Traveling Functions, Traveling Waves, and the Uncertainty Principle

R.M. Suter
Department of Physics, Carnegie Mellon University

Experimental observations have indicated that all quanta have a wave-like property and that we can relate those wave-like properties to properties more commonly associated with localized particles: momentum and energy. Classical waves may carry momentum and energy, but we have to think in terms of densities rather than well defined, discrete "lumps." The wave-particle relations are the de Broglie relations

$$E = h\nu = \hbar\omega \tag{1}$$

$$p = \frac{h}{\lambda} = \hbar k, \tag{2}$$

where the *angular frequency* $\omega = 2\pi\nu = \frac{2\pi}{T}$, T being the period of oscillation of the wave, and the analogous quantity for spatial variation is the *wavenumber*, $k = \frac{2\pi}{\lambda}$ with λ being the wavelength.

Now we are presented with a difficulty: presumably, whatever the wave represents, it should have an appreciable amplitude in the neighborhood of the object, it should have a negligible amplitude far from the object, and it should travel through space with the object. That is to say, the wave is not a true periodic function because it does not repeat. How then, do we define λ or T? How do we construct a "wavy" mathematical object that is localized to a region of space and time?

Below, we construct mathematical functions that travel. Then we come to the definition of localized waves that travel and we find that in order to make a localized wave, we have to include a range of ω's and the corresponding range of k's. These are found to correspond inversely to the size of the region of space and the size of the interval in time over which the wavy function has appreciable amplitude. This observation, together with the de Broglie relations, leads directly to the Heisenberg Uncertainty principle.

Traveling functions. Take an arbitrary function of one variable, $f(y)$. This can be made into a function that travels through space by letting $y = a(x - vt)$ so that $f(x,t) = f[a(x - vt)]$.

Take a simple parabola, $f(y) = y^2$. Then, $f(x,t) = [a(x - vt)]^2$. At $t = 0$, the minimum point is at $x = 0$. At a later time, t_1, the minimum has moved to $x = vt_1$. In fact any point on the function does the same thing: take a point where the value is f_o: at an arbitrary time,

$$a(x_o - vt) = f_o^{1/2} \tag{3}$$

$$x_o = \frac{f_o^{1/2}}{a} + vt \tag{4}$$

where x_o is the position at which the function equals f_o. If $v < 0$, the function travels in the opposite direction.

Side note: We can generalize the above to three dimensions by specifying a direction of travel with a unit vector, $\hat{\mathbf{n}}$:

$$y = a(\hat{\mathbf{n}} \cdot \mathbf{r} - vt). \tag{5}$$

This yields an argument that does not change along the direction perpendicular to $\hat{\mathbf{n}}$ since the dot product is the projection of $\mathbf{r}$ onto $\hat{\mathbf{n}}$. The resulting function is constant on such planes (at a fixed instant of time); it varies and propagates along the $\hat{\mathbf{n}}$ direction. To see that it propagates, consider any point on a fixed plane perpendicular to $\hat{\mathbf{n}}$ – this fixes the product $\hat{\mathbf{n}} \cdot \mathbf{r}$ and at time, t_1, fixes the argument of the function. As t increases, we have to move to a new plane with larger $\hat{\mathbf{n}} \cdot \mathbf{r}$ to find the same value of the argument and thus of the function. The "plane of constant function" moves along $\hat{\mathbf{n}}$ at speed v.

A particularly important traveling function is the traveling harmonic wave. One reason this is important is that with a sum of such waves, we can synthesize almost arbitrarily general traveling functions (this goes under the heading of Fourier synthesis – more on that later). Here, we will make the replacement $a(x - vt) \rightarrow kx - \omega t = k(x - v_p t)$, where $v_p = \frac{\omega}{k}$. One example looks like

$$\Psi(x, t) = A \cos(kx - \omega t + \delta), \tag{6}$$

from which we see that, at fixed t, when kx changes by 2π the function goes through one period and, thus, $k = \frac{2\pi}{\lambda}$ and similarly, at fixed position, when ωt increases by 2π, the same thing happens which means that $\omega = \frac{2\pi}{T} = 2\pi\nu$. The term δ in (6) is a dimensionless constant that adjusts the "phase" of the wave – at, say, $t = 0$, it shifts the positions at which the peaks (or nodes or minima) occur. (6) travels at $v_p = \frac{\omega}{k}$ (note the consistency of the units here); v_p is called the *phase velocity* since it gives the speed of travel of a point (any point) of constant phase. The square of this function is just another function that also travels at v_p:

$$\Psi^2(x, t) = A^2 \cos^2(kx - \omega t + \delta) = \frac{A^2}{2} \left[1 + \cos 2(kx - \omega t + \delta)\right]. \tag{7}$$

