Primer in Fourier Analysis

A.C. Machacek

June 24, 2024

Fourier analysis is a process of taking a time-dependent function and investigating the frequencies present in it. This uses an integral called the Fourier Transform. In this article, we shall look at the nature of how a signal's time dependence affects its constituent frequencies for simple examples before presenting the Fourier Transform.

After deriving relevant results, such as the Inverse Fourier Transform, the power and convolution theorems, we shall explore its use in diffraction, spectroscopy, and information transfer. The origin of the Uncertainty Principle in quantum mechanics will be shown to be a natrual consequence of using wave mechanics for quantum objects.

Finally, we shall explore the use of Fourier transforms to describe waves and wave equations. In this way, the Maxwell equations of electromagnetism will be reduced to linear simultaneous equations which are easy to solve to derive the speed of such waves. We shall also apply these principles to investigate the dispersion relationships for light and electrostatic waves in plasmas.

1 Terminology

In this article, unless specified otherwise, when 'frequency' is referred to, *angular frequency* is meant. This is denoted ω, is measured in radians/second, and is equal to $2\pi f$ where f is the conventional frequency in Hertz. Given that we shall use radians throughout, it means that a wave of frequency ω can be written $y = A\cos\omega t$. Thus, when t increases from 0 to the time period $T = 1/f$, the argument of the cosine, namely ωt, rises from 0 to 2π. This produces a whole cycle of the cosine wave.

We will typically use complex exponential notation for these signals, and will accordingly write

$$y = Ae^{i\omega t}. \tag{1}$$

.

Where a wave is written as a function of position at a particular time, we will use the *wavenumber* k which is defined as $2\pi/\lambda$. Thus as the position x changes from 0 to λ, the product kx will rise from 0 to 2π. Thus we can write the wave as $y = \cos kx$. We note that $\omega/k = f\lambda$ and accordingly gives the wave's speed v.

Using complex exponential notation, we would have

$$y = Ae^{ikx}. \tag{2}$$

.

For a wave travelling in the $+x$ direction, the disturbance $y(x,t)$ will be the same as the disturbance at the origin a time x/v earlier. Thus $y(x,t) = y(0, t - x/v)$. If $y(0,t) = A\cos\omega t$, then it follows that

$$\begin{aligned} y(x,t) &= A\cos\omega\left(t - \frac{x}{v}\right) \\ &= A\cos\left(\omega t - \frac{\omega x}{v}\right) \\ &= A\cos\left(\omega t - kx\right). \end{aligned} \tag{3}$$

We will typically use complex exponential notation for waves, and accordingly, a plane wave would be written

$$y = Ae^{i(\omega t - kx)}. \tag{4}$$

Notice that if we are considering a wave simply in distance, we write $y = Ae^{ikx}$, whereas if distance and time need considering, we write $y = Ae^{i(\omega t - kx)}$, and these two expressions have different signs for the distance aspect of the exponent.

2 Amplitude Modulation and Beats

The simplest way of sending sound by radio is called *amplitude modulation.* Here the high frequency radio wave (the *carrier wave*) has its amplitude varied according to the current state of the low frequency audio signal we wish to transmit. Suppose that the carrier wave is described $y_c = A\cos\omega t$, and the sound wave by $y_s = B\cos\Omega t$, then the two would be combined in such a way that the amplitude A was no longer constant but equal to $A + y_s$, where we ensure that $B < A$. The new wave is then

$$y = (A + B\cos\Omega t)\,\cos\omega t. \tag{5}$$

If we use the trignometrical formula $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$, it follows that $2\cos a \cos b = \cos(a+b) + \cos(a-b)$. Thus our radio wave can be written

$$y = A\cos\omega t + \frac{B}{2}\left\{\cos\left(\omega+\Omega\right)t + \cos\left(\omega-\Omega\right)t\right\}. \tag{6}$$

Thus, the addition of the audio signal of frequency Ω has caused the frequencies present in the wave to include not only ω but also $\omega \pm \Omega$. Conveying the extra information has increased the number (and range) of frequencies present. As we shall see more formally later, the transmission of information increases the *bandwidth* of the wave transmitted. To understand the demands on a communication channel, it is necessary to know the frequencies it will be required to transmit, and frequency analysis is accordingly vital.

If you have tuned a musical instrument, you will be familiar with the phenomenon of 'beats', that is the apparent pulsing of sound when two similar notes are played at the same time. To see this mathematically, imagine that you are tuning your violin or guitar. Your friend, whose instrument has already been tuned, is playing the frequency ω. You play the frequency $\omega + \delta$ at the same time. Let us write the mean frequency ϵ and the frequency difference 2η. The original frequencies are thus $\epsilon - \eta$ and $\epsilon + \eta$. Suppose you both play with the same amplitude A, the result will be

$$\begin{aligned}
y &= A\cos\left(\epsilon-\eta\right)t + A\cos\left(\epsilon+\eta\right)t \\
&= \left(A\cos\epsilon t\cos\eta t + A\sin\epsilon t\sin\eta t\right) \\
&\quad + \left(A\cos\epsilon t\cos\eta t - A\sin\epsilon t\sin\eta t\right) \\
&= 2A\cos\epsilon t\cos\eta t.
\end{aligned} \tag{7}$$

In this case, we will hear an overall sound with the average frequency ϵ, with amplitude given by $A|\cos\eta t|$. This will give full strength whenever $\eta t = 0$ or $\eta t = \pi$, and as such the time between successive full strength sounds will be $\pi/\eta = 2\pi/\delta$. This means that the angular frequency of the pulsing (or *beats*) will be $2\pi/t = \delta$. Musicians use this to help tune their instruments, adjusting the strings so that the frequency of this pulsing decreases until the volume is steady when the two strings are in tune with each other. This is another example of a combination of two frequencies producing what appears to be a modulated single frequency.

In Fourier Analysis, we aim to establish the frequencies present in any signal.

3 Fourier Transform

Assuming that we have a signal of the form $f(t)$, the frequencies present in it are found from the Fourier Transform function, which is defined as

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\, e^{-i\omega t}\, dt. \tag{8}$$

Before proving that this works, let us explore why this function is suitable. Suppose that our initial function did have a particular frequency, say $f(t) = f_0 e^{i\Omega t}$. When this is put into equation 8, we end up integrating $e^{i(\Omega-\omega)t}$ over a very wide range of values of t. When $\omega = \Omega$, the integrand becomes $e^0 = 1$, and the integral will yield a large answer. However, if a different value of ω is chosen, where $\delta = (\Omega - \omega) \neq 0$, then the integral becomes

$$F(\omega) = \int e^{i\delta t} dt = \left[\frac{e^{i\delta t}}{i\delta}\right]. \tag{9}$$

This indefinite integral gives the same value every time δt increases by 2π, and thus every time interval from t to $t + 2\pi/\delta$ contributes nothing to the total value of the integral. As the full integral will be made up of many of these time periods, each worth nothing on its own, the full integral will be zero (or close to it).

Thus our transform in equation 8 has correctly identified the frequency of the original signal.

To show this working more rigorously, we will first need to prove some integrals which we will refer to frequently.

