the angle that satisfies this relation to be Brewster's angle, θ_B:

$$\theta_B := \tan^{-1}\left(\frac{n_2}{n_1}\right) \tag{167}$$

At this angle, the angle between $\hat{k}_r \perp \hat{k}_t$. This geometry is relevant; the zero transmission is due to the anisotropic nature of the Larmor power radiation; at this angle of incidence, electrons will accelerate in a motion parallel to $\hat{k}_r$ and normal to $\hat{k}_t$, thus not radiating any in the direction of the reflected wave. This won't happen in the s-polarization case since the $\vec{E}$ wave is parallel to the interface, and thus will be accelerating electrons back and forth in the interface plane, so the region of no radiation will be in the interface plane and thus cannot be in the direction of the reflected wave.

In the case of random polarization (a mix of s- and p-polarization) hitting a surface at Brewster's angle, this means that the reflected light will be s-polarized and the transmitted light will be p-polarized; this can be used as a simple polarization filter and has its applications with camera glare (if the light reflected from a surface is heavily s-polarized, then blocking out that polarization of light will greatly reduce the reflection (glare) from that surface while affecting less of the randomly-polarized light from other materials).

---

7 Absorption and dispersion in dielectrics

In dielectrics, we will derive a frequency-dependent index of refraction, n. This means that light waves at different frequencies will move at different speeds through a material, and will reflect/refract differently; e.g., a rainbow or white light going through a prism are examples of dispersion. This is closely related to another phenomenon, absorption; we will see that with complex n values, there is some absorption of light in matter; because of dispersion, different frequencies cause different n values, which in turn cause differing amounts of absorption, which lead to our concept of the color of materials based on the intrinsic properties of the material (the fundamental frequencies of its molecules).

7.1 Improved spring model

Recall that the moment of a dipole is:

$$\vec{p} = q\vec{d} \tag{168}$$

and recall that in a dielectric material, $\vec{d}$ (charge separation) is represented by the motion of an electron on a spring, which is dependent on the fundamental frequency of the charge. In the previous spring model, we assumed that a particle only has one fundamental frequency. Let us consider all of the fundamental frequencies of an atom, represented by some distribution; let i denote the number of different characteristic frequencies of an atom, $\vec{p_i}$ denote the dipole moment of the i-th frequency of an atom, and f_i be the fraction of electrons within that atom with that fundamental frequency (i.e., $f_i = f(\omega_{i_0})$ is the p.m.f. of the fundamental frequencies). Then the polarization field may be generalized to the bulk polarization of this sum of dipole moments (contrast this with (Eq. 81)):

$$\vec{P} = \sum_i N f_i \vec{p}_i \tag{169}$$

We have already solved for the solution for a single electron, and this more general case is simply the sum (recall that $\vec{p} = q\vec{x}$, since the displacement of an electron is equal to its dipole moment):

$$\vec{P} = \sum_i \frac{N f_i q^2}{m((\omega_{0_i} - \omega)^2 + j\gamma_i\omega)} \vec{E}_{inc} \tag{170}$$

(Here we switch the convention to representing the pure imaginary number as $j = \sqrt{-1}$ to avoid confusion with the indexing variable.) For linear dielectric materials, we can rearrange (Eq. 82) to get:

$$\chi_e = \frac{\vec{P}}{\epsilon_0 \vec{E}} \tag{171}$$

We also know from (Eq. 123) that $n = \sqrt{\kappa}$ and from (Eq. 84) that $\kappa \coloneqq 1 + \chi_e$, so we can solve for n:

$$n^2(\omega) = 1 + \frac{1}{\epsilon_0} \sum_i \frac{N f_i q^2}{m((\omega_{0_i} - \omega)^2 + j\gamma_i\omega)} \tag{172}$$

Thus, n is the square root of a complex number and is thus itself complex (we ignore the technicalities of there being two square roots, just focus on one). Since $n = \frac{c}{\omega}k$, and c and ω are real constants (don't know how to interpret complex speed or frequency), this means that k is also complex. In particular, we see that for an EM wave passing through dielectric material, the amplitude of the wave decreases exponentially with distance if k has an imaginary component:

$$\vec{E} = \vec{E}_0 e^{-\Im(\vec{k}\cdot\vec{r})} e^{i(\Re(\vec{k}\cdot\vec{x})-\omega t)} \tag{173}$$

In other words, the real part of the wave number is like the ordinary wave number (the "space-frequency" of the wave), while the imaginary part represents damping (this is seen in the evanescent wave case; the wave number in that case is complex). Thus, if we wanted to solve for the damping coefficient, we could solve for $\Im(k)$:

$$\Im[k(\omega)] = \frac{\omega}{c}\Im[n(\omega)] \tag{174}$$

We can see from the equation for $n(\omega)$ that if $(\omega_{0_i}-\omega)^2 \gg 0$, then the numerator and denominator are mostly real (and positive), and the resulting $n^2(\omega)$ is mostly real. However, as $\omega_{0_i} - \omega \to 0$, then we have a real numerator divided by a mostly imaginary denominator, resulting in an imaginary quotient (and thus larger damping). Thus, we have damping (absorption) that is largest at the fundamental frequencies of the atom.

The frequencies that correspond to absorption peaks (peaks of $\Im(n)$ and $\Im(k)$) then correspond to our idea of color: the stronger the absorption, the weaker the reflection of any particular frequency of light.

7.2 A better model: $\vec{E}_{site}$ and the Clausius-Mossotti relation

Note that the previous equation for n works best for gases and other sparse dielectrics, since in denser materials we have external contributions from the $\vec{p}$ of other molecules; this model (i.e., the solution to the damped oscillator) assumes that the primary driving force is some external oscillator $\vec{E}_{ext}$ wave. A better model takes into account the imposed electric field. We define the electric field in a hole (vacuum) surrounded by uniform dielectric material as $\vec{E}_{site}$. Dependent on the shape and orientation of the hole, the boundary conditions cause the imposed $\vec{E}_{ext}$ field to be different. For example, in a long thin slot parallel to the $\vec{E}_{ext}$ field, then:

$$\vec{E}_{site} \approx \vec{E}_{ext} \tag{175}$$

Since the dominating boundary is parallel to the $\vec{E}_{ext}$ and $\vec{P}$ fields, there is almost no polarization field lines. If the slit is in the plane normal to the $\vec{E}_{ext}$ field, then the polarization field is strong everywhere in the narrow slit:

$$\vec{E}_{site} \approx \vec{E}_{ext} + \frac{1}{\epsilon_0}\vec{P} \tag{176}$$

In a spherical hole (a result we won't derive):

$$\vec{E}_{site} \approx \vec{E}_{ext} + \frac{1}{3\epsilon_0}\vec{P} \tag{177}$$

This is a better approximation for the actual external (driving) $\vec{E}$ field on an atom in a dielectric material. Using this result, we obtain the Clausius-Mossotti relation to more accurately calculate n (implicitly) in denser dielectric material:

$$\frac{n^2-1}{n^2+2} = \frac{Nq^2}{3\epsilon_0 m}\sum_i \frac{f_i}{(\omega_{0_i}-\omega)^2 + j\gamma_i\omega} \tag{178}$$

---

8 EM waves in conductors

8.1 Free charges and currents in Maxwell equations

We have dealt with dielectric material for a long time, and been able to make assumptions such as that $\rho_f = \vec{j}_f = 0$, which simplified the Maxwell equations. However, in conductors, this is not true anymore. In general, we can express current in this version of Ohm's law (i.e., integrating this over space gives us Ohm's law):

$$\vec{j}_f = \sigma\vec{E} \tag{179}$$

where σ is the conductance of a material (inversely proportional to resistance). The full Maxwell equations are in effect here (these were all stated previously in (Sec. 5.2.4), except we the current component of Ampere's law was disregarded because of no free charges in dielectrics):

$$\nabla\cdot\vec{D} = \rho_f \tag{180}$$

$$\nabla\cdot\vec{H} = 0 \tag{181}$$

$$\nabla\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \tag{182}$$

$$\nabla\times\vec{H} = \sigma\vec{E} + \frac{\partial\vec{D}}{\partial t} \tag{183}$$

We can start with the continuity equation (local conservation law) for charges (this is similar to any other conservation law, such as that for power flux/energy (Eq. 43)):

$$\frac{\partial\rho_f}{\partial t} + \nabla\cdot\vec{j}_f = 0 \tag{184}$$

and, by substituting Ohm's law and the general form of Gauss's electric law, we obtain:

$$\rho_f(t) = \rho_{f_0}\exp\left(-\frac{\sigma}{\epsilon}t\right) \tag{185}$$

where $\rho_{f_0} = \rho_f(0)$. Thus any free charge density in a conducting material will dissipate at an exponential rate, with the time constant $\tau = \epsilon/\sigma$ being inversely proportional to the material's conductance and directly proportional to its permittivity. For good conductors with high conductivity, $\tau \approx 10^{-19}$, so charges dissipate very quickly. Because of this, we can approximate that

$\rho_f \approx 0$ at any moment of time in a conductor, so we can leave the first Maxwell equation for conductors as we have seen it:

$$\nabla \cdot \vec{D} \approx 0 \tag{186}$$

8.2 Dispersion relation

We can solve for the wave equation like we did earlier in the semester:

$$\begin{aligned}
\nabla \times (\nabla \times \vec{E}) &= -\nabla \times \frac{\partial \vec{B}}{\partial t} \\
\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} &= -\frac{\partial}{\partial t}\left(\mu\sigma\vec{E} + \mu\epsilon\frac{\partial \vec{E}}{\partial t}\right) \\
\nabla^2 \vec{E} &= \mu\sigma\frac{\partial \vec{E}}{\partial t} + \mu\epsilon\frac{\partial^2 \vec{E}}{\partial t^2} \qquad (187)
\end{aligned}$$

The same form can be derived for $\vec{B}$ (or $\vec{H}$). If we substitute the general form for a plane wave solution:

$$\vec{E} = \vec{E}_0 e^{i(\omega t - \vec{k}\cdot\vec{r} + \phi_0)}$$

into this differential relation for $\vec{E}$, then by our rules of differential operators on plane waves (Sec. 1.7):

$$\begin{aligned}
-k^2\vec{E} &= \mu\sigma(i\omega)\vec{E} + \mu\epsilon(-\omega^2)\vec{E} \\
k^2 &= \mu\epsilon\omega^2 - i\mu\sigma\omega \qquad (188)
\end{aligned}$$

This is a disperson equation: from Wikipedia, “a dispersion relation relates the wavelength or wavenumber of a wave to its frequency.” Here we have again a complex wave number, just like in the case for the evanescent wave and imaginary wave numbers in dielectrics. (Note that as $\sigma \to 0$, then $k = \Re(k)$, and we approach the simple real k case.) Similarly, this means that we have damping/absorption in a metal, and as before folding distance is:

$$z_0 = \frac{1}{\Im(k)} \tag{189}$$

(This particular folding distance is called the “skin depth” of a conductor.) Since $\Im(k)$ is small, then the skin depth is usually pretty short, which means that the EM fields inside of a conductor are near zero (which agrees with our assumptions/observations last semester in PH213).

8.3 Additional notes: phase shifts and reflectivity

Also note that, for complex wave numbers, there is an additional phase shift between $\vec{E}$ and $\vec{B}$, which does not occur in the real case. In particular, if we write $\vec{k}$ as a complex vector $\left(|k|e^{i \arg k}\right)\hat{k}$ with Faraday's law applying our differential operators:

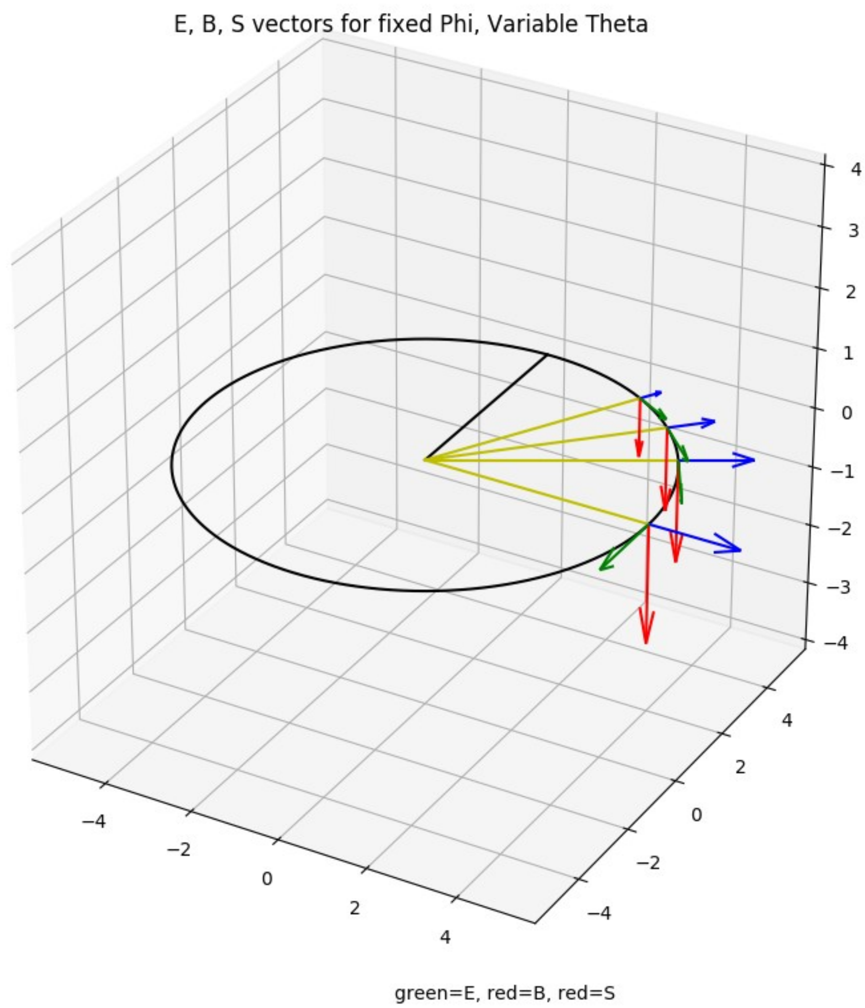
$$\begin{aligned}\nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ -i\vec{k} \times \vec{E} = -i\left(|k|e^{i \arg k}\right)\hat{k} \times \vec{E} &= -i\omega\vec{B} \\ |k|\hat{k} \times \vec{E} &= \omega\vec{B}e^{-i \arg k} \end{aligned} \tag{190}$$

This means that $\hat{k} \times \vec{E}$ is equal to some phase-shifted $\vec{B}$. (Again, this is also the case for evanescent and dielectric damping, but we choose to introduce it here, as it is not as important as previous results.)

