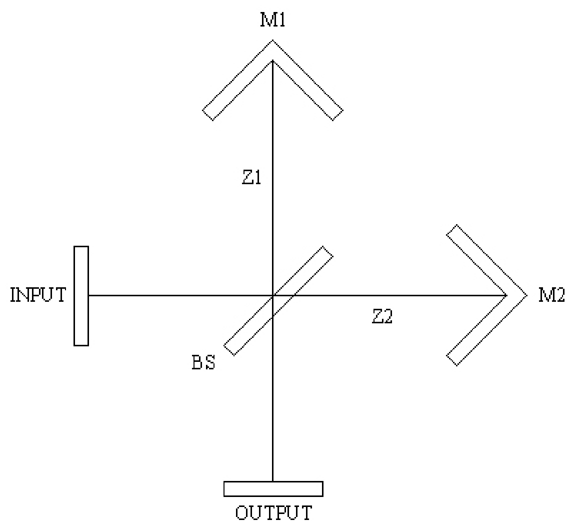
Mathematics of Fourier Transform Spectroscopy

Introduction:

We use a Martin-Puplett interferometer to obtain spectra of input sources. The first thing we are interested in is the expected power output of the detector given a monochromatic input source. From here it is easy to generalize to the more physically realistic case of a continuous input source. The next step is the mathematics of the Fourier transform itself, which allows us to determine the spectrum of the source given the value of the modulated power output at the detector. Finally, returning to the idealized monochromatic source we use a clever mathematical trick to obtain the spectral resolution of the interferometer setup.

Section 1: Derivation of Expected Power Output



Before we begin with the math let's quickly go over the experimental setup. The beam enters the interferometer at the input and travels to the beamsplitter. At the beamsplitter, a fraction of the beam is reflected toward Mirror 1 and the remaining fraction is transmitted toward Mirror 2. (Exactly which part is reflected and which is transmitted has to do with the polarization of the input source which is dealt with in Brit's paper and not here.) At the mirrors, each beam undergoes a 90 degree shift in polarization and reflects back toward the beamsplitter. The effect of this polarization shift is that the beam which was previously reflected by the BS (the one which traveled along Z1) is now transmitted and the part of the beam which was originally transmitted (the one that traveled along Z2) is now reflected. (Again, see Brit's paper for details on polarization). The result is that ALL of the beam exits through the output port. (Probably the easiest way to see this is simply to trace out the path on the diagram above.)

But why is this at all useful? Let's start with the case in which path Z1 is equal to path Z2. In that case, the two beams traveling along their respective paths will be *in phase* with each other at the output port and will therefore constructively interfere. If, on the other hand, the two paths differ by half a wavelength, the two beams will be *out of phase* with each other and will destructively interfere. The trick of the Fourier Transform Spectrometer is that one path (let's say Z1) is held constant while the other path (Z2) is varied. By observing how the intensity at the output port varies with the path difference we can effectively deduce the spectrum of the source. This is accomplished by a clever mathematical device known as the Fourier Transform.

Now for the math. Let's assume for simplicity that the input is a monochromatic source; later we'll generalize this result to the more physically realistic case of a continuous input spectrum. Our input source has the form:

$$E_{MS} = A_0 \ \mathrm{Sin}\ (2\,\pi\, f_o\, t)$$

The beam enters through the input port and travels toward the beamsplitter. The reflected part of the beam travels toward M1 and is reflected back toward the beamsplitter. Just before returning to the beamsplitter, the beam now has the form:

$$E_R = \left| r \right| A_0 \ \mathrm{Sin}\left(2\,\pi\, f_o\, t + \theta_r - \frac{2\,\pi\, f_o\, Z1}{c}\right)$$

where |r| is the reflection coefficient and $ø_r$ gives the phase change introduced upon reflection at the BS. The last term in the argument of the Sin function accounts for the phase change which occurs due to traveling along path Z1 and back.