4 Integrals for Reference

We start by stating, without proof, the value of the definite integral

$$\int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}. \tag{10}$$

It follows, if we use integration by substitution $y = \sqrt{a}x$, that

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-ax^2}\, dx &= \int_{-\infty}^{+\infty} e^{-y^2}\, \frac{dy}{\sqrt{a}} \\ &= \sqrt{\frac{\pi}{a}}. \end{aligned} \tag{11}$$

Next, we use the method of completing the square to evaluate, with $y = \sqrt{a}x$ as before,

$$
\begin{aligned}
\int_{-\infty}^{+\infty} e^{-ax^2}\, e^{-ikx}\, dx &= \int_{-\infty}^{+\infty} e^{-y^2}\, e^{-iky/\sqrt{a}}\, \frac{dy}{\sqrt{a}} \\
&= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-(y-ik/2\sqrt{a})^2}\, e^{-k^2/4a}\, dy \\
&= \frac{e^{-k^2/4a}}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-(y-ik/2\sqrt{a})^2}\, dy \\
&= \frac{e^{-k^2/4a}}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-z^2}\, dz \\
&= \sqrt{\frac{\pi}{a}}\, e^{-k^2/4a}, \qquad (12)
\end{aligned}
$$

where we used $z = y - ik/2\sqrt{a}$ in the final stage.

Notice that if we repeated this last integral with a function of x which was peaked at $x = X$ rather than $x = 0$, the result would be, after substituting $w = x - X$,

$$
\begin{aligned}
\int_{-\infty}^{+\infty} e^{-a(x-X)^2}\, e^{-ikx}\, dx &= \int_{-\infty}^{+\infty} e^{-aw^2}\, e^{-ik(X+w)}\, dw \\
&= e^{-ikX} \int_{-\infty}^{+\infty} e^{-aw^2}\, e^{-ikw}\, dw \\
&= e^{-ikX} \int_{-\infty}^{+\infty} e^{-y^2}\, e^{-iky/\sqrt{a}}\, \frac{dy}{\sqrt{a}} \\
&= e^{-ikX}\, \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-(y-ik/2\sqrt{a})^2}\, e^{-k^2/4a}\, dy \\
&= \frac{e^{-ikX}\, e^{-k^2/4a}}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-(y-ik/2\sqrt{a})^2}\, dy \\
&= \frac{e^{-ikX}\, e^{-k^2/4a}}{\sqrt{a}} \int_{-\infty}^{+\infty} e^{-z^2}\, dz \\
&= \sqrt{\frac{\pi}{a}}\, e^{-ikX}\, e^{-k^2/4a}. \qquad (13)
\end{aligned}
$$

Finally, we note that if the complex exponent were $k - k_0$ rather than simply k, equations 12 and 13 would become

$$
\int_{-\infty}^{+\infty} e^{-ax^2}\, e^{-i(k-k_0)x}\, dx = \sqrt{\frac{\pi}{a}}\, e^{-(k-k_0)^2/4a}, \qquad (14)
$$

and

$$
\int_{-\infty}^{+\infty} e^{-a(x-X)^2}\, e^{-i(k-k_0)x}\, dx = \sqrt{\frac{\pi}{a}}\, e^{-i(k-k_0)X}\, e^{-(k-k_0)^2/4a}. \qquad (15)
$$

5 The Delta Function

The integral $\delta(k) = \int e^{-ikx} dx/2\pi$ over all x has a useful property. To prove this, we shall work with the more polite form

$$\epsilon(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx}\, e^{-ax^2}\, dx, \tag{16}$$

where we shall let $a \to 0$ to achieve equality with $\delta(k)$.

Using equation 12, $\epsilon(k)$ evaluates to

$$\begin{aligned} \epsilon(k) &= \frac{1}{2\pi} \sqrt{\frac{\pi}{a}}\, e^{-k^2/4a} \\ &= \sqrt{\frac{1}{4a\pi}}\, e^{-k^2/4a} \end{aligned} \tag{17}$$

Firstly, note that if $k \neq 0$, then as we let $a \to 0$, equation 17 shows us that $\epsilon(k) \to 0$. If $k = 0$, on the other hand, ϵ becomes rather large as $a \to 0$.

That said, notice that

$$\begin{aligned} \int_{-\infty}^{+\infty} \epsilon(k) dk &= \sqrt{\frac{1}{4a\pi}} \int_{-\infty}^{+\infty} e^{-k^2/4a}\, dk \\ &= \sqrt{\frac{1}{4a\pi}} \sqrt{\frac{\pi}{1/4a}} = 1, \end{aligned} \tag{18}$$

where we have used the integral in equation 11. Thus $\epsilon(k)$ is normalized regardless of the value of a, and accordingly,

$$\int_{-\infty}^{+\infty} \delta(k)\, dk = 1. \tag{19}$$

Thus $\delta(k)$ a normalized function whose value is zero everywhere other than at $k = 0$.

It use comes in expressions like this one

$$\int_{-\infty}^{+\infty} f(k)\, \delta(k - K)\, dk. \tag{20}$$

First, let us imagine $f(k) = f_0$ is a constant. In this case, $f(k) = f_0$ can be brought outside the integral, and we have

$$\int_{-\infty}^{+\infty} f_0\, \delta(k - K)\, dk = f_0 \int_{-\infty}^{+\infty} \delta(k - K)\, dk = f_0, \tag{21}$$

because the $\delta(k)$ function integrates to 1.

Next, let us note that whatever $f(k)$ was when $k \neq K$ can not have influenced the final result, as $\delta(k-K) = 0$ in that case. Therefore, only the value of $f(k)$ when $k = K$ can influence the integral. Therefore the result was f_0 not because f_0 was constant, but because $f(k) = f_0$ specifically when $k = K$. It follows that

$$\int_{-\infty}^{+\infty} f(k)\,\delta(k-K)\,dk = f(K). \tag{22}$$

6 Inverse Fourier Transform

Suppose we are given $F(\omega)$ and we wish to recover $f(t)$. By equation 8, $F(\omega) = \int f(t)\,e^{-i\omega t}\,dt$. Given that the Fourier transform was performed by multiplication by $e^{-i\omega t}$ followed by integration, and that the inverse of multiplication by $e^{-i\omega t}$ is multiplication by $e^{i\omega t}$, let us see what happens if we multiply $F(\omega)$ by $e^{i\omega T}$ and integrate:

$$\begin{aligned}
\int_\omega F(\omega)\,e^{i\omega T}\,d\omega &= \int_\omega \int_t f(t)\,e^{-i\omega t}\,e^{i\omega T}\,dt\,d\omega \\
&= \int_t f(t) \int_\omega e^{-i(t-T)\omega}\,d\omega\,dt \\
&= \int_t f(t)\,2\pi\,\delta(t-T)\,dt \\
&= 2\pi f(T),
\end{aligned} \tag{23}$$

and we recover our original time-dependent function f, where we have used equation 22 and the definition of the δ function.

Thus, we write the inverse Fourier Transform function as

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)\,e^{i\omega t}\,d\omega. \tag{24}$$

Notice that this equation presents an argument in its own right that shows that any time-dependent function can be made out of a sum of lots of single-frequency oscillations $e^{i\omega t}$ each with its own amplitude $F(\omega)$.

7 Simple Pulse

Let us now see which frequencies we will need to send a simple pulse down a channel at a time T. To do this, we will use a signal

$$f(t) = A\,e^{-a(t-T)^2}, \tag{25}$$

where a gives a measure of how short the pulse is — the larger a, the shorter the pulse.

Using equation 13, the Fourier transform of our pulse can be evaluated as

$$F(\omega) = \int_t f(t)e^{-i\omega t}dt = A\sqrt{\frac{\pi}{a}}\, e^{-i\omega T}\, e^{-\omega^2/4a}. \tag{26}$$

The frequencies used are spread equally either side of $\omega = 0$. As we make a shorter pulse, we increase a, and we note that we need a greater range of frequencies.

