

Schrödinger's master equation of quantum mechanics

Notes on Quantum Mechanics

<http://quantum.bu.edu/notes/QuantumMechanics/TheSchrödingerEquation.pdf>
Last updated Thursday, October 7, 2004 13:46:37-05:00

Copyright © 2004 Dan Dill (dan@bu.edu)
Department of Chemistry, Boston University, Boston MA 02215

A student's search

When I first heard, as a student in college, that there was such a thing as quantum mechanics, it seemed like societally sanctioned science fiction. Everything seemed to depend on something called the Schrödinger equation, but that equation was a mystery to me. I was fascinated and eager to learn more, and so decided to browse in the college bookstore for something with "quantum mechanics" in the title.

I remember vividly the first book I picked up. Its title was simply "Quantum Mechanics" and I thought "How lucky to have found a book that will explain things." With great expectation I open to the first page. There were several lines of text, and then, wham!, the equation

$$H\psi = E\psi,$$

the very Schrödinger equation I had heard about.

Needless to say, *I had no idea whatsoever what such an equation could possibly mean!* That it appeared so early seemed cruel treatment, at best, and certainly daunting. I closed the book, put it back on the shelf, and decided that perhaps quantum mechanics was even more impenetrable than I thought.

It turns out the Schrödinger equation is the fundamental equation of quantum mechanics. It is called the Schrödinger equation, after its inventor, Erwin Schrödinger. I have spent many years working with this equation, and have learned that, with a little guidance, it is not so daunting after all.

I would like to introduce you to the Schrödinger equation in a way that hopefully will allow you to feel quite at home with it, to see where it comes from, how to "solve" it in pictorial way, how it leads to *quantized energies* E in atoms and molecules, and how the functions that solve it, the so-called *wavefunctions* ψ , describe *all* properties of atoms and molecules.

This then is a story of a way to understand the Schrödinger equation, the key equation of quantum mechanics. Let's begin by seeing why we need something like the Schrödinger equation at all.

The Schrödinger equation

The de Broglie hypothesis introduced the idea that particles are in some sense associated with waves, Born proposed that these are abstract, nonphysical waves describing fluctuations in probability amplitude. These ideas do not, however, tell us any details of the waves, other than their wavelengths. To proceed further, we need a way to determine the waves.

The way to do this was discovered by Schrödinger. He proposed that wavefunctions are solutions to a rather peculiar kind of equation,

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$$

Here $d^2\psi(x)/dx^2$ is the rate of change of the slope—the *curvature*—of the wavefunction at position x , $V(x)$ is the potential energy at position x , and E is the total energy; we have also used the conventional abbreviation $\hbar = h/2\pi$.

The first time I saw this equation, I didn't have the foggiest idea of where it came from or, even more worrisome, what it could possibly mean. A colleague tells me that Schrödinger himself was puzzled by the equation when he first wrote it down!

There are three steps we need to take to understand where Schrödinger equation comes from. The first step is the strangest. We need to learn about a key relation that tells us that while in our everyday experience the position and momentum of a particle seem to be simple numbers, it turns out that instead they are something different, for when we multiply them as $x p$ we get a different result than when we multiply them as $p x$. The difference turns out to be $i h/2\pi$. Because Planck's constant is so tiny, the difference is absolutely negligible in everyday experience. But for electrons in atoms and molecules, where both distances (x) and masses, and so momenta (p) are tiny, the difference becomes huge. This evident property of our world is known as the position-momentum commutation relation. It is the foundation on which all of quantum mechanics rests.

The next step is to see what changes we need to make to our concept of position and momentum in order to satisfy the commutation relation.

The third step is to incorporate our new way of representing position and momentum into the analysis of the behavior of electrons in atoms and molecules. It is this step that leads to Schrödinger equation.

■ Position-momentum commutator

The story begins with the discovery by Heisenberg of a peculiar relationship between the position of a particle in a particular direction, say x , and the momentum of the particle along that direction, p_x .