An even more important traveling function is the complex function $\Psi(x, t) = Ae^{i(kx - \omega t)}$. This is a traveling complex harmonic wave. It's called harmonic because both the real and imaginary parts are harmonic functions:

$$\Psi(x, t) = Ae^{i(kx - \omega t)} = A \left[\cos(kx - \omega t) + i \sin(kx - \omega t)\right]. \tag{8}$$

Note that if we let A be complex, we can write it in polar form as $|A|e^{i\delta}$ in which case $\Psi(x, t) = |A|e^{i(kx - \omega t + \delta)}$; so multiplying by a complex number changes the phase and amplitude of the wave. The square of the amplitude of this complex wave is given by $|\Psi|^2 = \Psi^* \Psi = A^* \; e^{-i(kx - \omega t)} \times Ae^{i(kx - \omega t)} = |A|^2$ – just a constant, independent of x and t! Compare this result to Eq. 7.

Both the real and complex harmonic waves are true periodic functions: they both obey the relations $\Psi(x + \lambda, t) = \Psi(x, t)$ and $\Psi(x, t + T) = \Psi(x, t)$, so the wavelength and period are well defined quantities. In fact, if one of these functions is supposed to describe the mysterious quantum waves, then we can use λ and T in the de Broglie relations to relate the wave properties to particle-like properties, p and E. The only drawback is that these functions, being periodic, necessarily go on forever in both time and space!

Dispersion relations. Finally, we need to note that ω and k (or their associated T and λ) are not independent quantities. If $\Psi(x, t)$ is to describe some physical situation and

if we know the wavelength and the speed of propagation, then we know the frequency of oscillation: the period, T is the time it takes one wavelength to go past a fixed point in space or $\lambda = v_p T$, which we can write as $\omega(k) = v_p k$. The relation between ω and k is called the *dispersion relation* of the wave. Because v_p can depend on k (or ω), different types of waves will have different dispersion relations, as we illustrate below.

Notice that the dispersion relation for a particular physical system couples the two de Broglie relations: k gives us $\omega(k)$ and

$$\begin{aligned} p &= \hbar k & (9)\\ E &= \hbar\omega(k). & (10) \end{aligned}$$

These equations can also be expressed in terms of ω if we can invert the dispersion relation.

Electromagnetic waves: Here, we know that $\nu = c/\lambda$ or $\omega(k) = ck$ (in vacuum) and the dispersion relation is linear since c is independent of the wavelength. The phase velocity is then $\frac{\omega}{k} = c$ and a point of constant phase moves at the speed of light. This seems satisfying and appropriate. Furthermore, using this dispersion relation, we can write the energy of a photon as $E = \hbar\omega = \hbar ck = pc$. This is the relation given by special relativity for the energy-momentum relationship of a massless object (and the photon has to have zero rest mass since it travels at c).

Particle waves: In this case, we have to figure out the dispersion relation for these waves that we do not yet know how to interpret. For the case of a free particle we can do this using the de Broglie relations. We can assume in this case that the energy is all kinetic: $E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m} = \hbar\omega(k)$. Then,

$$\omega(k) = \frac{\hbar k^2}{2m} \qquad (11)$$

(a quadratic dispersion relation) and the phase velocity of the wave is $v_p = \frac{\omega}{k} = \frac{\hbar k}{2m} = \frac{p}{2m}$. But, classically, $p = mv$, so we find that $v_p = \frac{1}{2} v_{classical}$! Now, we're not so happy! It looks like these waves will not move along with the expected position of the particle.

There is one consideration that will allow us to wiggle out of the above quandary. Because the harmonic, periodic functions we have used above go on forever, they are really not a satisfying type of wave to describe the motion of a localized object. Maybe once we develop a localized wavy function (a "wavepacket"), the localized region that has significant amplitude will move at the right speed even if the points of constant phase we've just worked with do not...