8 Bandwidth Theorem

When a variable is randomly distributed on the Normal distribution, we write the probability of it taking a value x as being proportional to $e^{-(x-\mu)^2/2\sigma^2}$ where μ is the mean and σ is the standard deviation.

If we have a signal with amplitude $f(t) = Ae^{-a(t-T)^2}$ as in equation 25, then the intensity will be proportional to the amplitude squared, and thus

$$I_t \propto |f(t)|^2 = A^2\, e^{-2a(t-T)^2}. \tag{27}$$

By comparing this with the Normal distribution, T represents the mean, and $2a$ represents $1/2\sigma^2$. Therefore if we define the 'width' of the pulse using the idea of a standard deviation applied to the intensity function, we have

$$\sigma_t = \sqrt{\frac{1}{4a}}. \tag{28}$$

We now look at the Fourier transform, as shown in equation 26. If we take the modulus square to find the 'intensity' of each frequency component, we have

$$I_\omega \propto |F(\omega)|^2 = \frac{A^2\pi}{a}\, e^{-\omega^2/2a}. \tag{29}$$

When we observe this frequency distribution, and compare it with the Normal distribution, we have a mean frequency of zero. As far as deviation goes, $2a$ is equivalent to $2\sigma_\omega^2$, and thus

$$\sigma_\omega = \sqrt{a}. \tag{30}$$

Notice that σ_ω and σ_t are inversely proportional to each other, and in particular

$$\sigma_t\, \sigma_\omega = \sqrt{\frac{1}{4a}} \times \sqrt{a} = \frac{1}{2}. \tag{31}$$

While we are not going to prove it here, it turns out that the exponential curve e^{-ax^2} has the smallest product possible of $\sigma_t\, \sigma_\omega$. Thus if you want to send a pulse of width σ_t, you will have to put up with a range of frequencies at least $\sigma_\omega = 1/2\sigma_t$.

This principle is known as the *bandwidth theorem*. The more information we wish to send down a channel each second, the shorter the pulses will need to be, and the larger the range of frequencies needed to send them.

9 Uncertainty Principle

The bandwidth theorem has applications in Quantum Mechanics, where we frequently wish to describe the location and state of an electron, or a photon, say, using a *wavefunction*. Here the σ values are known as uncertainties. For such particles, the kinetic energy E is related to the angular frequency ω by the equation $E = \hbar\omega$ where $\hbar = h/2\pi$ and h is known as the Planck constant.

As $E = \hbar\omega$, it follows that $\sigma_E = \hbar\sigma_\omega$. Notice that the consequence of equation 31 will be that

$$\sigma_E\sigma_t = \hbar\sigma_\omega\,\sigma_t \geq \frac{\hbar}{2}, \tag{32}$$

where equality holds for pulses in the shape of an e^{-ax^2} curve.

This is one form of *Heisenberg's Uncertainty Principle*. One consequence of this is that if an excited state of an atom or nucleus is long-lived, σ_t will be large. In consequence σ_E can be small, and the frequency of the radiation emitted when it de-excites will be precisely defined. Some of the longest lived states are the hyperfine transitions where the electronic and nuclear magnetic fields in an atom line up with respect to each other. The radiation emitted when they do so therefore has the most precisely defined photon energies, and hence frequencies, and this is why the radiation emitted from the hyperfine transition of the caesium atom was chosen to define the second.

Similar arguments involving k and x could be used in place of ω and t. With this being done, we have, in place of equation 31

$$\sigma_x\,\sigma_k \geq \frac{1}{2}, \tag{33}$$

with equality occurring in the optimum case of e^{-ax^2} style pulses.

In Quantum Mechanics, the momentum p is related to the wavelength λ by $p = h/\lambda$. Combining this with the result in section 1 that $k = 2\pi/\lambda$, we see

$$p = \frac{h}{\lambda} = \frac{h}{2\pi/k} = \frac{hk}{2\pi} = \hbar k. \tag{34}$$

Thus $\sigma_p = \hbar\sigma_k$, and therefore

$$\sigma_p\sigma_x = \hbar\sigma_k\,\sigma_x \geq \frac{\hbar}{2}, \tag{35}$$

which is another statement of Heisenberg's Uncertainty theorem.

10 Wave Packet

Usually, when we send a pulse down a channel, it will be a pulse of a carrier wave. Thus, rather than the $f(t) = Ae^{-a(t-T)^2}$ of equation 25, we will have

$$f(t) = A\,e^{-a(t-T)^2}\,e^{i\Omega t}, \tag{36}$$

where Ω is the frequency of the carrier wave.

The frequency distribution is then given by

$$\begin{aligned} F(\omega) &= \int_t A\,e^{-a(t-T)^2}\,e^{i\Omega t}\,e^{-i\omega t}\,dt \\ &= \int_t A\,e^{-a(t-T)^2}\,e^{-i(\omega-\Omega)t}\,dt. \end{aligned} \tag{37}$$

Using equation 15, this can be evaluated as

$$F(\omega) = A\sqrt{\frac{\pi}{a}}\,e^{-i(\omega-\Omega)T}\,e^{-(\omega-\Omega)^2/4a}. \tag{38}$$

The range of frequencies is as it was in equation 24, but now is centred on a new mean frequency Ω as we might expect. Notice that the timing of the centre of the pulse T only affects the phase of the Fourier Transform F.

In the limiting case of a very long pulse, where $a \to 0$, equation 38 gives us effectively a very narrow range of frequencies, and in the limiting case of $a \to 0$, the single frequency Ω.

11 Power

In this section, we look at the total power carried in a signal. As mentioned before equation 27, the power or intensity of a signal is related to the square of the amplitude. Let's add this up for all parts of a signal:

$$P = \int_t |f(t)|^2\,dt. \tag{39}$$

We now see how this could be calculated from the Fourier Transforms $F(\omega)$, remembering from equation 24 that $f(t) = \int_\omega F(\omega)e^{i\omega t}dt/2\pi$:

$$P = \int_t f^*(t)\,f(t)\,dt$$

$$
\begin{aligned}
&= \frac{1}{4\pi^2} \int_t \left(\int_{\omega'} F(\omega') e^{i\omega' t} d\omega' \right)^* \left(\int_\omega F(\omega) e^{i\omega t} d\omega \right) dt \\
&= \frac{1}{4\pi^2} \int_t \left(\int_{\omega'} F^*(\omega') e^{-i\omega' t} d\omega' \right) \left(\int_\omega F(\omega) e^{i\omega t} d\omega \right) dt \\
&= \frac{1}{4\pi^2} \int_\omega \int_{\omega'} F^*(\omega')\, F(\omega) \int_t e^{-it(\omega' - \omega)}\, dt\, d\omega'\, d\omega \\
&= \frac{1}{2\pi} \int_\omega \int_{\omega'} F^*(\omega')\, F(\omega)\, \delta(\omega' - \omega)\, d\omega'\, d\omega \\
&= \frac{1}{2\pi} \int_\omega F^*(\omega)\, F(\omega)\, d\omega \\
&= \frac{1}{2\pi} \int_\omega |F(\omega)|^2\, d\omega. \qquad (40)
\end{aligned}
$$

This result, also known as *Parseval's Theorem* shows that the total power of the signal can be evaluated either by adding up the energies in the frequencies, or by working from the energy at each time.

You might be uncomfortable with the fact that there is a factor of $1/2\pi$ in the final answer, rather than the two integrals being equal. This comes from the asymmetry of our defining $F(\omega) = \int_t f(t) e^{-i\omega t} dt$ while the Inverse Fourier Transform has a factor of $1/2\pi$ in front of it. It is possible to define both the Fourier Transform and its inverse with a factor of $1/\sqrt{2\pi}$. If this is done, then equation 40 does not have the factor of $1/2\pi$. In this article we have maintained the usual conventions for the definition of the Fourier transform for compatibility with other texts.

