

Fourier Analysis

by Dr. Colton, Honors 225 (last updated: Fall 2019)

This document is intended to replace *Roots* Ch 2 Appendix FOU. That appendix has a lot of important, even vital, information for this class; however, it also suffers from being quite hard to understand. In this document I am attempting to convey the same information but in a more understandable way.

What is a frequency spectrum?

You are probably familiar with stereo displays which graphically illustrate how much sound is coming out in treble and bass frequencies. And you've probably also seen "graphic equalizers" such as this one that can be used to adjust the amount of treble, bass, and mid-range frequencies present in an audio signal.

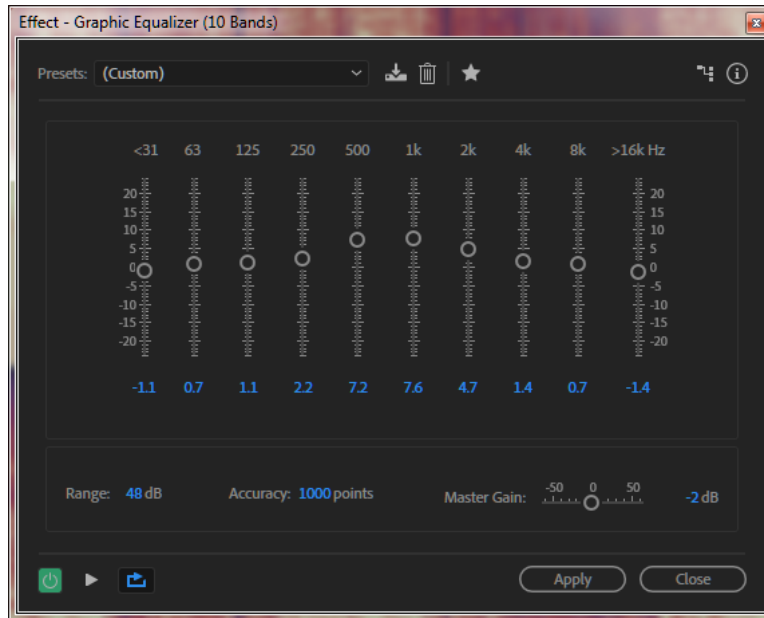


Figure 1. A 10-band graphic equalizer, from Adobe Audition software.

Inherent in displays and equalizers like that is the notion that sound waves—and indeed all waves—can be decomposed into their frequency components. The *frequency spectrum* of a wave is a plot describing how much of each frequency is present. More specifically, it indicates the amplitudes of each of the frequency components which add together to form the wave.

Here's an example of a frequency spectrum:

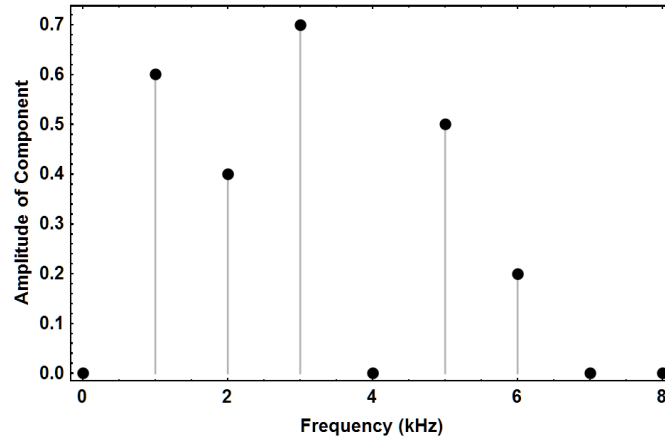


Figure 2. Frequency spectrum example.

For this particular case, the wave is formed by adding together only these five frequencies: 1, 2, 3, 5, and 6 kHz, with 3 kHz being the strongest frequency and 6 kHz being the weakest. No other frequencies are present in the wave. Four other points are explicitly plotted as zero amplitude (0, 4, 7, and 8 kHz), but there are an infinite number of frequencies not explicitly indicated, which are also zero.

In general, the frequency spectrum could include a finite number of discrete values like this case, or an infinite number of discrete values, or even an infinite number of values that form a continuum (a band of frequencies). When you have a *periodic* wave, meaning a wave that repeats its shape every so often in time, you get a set of discrete frequencies that are multiples of a what's called the *fundamental frequency* of the wave. The fundamental frequency is equal to the inverse of the period: $f_{fund} = 1/T$. The higher frequencies in that case are called the *overtones* or *higher harmonics*. (The fundamental frequency is defined to be the first harmonic.)

