Reflection and transmission at a change in medium

October 26, 2003

Cornell University

Department of Physics

Physics 214 October 26, 2003

Contents

List of Figures

1 Motivation

The previous set of notes discussed wave reflections generated from boundaries. The phenomenon of reflection is far more general and occurs (except under very special circumstances) whenever there is a change in the medium of propagation. Practical applications of this include RADAR, where microwaves propagating through the atmosphere encounter another material, such as the metallic body of an aircraft or even the flesh and bone of a flock of birds, whose propagation properties are different. The resulting reflection then returns to the RADAR source, where the presence of the reflection and thus the presence of something in the air is detected. Another application is in medical ultrasound applications. Here, high frequency sound waves traveling through one material, such as amniotic fluid, encounters another material which different propagation properties, such as a baby. Again, the change in propagation generates a reflection, which can be measured to study how the baby is developing.

At a general change in medium, not all of the wave is reflected. Some of it is transmitted into the new medium, where it continues to propagation. The purpose of these notes is to understand both processes, *transmission* and *reflection*.

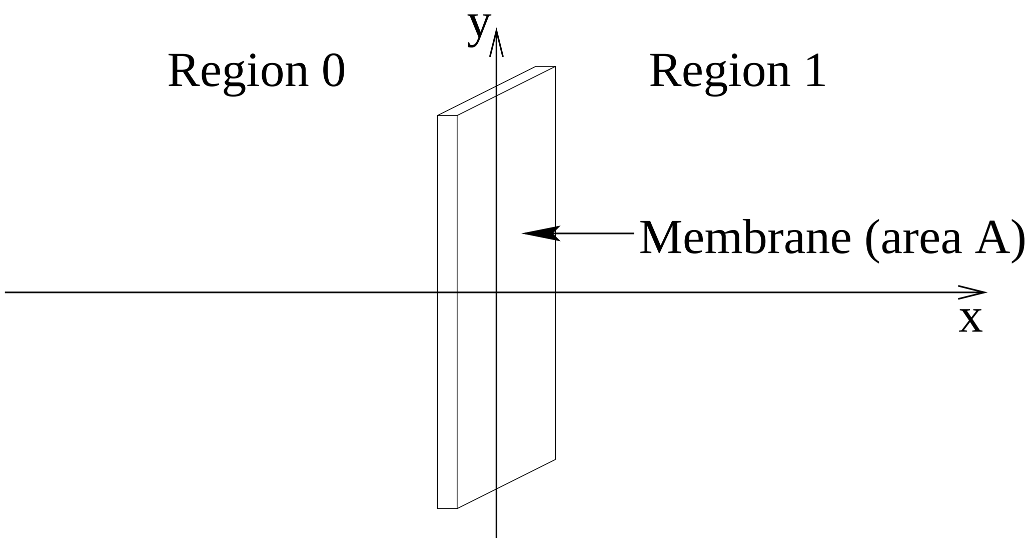

Figure 1: Physical realization for wave behavior at a change in medium: a massless membrane of cross-sectional area A separates regions filled with two different media.

2 Physical realization

As described above, transmission and reflection from a change in medium is quite general and applies to all types of waves. For concreteness, we focus on the propagation of sound as an example. As we move along in the analysis we will also describe briefly the analogous results for other types of waves in a series of footnotes.

Figure 1 illustrates our realization. The system consists of two regions, Region 0 (all points $x < 0$) and Region 1 (all points $x > 0$). The material in Region 0 may be a gas, fluid, or solid, and is characterized by an equilibrium mass density, bulk modulus and pressure of ρ_0, B_0, and P_0, respectively. Region 1 consists of some other medium with corresponding quantities ρ_1, B_1 and P_1. Finally, if necessary to keep the media from mixing (if they are gas or fluid), we may include a thin, *massless* membrane at $x = 0$. We let this membrane be completely free to move without friction.

Before any waves are introduced into this combined system, we allow the two media to reach equilibrium with each other. In equilibrium, the membrane will not move, and so the net force on it in the x-direction, $F = A \cdot (P_0 - P_1)$, where A is the total area of the membrane, must be zero. Hence, $P_0 = P_1$, and the pressures on either side are in balance. Because these two quantities are equal, we shall henceforth refer to them both as $P \equiv P_0 = P_1$[1].