Construction of wavepackets. Adding together harmonic waves of different wavelengths yields a function of more complex shape. Just adding two of unit amplitude yields

$$\begin{aligned} f(x) &= e^{ik_1 x} + e^{ik_2 x} & (12)\\ &= e^{i\frac{1}{2}(k_1+k_2)x}\left(e^{i\frac{1}{2}(k_1-k_2)x} + e^{i\frac{1}{2}(k_2-k_1)x}\right) & (13)\\ &= e^{ik_o x}\left(e^{i\frac{1}{2}(\delta k)x} + e^{-i\frac{1}{2}(\delta k)x}\right) & (14)\\ &= 2e^{ik_o x}\cos\left[\frac{1}{2}(\delta k)x\right], & (15) \end{aligned}$$

where k_o is the average wavenumber and $\delta k = k_1 - k_2$ is the difference (note that we could have done this in the time domain with exactly the same math – everywhere you see k, replace it with ω and everywhere you see x, replace it with t). If we are dealing with a real harmonic function, we can sum cosines by just taking the real part of each of the above equations (the sum of real parts is the real part of the sum – it's a linear operation!); we get

$$\mathcal{R}\{f(x)\} = 2\cos k_o x \ \cos\frac{1}{2}(\delta k)x. \tag{16}$$

Isn't that easier than remembering trigonometric identities? We can sum sines by taking the imaginary parts (note that that yields a sine times a cosine). And we can deal with arbitrary phase shifts by factoring the average and difference as we've done above (try it out – what do you get?).

Eq. 15 and its real counterparts are in a useful form: if we suppose, as is often the case, that $\delta k << k_o$, then we have a rapidly oscillating function (the first term) multiplying a slowly varying function (called the "envelope" function). The rapidly oscillating function with the nominal (or average) frequency is *modulated* by the slowly varying one. [We'll look at examples in class.]

Is (12) and thus, (15), a periodic function? The answer depends on whether k_1 and k_2 are rationally related. If so, then yes; if not, then no. In either case, however, the qualitative idea of a modulated rapid oscillation holds. As we make $\delta k/k_0$ small, there are many large amplitude oscillations between positions where the envelope function becomes small. In the limit as $\delta k \to 0$ (but not *at* $\delta k = 0$), the wavelength of the envelope function goes to ∞ – the modulation function has infinite period (it takes forever for the component waves to get out of phase with each other). Conversely, for large $\delta k/k_0$, the component waves rapidly de-phase and the envelope has a short period.

What we are looking for is a non-periodic or non-repeating function in which the component waves de-phase in some finite interval, say δx. The non-repeating requirement argues that we should include waves with very small differences δk, but the second requirement says we need to include waves with significantly different periods. We can have it both ways: we integrate (instead of doing a discrete sum) over a finite interval. The interval width sets the de-phasing length (or time) and the inclusion in the integration of infinitesimally different wavenumbers guarantees that the function will not be periodic.

We conclude that the general form for an appropriate function is

$$f(x) = \int dk \ A(k) \ e^{ikx}, \tag{17}$$

where $A(k)$ is the weight function that determines the range of k values that receive significant weight. This formula sums up harmonic waves that are all in-phase at $x = 0$, so we expect $f(0)$ to be the maximum possible value. For simplicity, let's require (for now) that

$$\int A(k) \ dk = 1. \tag{18}$$

This means that $f(0) = 1$ (just substitute $x = 0$ in (17)).

The simplest $A(k)$ is just a square box:

$$A(k) = \left\{ \begin{array}{l} \frac{1}{\delta k},\ k_o - \frac{1}{2}\delta k \le k \le k_o + \frac{1}{2}\delta k \\ 0,\ \text{otherwise.} \end{array} \right\} \tag{19}$$

It will be convenient to change the integration variable from k to the deviation of k from k_o: $k = k_o + \alpha$. The integration variable is now α and $dk = d\alpha$ since k_o is a constant. This allows us to factor the exponential into a constant term (constant with respect to the integration variable) times the term we have to integrate: $e^{ikx} = e^{ik_o x}e^{i\alpha x}$. For the square distribution (19), Eq. 17 becomes

$$f(x) = e^{ik_o x}\frac{1}{\delta k}\int_{-\frac{1}{2}\delta k}^{+\frac{1}{2}\delta k} d\alpha\ e^{i\alpha x} \tag{20}$$

$$= e^{ik_o x}\frac{1}{i(\delta k)x}\left[e^{i\frac{1}{2}(\delta k)x} - e^{-i\frac{1}{2}(\delta k)x}\right] \tag{21}$$

Use of the Euler identity, $\sin\theta = \frac{1}{2i}[e^{i\theta} - e^{-i\theta}]$, then yields

$$f(x) = e^{ik_o x}\ \frac{\sin\beta}{\beta}, \tag{22}$$

where $\beta = \frac{1}{2}(\delta k)x$.

