Vandy's Notes on Everything

vandy

July 29, 2024

Contents

Chapter 1

Introduction test

This document will contain ALLLLL the notes I take on everything I learn. I will try to keep it as organized as possible, but I can't promise anything.

1.1 Latex settings

Latex is a pain to use sometimes. Currently, with the setup i have, i have to do some magic rituals to get the subfiles to compile correctly. I have to have a `.latexmkrc` file in the root directory of the project, se next section, and the following sections settings in the `settings.json` file in vscode.

The main.tex file typically compiles fine, but the subfiles don't compile correctly after compiling the main file unless you do the following. (1) compile main file (ctrl-s will build automatically), (2) go to subfile and change the path from `\document class[../../main.tex]` to `\document class[../main.tex]{subfiles}`, the sub file will build and send output to a weird spot, and i think through an error, that's fine. (3) change the path back to `\documentclass[../../main.tex]` and the subfile will build correctly, with its pdf in the same folder as the subfile.tex.

This should be fixable, but i haven't figured it out yet.

1.1.1 latexmkrc

the following is the content of the `.latexmkrc` file (as of 6/26/2024):

```
    # .latexmkrc
use File::Basename;

# Custom subroutine to compile subfiles
add_cus_dep('tex', 'pdf', 0, 'subfile_pdf');
sub subfile_pdf {
    my ($base_name, $path) = fileparse(shift);
    my $dir = $path ? "$path/" : "";

    system("latexmk -pdf -aux-directory=out $dir$base_name.tex");
    return 0;
}

# Hook into LaTeXmk's clean routine to remove auxiliary files
add_cus_dep('tex', 'clean', 0, 'clean_aux_files');
sub clean_aux_files {
    my ($base_name, $path) = fileparse(shift);
    my $dir = $path ? "$path/" : "";
```

```
    system("rm -f out/$base_name.*");
    return 0;
}

# Ensure LaTeXmk knows to clean auxiliary files on 'latexmk -C'
push @clean_ext, qw(aux log out toc fdb_latexmk fls nav snm vrb xdv);

# Output directory for auxiliary files
$aux_dir = 'out';
```

that script placed in the root directory of the project will allow latexmk to compile subfiles and clean auxiliary files. the pdfs will be in the root directory, while the auxiliary files will be in the `out` directory. cleaning the project will remove all auxiliary files named "main" but not the pdfs. clean doesn't clean the subfiles auxiliary files currently. the preview pdf button in vscode will not work with the aux files with this setup, but the pdfs will update if opened manually.

1.1.2 settings.json

as of 6/26/2024, the following is the content of the `settings.json` file:

```
{
    "workbench.colorTheme": "Visual Studio Dark",
    "editor.showFoldingControls": "always",
    "security.workspace.trust.untrustedFiles": "open",
    "editor.inlineSuggest.enabled": true,
    "julia.symbolCacheDownload": true,
    "wolfram.command": [
        "`kernel`",
        "-noinit",
        "-noprompt",
        "-nopaclet",
        "-noicon",
        "-nostartuppaclets",
        "-run",
        "Needs[\"LSPServer`\"];LSPServer`StartServer[]"
    ],
    "wolfram.kernel": "C:\\Program Files\\Wolfram Research\\
       Mathematica\\13.3\\WolframKernel.exe",
    "git.openRepositoryInParentFolders": "always",
    "editor.minimap.enabled": false,
    "julia.enableTelemetry": false,
    "github.copilot.advanced": {},
    "jupyter.askForKernelRestart": false,
    "editor.tabCompletion": "on",
    "editor.formatOnPaste": true,
    "python.analysis.autoSearchPaths": true,
    "python.autoComplete.extraPaths": [
        "Z:\\home\\Vandy\\code\\vandyCode\\MyPackages"
    ],
    "python.analysis.extraPaths": [
        "Z:\\home\\Vandy\\code\\vandyCode\\MyPackages"
    ],
    "explorer.confirmDelete": false,
    "explorer.confirmDragAndDrop": false,
    "julia.lint.run": true,
```

```
    "security.allowedUNCHosts": [
        "srsrv1"
    ],
    "jupyter.debugJustMyCode": false,
    "python.defaultInterpreterPath": "C:\\Users\\micha\\anaconda3\\
       python.exe",
    "jupyter.stopOnFirstLineWhileDebugging": false,
    "jupyter.showVariableViewWhenDebugging": true,

    "latex-workshop.latex.recipes": [
        {
            "name": "latexmk",
            "tools": [
                "latexmk"
            ]
        }
    ],
    "latex-workshop.latex.tools": [
        {
            "name": "latexmk",
            "command": "latexmk",
            "args": [
                "-synctex=1",
                "-interaction=nonstopmode",
                "-file-line-error",
                "-pdf",
                "%DOC%"
            ]
        }
    ],
    "latex-workshop.latex.autoBuild.run": "onSave",
    "latex-workshop.viewer.pdf.tab.editorGroup": "both"
}
```

and things are working well.

1.2 Mathematica settings

To change mathematica's annoyingly large margins , you can fight with the inheritance system and options inspector, with the option CellMargins. Or you can edit the package stylesheet. you aren't allowed to edit the default stylesheet, so you have to make a new one and put it in the appropriate directory.

```
CopyFile @@ (FileNameJoin[{#, "SystemFiles", "FrontEnd",
  "StyleSheets",
  "Package.nb"}] & /@ {$InstallationDirectory, $UserBaseDirectory})
```

will copy the default package stylesheet to the user directory. you can then edit the stylesheet to change the margins. The sheet looks like this:

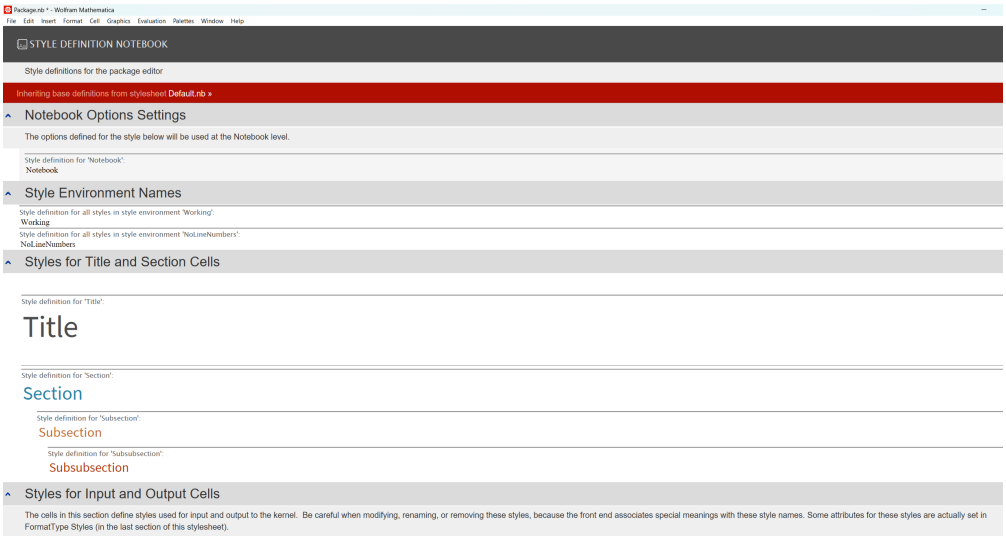


Figure 1.1: The stylesheet

To edit the style of a type of cell, select the cell, such as the "Title" cell, and Show the Cell Expression via ctrl+shift+E. You can then edit the cell's style. For example, to change the margins of the title cell, you can change the margins in the cell expression by adding the option CellMargins→{{left,right},{bottom,top}}.

Another fun fact, if

1.3 Git

you can use the following command to remove the specific safe directory directly from the terminal:

Chapter 2

Formulas

This document is for everything. all the formulas. forever. The path to this file from (from sr gradiometer) is `Z:\home\Vandy\Latex\Postdoc all formulas`.

2.1 Lande g-factors

the lande g-facotrs or g factors are a proportionality term for the energy of an atom in a magnetic field

$$g_J = g_L \frac{J(J+1) - S(S+1) + L(L+1)}{2J(J+1)} + g_S \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

Where g_L is the electrons orbital g-factor and g_S is its spin g factor. g_L is approximately 1, but Daniel Steck's notes also account for a finite nuclear mass to given $g_L = m_e/m_{nuc}$, but most sources just have it be one, and one source (https://iopscience.iop.org/article/10.1088/1402-4896/ab8f44) wondering if we ignore it. g_S 2 is also an approximation, $g_S = 2.00231930436256$ but it has been measured like nuts.

therefore, we often approximate

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

2.2 Doppler Broadening (from wikipedia)

For non relativistic thermal velocities the doppler shift is

$$\omega_v = \omega_0(1 + v/c)$$

$$\Delta\omega = kv$$

and for a system with a velocity distribution $P_v(v)$, meaning $P_v(v)dv$ is the probability (or fraction) of the system with velcoties v to $v + dv$. Then the corresponding distribution of frequencies would be

$$P_\omega(\omega)d\omega = P_v(v_\omega)\frac{dv}{d\omega}d\omega$$

where $v_\omega = c(\omega/\omega 0 - 1)$ is the corresponding velocity for the freuqnecy ω. We therefore have

$$P_\omega(\omega)d\omega = \frac{c}{\omega}P_v\left(c\left(\frac{\omega}{\omega_0} - 1\right)\right)d\omega$$

. We can also expressed in terms of wavelength, where in the nonrelativistic limit we may used that $\frac{\lambda-\lambda_0}{\lambda_0} \approx -\frac{\omega-\omega_0}{\omega_0}$ so see that

$$P_\lambda(\lambda)d\lambda = \frac{c}{\lambda_0}P_v\left(c\left(1 - \frac{\lambda}{\lambda_0}\right)\right)d\lambda$$

For a thermal, Maxwell-Boltzman distribution, we have the Maxwell distribution

$$P_v(V)dv = \sqrt{\frac{m}{2\pi k_b T}} \exp\left(-\frac{mv^2}{2k_b T}\right) dv, \tag{2.1}$$

and so we substitute as appropriate and similpy to get

$$P_\omega(V)d\omega = \sqrt{\frac{mc^2}{2\pi k_b T\omega_0^2}} \exp\left(-\frac{mc^2(\omega-\omega_0)^2}{2k_b T\omega_0^2}\right) d\omega$$

which is a Gaussian profile with

$$\sigma_\omega = \sqrt{\frac{k_b T}{mc^2}}\,\omega_0$$

, and since $FWHM = 2\sqrt{2\ln 2}\sigma$ we have

$$\Delta\omega_{FWHM} = \sqrt{\frac{8k_b T\ln 2}{mc^2}}\,\omega_0 \tag{2.2}$$

2.3 Optical cross-section

the resonant cross section for a transition is

$$\sigma_0 = \frac{3\lambda^2}{2\pi} \tag{2.3}$$

I have had a terrible time trying to remember the numerical prefactor, but i think i've figure out an intuitive way to remember! Consider that while the wavelength in nano meters and the frequency in MHz are what we use to describe it physically, for the math we use angular frequency $\omega = 2\pi f$, we basically don't use f alone in the math. Why should we behave differently with wavelength? I hate using $h = 2\pi\hbar$ and $k = \lambda^{-1} = 2\pi/\lambda$ is already our typical thing, so lets swap λ for λ. then, very intuitively, the cross-section is $\sigma 3 * 2 * \pi\lambda^2$ where the three is for three directions, and the two is for two polarizations. So thus,

$$\sigma_0 = 6\pi\lambda^2 = \frac{3\lambda^2}{2\pi} \tag{2.4}$$

which is going to be much easier to remember!

2.4 Gaussian beams

peak intensity of a gaussian beam is (remember the factor of two!)

$$P_0 = \frac{2P}{\pi\omega_x\omega_Y} \tag{2.5}$$

Chapter 3

MOTs

https://journals.aps.org/pra/pdf/10.1103/PhysRevA.70.063413 narrow line cooling and momentum space crystals

https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.93.073003 previous paper narrow line cooling finite photon recoil dynamics For the 689 line in strontium, the ratio of the recoil frequency $\omega_r = \frac{\hbar k^2}{2m} = 2\pi \times 4.8$ kHz to the natural linewidth $\Gamma = 2\pi \times 7.4$ kHz the transition is on the order of unity. this is what defines a narrow line MOT.

3.1 Finite photon recoil dynamics

The 1s0-3p1 transition has its single photon recoil is of comparable sixe to the natural linewidth. When an atom of mass m emits or absorbs a photon γ of frequency ω, then momentum of the photon $p_\gamma = h/\lambda = \hbar k$ is transfered to the atom giving it some velocity $v_{rec} = \hbar k/m$

3.2 Regimes of red-detuned MOT

Three regimes of behavior depend on the relative magnitudes of $|\Delta|$, Γ, and $\Gamma_E = \Gamma\sqrt{1+s}$ where Γ is the natural linewidth $s = I/I_{sat}$ is the saturation intensity which may broaden the line. The three regimes are (I) $\Gamma \ll \Gamma_E \ll |\Delta|$, (II), $|\Delta| < \Gamma_E$ and $\Gamma \ll \Gamma_E$ so $\Gamma < |\Delta| < \Gamma_e$ (but perhaps $\Gamma \not\ll |\Delta|$?), and (III), when $s \simeq 1$. and so $\Gamma \not\ll \Gamma_E$.

3.2.1 Red-detuned MOT Regime (I)

In the first regime, $\Gamma \ll \Gamma_E \ll |\Delta|$. For the 689 line in strontium, the ratio of the recoil frequency $\omega_r = \frac{\hbar k^2}{2m} = 2\pi \times 4.8$ kHz to the natural linewidth $\Gamma = 2\pi \times 7.4$ kHz the transition is on the order of unity. In this regime,

3.2.2 regime 2

$|\Delta| \lesssim \Gamma_E$ and $\Gamma \ll \Gamma_E$

this is more typical and thermodynamics remicent of ordinary doppler cooling. δ and s dependent teperature minima

3.2.3 regime 3

s approces unit, photon-tecoil-driven impulsive force dominates and temperature below photon-recoil limit $T_r = 2\hbar\omega_r/k_b$

$$\Gamma \simeq \Gamma_E$$

3.2.4 other regimes?

sing $\Gamma \sim \omega_r$ you get momentum packets. explained by analytic solution to the 1D semiclassical radiative force equation. for $\delta > 0$, and $\Gamma \sim \omega_r$ allow direct visualization of positive feedback acceleration that effeciently bunches the atoms into discrete, well-defined momentum packets. the gravity ratio R directly influences the blue detuned thermodynamics, enabling cooling around a velocyt where radiation pressure and gravity balance. aparently the underlying physics here is the same as regime I but manifest bvery differently.

3.2.5 method

uses frequency modulate redMOT, start at 3G/cm and go to 10 G/cm. Then transition to singlefrequency

Chapter 4

Cooling and Trapping atoms with t

I want to write a high level overview in my own words for how we use light to slow down and trap atoms. We will begin with discussing how an atom moving in one dimension is influenced by a single laser beam. This is the mechanism used to slow decellerate an atomic beam in a Zeeman slower. Then we will expand our discussion to the phenomenon of optical molasses, wherein the atom experiences a damping force due to the presence of more than one laser. However, it should be remembered that an optical molasses will not trap the atoms, they will continue to move in a random walk as photons are scattered. Then we will introduce trapping and confinement and put these ideas together to create the Magneto-Optical Trap (MOT).

I don't know where i want to discuss the idea of dissipation and increase of phase space density. The last section of chapter Five of the yellow book points out than if the system in question is Hamiltonian, i.e. it can be decrisbed by a Hamiltonian and Hamilton's equations, then Liouville's theorem demonstrates that the phase space density of the system cannot decrease. However dissipative systems, such as those including friction, cannot be modeled via a Hamiltonian. Systems with explicitly velocity-dependent forces are not expressable via a Hamiltonian. However, this can be misleading, since the velocity-dependent Lorentz force does not prevent a system of charged particles from being described with a Hamiltonian. More generally, it is conservative systemswhich may be described by a Hamiltonian, systems containing friction are not conservation, energy is lost as heat.

It may be possible to extend Hamiltonian techniques to incorporate such dissipation, but we will not consider those to be Hamiltonian systems.