3 General Analysis

3.1 Degrees of Freedom

The degrees of freedom for this system are the sound displacement for each point in either region and the displacement of the membrane. We denote the displacements in Region 0 as $s_0(x < 0, t)$ and in Region 1 as $s_1(x > 0, t)$. The membrane follows the motion of the material of the neighboring points on either side of it. Its position, $x_{\text{mem}}(t)$ thus can be determined equally well in two different ways, either by considering Region 0 or Region 1, $x_{\text{mem}}(t) = s_0(x = 0, t) = s_1(x = 0, t)$.

Thus, the degrees of freedom (minimal set of variables needed to describe the state of the system at any time t) for the combined system are the values $s_0(x \leq 0, t)$ and $s_1(x \geq 0, t)$, with the extra constraint

[1] The analogous result for waves on strings would be that the two applied tensions are in balance, $\tau_0 = \tau_1$. For electromagnetic waves, there is no analogous balance: ϵ and μ may take any value on either side of the boundary.

(because this description is slightly redundant) of *consistency*,

$$s_0(x=0,t) = s_1(x=0,t), \tag{1}$$

which, mathematically, is considered our first *boundary condition* at $x = 0$.[2]

3.2 Equations of Motion

As there are no long range forces in this problem, all points interior to each boundary feel forces only from their neighboring points, and so all interior points obey the same equation of motion we have already found for such media, namely the wave equation. Thus, we have

$$\begin{aligned} \frac{\partial^2 s_0(x,t)}{\partial t^2} &= c_0^2 \frac{\partial^2 s_0(x,t)}{\partial x^2} \\ \frac{\partial^2 s_1(x,t)}{\partial t^2} &= c_1^2 \frac{\partial^2 s_0(x,t)}{\partial x^2}, \end{aligned} \tag{2}$$

where $c_0 = \sqrt{B_0/\rho_0}$, $c_1 = \sqrt{B_1/\rho_1}$.[3]

The only remaining equation of motion is that for the membrane $x_{\mathrm{mem}}(t) = s_0(x=0,t) = s_1(x=0,t)$. Considering motion only in the x-direction and using the facts that the membrane is *massless* and that the equilibrium pressures are equal on either side, we find the second boundary condition at $x = 0$,

$$\begin{aligned} m a_x^{\mathrm{c\,of\,\mathbf{m}}} &= \sum F_x^{(\mathrm{ext})} \\ 0 \cdot a_x^{\mathrm{c\,of\,\mathbf{m}}} &= +A\left(P_0 - B_0 \left.\frac{\partial s_0(x,t)}{\partial x}\right|_{x=0}\right) - A\left(P_1 - B_1 \left.\frac{\partial s_1(x,t)}{\partial x}\right|_{x=0}\right) \\ 0 &= A\left(-B_0 \left.\frac{\partial s_0(x,t)}{\partial x}\right|_{x=0} + B_1 \left.\frac{\partial s_1(x,t)}{\partial x}\right|_{x=0}\right) \\ &\Rightarrow \\ B_0 \left.\frac{\partial s_0(x,t)}{\partial x}\right|_{x=0} &= B_1 \left.\frac{\partial s_1(x,t)}{\partial x}\right|_{x=0}. \end{aligned} \tag{3}$$

We shall refer to this second boundary condition as *force balance*.[4]

3.3 General Solution

3.3.1 Solution within each region

We already have the general solutions to (2),

$$\begin{aligned} s_0(x<0,t) &= t_0(x - c_0 t) + a_0(x + c_0 t) \\ s_1(x>0,t) &= t_1(x + c_1 t) + a_1(x - c_1 t). \end{aligned} \tag{4}$$

Here, to properly solve the respective wave equation, we use the wave speed appropriate to each region, c_0 and c_1, respectively. As a matter of notation, rather than labeling the forward and backward traveling pulse shape functions in each region as $f(x-ct)$ and $g(x+ct)$, respectively, we now name them so that $t_0(x - c_0 t)$ describes a pulse in Region 0 traveling *toward* the interface at speed c_0, $a_0(x + c_0 t)$ describes a pulse in Region 0 traveling *away* from the interface, and $t_1(x + c_1 t)$ and $a_1(x - c_1 t)$ describe pulses moving in Region 1 either *toward* or *away* from the interface, respectively. We introduce this labeling here because,

[2] Analogously, for strings we would have as the degrees of freedom $y_0(x \le 0,t)$, $y_1(x \ge 0,t)$ with the boundary condition $y_0(x=0,t) = y_1(x=0,t)$. For electromagnetic waves traveling along x, one finds as the degrees of freedom $\vec{E}_0(x \le 0,t)$, $\vec{E}_1(x \ge 0,t)$; $\vec{B}_0(x \le 0,t)$, $\vec{B}_1(x \ge 0,t)$, with constraint $\vec{E}_0(x=0,t) = \vec{E}_1(x=0,t)$.

