Lecture topics for 18376.

R. R. Rosales (MIT, Math. Dept.)

September 3, 2023

These are brief, and not-so-brief, 18.376 lecture summaries. Further details can be found in the course web page notes, books, etc. These summaries will be updated from time to time. Check the date.

Contents

List of Figures

1 Reflection and Transmission of waves

In this section we consider various topics related to transmission and reflection of waves in various contexts.

1.1 Properties of R and T in a 1-D example

Consider the following 1-D example dispersive equation, with localized variable coefficients

$$u_{tt} - (c^2\, u_x)_x + V\, u = 0, \tag{1.1}$$

where $\boldsymbol{c > 0}$, $\boldsymbol{c = c(x) \to 1}$ and $\boldsymbol{V = V(x) \to 0}$ (fast enough) as $\boldsymbol{|x| \to \infty}$. Imagine now that c and V are not known (beyond their asymptotic behavior), and we want to determine them via measurements "at infinity". Because for $|x| \gg 1$ the solutions can be written in terms of simple exponential waves, the question can be formulated in a mathematically precise wave via the *scattering problem* formulated below.

Define the left *reflection and transmission coefficients*, $\boldsymbol{R_1 = R_1(\omega)}$ and $\boldsymbol{T_1 = T_1(\omega)}$, via the solutions to the equation above characterized by

$$u \sim e^{i\,\omega\,(x-t)} + R_1\, e^{-i\,\omega\,(t+x)} \ \text{ for } x \ll -1, \quad \text{and} \quad u \sim T_1\, e^{i\,\omega\,(x-t)} \ \text{ for } x \gg 1. \tag{1.2}$$

This corresponds to sending a wave (from the left) into the system, and seeing how much comes back, and how much goes through. † *An interesting question is now: suppose that we know $R_1(\omega)$ for all values $-\infty < \omega < \infty$. Is this enough information to determine c and V?* The answer to this turns out to be: no. In fact, *even if we know that $c \equiv 1$, knowledge of R_1 is still not enough to determine V* — though, in this $c = 1$ case, it is known what the missing information is (and how to recover V once this extra information is provided). Proving these results is beyond the scope of these notes. However, in § 1.1.1 we state what the needed extra information is when $c = 1$.

† Note: the time dependence of all the waves must equal that of the input wave.

The question above is a "simple" example of a situation that arises in many applications: there is some closed system to which we do not have direct access, but the system can be probed by sending waves into it, and measuring what comes out after the waves go through. Can the structure of the system be ascertained from this *scattering data*? This is, quite often, a hard mathematical problem, with many less answers known than problems are out there.

Right *reflection and transmission coefficients*, $\boldsymbol{R_2 = R_2(\omega)}$ and $\boldsymbol{T_2 = T_2(\omega)}$, via the solutions characterized by

$$u \sim e^{-i\,\omega\,(t+x)} + R_2\, e^{i\,\omega\,(x-t)} \ \text{ for } x \gg 1, \quad \text{and} \quad u \sim T_2\, e^{-i\,\omega\,(t+x)} \ \text{ for } x \ll -1. \tag{1.3}$$

The right coefficients are the left coefficients for the
mirror problem: $c(x) \to c(-x)$ and $V(x) \to V(-x)$. (1.4)

Our objective in this subsection is to derive some properties of these coefficients, and to provide physical interpretations for their meaning (when one is known to me). Unfortunately, this will require the introduction of mathematical machinery whose purpose cannot be understood "a priori" — though I hope you will see, at least partly, the why after going through to the end of the subsection.

We begin by separating the time dependence in the equation, in the form $\boldsymbol{u = e^{-i\,\omega\,t}\,\phi(x)}$. Then the equation becomes the ode

$$(c^2\,\phi')' + (\omega^2 - V)\,\phi = 0, \tag{1.5}$$

while the transmission and reflection coefficients are defined by

$$\phi \sim e^{i\,\omega\,x} + R_1\, e^{-i\,\omega\,x} \ \text{ for } x \ll -1, \quad \text{and} \quad \phi \sim T_1\, e^{i\,\omega\,x} \ \text{ for } x \gg 1. \tag{1.6}$$

$$\phi \sim e^{-i\,\omega\,x} + R_2\, e^{i\,\omega\,x} \ \text{ for } x \gg 1, \quad \text{and} \quad \phi \sim T_2\, e^{-i\,\omega\,x} \ \text{ for } x \ll -1. \tag{1.7}$$

Technical detail: R_j and T_j are not defined for $\omega = 0$. Why: (1, math.) $e^{\pm\, i\,\omega}$ are no longer linearly independent solutions of the equation. (2, physics) the notion of wave traveling right/left is lost.

Define now the *left and right Jost functions* $J_L(x,\,\omega)$ and $J_R(x,\,\omega)$, $-\infty < \omega \neq 0 < \infty$, by

1. J_L solves (1.5) and $J_L \sim e^{-i\,\omega\,x}$ for $x \ll -1$. (1.8)

2. J_R solves (1.5) and $J_R \sim e^{+i\,\omega\,x}$ for $x \gg 1$. (1.9)

J_L and J_R are also defined for ω complex, with $Im(\omega) > 0$.
Further: it can be shown that they are both analytic there. (1.10)

These two functions are connected to the reflection and transmission coefficients via

$$T_1\, J_R = J_L^* + R_1\, J_L \qquad \text{and} \qquad T_2\, J_L = J_R^* + R_2\, J_R, \tag{1.11}$$

where the star denotes complex conjugates.

We now recall, from the theory of ode, that the Wronskian of any two solutions to (1.5) satisfies (use the equation to show $\frac{\mathrm{d}W}{\mathrm{d}x} = 0$)

$$W(\phi_1,\,\phi_2) = c^2\,\phi_1\,\phi_2' - c^2\,\phi_2\,\phi_1' = \text{constant}, \tag{1.12}$$

where $W = 0$ if and only if the solutions are linearly dependent. In particular, since the complex conjugate of a solution is a solution,

$$W(J_L,\,J_L^*) = 2\,i\,\omega = -W(J_R,\,J_R^*), \tag{1.13}$$

and (of course) $0 = W(J_R,\,J_R) = W(J_L,\,J_L)$. It then follows that

$$T_1\, W(J_L,\,J_R) = 2\,i\,\omega = T_2\, W(J_L,\,J_R), \quad \Longrightarrow \quad T_1 = T_2 = \frac{2\,i\,\omega}{W(J_L,\,J_R)} = T. \tag{1.14}$$

Proof: Compute $W(J_L,\,T_1\,J_R)$ and $W(T_2\,J_L,\,J_R)$ using (1.11) and (1.13).