12 Convolution

The convolution of two functions f and g is defined as

$$
f * g = \int_{t'} f(t')\, g(t - t')\, dt'. \qquad (41)
$$

The meaning of this function is easier to visualize in its two dimensional equivalent. The convolution of a square grid of dots and a picture of a single rabbit is a square grid of rabbits (one rabbit gets put on each dot).

In one dimension, $f(t)$ might be a complicated shape which a pulse might have, and $g(t)$ might be a delta function repeated ten times a second (effectively, ten dots each second). In this case, the convolution would be the complicated pulse replicated so that ten of them occur each second.

Let us work out the Fourier transform of the convolution of f and g, where we shall define $\tau = t - t'$ and so $t = \tau + t'$:

$$
\mathcal{F}(f * g) = \int_t \int_{t'} f(t')\, g(t - t')\, e^{-i\omega t}\, dt'\, dt
$$

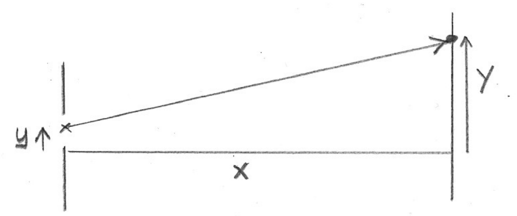

Figure 1: Light passes through an aperture, and we study the image on a distant screen.

$$\begin{aligned}
&= \int_\tau \int_{t'} f(t')\, g(\tau)\, e^{-i\omega(\tau+t')}\, dt'\, d\tau \\
&= \left(\int_\tau g(\tau)\, e^{-i\omega\tau}\, d\tau\right) \left(\int_{t'} f(t')\, e^{-i\omega t'}\, dt'\right) \\
&= G(\omega)\, F(\omega).
\end{aligned} \tag{42}$$

Thus the Fourier transform of a convolution of two functions is the product of the Fourier transforms of the original functions. In the case of our one-dimensional example, the Fourier transform of a signal containing ten identical complicated pulses each second is simply the product of the Fourier transform of one complicated pulse and the Fourier transform of a series of delta functions occuring ten times each second.

13 Fraunhofer Diffraction

Fourier transforms are closely linked to diffraction — indeed the far-field pattern produced when an object is illuminated by monochromatic light is the Fourier transform of the pattern on the object.

To see why this is so, consider the arrangement shown in figure 1. The amplitude passing a point on the object at point y is given by the function $a(y)$. We study the wave amplitude at a point on the screen a distance Y above the axis, where the screen is a distance X from the object. We also write $R = \sqrt{X^2 + Y^2}$, and $\sin\theta = Y/R$.

The distance travelled from the point $(0, y)$ to (X, Y) is given, where we use the Binomial expansion to first order, by

$$\begin{aligned}
L &= \sqrt{X^2 + (Y-y)^2} \\
&= \sqrt{X^2 + Y^2 - 2Yy + y^2} \\
&= \sqrt{R^2 - 2Yy + y^2} \\
&= R\left(1 - \frac{2Yy}{R^2} + \frac{y^2}{R^2}\right)^{\frac{1}{2}}
\end{aligned}$$

$$\approx \; R\left(1-\frac{Yy}{R^2}\right), \tag{43}$$

where the approximation is good providing that $y \ll R$, which is known as the *far field* or Fraunhofer limit.

The contribution to the amplitude at (X, Y) from the source at $(0, y)$ will be $a(y)\, e^{ikL}$, using the nomenclature of section 1. The total amplitude at the point will be given by integrating this over all parts of the object. This gives an amplitude of

$$\begin{aligned} A(Y) &= \int_y a(y)\, e^{ikL}\, dy \\ &= \int_y a(y)\, e^{ikR\left(1-Yy/R^2\right)}\, dy \\ &= e^{ikR} \int_y a(y)\, e^{-ikyY/R}\, dy \\ &= e^{ikR} \int_y a(y)\, e^{-iky\,\sin\theta}\, dy. \end{aligned} \tag{44}$$

Now if we define the Fourier transform of $a(y)$ as

$$F(K) = \int_y a(y)\, e^{-iKy}\, dy, \tag{45}$$

then our amplitude at (X, Y) will be given by

$$A(Y) = A(R\,\sin\theta) = e^{ikR}\, F(k\,\sin\theta), \tag{46}$$

as $k\,\sin\theta$ in equation 44 takes the place of K in equation 45. We see that the pattern indeed has an amplitude proportional to the Fourier transform of the object pattern $a(y)$.

14 Diffraction Examples

The three most common slit patterns used when first studying diffraction are the double slit, the multiple slit and the single wide slit. Let us work out the Fourier transforms of these patterns to see if they do correspond to the diffraction patterns we are familiar with.

14.1 Double Slit

For the double narrow slit, we have $a(y) = A\delta(y) + A\delta(y-s)$. That means that we get lots of light through if $y = 0$ or $y = s$, but none otherwise. The variable s represents the separation of the two slits. The Fourier transform is

$$F(K) \;=\; \int_y a(y)\, e^{-iKy}\, dy$$

$$
\begin{aligned}
&= \int_y A\left\{\delta(y) + \delta(y-s)\right\} e^{-iKy}\, dy \\
&= \int_y A\delta(y)\, e^{-iKy}\, dy + \int_y A\delta(y-s)\, e^{-iKy}\, dy \\
&= A\, e^0 + A\, e^{-iKs} \\
&= A\, e^{-ikS/2} \left(e^{iKs/2} + e^{-iKs/2}\right) \\
&= A\, e^{-ikS/2} \cos\left(\frac{Ks}{2}\right). \qquad (47)
\end{aligned}
$$

For this function, we wil have full constructive interference when $Ks/2 = n\pi$, and so $K = 2n\pi/s$. Remembering from equation 46 that $K = k \sin\theta = 2\pi \sin\theta/\lambda$, this means that our condition for constructive interference is

$$
\begin{aligned}
\frac{2\pi \sin\theta}{\lambda} &= \frac{2n\pi}{s} \\
s \sin\theta &= n\lambda, \qquad (48)
\end{aligned}
$$

as we expect for the double slit. Full destructive interference first occurs when $Ks/2 = \pi/2$, so $K = \pi/s$ and we therefore have destructive interference at the places half way in between our constructive interference points.

14.2 Multiple Slits

For the case of N narrow slits, each separated by distance s. The amplitude function $a(y)$ will therefore be

$$
a(y) = \sum_{n=0}^{N-1} A\,\delta(y - ns), \qquad (49)
$$

which has a Fourier transform

$$
\begin{aligned}
F(K) &= \int_y a(y)\, e^{-iKy}\, dy \\
&= \sum_{n=0}^{N-1} A \int_y \delta(y-ns)\, e^{-iKy}\, dy \\
&= \sum_{n=0}^{N-1} A\, e^{-iKns} \\
&= A\, \frac{e^{-iKNs} - 1}{e^{-iKs} - 1} \\
&= A\, \frac{e^{-iKNs/2}}{e^{-iKs/2}}\, \frac{e^{-iKNs/2} - e^{iKNs/2}}{e^{-iKs/2} - e^{iKs/2}} \\
&= A\, \frac{e^{-iKNs/2}}{e^{-iKs/2}}\, \frac{e^{iKNs/2} - e^{-iKNs/2}}{e^{iKs/2} - e^{-iKs/2}} \\
&= A\, \frac{e^{-iKNs/2}}{e^{-iKs/2}}\, \frac{2i \sin\left(KNs/2\right)}{2i \sin\left(Ks/2\right)}
\end{aligned}
$$

$$= A \frac{e^{-iKNs/2}}{e^{-iKs/2}} \frac{\sin(KNs/2)}{\sin(Ks/2)}, \tag{50}$$

where on the fourth line, we have used the standard result for summing a finite geometric series.