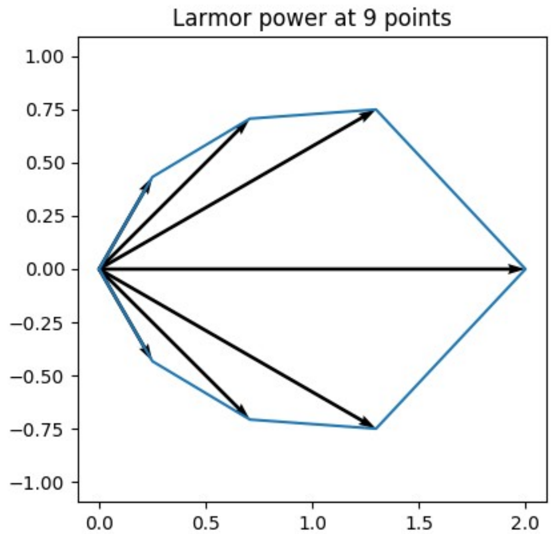
A final comment is that with free charges, we may also have surface free charges and surface free currents. In this case, a good conductor has a very high reflection coefficient, which makes it a good reflector (mirror). (Not derived here but involves rewriting boundary conditions involving Ampere's law with free currents.)

Jonathan Lam
Prof. Yecko
PH214
Optics and Modern
2 / 13 / 20

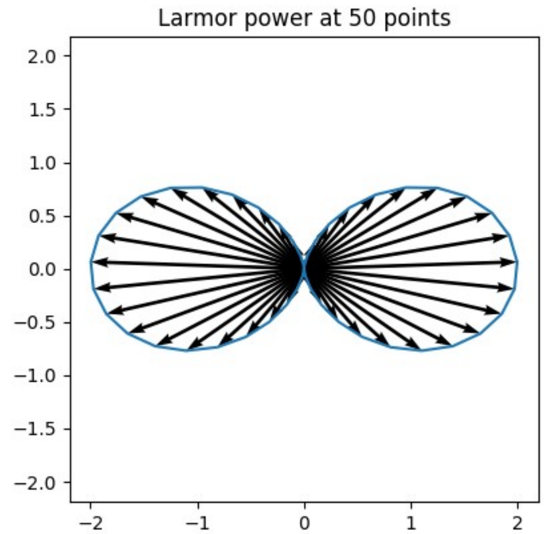
1. Everything Donut



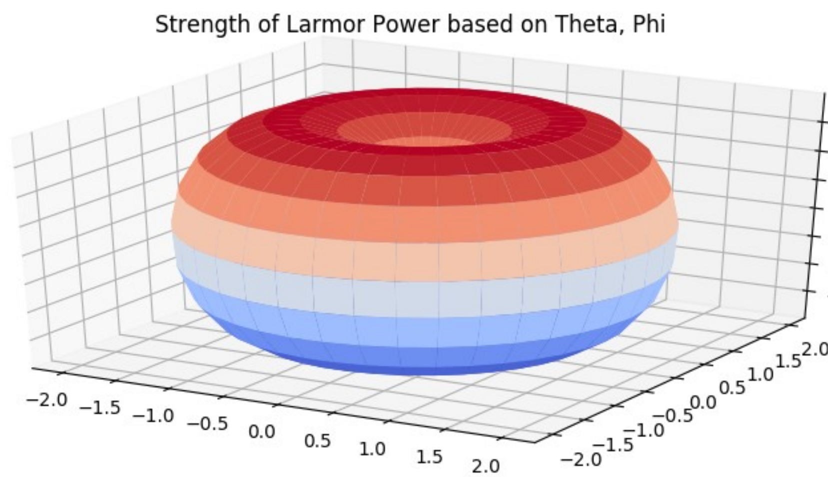
(scaled relative to each other)



Larmor power at 9 points
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0.0
0.5
1.0
1.5
2.0

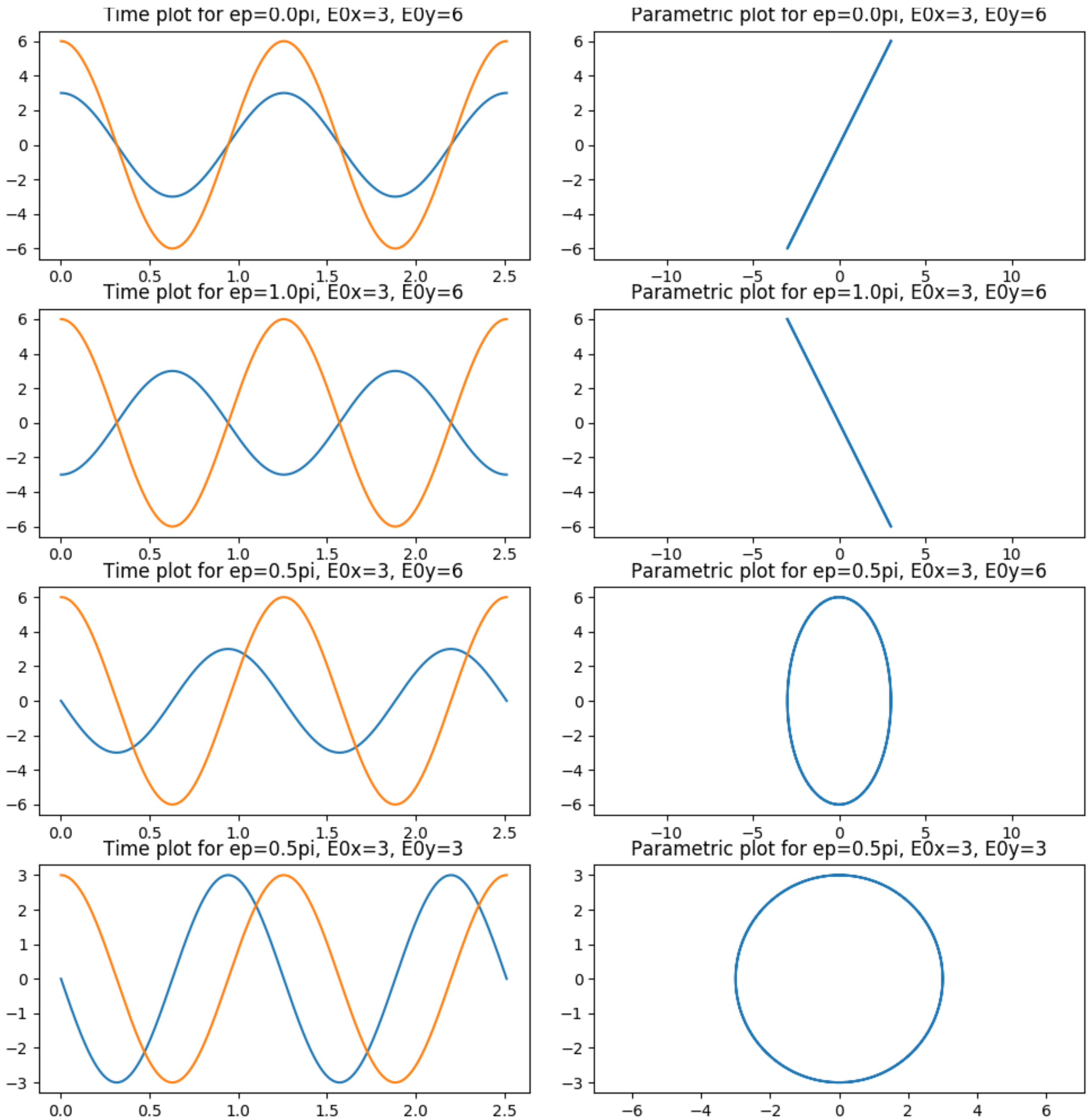


Larmor power at 50 points
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2
−1
0
1
2



Strength of Larmor Power based on Theta, Phi
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

4. Pol



Time plot for ep=0.0pi, E0x=3, E0y=6
Parametric plot for ep=0.0pi, E0x=3, E0y=6
Time plot for ep=1.0pi, E0x=3, E0y=6
Parametric plot for ep=1.0pi, E0x=3, E0y=6
Time plot for ep=0.5pi, E0x=3, E0y=6
Parametric plot for ep=0.5pi, E0x=3, E0y=6
Time plot for ep=0.5pi, E0x=3, E0y=3
Parametric plot for ep=0.5pi, E0x=3, E0y=3

PH214C – PSET3

Jonathan Lam

February 26, 2020

Tale of two sigmas The Thomson scattering cross-section is given by

$$\sigma_{th} = \frac{8\pi}{3}\left(\frac{q^2}{4\pi\epsilon_0 mc^2}\right)^2$$

Plugging in values:

$$q = \pm 1.602 \times 10^{-19}\,\mathrm{C}$$
$$\epsilon_0 = 8.854 \times 10^{-12}\,\mathrm{F/m}$$
$$c = 2.998 \times 10^{8}\,\mathrm{m/s}$$
$$m_e = 9.109 \times 10^{-31}\,\mathrm{kg},\ m_p = 1.673 \times 10^{-27}\,\mathrm{kg}$$

We get:

$$\sigma_e = 6.650 \times 10^{-29}\,\mathrm{m}^2,\ \sigma_p = 1.971 \times 10^{-35}\,\mathrm{m}^2$$

It makes sense that the cross-section of an electron's scattering area is larger than that of a proton, since the electron is free to move over a much larger area than the proton is; i.e., assuming an atom's center-of-mass is fixed, the volume (and therefore surface area) occupied by the nucleus is tiny compared to the amount of volume and surface area the electrons are allowed to roam, and therefore it makes sense that the electron scatters more radiation per the same overall area.

The "classical electron radius" is the term inside the square, i.e., $\frac{q^2}{4\pi\epsilon_0 mc^2}$:

$$r_{0_e} = 2.817 \times 10^{-15}\,\mathrm{m}$$

N_2 **blues** N_2 characteristic frequency ("transition") has $\lambda \approx 75\,\mathrm{nm}$ (ultraviolet). Number density per unit surface area "footprint" on Earth's surface of N_2 is $n = 1.68 \times 10^{25}\,\mathrm{cm}^{-2}$.

a) Blue sunlight has $\lambda \approx 450\,\mathrm{nm}$. The transmitted energy flux through the atmosphere

$$\langle S\rangle_{tr} = \langle S_0\rangle e^{-N\sigma z}$$

where S_0 is the initial unscattered energy flux, N is the scatterer number density (per unit volume), σ is the scattering cross-section, and z is the distance the light has to travel through. In this case, $Nz = n$ is the number of scatterers per cross-sectional area. Thus the percentage of light scattered is

$$\% \text{ light scattered} = \frac{\langle S_0 \rangle - \langle S_0 \rangle e^{-n\sigma}}{\langle S_0 \rangle} = 1 - e^{-n\sigma}$$

where σ is the Rayleigh scattering cross-section.

$$\omega = \frac{2\pi c}{\lambda} = 4.186 \times 10^{15} \text{rad/s}; \ \omega_0 = \frac{2\pi c}{\lambda_0} = 2.512 \times 10^{16} \text{rad/s}$$

$$\sigma_{ray} = \left(\frac{\omega^2}{\omega^2 - \omega_0^2} \right)^2 \sigma_{th}$$

$$= \left(\frac{(4.186 \times 10^{15} \text{ rad/s})^2}{(4.186 \times 10^{15} \text{ rad/s})^2 - (2.512 \times 10^{16} \text{ rad/s})^2} \right)^2 (6.650 \times 10^{-29} \text{ m}^2)$$

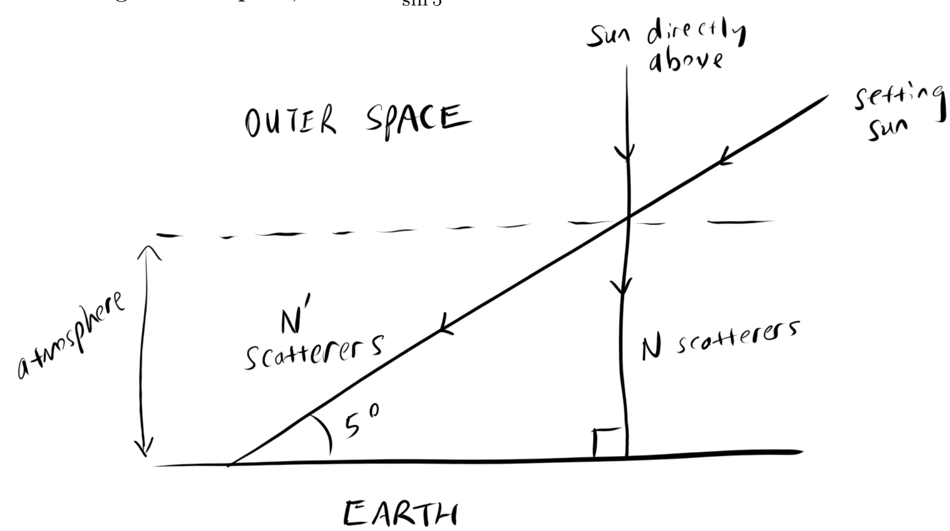
$$= 5.425 \times 10^{-32} \text{ m}^2$$

Thus

$$\% \text{ light scattered} = 1 - \exp(-(1.68 \times 10^{29} \text{ m}^{-2})(5.425 \times 10^{-32} \text{ m}^2))$$

$$= 0.907\%$$

b) The path taken by a wave of light from the Sun when the angle of elevation is 5° is much longer than that when it is shining directly overhead. Assuming that the Earth is a planar slab and the atmosphere is a constant height above the surface of the Earth, the ratio of the vertical path vs. the path at 5° elevation is $\sin 5°$. The number of electrons that scatter the blue light should be roughly proportional to the length of the path, so $n' = \frac{n}{\sin 5°}$.