3. This formula indicates that **the transmission coefficient depends only on the even part of the coefficients.** T is the same for both $[c(x),\,V(x)]$ and $[c(-x),\,V(-x)]$ — see (1.4).

4. $1/T$ is *analytic in the upper half plane,* as follows from (1.9), and the definition of W. Further: **The zeros of $1/T$ correspond to eigenvalues of the operator $\mathcal{L}\,\phi = -(c^2\,\phi')' + V\,\phi$ in (1.5). The eigenvalue is ω^2, and the eigenfunction is either J_L or J_R.**

Why? Because at a zero $W(J_L,\,J_R) = 0$, so that they are proportional to each other. Furthermore, if $\mathrm{Im}(\omega) > 0$, J_L decays exponentially as $x \to -\infty$ and J_R decays exponentially as $x \to \infty$. Hence at a zero of $1/T$ both decay exponentially as $|x| \to \infty$; that is, they are eigenfunctions.

It can be shown that:

(a) The zeros are purely imaginary, $-i\,\omega > 0$.

(b) There is only a finite number of zeros.

(c) The zeros are simple — i.e.: T has a simple pole there.

It should be clear that $W(T_1\,J_R, T_1^*\,J_R^*) = -2\,i\,\omega\,|T_1|^2$, using (1.13). On the other hand, from (1.11), $W(T_1\,J_R, T_1^*\,J_R^*) = W(J_L^* + R_1\,J_L,\, J_L + R_1^*\,J_L^*) = -2\,i\,\omega + 2\,i\,\omega\,|R_1|^2$. Thus

$$|T_1|^2 + |R_1|^2 = 1. \quad \text{Similarly} \quad |T_2|^2 + |R_2|^2 = 1. \tag{1.15}$$

This is the **conservation of energy** by the waves. We can also show that

$$\boldsymbol{T \neq 0 \text{ for any real } \omega \neq 0, \text{ hence } |R_j| < 1.} \tag{1.16}$$

In other words: total reflection is not possible.

Proof: If $T = 0$, then $J_L^* = -R_1\,J_L$, from (1.11). Not possible since J_L^* and J_L are linearly independent.

Finally, from (1.11) and (1.13),

$$T_1\,W(J_R,\,J_L^*) = W(T_1\,J_R,\,J_L^*) = 2\,i\,\omega\,R_1 \quad \text{and} \quad T_2\,W(J_L,\,J_R^*) = W(T_2\,J_L,\,J_R^*) = -2\,i\,\omega\,R_2.$$

Thus $(2\,i\,\omega\,R_1/T_1)^* = 2\,i\,\omega\,R_2/T_2$, that is

$$\frac{R_1^*}{T_1^*} + \frac{R_2}{T_2} = 0. \tag{1.17}$$

This, again, relates the left and right scattering data. But I do not know what it means physically.

1.1.1 The case $c = 1$

The case where $c \equiv 1$ is of particular interest. It arises not just as a special case for (1.1), but directly from the 1-D Schrödinger equation

$$i\,\psi_t = -\psi_{xx} + V\,\psi \tag{1.18}$$

scattering problem. Then the separated problem takes the form $\phi'' + \omega^2\,\phi = V\,\phi.$ (1.19)

This leads to the following Volterra integral equation for the Jost function J_R (a similar equation applies for J_L)

$$J_R = e^{i\,\omega\,x} - \int_x^\infty \frac{\sin(\omega(x-y))}{\omega}\,V(y)\,J_R(y)\,dy = e^{i\,\omega\,x} + \mathcal{S}\,J_R, \tag{1.20}$$

where the operator $\mathcal{S}$ is defined by the second equality. The solution to this equation is given by the convergent series $J_R = \sum_0^\infty \mathcal{S}^n(e^{i\,\omega\,x})$, which can be used to show that J_R is analytic for $\mathrm{Im}(\omega) > 0$.

In this $c = 1$ case *it can be show that the potential V can be recovered from knowledge of:*

(a) The reflection coefficient $R_1(\omega)$, for ω real.

(b) The poles and residues of the transmission coefficient $T_1(\omega)$, in $\mathrm{Im}(\omega) > 0$.

It can be shown that (b) is equivalent to knowing the eigenvalues, and normalization constants, for $\lambda\,\phi = \mathcal{L}\,\phi$ — where $\mathcal{L}\,\phi = -\phi'' + V\,\phi$. The normalization constants γ_n are defined as follows: let $\lambda_n = -\mu_n^2 < 0$ be an eigenvalue ($\mu > 0$), with corresponding eigenfunction ϕ_n normalized by $\phi_n \sim e^{\mu\,x}$ as $x \to -\infty$. Then

$$\gamma_n = \left(\int_{-\infty}^{\infty} \phi_n^2 dx\right)^{-1} \tag{1.21}$$

is the normalization constant corresponding to λ_n.

***Exercise 1:* Example of a nontrivial potential with $R = 0$**

Here we present an example of a vanishing reflection coefficient that corresponds to a nontrivial situation. Specifically, consider the equation

$$-\phi'' + V\,\phi = \omega^2\,\phi, \tag{1.22}$$

where $V = V(x)$ vanishes rapidly as $|x| \to \infty$. Define the transmission and reflection coefficients by [1]

$$\phi \sim e^{i\,\omega\,x} + R\,e^{-i\,\omega\,x} \ \text{ for } x \ll -1, \quad \text{and} \quad \phi \sim T\,e^{i\,\omega\,x} \ \text{ for } x \gg 1, \tag{1.23}$$

where $-\infty < \omega \neq 0 < \infty$. Then take

$$V = \frac{-2\,e^{-x}}{(1+e^{-x})^2} = -\frac{1}{4}\,\mathrm{sech}^2\left(\frac{x}{2}\right), \tag{1.24}$$

and **compute R, T, and the eigenvalues for (1.22)**.
Recall: the eigenvalues occur at the poles of T in $\mathrm{Im}(\omega) > 0$. In addition, **verify that $|R|^2 + |T|^2 = 1$**.

Hints:

(a) Calculate the right Jost function J, defined as the solution to (1.22) that satisfies $J \sim e^{+i\,\omega\,x}$ for $x \gg 1$ — here we suppress the sub-index R in J_R to simplify the notation. The solution in (1.23) is then $\phi = T\,J$, for a T selected so that the desired behavior occurs for $x \ll -1$.