4.1 Deceleration of an atom by a single beam

At the heart of laser cooling is the exchange of momentum between photons and atoms. Consider an atom with mass m at rest which emits a photon with frequency ω. After the emission of the photon, the total momentum of atom and photon together must remain zero. We know from special relativity that photons have momentum $p_\gamma = \hbar k$ (this follows from the fact that $E^2 = (pc)^2 + (mc^2)^2$ even for $m = 0$, which follows from the definition of the momentum four-vector and its behavior under Lorentz transformations). Therefore, the atom will also be moving with momentum $p_a = mv_r$ where m is the mass of the atom and $v_r = \frac{\hbar k}{m}$ is the recoil velocity of the atom. For optical transitions ($\lambda \sim 600$ nm) of single atoms or molecules (masses between 1 and 100 amu) this velocity ranges from 1 mm/s to 1 m/s. For the 689 transition of Sr88 it is 6.6 millimeters/second.

If the atom with momentum $\mathbf{p}$ begins in the ground state an absorbs a photon with energy $\hbar\omega$ then the atoms momentum will be modified from $\mathbf{p}' = \mathbf{p} + \mathbf{p}_\gamma$ until otherwise modified. If the photon is spontaneously released in a random direction, the emmision of the photon will on average not lead to any net change in momentum over many cycles of absorption and emission, but the asborption will always push the atom along the direction the photon was traveling. We can use this fact to decellerate an atom moving with some initial velocity. All that is needed to to shine a laser on it tuned near some transition so that the atom absorbs and emits many photons. Therefore the excited state must decay back to the ground state.

However, we know that an atom can only absorb and emit light at specific frequencies where the photon energy $\hbar\omega$ matches the energy difference between two distinct internal atomic states, or energy levels. If the

light is detuned away from the correct frequency, the probability of absorption decreases away from resonance:

$$Probability\ of\ absorption \propto \frac{1}{1+\delta^2}$$

where $\delta \propto \omega - \omega_0$ with $\omega_0 = \Delta E/\hbar$ being the energy splitting of the levels divided by $\hbar$ and ω being the frequency of the light *in the atom's rest frame*. Of course the rest frame of the atom changes after every absorbtion and emission event, and this means that the resonance condition is dependent upon the velocity of the atom.

Now we can begin to consider an atom moving (nonrelativistically) at some velocity v in the presence of a laser beam with frequency ω_0, with both v and ω_0 defined relative to the laboratory's intertial frame. The nonrelativistic Doppler shift of the light in the atom's from is $\Delta\omega_D = kv$ where $k = \frac{2\pi}{\lambda_0} = \omega_0/c$ (I'm so bad about the relationships between k,ω, ν, λ, and c, I'm just awful at it). As the atom cycles photons, its velocity will change and the probability of absorption will also change. Therefore we want a more quantitative idea of exactly how close the Doppler shifted light $\omega_D = \omega_0(1 + v/c)v/c$ need to be to ω_0 to drive the transition. Even better would be to have a quantitative idea of how fast, in absolute terms, we can actually cycle through photons.

The quantity intrinsic to the atom that determines both of these things is the linewidth of the transition, Γ. It is a fact of nature not derived here that when an excited state decays to a ground state over time according to $P_{excited} \propto \exp -t/\tau = \exp -\Gamma t$ that the steady-state rate (R_{scatt}) at which photons will be spntaneously emitted while illuminated by light with frequency ω and lifetime τ is given by

$$R_{scatt} = \frac{\Gamma}{2}\frac{I/I_{sat}}{1 + I/I_{sat} + (2\delta/\Gamma)^2} \tag{4.1}$$

where Γ is both the inverse lifetime and the FWHM of the resonance (in the low intensity limit), I_{sat} is the saturation intensity of the transition, I is the intensity of the light experiences by the atom, and $\delta = \omega - \omega_0$ is the detuning of the light from reonance. The saturation intensity

$$I_{sat} = \frac{\pi hc}{3\lambda^3\tau} = \frac{2\pi^2\hbar}{3\lambda^3}\Gamma = \frac{\hbar\omega^3}{12\pi c^2}\Gamma \tag{4.2}$$

is another measure of how easily the transition is driven due to its dependence on Γ. A transition is "strong" when the saturation intensity is high, or when the excited state decays very rapidly to the ground state. A "strong" transition is one that goes easily, and the excited state lifetime is a direct measure of this quantity. Strong transitions are also much wider, meaning that the propability of absorbing a photon remains high for a range of frequencies.

As a brief aside, an intuitive way of recalling the saturation intensity is that it is the energy per photon $\hbar\omega$ multiplied by the maximum scattering rate $\Gamma/2$ and then divided by an appropraite quantity of area. The nature measure of the area is the resonant scattering cross-section

$$\sigma_0 = \frac{3\lambda^2}{2\pi} = 6\pi\lambda^2.$$

The key to remembering the optical cross section is it is the a circular area the size of the reduced wavelength (we use reduced wavelength because we always use the versions with extra factors of 2π) then multiplied by a factor of six, where a factor of three comes from three dimensions and a factor of two from two polarizations. In terms of the resonant cross-section we have

$$I_{sat} = \frac{\Gamma\hbar\omega}{2\sigma_0}$$

The scattering rate formula of Eq. 4.1 is central to understanding how light and atoms interact. It is clear from the form of Eq. 4.1 that the maximum scattering rate saturates at $\Gamma/2$ as $I/I_{sat} \to \infty$. There is no lower bound for the scattering rate, which is obvious when one considers how many photons the atom will scatter when there is not light to scatter. The FWHM of the resonance approches the natural linewidth Γ

as $I/I_{sat} \to 0$, but the line actually broadens as $I/I_{sat} \to \infty$ and we may rewrite Eq. 4.1 to illuminate this mathematically, introducing the so-called saturation parameter $s_0 = I/I_{sat}$ to simplify:

$$R_{scatt} = \left(\frac{s_0}{1+s_0}\right)\left(\frac{\Gamma/2}{1+(2\delta/\Gamma')^2}\right) \tag{4.3}$$

where $\Gamma' = \Gamma\sqrt{1+s_0}$ is the power broadened linewidth. When $\delta = \Gamma'/2$ the scatering rate is half its value on resonance. The saturation parameter s_0 can be generalized to off-resonant light as well,

$$s = s_0/(1+(2\delta/\Gamma)^2)$$

and is minimimized for $\delta = 0$. Off-resonant scattering is slower, but also saturates at lower intensities. I personally find the off-resonant saturation paraeter and corresponding off-resonant saturation intensities to be more confusing than helpful, and generally even prefer to explicitly use I/I_{sat} in place of s_0.

Now we have the tools to consider slowing an atom move at velocity v. To slow down, it must cycle many photons from source, typically a laser, propogating opposite the atoms velocity, and the light must be close enough, inthe atom's frame, to be absorbed. The linewidth Γ is the obviously reasonable metric for how far from resonance the photon can be before the laser stops effectivly slowing the atom. More precisely it is the power broadened line width, though since there is no upper bound on Γ', we use the natural linewidth in defining the range of frequencies and thus atomic velocities which will be slowed by the laser. We say that atoms within one natural linewidth are able to be captured by the deceleration, while those with velocities exceeding this window are Doppler shifted too far from resonance to be reliably captured. Therefore the capture vel0city is

$$v_c = \frac{\Gamma}{k} = \frac{\omega\Gamma}{2\pi}. \tag{4.4}$$

A recurring theme in our discussion will be identifying the various velocity, temperature, and energy scales. A word of caution is warranted because although we routinely speak of the ensemble of atoms interacting with our cooling light as having a temperature, the atoms are not in thermal quilibrium with each other in a MOT. However, we are in a uniquely good position to simply define a temperature according the the expectation value of the atomic ensemble:

$$\frac{1}{2}k_bT = \langle E_k \rangle \tag{4.5}$$

where the $\frac{1}{2}k_bT$ is the standard expected energy for a given degree of freedom from the equipartition theorem. We can thus turn our capture velocity into a capture temperature T_c

$$T_c = \sqrt{\frac{mv_c^2}{k_b}} = \frac{\Gamma}{2\pi}\sqrt{\frac{m}{k_b}}. \tag{4.6}$$

We shold note that "caputure temperature" is not a particularly common thing to talk about, and the symbol T_c is often used to mean a "critical temperature", e.g. the transition temperature of a BEC.

The rate at which the atom is decellerated is simply the rate of photons being cycled, given by R_{scatt} multiplied by the average change in velocity per cycled photon, which is $\hbar k/m$. Therefore the force on the atom is

$$F_{scatt} = ma = \hbar k R_{scatt} \xrightarrow{R_{scatt}\to\frac{\Gamma}{2}} \frac{\hbar k}{2}\Gamma \tag{4.7}$$

4.2 Optical Molasses

The previous section discussed the very basics of how an atom's velocity can be slowed due to a presence of a single beam. Now we consider the effect of two counter-progating beams of the same frequency (we will add some effects of beams in other directions by hand in our discussion of the Doppler temperature). In all cases, the force associated on the atom from a given laser beam points in the same direction as the laser. However, while in the lab frame each beam is the same frequency, the moving frame of the atom will cause a differential Doppler shift. We can evaluate the force form either beam

$$F_\pm = \pm\frac{\hbar k\Gamma}{2}\frac{s_0}{1+s_0+(2\delta \mp \Delta\omega_D/\Gamma)^2} \tag{4.8}$$

where δ is the detuning from resonance for an atom in the laboratory frame and $\Delta\omega_D = kv$ is the Doppler shift contribution to the detuning seen differently in the moving atom's frame. These forces and the net force on the atom are shown in Fig. 4.1. We can see that for the particular combination of detuning and intensity in Fig. 4.1a there is a central region where the force is a pure damping force to a very good approximation. In this regime the optical molasses force may be approxmately written (Eq. 7.2 from Metcalf) as

$$\mathbf{F}_{OM} \simeq \frac{8\hbar k^2 \delta s_0 \mathbf{v}}{\Gamma(1 + s_0 + (2\delta/\Gamma)^2)^2} \equiv -\beta \mathbf{v}. \tag{4.9}$$

However, this expression does not apply generally, reducing the linewidth while holding the detuning and saturation parameter constant[1] shows how the force may deviate from a pure viscous force.

We can evaluate this damping by consider the evolution of the atom's kinetic energy:

$$\frac{d}{dt}\left(\frac{1}{2}mv^2\right) = mv\frac{dv}{dt} = vF_{OM} = -\beta v^2. \tag{4.10}$$

4.2.1 The Doppler cooling limit

In the abscence of any other phenomena, the optical molasses force would damp the atomic motion to zero, but of course reflecting on the physical orgin of the source as a series of discrete steps in momentum space shows how the atom will experience a random walk once it has been cooled near $T = 0$. Following Section 9.3.1 of Foot, we can derive the limit on cooling that this process imposes. We begin by writing the force on the atom in terms of its average and its fluctuating components for both absorption and emission.

$$\mathbf{F} = \overline{\mathbf{F}}_{abs} + \delta\mathbf{F}_{abs} + \bar{\mathbf{F}}_{spont} + \delta\mathbf{F}_{spont} \tag{4.11}$$

where $\overline{\mathbf{F}}_{abs} = F_{scatt}$ from Eq. 4.7 and $\overline{\mathbf{F}}_{spont} = 0$ due to isotropic spontaneous emmission. The fluctuating components are of interest now. Each photon spontaneously emitted causes the atom to recoil to conserve momentum, and this recoil is equivalent to a step of size

$$v_r = \frac{\hbar k}{m}, \tag{4.12}$$

which we call the recoil velocity. In a given period of time t the number of spontaneous emmision events is

$$N_{scatt} = 2R_{scatt}t. \tag{4.13}$$

The factor two results from our assumption of low intensity so we may neglect saturation and the scattering rates may simply be summed. After N steps of a random walk the expected displacement scales as $\sqrt{N}$ times the step length. In other words, the mean squared displacement grows as N times the step length squared.

$$\langle v^2 \rangle_{spont} = \eta_{spont} v_r^2 (2R_{scatt})t, \tag{4.14}$$

where the proportionality term $\eta_{spont} = 1/3$ comes from the fact that the photon emits in all directions isotropically so the size of the kick along the axis under consideration is $\hbar k \cos\theta$ and we average over $\langle \cos^2\theta \rangle = \eta_{spont} = 1/3$.

To incorporate the $\delta\mathbf{F}_{abs}$ we consider the kicks from both beams, where our random walk is realized because any given kick will be one of the two beams randomly. The argument is exactly as for $\delta\mathbf{F}_{spont}$ and we have

$$\langle v^2 \rangle_{abs} = \eta_{abs} v_r^2 (2R_{scatt}t), \tag{4.15}$$

where $\eta_{abs} = 1$ because all the kicks from absorption are along the axis under consideration. Now lets consider again the evolution of the kinetic energy as we did in Eq. 4.10, but now considering all three beam pairs:

$$\frac{1}{2}m\frac{d\overline{v^2}}{dt} = (1 + 3\eta_{spont})\frac{1}{2}mv_r^2(2R_{scatt}) - \beta\overline{v^2}. \tag{4.16}$$

[1]The saturation parameter is held constant, but the actual value of the saturation intensity, and thus the intensity being used, does differ in the two cases, but the detuning does not. The detuning as a function of the linewidth does change, but not the absolute value.

where the factor of 3η coming from the contribution to the $\hat{x}$ direction from $\hat{x}$, $\hat{z}$, and $\hat{z}$ directions. Setting that derivative equal to zero yields

$$\overline{v^2} = 2mv_r^2 \frac{R_{scatt}}{\beta}$$

and using our standard prescription for defining temperature scales via $\frac{1}{2}mv^2 = \frac{1}{2}k_bT_D$ gives:

$$\frac{1}{2}k_bT_D = m^2v_r^2 \frac{R_{scatt}}{\beta} \tag{4.17}$$

ok, so

$$R_{scatt}/\beta = -\frac{\Gamma^2 + 4\delta^2}{16\hbar k^2 \delta} \tag{4.18}$$

and we are restricting ourselves to the case where $\delta < 0$ so that the force is a damping force. This ratio then has a minimum at $\delta = -\Gamma/2$ where it assumes the value $2\Gamma^2/(16\hbar k^2\Gamma/2) = \Gamma/(4\hbar k^2)$. the term $m^2v_r^2$ is $\hbar^2k^2$ and so we have

$$\frac{1}{2}k_bT_D = \hbar\Gamma/4 \implies T_D = \frac{\hbar\Gamma}{2k_b} \tag{4.19}$$

which is the correct value according to the Yallow book section 7.2.

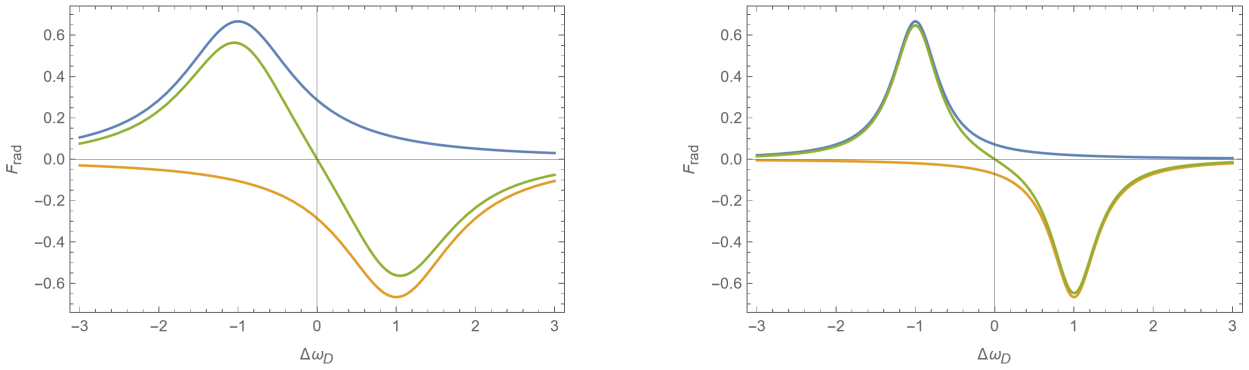
4.3 One dimensional magneto-optical trap

Theprevious section discussed the very basics of how an atom's velocity can be slowed due to a presence of two counter-propagating beams of equal frequency. The effect was to create an effective viscous force on the atom, which is a force proportional to the velocity of the atom. However, the atom will continue to move in a random walk as photons are scattered, there is no restoring force that serves to confine the atom. In this section we will introduce the idea of trapping and confinement and put these ideas together to create the Magneto-Optical Trap (MOT). It is also at this point we are forced to move beyond the two-level approximation implicit in the above discussion. This is because the restoring force the magneto-opticl effect causes derives from the differential radiative forces on the atom due to the presence of a magnetic field.