When $K \to 0$, we use the small angle approximation $\sin(KNs/2) = KNs/2$, and the amplitude of equation 50 becomes AN. We have similar constructive interference whenever the numerator and denominator both become zero, namely when $Ks/2 = n\pi$, and so $K = 2n\pi/s$ as with the double slit. However, we now also have destructive interference whenever $KNs/2 = n\pi$, that is, when $K = 2n\pi/Ns$. Accordingly we will have $N-1$ points of fully destructive interference between each adjacent pair of contructive interference places. In the limit of $N \to \infty$, $F(K)$ will be zero whenever we are not at a full constructive point — and this is approximated very well by a diffraction grating.

14.3 Single wide slit

Here, we take $a(y)$ to be a 'hat' function. Namely, $a(y) = A$ if $y > 0$ and $y < w$, and $a(y) = 0$ elsewhere, where w is the width of the slit. The Fourier transform here becomes

$$\begin{aligned}
F(K) &= \int_y a(y)\, e^{-iKy}\, dy \\
&= \int_{y=0}^{w} A\, e^{-iKy}\, dy \\
&= A \left[\frac{e^{-iKy}}{-iK}\right]_{y=0}^{w} \qquad (51) \\
&= \frac{A}{-iK}\left(e^{-iKw} - e^{0}\right) \\
&= \frac{A}{-iK} e^{-iKw/2}\left(e^{-iKw/2} - e^{iKw/2}\right) \\
&= \frac{A}{iK} e^{-iKw/2}\left(e^{iKw/2} - e^{-iKw/2}\right) \\
&= \frac{2A}{K} e^{-iKw/2} \sin(Kw/2) \\
&= Aw\, e^{-iKw/2} \frac{\sin(Kw/2)}{(Kw/2)}. \qquad (52)
\end{aligned}$$

Full, constructive interference only occurs at $K = 0$ (which means $\theta = 0$), when the total amplitude will be Aw. Destructive interference occurs when $Kw/2 = n\pi$, which means that $kw \sin\theta = 2n\pi$, which in turn means $w \sin\theta = n\lambda$, as we would expect.

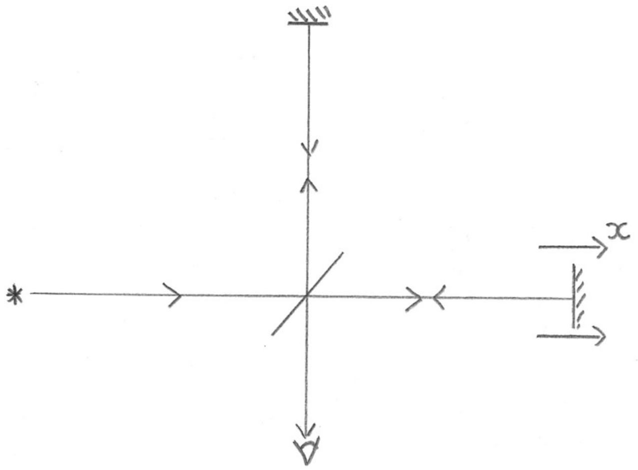

Figure 2: An interferometer. Light entering on the left will either bounce off the mirror at the top or the right before being combined. The mirror on the right can be moved in or out as desired.

14.4 Multiple wide slits

Working out the diffraction pattern for multiple wide slits is particularly easy, as in this case $a(y)$ is the convolution of the pattern for multiple narrow slits and one wide slit. This means that $F(K)$ will be the product of the transforms expected of the multiple narrow slits and the one wide slit, as proved in section 12 on the convolution theorem.

15 Fourier Transform Spectroscopy

The spectroscopy which uses a diffraction grating is very convenient for time-dependent or short-lived phenomena. The different frequencies (or colours) end up in different places on the image plane, and a camera can record their different intensities. This is known as spectroscopy by *division of wavefront* — different wavelengths are directed on different paths.

If the source is constant, but weak, an alternative method can be used using an interferometer as shown in figure 2. The amplitude $a(x)$ is recorded by the detector as a function of the position of the movable mirror, where $x = 0$ corresponds to equality of the time taken for light to reach the detector by either route and x represents the extra path which must be travelled by light passing to and from the movable mirror on its way to the detector.

To see how we can work out the wavelengths present in light using this technique, let us suppose that the amplitude of each wavenumber k in the light going to each of the two mirrors is given by $F(k)$. If the path length of light reflecting off the fixed mirror is L, the path length of the other route will be $L+x$. Using the terminology of section 1, the total disturbance on recombination from this frequency will be

$$F(k)\, e^{i\omega t-kL} + F(k)\, e^{i\omega t-ik(L+x)} = F(k)\, e^{i(\omega t-kL)} \left(1+e^{-ikx}\right), \tag{53}$$

and so the amplitude of this wavelength contributing to the signal in the sensor is $F(k)\,\left(1+e^{-ikx}\right)$.

The total amplitude from all of the frequencies present will be

$$a(x) = \int_k F(k)\,\left(1+e^{-ikx}\right)\,dk = \int_k F(k)\,dk + \int_k F(k)\,e^{-ikx}\,dk. \tag{54}$$

Notice that if we set $x=0$, then $e^{-ikx}=1$ for all values of k, and thus

$$a(0) = 2\int_k F(k)\,dk. \tag{55}$$

Equation 54 can be rewritten using this fact

$$\begin{aligned} a(x) &= \frac{a(0)}{2} + \int_k F(k)\,e^{-ikx}\,dk \\ a(x) - \frac{a(0)}{2} &= \int_k F(k)\,e^{-ikx}\,dk. \end{aligned} \tag{56}$$

If we now multiply by e^{iKx} and integrate over all x, we find

$$\begin{aligned} \int_x \left(a(x)-\frac{a(0)}{2}\right) e^{iKx}\,dx &= \int_x\int_k F(k)\,e^{-ikx}\,e^{iKx}\,dk\,dx \\ &= \int_k F(k)\int_x e^{-i(k-K)x}\,dx\,dk \\ &= \int_k F(k)\,2\pi\,\delta(k-K)\,dk \\ &= 2\pi\,F(K). \end{aligned} \tag{57}$$

Thus the amplitude of the wavenumber K in the light is given by

$$F(K) = \frac{1}{2\pi}\int_x \left(a(x)-\frac{a(0)}{2}\right) e^{iKx}\,dx. \tag{58}$$

In this type of spectroscopy, the amplitude of the light for the different positions x is measured over the course of time by recording the intensity reaching the detector as the mirror moves. These signals are given to a computer, which evaluates the integrals in equation 58 (effectively, it performs a Fourier transform) to determine the spectrum. For weak, but constant signals, this gives much better results than a diffraction grating would, as you are always working with the whole spectrum. This technique is a form of spectroscopy by *division of amplitude*.