W. Heisenberg, "Quantum-theoretical re-interpretation of kinematic and mechanical relations," Z. Phys. 33 (1925) 879-893. A translation of this paper into English, and also of most of the other key works on the creation of quantum mechanics is in B. L. van der Waerden, "Sources of Quantum Mechanics," (Dover, 1968, ISBN 0-486-61881-1).

The relationship, known as the *position-momentum commutation relation*, says that

$$x p_x - p_x x = i \hbar,$$

where i is the square root of -1 and \hbar is an abbreviation for Planck's constant divided by 2π . Heisenberg made the discovery in order to understand the differences between the spectrum of

two-electron systems (He, Li⁺, Be²⁺, etc.) and Bohr's model of one-electron systems (H, He⁺, Li²⁺, etc.).

On the face of it, this commutation relation makes no sense at all. For example, let's say the position of a particle is 5 m along x from the origin, and at that point its momentum is 7 kg m s⁻¹. With these values we reasonably expect that

$$x p_x = 35 \text{ kg m}^2 \text{ s}^{-1} = 35 \text{ J s},$$

$$p_x x = 35 \text{ kg m}^2 \text{ s}^{-1} = 35 \text{ J s},$$

and so that the commutator vanishes,

$$x p_x - p_x x = 35 \text{ J s} - 35 \text{ J s} = 0.$$

In fact, while the commutator is *nearly* zero on the scale of everyday experience, since Planck's constant is so small (on the order of 10⁻³⁴ J s), the difference is *not* exactly zero. This small deviation from zero is the very foundation of the quantum behavior of matter and light.

■ Three approaches to the position-momentum commutator

How can it be that $x p_x - p_x x$ is not exactly zero? The answer must be that x and p_x are *not simple numbers*, for if they were, their commutator *would be* exactly zero. Three different approaches have been taken to account for the non-zero difference.

Heisenberg's matrices

The first approach was that taken by Heisenberg in his analysis of the spectra of He. Heisenberg represented the position and momentum of the electrons as *matrices*. A property of matrices is that in general their multiplication gives a different result depending on the order of the factors. For example, the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}$$

have the different products

$$A.B = \begin{pmatrix} 8 & 5 \\ 20 & 13 \end{pmatrix} \text{ and } B.A = \begin{pmatrix} 13 & 20 \\ 5 & 8 \end{pmatrix}.$$

and so their commutator,

$$A.B - B.A = \begin{pmatrix} -5 & -15 \\ 15 & 5 \end{pmatrix},$$

is not zero.

Schrödinger's operators

The second approach was taken by Schrödinger. He assumed that position and momentum were *mathematical operators* chosen so that the result of their operation on a function amounted to the position momentum commutation relation. There are several ways to make this choice, and they all give equivalent results. The most common choice is that the operator for position is

$$x \rightarrow x \times,$$

that is, "multiply by position," and the operator for momentum is

$$p_x \rightarrow -i \hbar \frac{d}{dx}$$

that is, $-i \hbar$ multiplied by the slope (derivative) operator. But multiply x by what? And the slope of what? The answer is, some mathematical function that varies smoothly with position, so that its slope can be determined at each point. To see how this works, let's first confirm that the choices for the operators for x and p_x are consistent with the commutation relation, by evaluating the effect of

$$x \left(-i \hbar \frac{d}{dx} \right) - \left(-i \hbar \frac{d}{dx} \right) x$$

on an arbitrary function of position, $f(x)$. Here are the details:

$$\begin{aligned} & \left\{ x \left(-i \hbar \frac{d}{dx} \right) - \left(-i \hbar \frac{d}{dx} \right) x \right\} f(x) \\ &= -i \hbar \left[x \frac{df(x)}{dx} - \frac{d}{dx} \{ x f(x) \} \right] \\ &= -i \hbar \left\{ x \frac{df(x)}{dx} - f(x) - x \frac{df(x)}{dx} \right\} \\ &= i \hbar f(x) \end{aligned}$$

The way to read the equality of the first and last lines is that operating on $f(x)$ by $x(-i \hbar d/dx) - (-i \hbar d/dx)x$ is the same as simply multiplying $f(x)$ by $i \hbar$.