(b) To compute J, write $J = \psi(x)\,e^{i\,\omega\,x}$, where $\psi \sim 1$ for $x \gg 1$, and write the equation for ψ.

(c) V can be written in the form $V = z' = \dfrac{1}{2}\,z^2 - z$, where $z = \dfrac{2\,e^{-x}}{1+e^{-x}}$. Note that $z'' = (z-1)\,z'$.

(d) Since z is a strictly monotone function of x, with $z' < 0$, a change of variables from x to z is allowed. Write the equation for ψ obtained in (b) in terms of z. That is, think of ψ as a function of z, $\psi = \psi(z)$ — with $\psi(0) = 1$. The solution ψ that we are looking for, in this new variable, is *very simple*! You can find it by expanding $\psi = \sum_n a_n\,z^n$, with $a_0 = 1$.

---

2 Radiation damping

In this section we present examples where radiation damping plays a role.

2.1 Semi-infinite string with mass-spring at end

Consider a semi-infinite string under tension, with a mass-spring system attached at its end. Then

$$\rho\,u_{tt} - T\,u_{xx} = 0 \qquad \text{for } \ x > 0, \tag{2.1}$$

$$M\,u_{tt} + k_s\,u = T\,u_x \qquad \text{at } \ x = 0, \tag{2.2}$$

where we assume in-plane motion — see figure 2.1. Introduce the notation (the meaning of ν is made clear below)

$$\boldsymbol{c = \sqrt{\frac{T}{\rho}}}\ \text{(wave speed)}, \qquad \boldsymbol{\Omega = \sqrt{\frac{k_s}{M}}}\ \text{(mass-spring frequency)}, \qquad \text{and} \qquad \boldsymbol{\nu = \frac{T}{2\,c\,M} = \frac{\sqrt{\rho\,T}}{2\,M}}.$$

[1] These are the left coefficients, where we suppress the sub-index 1 to simplify the notation.

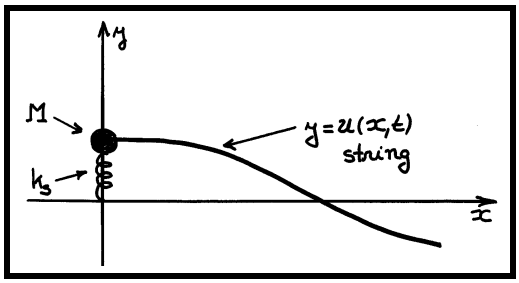


Figure 2.1: String with mass-spring system

Sketch of the string under tension with a mass-spring system attached at its end. The physical parameters are: (i) The string density $\rho > 0$ (assumed constant); (ii) The string tension $T > 0$; (iii) The mass at the end $M > 0$; and (iv) The spring constant for the mass-spring system, $k_s > 0$.

Then

$$u_{tt} - c^2\, u_{xx} = 0 \qquad \text{for } x > 0, \tag{2.3}$$

$$u_{tt} + \Omega^2\, u = 2\, c\, \nu\, u_x \qquad \text{at } x = 0. \tag{2.4}$$

Assume a situation where no waves "from infinity" are arriving at the mass-spring system. Then we can write

$$u = y(t - x/c), \tag{2.5}$$

where $\boldsymbol{y}$ = **position of the mass** $\boldsymbol{M}$. Substituting this into (2.4) yields

$$\ddot{y} + 2\,\nu\,\dot{y} + \Omega^2\, y = 0. \tag{2.6}$$

This is the equation for a **damped harmonic oscillator, with damping coefficient** $\boldsymbol{\nu}$. Its general solution is a linear combination of the two solutions $\boldsymbol{y = e^{\lambda_\pm\, t}}$, where

$$\lambda_\pm = -\nu \pm i\,\sqrt{\Omega^2 - \nu^2} \qquad \text{if } \Omega > \nu, \tag{2.7}$$

$$\lambda_\pm = -\nu \pm \sqrt{\nu^2 - \Omega^2} \qquad \text{if } \Omega < \nu. \tag{2.8}$$

Issue: if these solutions are plugged back into (2.5), the resulting *string perturbation blows up exponentially as* $x \to \infty$. This is **not physical!** **Why does this happen?**

Answer: The mass-spring system continuously looses energy to waves moving along the string towards infinity. For the system to have been oscillating "forever", its energy as $t \to -\infty$ must blow up. Since the waves along the string at $(x,\, t)$ originate at the mass-spring system at time $t - x/c$, it is then not surprising that their amplitude should grow exponentially with x. *We get a non-physical answer because a damped oscillator vibrating "forever" into the past is not physical.* The situation only makes sense if we start (2.6) at some given time t_0, so that y vanishes for $t < t_0$.

Remark 2.1 *Note that the damping coefficient goes down as the mass increases.* The interpretation is that the string is able to carry energy away at a rate that depends on the string only — hence a heavier mass looses a smaller fraction of its energy per unit time. Note also that the damping is larger for heavier strings, or a higher tension. ♣

A more realistic endeavor is to consider the initial value problem for (2.3–2.4), perhaps with a force applied to the mass, so that (2.4) is changed to:

$$u_{tt} + \Omega^2\, u = 2\, c\, \nu\, u_x + G(t), \tag{2.9}$$

where $F = M\, G$ is the applied force. Then we write ..... $\boldsymbol{u = y(t - x/c) + g(t + x/c)}$, where

1. $g = g(\tau)$, $\tau > 0$ is the wave moving left .....generated by the initial conditions.

2. $y = y(\tau)$, $\tau < 0$ is the wave moving right ..... generated by the initial conditions.

3. $y = y(\tau)$, $\tau > 0$ is the wave moving right ..... generated by the mass-spring system. Below we write an equation determining this wave.

Substituting the solution u above into (2.9), the equation for the mass-spring system results

$$\ddot{y} + 2\,\nu\,\dot{y} + \Omega^2\,y = G(t) - \ddot{g} + 2\,\nu\,\dot{g} - \Omega^2\,g = G_T(t), \tag{2.10}$$

where G_T is defined by the second equality, and g is evaluated at t — i.e.: $g = g(t)$, etc. This is the equation for a **damped harmonic oscillator, driven by the applied force and the wave arriving from the right (produced by the initial conditions).** Note that now

4. There is no exponential growth as $x \to \infty$.

5. The mass-spring system influences only the region $0 < x < ct$ of the string.

6. If $G = G(t)$ vanishes, and $g(\tau) \to 0$ as $\tau \to \infty$ (bounded initial disturbance on the string), then this equation reduces to (2.6) for large enough time.