Let us consider an atom with two manifolds, each with a single "sublevel", where the ground state has a single zeeman and the excited state has three Zeeman states which may be split via the Zeeman effect. This is essentially the case for the bosonic isotopes of Strontium, but in general the ground and excited manifold may have multiple sublevels. In fact, for atoms with nuclear spin, there are two hierarchies of sublevels, one from the electronic spin-orbit interactions, called the fine structure, and one from the nuclear spin-orbit interactions, called the hyperfine structure. Each sublevel may then have multiple Zeeman sublevels. So the general case is that the ground and excited manifolds have multiple sublevels, each of which may have multiple Zeeman sublevels. But typically these effects can be considered manually (like the possibility of a repump laser being required to return atoms from a dark state).

It should not come as a surprise that to understand the physics of the Magneto-optical trap, we ought to return to considering the scattering force, this time as a function of position instead of velocity. Let us label our three excited Zeeman sublevels with $|e,+\rangle$, $|e,0\rangle$, and $|e,-\rangle$, which correspond to atoms with azimuthal quantum number $m = +1$, $m = 0$, and $m = -1$ respectively. The ground state is labeled $|g,0\rangle$, and we will assume that the ground state has no magnetic substructure. We have not labeled the m quantum number with J or I to maintain generality at this stage. The Magneto part of the magneto-optical trap is provided by a quadrupole magnetic field, which is a magnetic field that varies linearly with position. The magnitude of the field is proportional to the displacement from the origin, through the direction may be inward or outwar (and in fact will be two of one and one of the other). As before, we consider two counter propagating beams of the same frequency. We note we must also now specify the polarization states of the beams, as a consequence of the dipole transition selection rules we are considering (though we may not have mentioned that fact yet).

4.4 Tables



(a) (b)

Figure 4.1: The force from the two counter-propagating beams and their combined force on the atom for two different values of Γ. In each case $s_0 = 2$ and $\delta = -1$ and the prefactor $\hbar k\Gamma/2$ has been set equal to unity. In (a) Γ=1 while in (b) $\Gamma = 0.4$. The linear regime near the center of (a) represents a viscous damping force since $\Delta\omega_D$ is directly proportional to the atom's velocity. In (b) the force it still a damping force but, not directly proportional to the velocity, so Eq. 4.8 does not hold for (b) while it does for (a).

In Strontium, two transitions are used to cool atoms, the strong 461 nm line coupling the $^1S_0 - ^1P_1$ states, and the much weaker 689 nm line coupling the $^1S_0 - ^3P_1$ states.

Out[ ]=

	Formula	Values
Wavelength	λ	460.862 meters nano
Linewidth	Γ	$\frac{2.01\times10^{8}}{\text{seconds}}$, 2π (31.9901 Hz Mega)
Frequency	$\omega = 2\pi\nu$	$\frac{4.08723\times10^{15}\ \text{radians}}{\text{seconds}}$, 2π (650.503 Hz Tera)
Recoil Velocity	$v = \frac{\hbar k}{m}$	$\frac{0.00984961\ \text{meters}}{\text{seconds}}$, $\frac{9.84961\ \text{meters micro}}{\text{milli seconds}}$
Lifetime	$\tau = \frac{1}{\Gamma}$	4.97512 nano seconds
Saturation intensity	$Isat = \frac{\hbar\omega\Gamma}{2\sigma 0} = \frac{2\pi^2\hbar c\Gamma}{3\lambda^3}$	$\frac{42.7158\ \text{milli Watts}}{\text{centi}^2\ \text{meters}^2}$
Optical cross–section	$\sigma 0 = \frac{3\lambda^2}{2\pi} = 6\pi\lambda\!\!\!^{-2}$	$0.101411\ \text{meters}^2\ \text{micro}^2$
Recoil Energy	$\frac{Erec}{\hbar} = \omega_r = \frac{\hbar k^2}{2m}$	67.1426 Hz Kilo, 2π (10.6861 Hz Kilo)
Recoil Ratio	$\frac{\omega_r}{\Gamma}$	0.000334043
Capture Velocity	$v_c = \frac{\Gamma}{k} = \frac{\Gamma\lambda}{2\pi}$	$\frac{14.743\ \text{meters}}{\text{seconds}}$
Doppler temperature	$T_D = \frac{\hbar\Gamma}{2KB}$	767.642 Kelvin micro
Recoil Temperature	$T_r = \frac{2\omega r\hbar}{KB}$	1.0257 Kelvin micro

Figure 4.2: Table of values for the 461 nm transition

Out[ ]=

	Formula	Values
Wavelength	λ	689.449 meters nano
Linewidth	Γ	$\frac{46900.}{\text{seconds}}$, 2π (0.00746437 Hz Mega)
Frequency	$\omega = 2\pi\nu$	$\frac{2.73211\times10^{15}\text{ radians}}{\text{seconds}}$, 2π (434.829 Hz Tera)
Recoil Velocity	$v = \frac{\hbar k}{m}$	$\frac{0.00658397\text{ meters}}{\text{seconds}}$, $\frac{6.58397\text{ meters micro}}{\text{milli seconds}}$
Lifetime	$\tau = \frac{1}{\Gamma}$	21 322. nano seconds
Saturation intensity	$Isat = \frac{\hbar\omega\Gamma}{2\sigma 0} = \frac{2\pi^2\hbar c\Gamma}{3\lambda^3}$	$\frac{0.00297695\text{ milli Watts}}{\text{centi}^2\text{ meters}^2}$
Optical cross–section	$\sigma 0 = \frac{3\lambda^2}{2\pi} = 6\pi\lambda\!\!\!^{-2}$	$0.226958\text{ meters}^2\text{ micro}^2$
Recoil Energy	$\frac{Erec}{\hbar} = \omega_r = \frac{\hbar k^2}{2m}$	30.001 Hz Kilo, 2π (4.77481 Hz Kilo)
Recoil Ratio	$\frac{\omega_r}{\Gamma}$	0.63968
Capture Velocity	$v_c = \frac{\Gamma}{k} = \frac{\Gamma\lambda}{2\pi}$	$\frac{0.0051463\text{ meters}}{\text{seconds}}$
Doppler temperature	$T_D = \frac{\hbar\Gamma}{2KB}$	0.179116 Kelvin micro
Recoil Temperature	$T_r = \frac{2\omega r\hbar}{KB}$	0.458309 Kelvin micro

Figure 4.3: Table of values for the 698 nm transition

Chapter 5

Math

5.1 Gaussian integrals

$$\int_{\infty}^{\infty} e^{-\frac{ax^2}{2}} dx = \sqrt{\frac{2\pi}{a}} \tag{5.1}$$

$$\int_{\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx = \sigma\sqrt{2\pi} \tag{5.2}$$

$$\int re^{-\frac{r^2}{2\sigma^2}} dr = \sigma^2 \tag{5.3}$$

$$\int_{\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \times \int_{\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy = 2\pi\sigma^2 \tag{5.4}$$

5.2 Green's Theorem

Following section 1.8 of Jackson. Green's theory follows from the divergence theorem, which states that the volume integral of the divergence of a vector field $\mathbf{A}$ is equal to the surface integral of the normal component of $\mathbf{A}$ over the surface enclosing the volume.

$$\int_V \nabla \cdot \mathbf{A} dV = \oint_S \mathbf{A} \cdot d\mathbf{a} \tag{5.5}$$

for any well-ehaved vector field $\mathbf{A}$ and volume V enclosed by surface S. Let $\mathbf{A} = \phi\nabla\psi$, where ϕ and ψ are arbirary scalar fields (i suppose subject to the constraint of producing a suitable vector field). We use the identity

$$\nabla \cdot (\phi\nabla\psi) = \phi\nabla^2\psi \cdot + \nabla\phi \cdot \nabla\psi \tag{5.6}$$

Alongside this, the normal component of $\mathbf{A}$ on the surface S is given by:

$$\phi\nabla\psi \cdot \mathbf{n} = \phi\frac{\partial\psi}{\partial n} \tag{5.7}$$

where $\mathbf{n}$ is the (outward) unit normal to the surface S. Substituting $\mathbf{A} = \phi\nabla\psi$ in the divergence theorem, we get:

$$\int_V (\phi\nabla^2\psi + \nabla\phi \cdot \nabla\psi) dV = \oint_S \phi\frac{\partial\psi}{\partial n} d\mathbf{a} \tag{5.8}$$

Consider 5.8 where wwe exchange the roles of ϕ and ψ and subtract the result from the original. We get

$$\int_V (\phi\nabla^2\psi - \psi\nabla^2\phi) dV = \oint_S \left(\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n}\right) d\mathbf{a} \tag{5.9}$$

which is Green's theorem. Ok. Cool. But how is this helpful for us?

What are Green's functions? Building up intuition (The following discussion actually is my own thoughts, obviously informed by texts, but this is mostly me, expect when i refer to definitonsand terms like fundamental solution or the boundary condisiton) A Green's function is a response function. A related idea is a so-called *fundamental solution*; all Green's functions are fundamental solutions, but not all fundamental solutions are Green's functions. A fundamental solution is a solution to a differential equation with a delta function source. A Green's function is a fundamental solution that is also symmetric and obeys boundary conditions.

What does that mean? Well, maybe this is bad intuition, but I think of a differential operator L as a thing which relates function to its derivatives, and therefore L is a thing that in some sense "propagates" the function from places where it is known, to the rest of the domain. L is not a *propagator*, its a connector, again i mean that colloquilally. It tells you how a particular distribution of the function will eolve as one moves within the domain.

A fundamental solution is a function F such that $LF \propto \delta$ (the coefficent of the delta function depends on the dimensionallity and coordinate system, the). In other words, F is a function that is zero everywhere except at the origin, where it is infinite. The delta function is a source, and F is the response to that source. It is often called an impulse response function. Importantly, a fundamental solution F does not depend on the boundary conditions of the problem, it is a property of the differential operator L alone. To reiterate, a fundamental solution (function) F of a differential operator L satisies

$$LF \sim \delta(x) \tag{5.10}$$

The δ function is a (impulsive) source, and it is infinitely sharp, short, and tall. A point particle with mass $m > 0$ is a impulse source of the gravitational potential, and the operator which connects that source to the potential is the Laplacian. The fundamental solution of the Laplacian is the Newtonian potential, which is the potential of a point particle with mass m at the origin. In this case, when we said "it is infinitely sharp, short, and tall" mean that the mass density is infinite, and the spatial extent is zero. the situation is literally exactly analogous for the electric potential of a point charge. An infinitely loud, brief acoustic signal would be another example of a delta function source.

As stated above, a Green's function is an example of a fundamental solution, but it is a special kind of fundamental solution. A Green's function is a fundamental solution that obeys the boundary conditions of the problem. In other words, it is a fundamental solution that is also a solution to the problem for an impulse source. but it is alsot the case that the Green's function always satisfies

$$LG(x, x') = \delta(x - x') \tag{5.11}$$

We have just referred to the boundary conditions of a problem; the boundary conditions are the conditions imposed on the solution at the boundary of the domain. The boundary conditions are going to be the other thing that determines the form of the Green's function we will leverage to provide our solution. Let's consider again Green's theorem, Eq. 5.9 when we have decided (just now, right htis moment) that the function ψ is the Green's function $G(x, x')$ and the function ϕ is the solution to the problem $f(x)$. What we mean by this is that in addition to 5.11, we also have

$$L\phi(x) = f(x) \tag{5.12}$$

and we want to solve for $\phi(x)$. We can evaluate Eq. 5.9 with these facts in mind to find that

$$\int_V \phi(x)\cancelto{\delta(x-x')}{\nabla^2 G(x,x')} - G(x)\cancelto{f(x')}{\nabla^2 \phi(x')}\, dV = \oint_S \left(\phi(x')\frac{\partial G(x,x')}{\partial n} - G(x,x')\frac{\partial \phi(x')}{\partial n}\right) d\mathbf{a}, \tag{5.13}$$

which resolves and rearranges to give an integral equation for $\phi(x)$

$$\phi(x) = \int_V G(x - x')f(x')d^n x' + \oint_S \left(\phi(x')\frac{\partial G(x,x')}{\partial n'} + G(x,x')\frac{\partial \phi(x')}{\partial n'}\right) d\mathbf{a}', \tag{5.14}$$

And here is where we begin to see the power of Green's functions. We have reduced the problem of solving a differential equation to the problem of solving an integral equation, which is often much easier. But *why* is it easier? Well, we have a lot of freedom in choosing the Green's function, and we can use that freedom to

make the integral equation easier to solve. In particular, we may choose Green's function to make one of the terms (in general you cannot supress both) in the surface integral vanish.

This brings us to the two primary types of boundary conditions we will encounter: Dirichlet and Neumann. Dirichlet boundary conditions are those where the value of the function is specified at the boundary, while Neumann boundary conditions are those where the value of the (normal) derivative of the function is specified at the boundary. Those terms refer to our knowledge of the function at the boundary, not the form of the Green's function, but they determine certain properties of the Green's function. If we know the value of the function at the boundary (Dirichlet), we can choose the Green's function to make the second term in Eq. 5.14 vanish. If we know the value of the (normal) derivative of the function at the boundary (Neumann), we can choose the Green's function to make the first term in Eq. 5.14 vanish.

Constructing Green's functions for Dirichlet boundary conditions

(Following in part Section 8.5 of "Modern Electrodynamics" by Andrew Zangwill (first edition i believe), also following https://home.cc.umanitoba.ca/ dtrim/BooksandNotes/PDE/Green.pdf (notes from chapter 13 of an unpublished revision of Donald Trim's "Applied Partial Differential equations", i have it downloaded to hogan labs file server if that link dies)) Of course all that is easier said than done, and we need to know how to construct Green's functions. One method is the *method of images* which uses image charges to construct the greens function. But it is only applicable to a limited set of problems. Another more general method is to use the *method of eigenfunction expansion*. This method is general, but it is also a bit abstract. However, it is a recipe for constructing Green's functions for any problem procedurally.

In the method of eigenfunction expansion, we begin by expanding the Green's function in terms of the eigenfunctions of the differential operator L. We will start with a general two dimensional problem in the hopes that generalization will be clear. We may make furtur assumptions which reduce the generality as we proceed. But for now, we remain general, though it should be noted that Trim uses a rectangular opening in a grounded plane to illustrate the method, so modifications to the equations derived from Trim's text will be aplenty). We begin by expanding the Green's function in terms of the eigenfunctions of the differential operator L:

$$\begin{aligned} Lu + \lambda^2 u &= 0 \quad \in V \\ u(x,y) &= 0 \quad \in S \end{aligned} \tag{5.15}$$

where λ is the eigenvalue and u is the eigenfunction defined over the domain V and its boundary S. Let us label the eigenfunctions $u_n(x,y)$ and the eigenvalues λ_n. We may expand the Green's function in terms of the eigenfunctions of the differential operator L:

$$G(x,y;x',y') = \sum_n \frac{u_n(x,y)u_n^*(x',y')}{\lambda_n^2} \tag{5.16}$$

similar (but more general) to Zangwill (8.5.3 eq 8.71). Trim instead expresses this in a less developed form in eq. 13.23. which i shall gneralize to

$$G(x,y,x',y') = \sum_n c_n u(x,y) \tag{5.17}$$

so we must calculate the coefficients c_n. We do this by substituting Eq. 5.17 into the necessarily true statement $LG(x,y;x',y') = \delta(x-x')\delta(y-y')$

$$\begin{aligned} LG(x,y;x',y') &= \delta(x-x')\delta(y-y') \\ \sum_n c_n L u_n(x,y) &= \delta(x-x')\delta(y-y') \\ \sum_n c_n(-\lambda_n^2)u_n(x,y) &= \delta(x-x')\delta(y-y') \\ \sum_n c_n \lambda_n^2 u_n(x,y) &= -\delta(x-x')\delta(y-y') \end{aligned} \tag{5.18}$$

now, we express the delta function as a sum of the eigenfunctions (this is an identity that is true generally? i think? its is intuitively obvious, but is expounded upon in section 7.4 of Zangwill, and i think it is also in Trim, but i dont have that part of the text)

$$\delta(x-x')\delta(y-y') = \sum_n u_n(x,y)u_n^*(x',y') \tag{5.19}$$

comparing the last line of Eq. 5.18 and Eq. 5.19 we see that the coefficients c_n are given by

$$c_n = -\frac{u_n^*(x',y')}{\lambda_n^2} \tag{5.20}$$

which is the result Zangwill implicitly uses.