Dirac's Hilbert space

These two approaches, Heisenberg's in terms of numerical matrices and Schrödinger's in terms of mathematical operators, seem very different, and when they were proposed it was of great concern whether both approaches could be correct. It was the third approach, invented by Dirac, that showed that the matrix and operator approaches were each special cases of a more general, purely algebraic approach.

The very best exposition of Dirac's approach is by Dirac himself, in his Principles of Quantum Mechanics, 4/e (Oxford University Press, 1982, ISBN 0198520115).

Dirac's approach is based on an abstract algebra that incorporates as its fundamental postulate the position-momentum commutation relation. Dirac's approach can be thought of as a multidimensional generalization of geometrical ideas familiar from three-dimensional space. This generalization is known a Hilbert space, and the "axes" of this space are the different quantum states that a system may be in. People (such as me!) who think visually tend to feel right at home in Hilbert space!

■ Schrödinger's approach

The three different approaches are thus equivalent (albeit with the nod for generality going to Dirac's). Which one is used depends on which is most convenient in a given circumstance. In the analysis of the properties of atoms and molecules, the Schrödinger approach is the most directly accessible, and so it is the one that we will use.

Schrödinger began by expressing the energy of a system as

$$T(x) + V(x) = E$$

the sum of the kinetic energy, $T(x)$, and potential energy, $V(x)$, at each position. He next used Newton's laws to write this expression, for the example of a single particle in one dimension, as

$$\frac{p^2}{2m} + V(x) = E$$

Finally, he made the key step of incorporating the position-momentum commutation relation. He did this by using his operator expressions for position and momentum, which we have the required commutator, to transform this equation into the *operator equation*

$$\left\{ \frac{1}{2m} \left(-i\hbar \frac{d}{dx} \right)^2 + V(x) \right\} \psi(x) = E \psi(x).$$

In this equation Schrödinger has introduced a function $\psi(x)$ for the operators to operate on. But what function do we use? The answer, it turns out, is just the wavefunction of the Born interpretation of matter waves. This equation is known as the Schrödinger equation and is usually written as

$$\left\{ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right\} \psi(x) = E \psi(x).$$

If the potential energy is 0, we say we have a free particle. Show that the function $\sin(2\pi x/\lambda)$ satisfies the Schrödinger equation for a free particle, and that its energy is $p^2/2m$.

Schrödinger's equation can be written even more compactly by making the definition of what is called the hamiltonian operator,

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x).$$

With this definition, the equation becomes

$$H \psi = E \psi$$

This is the very equation that I encountered in my search in the university bookstore so many years ago!

Now that we know a little about where the Schrödinger comes from, we can next see what it means.

What the Schrödinger equation means

The key to beginning to make sense of the Schrödinger equation is to rewrite it in a special way,

$$\text{curvature of } \psi \text{ at } x \propto - \text{kinetic energy at } x \times \psi \text{ at } x$$

We will see that using this form of the Schrödinger equation we can develop a qualitative understand all of the key features of wavefunctions: that in regions of high kinetic energy wavefunctions oscillate rapidly and so have small wavelengths; that in regions of low kinetic energy that wavefunctions oscillate slowly and so have large wavelengths; that in regions of *negative* (!) kinetic energy, instead of oscillating, wavefunction diverge away from zero; and, finally, that the observed (quantized) energies, are just those energies for which the wavefunction is prevented from diverging to infinity.

We will be able to do all of this simply by making simple sketches of the form of the wavefunction required by the Schrödinger equation. It is easy to make these sketches, once we learn what the rules are, and these sketches allow us to accurately predict what wavefunctions look like without doing any calculations at all.