The oscillatory pattern which is obtained by adding together these five discrete harmonics from the frequency spectrum above is depicted as the “Clarinet tone” in part (b) of this figure from *Roots*. It corresponds to a clarinet playing a note just a little sharp of B5 (almost two octaves above middle C)

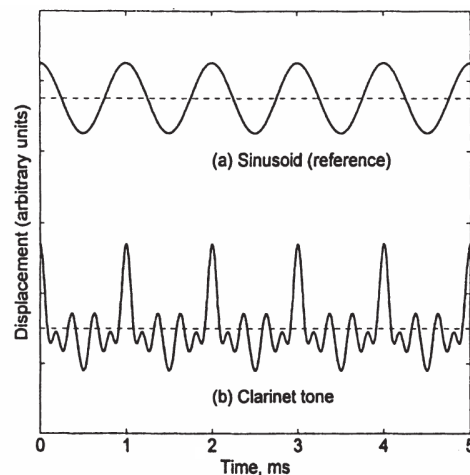


Figure 3. Tone of a clarinet showing harmonic structure (b), compared with a cosine wave at the same frequency (a). The “clarinet tone” is the oscillatory wave whose frequency spectrum is shown in Fig. 2. It has a period of 1 ms and a fundamental frequency of 1 kHz. From *Roots*, Fig. A2.1.

A pure sine or cosine wave, such as what is plotted in Fig. 3(a), has only a single frequency component. Its frequency spectrum has a single spike at that frequency (e.g. 1 kHz) and is zero everywhere else.

Any periodic wave can be decomposed into a sum of sine and cosine wave components; determining those components for a given wave is called *Fourier analysis*, and there are mathematical formulas (beyond the scope of this class) which allow you to do those calculations. Technically you actually need two graphs to depict the frequency spectrum of a wave—one for the amplitudes of the sine components and one for the cosines—but in cases where the wave has odd or even symmetry only one of the two is needed. For the clarinet example above, the time-dependent signal has even symmetry so only cosine components are used.

A different clarinet, or a different instrument altogether, would have a different oscillatory pattern, and also a different frequency spectrum. That explains how two instruments playing the same pitch can have a difference in tone quality: the pitch is determined by the fundamental frequency and the tone quality by the higher harmonics. The overall repetition period of the two sounds will be the same but the fine details of the oscillatory pattern will be different.

As mentioned above, the frequency spectrum depicts the *amplitude* of the frequency components needed to be combined together to form the wave. This could be referring to pressure amplitude in the case of a sound wave, voltage amplitude in the case of the electrical signal going to the speakers to create that sound wave, displacement amplitude in the case of a wave on a rope, and so forth.

In contrast, the amount of *energy* or *power* (power is energy per time) involved in creating the wave depends not just on the amplitude but on the amplitude squared. All other things being equal, a 2 V amplitude electrical wave requires 4× the energy to create compared to a 1 V amplitude electrical wave. Therefore a frequency *power spectrum* is often plotted instead of the amplitude spectrum. A power spectrum depicts the amount of power present in each frequency component; each point in the power spectrum is proportional to the amplitude of that frequency, squared.

The power spectrum of the clarinet tone from above looks like this; compare with Fig. 2. There are again only five frequencies with non-zero powers.

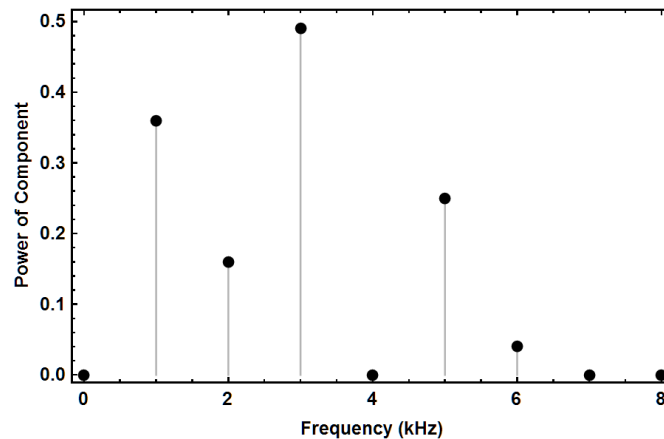


Figure 4. Power spectrum of clarinet tone.