As long as δk is small compared to k_o (typically the limit in which we'll be interested), the conclusion is that we have a rapidly oscillating complex wave, $e^{ik_o x}$, multiplying a more slowly varying "envelope" function, $\frac{\sin\beta}{\beta}$. Note that the factoring trick used above can be used with arbitrary weight functions, $A(k)$. Pick the average k, call it k_o and write $A(\alpha)$. Particularly if $A(\alpha)$ is a symmetric function, this generally makes the integration simpler and it shows the separation of the phase oscillation from the envelope function.

The function $\frac{\sin\beta}{\beta}$ has a large lobe around $\beta = 0$ and then smaller and smaller lobes at larger $|\beta|$ (plots on the next page). We identify the central lobe as the region where the particle is most likely to be found.[1] The size of this region is $\delta\beta = 2\pi$ since the sine function goes to zero at $\pm\pi$. We then have $\delta\beta = \frac{1}{2}(\delta k)(\delta x) = 2\pi$ or

$$\delta k\ \delta x = 4\pi. \tag{23}$$

This says that including a range δk of wavenumbers generates a function with a central lobe of width δx. This is an understandable mathematical observation: we've set up the calculation so that all the component waves in (17) are in-phase at $x = 0$. The broader the range of component waves we include (i.e., the larger is δk), the more rapidly these component waves will get out of phase with each other and the smaller will be the interval δx.

[1] The meaning in this case is that the particle is more likely to be found in this interval than in any other interval of the same size. More on this later when we establish a firm interpretation of the quantum wavefunction.

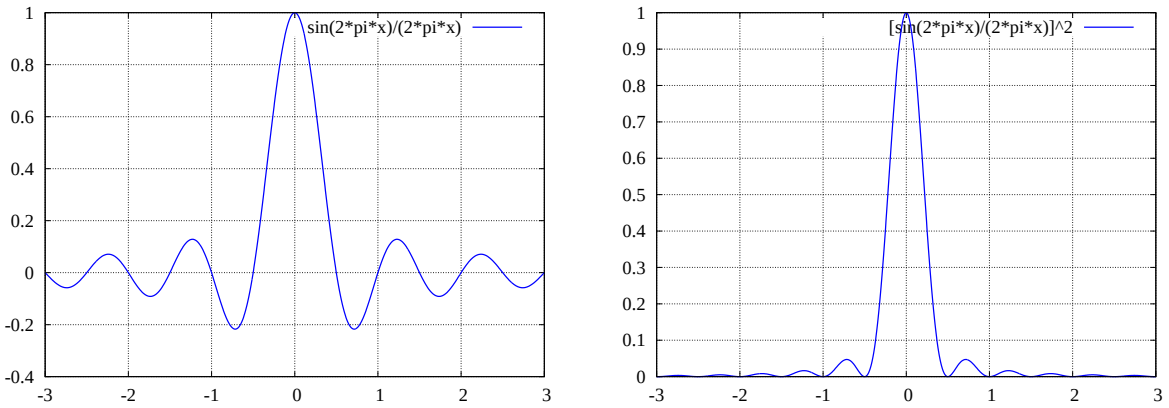

Figure: Wavepacket and its square resulting from a square distribution of frequencies.

We started off saying that we had to assign a wave character to all quanta (through the de Broglie relations) and that whatever was waving should be large "near" where we expect the particle to be and should be small "far away". We find we can make such a function by summing harmonic waves over some range of wavenumbers. So, now we find that appropriate functions do not have well-defined wavenumbers to plug into the de Broglie relation $p = \hbar k$. We can define an "average" momentum through $p_o = \hbar k_o$ but we have to also say that there is a range of momentum components, $\delta p = \hbar \delta k$. For the square distribution, multiplying both sides of (23) by $\hbar$ yields,

$$\delta p \ \delta x = 4\pi\hbar. \tag{24}$$

This is related to one of the Heisenberg Uncertainty relations.