■ Kinetic energy variation with position

The first step toward understanding the Schrödinger equation is to understand how kinetic energy changes with position. The kinetic energy is determined by the total energy, E , and the potential energy, $V(x)$. The total energy is a constant but the potential energy in general changes with position. This means that the kinetic energy changes with position.

A nice example illustrating changing potential energy of a skateboarder traversing a parabolic track. At the top of the track the skateboarder has the greatest (gravitational) potential energy but no kinetic energy; her total energy is entirely potential energy. Once she starts to roll down the track, she picks up speed as her potential energy is converted to kinetic energy. At the bottom of the track she is moving the fastest and her potential energy is at its lowest value. She continues to roll up the other side of the track, rapidly at first but slowing down more and more, as her kinetic energy is converted back into potential energy. The potential energy reaches its maximum at the top of the other side of the track, where the skateboarder comes to a stop and so has zero kinetic energy.

We can develop this example quantitatively. We begin by expressing the shape of the track in terms of its vertical height relative to the bottom of the track. Let's make the track have a parabolic profile,

$$h(x) = x^2 \text{ meter.}$$

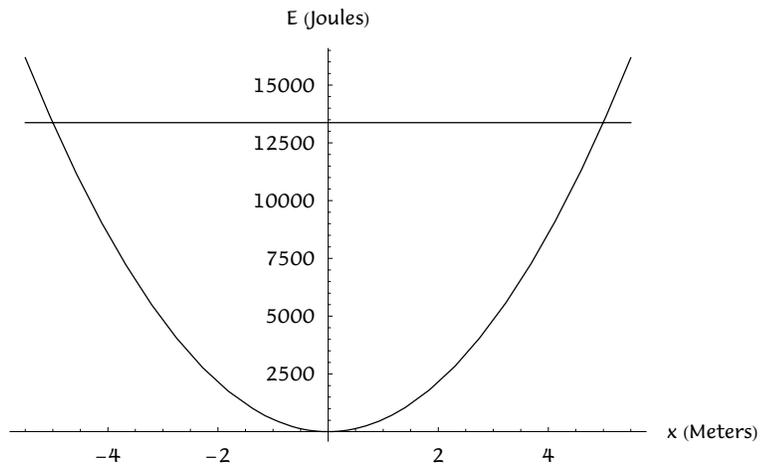
We can specify how high up the track the skateboarder begins the traverse by setting her total energy equal to her gravitational potential energy, relative to the bottom of the track, at that height. The gravitational potential energy is determined by the skateboarder's mass and by the height,

$$V(x) = m g h.$$

Here m is the skateboarder's mass, g is the acceleration due to gravity, and h is the height of the skateboarder above the bottom of the track (note, h is *not* Planck's constant here!). Let's say the horizontal length of the track, from the starting edge across to the finishing edge, is 10 meters and that the skateboarder weighs 120 lbs. The means that height of the track is 25 m and that the skateboarder's total energy is 13,000 J.

Confirm these values for the height of the track and the skateboarder's total energy.

Here is a graph of the skateboarder's potential energy, $V(x)$, along the traverse.



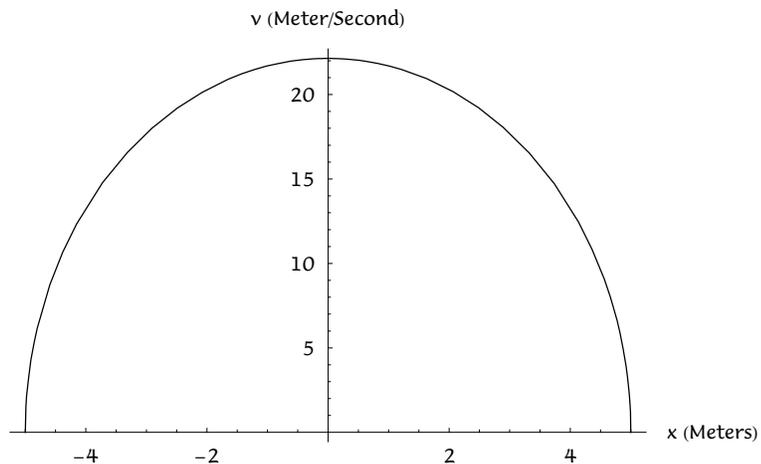
Gravitational potential energy $V(x)$ along a parabolic track. The horizontal line indicates the total energy E .