Example: a “shuttered” wave

Let’s depart from sound at this point and move on to an important example from optics. Suppose you have a laser which produces a light wave (where the electric field oscillates sinusoidally in time), and you use a shutter to rapidly block and unblock the laser to form pulses. The wave might look like this:

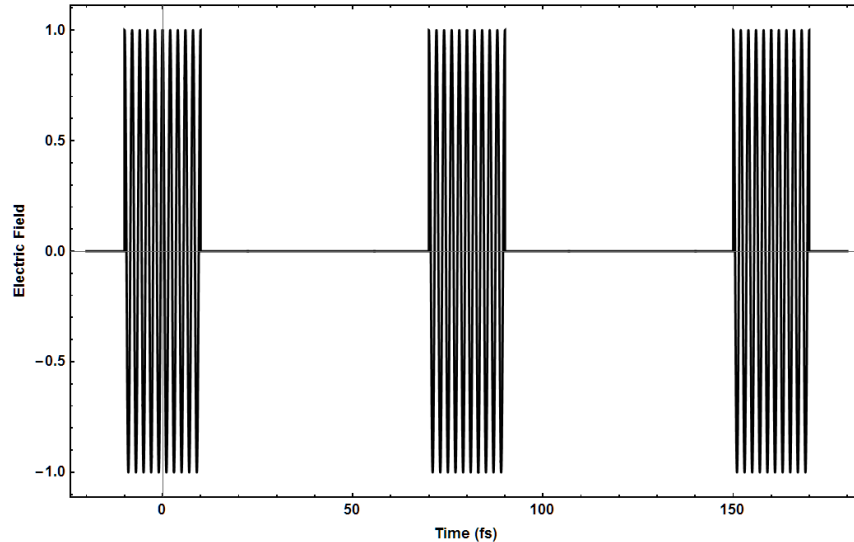


Figure 5. A shuttered wave example. Since the electric field in a laser beam oscillates incredibly quickly, the times are given in femtoseconds ($1 \text{ fs} = 10^{-15} \text{ s}$). A mechanical shutter could not possibly open and close this quickly, but let’s not worry about that.

Incidentally the envelope of a wave like that, where it switches from on and off (and vice versa) with instantaneous ramping up or ramping down, is called a “square pulse”.

There are three important times involved:

- Period of oscillation of the laser’s electric field
 - Let’s call it t_{fast} because it’s the fastest of the three times.
- Pulse width time
 - Let’s call it Δt because it represents the difference between the moment in time when the shutter opens and the moment when it closes.
- Overall repetition period
 - Let’s call it T because the symbol T is commonly used to designate periods. It’s the longest of the three times.

For the specific case plotted above, I used the following values; make sure you can identify how all three times impact the graph:

- $t_{fast} = 2 \text{ fs}$
- $\Delta t = 20 \text{ fs}$
- $T = 80 \text{ fs}$

A t_{fast} value of 2 fs corresponds to a red laser whose wavelength is 600 nm; its frequency (electric field oscillation rate) would be $5 \times 10^{14} \text{ Hz}$. Here is a plot of the power spectrum for that shuttered wave:

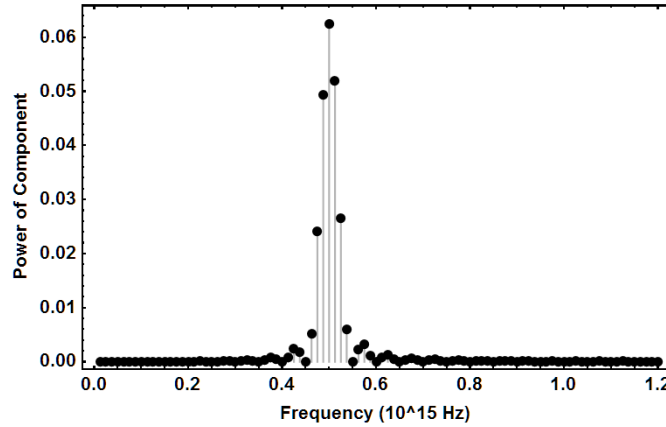


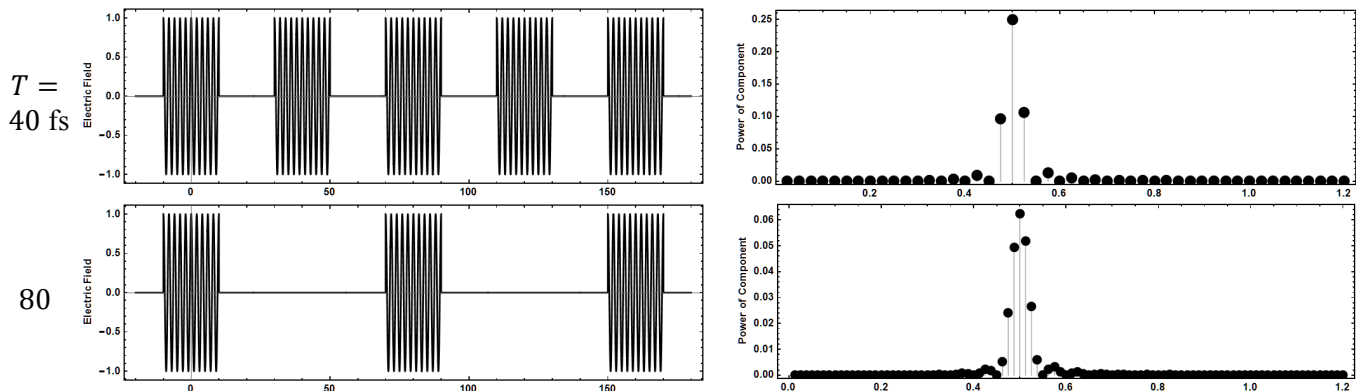
Figure 6. Power spectrum of the shuttered wave example.

Recall that all of the frequencies which are not plotted, in between the points which are plotted, have zero power. You can see that the largest amount of power, the largest *Fourier component* as it's called, is at 0.5×10^{15} Hz. That's the frequency of the laser's electric field. If there were no shutter, that is the only frequency that would exist. Strangely, pulsing the laser on and off in this manner has substantially changed its power spectrum! Instead of being a single peak at 0.5×10^{15} Hz, as would be the case for an unshuttered laser, several other frequency components are present—some that are higher than the laser's own frequency and some that are lower. This is a real effect that actually is used in some experiments I'm familiar with! Photons really do exist at those other frequencies. Pretty crazy.

Changing the overall repetition period

Now let's investigate what happens if we change the overall repetition period. I'll plot things for three different T values (the middle of which is the same as Figs. 5 and 6), leaving everything else the same.

- $t_{fast} = 2$ fs
- $\Delta t = 20$ fs
- $T = 40, 80, 160$ fs



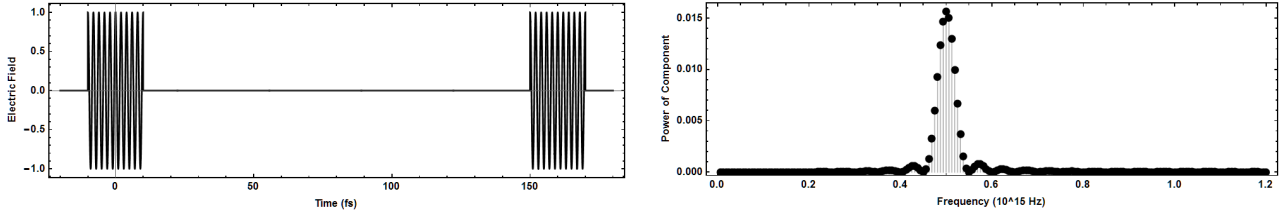


Figure 7. Waves and power spectra for three different periods. Notice how the spacing between frequency points gets smaller and smaller as the period gets longer and longer. The wave with the longest T contains the most closely spaced frequencies.

Key observation: look at the power spectra. Notice how the shortest T (top graph) has the widest gaps between frequency points and the longest T (bottom graph) has the smallest gaps between frequency points. **The spacing between frequency points is inversely proportional to the repetition period.** In fact, if you examine the power spectra really closely, you can deduce this fact, which I will box:

$$\text{frequency spacing} = \frac{1}{T}$$

Test it out! Is the spacing between frequency points in the top graph equal to $\frac{1}{40 \times 10^{-15}}$ Hz?