The interpretation of the above observations is that the function that we are describing, that has the wave property, is related to a probability for where the particle will be found and, at the same time, it tells us the probability for the particle to have a particular value of momentum. We will later develop a general definition of the widths of these probability distributions analogous to (but not identical to) δp and δx used here. A highly localized wave has to have a broad distribution of k's and therefore of p's or momenta.

The factor of 4π in (24) is independent of the width of the distribution. However, if we choose a different shape for $A(k)$, we will get a relation just like (24) but with a different numerical constant.

Traveling wavepackets. By analogy with (17), we can generate a traveling wavepacket by doing an integral like

$$\Psi(x,t) = \int dk \ A(k) \ e^{i[kx - \omega(k)t]} \tag{25}$$

The complication is that we have to keep track of how ω changes with k. The first approximation is to assume the range of k's to be narrow enough that we can approximate $\omega(k)$ by a straight line:

$$\omega(k) = \omega(k_o) + \frac{d\omega}{dk}|_{k_o}\alpha = \omega_o + v_g\alpha, \tag{26}$$

Where again, we have used $k = k_o + \alpha$ and on the right $\omega_o = \omega(k_o)$ and $v_g = \frac{d\omega}{dk}$ is called the group velocity. In this approximation,

$$\Psi(x,t) = e^{i(k_o x - \omega_o t)} \int_{-\frac{1}{2}\delta k}^{\frac{1}{2}\delta k} d\alpha \; A(\alpha) \; e^{i\alpha(x - v_g t)}. \tag{27}$$

For the square distribution (19), this yields the same math as before (see Eq. 20) if we just replace x with $x - v_g t$:

$$\Psi(x,t) = e^{i(k_o x - \omega_o t)} \; \frac{\sin[\frac{1}{2}(\delta k)(x - v_g t)]}{\frac{1}{2}(\delta k)(x - v_g t)} \tag{28}$$

Now we have a product of two terms where the envelope function travels at the speed $v_g = \frac{d\omega}{dk}$ instead of $v_p = \frac{\omega}{k}$. For electromagnetic waves in vacuum, $\omega = ck$ and $v_p = v_g = c$: both the rapid oscillation and the envelope travel at the same speed. For masses with the dispersion relation $\omega = \frac{\hbar k^2}{2m}$, $v_p = \frac{\hbar k}{2m} = \frac{1}{2} v_{classical}$ whereas $v_g = \frac{\hbar k}{m} = v_{classical}$. In both cases, the envelope that contains the rapid oscillations travels with the appropriate speed to move with the object. In this approximation, the envelope maintains a constant shape as it propagates and the uncertainty product, $\delta p \; \delta x$ is constant and is again equal to 4π for this case (the square distribution). Furthermore, the time it takes the central region to pass a point in space can be called δt; requiring x to be constant, we ask for the argument of the sine function in (28)to change by 2π: $\delta\beta = 2\pi = \frac{1}{2}(\delta k) v_g \delta t$. The quantity $v_g \delta k = \delta\omega$ (by the definition of v_g), with $\delta\omega$ being the range of frequencies included in the integral, so we have the result that

$$\delta\omega \delta t = 4\pi, \tag{29}$$

which is analogous to (23). This is equivalent (with de Broglie) to

$$\delta E \; \delta t = 4\pi\hbar, \tag{30}$$

which says that if the event (ex., an electron striking a detector) can be said to be localized in time within δt, then we do not know the energy of the object to better than δE.

Once again, changing the form of $A(k)$ will change the numerical coefficient but not the fact that the "uncertainty product" has to be some finite number. In fact, before we can really make this claim, we need a more generally applicable definition for the width that we've called δk here (and all the other δ quantities). That definition will turn out to be a suitable standard deviation of the distribution for each quantity and we will denote each with a "Δ." Then, we will be able to make the following statement: for any specific functional form for $A(k)$ (square versus Gaussian versus Lorentzian,...) the products $\Delta k \; \Delta x$ and $\Delta\omega \; \Delta t$ are constant. If we look over all functional forms, it turns out that the Gaussian distribution for $A(k)$ corresponds to the smallest coefficient and it is 1/2. So, finally, we can write the Heisenberg uncertainty principle as

$$\Delta p \; \Delta x \;\; \geq \;\; \frac{1}{2} \; \hbar \tag{31}$$

$$\Delta E \; \Delta t \;\; \geq \;\; \frac{1}{2} \; \hbar. \tag{32}$$