At this point you may be wondering why I have not also included the phase change which occurs as a result of traveling from the input source to the BS. The reason is that both the reflected and transmitted beams travel this path, and since ultimately I am concerned with the *difference* between the two beams, I don't need to account for this phase change. (I could include a term for this phase change as well but it would end up subtracting out in the end, so let's leave it out for simplicity. I will do the same thing with the path from the BS to the output port.)

Similarly, the part of the beam that is transmitted, after traveling to Mirror 2 and returning to the beamsplitter will produce an E-field:

$$E_T = \left| t \right| A_0 \ \mathrm{Sin}\left(2\,\pi\, f_o\, t + \theta_t - \frac{2\,\pi\, f_o\, Z2}{c}\right)$$

where again, |t| denotes the transmission coefficient and $ø_t$ denotes the phase change introduced by the beamsplitter.

Recall that the effect of the mirrors is to introduce a 90 degree polarization shift. This means that the beam that was originally reflected (the one which traveled along path Z1 to Mirror1 and back) will now be transmitted through the beamsplitter and travel toward the detector, giving rise to an E-field at the output detector:

$$E_{RT} = |r|\ |t|\ A_0 \operatorname{Sin}\left(2\pi f_o t + \theta_r + \theta_t - \frac{2\pi f_o Z1}{c}\right)$$

(Really all I have done to obtain this expression is to multiply E_R by |t| and include the phase shift $\varnothing_t$).

Similarly, the beam that was originally transmitted (the one which traveled along path Z2 and back) will now be reflected:

$$E_{TR} = |t|\ |r|\ A_0 \operatorname{Sin}\left(2\pi f_o t + \theta_t + \theta_r - \frac{2\pi f_o Z2}{c}\right)$$

The total E-field will be the sum of these two fields. Before proceeding, let's make a couple of assumptions for simplicity.

1) $|r|^2 = |t|^2 = 1/2$

This is justified because we know that $|r|^2 + |t|^2 = 1$, since we are dealing with a loss-free system, combined with the fact that the efficiency of a beamsplitter is maximized when it divides the beam evenly, ie. when $|r| = |t|$.

2) $\varnothing_r = \varnothing_t = 0$

The power output, which represents a time average of the modulating field, will ignore these phase changes anyway, so we might as well set them to zero.

Substituting these values into the expressions for E_{TR} and E_{RT}, we can add the two and use standard trigonometric identities to obtain:

$$E_{total} = A_0 \operatorname{Sin}\left(2\pi f_o t - \frac{\pi f_o (Z1 + Z2)}{c}\right) \operatorname{Cos}\left(\frac{\pi f_o (Z1 - Z2)}{c}\right)$$

To get the time-dependent power we simply square this expression. The next step is to average over the time-dependent Sin term in order to obtain the mean power at the detector. The result is:

$$P_{ave}(\Delta Z) = \frac{A_0^2}{4}\left(1 + \operatorname{Cos}\left(\frac{2\pi f_o \Delta Z}{c}\right)\right)$$

where I have again used standard trigonometric identities to get rid of the Cos^2 term and I have replaced the path difference Z1-Z2 with ΔZ.

The above expression depends on three values: the two constants A_0 and f_0 which are determined by the input source, and the variable path length ΔZ.

The expression above is exactly what we are looking for: the mean power output at the detector given a monochromatic source. Using the same argument as above, it is easy[1] to see that a source with a continuous spectrum of the form:

$$E_{cont} = \int_{f1}^{f2} A(f)\, \mathrm{Sin}(2\pi f t)\, df$$

will produce the following power output:

$$P_{ave}(\Delta Z) = \int_{f1}^{f2} \frac{A^2(f)}{4} \left(1 + \mathrm{Cos}\left(\frac{2\pi f \Delta Z}{c}\right)\right) df$$

The difference here is that the amplitude A(f) is now non-zero for a range of frequency values instead of just for one frequency, which means that we must perform an integration over this range to obtain the mean power output. Recall that the mean power output for a monochromatic source depends on two *constants* A and f_0. The generalized power function directly above depends on some *function* $A^2(f)$, which is precisely the function that gives the *spectrum* of the input source. This spectrum is exactly what we are looking for, since after all we are performing *spectroscopy*. We will soon see that by measuring the mean power output $P(\Delta Z)$ we can work backwards (via the magic of Fourier transformation) to obtain the spectrum function $A^2(f)$.