$=$

Thus the relative amount of light scattered is

$$\% \text{ light scattered} = 1 - e^{-n'\sigma} = 1 - \exp\left(-\frac{n\sigma}{\sin(5^\circ)}\right) = 9.93\%$$

v and c The Maxwell equations for materials:

$$\nabla \cdot \vec{D} = 0,\ \nabla \cdot \vec{B} = 0, \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t},\ \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t}$$

$$\vec{D} = \epsilon \vec{E},\ \vec{B} = \mu \vec{H}$$

a) Solving the wave equation for $\vec{E}$ and $\vec{B}$:

$$\nabla \times (\nabla \times \vec{E}) = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = \nabla \times \left(-\frac{\partial \vec{B}}{\partial t}\right)$$

$$\nabla(0) - \nabla^2 \vec{E} = -\mu \frac{\partial}{\partial t}(\nabla \times \vec{H})$$

$$\nabla^2 \vec{E} = -\mu \frac{\partial}{\partial t}\frac{\partial \vec{D}}{\partial t} = \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2}$$

This is the wave equation for $\vec{E}$ with $v = \frac{1}{\sqrt{\mu\epsilon}}$. Similarly, for $\vec{B}$:

$$\nabla \times \vec{H} = \frac{1}{\mu}(\nabla(\nabla \cdot \vec{B}) - \nabla^2 \vec{B}) = \nabla \times \left(\frac{\partial \vec{D}}{\partial t}\right)$$

$$\frac{1}{\mu}(\nabla(0) - \nabla^2 \vec{B}) = \epsilon \frac{\partial}{\partial t}(\nabla \times \vec{E})$$

$$-\frac{1}{\mu}\nabla^2 \vec{B} = \epsilon \frac{\partial}{\partial t}\left(-\frac{\partial \vec{B}}{\partial t}\right)$$

$$\nabla^2 \vec{B} = \mu\epsilon \frac{\partial^2 \vec{B}}{\partial t^2}$$

This is the wave equation for $\vec{B}$ with the same v. Since $\vec{B}$ and $\vec{E}$ are scaled versions of $\vec{D}$ and $\vec{H}$ and the wave equation is linear, these fields also are solutions to the wave equation.

b) $\kappa_e = \epsilon/\epsilon_0 = 1 + \chi_e$ (result from lecture). Let $\mu \approx \mu_0$. Then

$$n = \frac{c}{v} = \frac{(\mu_0\epsilon_0)^{-1/2}}{(\mu_0\epsilon)^{-1/2}} = \sqrt{\frac{\mu_0\epsilon}{\mu_0\epsilon_0}} = \sqrt{\kappa_e}$$

Dielectric raindrops and rods The subscript $_1$ indicates entities just outside the droplet, and $_2$ indicates those inside the raindrop. The $\vec{E}$-field within the raindrop is uniform and horizontal. Boundary conditions imposed by material Maxwell's equations:

$$\hat{n} \cdot (\vec{D}_2 - \vec{D}_1 = 0), \ \hat{n} \times (\vec{E}_2 - \vec{E}_1) = 0$$

Normal component of $\vec{E}_1$:

$$\vec{E}_1^{\perp} = \frac{1}{\epsilon_1}\vec{D}_1^{\perp} = \frac{1}{\epsilon_1}\vec{D}_2^{\perp} = \frac{1}{\epsilon_1}\vec{D}_2 \cos\theta = \frac{\epsilon_2}{\epsilon_1}\vec{E}_2 \cdot \hat{r}$$

Tangential component of $\vec{E}_1$:

$$\vec{E}_1^{||} = \vec{E}_2^{||} = \vec{E}_2 \times \hat{r}$$

Let $\epsilon_2/\epsilon_1 = 2$. Thus, the vector $\vec{E}_1$ in $(\hat{r}, \hat{\theta})$ coordinates is

$$\left(2\vec{E}_2 \cdot \hat{r}, \vec{E}_2 \times \hat{r}\right) = (2E_2 \cos\theta, E_2 \sin\theta)$$

In Cartesian coordinates $(\hat{x}, \hat{y})$, this is:

$$\left(2E_2 \cos^2\theta + E_2 \sin\theta \cos\left(\theta - \frac{\pi}{2}\right), 2E_2 \cos\theta \sin\theta + E_2 \sin\theta \sin\left(\theta - \frac{\pi}{2}\right)\right)$$

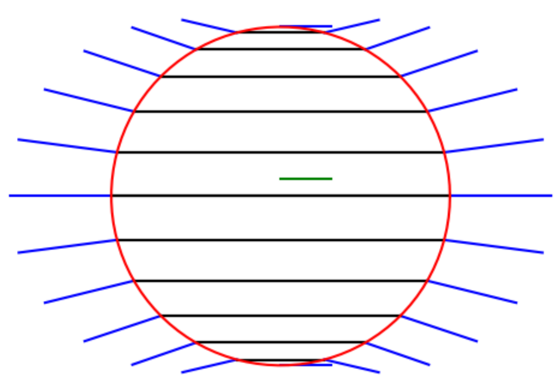


Figure 2. Droplet and E-fields. Black lines indicate the horizontal orientation of the $\vec{E}$ field. The blue vectors indicate the $\vec{E}$ field directly outside the droplet. The green vector indicates the magnitude of the (uniform) $\vec{E}$ within the droplet. The relative magnitude of the blue vectors and the green vectors are drawn to scale, with $\epsilon_2/\epsilon_1 = 2$. (The green vector and blue vectors are all pointing right, but arrowheads are not shown.)

PSET 4

Jonathan Lam

March 11, 2020

(Dielectric) Raindrops keep falling on my head

$$\phi_1 = -E_0 r\cos\theta + \frac{1}{4\pi\epsilon_1}\frac{A_1}{r^2}\cos\theta,\ \phi_2 = \frac{1}{4\pi\epsilon_2}B_1 r\cos\theta$$

(a) *Find the radial electric fields.*

$$E_1 = -\frac{\partial\phi_1}{\partial r} = E_0\cos\theta + \frac{1}{2\pi\epsilon_1}\frac{A_1}{r^3}\cos\theta$$

$$E_2 = -\frac{\partial\phi_2}{\partial r} = -\frac{1}{4\pi\epsilon_2}B_1\cos\theta$$

(b) *Find the unknown coefficients A_1 and B_1 by imposing the two matching conditions at $r = R$.*

Continuity of the normal D field:

$$\epsilon_1 E_1(R) = \epsilon_2 E_2(R)$$

$$\epsilon_1\left(E_0\cos\theta + \frac{1}{2\pi\epsilon_1}\frac{A_1}{R^3}\cos\theta\right) = \epsilon_2\left(-\frac{1}{4\pi\epsilon_2}B_1\cos\theta\right)$$

Continuity of the potential (continuity of the tangential E field):

$$\phi_1(R) = \phi_2(R)$$

$$-E_0 R\cos\theta + \frac{1}{4\pi\epsilon_1}\frac{A_1}{R^2}\cos\theta = \frac{1}{4\pi\epsilon_2}B_1 R\cos\theta$$

$$\begin{cases}\epsilon_1 E_0 + \frac{1}{2\pi}\frac{A_1}{R^3} = -\frac{\epsilon_2}{4\pi\epsilon_2}B_1\\ -E_0 R + \frac{1}{4\pi\epsilon_1}\frac{A_1}{R^2} = \frac{1}{4\pi\epsilon_2}B_1 R_1\end{cases}$$

$$\begin{cases}\epsilon_1 E_0 + \frac{1}{2\pi}\frac{A_1}{R^3} = -\frac{1}{4\pi}B_1\\ -\epsilon_2 E_0 + \frac{\epsilon_2}{4\pi\epsilon_1}\frac{A_1}{R^3} = \frac{1}{4\pi}B_1\end{cases}$$

$$(\epsilon_1 - \epsilon_2)E_0 + \frac{2\epsilon_1+\epsilon_2}{4\pi\epsilon_1}\frac{A_1}{R^3} = 0$$

$$A_1 = \frac{4\pi\epsilon_1 R^3}{2\epsilon_1 + \epsilon_2}(\epsilon_2 - \epsilon_1)E_0$$

Plugging A_1 back into an earlier equation to solve for B_1:

$$\epsilon_1 E_0 + \frac{1}{2\pi}\frac{A_1}{R^3} = \epsilon_1 E_0 + \frac{2\epsilon_1 R^3}{2\epsilon_1 + \epsilon_2}(\epsilon_2 - \epsilon_1)E_0 = -\frac{1}{4\pi}B_2$$

$$B_1 = -4\pi\epsilon_1 E_0 + \frac{8\pi\epsilon_1 R^3}{2\epsilon_1 + \epsilon_2}(\epsilon_1 - \epsilon_2)E_0 = 4\pi\epsilon_1 \left(\frac{2R^3}{2\epsilon_1 + \epsilon_2}(\epsilon_1 - \epsilon_2) - 1\right)E_0$$

...simplify a little...

$$B_1 = \frac{-12\pi\epsilon_1\epsilon_2 E_0}{2\epsilon_1 + \epsilon_2}$$

(c) *Show that A_1 is a dipole moment (the "effective dipole moment" of the sphere).*

$$\phi_{dip} = \frac{1}{4\pi\epsilon_0}\frac{\vec{p}\cdot\vec{r}}{r^3} = \frac{1}{4\pi\epsilon_0}\frac{p}{r^2}\cos\theta$$

Looking at the voltage equation outside of the sphere, the second term (the voltage caused by the sphere and not the external electric force) is:

$$\phi_{sph} = \frac{1}{4\pi\epsilon_1}\frac{A_1}{r^2}\cos\theta$$

Thus, A_1 is analogous to p in a dipole.

$p_{eff} = KR^3E_0$. *Define the Clausius-Mossotti function: $K = K(\epsilon_1, \epsilon_2)$ What do we know about K?*

$$K = \frac{4\pi\epsilon_1}{2\epsilon_1 + \epsilon_2}(\epsilon_2 - \epsilon_1)$$

$\epsilon_2 > \epsilon_1 \Rightarrow K > 0$, and vice versa; and $\epsilon_1 = \epsilon_2 \Rightarrow K = 0$.

T and R *Based on derivations from lecture, we get the reflected and transmitted E field amplitudes at normal incidence:*

$$E_r = E_i\left(\frac{n_2 - n_1}{n_2 + n_1}\right),\ E_t = E_i\left(\frac{2n_1}{n_2 + n_1}\right)$$

(a) *Find the reflected and transmitted energy flux density $\vec{S}$ and then the time average $\langle\vec{S}\rangle$.*

(Non-vectorized fields indicate their magnitudes: e.g., $E_i = |\vec{E}_i|$)

$$E = Bv = \mu H\frac{1}{\sqrt{\mu\epsilon}} \Rightarrow H = \sqrt{\frac{\epsilon}{\mu}}E = \frac{1}{Z}E$$

$$\vec{S} = \vec{E}\times\vec{H} = EH\hat{k} = \frac{1}{Z}E^2\hat{k} = \frac{1}{Z}(\vec{E}\cdot\vec{E})\vec{k}$$

Reflected Poynting vector:

$$\vec{S}_r = \frac{1}{Z_1}\left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 E_i^2 \hat{k},\ \langle \vec{S}_r \rangle = \frac{1}{2Z_1}\left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 E_{i_0}^2 \hat{k}$$

Transmitted Poynting vector:

$$\vec{S}_t = \frac{1}{Z_1}\left(\frac{2n_1}{n_2 + n_1}\right)^2 E_i^2 \hat{k},\ \langle \vec{S}_t \rangle = \frac{1}{2Z_1}\left(\frac{2n_1}{n_2 + n_1}\right)^2 E_{i_0}^2 \hat{k}$$

(b) *Show that energy is conserved.*
To show this, need to show that $E_i = E_r + E_t$.

$$E_i = E_i \frac{n_2 - n_1 + 2n_1}{n_2 + n_1} = E_i \frac{n_2 - n_1}{n_2 + n_1} + E_i \frac{2n_1}{n_2 + n_1} = E_r + E_t$$

(c) *Compute the reflection and transmission coefficients, R and T, for power and show that $R + T = 1$.*

$$R = \frac{\langle S_r \rangle}{\langle S_i \rangle} = \frac{\frac{1}{2Z_1}\left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 E_{i_0}^2}{\frac{1}{2Z_1} E_{i_0}^2} = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2$$

$$T = \frac{\langle S_t \rangle}{\langle S_i \rangle} = \frac{\frac{1}{2Z_2}\left(\frac{2n_1}{n_2 + n_1}\right)^2 E_{i_0}^2}{\frac{1}{2Z_1} E_{i_0}^2} = \frac{Z_1}{Z_2}\left(\frac{2n_1}{n_2 + n_1}\right)^2 = \frac{n_2}{n_1}\left(\frac{2n_1}{n_2 + n_1}\right)^2$$

(Note that $n = c/v = \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}} \approx \sqrt{\epsilon/\epsilon_0} \propto \sqrt{\epsilon/\mu} = Z^{-1}$.)