I want to make clear that i'm modifying the source text to fit with each other better, and so comparison to the source text will require some parsing to see how the equations relate.

This is the general form of the Green's function expansion, the only other detail is to determine the eigenfunctions and eigenvalues. We do this by staring at the problem and guessing. For example, if we have a rectangular opening in a grounded plane, we might guess that the eigenfunctions are the sines and cosines of the form $u(x,y) = \sin(k_x x)\sin(k_y y)$, and we would be right. We would then solve for the eigenvalues by substituting the eigenfunctions into the eigenvalue problem, Eq. 5.15, and solving for λ. This is the sad reality of differential equations.

Constructing Green's functions for Neumann boundary conditions

meh. i dont want to do this right now.

5.3 trig

5.3.1 hyperbolic trig

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \tag{5.21}$$

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \tag{5.22}$$

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{5.23}$$

$$\cosh^2(x) - \sinh^2(x) = 1 \tag{5.24}$$

the inverse hyperbolic trig functions are defined as

$$\begin{aligned}
\operatorname{arsinh} x &= \ln\left(x + \sqrt{x^2+1}\right) && -\infty < x < \infty, \\
\operatorname{arcosh} x &= \ln\left(x + \sqrt{x^2-1}\right) && 1 \le x < \infty, \\
\operatorname{artanh} x &= \frac{1}{2}\ln\frac{1+x}{1-x} && -1 < x < 1, \\
\operatorname{arcsch} x &= \ln\left(\frac{1}{x} + \sqrt{\frac{1}{x^2}+1}\right) && -\infty < x < \infty, x \neq 0, \\
\operatorname{arsech} x &= \ln\left(\frac{1}{x} + \sqrt{\frac{1}{x^2}-1}\right) && 0 < x \le 1, \\
\operatorname{arcoth} x &= \frac{1}{2}\ln\frac{x+1}{x-1} && -\infty < x < -1 \text{ or } 1 < x < \infty.
\end{aligned} \tag{5.25}$$

composition with the hyperbolic trig functions yields the following identities

$$\begin{aligned}
\sinh(\operatorname{arcosh} x) &= \sqrt{x^2-1} \quad \text{for} \quad |x|>1 \\
\sinh(\operatorname{artanh} x) &= \frac{x}{\sqrt{1-x^2}} \quad \text{for} \quad -1<x<1 \\
\cosh(\operatorname{arsinh} x) &= \sqrt{1+x^2} \\
\cosh(\operatorname{artanh} x) &= \frac{1}{\sqrt{1-x^2}} \quad \text{for} \quad -1<x<1 \\
\tanh(\operatorname{arsinh} x) &= \frac{x}{\sqrt{1+x^2}} \\
\tanh(\operatorname{arcosh} x) &= \frac{\sqrt{x^2-1}}{x} \quad \text{for} \quad |x|>1
\end{aligned} \tag{5.26}$$

5.4 Derivatives

5.4.1 Derivative of a^x

To find the derivative of a^x with respect to a, we employ logarithmic differentiation. Let's start by taking the natural logarithm of both sides:

Differentiating both sides with respect to a gives:

Chapter 6

Optics1

6.1 summary

6.1.1 Diffraction

One form of the Kirchhoff integral is (Eq. 8.3.7 Born and Wolf, 1999)

$$U(P) = \frac{1}{4\pi} \oint_S \left[U \frac{\partial}{\partial n} \left(\frac{e^{iks}}{s} \right) - \frac{e^{iks}}{s} \frac{\partial U}{\partial n} \right] dS \tag{6.1}$$

This is an approximate solutions to the Helmholtz equation

$$\nabla^2 U + k^2 U = 0 \tag{6.2}$$

Sections 8.2-8.3, and particularly figures 8.1 through 8.5 of Born and Wolf (1999) are very helpful in understanding the Kirchhoff integral. a simplification for source and observations points far from the aperture is (Eq. 8.3.17)

$$U(P) = -\frac{iA}{2\lambda} \int_{\mathcal{W}} \frac{e^{ik(r_0+s)}}{r_0 s} [1 - \cos(\chi)] dS \tag{6.3}$$

A further simplification is (Eq. 8.3.18) for a point source, another clever geometric argument is made to show that the above integral is equal to

$$U(P) = -\frac{iA}{2\lambda} \int_{\mathcal{W}} \frac{e^{ik(r_0+s)}}{r_0 s} [1 - \cos(\chi)] dS \tag{6.4}$$

where $\chi - (r_0, s)$ is π minus the (nearly straight) angle at a point Q on the portion of a spherical surface $\mathcal{W}$ centered at the source P_0 and filling the aperture, which some edge bits we neglect. See Figure 8.4 of Born and Wolf, which is helpful.the amplitude factor $K(\chi)$ is given by

$$K(\chi) = -\frac{i}{2\lambda} [1 + \cos(\chi)] \tag{6.5}$$

Furthur reduction permits

$$U(P) \approx -\frac{iA}{\lambda} \frac{\cos\delta e^{ik(r'+s')}}{r's'} \int_{\mathcal{A}} e^{-ikf(\xi,\eta)} d\xi d\eta. \tag{6.6}$$

where

$$f(\xi,\eta) = -\frac{x_0\xi + y_0\eta}{r'} - \frac{x\xi + y\eta}{s'} + \frac{\xi^2+\eta^2}{2r'} + \frac{\xi^2+\eta^2}{2s'} + \frac{(x_0\xi + y_0\eta)^2}{2r'^3} - \frac{(x\xi + y\eta)^2}{2s'^3} + \cdots . \tag{6.7}$$

When the quadratic and higher terms in 6.7 may be neglected, we have Fraunhofer diffraction, when they cannot we have Fresnel diffraction.

6.2 Beam Shaping

This section is from "Laser Beam Shaping: Theory and Techniques" by Fred M. Dickey and Scott C. Holswade (2014). I have paraphrased and editted it down for my own study purposes. This is not my original work.

In analogy with the Reynolds number in fluid dynamics, the Fresnel beta number is a dimensionless parameter that describes the relative importance of diffraction effects to the effects of geometric optics. The Fresnel beta number is defined as

$$\beta = \frac{2\pi RD}{\lambda f} \tag{6.8}$$

where R is the characteristic length scale of input beam, D is the characteristic length scale of the output beam, λ is the wavelength of the light, and f is propagation distance (suggestively given as the focal length of a lens). The Fresnel number is a measure of the relative importance of diffraction effects to the effects of geometric optics. When $F \ll 1$ the beam is said to be in the Fresnel regime, and when $F \gg 1$ the beam is said to be in the Fraunhofer regime.

We choose to expres out irradiance distribution in dimensionaless forms using lowercase letters for input functions, e.g. $g(x/R, y/R)$ and uppercase for output functions, e.g. $G(x/D, y/D)$. When β is large, geometric optics is a good approximation and the irradiance distribution is given by the Fourier transform of the input irradiance distribution. When β is small, diffraction effects are important. Generally speaking, our beam shaping task will be easier with large β and more difficult with small β. In the geometrical limit, we can reshape to any output we want, while this may become impossible for small β.

The continuity of the a beam shaping lens is also relevant,smooth lens allow good beam shaping, while discontinuities in the lens (as in a Fresnel lens) will require a much larger β to achieve the same beam shaping.

6.2.1 Fourier Theory

The Fourier transform of a function $f(x)$ is defined as

$$F(k) = T[f(x)] = \int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{6.9}$$

and the inverse Fourier transform is

$$f(x) = T^{-1}[F(k)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(k)e^{ikx}dk \tag{6.10}$$

The Fourier transform of a function $f(x, y)$ is defined as

$$F(k_x, k_y) = T[f(x, y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)e^{-i(k_x x + k_y y)}dxdy \tag{6.11}$$

and the inverse Fourier transform is

$$f(x, y) = T^{-1}[F(k_x, k_y)] = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(k_x, k_y)e^{i(k_x x + k_y y)}dk_x dk_y \tag{6.12}$$

Parseval's Theorem

$$2\pi\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(k)|^2 dk \tag{6.13}$$

and in two dimensions we have

$$(2\pi)^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x, y)|^2 dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |F(k_x, k_y)|^2 dk_x dk_y \tag{6.14}$$

Convolution Theorem Skipping directly to the two dimensional case, the convolution theorem states that if $F(k_x, k_y)$ is the Fourier transform of $f(x, y)$ and $G(k_x, k_y)$ is the Fourier transform of $g(x, y)$ then the Fourier transform of the convolution of f and g is given by the product of the Fourier transforms of f and g:

$$T[\{f(x, y) * g(x, y)\}] = F(k_x, k_y)G(k_x, k_y) \tag{6.15}$$

or, in terms of the convolution integral explicitly, we have

$$T^{-1}[F(k_x, k_y)G(k_x, k_y)] = \int_{\infty}^{\infty} f(x, y)g(x' - x, y' - y)dx'dy' \tag{6.16}$$

where to keep the reader on their toes, we have flipped the sides and applied the inverse Fourier transform to both sides.

Transforms of derivatives The Fourier transform of the derivative of a function is given by

$$T\left[\frac{\partial f}{\partial x}\right] = ik_x F(k_x, k_y) \tag{6.17}$$

Cauchy-Schwarz Inequality While not technically a theorem of Fourier analysis, the Cauchy-Schwarz inequality is useful in the context of Fourier transforms. It states that for any two functions $f(x)$ and $g(x)$

$$\left|\int_{-\infty}^{\infty} f(x)g^*(x)dx\right|^2 \leq \int_{-\infty}^{\infty} |f(x)|^2 dx \int_{-\infty}^{\infty} |g(x)|^2 dx \tag{6.18}$$

where g^* is the complex conjugate of g.

6.2.2 Uncertainty principle and Space-Bandwidth Product

We shall see in later sections that the space-bandwidth product is related to the β paremeter. We shall use the uncertainty principle to prove that for small β it is not possible to reshape a beam to an arbitrary shape.

Of course the Heisenberg uncertainty principle is a statement about the uncertainty in the position and momentum of a particle. it states

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{6.19}$$

where Δx is the uncertainty in the position of the particle and Δp is the uncertainty in the momentum of the particle. By uncertainty we mean the standard deviation of the probability distribution of the position and momentum. Let us define the uncertainty in $f(x)$ and $F(k)$ as

$$\Delta_f = \sqrt{\frac{\int_{-\infty}^{\infty} x^2 |f(x)|^2 dx}{\int_{-\infty}^{\infty} |f(x)|^2 dx}} \tag{6.20}$$

and for Δ_F we simply replace x with k. Defining the space-bandywidth produce

$$SBP = \Delta_f \Delta_F \tag{6.21}$$

we shall state the uncertainty principle in terms of the space-bandwidth product as

$$SBP \geq \frac{1}{2} \tag{6.22}$$

See Eq (2.22) through (2.28) of Dickey and Holswade for the proof. It is a simple application of the Cauchy-Schwarz inequality. It may also be good to note the (perhaps intuitively obvious fct) that rescaling of $f(x) \to af(bx)$ does not change the space-bandwidth product.

Gaussian's minimize the space-bandwidth product We only have $SBP = 1/2$ for a (real) Gaussian function $f(x) = Ae^{-\alpha x^2}$ $\alpha \in \mathbb{R}$. See Lemma 2 in 2.2.2 of Dickey and Holswade for the proof.

Theorem of minimization The fnction $q(x)$ which minimizes the space-bandwidth product for a given $f(x)e^{iq(x)}$ is the one which makes the phase of $f(x)e^{iq(x)}$ constant. See Theorem 11 in 2.2.2 of Dickey and Holswade for the proof.

In two dimensions uncertainty is defined as

$$\Delta_f = \sqrt{\frac{\int_{-\infty}^{\infty}(x^2+y^2)|f(x,y)|^2dxdy}{\int_{-\infty}^{\infty}|f(x,y)|^2dxdy}} \tag{6.23}$$

and similarly for Δ_F. The space-bandwidth product is bounded below by unity

$$SBP \geq 1 \tag{6.24}$$

6.2.3 Cylindrical symmetry and Hankel transforms

Let's provide some results about cylindrical symmetry and Hankel transforms. A handy identity from the theory of Bessel functions is the identity

$$J_k(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(r\cos\theta - k\theta)}d\theta \tag{6.25}$$

which looks suspiciously similar to the integral representation of the Bessel functions given at and around Eq. 9.1.21 of Abramowitz and Stegun, though not quite the same.

Suppose we instead use polar coordinates

$$(x,y) = (r\cos\theta, r\sin\theta)$$

$$(k_x,k_y) = (\alpha\cos\phi, \alpha\sin\phi)$$

then Eq. 6.11 becomes

$$F(\alpha,\phi) = \int_0^{\infty}\int_{-\pi}^{\pi} f(r,\theta)e^{-i\alpha r\cos(\theta-\phi)}rdrd\theta \tag{6.26}$$

and supposeing that $f(r,\theta)$ is cylindrically symmetric, i.e. $f(r,\theta) = f(r)$, then we have

$$F(\alpha,\phi) = \int_0^{\infty}\int_{-\pi}^{\pi} f(r)e^{-i\alpha r\cos(\theta)}rdrd\theta \tag{6.27}$$

and using the identity above we have

$$F(\alpha) = 2\pi\int_0^{\infty} f(r)J_0(\alpha r)rdr \tag{6.28}$$

The function $F(\alpha)$ is called the Hankel transform of $f(r)$. The inverse Hankel transform is given by

$$f(r) = \int_0^{\infty} F(\alpha)J_0(\alpha r)\alpha d\alpha \tag{6.29}$$

6.3 Stationary Phase

The method of stationary phase is an approximation method for evaluating integrals with rapidly oscillating integrands. The method is based on the observation that the integral of a rapidly oscillating function is dominated by the contributions from the stationary points of the phase of the integrand.. It is relevant in the theory of dispersice wave propagation where it motivates the concept of groupvelocite and it ay be used to derive the geometrrical optics limit from Fresnel diffraction theory and to provide bounds for when geometrical optics are valid.

Suppose we have an integral of the form

$$H(\gamma) = \int_{-\infty}^{\infty} f(\xi)e^{i\gamma q(\xi)}d\xi \tag{6.30}$$

Because the integral will be dominated by locations where the phasevaries slowly or not at all we may approximate the integral by expanding the phase about the stationary points of the phase. We may expand the phase about the stationary points as

$$H(\gamma) \approx f(\xi_0) e^{i\gamma q(\xi_0)} \int_{-\infty}^{\infty} e^{i\gamma \frac{q''(\xi_0)}{2}(\xi-\xi_0)^2} d\xi \tag{6.31}$$

where ξ_0 is the stationary point of the phase, i.e. $q'(\xi_0) = 0$. The integral is a Gaussian integral and may be evaluated exactly. The result is

$$H(\gamma) \approx f(\xi_0) e^{i\mu\pi/4} e^{i\gamma q(\xi_0)} \sqrt{\frac{2\pi}{\gamma |q''(\xi_0)|}} \tag{6.32}$$

Where

$$\mu = \operatorname{sgn}\left(\left. \frac{d^2 q}{d\xi^2} \right|_{\xi_0} \right) \tag{6.33}$$

is the Maslov index. The Maslov index is a topological invariant of the phase function $q(\xi)$ and is equal to the number of times the phase function winds around the origin as ξ is varied from $-\infty$ to ∞. The Maslov index is also equal to the number of stationary points of the phase function.(githp copilot wrote that idk what it is its not mentioned at that spot in Dickey and Holswade, its related to the absolute value in the quare root, it's also on wikipedia, i'm gonna roll with it)

6.3.1 convergence of the stationary phase approximation

We shall see that using the loewest order term only in the stationary phase approximation give the geomteric optics approximation. In this case, β will serve as the large parameter in the phase o fht integrand. While it is not important for our purposes to have exact expressions for the higher order terms, it is important to know when the approximation is valid. Before our somewhat technicl discussion, lets sumarize what will be our main results. In the cotent of beam shaping will will have another parameters in our phase functions so our integrals will be of the form

$$H(x, \gamma) = \int_{-\infty}^{\infty} f(\xi) e^{i\gamma q(\xi, x)} d\xi \tag{6.34}$$

where ξ is a point on the aperture and x is a point at the focal plane. The function $q(\xi, x)$ will be proportional to the travel time from ξ to x. While both ξ and x will be 2D vectors we use the one dimensional case to simlif the discussion.