In the limit as the repetition period goes infinite—that is, if you have a single pulse that is never followed by a second one—the frequency points become spaced infinitely close together, and the wave now possesses a *continuum* of frequencies rather than multiple discrete frequencies.

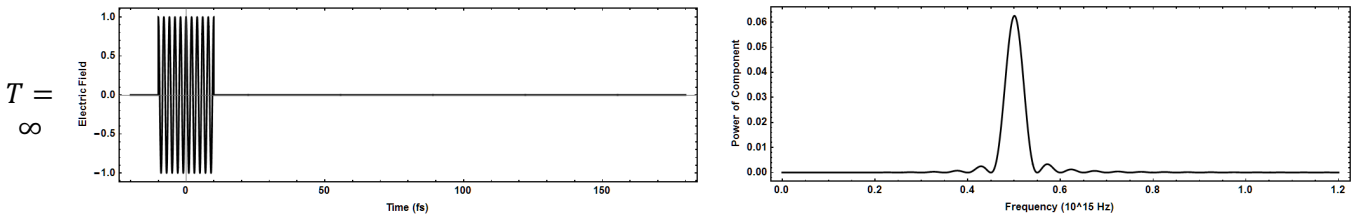
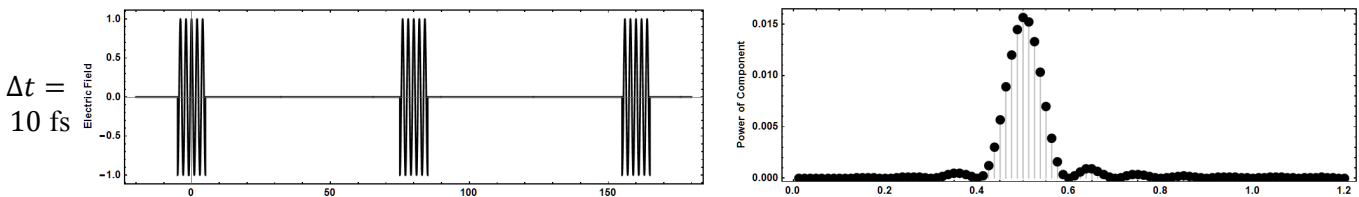


Figure 8. Wave and power spectrum for a *single* laser pulse instead of a train of pulses. The frequencies exist in a continuum instead of only at discrete points.

Changing the pulse width

Finally, let's investigate what happens if we change the pulse width (shutter open time). Here are three different Δt values, the middle of which is the same as in Figs. 5–8.

- $t_{fast} = 2$ fs
- $\Delta t = 10, 20, 40$ fs
- $T = 80$ fs



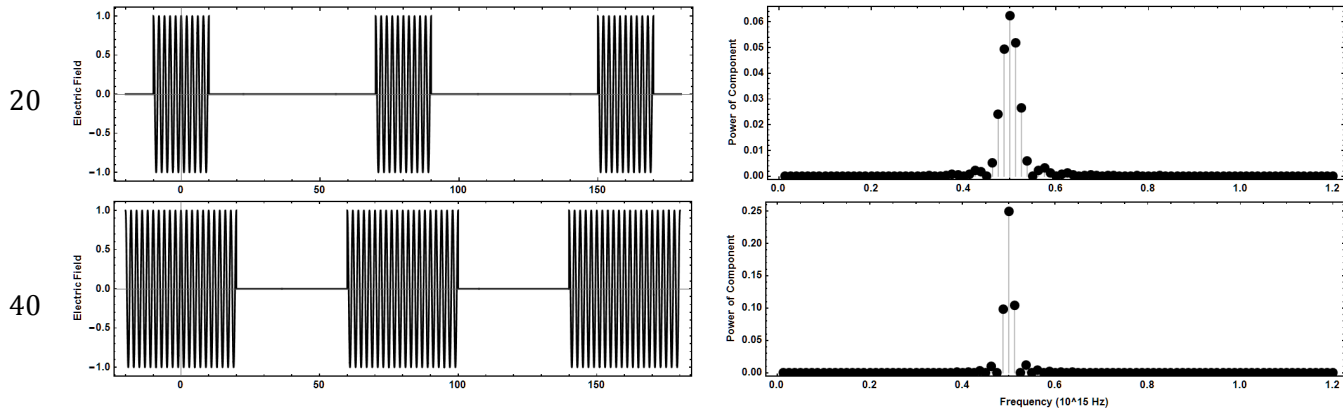


Figure 9. Waves and power spectra for three different pulse widths. Notice how the width of the frequency spectrum gets smaller and smaller as the pulse width gets larger and larger. The wave with the shortest pulse width contains the largest spread of frequencies.