[3] Analogously, for strings and electromagnetic waves, we find also our standard wave equations for all points $x \ne 0$.

[4] For strings, the analogous quantity to pressure is the y-component of the tension T_y, and we find $\tau \partial y_0/\partial x = \tau \partial y_1/\partial x$ at $x = 0$. In electromagnetic theory, we find that $(1/\mu_0)\vec{B}_0 = (1/\mu_1)\vec{B}_1$.

ultimately, we are interested in knowing what reflections and transmissions comes out from the interface, given what we send into it.

There are two reasons why (2) is not a general solution to the overall set of equations of motion. First, it has too many arbitrary parameters. Each point on either side of the interface (or molecule, if you prefer to count discrete objects) obeys a second order (in time) equation of motion. The general solution thus must have two arbitrary parameters for each point in space. However, there are *four* arbitrary functions in (4), each of which can take any value for each value of its argument. Thus, there are four adjustable parameters per point in space, rather than just two. We must therefore be able to eliminate two of the four functions in (4) using some other constraints in order to leave ourselves with just two adjustable functions.

The second difficulty with (4) gives us the additional constraints to resolve the first difficulty. Although (4) solves the equations of motion for the interior points, it does not necessarily solve the equations of motion for the membrane (1,3). These conditions translate into relations among the $t()$'s and $a()$'s which will allow us to eliminate the extra freedom while satisfying the equation of motion for all points in the system.

3.3.2 Imposing conditions at the boundary

Substituting (4) into the consistency boundary condition (1) gives,

$$\begin{aligned} t_0((x=0)-c_0t)+a_0((x=0)+c_0t) &= t_1((x=0)+c_1t)+a_1((x=0)-c_1t) \\ t_0(-c_0t)+a_0(c_0t) &= t_1(c_1t)+a_1(-c_1t), \end{aligned} \tag{5}$$

where we have used the fact that (1) applies only at the membrane point $x=0$. This gives one equation for the four adjustable functions, with which we could eliminate one function, leaving three adjustable functions. We thus require only one more relation among the four unknown functions.

To generate this relation, we next substitute (4) into the force balance boundary condition (3),

$$\begin{aligned} B_0\left(t_0'(x-c_0t)+a_0'(x+c_0t)\right|_{x=0} &= B_1\left(t_1'(x+c_1t)+a_1'(x-c_1t)\right|_{x=0} \\ B_0\left(t_0'(-c_0t)+a_0'(c_0t)\right) &= B_1\left(t_1'(c_1t)+a_1'(-c_1t)\right), \end{aligned} \tag{6}$$

where $f'(u)$ indicates the derivative of the function $f(u)$ with respect to its argument. We now have one additional equation, but have introduced four additional functions, the derivatives t_0', a_0', t_1', and a_1'. This leaves us in the undesirable position of having two equations in *eight* unknown functions. But, if we integrate both sides of (6) with respect to time, we can eliminate the derivatives, and have the second equation which we need among just the original four adjustable functions. To integrate, we guess at the result, check our guess by taking the derivative, and then insert constants as needed to make the result match the original equation. This procedure gives

$$B_0\left(\frac{1}{-c_0}t_0(-c_0t)+\frac{1}{c_0}a_0(c_0t)\right)=B_1\left(\frac{1}{c_1}t_1(c_1t)+\frac{1}{-c_1}a_1(-c_1t)\right)+\mathcal{C}, \tag{7}$$

where $\mathcal{C}$ is the constant of integration and the factors $\pm 1/c_{0,1}$ have been inserted to ensure that the derivative of (7) with respect to time is precisely (6). To determine the value of the integration constant, we consider what happens before any waves are introduced, $t\to-\infty$. For pulses of finite width, $t(\pm\infty), a(\pm\infty)\to 0$. Thus, the integration constant above must be $\mathcal{C}=0$, and we may ignore it safely.

As a final simplification, we note that we can write both sides in terms of the combinations of constants

$$\begin{aligned} Z_0 &= \frac{B_0}{c_0} \\ Z_1 &= \frac{B_1}{c_1}. \end{aligned} \tag{8}$$

Because this combination of constants appears so frequently in the study of waves, it has a special name, the *impedance*. Because $c=\sqrt{B/\rho}$, the impedance is often written in the equivalent form $Z=B/c=(\rho c^2)/c=\rho c$.[5] With these definitions, we then rewrite (7) compactly as

$$Z_0\left(-t_0(-c_0t)+a_0(c_0t)\right)=Z_1\left(t_1(c_1t)-a_1(-c_1t)\right). \tag{9}$$

[5] For sound we find the combination $Z\equiv\tau/c=(\mu c^2)/c=\mu c$ on each side of the equation. In electromagnetic theory, we find $Z\equiv(1/\mu)/c=(1/\mu)/\sqrt{1/(\epsilon\mu)}=\epsilon c$.