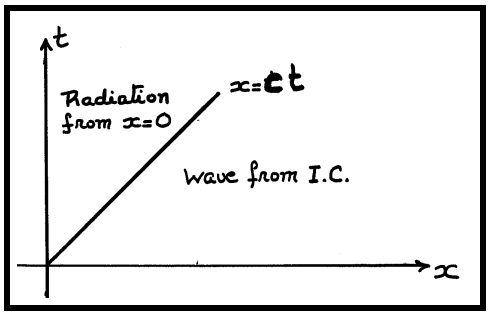


Figure 2.2: Waves in space-time

Space-time diagram for the simple situation where there is no forcing and the initial conditions generate no wave going left. That is: $G = 0$ and $g = 0$ in (2.9–2.10); in particular the initial conditions satisfy $u_t^0 + c\,u_x^0 = 0$. Then $u = y(t - x/c)$ where: (i) $y(\tau)$ solves (2.6) for $\tau > 0$. (ii) $y(\tau) = u^0(-c\,\tau)$ for $\tau < 0$.

2.2 Semi-infinite string over elastic bed with mass-spring at end

Consider a semi-infinite string over an elastic bed, under tension, with a forced mass-spring system attached at its end (assume small, in-plane, motion). The governing equations are

$$\rho\,u_{tt} - T\,u_{xx} + k_b\,u = 0 \qquad \text{for } x > 0, \tag{2.11}$$

$$M\,u_{tt} + k_s\,u = T\,u_x + F \qquad \text{at } x = 0, \tag{2.12}$$

where $\rho > 0$ and $T > 0$ are the string density and tension respectively, $k_b > 0$ is the bed elastic constant, $M > 0$ is the mass at the end, $k_s > 0$ is the spring constant for the mass-spring system, and $F = F(t)$ is the applied force. Define now

$$\boldsymbol{c = \sqrt{\tfrac{T}{\rho}}}, \qquad \boldsymbol{m = \sqrt{\tfrac{k_b}{\rho}}}, \qquad \boldsymbol{\Omega = \sqrt{\tfrac{k_s}{M}}}, \qquad \boldsymbol{\nu = \tfrac{T}{2\,c\,M}}, \qquad \text{and} \qquad \boldsymbol{G = \tfrac{F}{M}}, \tag{2.13}$$

where c is the wave speed, m is the bed frequency, Ω is the mass-spring frequency, ν is a damping coefficient whose meaning is made clear later, and G is the acceleration applied by the force. Then

$$u_{tt} - c^2\,u_{xx} + m^2\,u = 0 \qquad \text{for } x > 0, \tag{2.14}$$

$$u_{tt} + \Omega^2\,u = 2\,c\,\nu\,u_x + G \qquad \text{at } x = 0. \tag{2.15}$$

The dispersion relation is $\boldsymbol{\omega^2 = c^2\,k^2 + m^2}$, with group speed $\boldsymbol{c_g = \frac{d\omega}{dk} = \frac{c^2\,k}{\omega}}$ ($\boldsymbol{|c_g| < c}$ and $\boldsymbol{|c_g| \to c}$ as $\boldsymbol{|k| \to \infty}$). The bed frequency m is the minimum frequency which waves can have.

To simplify matters, we will now select units for which $c = m = 1$. That is, **non-dimensionalize the equations,** so that

$$u_{tt} - u_{xx} + u = 0 \qquad \text{for } x > 0, \tag{2.16}$$

$$u_{tt} + \Omega^2\,u = 2\,\nu\,u_x + G \qquad \text{at } x = 0. \tag{2.17}$$

Remark 2.2 Units and a-dimensional variables. (i) The units for (2.11–2.12) are: $[\rho]$ = mass/length, $[u]$ = length, $[T]$ = force, $[k_b]$ = force/length^2, $[M]$ = mass, $[k_s]$ = force/length, and $[F]$ = force. Note that k_b and k_s do not have the same units, because k_b is a spring constant per unit length. (ii) It follows that the units for (2.13) are: $[c]$ = velocity, $[m]$ = frequency, $[\Omega]$ = frequency, $[\nu]$ = 1/decay time, and $[G]$ = acceleration. (iii) Finally, to obtain (2.16–2.17), introduce a-dimensional variables and constants by: $\tilde{u} = u/\ell_e$, $\tilde{t} = t\,m$, $\tilde{x} = x\,m/c$, $\tilde{\Omega} = \Omega/m$, $\tilde{\nu} = \nu/m$, and $\tilde{G} = G/(m^2\,\ell_e)$ — where ℓ_e is some length scale. Substituting this into (2.14–2.15), and then dropping the tildes yields (2.16–2.17) — since we will no longer use the dimensional variables, this should not cause confusion.
Note: the approximations used to obtain (2.11) require $\ell_e \ll$ typical wavelength $\approx c/m$. Further: we must assume $G/\Omega^2 \ll c/m$, so the forcing does not generate large amplitudes that violate the regime of validity. ♣

2.2.1 Free trapped modes

We now look for solutions to (2.16–2.17), in the absence of forcing (i.e.: $G = 0$), of the form

$$\boldsymbol{u = a\,e^{s\,t - \ell\,x}}, \quad \text{where } \boldsymbol{a},\ \boldsymbol{s} \text{ and } \boldsymbol{\ell} \text{ are constants.} \tag{2.18}$$

This is a solution if and only if

$$\text{(a)}\ \ s^2 = \ell^2 - 1 \quad \text{and} \quad \text{(b)}\ \ s^2 + \Omega^2 + 2\,\nu\,\ell = 0. \tag{2.19}$$

Substituting (a) into (b) to eliminate s yields

$$\ell^2 + 2\,\nu\,\ell + \Omega^2 - 1 = 0. \tag{2.20}$$

Below we examine the various possible cases, and show that ℓ can never be purely imaginary; thus a "pure wave" situation is not allowed.

Case 1. $\Omega^2 \geq 1 + \nu^2$ ..... *no physical solutions.*

Then $\ell_\pm = -\nu \pm i\,\sqrt{\Omega^2 - 1 - \nu^2}$. Thus the roots † have negative real parts, from which it follows that u in (2.18) grows exponentially as $x \to \infty$. Hence these solutions are not physically acceptable.

† There is only one root, negative, for $\Omega^2 = 1 + \nu^2$.