16 Derivatives

Suppose we have a function $f(t)$, whose Fourier transform is $F(\omega)$. We now want to see what the Fourier transform of df/dt looks like. To do this, we assume that $f(t)$ is a well behaved function which reaches zero as $t \to \pm\infty$, as does its derivatives, and we integrate by parts:

$$\begin{aligned}
\mathcal{F}\left(\frac{df}{dt}\right) &= \int_t \frac{df}{dt}\, e^{-i\omega t}\, dt \\
&= \left[f(t)\, e^{-i\omega t}\right]_{-\infty}^{+\infty} - \int_t f(t)\, \frac{de^{-i\omega t}}{dt}\, dt \\
&= 0 + \int_t f(t)\, i\omega\, e^{-i\omega t}\, dt \\
&= i\omega \int_t f(t)\, e^{-i\omega t}\, dt \\
&= i\omega\, F(\omega).
\end{aligned} \tag{59}$$

.

Similarly, if we had a function $f(x)$ which was a function of x alone, then the Fourier transform of this function would be $F(k) = \int_x f(x) e^{-ikx} dx$, and we would find that the Fourier transform of df/dx would be $ikF(k)$.

17 Travelling waves

If we are working with a travelling wave $y = a(x,t)e^{i(\omega t - kx)}$ as in equation 4, then sometimes we wish to perform the Fourier transform so that it affects both of the x and t dependance. Effectively, we wish to Fourier transform the amplitude function $a(x,t)$ into a two-dimensional Fourier transform $A(k,\omega)$. This is done with a double integral:

$$A(k,\omega) = \int_t \int_x a(x,t)\, e^{-i(\omega t - kx)}\, dx\, dt. \tag{60}$$

Notice that in this context, the sign of the complex exponent of x has switched sign compared to our earlier Fourier transforms.

When we are in a three dimensional situation, we define a wavevector $\mathbf{k}$ which has magninude $k = 2\pi/\lambda$, and which points in the direction of propagation. It will have components parallel to the x, y and z axes, which are labelled k_x, k_y, k_z respectively. The disturbance g caused by a travelling plane wave would then be written

$$g = a(x,y,z,t)\, e^{i(\omega t - k_x x - k_y y - k_z z)} = a(x,y,z,t)\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{61}$$

where we can simplify the notation using the vector 'dot' or scalar product, and use $\mathbf{r}$ to represent the position vector of the point in question.

With this approach, we can now take a four-dimensional Fourier transformation of an amplitude in space and time

$$A(\mathbf{k},\omega) = \int_t \int_x \int_y \int_z a(\mathbf{r},t)\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{r})}\, dx\, dy\, dz\, dt. \tag{62}$$

18 Derivatives with Four Dimensional Transforms

Following the reasoning of equation 59, the transform of $\partial a(\mathbf{r},t)/\partial t$ will be $i\omega A(\mathbf{k},\omega)$. We can also calculate the transforms of the derivatives with respect to x, y or z using integration by parts:

$$\begin{aligned}
\mathcal{F}\left(\frac{\partial a}{\partial x}\right) &= \int_{\mathbf{r},t} \frac{\partial a}{\partial x}\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{r})}\, d^3_{\mathbf{r}}\, dt \\
&= \int_{y,z,t} \left[a(\mathbf{r},t)\, e^{-i(\omega t - \mathbf{k}\cdot\mathbf{r})}\right]_x\, dy\, dz\, dt \\
&\quad - \int_{\mathbf{r},t} a(\mathbf{r},t)\,(ik_x)\, e^{-i(\omega t - k_x x - k_y y - k_z z)}\, d^3_{\mathbf{r}}\, dt \\
&= 0 - ik_x\, A(\mathbf{k},\omega)
\end{aligned} \tag{63}$$

Similarly, the Fourier transform of $\partial a/\partial y$ will be $-ik_y A(\mathbf{k},\omega)$, and the transform of $\partial a/\partial z$ will be $-ik_z A(\mathbf{k},\omega)$.

The ∇ function is defined as a vector operator with components

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{64}$$

When we transform ∇a, we will gain the result

$$\begin{aligned}
\mathcal{F}\nabla a(\mathbf{r},t) &= -\,(ik_x,\, ik_y,\, ik_z)\; A(\mathbf{k},\omega) \\
&= -i\mathbf{k}\, A(\mathbf{k},\omega).
\end{aligned} \tag{65}$$

Up until now, we have assumed that $a(\mathbf{r},t)$ is a scalar, however we can also transform the three components of a vector field in the same way. Accordingly $\mathbf{b}(\mathbf{r},t)$ would transform to $\mathbf{B}(\mathbf{k},\omega)$.

Now that we have a vector field, we can work out the Fourier transforms of $\nabla\cdot\mathbf{B}$ and $\nabla\times\mathbf{B}$. Using the principle of equation 63, we have

$$\begin{aligned}
\mathcal{F}\nabla\cdot\mathbf{B} &= \left(\mathcal{F}\frac{\partial B_x}{\partial x} + \mathcal{F}\frac{\partial B_y}{\partial y} + \mathcal{F}\frac{\partial B_z}{\partial z}\right) \\
&= -i\,\{k_x\, B_x(\mathbf{k},\omega) + k_y\, B_y(\mathbf{k},\omega) + k_z\, B_z(\mathbf{k},\omega)\} \\
&= -i\mathbf{k}\cdot\mathbf{B}(\mathbf{k},\omega).
\end{aligned} \tag{66}$$

And, in similar vein, we have

$$\mathcal{F}\nabla\times\mathbf{B} = -i\mathbf{k}\times\mathbf{B}(\mathbf{k},\omega). \tag{67}$$

We conclude the article by giving three examples of the use of Fourier transformed equations of electromagnetism to show how such waves may progress in uniform material and in a plasma, and also to show how we might establish the nature of electrostatic waves in a plasma. If you have not seen these arguments before, they may look strange, and even complicated in places. However I think you will find them much simpler than the equivalents written as differential equations.

19 Maxwell's Equations

The power of the method can be shown in demonstrating the solution of Maxwell's equations without having to solve differential equations. The four equations, expressed in terms of the electric field $\mathbf{E}(\mathbf{r}, t)$ and the magnetic flux density $\mathbf{B}(\mathbf{r}, t)$ are

$$\nabla \cdot \mathbf{E} = \rho \tag{68}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{69}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{70}$$

$$\nabla \times \mathbf{B} = \mu \mathbf{J} + \mu\epsilon \frac{\partial \mathbf{E}}{\partial t}, \tag{71}$$

where ϵ is the electric permittivity of the medium, μ is the magnetic permeability, ρ is the free charge density, and $\mathbf{J}$ is the conduction current density.

We begin by assuming that we are in a region of space where there are no free charges or conduction currents, and so $\rho = 0$ and $\mathbf{J} = 0$.

If we Fourier transform these equations, and use the notation that the transform of $\mathbf{E}(\mathbf{r}, t)$ is $\tilde{\mathbf{E}}(\mathbf{k}, \omega)$, then we have

$$-i\mathbf{k} \cdot \tilde{\mathbf{E}} = 0 \tag{72}$$

$$-i\mathbf{k} \cdot \tilde{\mathbf{B}} = 0 \tag{73}$$

$$-i\mathbf{k} \times \tilde{\mathbf{E}} = -i\omega\tilde{\mathbf{B}} \tag{74}$$

$$-i\mathbf{k} \times \tilde{\mathbf{B}} = i\mu\epsilon\omega\tilde{\mathbf{E}}. \tag{75}$$

Equations 72 and 73 respectively tell us that there is no component of $\mathbf{E}$ or $\mathbf{B}$ parallel to the wave vector and therefore any electric or magnetic waves must be transverse.

Equation 74 tells us that $\tilde{\mathbf{B}}$ is perpendicular to both $\mathbf{k}$ and $\tilde{\mathbf{E}}$, which means that any electric wave must be polarized in a plane perpendicular to any magnetic wave. Furthermore, if the wave were propagating in the $+x$ direction, and the electric wave were polarized in the y direction, then the magnetic wave must be polarized in the z direction as $\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}}$.