The horizontal line is her total energy, E . The kinetic energy at any point along the traverse is the vertical distance, $E - V(x)$, from the total energy line down to the potential energy curve. The value of the skateboarder's maximum kinetic energy, at the bottom of the traverse, is 13,000 J. Her initial potential energy has been converted entirely to kinetic energy there.

The speed of the skateboarder at any point along her traverse is proportional to the square root of her kinetic energy at that point,

$$v(x) = \sqrt{2T(x)/m}.$$

Here is a graph of her speed along the traverse.



Velocity along a parabolic skateboard track, starting from rest at its top at $x = \pm 5$ m.

Her maximum speed, at the bottom of the traverse, is 22 m/s.

Use the graphs to determine the kinetic energy, potential energy and speed of the skateboarder at position $x = -3$ meters, having started down the track at rest from the left.

We can apply the same ideas to understand how the kinetic energy of a particle with a total energy E changes due to changing values of its potential energy, $V(x)$.

■ Effect of kinetic energy on wavefunction curvature

The Schrödinger equation is a relation between the curvature of the wavefunction, the kinetic energy of the particle, and the value of the wavefunction. To see precisely what this relation means, let's consider the Schrödinger equation in detail for the possible combinations of kinetic energy and sign of the wavefunction.

In the skateboarder example, the kinetic energy increases from zero at the start and end of the traverse to its maximum value at the midpoint. The skateboarder's kinetic energy can never be negative. In the Schrödinger equation, in contrast, the kinetic energy may be either positive *or* negative, according to whether the value of the potential energy is smaller or larger than the total energy. Regions of x where kinetic energy is positive or zero are called *classically allowed regions*, or more simply, allowed regions; regions of x where kinetic energy is negative are called *classically forbidden regions*, or more simply, forbidden regions. Boundaries between allowed and forbidden regions are called *classical turning points*.

It is quite peculiar that kinetic energy may be negative, since for an ordinary object this would mean the particle velocity would have to be *imaginary*, since a negative kinetic energy means $v^2 < 0$. The reason this is possible in quantum systems is that we are dealing with probability amplitudes rather than localized particles. The squared modulus of the probability amplitude, $|\psi(x)|^2$, tells us the probability per unit length of finding the particle in any particular region. To compute the kinetic energy of the particle, we need to add up its kinetic energy in each region, weighting each contribution by the probability that the particle will be in that region. The result is that the negative contributions to the kinetic energy from the forbidden regions are always more than offset by the positive contributions to the kinetic energy from the allowed regions, so that the total kinetic energy is always positive.

Forbidden regions turn out to be crucial in understanding why only certain values of the energy may occur; that is, forbidden regions are the source of energy quantization. Energy quantization *always* means that particle is *confined to a limited region of space by classically forbidden regions*.

There are four possible combinations of the sign of the kinetic energy and the sign of the wavefunction: positive kinetic energy and either positive or negative wavefunctions, and negative kinetic energy and either positive or negative wavefunctions. Let's look at these combinations separately for the case of positive and negative kinetic energy.