We will show that if the function $q(\xi, x)$ and $f(\xi)$ are analytic at the stationary point ξ_0 and the second derivative $\partial^2 q(\xi, x)/\partial \xi^2$ is nonzero at ξ_0 then the next order correction dies down like $1/\gamma^{3/2}$. This give the expansion of the form

$$H(x, \gamma) = \frac{A(x)}{\gamma^{1/2}} + \frac{B(x)}{\gamma^{3/2}} + \cdots \tag{6.35}$$

And the relative error between the first order term and the exact integral is of order $1/\gamma$. If $f(\xi)$ is real, then the functions A and B will have the same phase and the error of $|H(x, \gamma)|^2$ is $\mathcal{O}(1/\gamma^2)$.

The discussio proceeds with les well behaved functions resulting in slower convergence and such. i might get back to it later. skippnig the reset of 2.3.

6.4 Maxwell's equations

thi si ssection 2.4 of Dickey and Holswade. I'm not gonna write it all out. they indroduce the wave equation here.

6.5 Geometrical optics

I will skip section 2.5.1 about fermats principle for now, but it is important. Several examples are given in the text which may be helpful.

6.5.1 The eikonal equation

The laws of geometrical optics may be derived as a high-frequency approximations to the solutions of Maxwell's equations (ME). Before considering that limit, lets consider the high-frequency limit of the scalar wave equation, supposing whe have a solution $p(\mathbf{x}, \omega)$ to

$$\nabla^2 p - \frac{\omega^2}{c(\mathbf{x})^2} p = 0 \tag{6.36}$$

This is the time-harmonic wave equation also known as the reduced wave equation. we are interesting in the behavior for large ω, and in particular for olutions from point source. The theory ofGreen's functions tells us that the general high-frequency limit (not just for point sources!) can be built by integrating over many such point sources. For constant velocity (of what?) the point source solutions may be written as

$$p(\mathbf{x}, \omega) = A \frac{e^{i\omega r/c}}{r} \tag{6.37}$$

This solution has a radipldy varying phase and slowly varying amplitude (that is independent of ω). In the case of variable $c(\mathbf{x})$ we assume that while the amplitde may not be completely independent of ω, it is at least slowly varying. We may then write the solution as

$$p(\mathbf{x}, \omega) = A(\mathbf{x}) e^{i\omega\phi(\mathbf{x})} \tag{6.38}$$

this is only the first term in an asymptotic expansion in powers of $1/\omega$. The genera solution must include corredtion to the amplitude from the ω dependence.

$$\nabla p = i \left(\nabla A + A i\omega \nabla\phi \right) e^{i\omega\phi} \tag{6.39}$$

and

$$\nabla^2 p = i \left(\nabla^2 A + 2i\omega \nabla A \cdot \nabla\phi + i\omega A \nabla^2\phi - A\omega^2 (\nabla\phi)^2 \right) e^{i\omega\phi} \tag{6.40}$$

Substitution into Eq. 6.36 gives

$$\nabla\phi|^2 = 1/c(\mathbf{x})^2 \tag{6.41}$$

This is the eikonal equation. The next term in the asymptotic expansion is the transport equation (claims AI) and is given by

$$2\nabla A \cdot \nabla\phi + A\nabla^2\phi = \nabla \cdot (A^2 \nabla\phi) \tag{6.42}$$

The divergence form suggests that $A^2\nabla\phi$ is a is a flux of some conserved quantity. this will turn out to be the energy flux.

We mentioned that the gneral high-frequency approximation may be built from many point sources. If the field emenatess from two point sources the HF limit of the wave field looks like

$$p(\mathbf{x}, \omega) = A_1(\mathbf{x}) e^{i\omega\phi_1(\mathbf{x})} + A_2(\mathbf{x}) e^{i\omega\phi_2(\mathbf{x})} \tag{6.43}$$

Even in the case of a single point source, there maybe be locations wherehte HF limit consists of several terms from the same source, corresponding to diferent ray from the source to that point. the surfaces which separate the regions of space where the HF limit consists of different terms from the same source are called caustics or caustic surfaces.

6.5.2 The eikonal equation and Maxwell's equations

In the last section we derived the eikonal equation as the high-frequency limit of the scalar wave equation. We now consider the high-frequency limit of Maxwell's equations. We will see that the eikonal equation is the high-frequency limit of the wave equation for the electric field. this is natural considering that each component of the electric field satisfies the wave equation. I will skip several steps here. We assume a solutions to the wave equation of the form

$$\mathbf{E}(\mathbf{x}, \omega) = \mathbf{E_0}(\mathbf{x}) e^{i\omega\phi(\mathbf{x})} \tag{6.44}$$

and we shall arrive at the result

$$\mathbf{E_0}\left(\frac{1}{c(\mathbf{x})^2} - \nabla\phi\cdot\nabla\phi\right) = 0 \tag{6.45}$$

where the speed of light is given by

$$c(\mathbf{x}) = \frac{c_0^2}{\mu\epsilon} \tag{6.46}$$

which is the eikonal equation. The energy flux is given by the Poynting vector

$$\mathbf{S} = \frac{1}{2}\mathbf{E}\times\mathbf{H} \tag{6.47}$$

where $\mu(\mathbf{x})\mathbf{H} = \mathbf{B}$. Using that $\mathbf{E}$ and $\mathbf{H}$ are orthogonal to each other and also to ϕ we can conclude that the Poynting vector is in the direction of the gradient of the phase.

I will skip all of sections 2.5.4, 2.5.5, ad 2.5.6 of Dickey and Holswade for now. they cover the development of geometrical optics and fermat's princible with and without reflections.

6.6 Fourier Optics 1

6.6.1 Fresnel diffraction introduction

The Fresnel approximation is cocenerd with the wave field for $z > 0$ when an incoming wave field is incident on a planar aperture at $z = 0$. The approach can be outlined n three steps:

1. Write down an exact expression for the wave field at $z > 0$ in terms of the wave field at the aperture.
2. Use a paraxial approximation to simplify the exact expression assuming the observation point is close to the optical axis.
3. Compute the wave feld aware from the apertre using the first two steps with under the following assumption: that at the aperture the field is equal to the undistred field of the incoming way (modified by any optical element inside the aperture) and that the field is zero outside the aperture.

The first step is rogorous, the second step straightforward to justify. The third step is the hardest to justify, but is plausible provided the aperture is large compared to the wavelength.

Chapter 7

Optics2

7.1 Fourier optics Born and Wolf chapter 8

the text i was following by Dickey and Holswade is not as thorough. I will switch to Born and Wolf chapter 8, which is likely to be far more thorough than i need, but is the seminal text. I am using the seventh edition, printed in 1999.

7.1.1 The Huygens-Fresnel principle

From section 8.2 of Born and Wolf. Let us consider a spherial wavefront S proceeding from a point source P_0 and let Q be a point on S, and let our observation point be P. (Figure 8.1 of Born and Wolf is helpful here). The field at Q (r_0 from P_0) is Ae^{ikr_0}/r_0 where A is the amplitude of the field at a distance of unity. The contribution fom the element of the wavefront dS at Q to the field at P is denoted $dU(P)$ and is given by

$$\frac{Ae^{ikr_0}}{r_0}K(\chi)\frac{e^{iks}}{s}dS \tag{7.1}$$

where s is the distance from Q to P and $K(\chi)$ is the obliquity or inclination factor and χ is the angle of diffraction. $K(\chi)$ describes the reduction in amplitude due to the fact that the wavefront is not normal to the line QP. Following Fresnel, we shall assume that $K(\chi)$ is given by a function which is maximal for $\chi = 0$ and decreases monotonically as χ increases until it is zero at $\chi = \pi/2$. One obvious choice is

$$K(\chi) = \cos\chi \tag{7.2}$$

however, Born and Wolf do not actually choose a functional form for K at this time. We also assume, following Fresnel, that only those portion S' visible from P contributes to the field at P as there may be obstacles. The field at P is then given by integrating over S

$$U(P) = \frac{Ae^{ikr_0}}{r_0}\int_{S'} K(\chi)\frac{e^{iks}}{s}dS \tag{7.3}$$

This seems to be the core idea of the Huygens-Fresnel principle. The evalutation of Eq. 7.3 is not trivial. To proceed we shall follow the *zone construction* of Fresnel. This is done by constructing spheres centered at P (not the source P_0) with radii

$$b, \quad b+\frac{\lambda}{2}, \quad b+\frac{2\lambda}{2}, \quad \cdots, \quad b+\frac{j\lambda}{2}, \quad \cdots$$

where b is the distance from P to the nearest point on S and j is an integer. In other words, if C is the point intersection of S and P_0P, then $b = CP$. These spheres divide S into zome Z_j.

We assume that bot r_0 and b are large compared to λ. This mean that within a particular zone the value of the inclination factor $K(\chi)$ is approximately constant. We may then write the contribution of the jth sone

to $U(P)$ as

$$U_j(P) = 2\pi \frac{Ae^{ikr_0}}{r_0 + b} K_j \int_{b+(j-1)\lambda/2}^{b+j\lambda/2} e^{iks} ds \tag{7.4}$$

where we've done some variable algebra and a change of integration variable. The integral is easily evaluated and we have

$$U_j(P) = \frac{2\pi A}{ik} \frac{e^{ik(r_0+b)}}{r_0 + b} K_j e^{ik\lambda/2} \left(1 - e^{-ik\lambda/2}\right) \tag{7.5}$$

but since $k\lambda = 2\pi$ the last two factors become $2(-1)^j$.

$$U_j(P) = 2i\lambda \frac{e^{ik(r_0+b)}}{r_0 + b} K_j (-1)^{j+1} \tag{7.6}$$

The total field at P is then given by summing over all zones

$$U(P) = 2i\lambda \frac{e^{ik(r_0+b)}}{r_0 + b} \sum_{j=1}^{n} (-1)^{j+1} K_j \tag{7.7}$$

Some clever math is done in Eqs. 8.2.5 to 8.2.9 of Born and Wolf to show that the sum may be written as

$$\begin{aligned} \Sigma &= \frac{K_1}{2} + \frac{K_n}{2} \text{if } n \text{ is even} \\ \Sigma &= \frac{K_1}{2} - \frac{K_n}{2} \text{if } n \text{ is odd} \end{aligned} \tag{7.8}$$

The clever math means that Eq. 7.8 is valid unless varius terms in the series switch sign enough to cause appreciable error. Ignoreing that situation, we have

$$U(P) = i\lambda (K_1 \pm K_n) \frac{e^{ik(r_0+b)}}{r_0 + b} \tag{7.9}$$

using Eq. 7.7 we can express the sum as

$$U(P) = \tfrac{1}{2}[U_1(P) + U_n(P)] \tag{7.10}$$

but as we already mentioned, the last zone will have a vanishing contribution and so we have

$$U(P) = i\lambda K_1 \frac{e^{ik(r_0+b)}}{r_0 + b} = \tfrac{1}{2} U_1(P) \tag{7.11}$$

demonstrating that the totae field at P is equal to hlf the disturbance of the first zone, a non-obvious result. When further see that Eq. 7.11 provides the correct expression for a spherical wave if we have $i\lambda K_1 = 1$ and hence that

$$K_1 = \frac{e^{-i\pi/2}}{\lambda}. \tag{7.12}$$

(which i didn't expect). The quarter phase in Eq. 7.12 may be accounted for by assuming that th seondary waves oscillate a quatryer of a period out of phase with the primary wave while the amplitude factor is accounted for by assuming that the secondary waves have an amplitude a factor of λ smaller than the primary wave. We can thus conclude that with those assumptions the Huygens-Fresnelprinciple produces the correct expression for a spherical wave. However, these additional assumptions are not required and a better justification for the factor of Eq. 7.12 is given in later secions.

Still following Fresnel, we can use Eq.7.6 to make predictions about the field produces when obstructions are present. Non trivial results are immediately apparent, such as if all but one half of the first zone is covered, the field at P is the same as if no screen were present. If the screen covers all but the entire first zone, the field is twice as large (and so the intensity four ties larger) as if no screen were present.When only zone 1 and zone 2 are not covered, the field at P is nearly zero since the terms of Eq. 7.6 nearly cancel (since K is slowly varying). Even the spot of Arago is predicted successfully Fresnel theory.

7.1.2 The Kirchhoff diffraction theory

From section 8.3 of Born and Wolf.

The Kirchhoff integral theorem

Kirchoff showd that the Huygens-Fresnel principle may be derived from an integral thorrem which expresses the solution of the homogeneous wave equation -at an arbitrary point-in terms of the values of the field and its first derivatis at all points on an arbitrary closed surface surrounding the point P.

We begin by considering a monochromatics scalar wave

$$V(x,y,z,t) = U(x,y,z)e^{-i\omega t} \tag{7.13}$$

where the spatial part U satisfies (in a vacuum) the time-independent wave equation

$$(\nabla^2 + k^2)U = 0 \tag{7.14}$$

where $k = \omega/c$ is the wave number. We shall assume that U and its first and second order partial derivatives are continuous within a volume v bounded by a closed surface S. We shall also assume that U and its first- and second-order partial are continuous on and within this surface. If U' is any other function satisfying the same conditions, then we may then apply Gauss's theorem to obtain

$$\int_v (U\nabla^2 U' - U'\nabla^2 U)dv = -\oint_S \left(U\frac{\partial U'}{\partial n} - U'\frac{\partial U}{\partial n}\right) dS \tag{7.15}$$

where the differentiation is along the *inward* normal to S (which is the opposite of the normal typically used in Greens theorem). If U' additionally obeys the Helmholtz equation then the left hand side of the above equation vanishes. Let us make the ansatz $U'(x,y,z) = e^{iks}/s$ where s is the distance from the point (x,y,z) to the point P (the observation point). However, at the orgin $s = 0$ and so we must exclude the origin from the volume v. Therefore we surround P by a small sphere S_0 of radius ϵ and let the integration be over the volume bounded by S and S_0. We then have

$$\begin{aligned}\oint_S \left[U\frac{\partial}{\partial n}\left(\frac{e^{iks}}{s}\right) - \frac{e^{iks}}{s}\frac{\partial U}{\partial n}\right] dS &= -\oint_{S'} \left[U\frac{e^{iks}}{s}\left(ik - \frac{1}{s}\right) - \frac{e^{iks}}{s}\frac{\partial U}{\partial n}\right] dS \\ &= -\oint_\Omega \left[U\frac{e^{ik\varepsilon}}{\varepsilon^2}\left(ik - \frac{1}{\varepsilon}\right) - \frac{e^{ik\varepsilon}}{\varepsilon}\frac{\partial U}{\partial s}\right]\varepsilon^2 d\Omega\end{aligned} \tag{7.16}$$

The LHS does not depend on ε and so we may take the limit $\varepsilon \to 0$. The RHS is then equal to $4\pi U(P)$. We thus have

$$U(P) = \frac{1}{4\pi}\oint_S \left[U\frac{\partial}{\partial n}\left(\frac{e^{iks}}{s}\right) - \frac{e^{iks}}{s}\frac{\partial U}{\partial n}\right] dS \tag{7.17}$$

This is the Kirchhoff integral theorem. It expresses the value of the field at P in terms of the values of the field and its normal derivative on the surface S. Born and Wolf proceed to generalize to the case of non-monochromatic fields, but i shal skip that for now.

Let's take a moment to compare this expression for the Kirchoff integral from Born and Wolf wth the treatment from Jackson (third edition, section 10.5)

Kirchoff's diffraction theory

(section 8.3.2 of Born and Wolf). While the integral theorem embidies the spirit of te Huygens-Fresnel principle, the laws governingthe contributions from the various elements of the surface S are more complicated Fresnel assumed. Kirchoff showed, however that in many cases the theorem may be reduced ito an approximate bute simpler form which is esentially quivalent to the formulat of Fresnel, but which additionally provides an explicit expression for the amplitude factor $K(\chi)$, which remained undetermined in Fresnel's theory.

Consider a monochromatic wave, from a source P_0propagated through an opening in a creen, and let P be the point where the field is to be determined (Figure 8.3 of Born and Wolf). We shall assume that

the opening is small compared to the distance from the source to the opening, that the opening is large compared to the wavelength, and that the opening is small compared to the distance from the opening to the observation point.