$$R + T = \frac{(n_2 - n_1)^2}{(n_2 + n_1)^2} + \frac{4n_1 n_2}{(n_2 + n_1)^2} = \frac{n_2^2 - 2n_1 n_2 + n_1^2 + 4n_1 n_2}{(n_1 + n_2)^2}$$

$$= \frac{(n_1 + n_2)^2}{(n_1 + n_2)^2} = 1$$

Evanescence *Find the e-folding distance, z_0 for the evanescent transmitted wave of TIR. Make a sketch of the moving "truncated plane wave." What about the $\vec{B}$ field for this situation?*

$$\vec{E}_t = \vec{E}_{t_0} \exp(i(\omega t - \vec{k_2} \cdot \vec{r})) = \vec{E}_{t_0} \exp(i(\omega t - k_2 y \sin\theta_2 - ik_2 z \cos\theta_2))$$

Making the substitutions:

$$\cos\theta_2 = \sqrt{1 - \sin^2\theta_2},\ \sin\theta_2 = \frac{n_1}{n_2}\sin\theta_1,\ \frac{n_2}{n_1} = \sin\theta_{crit}$$

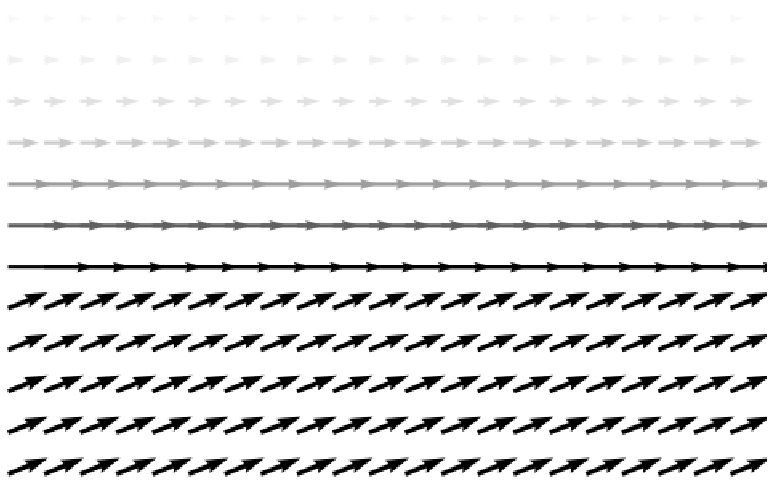
Then:

$$\vec{E}_t = \vec{E}_{t_0} \exp\left(-k_2 z \sqrt{\left(\frac{\sin\theta_1}{\sin\theta_{crit}}\right)^2 - 1}\right) \exp\left(i\left(\omega t - k_2 y \frac{n_2}{n_1}\sin\theta_1\right)\right)$$

The first exponential is real and denotes the magnitude (and thus the damping). Since the exponent is $-z/z_0$, then

$$z_0 = \frac{1}{k_2\sqrt{\left(\frac{\sin\theta_1}{\sin\theta_{crit}}\right)^2 - 1}}$$

Figure 1: Vectors indicate the direction of $\vec{k}$, scaled to the relative magnitude E, near an interface. Opacity is is proportional to length (to make seeing the exponential decay of E easier).



Since this evanescent wave is traveling completely in the y direction, this wave is traveling parallel to the interface. Thus $\vec{E}$ and $\vec{B}$ of the evanescent wave both lie on planes perpendicular to the surface and perpendicular to $\hat{k}$. Since $E_0 \propto B_0$, the $\vec{B}$ field also diminishes exponentially w.r.t. distance from the interface like the $\vec{E}$ field.

Don't drink it? Given a material in which $\vec{M} = s\dot{\vec{P}}$:

(a) *Show that $E \cdot \dot{E} = 0$.*

$$\vec{M} = \chi_m \vec{H} = s\dot{\vec{P}} = s\frac{\partial}{\partial t}(\epsilon_0 \chi_e \vec{E}) = s\epsilon_0\chi_e\dot{\vec{E}}$$

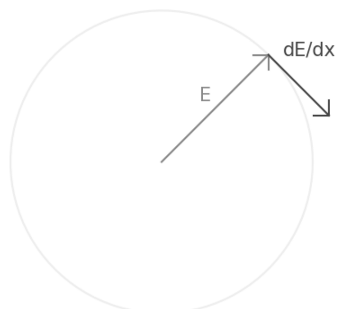
$$\dot{\vec{E}} = c\vec{H}, \ c = \frac{\chi_m}{s\epsilon_0\chi_e}$$

$$\vec{E}\cdot\dot{\vec{E}} = c(\vec{E}\cdot\dot{\vec{H}}) = c(0) = 0$$

(b) *Give a physical explanation for this result.*

This means that the vector $\vec{E}$ is rotating only (i.e., not changing magnitude).

Figure 2: $\vec{E} \perp \dot{\vec{E}}$

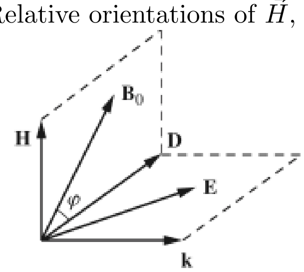


Big magnets and powerful lasers! *The following constituent equations are given for the non-simple optical material, Leelanium:*

$$\vec{B} = \mu\vec{H}, \ \vec{D} = \epsilon\vec{E} - i\gamma\vec{B}\times\vec{E}$$

Treat the wave as a plane wave with wave vector $\vec{k}$, where ϕ is the angle between the field of the magnet and the $\vec{D}$ field of the plane wave.

Figure 3: Relative orientations of $\vec{H}$, $\vec{B}_0$, $\vec{D}$, $\vec{E}$, $\hat{k}$



(a) *You have a very strong 7.5 W laser and a strong magnetic field $\vec{B}_0$ of 1 T. Find the magnitude of the laser's $\vec{B}$ field, assuming the beam has a diameter of 2mm. How large is B_0/B?*

Since $\vec{S}$ is an energy flux (energy per unit area per unit time, or power per unit area):

$$\langle S\rangle = \frac{P}{A} = \frac{7.5\text{W}}{\pi(1\times10^{-3}\text{m})^2} = 2.39\times10^6\ \text{W/m}^2$$

But also

$$\langle S\rangle = \langle u\rangle c = \frac{1}{2\mu_0}B^2 c$$

Thus

$$B = \sqrt{\frac{2\mu_0\langle S\rangle}{c}} = \sqrt{\frac{2(4\pi\times10^{-7}\text{H/m})(2.39\times10^6\text{W/m}^2)}{3.00\times10^8\text{m/s}}}$$

$$= 1.42\times10^{-4}\text{T}$$

Thus

$$\frac{B_0}{B} = \frac{1\text{T}}{1.42\times10^{-4}\text{T}} = 7.07\times10^4$$

Thus it is safe to assume $B_0 \gg B$.

(b) *Solve for the vector $\vec{H}$ in terms of $\vec{k}$ and $\vec{E}$. Then find the Poynting vector. Simplify and express the wave-vector as $\vec{k} = k\hat{k}$.*

Finding $\vec{H}$:

$$\vec{B} = \mu\vec{H}$$

$$\nabla\times\vec{E} = -\frac{\partial}{\partial t}(\mu\vec{H})$$

$$i\vec{k}\times\vec{E} = i\omega\mu\vec{H}$$

$$\vec{H} = \frac{1}{\mu\omega}\vec{k}\times\vec{E}$$

Finding $\vec{S}$:

$$\vec{S} = \vec{E}\times\vec{H} = \vec{E}\times\left(\frac{1}{\mu\omega}\vec{k}\times\vec{E}\right) = \frac{1}{\mu\omega}\left(\vec{k}(\vec{E}\cdot\vec{E}) - (\vec{k}\cdot\vec{E})\vec{E}\right)$$

$$= \frac{k}{\mu\omega}\left(E^2\hat{k} - (\hat{k}\cdot\vec{E})\vec{E}\right)$$

(c) *Use $\nabla\cdot\vec{D}$ to show that $\hat{k}\cdot\vec{E} = -i\gamma/\epsilon^2 B_0 D\sin\phi$. Assume (i) $B_0 \gg B$, and (ii) $\vec{E}\approx\vec{D}/\epsilon$.*

$$\nabla\cdot\vec{D} = i\vec{k}\cdot\vec{D} = 0 \Rightarrow \hat{k}\cdot\vec{D} = 0$$

$$\hat{k}\cdot(\epsilon\vec{E} - i\gamma\vec{B}\times\vec{E}) = \epsilon\hat{k}\cdot\vec{E} - i\gamma\hat{k}\cdot(\vec{B}\times\vec{E}) = 0$$

Using the approximations, we can replace $\vec{E} = \vec{D}/\epsilon$, and since almost all of the $\vec{B}$ field comes from $\vec{B_0}$, replace $\vec{B} = \vec{B_0}$.1

$$\hat{k} \cdot \vec{E} = \frac{i\gamma}{\epsilon} \hat{k} \cdot (\vec{B} \times \vec{E}) = \frac{i\gamma}{\epsilon^2} \hat{k} \cdot (\vec{B_0} \times \vec{D})$$

From (Figure 3), we can tell that $\vec{B_0} \times \vec{D}$ is in the direction $-\hat{k}$. Thus this further simplifies to

$$\hat{k} \cdot \vec{E} = \frac{i\gamma}{\epsilon^2} \hat{k} \cdot (B_0 D \sin\phi(-\hat{k})) = \frac{i\gamma}{\epsilon^2} B_0 D \sin\phi(\hat{k} \cdot (-\hat{k})) = -\frac{i\gamma}{\epsilon^2} B_0 D \sin\phi$$

(d) *Use (c) to simplify the expression for $\vec{S}$ from (b). Explain why this result shows the laser bends away from the direction it would travel in simple material. At what angle ϕ is the bending at a maximum?*

$$\vec{S} = \frac{k}{\mu\omega} \left(E^2 \hat{k} - \left(-\frac{i\gamma}{\epsilon^2} B_0 D \sin\phi \right) \vec{E} \right) = \frac{k}{\mu\omega} \left(E^2 \hat{k} + \frac{i\gamma}{\epsilon^2} B_0 D \sin\phi \vec{E} \right)$$

In simple material (or without a magnetic field), $\vec{S}$ should be only in the direction of $\hat{k}$ (in the direction of wave propagation). Here, we see some component of the Poynting vector in the direction of $\vec{E}$, which is almost perpendicular to $\vec{k}$.

The bending of $\vec{S}$ is greatest when $\sin\phi = 1$, i.e., when $\phi = \pi/2$, i.e., when the external (strong) $\vec{B_0}$ field is perpendicular to the $\vec{E}$ of the wave.

Pset 5 – PH214C

Jonathan Lam

April 16, 2020

Don't box me in *We wish to derive the final coefficient in the classical formula for energy density,*

$$u = kT\left(\frac{8\pi\nu^2}{c^3}\right)$$

The quantity we are trying to find is the density of states w.r.t. frequency and space. In other words, if N represents number of states, then we are trying to find

$$\frac{\partial^2 N}{\partial V \partial \nu}$$

The modes must all have nodes in the x, y, and z directions. Since these nodes are stationary, the EM wave must be a standing wave:

$$\vec{E} = \vec{E}_0 \sin(k_x x)\sin(k_y y)\sin(k_z z)\sin(\omega t)$$

with the node imposing the condition (assume the box is a cube with length L):

$$\sin(k_x L) = 0 \Rightarrow k_x L = \pi n_x \ (n_x \in \mathbb{Z}_+)$$

in the x, y, and z directions. By substituting $k = \frac{2\pi}{\lambda} = \frac{2\pi c}{\nu}$,

$$\nu_x = n_x \frac{c}{2L} \ (n_x \in \mathbb{Z}_+)$$

This is a set of discrete frequencies in one direction. More generally, a three-dimensional mode can be uniquely defined by its tuple of (n_x, n_y, n_z) coordinates, and so we define this as the phase space. In this space, there is clearly one mode at every (discrete integer-tuple) coordinate (but only existing in the first octant), so the number density δ is:

$$\delta = \frac{1 \text{ mode}}{1 \text{ (unit volume)}}$$

where, if N is a number of modes and V is a (dimensionless) "volume" in this coordinate space, then

$$N = Vd$$

In other words, the number of modes is numerically equal to the volume in this space. Let

$$n \coloneqq \sqrt{n_x^2 + n_y^2 + n_z^2}$$

indicate the magnitude of the position vector in this space and

$$\nu \coloneqq \sqrt{\nu_x^2 + \nu_y^2 + \nu_z^2}$$

indicate the magnitude of the frequency. To find the "density" of this space with respect to frequency, we can find the number of modes with a frequency (magnitude) between ν_0 and $\nu_0 + d\nu$ and divide by $d\nu$ (i.e., differentiating w.r.t. frequency). The change in volume in the first octant is an eighth of a thin spherical shell with thickness ν. Since the number density is equal to the volume in the n coordinates, we begin by stating the volume of such a shell in n coordinates and then convert it to ν.

$$dN = dV = \frac{1}{8}\left(4\pi n^2\right)\, dn$$

$$n = \frac{2L}{c}\nu,\ dn = \frac{2L}{c}d\nu$$

$$\frac{dN}{d\nu} = \frac{4\pi L^3 \nu^2}{c^3}$$

Now, if we wish to find the density w.r.t. volume (in the regular coordinate system), we can divide by the volume of the cube, L^3. (It is a derivative in the limit of $L \to 0$.)