Next, with the directions taken care of, equation 74 tells us that $k\tilde{E} = \omega\tilde{B}$, and thus that $\tilde{E} = \tilde{B}\omega/k = v\tilde{B}$ where v is the speed of the wave.

Finally, looking at the magnitudes of the quantities in equation 75, $k\tilde{B} = \mu\epsilon\omega\tilde{E}$. Given that $\omega = vk$ and that $\tilde{E} = v\tilde{B}$, it follows that

$$\begin{aligned} k\tilde{B} &= \mu\epsilon \times vk \times v\tilde{B} \\ 1 &= \mu\epsilon v^2 \\ v^2 &= \frac{1}{\mu\epsilon}. \end{aligned} \tag{76}$$

We can thus see that waves of all frequencies are permitted, and will all have the same speed. The use of the Fourier transform changed the partial differential equations into linear equations which were much easier to solve.

If the wave is in air, then the permittivity and permeability take their 'free space' values ϵ_0 and μ_0. As these are constants, it means that the speed $c = 1/\sqrt{\epsilon_0\mu_0}$ is the same for all frequencies of light. This is an unusually simple situation for a wave. More normally, the speed will depend on the frequency, and the formula which relates the frequency to the wavelength is known as the dispersion relationship.

20 Electromagnetic Waves in Plasma

Our first 'proper' dispersion relation is the one for an electromagnetic wave in a plasma. Plasmas contain positive and negative particles in balanced numbers. On the whole, the plasma is neutral, however the electric field of an electromagnetic wave can move the particles. For the moment, we shall neglect the motion of the positive nuclei, as they are much heavier, and tend not to move much in the time in which the waves pass.

Also, to declutter the notation, as we shall be working only in terms of Fourier transformed fields from now on, we shall omit the $\tilde{}$ above the fields such as $\tilde{\mathbf{E}}$. The fact that our equations involve $\mathbf{k}$ and ω as constants does imply we are working the amplitudes of particular frequency components, and hence with the Fourier transform of the original field.

We can use the same working as the last section, but have to make one modification. Funnily enough, we don't need to change our assumption that $\rho = 0$, for although one electron might move upwards in response to the electric field, another will move into its old position from below, so the charge at the point remains zero.

We can't assume that $\mathbf{J} = \mathbf{0}$, as the moving charges make a current. If the electrons each have a charge of q_e, and they have a number density of N_e per cubic metre, and they have an average velocity of $\mathbf{V}_e$, then we find that $\mathbf{J} = N_e q_e \mathbf{V}_e$.

This means that equation 71 now needs to be re-written as

$$-i\mathbf{k} \times \mathbf{B} = \mu_0 N_e q_e \mathbf{V}_e + \mu_0 \epsilon_0 i\omega \mathbf{E}. \tag{77}$$

We combine this with the information from equation 74 that $\omega \mathbf{B} = \mathbf{k} \times \mathbf{E}$, and so

$$\begin{aligned} -i\mathbf{k} \times \mathbf{B} &= -i\frac{\mathbf{k} \times (\mathbf{k} \times \mathbf{E})}{\omega} \\ &= -i\frac{(\mathbf{k} \cdot \mathbf{E})\,\mathbf{k} - (\mathbf{k} \cdot \mathbf{k})\,\mathbf{E}}{\omega} \\ &= i\frac{k^2}{\omega}\mathbf{E}, \end{aligned} \tag{78}$$

where we have used the vector identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\,\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\,\mathbf{c}$, and we remember from equation 72 that $\mathbf{k} \cdot \mathbf{E} = 0$. Putting equations 77 and 78 together, we obtain

$$k^2 \mathbf{E} = \frac{\mu_0 N_e q_e \omega \mathbf{V}_e}{i} + \mu_0 \epsilon_0 \omega^2 \mathbf{E}. \tag{79}$$

To finish off our reasoning, we need to relate the $\mathbf{V}_e$ to $\mathbf{E}$. To do this, we use Newton's second law. The vector equation for the force on an electron is $\mathbf{F} = q_e(\mathbf{E} + \mathbf{V}_e \times \mathbf{B})$. However, we are currently assuming that the electromagnetic wave is not terribly strong, and accordingly that the electron won't be going an appreciable fraction of the speed of light. Equation 74 gives us $kE = \omega B$, which means that $B/E = k/\omega = 1/c$. So as long as the electron is slow, the force caused by the magnetic part of the electromagnetic wave can be ignored, and we simply write $\mathbf{F} = q_e \mathbf{E}$. The acceleration of the electron will be $\mathbf{A}_e = \mathbf{F}/m_e = q_e \mathbf{E}/m_e$.

Now the acceleration is also given by the time derivative of velocity, so $\mathbf{A}_e = \partial \mathbf{V}_e / \partial t = i\omega \mathbf{V}_e$. Putting together these two expressions for the acceleration gives us

$$i\omega \mathbf{V}_e = \frac{q_e \mathbf{E}}{m_e} \tag{80}$$

$$\mathbf{V}_e = \frac{q_e \mathbf{E}}{i\omega m_e}. \tag{81}$$

Putting together equations 79 and 81, we achieve

$$\begin{aligned} k^2\mathbf{E} &= \frac{\mu_0 N_e q_e \omega}{i} \frac{q_e \mathbf{E}}{i\omega m_e} + \mu_0\epsilon_0\omega^2\mathbf{E} \\ &= \frac{\mu_0 N_e q_e^2 \omega \mathbf{E}}{i^2 \omega m_e} + \mu_0\epsilon_0\omega^2\mathbf{E}. \end{aligned}$$

We now remember that $\mu_0\epsilon_0 = c^{-2}$ and divide all terms by this factor, and also by $\mathbf{E}$. We also note that therefore $c^2\mu_0 = \epsilon_0$, and remember that $i^2 = -1$ by definition. This gives

$$\begin{aligned} c^2k^2 &= -\frac{c^2\mu_0 N_e q_e^2 \omega}{\omega m_e} + \omega^2 \\ \omega^2 &= \frac{N_e q_e^2}{\epsilon_0 m_e} + c^2k^2 \\ \omega^2 &= \omega_p^2 + c^2k^2, \end{aligned} \tag{82}$$

where we define the electron plasma frequency $\omega_p = \sqrt{N_e q_e^2/\epsilon_0 m_e}$. This is the frequency at which the electrons will naturally rock back and forth in a plasma, and accordingly, if you wish to form a wave, you will need to add more energy to allow the oscillation to move forward. This necessitates a frequency higher than ω_p. Usually, however, electromagnetic radiation doesn't simply just find itself in a plasma with a high ω_p. Instead, there is a 'density gradient' with the plasma having a soft edge. The light reaches the plasma, and the wavelength gets bigger as the plasma gets denser (N_e goes up, so ω_p goes up, and so for the same ω, the value of k has to go down, causing an increase in λ). If it gets too dense, the light reflects in a process analagous to total internal reflection. The maximum density which the light can reach, where ω_p is equal to the frequency of the light, is called the *critical density* N_c. Given that $\omega_p \propto \sqrt{N_e}$, it follows that $N_e/N_c = \omega_p^2/\omega^2$. You will hear plasma physicists talking about the 'quarter critical' surface, where $N_e = \frac{1}{4}N_c$, and accordingly $\omega_p = \frac{1}{2}\omega$.