Allowed regions

In allowed regions, the kinetic energy is never negative. This means that the sign of the curvature of the wavefunction is always opposite to the sign of the wavefunction.

$$\begin{aligned} &\text{sign of curvature of } \psi \text{ at } x \\ &= - \times \text{sign of kinetic energy at } x \times \text{sign of } \psi \text{ at } x \\ &= - \times + \times \text{sign of } \psi \text{ at } x \\ &= - \times \text{sign of } \psi \text{ at } x \end{aligned}$$

If the wavefunction is positive, then its curvature will be negative. A negative curvature means that the slope of the wavefunction is becoming more negative (less positive); that is, the wavefunction is curving *downward* from positive values toward the zero line.

If the wavefunction is negative, then its curvature will be positive. A positive curvature means that the slope of the wavefunction is becoming more positive (less negative); that is, the wavefunction is curving *upward* from negative values toward the zero line.

The net result of this behavior is that

in classically *allowed regions*, the wavefunction *oscillates* back and forth across its zero line.

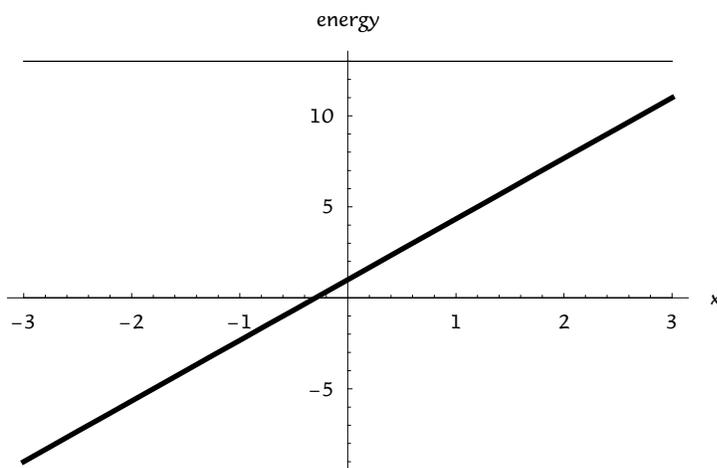
In this way we see that Schrödinger equation predicts an oscillatory probability amplitudes in allowed regions.

Since the curvature — how quickly the slope of the wavefunction is changing — is proportional to the kinetic energy, where the kinetic energy is large, the oscillations will be rapid; that is, each *loop* of the wavefunction (wavefunction variation between two adjacent zero values) will occupy a relatively small range of x and the maximum value of the wavefunction will never have a chance to get very large. Conversely, where the kinetic energy is small, the oscillations will be slow; that is, each loop of the wavefunction will occupy a relatively large region of x and the maximum value of the wavefunction will be able to become relatively large.

The picture that emerges is consistent with de Broglie's relation, $\lambda = h / p$. When kinetic energy increases, the momentum, p , gets larger and the wavelength of the wavefunction becomes smaller. When kinetic energy decreases, the momentum, p , gets smaller and the wavelength of the wavefunction becomes larger. (This connection with the de Broglie behavior is no accident, for de Broglie's relation is built into the Schrödinger equation, through the equation's dependence on the position-momentum commutator, $[x, p] = i\hbar$.)

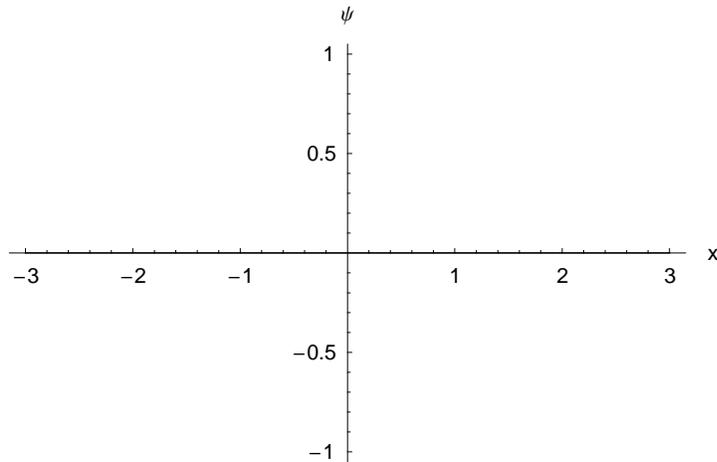
Allowed region example

To get experience with these ideas, consider the following kinetic energy variation, given by the distance between the potential energy curve (thick line) and the constant total energy (thin horizontal line).