$$\frac{\partial^2 N}{\partial V \partial \nu} = \frac{4\pi \nu^2}{c^3}$$

We can follow a similar approach to find the analogue w.r.t. λ. Make the substitutions $\nu = c/\lambda$ and $d\nu = -c/\lambda^2\, d\lambda$ to arrive at the analogous density w.r.t. space and wavelength:

$$\frac{\partial^2 N}{\partial V \left(-\frac{c}{\lambda^2}\partial\lambda\right)} = \frac{4\pi\left(\frac{c}{\lambda}\right)^2}{c^3}$$

$$\frac{\partial^2 N}{\partial V \partial \lambda} = -\frac{4\pi}{\lambda^4}$$

Note that, in this derivative formulation, the "density" of modes w.r.t. λ is negative. To get the desired factor of 8, we note that there are actually two valid polarizations for each (n_x, n_y, n_z) tuple, so we double both of these results:

$$\frac{\partial^2 N}{\partial V \partial \nu} = \frac{8\pi \nu^2}{c^3}$$

$$\frac{\partial^2 N}{\partial V \partial \lambda} = -\frac{8\pi}{\lambda^4}$$

Certainly! *Reformulate the wave packet-based derivation in the notes so that we get, for the uncertainty product:*

$$\Delta k \Delta x \geq \frac{1}{2}$$

Using a Gaussian for $A(k)$ and the given measure of spread (the width of the Gaussian at height e^{-1}) gives the constant 4. Thus, if we reformulate the problem to use a different measure of spread (i.e., to standard deviation), we can change the constant that the uncertainties of wave number and position multiply to. Start with the same Gaussian wave number amplitude distribution:

$$A(k) = e^{-\frac{\alpha}{2}(k-k_0)^2}$$

The inverse Fourier transform of this wave number (at $t = 0$) is the waveform w.r.t. position (which is also a Gaussian amplitude distribution):

$$\Psi(x, 0) = \int_{-\infty}^{\infty} dk\, A(k) e^{ikx} = \sqrt{\frac{2\pi}{\alpha}} e^{ik_0 x} e^{-\frac{x^2}{2\alpha}}$$

(We neglect the tedious details of the inverse Fourier transform calculation here and use the result from the lecture notes.) Noting that the squares of A and Ψ form the PDF $f(k)$ of the wave number and PDF $g(x)$ of the position, respectively:

$$f(k) = A^2(k) = e^{-\alpha(k-k_0)^2}$$

$$g(x) = |\Psi^2(x, 0)| = \frac{2\pi}{\alpha} e^{-\frac{x^2}{\alpha}}$$

The general form of a Gaussian PDF is

$$h(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Since f and g are clearly in the form of a Gaussian (although not normalized, but this doesn't affect spread), we can see by matching the exponent to the Gaussian form that:

$$\sigma_k = \frac{1}{\sqrt{2\alpha}}$$

$$\sigma_x = \sqrt{\frac{\alpha}{2}}$$

Using this measure of spread instead of the Δ-spread defined in the lecture notes, we get:

$$\sigma_k \sigma_x = \sqrt{\frac{1}{2\alpha}} \sqrt{\frac{\alpha}{2}} = \frac{1}{2}$$

which saturates the stated lower bound of $\frac{1}{2}$. (This doesn't show the inequality, but we obtain a product lower than 4 from the lecture notes.)

PH214C – Pset 6

Jonathan Lam

April 16, 2020

1 A Bohr result

Suppose that an integer number of de Broglie wavelengths fit around a circular Bohr orbit ($n \in \mathbb{Z}_+$). Find an expression for the quantized angular momentum for general n and sketch the case for $n = 4$.

$$l = mvr \qquad \text{(angular momentum)}$$

$$\lambda = \frac{h}{mv} \qquad \text{(de Broglie wavelength)}$$

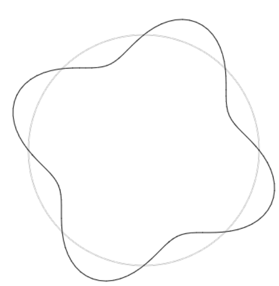
$$2\pi r = n\lambda \qquad \text{(quantized orbit)}$$

Thus

$$l = \frac{h}{\lambda} r = \frac{hn}{2\pi r} r = n\bar{h}$$

In the case where $n = 4$,

$$l_4 = 4\bar{h} = 4.05 \times 10^{-34}\,\mathrm{kg\,m^2\,s^{-1}}$$



2 Wave packet

Show that our definition of the wave function of a wave packet

$$\psi(x,t) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dp\, \phi(p) e^{i(px-Et)/\hbar}$$

solves the free-particle Schödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}$$

but not the generic wave equation:

$$\frac{\partial^2\psi}{\partial t^2} = a^2\frac{\partial^2\psi}{\partial x^2}$$

2.1 Free-particle Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = i\hbar\left(-\frac{iE}{\hbar}\psi\right) \qquad = E\psi \qquad \text{(LHS)}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = -\frac{\hbar^2}{2m}\left(\left(\frac{ip}{\hbar}\right)^2\psi\right) = \frac{p^2}{2m}\psi \qquad = E\psi \qquad \text{(RHS)}$$

2.2 Generic wave equation

$$\frac{\partial^2\psi}{\partial t^2} = \left(\frac{-iE}{\hbar}\right)^2\psi \qquad = -\frac{E^2}{\hbar^2}\psi \qquad \text{(LHS)}$$

$$a^2\frac{\partial^2\psi}{\partial x^2} = a^2\left(\frac{ip}{\hbar}\right)^2\psi \qquad = -\frac{a^2p^2}{\hbar^2}\psi \qquad \text{(RHS)}$$

For ψ to satisfy the wave equation, both sides must be equal:

$$-\frac{E^2}{\hbar^2}\psi = -\frac{a^2p^2}{\hbar^2}\psi$$

$$E^2 \propto p^2$$

$$\text{but } E = \frac{p^2}{2m} \Rightarrow \text{contradiction}$$

3 Probability

We found the peak probability and expectation value $\langle x\rangle$ of the wave function

$$\psi(x) = 2a^{3/2}xe^{-ax}, \text{ for } x > 0, a = \text{constant}$$

Verify these results and then find the probability that the particle will be found between $x = 0$ and $x = 1/a$.

3.1 Peak probability

This occurs when $P = |\psi|^2$ is maximized. (Assume $a > 0$.)

$$\begin{aligned}
P(x) &= |\psi(x)|^2 = 4a^3x^2e^{-2ax} \\
\frac{dP}{dx} &= 8a^3x(1-ax)e^{-2ax} = 0 \\
\frac{d^2P}{dx^2} &= 8a^3(1-4ax+2a^2x^2)e^{-2ax}
\end{aligned}$$

This equation is satisfied when $x = 0$ or $x = 1/a$. Using the second derivative:

$$\begin{aligned}
\left.\frac{d^2P}{dx}\right|_{x=0} &= 8a^3 > 0 \\
\left.\frac{d^2P}{dx}\right|_{x=1/a} &= -8a^3e^{-2} < 0
\end{aligned}$$

Thus the only local maximum occurs at $x = 1/a$. (Also, since the probability dies out as $x \to \infty$, this must be the global maximum in the interval $[0, \infty)$.)

3.2 Expectation value

Since $\psi = 0$ for negative x, we can begin the integration at $x = 0$. Since ψ and x are real, $P = |\psi|^2 = \psi^*\psi = \psi^2$, and $\psi^*x\psi = x\psi^2 = xP$.

$$\begin{aligned}
\langle x \rangle &= \int_{-\infty}^{\infty} dx\, \psi^* x \psi \\
&= \int_0^{\infty} dx\, x(4a^3x^2e^{-2ax}) \\
&= \frac{1}{4a}\int_0^{\infty} dy\, y^3e^{-y} && (y = 2ax) \\
&= \frac{1}{4a}\left[e^{-y}(-y^3-3y^2-6y-6)\right|_{x=0}^{\infty} && \text{(a lot of integration by parts)} \\
&= \frac{3}{2a}
\end{aligned}$$

3.3 Probability of position between 0 and $1/a$

This is just an integral of probability over some interval of x.

$$\begin{aligned}
P\left(0 \leq x \leq \frac{1}{a}\right) &= \int_0^{1/a} dx\, P(x) \\
&= \int_0^{1/a} dx\, (4a^3 x^2 e^{-2ax}) \\
&= \frac{1}{2}\int_0^2 dy\, y^2 e^{-y} && (y = 2ax) \\
&= \frac{1}{2}\left[e^{-y}(-y^2 - 2y - 2)\right|_0^2 && \text{(integration by parts)} \\
&= \frac{1}{2}(e^{-2}(-4-4-2) - e^0(-2)) \\
&= 1 - 5e^{-2} \approx 0.32
\end{aligned}$$

4 Momentum keeps it real

Show that $\langle p\rangle - \langle p\rangle^* = 0$ to show that the operator for p is Hermitian.

$$\hat{p} \coloneqq \frac{\hbar}{i}\frac{\partial}{\partial x}$$

$$\begin{aligned}
\langle p\rangle - \langle p\rangle^* &= \left[\int_{-\infty}^{\infty} dx\, \psi^* \left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\psi\right] - \left[\int_{-\infty}^{\infty} dx\, \psi^* \left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\psi\right]^* \\
&= \frac{\hbar}{i}\left[\int_{-\infty}^{\infty} dx\, \psi^* \frac{\partial\psi}{\partial x}\right] - \left(\frac{\hbar}{-i}\right)\left[\int_{-\infty}^{\infty} dx\, \psi \frac{\partial\psi^*}{\partial x}\right] \\
&= \frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \psi^* \frac{\partial\psi}{\partial x} + \psi\frac{\partial\psi^*}{\partial x} \\
&= \frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \frac{\partial}{\partial x}(\psi^*\psi) = \frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \frac{\partial|\psi|^2}{\partial x} \\
&= \frac{\hbar}{i}\lim_{b\to\infty} |\psi|^2\Big|_{x=-b}^{b} = \frac{\hbar}{i}\,[0-0] = 0
\end{aligned}$$

Notes about the math:

- Any valid PDF (e.g., $|\psi|^2$) must approach the value 0 in the limit as the variable goes to $\pm\infty$.
- In general,
  $$\left[\int f(x)\,dx\right]^* = \int [f(x)]^*\,[dx]^*$$
  Since x and dx are real, then the conjugate of the integral is the integral of the conjugate of the integrand.

PH214C – Pset 7

Jonathan Lam

April 23, 2020

Infinite well → positive E

For the infinite square well problem, show that in solving the spatial problem, for the eigenfunctions $U_n(x)$, eigenvalues E_n, that no solutions exist for any $E_n \leq 0$ and show why.

The spatial component $U(x)$ of the wavefunction should satisfy the (one-dimensional) time-invariant Shrödinger equation in the well.

$$-\frac{\hbar^2}{2m}\frac{d^2U}{dx^2} + V(x)U = EU$$

$$V(x) = \begin{cases} \infty & x < 0 \text{ or } x > a \\ 0 & 0 < x < a \end{cases}$$

Since we are looking at the solution to the ODE in the well, where $V = 0$, this reduces to a familiar constant-coefficient second-order homogeneous ODE:

$$\frac{\hbar^2}{2m}\frac{d^2U}{dx^2} + EU = 0$$

The solution will be of the form: $U(x) = Ae^{s_1 x} + Be^{s_2 x}$. Solving normally:

$$\frac{\hbar^2}{2m}s^2 + E = 0 \qquad \text{(auxiliary equation)}$$

$$s = \left(\frac{-2mE}{\hbar^2}\right)^{1/2}$$

$$U(x) = A\exp\left(\frac{\sqrt{-2mE}}{\hbar}x\right) + B\exp\left(-\frac{\sqrt{-2mE}}{\hbar}x\right) \tag{1}$$

With voltage being infinite past the boundaries of the well:

$$\lim_{V\to\infty}\left[-\frac{\hbar^2}{2m}\frac{d^2U}{dx^2} + VU = EU\right]$$

$$\Rightarrow \lim_{V\to\infty}\left[-\frac{\hbar^2}{2mV}\frac{d^2U}{dx^2} + U = \frac{E}{V}U\right]$$

$$\Rightarrow U = 0 \qquad (x < 0 \text{ or } x > a)$$

we obtain the following boundary conditions for the region in the square well:

$$U(0) = 0$$
$$U(a) = 0$$

Looking back at the form of U, we see that it is the sum of two exponentials. If the exponents are real (if $E \leq 0$), then neither exponential will ever be zero (and neither will their sum), so the wavefunction will never satisfy the boundary conditions. Thus, $E > 0$ for any particular solution to the infinite square well. More generally, the set of all particular solutions (eigenfunctions) $\{U_n\}$ will have corresponding energies (eigenvalues) $\{E_n\}$, where $E_i > 0$. This will still be true for any infinite square well problem, including shifted ones (e.g., the next problem).

Shifted well

Repeat the infinite square well problem, finding solutions (eigenvalues and eigenfunctions), but for a well in the interval $-a/2 < x < a/2$. Explain how your results relate to the original well.

The only change this makes to solving the problem is the boundary conditions. The solution in the square well is still of the same form as (1), but we have to match the shifted boundary conditions:

$$U\left(-\frac{a}{2}\right) = 0$$
$$U\left(\frac{a}{2}\right) = 0$$

The general form of $U(x)$ is given by (1). Reexpressing it with $E > 0$ (which must be true according to question 1), expanding using Euler's formula, simplifying, and letting $k = \sqrt{2mE}/\hbar$:

$$U(x) = C\cos kx + D\sin kx$$

Plugging in the boundary conditions, we obtain:

$$C\cos\left(-\frac{ak}{2}x\right) + D\sin\left(-\frac{ak}{2}x\right) = 0$$

$$C\cos\left(\frac{ak}{2}x\right) - D\sin\left(\frac{ak}{2}x\right) = 0 \tag{2}$$

$$C\cos\left(\frac{ak}{2}x\right) + D\sin\left(\frac{ak}{2}x\right) = 0 \tag{3}$$

$$2C\cos\left(\frac{ak}{2}x\right) = 0 \tag{(2)+(3)}$$

$$2D\sin\left(\frac{ak}{2}x\right) = 0 \tag{(3)-(2)}$$

From the latter two equations, we see that we can cosine functions in the solution if $k = (2n+1)\pi/a$, or sine functions in the solution if $k = 2n\pi/a$, where $n \in \mathbb{Z}$. Note that, since sine and cosine have different zeros, and therefore C and D cannot be simultaneously nonzero. Thus U_n (at a certain energy) must be a pure cosine or sine function. Normalizing (assuming pure cosine):

$$\begin{aligned}
1 &= \int_{-a/2}^{a/2} P(x)\,dx \\
&= \int_{-a/2}^{a/2} U^*(x)U(x)\,dx \\
&= \int_{-a/2}^{a/2} C^2\cos^2\left(\frac{(2n+1)\pi}{a}x\right)\,dx \\
&= \frac{C^2}{2}\int_{-a/2}^{a/2} 1+\cos\left(\frac{(2n+1)\pi}{a}x\right)\,dx \\
&= \frac{C^2}{2}(a+0)
\end{aligned}$$

Thus $C = \sqrt{2/a}$. Similarly, for pure sines, we get the same result: $D = \sqrt{2/a}$. Thus, the eigenfunctions and eigenvalues of this shifted infinite square well are:

$$U_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{2n\pi}{a}x\right) \qquad E_n = \frac{(2n+1)^2\pi^2\hbar^2}{2ma^2}$$

or

$$U_n(x) = \sqrt{\frac{2}{a}}\cos\left(\frac{(2n+1)\pi}{a}x\right) \qquad E_n = \frac{2n^2\pi^2\hbar^2}{ma^2}$$

These are similar to the particular solutions of the infinite square well from $0 < x < a$ in that their amplitudes are the same and they are all pure sinusoids. However, we also allow cosine functions because these new boundary conditions allow for it. However, since the discrete sine and cosine are now spaced at with wave numbers $2n\pi/a$ apart, it looks the same as the infinite square well for $0 < x < a$ with sine waves with wave numbers spaced $n\pi/a$ apart. (I.e., The introduction of cosines is simply an artifact of the parity of sine and cosine and doesn't reflect a change in the shape of the solutions.)