Section 2: Mathematics of Fourier Transforms

Before we can apply the Fourier transform, we need to obtain the power output in a slightly different form. First, we note that the interferometer setup is symmetrical; that is, the choice of input and output ports is determined by convention and is not inherent in the design. The effect of this symmetry is to change the power output to the following form:

$$P_{AVE}(\Delta Z) = \int_{f1}^{f2} \frac{A^2(f)}{4} \left(1 + F\, \mathrm{Cos}\left(\frac{2\pi f \Delta Z}{c}\right)\right) df$$

where F equals plus or minus one depending on the choice of input and output ports.
Next we take the difference between these two formulations:

$$P(\Delta Z) = P_{AVE}(F = +1) - P_{AVE}(F = -1)$$

$$P(\Delta Z) = \int_{f1}^{f2} \frac{A^2(f)}{2}\, \mathrm{Cos}\left(\frac{2\pi f_o \Delta Z}{c}\right) df$$

---

[1] Well, maybe it's not *easy* to see, per se. But it makes sense if you consider a continuous spectrum to be the sum of a series of monochromatic sources each with a different amplitude A(f).

Now that we have the power output in this form we are ready to perform the Fourier transform[2].

If some function A(x) can be expressed as:

$$A(x) = \int_{y1}^{y2} B(y)\, \mathrm{Cos}(ky)\, dy$$

then we can apply the cosine Fourier transform to obtain:

$$B(y) = \lim (D \to \infty)\ \frac{1}{D} \int_{-D}^{D} A(x)\, \mathrm{Cos}(kx)\, dx$$

Applying the cosine Fourier transform to the expression for P(ΔZ) we obtain:

$$A^2(f) = \frac{2}{D} \int_{-D}^{D} P(\Delta Z)\, \mathrm{Cos}\left(\frac{2\pi f \Delta Z}{c}\right) d\Delta Z$$

Bingo- this is exactly what we need to calculate the spectrum of the source, which is given by A^2(f) (the power intensity of each spectral component f) in terms of P(ΔZ) (the measured power output of the detectors).

To obtain the spectrum function A^2(f) we do the following:

1) Measure the mean power output (PΔZ) over as large a range of path difference as possible. This corresponds to making D as large as possible, as we shall soon see.

2) Plot the power output as a function of path difference ΔZ.

3) Fit a function to the results. Input this function into P(ΔZ) in the above expression and integrate. This gives the desired output, which is the spectrum function A^2(f).

At this point you may be wondering what "D" represents. Since it appears as the limit of integration, it represents the total range of path length difference used in the experiment. In other words, D is the total distance that the second mirror is moved while a measurement of the power output is being taken. The exact Fourier transform (given above for B(y)) assumes that D tends to infinity, which is of course impossible to achieve in a physical experiment. We want D to be as large as possible however, and we will soon see that the effect of making D large is to increase the spectral resolution of the spectrometer (more on this in the next section)

---

[2] To be perfectly honest, I am not sure why this step is justified, by which I mean that I do not really understand why it corresponds to anything physically meaningful. I do understand, however, why it is necessary, since it is possible to perform a Fourier transform on the expression directly above, but the previous expression (the one that included a 1+Cos term) is not Fourier transformable. So think of this simply as a necessary mathematical manipulation which we use to get the expression in a specific form on which we can perform the Fourier transform.