Using the axes below, sketch three loops of a wavefunction for this kinetic energy variation, as follows. Start your wavefunction with value at $x = -3$ equal to $\psi(-3) = 0$, have the left-most loop of the wavefunction extend from $x = -3$ to $x = -2$, and have the amplitude of this first loop be $+0.2$. Then sketch in the next two loops, extending toward

more positive values of x . The key parts to answering this question is deciding, first, whether the widths of the two additional loops decrease, are the same as, or increase relative to the first, left-most loop, and, second, whether the amplitudes of the two additional loops decrease, are the same as, or increase relative to the first, left-most loop.



Forbidden regions

In forbidden regions, the kinetic energy is always negative. This means that the sign of the curvature of the wavefunction is always the same as the sign of the wavefunction.

$$\begin{aligned}
 &\text{sign of curvature of } \psi \text{ at } x \\
 &= - \times \text{sign of kinetic energy at } x \times \text{sign of } \psi \text{ at } x \\
 &= - \times - \times \text{sign of } \psi \text{ at } x \\
 &= + \times \text{sign of } \psi \text{ at } x
 \end{aligned}$$

If the wavefunction is positive, then its curvature will be positive. A positive curvature means that the slope of the wavefunction is becoming more positive (less negative); that is, positive wavefunctions curve *away* from the zero line toward more positive values.

If the wavefunction is negative, then its curvature will be negative. A negative curvature means that the slope of the wavefunction is becoming more negative (less positive); that is, the negative wavefunctions again curve *away* from the zero line toward more negative values.

The net result of this behavior is that

in classically *forbidden regions*, the wavefunction *curves away* from its zero line.

In this way we see that Schrödinger equation predicts that probability amplitudes tend to diverge to large, possibly infinite values in forbidden regions.

As in allowed regions, since the curvature is proportional to the kinetic energy, where the negative kinetic energy is large, the curving away from the zero line will be rapid; that is, over a relatively small distance x the wavefunction value can move relatively far from the zero line. Further, this movement away from the zero line will be amplified as the wavefunction value increases, since it multiplies the kinetic energy to determine the curvature. Conversely, where the negative kinetic energy is small, the curving away from the zero line will be slow; that is, a relatively larger distance in x will be needed for the wavefunction to change appreciably. But again, because curvature

depends on the product of kinetic energy and wavefunction, as the wavefunction becomes larger, the movement away from the zero line will accelerate.

■ Quantization

The different behavior of wavefunctions in allowed and forbidden regions is the origin of *energy quantization*. To see why this is so, we need to recall the rule which associates wavefunctions with particles.

$|\psi(x)|^2 dx$ is probability of particle being within dx of x

Now, the total probability — the probability of the particle being *somewhere* — must be unity,

$$1 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx$$

and this so-called *normalization requirement* can be satisfied only if the wavefunction, $\psi(x)$, does not diverge to infinite positive or negative values. Since we know that the wavefunction curves away from the zero line in the forbidden region, we can insure that it does not grow to infinite values only by arranging for it to enter forbidden regions with just the right curvature so that it curves away from the axes exponentially slowly.