Case 2. $1 < \Omega^2 < 1 + \nu^2$ ..... *no physical solutions.*

Then $\ell_\pm = -\nu \pm \sqrt{1 + \nu^2 - \Omega^2}$. Thus both roots are real and negative, † from which it follows that u in (2.18) grows exponentially as $x \to \infty$. Hence these solutions are not physically acceptable.

† In the limit case $\Omega^2 = 1$, $\ell_+ = 0$. Not physical, since it corresponds to $s = \pm i$ and $u \propto \cos(t - t_0)$ — the whole string, all the way up to infinity, oscillating rigidly up and down at the bed frequency.

Case 3. $\boldsymbol{\Omega^2 < 1}$ ..... ***free trapped oscillatory mode.***

Then $\boldsymbol{\ell_+ = -\nu + \sqrt{1+\nu^2-\Omega^2} > 0}$ gives rise to a solution that decays exponentially as $x \to \infty$. Note also † that $\ell_+ = -\nu + \sqrt{1+\nu^2-\Omega^2} < 1$. Thus $\boldsymbol{s = \pm i\sqrt{1-\ell_+^2} = \pm i\,\omega}$ is pure imaginary. In fact $\boldsymbol{\Omega < \omega < 1}$, since $\omega^2 = 1-\ell_+^2 = \Omega^2 + 2\,\nu\,\ell_+ > \Omega^2$.

† Write the inequality as $\sqrt{1+\nu^2-\Omega^2} < 1+\nu$, note that both sides are positive, and take squares.

A **physical explanation/interpretation for the free trapped mode** is: If the mass-spring system frequency is below that of the string bed, then the mass-spring system cannot generate waves that propagate into the string — the dispersion relation for (2.16) is $\omega = \pm\sqrt{1+k^2}$, there are no waves with wave-frequency below one. Thus the mass-spring system cannot loose energy via radiation. All the energy stays near the mass, and the system oscillates at a frequency between that of the mass-spring system and that of the bed. This last because what oscillates is the mass plus a chunk of the string, so that **a "compromise" frequency arises.**

Note: $\omega^2 = \Omega^2 + 2\,\nu\,\ell_+ = \Omega^2 - 2\,\nu^2 + 2\,\nu\sqrt{1+\nu^2-\Omega^2}$ is an increasing function of Ω^2 for $0 < \Omega^2 < 1$. Thus the *slowest free trapped mode occurs for* $\Omega = 0$ — there is a *mass at the end, but no spring.*

Why are there no physical solutions when $\boldsymbol{\Omega^2 > 1}$? In this case the mass-spring system generates waves that propagate into the string and carry energy away. For a solution of the form in (2.18) to exist it would require the mass-spring system to have infinite energy in the far past ($t \to -\infty$). Of course, these would have generated very large amplitude waves into the string, propagating towards $x = \infty$. At the current time these waves would be at a large value of x, the larger that further into the past they originated. Hence the blow up as $x \to \infty$ of these solutions.

Proviso: The "explanation" in the prior paragraph makes sense only for solutions where a root of the equation $s^2 = \ell^2 - 1$ with negative real part (decay) is taken, $\mathrm{Re}(s) < 0$. Obviously, there are also solutions with a positive real part (if s is a solution, so is $-s$). *What about these solutions?* Well, it can be shown that they correspond to situations where there is an infinite amount of energy in the spring, moving from $x = \infty$ towards the mass-spring system, not away from it!

Remark 2.3 The arguments in the prior two paragraphs indicate how tricky it can be to deal with solutions that are not waves — i.e.: either ℓ or s in (2.18) is not pure imaginary. For example: what does it mean for such solutions to carry energy to the left, or the right? While it turns out that these concepts can be given a meaning, we will not do this here. Instead, in what follows we consider approaches to radiation damping that avoid these type of questions. ♣

Remark 2.4 From the results above, we can see that the pde in (2.16–2.17) is an example illustrating the fact that **separation of variables does not always work,** even in 1-D problems with constant coefficients. On the other hand, **the Laplace Transform approach does work** for (2.16–2.17) — see §2.2.3. Unfortunately, the solution it provides is rather complicated, and hard to interpret. ♣

2.2.2 Harmonic forcing

We look for solutions to (2.16–2.17), in the presence of an harmonic force $\boldsymbol{G = e^{-i\,\omega\,t}}$, where $\boldsymbol{\omega > 0}$ is real (the case $\omega < 0$ follows by taking complex conjugates). We concentrate on the part of the solution produced by the forcing (ignore initial conditions). Then the time dependence can be separated as $e^{-i\,\omega\,t}$. The problem then becomes an ode in $0 < x < \infty$, with an unusual boundary condition at $x = 0$.

Case 1. $\omega > 1$ ..... ***frequency dependent damping.***

In this case the forcing can generate propagating waves in the string. The solution takes the form

$$u = a\, e^{i\,(k\,x - \omega\,t)} \quad \text{for} \quad 0 < x < \infty, \quad \text{where} \quad k = \sqrt{\omega^2 - 1} \tag{2.21}$$

and a is a constant to be determined. Note that we have selected the sign of k so that the group speed $c_g = k/\omega$ is positive — the waves should move away from the mass-spring system. Substituting this into (2.17) yields $\left(\Omega^2 - \omega^2 - i\,2\,\nu\,k\right)\, a = 1$. That is

$$a = \frac{1}{\Omega^2 - \omega^2 - i\,2\,\nu\,k}. \tag{2.22}$$

We note now that, in terms of the oscillator position $y = u(0,\,t) = a\, e^{-i\,\omega\,t}$, this corresponds to the equation

$$\ddot{y} + 2\,\nu\,c_g\,\dot{y} + \Omega^2\,y = G. \tag{2.23}$$

Compare this with (2.6). The **damping coefficient here, $\nu\, c_g$, is now frequency dependent.** Note that *the damping coefficient is proportional to c_g.* This has a very clear physical interpretation: the wave energy flux away from the mass-spring is proportional to the group speed. † Since this flux is what causes the damping, the damping ends up being proportional to the group speed.

† As shown in §4 (e.g.: (4.35)) and problem sets, the average wave energy density and flux satisfy $\mathcal{F}_{av} = c_g\,\mathcal{E}_{av}$.