Section 3: Calculation of Spectral Resolution

To calculate the spectral resolution of the FTS we will use a clever trick. Basically what we need is the spectrum (given of course by $A^2(f)$) that we would obtain for an idealized monochromatic source. If Fourier Transform Spectroscopy were perfect, then a monochromatic source would yield a spectrum which is non-zero only at the frequency f_0 (a Dirac spike). But because the total path length difference D is necessarily finite, what we actually obtain is a power spectrum peaked at f_0 which falls off quickly as you move away from the central frequency.

First let's figure out what the measured spectrum $A^2(f)$ will look like given a monochromatic source. Recall that the power output corresponding to a monochromatic source can be expressed as:

$$P(\Delta Z) = \frac{A_0^2}{2} \cos\left(\frac{2\pi f_o \Delta Z}{c}\right)$$

Plugging this in to:

$$A^2(f) = \frac{2}{D} \int_{-D}^{D} P(\Delta Z) \cos\left(\frac{2\pi f \Delta Z}{c}\right) d\Delta Z$$

yields:

$$A^2(f) = \frac{A_0^2}{D} \int_{-D}^{D} \cos\left(\frac{2\pi f_o \Delta Z}{c}\right) \cos\left(\frac{2\pi f \Delta Z}{c}\right) d\Delta Z$$

Applying standard trig identities and integrating we obtain:

$$A^2(f) = \frac{A_0^2}{2D}\left[\frac{c}{2\pi(f_0 - f)} \sin\left(\frac{4\pi D(f_o - f)}{c}\right) + \frac{c}{2\pi(f_0 + f)} \sin\left(\frac{4\pi D(f_o + f)}{c}\right)\right]$$

By this point may have lost track of the definitions of f_0 and f. f_0 is a constant denoting the *actual* frequency of the monochromatic source, and f is a variable denoting the *measured* frequency range. Since f_0 is approximately equal to f (otherwise our spectrometer wouldn't be of much use) we notice that

$$\frac{c}{2\pi(f_0 - f)} \gg \frac{c}{2\pi(f_0 + f)}$$

so the first term in brackets dominates and we can neglect the second term.

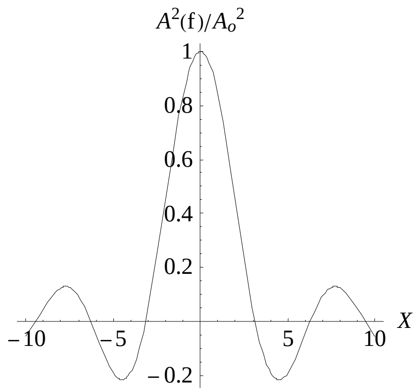
If we let

$$x = \frac{4\,\pi\,D\,(f_o - f)}{c}$$

we have:

$$A^2\,(f) = A_0{}^2\,\frac{\mathrm{Sin}\,(x)}{x}$$

or:



The Sin(x)/x term acts as an envelope function, since the function is greatest in the range:

$$|x| < \frac{\pi}{2}$$

Outside of this range the function goes quickly to values which oscillate with small amplitude about zero. So, as a first-order approximation, we can assume that A^2(f)=0 for x outside of this range. In other words, if the following condition is satisfied:

$$\left|\frac{4\,\pi\,D\,(f_o - f)}{c}\right| > \frac{\pi}{2} \quad \rightarrow \quad |f_o - f| > \frac{c}{8\,D}$$

then A^2(f)=0. If this condition is not satisfied and f is closer than c/8D to f_0, then we will obtain a non-zero value of the spectrum A^2(f).

The conclusion is that a monochromatic input source will yield a measured spectrum of width c/4D centered at f_0. This means that the spectrometer will be unable to resolve two spectral components which are closer to each other than c/4D. In other words, c/4D is the *spectral resolution* of the FTS.

$$R = \frac{c}{4\,D}$$