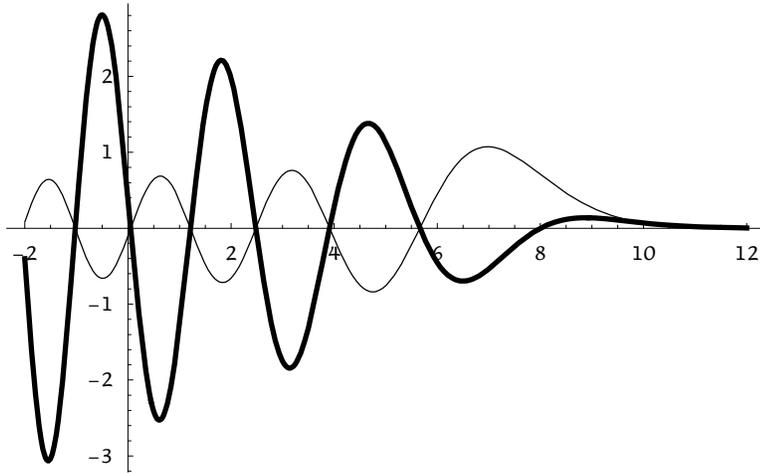
For example, say the wavefunction enters a forbidden region with a positive value but negative slope. Since its slope entering the forbidden region is downward, its value will initially continue to decrease in the forbidden region. If its curvature is too great, the decrease will slow to a stop and then increase toward infinite positive values. If its curvature is too small, the decrease will carry the function to the zero line, then to negative values, and thereafter the function will increase towards infinite negative values. Only if the curvature is just right will the function decrease toward the zero line at an ever slowing rate (since the function itself is decreasing and the curvature is proportional to the value of the function), never getting to the zero line.

The way to adjust the curvature so that exponential decay is achieved is by adjusting the total energy, E . The special values of the energy for which the wavefunction decays to zero in forbidden regions are precisely the possible energies of the system.

■ Summary and study problem

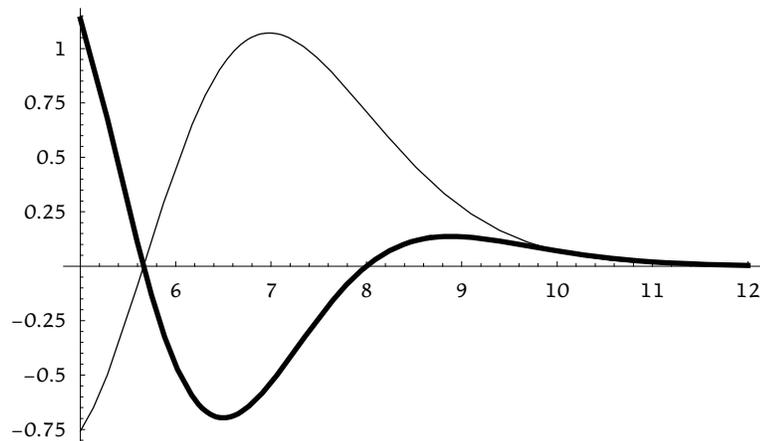
We have seen that wavefunctions ψ that satisfy the Schrödinger equation are those that oscillate in allowed regions and curve away from the zero line in forbidden regions. The oscillations in the allowed region are slow and with large amplitude where the kinetic energy is small, and are rapid and with small amplitude where the kinetic energy is large. The curving away from the zero line usually leads to wavefunctions that diverge to infinite values. The normalization conditions means that such wavefunctions, while mathematically acceptable, are physically unacceptable. Only at special values of the energy does the divergence in the forbidden region become an exponential decay to the zero line and so are physically acceptable wavefunctions possible. These special values of the energy are the quantized energies of the system.

Airy functions are physically acceptable solutions of the one dimensional Schrödinger equation for a particular potential energy variation, $V(x)$. Here is a plot of an Airy function (thin line) and its curvature (thick line).



Airy function (thin line) and its curvature (thick line).

Here is a more detailed view of the function and its curvature in the region $5 \leq x \leq 12$.



Airy function (thin line) and its curvature (thick line).

Use the ideas developed in these notes and the details in these plots to make a sketch of the corresponding potential energy variation, $V(x)$, and indicate on your sketch the total energy and the position of the classical turning point.