We choose our surface of inegration to be composed of three regions, the opening $\mathcal{A}$, a portion $\mathcal{B}$ of the dark side of the screen, and portion $\mathcal{C}$ of the surface of a sphere of radius R centered at P. The surface integral is then over the closed surface formed by the union of $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$.

the difficulty is that to apply Kirchoff's integral theorem Eq. 7.17 we must know the values of the field and its normal derivative on the surface. However, we never (or rarely) know the exact field or its derivatives on $\mathcal{A}$, $\mathcal{B}$, or $\mathcal{C}$. However, we may make some assumptions. On $\mathcal{A}$ we assume that the field and its derivative is equal to those of the incoming wave (potentially modified by any optical element inside the aperture? similar to Dickey?), an assumption which should hold extremely well everywhere but in the immediate vicinity of the edge of the aperture. On $\mathcal{B}$ we assume that the field and its derivative is zero.

The situation for $\mathcal{C}$ is more complicated it would seem, but here we just straight up cheat and take advantage of the fact that the source has only been radiating for a finite amount of time. then we simply make the radius of the sphere R large enough that the field and its derivative are zero on $\mathcal{C}$ because of the finite speed of light.

Now, using the fact that

$$U^{(i)} = \frac{Ae^{ikr}}{r}, \quad \frac{\partial U^{(i)}}{\partial n} = \frac{Ae^{ikr}}{r}\left(ik - \frac{1}{r}\right)\cos(n, r) \tag{7.18}$$

where $\cos(n, r)$ is the cosine of the angle angle between the normal to the surface and the radius vector from the source to the point on the surface, we may evaluate the integral in Eq. 7.17 and neglect the terms $1/r$ and $1/s$ in the field derivatives to obtain

$$U(P) = -\frac{iA}{2\lambda}\int_{\mathcal{A}} \frac{e^{ik(r+s)}}{rs}[\cos(n, r) - \cos(n, s)]dS \tag{7.19}$$

For a point source, another clever geometric argument is made to show that the above integral is equal to

$$U(P) = -\frac{iA}{2\lambda}\int_{\mathcal{W}} \frac{e^{ik(r_0+s)}}{r_0 s}[1 - \cos(\chi)]dS \tag{7.20}$$

where $\chi - (r_0, s)$ is π minus the (nearly straight) angle at a point Q on the portion of a spherical surface $\mathcal{W}$ centered at the source P_0 and filling the aperture, which some edge bits we neglect. See Figure 8.4 of Born and Wolf, which is helpful.

Comparison of Eq. 7.20 with Eq. 7.3 shows that the amplitude factor $K(\chi)$ is given by

$$K(\chi) = -\frac{i}{2\lambda}[1 + \cos(\chi)] \tag{7.21}$$

which for $\chi = 0$ gives the correct value of $K(\chi)$ from Fresnels theory, but does not vanish at $\chi = \pi/2$ as Fresnel assumed.

Poincare showed that the assumed values of the field in the plane of the aperture are not reproduced by the Kirchhoff integral theorem. This is not suprising given the assumptions about the relative sizes of the aperture and the distance from the aperture to the observation point. However, the field in the plane of the aperture is not of interest in most cases. but i think we should be able to figure it out if we want to, i hope it is addressed.

7.1.3 Kirchoff theory: Fraunhofer and Fresnel diffraction

Initial approximations

Let us return to the main version of the Fresnel-Kirchoff diffraction formula

$$U(P) = -\frac{iA}{2\lambda}\int_{\mathcal{A}} \frac{e^{ik(r+s)}}{rs}[\cos(n, r) - \cos(n, s)]dS \tag{7.22}$$

and consider the integrand as the element dS explores the domain of integration. It is clear that the quantities r and s will vary by many wavelengths assuming the aperture is large compared to the wavelength; therefore the phase factor will oscillate rapidly. in constrast the cosine terms will vary slowly. Given our assumptions about the aperture, we may then approximate the cosine terms by replacing them with $2\cos\delta$ where is the angle beteween P_0P and the normal to the aperture. We may similarly replace $1/rs$ with $1/r's'$ where r' and s' are the distances from P_0 and P to the aperture. We then have

$$U(P) \approx -\frac{iA}{\lambda}\frac{\cos\delta}{r's'}\int_{\mathcal{A}} e^{ik(r+s)}dS. \tag{7.23}$$

Our coordinate system places the origin within the aperture and the x and y axes in the plane of the aperture. Setting (x_0, y_0, z_0) to be the coordinates of the source, (x, y, z) to be the coordinates of the observation point, and (ξ, η) to be the coordinates of the point on the aperture, we have

$$\begin{aligned} r^2 &= (x_0-\xi)^2 + (y_0-\eta)^2 + z_0^2 \\ s^2 &= (x-\xi)^2 + (y-\eta)^2 + z^2 \\ r'^2 &= x_0^2 + y_0^2 + z_0^2 \\ s'^2 &= x^2 + y^2 + z^2 \end{aligned} \tag{7.24}$$

Hence

$$\begin{aligned} r^2 &= r'^2 - 2x_0\xi - 2y_0\eta + 2 + \eta^2 \\ s^2 &= s'^2 - 2x\xi - 2y\eta + \xi^2 + \eta^2 \end{aligned} \tag{7.25}$$

The Fraunhofer and Fresnel approximations

Due to our assuptions of the aperture being small, we may expand both r and s in powers of ξ/r' and η/r' and ξ/s' and η/s' respectively up to second order. We then have

$$\begin{aligned} r &\sim r' - \frac{x_0\xi + y_0\eta}{r'} + \frac{\xi^2+\eta^2}{2r'} - \frac{(x_0\xi + y_0\eta)^2}{2r'^3} + \cdots, \\ s &\sim s' - \frac{x\xi + y\eta}{s'} + \frac{\xi^2+\eta^2}{2s'} - \frac{(x\xi + y\eta)^2}{2s'^3} + \cdots. \end{aligned} \tag{7.26}$$

We may rewrite 7.23 using 7.26 in terms of a function $f(\xi, \eta)$ as

$$U(P) \approx -\frac{iA}{\lambda}\frac{\cos\delta e^{ik(r'+s')}}{r's'}\int_{\mathcal{A}} e^{-ikf(\xi,\eta)}d\xi d\eta. \tag{7.27}$$

where

$$f(\xi,\eta) = -\frac{x_0\xi + y_0\eta}{r'} - \frac{x\xi + y\eta}{s'} + \frac{\xi^2+\eta^2}{2r'} + \frac{\xi^2+\eta^2}{2s'} + \frac{(x_0\xi + y_0\eta)^2}{2r'^3} - \frac{(x\xi + y\eta)^2}{2s'^3} + \cdots. \tag{7.28}$$

let us define

$$\begin{aligned} l_0 &= -\frac{x_0}{r'}, \qquad l = \frac{x}{s'} \\ m_0 &= -\frac{y_0}{r'}, \qquad m = \frac{y}{s'} \end{aligned} \tag{7.29}$$

Which the text describes (l_0, m_0) and (l, m) as the direction cosines of the rays from the source to the observation point and from the aperture to the observation point respectively. but i do not understand this. looking at the text, the x cosine from P_0 would be ξ/r' and the x cosine from P would be ξ/s', but i do not see how that is the same as the above. let me think this through. letting θ_0 be the angle between the ray from OP_0 where O is the origin (in center of the aperture) and the z axis (which is normal to the aperture), we have $l_0 = \sin\theta_{0x}$. similarly, letting θ be the angle between the ray from OP and the z axis, we have $l = \sin\theta_x$. I've just been assuming $x_0 = y_0 = 0$ and $x = y = 0$ for some reason. but i still find it odd that they would identify these angles as sines of a small angle, rather than cosies of a large angle (which is what i would have done).

With those comments, we may rewrite Eq. 7.28 as

$$f(\xi,\eta) = (l_0 - l)\xi + (m_0 - m)\eta + \frac{1}{2}\left[\left(\frac{1}{r'} - \frac{1}{s'}\right)(\xi^2+\eta^2) - \frac{(l_0\xi + m_0\eta)^2}{r'} - \frac{(l\xi + m\eta)^2}{s'}\right]\cdots . \quad (7.30)$$

Therefore one an identify the condition needed to neglect the higher order terms by checking if

$$\frac{k}{2}\left|\left(\frac{1}{r'} - \frac{1}{s'}\right)(\xi^2+\eta^2) - \frac{(l_0\xi + m_0\eta)^2}{r'} - \frac{(l\xi + m\eta)^2}{s'}\right| \ll 2\pi \quad (7.31)$$

which is satisfied when

$$|r'| \gg \frac{(\xi^2+\eta^2)_{max}}{\lambda} \quad \text{and} \quad |s'| \gg \frac{(\xi^2+\eta^2)_{max}}{\lambda} \quad (7.32)$$

or if

$$l_0^2, m_0^2, l^2, m^2 \ll \frac{|r'|\lambda}{(\xi^2+\eta^2)_{max}} \quad (7.33)$$

7.1.4 Fraunhofer diffraction

In this section we cover the second half of section 8.3.3 of Born and Wolf. As a reminder, we are still considering point sources and our starting integral is Eq. 7.27, which one can see explicitly has pherical wavefronts. There is a lengthy discussion belw Eq 8.3.34 which i will skip for now, partially because i do not understand it. i kinda feel called out by github copilot for writing that last sentence for me.

In the Fraunhofer approximation, the four quantities l_0, m_0, l, m are all small enough that we drop the quadratic terms in Eq. 7.30 and they only enter the integral in the combinations

$$p = l_0 - l, \quad q = m_0 - m.$$

We may then write a Fraunhofer approximation to Eq. 7.27 as

$$U(P) = C\int_{\mathcal{A}} e^{-ik(p\xi+q\eta)}d\xi d\eta. \quad (7.34)$$

Where C is the constant prefactor in Eq. 7.27. It is often convenient to write Eq. 7.34 in terms of the aperture function $A(\xi,\eta)$, which is defined as being constant within the aperture and zero outside the aperture. We then have

$$U(p,q) = \int_{-\infty}^{\infty} A(\xi,\eta)e^{-ik(p\xi+q\eta)}d\xi d\eta. \quad (7.35)$$

The integration now extend over all space, and its true nature as a Fourier transform is clear. The aperture function is also called the pupil function.

Chapter 8

Fourier Optics, following Goodman

These will be my notes on fourier optics following the presentation of Joseph Goodman's widely used text "Introduction to Fourier Optics". I will be using the fourth edition, from 2017.

8.1 Introduction and chapter 2

The text opens (in hapter two) with a discussion of linear systems, which is an emphasis that hasn't beenas clear in other treatments of the text, thogh obvious lies at the heart of optical physics. The text also immediatly appeals to the language of linear systems theory, which is a bit of a barrier to entry. But basically there is a focus on reducing the problem so how a system responds to what Goodman refers to as 'elementary' stimuli. One can then use the linearity of the system to build up a response to more complicated stimuli.

8.1.1 Fourier 2d

We repeat Goodmans definitions of Fourier analysis in two dimensional systems to estblish the notation. we defife the Fourier transform of a function $g(x, y)$ as $\mathcal{F}\{g\}$ to be

$$\mathcal{F}\{g\} = \int_{-\infty}^{\infty} g(x, y) e^{-i2\pi(f_x x + f_y y)} dx dy \tag{8.1}$$

where apparently Goodman uses j for i because he is a monster (copilot wrote that monster comment Liouville), but i will probably use i? i haven't decided lol. The inverse Fourier transform $\mathcal{F}^{-1}\{G\}$ is then defined as

$$\mathcal{F}^{-1}\{G\} = \int_{-\infty}^{\infty} G(f_x, f_y) e^{i2\pi(f_x x + f_y y)} df_x df_y \tag{8.2}$$

where $G(f_x, f_y)$ is the Fourier transform of $g(x, y)$. Notice that this convention includes the 2π in the exponent of the forward transform, which is not always the case, but its objectively the best convention. For 8.2 and 8.1 to be meaningfull, we must have some conditions on g

1. g must be absolutely integrable over the entire $x - y$ plane
2. g must be continuous everywhere except at a finite number of finite discontinuities and a finites number of extrema in any finite rectangle.
3. g must have no infinite discontinuities.

These conditions are a common set to be used, but one can weaken of strengthen them as needed, apparently. A nice quote from Bracewell is given here "physical possibility is a valid sufficient condition for the existence of a Fourier transform". But certain functions are weird, like this limit form of the Dirac delta function

$$\delta(x, y) = \lim_{N \to \infty} N^2 \exp\left[-N^2 \pi (x^2 + y^2)\right] \tag{8.3}$$

but we can define generalized transforms for the functions in the defining sequence, and then take the limit of the transforms. for the above, we have

$$\mathcal{F}\{\delta\} = \lim_{N\to\infty} \mathcal{F}\{N^2 \exp\left[-N^2\pi(f_x^2+f_y^2)\right]\} = \lim_{N\to\infty} \exp\left[-\frac{\pi(f_x^2+f_y^2)}{N^2}\right] = 1. \tag{8.4}$$

Goodman notes that it is useful to think of Fourier tranforms as decomposiiton into elemenary functions of the form

$$\exp\left[-i2\pi(f_x x + f_y y)\right] \tag{8.5}$$

This form shows how the elemtary functions are plane waves in the $x-y$ plane, with the dirction of the wave ggiven by the angle

$$\theta = \tan^{-1}\left(\frac{f_y}{f_x}\right) \tag{8.6}$$

and the wavelength given by

$$\lambda = \frac{1}{\sqrt{f_x^2+f_y^2}} \tag{8.7}$$

8.1.2 theorem and identities

Several theorems are given, i wont even list the linearity conditions here, but also have the Similarity theorem which is

$$\mathcal{F}\{g(ax,by)\} = \frac{1}{|ab|}\mathcal{F}\{g(x,y)\} \tag{8.8}$$

which shows how scaling behaves recipricolly in real and Fourier space. The shift theorem is:

$$\mathcal{F}\{g(x-x_0,y-y_0)\} = e^{-i2\pi(f_x x_0 + f_y y_0)}\mathcal{F}\{g(x,y)\} \tag{8.9}$$

which shows how a shift in real space is a phase shift in Fourier space (i'm not sure i appreciated that before). Raylieghs or Parsevals theorem is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |g(x,y)|^2 dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |G(f_x,f_y)|^2 df_x df_y \tag{8.10}$$

which is a statement of conservation of energy. The convolution theorem is

$$\mathcal{F}\{g \otimes h\} = G(f_x,f_y)H(f_x,f_y) \tag{8.11}$$

which is a statement of the fact that convolution in real space is multiplication in Fourier space. The autocorrelation theorem is

$$\mathcal{F}\{\int_{\infty} g(\xi,\eta)g^*(\xi-x,\eta-y)dxdy\} = |G(f_x,f_y)|^2 \tag{8.12}$$

which is a special case of the convolution theorem for the case where $h = g^*(x,y)$. The differentiation theorem is

$$\mathcal{F}\left\{\frac{\partial^n g}{\partial x^n}\right\} = (i2\pi f_x)^n G(f_x,f_y)??? \tag{8.13}$$

idk this. The rotation theorem for $\mathcal{F}\{g(r,\theta)\} = G(\rho,\phi)$ is that $g(r,\theta+\theta_0)$ is a rotation of $G(\rho,\phi+\theta_0)$ by an identical angle θ_0. in rectangular coordinates, this is

$$\mathcal{F}\{g(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)\} = e^{-i2\pi f_x x_0}G(f_x\cos\theta + f_y\sin\theta, f_y\cos\theta - f_x\sin\theta) \tag{8.14}$$

The shear theorem is

$$\mathcal{F}\{g(x+ay,y)\} = e^{-i2\pi a f_y y}G(f_x,f_y) \tag{8.15}$$

and similarly for the other shear. Goodman also points out that sperability is a useful property, where if $g(x,y) = f(x)h(y)$ then $G(f_x,f_y) = F(f_x)H(f_y)$.