Finally, we can use the two equations (5) and (9) to eliminate two of the adjustable functions. Because, as mentioned above, we generally study what comes out from the boundary given what goes into it, the standard choice is to use the boundary conditions to solve for the two "away" functions in terms of the two "toward" functions. To solve for a_0, we choose a linear combination of (5) and (9) which eliminates a_1, $Z_1\times(5)+(9)$:

$$\begin{aligned}(Z_1 - Z_0)t_0(-c_0 t) + (Z_1 + Z_0)a_0(c_0 t) &= 2Z_1 t_1(c_1 t)\\ &\Rightarrow\\ a_0(c_0 t) &= \frac{Z_0 - Z_1}{Z_0 + Z_1}t_0(-c_0 t) + \frac{2Z_1}{Z_0 + Z_1}t_1(c_1 t)\end{aligned} \tag{10}$$

Similarly, to find a_1, we take the linear combination which eliminates a_0, $Z_0\times(5)-(9)$:

$$\begin{aligned}2Z_0 t_0(-c_0 t) &= (Z_0 - Z_1)t_1(c_1 t) + (Z_0 + Z_1)a_1(-c_1 t)\\ &\Rightarrow\\ a_1(-c_1 t) &= \frac{Z_1 - Z_0}{Z_0 + Z_1}t_1(c_1 t) + \frac{2Z_0}{Z_0 + Z_1}t_0(-c_0 t).\end{aligned} \tag{11}$$

3.3.3 Interpretations

Eqs. (10,11) give all the information about the "away' traveling functions. If, for instance, we wanted to know the value $a_0(3\,\mathbf{m})$, we could look at (10) at time $t = (3\,\mathbf{m})/c_0$, to find that

$$\begin{aligned}a_0(3\,\mathbf{m}) &= \frac{Z_0 - Z_1}{Z_0 + Z_1}t_0\left(-c_0\frac{3\,\mathbf{m}}{c_0}\right) + \frac{2Z_1}{Z_0 + Z_1}t_1\left(c_1\frac{3\,\mathbf{m}}{c_0}\right)\\ &= \frac{Z_0 - Z_1}{Z_0 + Z_1}t_0\left(-3\,\mathbf{m}\right) + \frac{2Z_1}{Z_0 + Z_1}t_1\left(\frac{c_1}{c_0}\cdot 3\,\mathbf{m}\right).\end{aligned}$$

In more general terms, if we wanted to learn something about the value of $a_0(u)$ for any point u, we would look at (10) at the time $t = u/c_0$, to find

$$a_0(u) = \frac{Z_0 - Z_1}{Z_0 + Z_1}t_0(-u) + \frac{2Z_1}{Z_0 + Z_1}t_1(\frac{c_1}{c_0}u). \tag{12}$$

Note that this analysis corresponds precisely to the mathematical "trick" introduced in lecture of defining $u = c_0 t$. Following the same logic, to find the value of $a_1(v)$ for any point v, we should look at time $t = -v/c_1$ to learn from (11) that

$$a_1(v) = \frac{Z_1 - Z_0}{Z_0 + Z_1}t_1(-v) + \frac{2Z_0}{Z_0 + Z_1}t_0(\frac{c_0}{c_1}v). \tag{13}$$

To interpret the results (12-13),[6] we note that the pulse a_0 (a_1) traveling away from the interface into Region 0 (Region 1) has two contributions. The first comes from t_0 (t_1), the pulse originally traveling in Region 0 (Region 1) toward the interface, which evidently has been reflected by the interface back into Region 0 (Region 1). The "$-u$" ("-v") in the argument of this pulse tells us that the reflected pulse is flipped horizontally, corresponding to the physical observation that the leading edge of a pulse should be the first part that returns in a reflection. The reflected pulse is also scaled vertically by a "reflection amplitude" which we denote $R_{0\to 1} = (Z_0 - Z_1)/(Z_0 + Z_1)$ ($R_{1\to 0} = (Z_1 - Z_0)/(Z_0 + Z_1)$) to indicate reflection of a pulse traveling from Region 0 toward Region 1 (from Region 1 toward Region 0).