21 Electrostatic Electron waves in Plasma

After the electromagnetic wave, which is extremely important in laser-produced plasma, the next most important kind of wave is the electrostatic electron wave sometimes known as a Langmuir wave. This is where you get bands in the plasma, some with a net positive charge (too few electrons) and some with net negative charge. There is a restoring force pulling the electrons back to the positive regions just as gravity tries to pull sea water down from the top of surf wave crests to the troughs, and if you set it up correctly, a wave will travel. Given that the electric field is responsible for this force, and this field pulls the electrons back and forth in the direction of the wave, this is a longitudinal wave.

We can use a very similar analysis to our last section, but as before, we do need to amend our assumptions. Firstly, we can no longer assume that $\rho = 0$. We use the letter n_e to represent the electron density in a region (in electrons per cubic metre) relative to the density needed for neutrality. Therefore $n_e \neq N_e$. If the charge on each electron is q_e, then it follows that $\rho = q_e n_e$, and this makes equation 72 become

$$-i\epsilon_0 \mathbf{k} \cdot \mathbf{E} = q_e n_e. \tag{83}$$

But if this makes things harder, there is a consequent simplification. Equation 74 tells us that if $\mathbf{k}$ is to be parallel to $\mathbf{E}$ as is required for our longitudinal wave, then we can expect $\mathbf{B} = 0$.

Accordingly, equation 79 loses its initial k^2 term, and instead looks like this:

$$0 = \frac{\mu_0 N_e q_e \omega \mathbf{V}_e}{i} + \mu_0 \epsilon_0 \omega^2 \mathbf{E}. \tag{84}$$

We do, however, have more complications when we start considering Newton's second law for the electrons. Now they are not only forced by the electric field, they also respond to their own mutual pressure which won't let them all bunch up in one place. In addition, the electric field is now self generated from the variations in n_e rather than imposed by an electromagnetic wave.

Let's start with that pressure. Just as with an ideal gas, we can write $PV = Nk_BT$ where k_B is the Boltzmann constant of 1.38×10^{-23} J/K (and has no relation to any wave vectors at all), and N is the total number of particles. We prefer to write $N_e = N/V$ as a number density, and accordingly, the expected pressure $P = N_e k_B T$. Another preference of plasma physicists is to write the pressure in terms of the root mean square speed of the electrons $v_{e,th}$ due to the thermal motion, where we define $m_e v_{e,th}^2 = k_B T$. The upshot of these conventions is that $P = N_e m_e v_{e,th}^2$.

Of course, the actual pressure will go up and down as the wave passes, and we use the letter p to represent the pressure at a particular place above the equilibrium pressure P. It is p which forces the electrons back and forth. Newton's second law for the situation is $\mathbf{F} = m_e \mathbf{A} = q_e \mathbf{E} - \nabla p / N_e$, where the negative gradient function 'pushes' the electrons from regions of large to small p. The division by N_e is necessary as ∇p gives the force on a whole cubic metre of plasma, so we divide by N_e to get the force on one particle. As before, $\mathbf{A} = i\omega \mathbf{V}_e$, and so

$$im_e \omega \mathbf{V}_e = q_e \mathbf{E} + \frac{i\mathbf{k}p}{N_e}$$

$$\mathbf{V}_e = \frac{q_e}{im_e\omega}\mathbf{E} + \frac{\mathbf{k}}{m_e\omega N_e}p. \tag{85}$$

We can't put equation 85 together with 84 to form a dispersion relation until we have related p to $\mathbf{E}$. We do this via the electron density n_e. Firstly, for any given situation, it transpires that $PV^\gamma \propto PN_e^{-\gamma}$ is a constant for some value of γ. If things are happening really slowly, we would expect temperatures to be even, and given that this requires PV to be constant, $\gamma = 1$. This is called the isothermal limit. The opposite extreme, where thermal energy doesn't have time to flow down the wave from one part to another is called the adiabatic limit, and for this $\gamma = 3$, a fact I am not going to prove here.

Let's imagine we start with the equilibrium situation with pressure P and number density N_e. Now we increase the number density to $N_e + n_e$, and see what effect this has on the pressure. If $PN_e^{-\gamma}$ is a constant, we can write the new pressure $P + p$ as

$$\begin{aligned} P + p &= \frac{PN_e^{-\gamma}}{(N_e + n_e)^{-\gamma}} \\ &= \frac{P}{\left(1 + \frac{n_e}{N_e}\right)^{-\gamma}} \\ &= P\left(1 + \frac{n_e}{N_e}\right)^{\gamma} \\ &\approx P\left(1 + \gamma\frac{n_e}{N_e}\right), \end{aligned}$$

where we use the first two terms of the binomial expansion to produce the approximation. We can now calculate the pressure increase p as

$$\begin{aligned} p &\approx \frac{\gamma n_e}{N_e}P \\ &\approx \frac{\gamma n_e}{N_e}N_e m_e v_{e,th}^2 \\ &\approx \gamma m_e v_{e,th}^2 n_e. \end{aligned} \tag{86}$$

Combining equations 85 and 86 we have

$$\mathbf{V}_e = \frac{q_e}{im_e\omega}\mathbf{E} - \frac{\mathbf{k}}{m_e\omega N_e}\gamma m_e v_{e,th}^2 n_e. \tag{87}$$

We now use equation 83 to replace n_e with a function of $\mathbf{E}$, giving

$$\begin{aligned} \mathbf{V}_e &= \frac{q_e}{im_e\omega}\mathbf{E} - \frac{\mathbf{k}}{m_e\omega N_e}\gamma m_e v_{e,th}^2 \frac{i\epsilon_0(\mathbf{k}\cdot\mathbf{E})}{q_e} \\ &= \frac{q_e}{im_e\omega}\mathbf{E} - \frac{i\epsilon_0\gamma v_{e,th}^2(\mathbf{k}\cdot\mathbf{E})}{q_e\omega N_e}\mathbf{k}. \end{aligned} \tag{88}$$

Now, this expression for $\mathbf{V}_e$ is substituted into 84 (at last) to give

$$\begin{aligned}
0 &= \frac{\mu_0 N_e q_e \omega}{i}\left(\frac{q_e}{i m_e \omega}\mathbf{E} - \frac{i\epsilon_0 \gamma v_{e,th}^2(\mathbf{k}\cdot\mathbf{E})}{q_e \omega N_e}\mathbf{k}\right) + \mu_0\epsilon_0\omega^2\mathbf{E} \\
\mu_0\epsilon_0\omega^2\mathbf{E} &= -\frac{\mu_0 N_e q_e^2}{i^2 m_e}\mathbf{E} + \mu_0\epsilon_0\gamma v_{e,th}^2(\mathbf{k}\cdot\mathbf{E})\mathbf{k} \\
\omega^2\mathbf{E} &= \frac{N_e q_e^2}{\epsilon_0 m_e}\mathbf{E} - \gamma v_{e,th}^2(\mathbf{k}\cdot\mathbf{E})\mathbf{k} \\
\omega^2\mathbf{E} &= \omega_p^2\mathbf{E} + \gamma v_{e,th}^2(\mathbf{k}\cdot\mathbf{E})\mathbf{k}. \qquad (89)
\end{aligned}$$

To complete the dispersion relation, we remember that as a longitudinal wave, $E_\perp = 0$ and accordingly, $(\mathbf{k}\cdot\mathbf{E})\mathbf{k} = k^2 E_\parallel$. It follows that we can cancel the factor of $E = E_\parallel$ leaving us with

$$\omega^2 = \omega_p^2 + \gamma v_{e,th}^2 k^2, \qquad (90)$$

which is called the Bohm-Gross dispersion relationship.