Well expectations

Compute $\langle x\rangle$, $\langle x^2\rangle$, $\langle p\rangle$, $\langle p^2\rangle$, σ_x and σ_p for the n-th state of the infinite square well. Do these states satisfy the uncertainty principle? For what n is the product of uncertainties smallest?

$$\psi_n(x,t) = \frac{\sqrt{2}}{a}\sin\left(\frac{n\pi}{a}x\right)e^{-iEt/\hbar}$$

$$
\begin{aligned}
\langle f(x)\rangle &= \int_{-\infty}^{\infty} dx\, \psi^*(x,t)\, f(x)\, \psi(x,t) \\
&= \int_0^a dx\, \psi^* f \psi && \text{(infinite square well case)}
\end{aligned}
$$

Position expectations

$$
\begin{aligned}
\langle x\rangle_n &= \int_0^a dx, \psi^* x \psi \\
&= \frac{2}{a}\int_0^a dx\, x \sin^2\left(\frac{n\pi}{a}x\right) e^0 \\
&= \frac{2a}{n^2\pi^2}\int_0^{n\pi} du\, u \sin^2 u && (u = \tfrac{n\pi}{a}x) \\
&= \frac{a}{n^2\pi^2}\int_0^{n\pi} du\, u(1-\cos(2u)) \\
&= \frac{a}{n^2\pi^2}\left[\left.\frac{u^2}{2}\right|_0^{n\pi} - \int_0^{n\pi} du\, u\cos(2u)\right] \\
&= \frac{a}{n^2\pi^2}\left[\frac{n^2\pi^2}{2} - 0\right] && \text{(integration by parts)} \\
&= \frac{a}{2}
\end{aligned}
$$

$$
\begin{aligned}
\langle x^2\rangle_n &= \int_0^a dx, \psi^* x^2 \psi \\
&= \frac{2}{a}\int_0^a dx\, x^2 \sin^2\left(\frac{n\pi}{a}x\right) e^0 \\
&= \frac{2a^2}{n^3\pi^3}\int_0^{n\pi} du\, u^2 \sin^2 u && (u = \tfrac{n\pi}{a}x) \\
&= \frac{a^2}{n^3\pi^3}\int_0^{n\pi} du\, u^2(1-\cos(2u)) \\
&= \frac{a^2}{n^3\pi^3}\left[\left.\frac{u^3}{3}\right|_0^{n\pi} - \int_0^{n\pi} du\, u^2\cos(2u)\right] \\
&= \frac{a^2}{n^3\pi^3}\left[\frac{n^3\pi^3}{3} - \left[\frac{x^2}{2}\sin 2x + \frac{x}{2}\cos 2x - \frac{1}{4}\sin 2x\right|_0^{n\pi}\right] \\
&= \frac{a^2}{n^3\pi^3}\left[\frac{n^3\pi^3}{2} - \frac{n\pi}{2}\right] \\
&= a^2\left(\frac{1}{3} - \frac{1}{2n^2\pi^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{x_n}^2 &= \langle x^2\rangle_n - \langle x\rangle_n^2 \\
&= a^2\left(\frac{1}{3} - \frac{1}{2n^2\pi^2} - \left(\frac{a}{2}\right)^2\right) \\
&= a^2\left(\frac{1}{12} - \frac{1}{2n^2\pi^2}\right)
\end{aligned}
$$

Momentum expectations

$$
\hat{p} = -i\hbar\frac{\partial\psi}{\partial x}
$$

$$
\begin{aligned}
\langle p\rangle_n &= \int_0^a dx\,\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)\left[-i\hbar\left(-\frac{n\pi}{a}\right)\sqrt{\frac{2}{a}}\cos\left(\frac{n\pi}{a}x\right)\right] \\
&= \frac{2i\hbar}{a}\int_0^a dx\,\frac{n\pi}{a}\sin\left(\frac{n\pi}{a}x\right)\cos\left(\frac{n\pi}{a}x\right) \\
&= \frac{2i\hbar}{a}\int_0^a d\left(\sin\left(\frac{n\pi}{a}x\right)\right)\,\sin\left(\frac{n\pi}{a}x\right) \\
&= \frac{2i\hbar}{a}\left[\frac{1}{2}\sin^2\left(\frac{n\pi}{a}x\right)\right|_0^a \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\langle p^2\rangle_n &= \int_0^a dx\,\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)\left[-\hbar^2\left(-\frac{n^2\pi^2}{a^2}\right)\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)\right] \\
&= \frac{2}{a}\left(\frac{\hbar n\pi}{a}\right)^2\int_0^a dx\,\sin^2\left(\frac{n\pi}{a}x\right) \\
&= \frac{1}{a}\left(\frac{\hbar n\pi}{a}\right)^2\int_0^a dx\,1-\cos\left(\frac{n\pi}{a}x\right) \\
&= \frac{1}{a}\left(\frac{\hbar n\pi}{a}\right)^2(a-0) \\
&= \left(\frac{\hbar n\pi}{a}\right)^2
\end{aligned}
$$

$$
\sigma_{p_n}^2 = \langle p^2\rangle_n - \langle p\rangle_n^2 = \left(\frac{\hbar n\pi}{a}\right)^2
$$

Uncertainty

$$\begin{aligned}\sigma_{x_n}\sigma_{p_n} &= \sqrt{a^2\left(\frac{1}{12}-\frac{1}{2n^2\pi^2}\right)}\sqrt{\left(\frac{\hbar n\pi}{a}\right)^2} \\ &= \hbar n\pi\sqrt{\frac{1}{12}-\frac{1}{2n^2\pi^2}}\end{aligned}$$

This has the same uncertainty characteristic: the product of uncertainties of position and momentum is a fixed number. (In other words, decreasing uncertainty for either position or momentum will increase uncertainty of the other.) We can see that this is strictly monotonically increasing w.r.t. n, so the minimum uncertainty is at $n = 1$, which gives the value:

$$\sigma_{x_1}\sigma_{p_1} = \hbar\pi\sqrt{\frac{1}{12}-\frac{1}{2\pi^2}} = 0.568\hbar = 5.99\times 10^{-35}\ \text{J s}$$

Mixed well

A particle in an infinite square well has the initial wave function

$$\psi(x,0) = A[U_1(x) + U_2(x)]$$

Normalize ψ

Normalize ψ to find A. Will $\psi = \psi(x,0)$ remain normalized for $t > 0$?

Since ψ is real, we can simply square it to get its magnitude. We also know that $\{U_n(x)\}$ is an orthonormal set, so $\int_{-\infty}^{\infty} dx\, U_n U_m = \delta_{nm}$.

$$\begin{aligned}1 &= \int_0^a dx\,\psi^2(x,0) \\ &= A^2\int_0^a dx\ \left[U_1^2 + 2U_1U_2 + U_2^2\right] \\ &= A^2\left[\int_0^a dx\,U_1^2 + 2\int_0^a dx\,U_1U_2 + \int_0^a dx\,U_2^2\right] \\ &= A^2\left[1 + 2(0) + 1\right] \\ A &= \frac{1}{\sqrt{2}}\end{aligned}$$

(Intuitively, this makes sense: since U_1 and U_2 are normalized, the integral of their PDFs [the PDF of ψ] should sum to 2. Thus $A = 1/\sqrt{2}$.) The wavefunction will remain normalized for $t > 0$. Multiplying in the time-dependent part, the full wavefunction is of the form:

$$\psi(x,t) = \frac{1}{\sqrt{2}} \left[U_1(x) + U_2(x) \right] e^{iEt/\hbar}$$

So

$$\begin{aligned}
&\int_0^a dx\, \psi^* \psi \\
&= \int_0^a dx\, \left[A \left[U_1 + U_2 \right] e^{-iEt/\hbar} \right] \left[A \left[U_1 + U_2 \right] e^{iEt/\hbar} \right] \\
&= \int_0^a dx\, A^2 \left[U_1 + U_2 \right]^2 \\
&= \int_0^a dx\, \psi^2(x,0) \\
&= 1
\end{aligned}$$

PDF

Find $\psi(x,t)$ and $|\psi(x,t)|^2$.

As stated above, the wavefunction as a function of position and time can be obtained by multiplying in the time-dependent part (since we solved the wavefunction PDE using separation of variables):

$$\psi(x,t) = \frac{1}{\sqrt{2}} \left[U_1(x) + U_2(x) \right] e^{iEt/\hbar}$$

Finding the PDF (magnitude-squared distribution) is straightforward:

$$\begin{aligned}
|\psi(x,t)|^2 &= \psi^* \psi \\
&= \left[\frac{1}{\sqrt{2}} \left[U_1 + U_2 \right] e^{-iEt/\hbar} \right] \left[\frac{1}{\sqrt{2}} \left[U_1 + U_2 \right] e^{iEt/\hbar} \right] \\
&= \frac{1}{2} \left[U_1(x) + U_2(x) \right]
\end{aligned}$$

Momentum and position

Find $\langle x \rangle$ and $\langle p \rangle$.

Position

$$
\begin{aligned}
\langle x\rangle &= \int_0^a dx\,\psi^*(x,t)\,x\,\psi(x,t)\\
&= \int_0^a dx\,A^2x\left[U_1e^{-iE_1t/\hbar}+U_2e^{-iE_2t/\hbar}\right]\left[U_1e^{iE_1t/\hbar}+U_2e^{iE_2t/\hbar}\right]\\
&= A^2\left[\int_0^a dx\,xU_1^2+\int_0^a dx\,xU_1U_2\left[e^{i(E_1-E_2)t/\hbar}+e^{i(E_2-E_1)t/\hbar}\right]+\int_0^a dx\,xU_2^2\right]\\
&= A^2[I_1+I_2+I_3]\\
I_1 &= \int_0^a dx\,xU_1^2\\
&= \frac{2}{a}\int_0^a dx\,x\sin^2\left(\frac{\pi}{a}x\right)\\
&= \frac{1}{a}\left[\int_0^a dx\,x-\int_0^a dx\,\cos\left(\frac{2\pi}{a}x\right)\right]\\
&= \frac{1}{a}\left[\frac{x^2}{2}\bigg|_0^a-0\right] && \text{(integration by parts)}\\
&= \frac{a}{2}\\
I_3 &= \frac{a}{2} && \text{(works out same as } I_1\text{)}\\
I_2 &= 2\int_0^a dx\,xU_1U_2\left[e^{i(E_1-E_2)t/\hbar}+e^{i(E_2-E_1)t/\hbar}\right]\\
&= \frac{4}{a}\int_0^a dx\,x\sin\left(\frac{\pi}{a}x\right)\sin\left(\frac{2\pi}{a}x\right)\left[2\cos\left(\frac{E_2-E_1}{\hbar}t\right)\right]\\
&= \frac{4}{a}\cos\left(\frac{E_2-E_1}{\hbar}t\right)\int_0^a dx\,x\left[\cos\left(\frac{\pi}{a}x\right)-\cos\left(\frac{3\pi}{a}x\right)\right]\\
&= \frac{4}{a}\cos\left(\frac{E_2-E_1}{\hbar}t\right)\left[-\frac{2a^2}{\pi^2}+\frac{2a^2}{9\pi^2}\right] && \text{(integration by parts)}\\
&= -\frac{64a}{9\pi^2}\cos\left(\frac{E_2-E_1}{\hbar}t\right)\\
\langle x\rangle &= A^2[I_1+I_3+I_2]\\
&= \frac{1}{2}\left[a-\frac{64a}{9\pi^2}\cos\left(\frac{E_2-E_1}{\hbar}t\right)\right]\\
&= a\left(\frac{1}{2}-\frac{32}{9\pi^2}\cos\left(\frac{E_2-E_1}{\hbar}t\right)\right)\\
&= a\left(\frac{1}{2}-\frac{32}{9\pi^2}\cos\left(\frac{3\hbar^2\pi^2}{2ma^2}t\right)\right)
\end{aligned}
$$