The second contribution to a_0 (a_1) comes from the pulse t_1 (t_0) originally traveling in Region 1 (Region 0) toward the interface, which is evidently then transmitted into Region 0 (Region 1). The "$(c_1/c_0)u$" ("$(c_0/c_1)v$") in the argument of this pulse tells us that the transmitted pulse is horizontally stretched out by a factor of c_0/c_1 (c_1/c_0). This corresponds to the physical observation that a pulse entering a region of higher wave speed will be stretched out as the leading edge of the pulse rushes ahead into the new region where

[6] Rather than writing the same two paragraphs twice, once for a_0 and once for a_1, we have written this paragraph just once for a_0, including everything that we would say differently for a_1 in parentheses.

the waves move faster. Finally, the transmitted pulse is scaled vertically by a "transmission amplitude" for pulses which we shall denote $T_{1\to 0} = (2Z_1)/(Z_0 + Z_1)$ ($T_{0\to 1} = (2Z_0)/(Z_0 + Z_1)$) to indicate transmission of a pulse from Region 1 into Region 0 (from Region 0 into Region 1),

The primary lessons of the above two paragraphs are that, in traveling from medium a to medium b, regardless of whether travel is from left to right or from right to left, reflected pulses always flip horizontally, and transmitted pulses always stretch horizontally by a factor c_b/c_a, where c_a and c_b are the respective wave speeds. (Note that, the pulse is actually compressed if this factor is a number less than one, like 0.8.) Moreover, reflected pulses scale vertically by a reflection amplitude

$$R_{a\to b} = \frac{Z_a - Z_b}{Z_a + Z_b}, \tag{14}$$

and transmitted pulses scale vertically by a transmission amplitude

$$T_{a\to b} = \frac{2Z_a}{Z_a + Z_b}, \tag{15}$$

where Z_a and Z_b are the impedances of the respective materials.

3.3.4 Final form for general solution

Finally, we may substitute the forms (12,13) for the outward moving pulses into (4) to find a valid general solution, one which satisfies the internal equations of motion for each region and for the membrane and which contains the required number of adjustable parameters (two adjustable functions). Using the definitions (14,15) of the various reflection/transmission amplitudes, the final result is

$$\begin{aligned} s_0(x \le 0, t) &= t_0(x - c_0 t) + R_{0\to 1} t_0\left(-(x + c_0 t)\right) + T_{1\to 0} t_1\left(\frac{c_1}{c_0}(x + c_0 t)\right) \\ s_1(x \ge 0, t) &= t_1(x + c_1 t) + R_{1\to 0} t_1\left(-(x - c_1 t)\right) + T_{0\to 1} t_0\left(\frac{c_0}{c_1}(x - c_1 t)\right). \end{aligned} \tag{16}$$

4 Particular Solutions

From the final general solution (16), it is relatively easy to determine particular solutions for when a single pulse is incident either from Region 0 or Region 1, thereby confirming the interpretations giving in Section 3.3.3. For instance, if a pulse of shape $f(u)$ comes in from Region 0, then $t_0(u) = f(u)$ and $t_1(v) = 0$. Thus,

$$\begin{aligned} s_0(x \le 0, t) &= f(x - c_0 t) + R_{0\to 1} f\left(-(x + c_0 t)\right) \\ s_1(x \ge 0, t) &= T_{0\to 1} f\left(\frac{c_0}{c_1}(x - c_1 t)\right), \end{aligned} \tag{17}$$

Thus, we have a reflected pulse, horizontally flipped and vertically scaled by $R_{0\to 1}$, traveling to the left in Region 0 with speed c_0, and we have a transmitted pulse, stretched horizontally by c_1/c_0 and scaled vertically by $T_{0\to 1}$, traveling to the right in Region 1 at speed c_1. Alternately, if a pulse of shape $f(u)$ comes in from Region 1, then $t_1(u) = f(u)$ and $t_0(v) = 0$, and

$$\begin{aligned} s_0(x \le 0, t) &= T_{1\to 0} f\left(\frac{c_1}{c_0}(x + c_0 t)\right) \\ s_1(x \ge 0, t) &= f(x + c_1 t) + R_{1\to 0} f\left(-(x - c_1 t)\right). \end{aligned} \tag{18}$$

In this case, we then have a reflected pulse, horizontally flipped and vertically scaled by $R_{1\to 0}$, traveling to the right in Region 1 at speed c_1, and a transmitted pulse, horizontally stretched out by a factor c_0/c_1 and vertically scaled by $T_{1\to 0}$, traveling to the left in Region 0 at speed c_0.