Case 2. $0 < \omega < 1$ ..... ***trapped modes.***

In this case the forcing cannot generate propagating waves in the string. The solution takes the form

$$u = a\, e^{-\ell\,x - i\,\omega\,t} \quad \text{for} \quad 0 < x < \infty, \quad \text{where} \quad \ell = \sqrt{1 - \omega^2} \tag{2.24}$$

and a is a constant to be determined. Hence *all the energy is trapped within some distance from the mass-spring oscillator.* Substituting this into (2.17) yields $\left(\Omega^2 - \omega^2 + 2\,\nu\,\ell\right)\, a = 1$. That is

$$a = \frac{1}{\Omega^2 + 2\,\nu\,\ell - \omega^2}. \tag{2.25}$$

Note that, in terms of the oscillator position $y = u(0,\,t) = a\, e^{-i\,\omega\,t}$, this corresponds to the equation

$$\ddot{y} + \left(\Omega^2 + 2\,\nu\,\ell\right)\, y = G. \tag{2.26}$$

There is no damping. This is not surprising: no waves cannot radiated below the bed frequency, so there cannot be any energy loss. The resulting natural frequency $\sqrt{\Omega^2 + 2\,\nu\,\ell}$ reflects these facts: (i) what oscillates here is not just the mass M, but the string as well; (ii) the restoring force is provided by the combined effect of the spring, the bed, and the string tension; (iii) how much of the string is involved depends on ω, via ℓ; (iv) even though the frequency is the same, the oscillation amplitude varies with distance from the origin. These factors enter in a complicated way, that produces a counter-intuitive result. For example, as $\ell \to 0$ and more and more of the string is involved, one would expect the response to be fully controlled by the string — yet exactly the opposite happens.

Resonance. Finally, note that a resonance occurs for $\omega^2 = \Omega^2 + 2\,\nu\,\ell$. Since $\omega^2 = 1 - \ell^2$, this can happen only if $\Omega^2 < 1$ and $0 = \ell^2 - 1 + \Omega^2 + 2\,\nu\,\ell$ — which is exactly (2.20). Thus *the resonance occurs when the forcing frequency is that of the free trapped mode.*

Question to the reader: Let $\omega^2 = \Omega^2 + 2\,\nu\,\ell$. Then (2.24–2.25) fails. **What happens in this case?**

2.2.3 Laplace Transform and radiation damping

A good way to avoid the issues pointed out in remark 2.3, is to solve (2.16–2.17) using the Laplace Transform. Let

$$\mathcal{U}(x,\,s)=\int_0^\infty e^{-s\,t}\,u(x,\,t)\,dt \quad \text{and} \quad \mathcal{G}(s)=\int_0^\infty e^{-s\,t}\,G(t)\,dt, \tag{2.27}$$

where s is complex valued, with $\mathrm{Re}(s)$ large enough. Then (2.16–2.17) becomes

$$\begin{aligned} s^2\,\mathcal{U}-\mathcal{U}_{xx}+\mathcal{U} &= u_t(x,\,0)+s\,u(x,\,0) \qquad \text{for } \; x>0, &(2.28)\\ (s^2+\Omega^2)\,\mathcal{Y}-2\,\nu\,\mathcal{U}_x &= \mathcal{G}+\dot{y}(0)+s\,y(0) \qquad \text{at } \; x=0, &(2.29)\end{aligned}$$

where $\boldsymbol{y(t)=u(0,\,t)}$ is the position of the mass M, and $\boldsymbol{\mathcal{Y}(s)=\mathcal{U}(0,\,s)}$ is its Laplace Transform.

Solving (2.28–2.29) in full is a bit cumbersome — compare this with the simplicity of the solution to the problem without dispersion in (2.9–2.10). However, the main point can be illustrated by the **solution to the simple case with $u_t(x,\,0)=0=u(x,\,0)$** — hence $\dot{y}(0)=0=y(0)$ as well. Then

$$\mathcal{U}=\mathcal{Y}(s)\,e^{-\ell\,x} \quad \text{for } \; 0<x<\infty, \quad \text{where} \quad \ell=\sqrt{1+s^2}. \tag{2.30}$$

Note that, **since s is complex, it is important that we define $\ell=\ell(s)$ carefully.** In particular: the Laplace Transform should be defined and analytic for $\mathrm{Re}(s)$ large enough, and should vanish as $|s|\to\infty$ in this realm. Thus **we define ℓ as follows:** †

1. ℓ is real and positive for s real and positive.
2. ℓ has a branch cut along the imaginary axis, from $-i$ to i.

† The Laplace Transform is uniquely defined for $\mathrm{Re}(s)$ large enough. Its continuation to the rest of the complex plane may not be unique, or even exist. For example, the branch cut for ℓ could be selected differently

It follows that ℓ has the following further properties

3. ℓ is analytic at infinity, and has a Laurent expansion there, with $\ell=s+\frac{1}{2\,s}+O\left(s^{-3}\right)$.
4. ℓ is real and negative for s real and negative.
5. ℓ is real and positive on the right side of the branch cut.
6. ℓ is real and negative on the left side of the branch cut.
7. ℓ has the form $\ell=+i\,k,\ k>0$, for $s=i\,\omega,\ \omega>1$.
8. ℓ has the form $\ell=-i\,k,\ k>0$, for $s=i\,\omega,\ \omega<-1$.
9. Introduce polar coordinates centered at i and $-i$, so that $s=i+r_1\,e^{i\,\theta_1}=-i+r_2\,e^{i\,\theta_2}$ — where $-\pi/2\le\theta_1,\,\theta_2\le 3\,\pi/2$. Then ..... $\boldsymbol{\ell=\sqrt{r_1\,r_2}\,e^{i\,\frac{1}{2}\,(\theta_1+\theta_2)}}$.

   The θ_j are defined so that, as the imaginary axis is crossed from right to left: (i) $\theta_1+\theta_2$ does not jump for crossings above i, (ii) $\theta_1+\theta_2$ jumps by $2\,\pi$ for crossings between $-i$ and i, and (iii) $\theta_1+\theta_2$ jumps by $4\,\pi$ for crossings below $-i$. This produces the branch cut/discontinuity as selected in item **2**.

Substituting (2.30) into (2.29) yields $$(s^2+\Omega^2+2\,\nu\,\ell)\,\mathcal{Y}=\mathcal{G}, \tag{2.31}$$
which does not correspond to an ode (e.g.: as in (2.6)).
This generalizes the situation observed in § 2.2.2, with a frequency dependent damping (which can become imaginary, to produce a frequency correction, as in (2.26)).