Momentum

$$\begin{aligned}
\langle p \rangle &= \int_0^a dx \, \left[A \left[U_1 e^{-iE_1 t/\hbar} + U_2 e^{-iE_2 t/\hbar} \right] \right] \\
&\quad \times \left[\left(-i\hbar \frac{\partial}{\partial x} \right) \left(A \left[U_1 e^{iE_1 t/\hbar} + U_2 e^{iE_2 t/\hbar} \right] \right) \right] \\
&= -i\hbar A^2 \int_0^a dx \, \left[\sqrt{\frac{2}{a}} \sin\left(\frac{\pi}{a}x\right) e^{-iE_1 t/\hbar} + \sqrt{\frac{2}{a}} \sin\left(\frac{2\pi}{a}x\right) e^{-iE_2 t/\hbar} \right] \\
&\quad \times \left[\frac{\pi}{a}\sqrt{\frac{2}{a}} \cos\left(\frac{\pi}{a}x\right) e^{iE_1 t/\hbar} + \frac{2\pi}{a}\sqrt{\frac{2}{a}} \cos\left(\frac{2\pi}{a}x\right) e^{iE_2 t/\hbar} \right] \\
&= -\frac{i\hbar}{a} \left[\int_0^a dx \, \frac{\pi}{a} \sin\left(\frac{\pi}{a}x\right) \cos\left(\frac{\pi}{a}x\right) \right. \\
&\qquad + \int_0^a dx \, \frac{2\pi}{a} \sin\left(\frac{\pi}{a}x\right) \cos\left(\frac{2\pi}{a}x\right) e^{i(E_1 - E_2)t/\hbar} \\
&\qquad + \int_0^a dx \, \frac{\pi}{a} \sin\left(\frac{2\pi}{a}x\right) \cos\left(\frac{\pi}{a}x\right) e^{i(E_2 - E_1)t/\hbar} \\
&\qquad \left. + \int_0^a \frac{2\pi}{a} \sin\left(\frac{2\pi}{a}x\right) \cos\left(\frac{2\pi}{a}x\right) \right] \\
&= -i\hbar A^2 [I_1 + I_2 + I_3 + I_4] \\
I_1 &= \int_0^a dx \, \frac{\pi}{a} \sin\left(\frac{\pi}{a}x\right) \cos\left(\frac{\pi}{a}x\right) \\
&= \int_0^\pi d\left(\sin\left(\frac{\pi}{a}\right)\right) \sin\left(\frac{\pi}{a}x\right) \\
&= \frac{1}{2} \left[\sin\left(\frac{\pi}{a}x\right) \right|_0^\pi \\
&= 0 \\
I_4 &= 0 \text{ (by analogous calculation)} \\
I_2 &= \frac{2\pi}{a} \int_0^a dx \, \sin\left(\frac{\pi}{a}x\right) \cos\left(\frac{2\pi}{a}x\right) e^{i(E_1 - E_2)t/\hbar} \\
&= \frac{\pi}{a} e^{i(E_1 - E_2)t/\hbar} \int_0^a dx \, \sin\left(\frac{3\pi}{a}x\right) - \sin\left(\frac{\pi}{a}x\right) \\
&= \frac{\pi}{a} e^{i(E_1 - E_2)t/\hbar} \left[-\frac{a}{3\pi} \cos\left(\frac{3\pi}{a}x\right) + \frac{a}{\pi} \cos\left(\frac{\pi}{a}x\right) \right|_0^a \\
&= -\frac{2}{3} e^{i(E_1 - E_2)t/\hbar} \\
I_3 &= \frac{2}{3} e^{i(E_2 - E_1)t/\hbar} \text{ (by similar calculation)}
\end{aligned}$$

$$
\begin{aligned}
\langle p \rangle &= -i\hbar A^2 [I_1 + I_2 + I_3 + I_4] \\
&= -i\hbar A^2 \left[0 - \frac{2}{3} e^{i(E_1 - E_2)t/\hbar} + \frac{2}{3} e^{i(E_2 - E_1)t/\hbar} + 0 \right] \\
&= \hbar A^2 \left[\frac{4}{3} \left(\frac{e^{i(E_2 - E_1)t/\hbar} - e^{-i(E_2 - E_1)t/\hbar}}{2i} \right) \right] \\
&= \frac{4\hbar A^2}{3} \sin\left(\frac{E_2 - E_1}{\hbar} t \right) \\
&= \frac{4\hbar A^2}{3} \sin\left(\frac{3\hbar^2 \pi^2}{2ma^2} t \right)
\end{aligned}
$$

Expectation of H

Find the expectation value of H and compare it to E_1 and E_2.

$$
\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}
$$

$$
\begin{aligned}
\langle H \rangle &= \int_0^a dx \left[\frac{1}{\sqrt{2}} \left[U_1 + U_2 \right] e^{-iEt/\hbar} \right] \left[\left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right) \left(\frac{1}{\sqrt{2}} [U_1 + U_2] e^{iEt/\hbar} \right) \right] \\
&= \frac{\hbar^2}{4m} \int_0^a dx \left[U_1 + U_2 \right] \left[\left(\frac{\pi}{a} \right)^2 U_1 + \left(\frac{2\pi}{a} \right)^2 U_2 \right] \\
&= \frac{\hbar^2}{4m} \int_0^a dx \left[\left(\frac{\pi}{a} \right)^2 U_1^2 + \left(\frac{\sqrt{5}\pi}{a} \right)^2 U_1 U_2 + \left(\frac{2\pi}{a} \right)^2 U_2^2 \right] \\
&= \frac{\hbar^2}{4m} \left[\left(\frac{\pi}{a} \right)^2 \int_0^a dx \, U_1^2 + \left(\frac{\sqrt{5}\pi}{a} \right)^2 \int_0^a dx \, U_1 U_2 + \left(\frac{2\pi}{a} \right)^2 \int_0^a dx \, U_2^2 \right] \\
&= \frac{\hbar^2}{4m} \left[\left(\frac{\pi}{a} \right)^2 (1) + \left(\frac{\sqrt{5}\pi}{a} \right)^2 (0) + \left(\frac{2\pi}{a} \right)^2 (1) \right] \\
&= \frac{5\pi^2 \hbar^2}{4ma^2}
\end{aligned}
$$

We know that the energies of a given particular solution are:

$$
E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}
$$

so

$$
\begin{aligned}
\langle H \rangle &= \frac{5\pi^2\hbar^2}{4ma^2} \\
&= \frac{1}{2}\left[\frac{1^2\pi^2\hbar^2}{2ma^2} + \frac{2^2\pi^2\hbar^2}{2ma^2}\right] \\
&= A^2\left[E_1 + E_2\right] \\
&= A^2\left[\langle H \rangle_1 + \langle H \rangle_2\right]
\end{aligned}
$$

as expected.

PH214C – Pset 8

Jonathan Lam

May 1, 2020

1 P-current

Using the definition for probability density $P = |\psi^2| = \psi^*\psi$, calculate $\partial P/\partial t$ to find that

$$\frac{\partial P}{\partial t} + \frac{\partial \mathbb{J}}{\partial x} = 0$$

in this way defining the probability current $\mathbb{J}$.

We know that $\mathbb{J}$ is defined as:

$$\mathbb{J} \coloneqq \frac{i\hbar}{2m}\left(\psi\frac{\partial\psi^*}{\partial x} - \psi^*\frac{\partial\psi}{\partial x}\right)$$

So:

$$\begin{aligned}
\frac{\partial}{\partial t}P &= \frac{\partial}{\partial t}|\psi|^2 = \frac{\partial}{\partial t}\psi^*\psi = \psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t} \\
&= \psi^*\left[\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi\right)\right] + \psi\left[-\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi\right)\right] \\
&= \psi^*\left[-\frac{\hbar}{2mi}\frac{\partial^2\psi}{\partial x^2}\right] + \frac{V}{i\hbar}\psi^*\psi + \psi\left[\frac{\hbar}{2mi}\frac{\partial^2\psi}{\partial x^2}\right] - \frac{V}{i\hbar}\psi^*\psi \\
&= \frac{i\hbar}{2m}\left[\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right] \\
&= \frac{i\hbar}{2m}\left[\left(\frac{\partial\psi}{\partial x}\frac{\partial\psi^*}{\partial x} + \psi^*\frac{\partial}{\partial x}\left(\frac{\partial\psi}{\partial x}\right)\right) - \left(\frac{\partial\psi^*}{\partial x}\frac{\partial\psi}{\partial x} + \psi\frac{\partial}{\partial x}\left(\frac{\partial\psi^*}{\partial x}\right)\right)\right] \\
&= \frac{i\hbar}{2m}\left[\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x}\right) - \frac{\partial}{\partial x}\left(\psi\frac{\partial\psi^*}{\partial x}\right)\right] \\
&= \frac{\partial}{\partial x}\left[\frac{i\hbar}{2m}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right)\right] \\
&= -\frac{\partial}{\partial x}\mathbb{J} \\
&\Rightarrow \frac{\partial P}{\partial t} + \frac{\partial \mathbb{J}}{\partial x} = 0
\end{aligned}$$

2 0-current

For the potential step problem, in the case that $E < V_0$, there is a finite probability that a particle is found in the forbidden region, $x > 0$, where $E < V_0$. Yet the probability current in the transmitted region, $\mathbb{J}_{trans} = 0$. Show these results.

The time-independent solution U to the finite step function was governed by the following equations. Note that U_L denotes the time-independent solution to the Schödinger equation for $x < 0$ ($V(x) = 0$), and U_R denotes the like for $x > 0$ ($V(x) = V_0$).

$$\begin{aligned}
\frac{d^2U_L}{dx^2} &= -\frac{2mE}{\hbar^2}U_R & \qquad \frac{d^2U_R}{dx^2} &= -\frac{2m(E-V_0)}{\hbar^2}U_R \\
U_L(x) &= e^{ikx} + Re^{-ikx} & U_R(x) &= Te^{i\kappa x} \\
k^2 &= \frac{2mE}{\hbar^2} & \kappa^2 &= \frac{2m(E-V_0)}{\hbar^2} \\
R &= \frac{k-\kappa}{k+\kappa} & T &= \frac{2k}{k+\kappa}
\end{aligned}$$

We are interested in the solution in the positive region, where $E < V_0$. This would suggest that κ becomes (pure) imaginary, so that:

$$\kappa = i\frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

$$U_R = T\exp\left(-\frac{2m(V_0 - E)}{\hbar}x\right)$$

This is a decaying exponential with coefficient $T \neq 0$, so $U_R(x) \neq 0$ in the forbidden region. Since $S(t)$ is also an exponential (and thus not identically zero), then $P = |\psi^2| = |U(x)S(t)| \neq 0$, there is a nonzero probability that a particle is found in the forbidden region. However, if we were to calculate the probability current, it becomes zero now that U_R becomes pure real:

$$\begin{aligned}
\mathbb{J}_R &= \frac{i\hbar}{2m}\left(\psi\frac{\partial\psi^*}{\partial x} - \psi^*\frac{\partial\psi}{\partial x}\right) \\
&= \frac{i\hbar}{2m}\left([U_R(x)S(t)]\frac{\partial}{\partial x}[U_R(x)S^*(t)] - [U_R(x)S^*(t)]\frac{\partial}{\partial x}[U_R(x)S(t)]\right) \\
&= \frac{i\hbar}{2m}\left(U_R S S^*\frac{\partial U_R}{\partial x} - U_R S^* S\frac{\partial U_R}{\partial x}\right) \\
&= \frac{i\hbar}{2m}(0) \\
&= 0
\end{aligned}$$

In other words, the probability distribution of the particles in the forbidden region is not evolving over time.

How low can you go?

Use the property of the lowering operator and ground state of the QMSHO:

$$\hat{a}^- \psi_0 = 0$$

to find the ground state wave function ψ_0. Now raise ψ_0 to get ψ_1 and then lower ψ_1 back to ψ_0. Do you get back to your starting point or is there a "leftover"?

Definitions:

$$\hat{a}^- := \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right)$$
$$\hat{a}^+ := \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i\hat{p}}{m\omega} \right)$$

$$\hat{p} := -i\hbar \frac{\partial}{\partial x}$$

Solving the ODE:

$$\begin{aligned}
\sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i}{m\omega} \left(-i\hbar \frac{\partial}{\partial x} \right) \right) \psi_0 &= 0 \\
\frac{\hbar}{m\omega} \frac{\partial \psi_0}{\partial x} + x\psi_0 &= 0 \\
\frac{d\psi_0}{\psi_0} &= -\frac{m\omega x}{\hbar}\, dx \\
\psi_0 &= A \exp\left(-\frac{m\omega x^2}{2\hbar} \right)
\end{aligned}$$

Normalizing:

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} dx\, \psi^* \psi \\
&= \int_{-\infty}^{\infty} dx\, A^* \exp\left(-\frac{m\omega x^2}{2\hbar} \right) A \exp\left(-\frac{m\omega x^2}{2\hbar} \right) \\
&= |A|^2 \int_{-\infty}^{\infty} dx\, \exp\left(-\frac{m\omega x^2}{\hbar} \right) \\
|A| &= \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \\
\psi_0 &= \sqrt[4]{\frac{m\omega}{\pi\hbar}} \exp\left(-\frac{m\omega x^2}{2\hbar} \right)
\end{aligned}$$

To save some typing, let $\alpha = m\omega\hbar^{-1}$. "Raising" ψ_0 to ψ_1' (non-normalized):

$$\begin{aligned}
\psi_1' &= \hat{a}^+\psi_0 \\
&= \sqrt{\frac{\alpha}{2}}\left(\hat{x} + \frac{i}{m\omega}\left(-i\hbar\frac{\partial}{\partial x}\right)\right)\psi_0 \\
&= \sqrt{\frac{\alpha}{2}}\left(\hat{x} - \frac{1}{\alpha}\frac{\partial}{\partial x}\right)\left[\sqrt[4]{\frac{\alpha}{\pi}}\exp\left(-\frac{\alpha x^2}{2}\right)\right] \\
&= \sqrt{\frac{\alpha}{2}}\sqrt[4]{\frac{\alpha}{\pi}}\left(x\exp\left(-\frac{\alpha x^2}{2}\right) - \frac{1}{\alpha}\left(-\frac{2\alpha x}{2}\right)\exp\left(-\frac{\alpha x^2}{2}\right)\right) \\
&= \sqrt{\frac{\alpha}{2}}\sqrt[4]{\frac{\alpha}{\pi}}(2x)\exp\left(-\frac{\alpha x^2}{2}\right) \\
&= \sqrt{2\alpha}x\psi_0
\end{aligned}$$