8.1.3 Local spatial frequency

Gonsider the function

$$g(x,y) = a(x,y)\exp\left[i\phi(x,y)\right] \tag{8.16}$$

where $a(x,y)$ is the slowly varying amplitude and $\phi(x,y)$ is the phase. The local spatial frequency is defined as the gradient of the phase

$$\vec{f}^{(l)} = \frac{1}{2\pi}\left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}\right) = \frac{1}{2\pi}\nabla\phi \tag{8.17}$$

for the case $g(x,y) = \exp\left[i2\pi(f_x x + f_y y)\right]$, we have $\vec{f}^{(l)} = (f_x, f_y)$, which is the local spatial frequency. The local spatial frequency is a vector, and the magnitude of the vector is the local spatial frequency magnitude. in this simple case its exactly as we would naively guess. Next we consider a space limited function $g(x,y)$, which is zero outside of some region

$$g(x,y) = \exp\left[i\pi\beta(x^2+y^2)\right]\text{rect}\left[\frac{x}{L_x}\right]\text{rect}\left[\frac{y}{L_y}\right] \tag{8.18}$$

where rect$[x]$ is the rectangular function, which is 1 for $|x| < 1/2$ and zero otherwise. The local spatial frequencies are then

$$f_x^{(l)} = \beta\, x\,\text{rect}\left[\frac{x}{L_x}\right], \quad f_y^{(l)} = \beta\, y\,\text{rect}\left[\frac{y}{L_y}\right] \tag{8.19}$$

This time the local spatial frequency is not constant, but varies linearly with position, and is confined to a rectangle in the $x-y$ plane.

It is tempting to think tht since the local spatial frequencies are bounded to a rectangle, that the Fourier transform of $g(x,y)$ will be confined to a rectangle in Fourier space. However, this is not the case! The shape of the spectrum depends fundamentally on the product $\frac{L_x L_y}{4}\beta$ which will be called the Fresnel number. We shall see that for Fresnel numbers greater than unity, the local spatial frequency distribution can yield good results about the shape nd extent of the spectrum, but for Fresnel numbers less than unity, the local spatial frequency distribution is not a good approximation to the spectrum.

The Fourier transform of 8.18 is , according to the definition,

$$G(f_x,f_y) = \int_{-L_x/2}^{L_x/2}\int_{-L_x/2}^{L_y/2}\exp\left[i\pi\beta(x^2+y^2)\right]e^{-i2\pi(f_x x + f_y y)}dxdy \tag{8.20}$$

blah blah blah, i'm not going to do the math here, i dont think the result is mporantant yet either. but the point is that the spectrum is not confined to a rectangle in Fourier space, but rather is a function of the Fresnel number.

8.1.4 chapter 2

theres a whole thing about the Wigner distribution function, which is a way to represent the local spatial frequency distribution. i've never heard of it before. i will ignore it for now. then theres a section about linear systems, which brings up impulse response AKA point spread functions.

Goodman brings up invariant linear systems, which are systems where the impulse response $h(t;\tau)$ is time invariant, that is, the response at τ following an impulse at t depends only on $\tau - t$. This is typically how electrical circuits behave. in optics, however, the impulse response is not time invariant, but rather space invariant. That is, the response at x, y following an impulse at x_0, y_0 depends only on $x - x_0$ and $y - y_0$. This is because the speed of light is finite.

That leads to a discussion on the utility of transfer functions, which are how the system behaves in the frequency domain and arehelpful for simplifying the math.

Then follows a lengthy section on sampling, and reconstructtion of a signal from its samples. This is a very important topic, but i will skip it for now. But the Whittaker-Shannon sampling theorem is important. as is the concept of the Nyquist frequency. they also discuss he space badwidth product.

Then section 2.5 is about the discrete Fourier transform, which is a way to numerically calculate the Fourier transform of a discrete set of data. This is important, but i will skip it for now. 2.6 covers the projection slice theorem, which i think is intuitively obvious. section 2.7 very breifly covers phase retrieval.

8.2 chapter 3, Scalar diffraction theory

This chapter is about the scalar diffraction theory, which is the theory of diffraction of scalar waves. This is a good approximation for light when the wavelength is much smaller than the smallest dimension of the aperture. This is a good approximation for most optical systems.

I will skip the first four sections, since they are about the Kirchoff integral theorem, which i have already covered in the optics2 notes. Also the treatment of Born and Wolf is much more clearly written. most because of the figures.

8.2.1 3.5 The Rayliegh Sommerfeld diffraction theory

The assumptions of Kirchoff in deriving the Kirchoff integral theorem include that the field and its dervitive are simultaneously zero for all points on the screen. This mathematically means the field must be zero everywhere for an analytic function, a basic theorem of complex analysis. Additionally, the boundary conditions at the edge of the aperture mean the field is not identical to the n perturbed field. However, the criticism that motivate Sommerfeld were those raised by Poincare, namely that the theory itself is not self consistent, as onemoves the observation point closer to the aperture, the field diverges from its assumed value. but whevver its fine.

Consider the Kirchoff integral theorem as given by Goodman (3-30):

$$U(P_0) = \frac{1}{4\pi} \int_{\mathcal{A}} \left[G \frac{\partial U(P)}{\partial n} - U(P) \frac{\partial G}{\partial n} \right] dS \tag{8.21}$$

where G is the appropriate greens function (remember i have notes on greens functions in the math notes). Sommerfeld pointed out that with a sitable choice of the Green's function, one of the terms in the integrand can be eleminated. Via he method of images, one can show that the Greens functions which kill the first and second term in the integrand are, respectively

$$G_- = \frac{e^{ikr}}{r} - \frac{e^{ikr'}}{r'} \tag{8.22}$$

and

$$G_+ = \frac{e^{ikr}}{r} + \frac{e^{ikr'}}{r'} \tag{8.23}$$

where r is the distance from P_0 to P on the apertureand r' is the distance from the fictitious image point P_0' to the point P on the aperture. clearly G_- and G_+ are the Greens functions which kill off the first and second terms in the integrand of 8.21 respectively. doing some math yields the first Sommerfeld solution for the field at P_0 as

$$U_I(P_0) = -\frac{1}{2\pi} \int_{\mathcal{A}} U(P) \left[ik - \frac{1}{r} \right] \frac{e^{ikr}}{r} \cos(\mathbf{n}, \mathbf{r}) dS \tag{8.24}$$

where $\cos(\mathbf{n}, \mathbf{r})$ is the cosine of the angle between the normal to the aperture and the distance vector from the source to the point on the aperture. This is the first Sommerfeld solution in its full form (3-35 in Goodman).

---

this can be written as a convolution:

$$U_I(x, y, z) = h(x, y, z) * U(x, y, 0) \tag{8.25}$$

where $h(x, y, z)$ is the impulse response of the system which we may write as

$$h(x, y, z) = \frac{z}{2\pi r} \left[ik - \frac{1}{r} \right] \frac{e^{ikr}}{r} \tag{8.26}$$

where z/r is the relevant cosine term.

We could find a solution $U_{II}(P_0)$ by using the other Greens function, but Dirichlet boundary conditions are generally easier and more intuitive to work with, so we will use the first Sommerfeld solution. The second

Sommerfeld solution is given by equations 3-39, 3-40 in Goodman. They are derived using the same method of images, but with the Greens function G_+. It is given by

$$U_{II}(P_0) = \frac{1}{2\pi}\int_{\mathcal{A}} \frac{\partial U(P)}{\partial n}\frac{e^{ikr}}{r}dS \tag{8.27}$$

Rayliegh-Sommerfeld diffraction formula

The full expression for the first Sommerfeld solution is is given by eq. (8.24), but we now simplify t by assuming the far field, which kills the $1/r^2$ term in the integrand. This gives us:

$$U_I(P_0) = \frac{1}{i\lambda}\int_{\mathcal{A}} U(P)\frac{e^{ikr}}{r}\cos(\mathbf{n},\mathbf{r})dS \tag{8.28}$$

At this point in the derivations of the Sommerfeld solutions we have not specified the form of $U(P)$, but we will do so now. We will assume that the field on the aperture is given by a spherical wave, which is a good approximation for the case of a point source at a large distance from the aperture. The field on the aperture is then given by:

$$U(P) = A\frac{e^{ikr_{21}}}{r_{21}} \tag{8.29}$$

Then eq. (8.28) becomes

$$U_I(P_0) = \frac{A}{i\lambda}\int_{\mathcal{A}} \frac{e^{ik(r_{21}+r_{01})}}{r_{21}r_{01}}\cos(\mathbf{n},\mathbf{r}_{01})dS \tag{8.30}$$

eq. (8.27) reuires computeing

$$\begin{aligned}\frac{\partial U(P)}{A\partial n} &= ik\frac{\cos(\mathbf{n},\mathbf{r}_{21})}{r_{21}}e^{ikr_{21}} - \frac{\cos(\mathbf{n},\mathbf{r}_{21})}{r_{21}^2}e^{ikr_{21}} \\ &= (ik - \frac{1}{r_{21}})\frac{\cos(\mathbf{n},\mathbf{r}_{21})}{r_{21}}e^{ikr_{21}}\end{aligned} \tag{8.31}$$

We kill the second term by assuming large distance from the source and we thus get Then eq. (8.27) becomes

$$U_II(P_0) = -\frac{A}{i\lambda}\int_{\mathcal{A}} \frac{e^{ik(r_{21}+r_01)}}{r_{21}r_01}\cos(\mathbf{n},\mathbf{r}_{21})dS \tag{8.32}$$

where the difference between the two is the overall sign and the cosine term. The cosine terms is generally nearly -1 so they are quite similar.

8.2.2 3.6 Kirchoff vs Sommerfeld

By the way, so far we have been assuming a spherical wave, name that on the aperture the field U is given by

$$U(P) = A\frac{e^{ikr_0}}{r_0} \tag{8.33}$$

where r_0 is the distance from the source to the point on the aperture.Then, For the three case of the Kirchoff theory, Sommerfeld solution I and Sommerfeld solution II, the field at P_0 is given by

$$U(P_0) = \frac{A}{i\lambda}\int_{\mathcal{A}} \frac{e^{ik(r+r_0)}}{rr_0}\psi dS \tag{8.34}$$

where the obliquity factor ψ is given in each case by

$$\psi = \begin{cases}\frac{1}{2}[\cos(\mathbf{n},\mathbf{r}) - \cos(\mathbf{n},\mathbf{r_0})] & \text{Kirchoff} \\ \cos(\mathbf{n},\mathbf{r}) & \text{Sommerfeld I} \\ -\cos(\mathbf{n},\mathbf{r_0}) & \text{Sommerfeld II}\end{cases} \tag{8.35}$$

and we can see that the Sommerfeld solutions are the same as the Kirchoff solution except for the obliquity factor. In fact, the Kirchhoff solution is the average of the two Sommerfeld solutions. For the case of an infinitely distance point sorce, things simplify even more. Goodman notes that the Kirchoff solution is lets restricted in that it does not require a planar aperture, but the Sommerfeld solutions do. this is not a big deal, but it is a difference.

A note about the distances. Goodman defines the observation point to be P_0, the point on the aperture to be P_1 and the source to be P_2, (figure 3.7 in Goodman). The distance from P_0 to P_1 is $\mathbf{r}_{01} = \mathbf{r}_0$ in my language (pointing from P_1 to P_0 in his), the distance from P_2 to P_1 is $\mathbf{r}_{21} = \mathbf{r}$. The vector $\mathbf{n}$ points from source to aperture, and is normal to the aperture.

Goodman proceeds by choosing to focus on the first Rayliegh-Sommerfeld soution in the fair field form given by eq. (8.28). The spherical wave source is not necessarily presumed in later portions. The theories are only differ in their results very close to the aperture, which is unsurprising, although Goodman doesn't areticulate precisly when version of the RS solutions are used, i believe it is the most simplified ones. Goodman also notes that while Kirchoff theory is internally inconsitent (RS are not), the RS theory requires planar screens, which Kirchhoff does not. It's not clear to me if the theories suffer at large angles $\cos(\mathbf{n}, \mathbf{r}_{21})$.

8.2.3 3.7 Furthur discussion of the Huygens Fresnel principle

The Huygens Fresnel principle is the familiar statement of Huygens principle, with extra details. Goodman provides a "quasi-physical" intepretation, based off this reepression of Sommerfeld I.

$$U(P_0) = \frac{1}{i\lambda} \int_{\mathcal{A}} U(P) \frac{e^{ik(r)}}{r} \cos(\mathbf{n}, \mathbf{r}) dS \tag{8.36}$$

That interpretation is that the field at P_0 is the sum of the contributions from all points on the aperture, where each contribution is a secondary source spherical wave with the following four properties:

1. the amplitude of the secondary wave is proportional to the amplitude of the primary wave at the point on the aperture
2. the amplitude is inversely proportional to the wavelength
3. the phase of the secondary wave *leads* the phase of the primary wave by a quarter period
4. the secondary wave is attenuated by the factor $\cos(\mathbf{n}, \mathbf{r})$

I still dont get the leading phase thing. motiate it fromt he derivative of the field at P on the aperture. Goodman writes: "Since or basic monochromatic field disturbance is a clockwise rotating phasor of the form $\exp(-i2\pi\nu t)$, the derivative of this fucntion is proportional to both ν and $-i = 1/i$". I dont get it. oh wait... i do get it, i had a sign error in my head about the time relationship.

8.3 Goodman Chapter 4, Fresnel and Fraunhofer

8.3.1 Goodman 4.1.2, starting point before Fresnel

Skipping sectionmuch of 4.1 which is about physics, but section 4.1.2 defines the geometry. I do not understand why Googman has chosen the conventions for labeling points and directions as he has, but oh well. It is the same as in Goodman section 3.5.

The formula we are using from the previous chapter of Goodman is, as previously stated, the first RS formula,

$$U_I(P_0) = \frac{1}{i\lambda} \int_{\mathcal{A}} U(P) \frac{e^{ikr}}{r} \cos(\mathbf{n}, \mathbf{r}) dS. \tag{8.37}$$

The cosine term is trivially $\cos(\mathbf{n}, \mathbf{r}) = \frac{z}{r_{01}}$ and thus we can write our starting point before invoking the Fresnel approximation:

$$U(x, y, z) = \frac{z}{i\lambda} \int_{\mathcal{A}} U(\xi, \eta) \frac{e^{ikr_{01}}}{r_{01}^2} d\xi d\eta. \tag{8.38}$$

where ξ and η are the coordinates of the aperture and the distance r_{01} is precisely

$$r_{01} = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2} \tag{8.39}$$

This is our starting point before making the Fresnel approximations. We have first invoked the inherent approximation of the scalar theory, namely that the aperture is large which respect to the wavelength, see the discussion at the end of Goodman 3.2 for refresher; this has been fundamental in our entire discussion of the scalar theory. We have also invoked the far-field assumption, namely that $r_{01} \gg \lambda$. This was done in section Goodman 3.5.2.

8.3.2 Goodman 4.2, The Fresnel Approximation

With eq. (8.38) as our starting point, we apply the binomial expansion and approximate:

$$r_{01} \approx z\left[1 + \frac{1}{2}\left(\frac{x-\xi}{z}\right)^2 + \frac{1}{2}\left(\frac{y-\eta}{z}\right)^2\right] \tag{8.40}$$

For the r_{01} which appears in the denominator, we are typically quite safe in furthur approximating $r_01 \approx z$. However, in the exponent changes on the order of the wavelength are significant, and we shall therefore keep thehigher ordr terms in eq. (8.40). Factoring out the z contribution to the integrand gives:

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z}\int U(\xi,\eta,0)\exp\left\{\frac{ik}{2z}\left[(x-\xi)^2 + (y-\eta)^2\right]\right\} \tag{8.41}$$

This is our first form of the *Fresnel diffraction integral* or the Fresnel diffraction formula.

Equation eq. (8.41) is quite clearly a convolution, and we are going to skip that for now because it doesn't matter to this portion.

By factoring out the portion of the exponential which is independent of the integration variables we find anothe form of the Fresnel diffraction formula

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z}e^{\frac{ik}{2z}(x^2+y^2)}\int U(\xi,\eta,0)e^{\frac{ik}{2z}(\xi^2+\eta^2)}e^{-\frac{i2\pi}{\lambda z}(x\xi+y\eta)}d\xi d\eta \tag{8.42}$$

This second form of the Fresnel difrction formula shows how the obsered field $U(x,y,z)$ is the Fourier transform of the quantity $U(\xi,\eta,0)e^{\frac{ik}{2z}(\xi^2+\eta^2)}$. In other words, it is the transform of the field at the aperature with an additional quadratic phase. Therefore we can write it as

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z}e^{\frac{ik}{2z}(x^2+y^2)}\mathcal{F}\left[U(\xi,\eta,0)e^{\frac{ik}{2z}(\xi^2+\eta^2)}\right]\left(\frac{x}{\lambda z},\frac{y}{\lambda z}\right) \tag{8.43}$$

Equation's eqs. (8.41) to (8.43) are all equivalent expressions of the Fresnel diffraction formula. It's approximations are (1) large aperture vs wavelength, (2) far field vs wavelength, and (3) the approximation of the binomial theorem in this section, which is related to the stationary phase approximation.