The question is now: **what does (2.31) correspond to in terms of the variable y?** To answer this question, **let $\mu = \mu(t)$ be the inverse Laplace Transform of $\ell - s$** (see remarks 2.5 and 2.6). That is:

$$\mu(t) = \frac{1}{2\,\pi\,i}\,\int_\Gamma (\ell(s) - s)\,e^{s\,t}\,ds, \tag{2.32}$$

where Γ is a path of the form $s = \alpha + i\,k$, with k real going from $-\infty$ to ∞, and some/any $\alpha > 0$. Then, using the fact that the Laplace Transform of a convolution of two functions is the product of two functions, we can write the following equation

$$\ddot{\boldsymbol{y}} + 2\,\boldsymbol{\nu}\,\dot{\boldsymbol{y}} + \boldsymbol{\Omega^2}\,\boldsymbol{y} + 2\,\boldsymbol{\nu}\,\int_0^t \boldsymbol{\mu(t-\tau)}\,\boldsymbol{y(\tau)}\,\boldsymbol{d\tau} = \boldsymbol{G}, \tag{2.33}$$

which is **equivalent to (2.31) provided that $y(0) = \dot{y}(0) = 0$.** Unlike (2.6), or (2.10), this is not an ode, and has an extra **memory term** in it. The *physical meaning of this is as follows:*

> As the mass in the mass-spring system vibrates, it radiates waves that propagate into the string. These waves eventually escape to infinity, and produce the damping term — the same as in (2.6) or (2.10). However, as the waves moves through the string, they interact with the bed, and produce waves that are reflected back to the mass-spring system. Hence the state of the mass at any given time does not just directly affect the immediate future, but any time in its future — via a wave going in some distance and reflecting back something that arrives later. This is the mechanism that the "memory" effect given by the integral term in the equation describes. ♣

We should expect this behavior to be generic for any radiation damping situation where reflections back to the "radiator" are produced by the media in which the waves propagate.

Remark 2.5 ***Why define μ as the inverse Laplace Transform of $\ell - s$, instead of just ℓ?*** This is because the Laplace Transform of a function must decay for $|s| \to \infty$ and $\mathrm{Re}(s)$ large enough, but ℓ does not. Hence we subtract the "bad" behavior of ℓ at infinity, using item **3** above. ♣

Remark 2.6 The integral defining μ in (2.32) can be manipulated to obtain a better description of the function — in particular, an integral representation that is not as poorly convergent as (2.32). For this purpose, move the integration contour Γ to the left. As α crosses zero, a contribution from the branch cut is picked, while the reminder vanishes as $\alpha \to -\infty$ (by Jordan's lemma). Thus

$$\mu(t) = \frac{1}{2\,\pi\,i}\,\oint \ell(s)\,e^{s\,t}\,ds = \frac{1}{\pi}\,\int_{-1}^{1} \sqrt{1-\zeta^2}\,e^{i\,\zeta\,t}\,d\zeta = \frac{2}{\pi}\,\int_0^1 \sqrt{1-\zeta^2}\,\cos(\zeta\,t)\,d\zeta. \tag{2.34}$$

Hence † $$\boldsymbol{\mu(t) = t^{-1}\,J_1(t)}, \tag{2.35}$$

where J_1 is a Bessel function of the first kind. ♣

† Item 9.1.20, Abramowitz M. and Stegun I. A., *Handbook of Math. Functions,* 9th ed., Dover, NYC, 1970.

---

3 Stationary phase and the far field approximation

In this section we consider the behavior of the solutions to linear, homogeneous, dispersive systems, in the "far field limit" — that is, far away (in both time and space) from the sources that generate the waves.

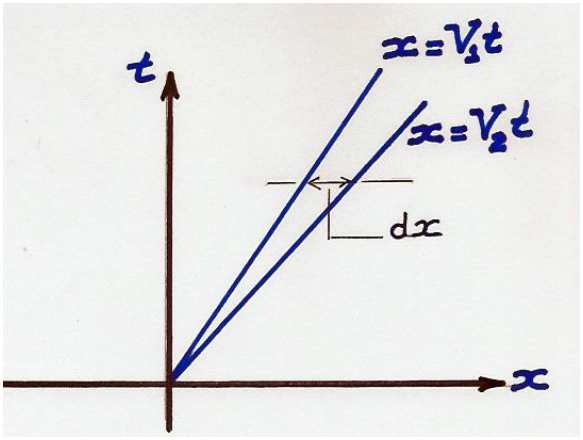
3.1 Large time for a 1-D scalar equation

In this subsection we study the large time behavior of the solutions to a 1-D scalar dispersive system, from the point of view of an observer moving at some constant speed V. That is:

$$u = u(x,\, t) = \int_{-\infty}^{\infty} U(k)\, e^{i\,(k\,x - \omega\, t)}\, \mathrm{d}k, \quad \omega = \Omega(k), \tag{3.1}$$

in the **limit $\boldsymbol{x = Vt}$**, with $\boldsymbol{t \gg 1}$ — see figure 3.1, where $\omega = \Omega$ is the dispersion relation.

Note: We assume here a single-branch dispersion relation. The general problem has solutions that consist of a sum of integrals like this, one per branch of the dispersion relation.



Large time for a 1-D scalar equation. Each velocity corresponds to a wave-number, $V_j = c_g(k_j)$. If $k_2 = k_1 + dk$, then $dx = (V_2 - V_1)\, t = |\Omega''(k_1)|\, t\, dk$.

Figure 3.1: Large time for a 1-D scalar equation

What we expect to see, because wave properties propagate at the group speed, [2] is

1. The answer should look as if everything arises from a point source. From sufficiently far away, even an extended source looks like a point.
2. Along the path $x = V\, t$, the wave determined by $c_g(k) = V$ is what the observer should see.
3. Consider two nearby paths, $x = V_j\, t$, with $V_j = c_g(k_j)$ and $k_2 = k_1 + dk$. It is easy to see then that $\boldsymbol{dx = |V_2 - V_1|\, t = |\Omega''(k)|\, t\, dk}$ — see figure 3.1. Now: the wave energy density is proportional to the square of the wave amplitude, $e = \gamma\, |A|^2$, for some $\gamma = \gamma(k)$. Then, since the energy is carried by the group speed, it should be that $\boldsymbol{\gamma\, |A|^2\, dx = E(k)\, dk}$, where E is the energy density per wave-number in the initial conditions. Putting these two things together yields

$$\boldsymbol{\gamma\, |A|^2\, |\Omega''(k)|\, t = E(k).} \tag{3.2}$$

Next we confirm these expectations by taking the limit stated in (3.1). The equation for u has the form

$$u = \int_{-\infty}^{\infty} U(k)\, e^{i\,\phi(k)\, t}\, \mathrm{d}k, \quad \text{where} \quad \phi = k\, V - \omega \quad \text{and} \quad t \gg 1. \tag{3.3}$$

[2] Truthfully, we have not yet shown this completely. The calculation here will be, actually, the proof of this.