"Lowering" ψ_1' to ψ_0':

$$\begin{aligned}
\psi_0' &= \hat{a}^-\psi_1' \\
&= \sqrt{\frac{\alpha}{2}}\left(\hat{x} - \frac{i}{m\omega}\left(-i\hbar\frac{\partial}{\partial x}\right)\right)\psi_1' \\
&= \sqrt{\frac{\alpha}{2}}\sqrt{2\alpha}\sqrt[4]{\frac{\alpha}{\pi}}\left(\hat{x} + \frac{1}{\alpha}\frac{\partial}{\partial x}\right)\left[x\exp\left(-\frac{\alpha x^2}{2}\right)\right] \\
&= \alpha\sqrt[4]{\frac{\alpha}{\pi}}\left(x^2\exp\left(-\frac{\alpha x^2}{2}\right) + \frac{1}{\alpha}\left(-\frac{2\alpha x}{2}x\exp\left(-\frac{\alpha x^2}{2}\right) + \exp\left(-\frac{\alpha x^2}{2}\right)\right)\right) \\
&= \alpha\sqrt[4]{\frac{\alpha}{\pi}}\left(\frac{1}{\alpha}\right)\exp\left(-\frac{\alpha x^2}{2}\right) \\
&= \psi_0
\end{aligned}$$

We see that $\psi_0' = \psi_0$, so there is no "leftover" scaling factor. Or, if we know the scaling factor caused by the raising and lowering operators (shown in the notes):

$$\begin{aligned}
\hat{a}^+\psi_n &= \sqrt{n+1}\psi_{n+1} \\
\hat{a}^-\psi_n &= \sqrt{n}\psi_{n-1}
\end{aligned}$$

then we get the same result:

$$\hat{a}^-\hat{a}^+\psi_0 = \hat{a}^-(\sqrt{0+1}\psi_{0+1}) = \hat{a}^-\psi_1 = \sqrt{1}\psi_{1-1} = \psi_0$$

Operator, operator

Find the position and momentum operators, $\hat{x}$ and $\hat{p}$, for the SHO potential in terms of the raising and lowering operators, $\hat{a}^+$ and $\hat{a}^-$.

We defined the raising and lowering operators partially in terms of $\hat{p}$ and $\hat{x}$, so we just need to solve in reverse. Solving for $\hat{p}$:

$$\hat{a}^{-} - \hat{a}^{+} = \sqrt{\frac{m\omega}{2\hbar}}\left(2\frac{i\hat{p}}{m\omega}\right) = \sqrt{\frac{2}{m\omega\hbar}}i\hat{p}$$

$$\hat{p} = -i\sqrt{\frac{m\omega\hbar}{2}}(\hat{a}^{-} - \hat{a}^{+})$$

Solving for $\hat{x}$:

$$\hat{a}^{-} + \hat{a}^{-} = \sqrt{\frac{m\omega}{2\hbar}}(2\hat{x}) = \sqrt{\frac{2m\omega}{\hbar}}\hat{x}$$

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^{-} + \hat{a}^{+})$$

Great expectations

Find $\langle x\rangle$, $\langle p\rangle$, $\langle x^2\rangle$, $\langle p^2\rangle$, and $\langle K\rangle$ (where K is the kinetic energy) for the n-th state of the SHO.

We use the linearity of the bra-ket (expectation value, or inner product) and the orthonormality of solutions ($\langle\psi_n|\psi_m\rangle = \delta_{nm}$) implicitly.

$$\begin{aligned}
\langle x\rangle_n &= \langle\psi_n|\hat{x}\psi_n\rangle \\
&= \left\langle\psi_n\middle|\sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^{-} + \hat{a}^{-})\psi_n\right\rangle \\
&= \sqrt{\frac{\hbar}{2m\omega}}\left[\langle\psi_n|\hat{a}^{-}\psi_n\rangle + \langle\psi_n|\hat{a}^{+}\psi_n\rangle\right] \\
&= \sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{n}\,\langle\psi_n|\psi_{n-1}\rangle + \sqrt{n+1}\,\langle\psi_n|\psi_{n+1}\rangle\right] \\
&= \sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{n}(0) + \sqrt{n+1}(0)\right] \\
&= 0
\end{aligned}$$

$$
\begin{aligned}
\langle p \rangle_n &= \langle \psi_n | \hat{p} \psi_n \rangle \\
&= \left\langle \psi_n \middle| -i\sqrt{\frac{m\omega\hbar}{2}} (\hat{a}^- - \hat{a}^+) \psi_n \right\rangle \\
&= -i\sqrt{\frac{m\omega\hbar}{2}} \left[\langle \psi_n | \hat{a}^- \psi_n \rangle - \langle \psi_n | \hat{a}^+ \psi_n \rangle \right] \\
&= i\sqrt{\frac{m\omega\hbar}{2}} \left[\sqrt{n} \langle \psi_n | \psi_{n-1} \rangle - \sqrt{n+1} \langle \psi_n | \psi_{n+1} \rangle \right] \\
&= i\sqrt{\frac{m\omega\hbar}{2}} \left[\sqrt{n}(0) - \sqrt{n+1}(0) \right] \\
&= 0
\end{aligned}
$$

Intuitively, the expectations for position and momentum can be seen by the symmetry of the SHO potential and an intuitive understanding of displacement and momentum in the classical sense. The variances are less obvious.

$$
\begin{aligned}
\langle x^2 \rangle_n &= \langle \psi_n | \hat{x}^2 \psi_n \rangle \\
&= \left\langle \psi_n \middle| \frac{\hbar}{2m\omega} (\hat{a}^- + \hat{a}^+)^2 \psi_n \right\rangle \\
&= \frac{\hbar}{2m\omega} \left[\left\langle \psi_n \,\middle|\, (\hat{a}^-)^2 \psi_n \right\rangle + \left\langle \psi_n \,\middle|\, \hat{a}^- \hat{a}^+ \psi_n \right\rangle + \left\langle \psi_n \,\middle|\, \hat{a}^+ \hat{a}^- \psi_n \right\rangle + \left\langle \psi_n \,\middle|\, (\hat{a}^+)^2 \psi_n \right\rangle \right] \\
&= \frac{\hbar}{2m\omega} \left[\sqrt{n(n-1)} \langle \psi_n | \psi_{n-2} \rangle + (n+1) \langle \psi_n | \psi_n \rangle + n \langle \psi_n | \psi_n \rangle + \sqrt{(n+1)(n+2)} \langle \psi_n | \psi_{n+2} \rangle \right] \\
&= \frac{\hbar}{2m\omega} \left[\sqrt{n(n-1)}(0) + (n+1)(1) + n(1) + \sqrt{(n+1)(n+2)}(0) \right] \\
&= \frac{\hbar}{2m\omega} (2n+1)
\end{aligned}
$$

$$
\begin{aligned}
\langle p^2 \rangle_n &= \langle \psi_n | \hat{p}^2 \psi_n \rangle \\
&= \left\langle \psi_n \middle| -\frac{m\omega\hbar}{2} (\hat{a}^- - \hat{a}^+) \psi_n \right\rangle \\
&= -\frac{m\omega\hbar}{2} \left[\left\langle \psi_n \,\middle|\, (\hat{a}^-)^2 \psi_n \right\rangle - \left\langle \psi_n \,\middle|\, \hat{a}^- \hat{a}^+ \psi_n \right\rangle - \left\langle \psi_n \,\middle|\, \hat{a}^+ \hat{a}^- \psi_n \right\rangle + \left\langle \psi_n \,\middle|\, (\hat{a}^+)^2 \psi_n \right\rangle \right] \\
&= -\frac{m\omega\hbar}{2} \left[\sqrt{n(n-1)}(0) - (n+1)(1) - n(1) + \sqrt{(n+1)(n+2)}(0) \right] \\
&= \frac{m\omega\hbar}{2} (2n+1)
\end{aligned}
$$

$$
\langle K \rangle_n = \left\langle \frac{p^2}{2m} \right\rangle = \frac{1}{2m} \langle p^2 \rangle = \frac{\omega\hbar}{4} (2n+1)
$$

Particle in a 3D box – Jonathan Lam PH214C Final Pset

Potential function:

$$V(\vec{r}) = V(x,y,z) = \begin{cases} 0 & 0<x<L_x, 0<y<L_y, 0<z<L_z \\ \infty & \text{otherwise} \end{cases}$$

The time-independent Schrödinger equation is given by:

$$\lim_{V\to\infty}\left[-\frac{\hbar^2}{2m}\nabla^2 U + V(\vec{r})U = EU\right]$$

Dividing each term by V leads to the boundary conditions $U(\vec{r}\text{ on boundary}) = 0$. Since $V=0$ in the box, we have:

$$-\frac{\hbar^2}{2m}\nabla^2 U + 0 = EU \Rightarrow \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)U = EU$$

To solve this, assume U is separable, i.e., $U(\vec{r}) = U_1(x)U_2(y)U_3(z)$. Thus:

$$-\frac{\hbar^2}{2m}\left(U_2U_3\frac{d^2U_1}{dx^2}+U_1U_3\frac{d^2U_2}{dy^2}+U_1U_2\frac{d^2U_3}{dz^2}\right) = EU \Rightarrow \frac{\hbar^2}{2m}\left(\frac{1}{U_1}\frac{d^2U_1}{dx^2}+\frac{1}{U_2}\frac{d^2U_2}{dy^2}+\frac{1}{U_3}\frac{d^2U_3}{dz^3}\right) = -E \tag{1}$$

This is always be true in the box, in particular when $d^2U_2/dy^2 = d^2U_3/dz^2 = 0$ (on a line parallel to the x-axis). In this case, let the energy eigenvalue be E_1:

$$\frac{\hbar^2}{2m}\left(\frac{1}{U_1}\frac{d^2U_1}{dx^2}+0+0\right) = \frac{\hbar^2}{2m}\frac{d^2U_1}{dx^2} = -E_1U_1 \qquad (U_1(0)=U_1(L_x)=0)$$

This is exactly the one-dimensional particle-in-a-box setup and differential equation (including boundary conditions), so the solution to U_1 should be the same as the 1D case.

$$U_1(x) = A_x\sin k_1x,\ k_1 = \frac{\sqrt{2mE_1}}{\hbar} = \frac{\pi n_1}{L_x},\ E_1 = \frac{\pi^2\hbar^2n_1^2}{2mL_x^2} \qquad (n_1\in\{1,2,3,\dots\})$$

The analogous result is true for U_2, U_3. Using (1), we obtain the discrete set of total energy eigenvalues for U (which unsurprisingly is the sum of the component energies, each of which is a discrete set):

$$\frac{1}{U_1}\frac{d^2U_1}{dx^2} = \frac{1}{A_1\sin k_1x}(-A_1k_1^2\sin k_1x) = -k_1^2 = -\frac{2mE_1}{\hbar^2} \qquad (\text{analogous results for } U_2,\ U_3)$$

$$E = -\frac{\hbar^2}{2m}\left(\frac{1}{U_1}\frac{d^2U_1}{dx^2}+\frac{1}{U_2}\frac{d^2U_2}{dy^2}+\frac{d^2U_3}{dz^2}\right) = -\frac{\hbar^2}{2m}\left(-\frac{2mE_1}{\hbar^2}-\frac{2mE_2}{\hbar^2}-\frac{2mE_3}{\hbar^2}\right) = E_1+E_2+E_3$$

By our assumption of separability, we have an explicit set of solutions for U (purely real) and ψ and can normalize:

$$U = U_1U_2U_3 = A\sin(k_1x)\sin(k_2y)\sin(k_3z) \Rightarrow \psi(\vec{r},t) = U(\vec{r})\exp(-iEt/\hbar)$$

$$1 = \int_V dV\,\psi^*\psi = \int_0^{L_x}dx\int_0^{L_y}dy\int_0^{L_z}dz\ [U\exp(iEt/\hbar)]\,[U\exp(-iEt/\hbar)] = \int_0^{L_x}dx\int_0^{L_y}dy\int_0^{L_z}dz\,U^2$$

$$= A^2\int_0^{L_x}dx\sin^2k_1x\int_0^{L_y}dy\sin^2k_2y\int_0^{L_z}dz\sin^2k_3z$$

$$\left[\text{example integral: }\int_0^{L_x}dx\sin^2k_1x = \int_0^{L_x}dx\left(\frac{1}{2}-\frac{\cos 2k_1x}{4k_1}\right) = \frac{L_x}{2}-\frac{1}{4k_1}\sin(2k_1L_x(=2\pi))\Big|_0^{L_x} = \frac{L_x}{2}\right]$$

$$= \frac{1}{8}A^2L_xL_yL_z \Rightarrow A = \sqrt{\frac{8}{L_xL_yL_z}}$$

In summary:

$$\psi(\vec{r},t) = \sqrt{\frac{8}{L_xL_yL_z}}\sin\left(\frac{\pi n_1x}{L_x}\right)\sin\left(\frac{\pi n_2y}{L_y}\right)\sin\left(\frac{\pi n_3z}{L_z}\right)\exp(-iEt/\hbar)$$

$$E = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_1^2}{L_x^2}+\frac{n_2^2}{L_y^2}+\frac{n_3^2}{L_z^2}\right)\ (n_1,n_2,n_3\in\{1,2,3,\dots\})$$

In the case where $L_x = L_y = L_z = L$ (a cube):

$$\psi(\vec{r},t) = \sqrt{\frac{8}{L^3}}\sin\left(\frac{\pi n_1x}{L}\right)\sin\left(\frac{\pi n_2x}{L}\right)\sin\left(\frac{\pi n_3x}{L}\right)\exp\left(-\frac{iEt}{\hbar}\right),\ E = \frac{\pi^2\hbar^2}{2mL^2}(n_1^2+n_2^2+n_3^2)$$

$$(n_1,n_2,n_3\in\{1,2,3,\dots\})$$