Key observation: look at the power spectra. Now the spacing between frequency points is not changing, but notice how the shortest Δt (top graph) has the widest peak in the frequency graph. The longest Δt (bottom graph) has the narrowest frequency peak. **The width of the frequency peak is inversely proportional to the laser pulse width.**

That is important enough that I will restate it like this and put it in a box:

**The shorter the pulse in time, the wider the frequency peak will be.
The longer the pulse in time, the narrower the frequency peak will be.**

Let's use the symbol Δf to represent the width of the frequency peaks. If we define the width to be the frequency of the low point just to the right of the peak minus the frequency of the low point just to the left of the peak, then you can see that this is true:

$$\Delta f = \frac{2}{\Delta t} \quad (\text{specifically for a square pulse})$$

Or it could be more elegantly written like this:

$$\Delta t \Delta f = 2 \quad (\text{specifically for a square pulse})$$

Notice that the units of time and frequency cancel out, so the 2 on the right hand side is dimensionless.

Test it out! Is the width of the frequency peak for the top graph equal to $\frac{2}{10 \times 10^{-15}}$ Hz?

This particular equation is only true for a square pulse. Although Δt and Δf will always have an inverse relationship, in general the equation relating Δt and Δf for a given laser pulse will depend heavily on three things:

- What shape is the pulse in time? That will determine the shape of the frequency spectrum.
- How do you decide to define the width of the pulse in time? In this example it seems obvious how to define Δt , but if you have a more complicated pulse shape it might not be so obvious.
- How do you decide to define the width of the peak in the frequency spectrum, Δf ?

It turns out that if you define the widths to be the statistical *standard deviation* of the pulses/peaks, then with a lot of math (well beyond the scope of this class) it can be shown that regardless of the shape of the pulse in time, the product of Δt and Δf can never be smaller than $1/(4\pi)$. That is important enough I will box the equation and give it a name:

$$\Delta t \Delta f \geq \frac{1}{4\pi}$$

Fourier Uncertainty Principle;
true for any pulse

This is called the *uncertainty principle* of Fourier analysis. It is the first major instance of uncertainty in the physics of this class. Simply put it says: if you want a wave with a well-defined frequency (a narrow frequency spectrum), it will need to be long in duration. If you want a wave that is localized in time (a short pulse), it will need to have a wide frequency spectrum. You cannot have a wave with a well-defined frequency that is also localized in time.

Quantum mechanical uncertainty

The frequency of a photon is related to its energy via $E = hf$, where h is Planck's constant. Since h is a constant, this means that $\Delta E = h\Delta f$, where ΔE and Δf refer to the *uncertainties* in energy and frequency which can be thought of as the widths (standard deviations) of the statistical distributions of the photons. Putting that information into the Fourier Uncertainty Principle leads to one of Heisenberg's Uncertainty Principles, which you may have seen before:

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

Heisenberg Uncertainty Principle #1

Often we use the symbol \hbar ("h-bar") in place of $h/(2\pi)$, so it can be written more compactly like this:

$$\Delta E \Delta t \geq \frac{\hbar}{2}$$

HUP #1, more compact form

You cannot have a particle with a well-defined energy that is also localized in time.

MOREOVER, all of this mathematics applies to waves in space as well as to waves in time, as long as we talk about spatial extent of a wave (symbol Δx) along with the "spatial frequencies" which must be present to make up the wave. We will see that a fundamental part of quantum mechanics is *wave-particle duality*, whereby all particles also have a wave nature. The spatial frequencies relate to the particle's momentum, p , and those concepts all combine to form Heisenberg's second Uncertainty Principle:

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

Heisenberg Uncertainty Principle #2

If you want a particle with a well-defined location in space (small Δx) then its momentum will have a large uncertainty (Δp will be large if Δx is small). If you want a particle with a well-defined momentum (small Δp), then it cannot have a very well-defined location in space (Δx will be large if Δp is small). You cannot have a wave with a well-defined momentum that is also localized in space.