Goodman 4.2.1, Postive and negative Phases

this section is dedicated to discussing the annoying sign confusion of the problem, which depend on the convention used by the phasors (goodman chooses Clockwise, i.e. fields oscillate at $\exp -i\omega t$). This is a helpful conversation but not required for my present purpose yet. It also jumps the gun in its discussion of the relationship between the quadratic phase and spherial wavelets.

Goodman 4.2.2 and 4.2.3, Accuracy of the Fresnel Approximtion

The derivation of the Fresnel diffraction formulas has relied upon the Huygens-Fresnel princible, which treats every point of a field as a source for an outgoing spherical wavelet, and the field at a given obervaton ppoint is the sum of contributions from all the other points. Inspection of the Fresnel diffraction formulas shows that these spherical secondary wavelets have been replaced by wavelets with quadratic-phase wavefronts. (I think

this is most directly seen in the form of eq. (8.41); i think that the prefactor in eqs. (8.42) and (8.43) are not fundamentlly the replacement being made). This was done specifically when the binomial approximation for the distance to the observation point from the aperture was invoked.

A naive measure of the accuracy of this approach thus requires that the higher order terms dropped from the binomial expansion mus be small. This turns out to be much too pessimistic. Rather, the requirement for validity of the approximatiion is that the value of the integral not change significantly with the addition of higher terms. Goodman shows how the behavior of this integral (which is quite appropriately given in terms of Fresnel integrals) is not so sensitive to higher order contributions. Goodman furthur describes how much of the aperture region contributes the main portion of the integral. This concept is alluded to be closely related t the *princible of stationary phase*.

Goodman 4.2.4 through 4.2.6

the next three sections deal with other details, 4.2.4 in particular returns to the convolution insight I skipped earlier, and discusses the power spectrom.

8.3.3 Goodman 4.3 the Fraunhofer Approximation

The Fraunhofer aproximation builds upon the Fresnel approximation by further supposing that the quadratic phase ($e^{\frac{ik}{2z}(\xi^2+\eta^2)}$) added to the field $U(\xi,\eta,0)$ at the aperture may be replaced with unity. It is the Fraunhofer approximation which is the most pure manifestation of the concept that the field in the far field is the Fourier transform of the field over the aperture.

Goodman has several insightful comments which I will not delve into now. They include the (more stringent) limitations of regimes of validity and how these are ameliorated by ceratin geometries or via a lens. The effect of lens's is in Chapter 6 of Goodman.

8.4 Thin lens as a Fourier Transformer

Here we will derive the fact that a thin lens acts as a Fourier transformer. This section demonstrates the special case of a result derived in slightly more generality in the chapter 6 of Goodman, specifically section 6.2.2. Let's begin with the expression for the field $U(x,y,z)$ due to a diffraction through an aperture at the origin, where the field has value $U(\xi,\eta,0)$. Scalar difrraction theory gives the field under the Fresnel approximation (cite Goodman, equation 4-17) as

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ik}{2z}(x^2+y^2)} \int U(\xi,\eta,0) e^{\frac{ik}{2z}(\xi^2+\eta^2)} e^{-\frac{i2\pi}{\lambda z}(x\xi+y\eta)} d\xi d\eta \tag{8.44}$$

or, equivalently:

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ik}{2z}(x^2+y^2)} \mathcal{F}\left[U(\xi,\eta,0) e^{\frac{ik}{2z}(\xi^2+\eta^2)}\right]\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right) \tag{8.45}$$

where we are the unitary Fourier transform operator $\mathcal{F}$. This approximation is valid when the aperture is small compared to the distance from the aperture to the observation plane, wen the distance beyond the aperture is large compared to the wavelength, and when the stationary phase approximation is valid. It is thus perfectly suitable for describing the propagation of a wavefront leaving a spatial light modulator (SLM) and propagating to a lens, and then to a camera. The effect of the lens is to introduce a quadratic phase $e^{-i\frac{k}{2f}(x^2+y^2)}$ to the field, where f is the focal length of the lens (Goodman eq 6.10). Dickey also has this formula , at the beginning of section 2.6.3 but the supplemental material lacks the negative sign both Goodman and Dickey have. The electric field at the camera is thus two applications of the Fresnel formula, one for the propagation from the SLM to the lens, and one for the propagation from the lens to the camera,

with the phase of the lens in between. The field at the camera is thus given by:

$$
\begin{aligned}
U(x,y,2f) &= \left(\frac{e^{ikf}}{i\lambda f}\right) e^{\frac{ik}{2f}(x^2+y^2)} \int \left[U(\xi,\eta,f) \overbrace{\cancel{e^{-\frac{ik}{2f}(\xi^2+\eta^2)}}}^{lens} \right] \overbrace{\cancel{e^{\frac{ik}{2f}(\xi^2+\eta^2)}}}^{Fresnel} e^{-\frac{i2\pi}{\lambda f}(x\xi+y\eta)} d\xi d\eta \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 e^{\frac{ik}{2f}(x^2+y^2)} \int e^{\frac{ik}{2f}(\xi^2+\eta^2)} \Bigg\{ \int U(x',y',0) \\
&\quad \times e^{\frac{ik}{2f}(x'^2+y'^2)} e^{-\frac{i2\pi}{\lambda f}(\xi x'+\eta y')} dx'dy' \Bigg\} e^{-\frac{i2\pi}{\lambda f}(x\xi+y\eta)} d\xi d\eta \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 e^{\frac{ik}{2f}(x^2+y^2)} \int U(x',y',0) \\
&\quad \times \left\{ \int e^{-\frac{i2\pi}{2\lambda f}(2x\xi+2y\eta)} e^{-\frac{i2\pi}{2\lambda f}(2\xi x'+2\eta y')} e^{\frac{i2\pi}{2\lambda f}(\xi^2+\eta^2)} e^{\frac{i2\pi}{2\lambda f}(x'^2+y'^2)} d\xi d\eta \right\} dx'dy' \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 \int U(x',y',0) \\
&\quad \times \left\{ \int e^{\frac{i2\pi}{2\lambda f}(\xi^2+\eta^2-2x\xi-2y\eta-2\xi x'-2\eta y'+x'^2+y'^2+x^2+y^2)} d\xi d\eta \right\} dx'dy' \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 \int U(x',y',0) \underbrace{\left\{ \int e^{\frac{i2\pi}{2\lambda f}((\xi-x-x')^2+(\eta-y-y')^2)} d\xi d\eta \right\}}_{i\lambda f} e^{\frac{i2\pi}{2\lambda f}(2x'x-2y'y)} dx'dy' \\
&= -\frac{ie^{i4\pi f/\lambda}}{\lambda f} \mathcal{F}[U(x',y',0)] \left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right)
\end{aligned}
\tag{8.46}
$$

where we've used completing the square, flipping the order of integration, and evaluated an internal Gaussian integral. The field at the camera is thus the Fourier transform of the field at the SLM, multiplied by a phase factor. We may drop the irrelevant phase factor $e^{i4\pi f/\lambda}$, and the field at the camera is thus precisely the Fourier transform of the field at the SLM. (see goodman section 6.2.2)

8.4.1 with different variables

paper: slm plane is $u(x,y,0)$ lens plane is $u(v,w,f)$ and image plane is $u(X,Y,2f)$ me: slm plane is $u(x,y,0)$ lens plane is $u(v,w,f)$ and image plane is $u(X,Y,2f)$

$$
\begin{aligned}
U(X,Y,2f) &= \left(\frac{e^{ikf}}{i\lambda f}\right) e^{\frac{ik}{2f}(X^2+Y^2)} \int \left[U(v,w,f) \overbrace{\cancel{e^{-\frac{ik}{2f}(v^2+w^2)}}}^{lens} \right] \overbrace{\cancel{e^{\frac{ik}{2f}(v^2+w^2)}}}^{Fresnel} e^{-\frac{i2\pi}{\lambda f}(Xv+Yw)} dvdw \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 e^{\frac{ik}{2f}(X^2+Y^2)} \int e^{\frac{ik}{2f}(v^2+w^2)} \left\{ \int U(x,y,0) \right. \\
&\quad \left. \times e^{\frac{ik}{2f}(x^2+y^2)} e^{-\frac{i2\pi}{\lambda f}(vx+wy)} dxdy \right\} e^{-\frac{i2\pi}{\lambda f}(Xv+Yw)} dvdw \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 e^{\frac{ik}{2f}(X^2+Y^2)} \int U(x,y,0) \\
&\quad \times \left\{ \int e^{-\frac{i2\pi}{2\lambda f}(2Xv+2Yw)} e^{-\frac{i2\pi}{2\lambda f}(2vx+2wy)} e^{\frac{i2\pi}{2\lambda f}(v^2+w^2)} e^{\frac{i2\pi}{2\lambda f}(x^2+y^2)} dvdw \right\} dxdy \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 \int U(x,y,0) \\
&\quad \times \left\{ \int e^{\frac{i2\pi}{2\lambda f}(v^2+w^2-2Xv-2Yw-2vx-2wy+x^2+y^2+X^2+Y^2+2xX+2yY)} e^{\frac{i2\pi}{2\lambda f}(-2xX-2yY)} dvdw \right\} dxdy \\
&= \left(\frac{e^{ikf}}{i\lambda f}\right)^2 \int U(x,y,0) \underbrace{\left\{ \int e^{\frac{i2\pi}{2\lambda f}((v-X-x)^2+(w-Y-y)^2)} dvdw \right\}}_{i\lambda f} e^{\frac{-i2\pi}{2\lambda f}(2xX+2yY)} dxdy \\
&= -\frac{ie^{i4\pi f/\lambda}}{\lambda f} \mathcal{F}[U(x,y,0)] \left(\frac{X}{\lambda f}, \frac{Y}{\lambda f}\right)
\end{aligned}
\tag{8.47}
$$

8.5 an aside about numerical diffraction packages, move elsewhere

In learning all this about diffraction, I was motivated in part by the desire to understand the numerical diffraction packages that i have been using.

The one i'm the most familiar with right now is the LightPipes package for python. this package is based on code going back to at least the early 90's, and therefore may not incorporatte more recent advances in the field. But it is a good place to start, and it has a particularly nice feature where the coordinate system can be transformed from rectilinear to spherical, which lets the grid contract around the focal point of a lens. but its also not quite as fast as i would like. it lso isn't the easiest to use and understand, particularly the coordinate system stuff.

I also found the pyoptica package, which is nice and fast, and well documented down to the level of the math. but it doesn't have the coordinate system stuff, and it doesn't seem to be as well maintained. the lasst commmit was two years ago as of this writing. but the code is very well documented, and it is very fast, though much of that is wastted when propagating to the focus of a lens, becuse the grid isn't adaptive.

Probably the most impressive package i found was the Diffractio package, which has a lot of features for field in varius dimensions, and its well documented, with plotting features that are very nice. I haven't used it much yet, but i'm excited to try it out, though i haven't come accross the coordinate system stuff yet, which is a bummer. the algorithms references are pretty recent and up to date, which is nice. its possible that some of the methods i haven't tried out yet have the coordinate system stuff, like the CZT method. i looks like it does actually... i have to try this out.

so i want to answer the question of how thick an optical element with with a given (differential) index of refraction must be to introduce a phase shift of 2π for light of a certain wavelength λ. Well, the phase shift is

given by

$$\phi = \frac{2\pi}{\lambda} n \Delta L \tag{8.48}$$

right? well if $n = 2$ the wavelength in the material is $\lambda/2$, so a layer of thickness λ will introduce a phase shift of 2π since it will contain two full wavelengths compared to a wavefront propagating through a vacuum, which will have one wavelength.

Chapter 9

Camera Calibration

9.1 Camera calibration

We would like to know how many atoms we have from a picture of them. this will depend on knowing how many photons we have for a given camera signal. This will depend on varius settings of the camera potentially.

We are currently using cameras from Flir, where this page is helpful in outlined the terms and definitions: https://www.flir.com/discover/iis/machine-vision/how-to-evaluate-camera-sensitivity/

We begin with the photon hitting a sensor. The photon may be converted to an electron, which is stored in the 'Well'. The number of electrons in the Well is then amplified with some Gain (which may be unity) and then the resulting signal is sent to the ADC where it is translated into some digital number. That is then saved to the computer.That's a high-level overview, but the devil is in the details.

Let's consider an image taken with exposure time t_{exp}. Let's begin by assuming we have N_γ photons incident upon our sensor during the exposure time. The probability that this photon will contribute an electron to the Well is called the Quantum Efficiency (QE). The value of the QE is generally wavelength dependent, and varies from detector to detector; let's denote the QE by the symbol Q_{eff}. However, due to thermal energy in the electronic there will also bea slow buildup of thermal electrons in the well, and these are indistinguishable from the photoelectrons. The rate at which these build up we shall denote $I_d(T)$ which has units of e^-/s. The dark current typically is negligible for short exposure times (as for us), and can be ameliorated via cooling the sensor. Therefore, after the exposure the Well has a number of electrons:

$$N_{well\ e^-} = Q_{eff} N_\gamma + I_d(T) t_{exp} \tag{9.1}$$

However, the Well can only hold a limited number of electrons, called the well depth D_{e^-}. If this amount is exceded the pixel will be saturated. In CCD cameras the well can overflow to neighboring pixels, and effect called blooming. The electrons in the well are an analog signal, and it is at this point that the application of the camera's amplifying gain may be be applied to the signal. The gain of an electronic signal is given by

$$Gain = 10 \log_{10} \frac{P_{out}}{P_{in}} \tag{9.2}$$

and has units of dB. When the input and output impedances are the same, we can expressed the voltage gain G_V as

$$G_V = 20 \log_{10} \frac{V_{out}}{V_{in}} \tag{9.3}$$

and so we immediatly see

$$V_{out} = V_{in} 10^{V_G/20}. \tag{9.4}$$

The electrons in the Well are simply analog voltages, and are amplified in the same way. The resulting quantity of electrons is then fed to the ADC, where the electrons are converted to Analog-Digital Units (ADU), also often called "counts". This is the unit that is ultimately stored in the TIFF file. The process of

reading out the analog signal will introduce noise, called the read noise. The read noise is generally given in units of electrons, and we shall denote it by R_{e^-}. The total number of counts in the image is then given by

$$N_{counts} = G_V N_{well\ e^-} + R_{e^-} = G_V(Q_{eff}N_\gamma + I_d(T)t_{exp}) + R_{e^-}. \tag{9.5}$$

The conversion from electrons to counts is called the gain, and is given in units of counts per electron. This is a different gain from the amplification gain, and we shall denote it by G_{ADU}. The gain is generally a constant for a given camera, but can be changed in the camera settings sometimes. The gain is generally given in units of counts per electron or the inverse, but is really a conversation factor, not an actual gain.

At this point, one final amplification of the signal can occur, called digital gain, which is simply a multiplication of the counts by a constant. This is generally done in the camera settings, and is generally not recommended as it will amplify the read noise as well. its not real gain, and is really more an interolation and upscaling of the counts.

9.1.1 Additional factors

There are other factors and processes which may occur in the image taking process. (has the lower limit been an issue? unlikely). The Flir cameras have a gamma setting, which seems to be intended to make the image look better to the human eye. This is a non-linear transformation of the counts, and so will affect the number of counts in the image. The default value for the gamma setting in the Flir cameras seems to be 0.8, but this can be changed or disabled. I'm not certain, but i think that the gamma setting is applied prior to the ADC?

Another factor is the image offset, or black level. This is a constant offset added to the signal both before and after the ADC. This is generally done to remove the read noise from the image, and is generally a small number of counts.

$$P_{out} = P_{max}\left(\frac{P_{in}}{P_{max}}\right)^\gamma \tag{9.6}$$

Chapter 10

Books

I want these books.

1. Kedar Khare (2016) Fourier Optics and Computational Imaging, Wiley and Sons Ltd.
2. David Voelz (2011) Computational Fourier Optics Matlab Tutorial, Spie Press
3. Visual Differential Geometry and Forms: A Mathematical Drama in Five Acts
4. Analytical Mechanics by Lemos