This is ready made for the method of the stationary phase [3] — see § 3.2. Stationary points occur at the values of $k = k_s$ where

$$\boldsymbol{\phi'(k_s) = 0 \iff c_g(k_s) = V = \frac{x}{t}.} \tag{3.4}$$

This equation yields k_s as a function of $\frac{x}{t}$ $$k_s = k_s\left(\frac{x}{t}\right). \tag{3.5}$$

Assume now simple stationary points: $$\boldsymbol{\Omega''(k_s) = -\phi''(k_s) \neq 0.} \tag{3.6}$$

Then we can write — see § 3.2 — for the far field limit of (3.1)

$$\boldsymbol{u \sim \sum_s \sqrt{\frac{2\,\pi}{t\,|\Omega''(k_s)|}}\, U(k_s)\, \exp\left(i\,\left(k_s\,x - \omega_s\,t + \nu_s\,\frac{\pi}{4}\right)\right),} \tag{3.7}$$

where k_s is given by (3.4), $$\boldsymbol{\omega_s = \Omega(k_s), \quad \nu_s = -\mathbf{sign}\left(\Omega''(k_s)\right),} \tag{3.8}$$
and *the sum is over all the stationary points.*

What happens if there are no stationary points? As shown in § 3.2, the absence of stationary points corresponds to u being very small — much smaller than any power t^{-n} if everything is smooth. [4] That is:

When moving at a velocity V that does not correspond to the group velocity of any wave-number, and observer sees no waves. (3.9)

What happens if there are stationary points that do not satisfy (3.6)? For this case see § 3.3.

Example 3.1 *Consider the equation* $u_{tt} - u_{xx} + u = 0.$
Then $\Omega = \pm\sqrt{1+k^2}$, $c_g = k/\Omega$, *and* $\Omega'' = (1 - c_g^2)/\Omega.$
Thus:
For $-1 < V < 1$, $c_g = V$ yields: $k_s = \dfrac{\pm V}{\sqrt{1-V^2}}$, $\omega_s = \dfrac{\pm 1}{\sqrt{1-V^2}}$, *and* $\Omega''(k_s) = \pm\,(1-V^2)^{3/2} \neq 0.$

For $V^2 \geq 1$ there are no stationary points. The solution is very small, with no waves. In fact, the equation is hyperbolic, nothing can propagate faster than speed 1. If the initial data is compact support (zero outside an interval), the solution vanishes for $V^2 > 1$ and $t \gg 1$.

As $V^2 \uparrow 1$, k_s and ω_s blow up, while $\Omega''(k_s) = O(k_s^{-3})$ vanishes. For $0 < 1 - V^2 \ll 1$, the solution is dominated by high frequencies. The amplitude need not be large, as (3.7) suggests for a small $\Omega''(k_s)$, since generally U vanishes as $|k_s| \to \infty$ (quite rapidly if the IV are smooth enough). ♣

3.1.1 Amplitude-phase notation

Each of the terms in the sum in (3.7) can be written in the form

$$\boldsymbol{u = A\,e^{i\,\theta}} \quad \text{where} \quad \boldsymbol{A = \sqrt{\frac{2\,\pi}{t\,|\Omega''(k_s)|}}\, U(k_s)\, e^{i\,\frac{\pi}{4}\,\nu_s}} \quad \text{and} \quad \boldsymbol{\theta = k_s\,x - \omega_s\,t,} \tag{3.10}$$

with k_s, ω_s, and ν_s, defined by (3.4–3.5) and (3.8). In particular, note that

$$|A^2|\,|\Omega''(k_s)|\,t = 2\,\pi\,|U^2(k_s)|, \tag{3.11}$$

which is (3.2). Thus (3.7) satisfies the expectations listed in items **1-3** above (3.2).

[3] The method of the stationary phase was developed precisely to address "far-field" type questions in wave propagation — by G. G. Stokes and Lord Kelvin, in the late 19-th century.

[4] It can be shown that: if U and Ω are analytic, the decay is (at least) exponential.

3.1.2 Preview to modulation theory

A and θ in (3.10) satisfy the equations (with $\boldsymbol{\omega = \Omega(k)}$)

$$\boldsymbol{\theta_x = k_s} \quad \text{and} \quad \boldsymbol{\theta_t = -\omega_s} \quad \Longrightarrow \quad \boldsymbol{(k_s)_t + (\omega_s)_x = 0}, \tag{3.12}$$

(conservation of waves) and

$$\boldsymbol{\left(A^2\right)_t + \left(c_g(k_s)\, A^2\right)_x = 0}, \tag{3.13}$$

$$\text{or, equivalently} \qquad \boldsymbol{A_t + \frac{1}{2}\,\left(c_g(k_s)\right)_x\, A + c_g(k_s)\, A_x = 0}, \tag{3.14}$$

where the first equation states that *waves/bumps are conserved*, while the second shows that the *wave energy propagates at the group speed.* Furthermore, note that (see remark 3.1)

$$\textbf{\textit{A}, } \boldsymbol{k_s} \textbf{ and } \boldsymbol{\omega_s} \textbf{ are slow functions of } \boldsymbol{x} \textbf{ and } \boldsymbol{t}\textbf{, but } \boldsymbol{\theta} \textbf{ is not.} \tag{3.15}$$

This last is a consequence of the assumption that $t \gg 1$.[5] In § 4.1 we consider the general situation where (3.15) applies, of which (3.10) is a special case.

Proof of (3.12–3.13).

(1) $\theta_x = k_s + (x - c_g(k_s)\,t)\,(k_s)_x = k_s$; (2) $\theta_t = -\omega_s + (x - c_g(k_s)\,t)\,(k_s)_t = -\omega_s$; and

(3) A^2 has the form $A^2 = \dfrac{1}{t}\, f(k_s)$, for some function f. Thus

$$(A^2)_t = -\frac{1}{t^2}\, f + \frac{1}{t}\, f'(k_s)\,(k_s)_t \qquad \text{and} \qquad \left(c_g(k_s)\, A^2\right)_x = \left(\frac{x}{t^2}\, f(k_s)\right)_x = \frac{1}{t^2}\, f(k_s) + \frac{1}{t}\,\frac{x}{t}\, f'(k_s)\,(k_s)_x.$$

But $(k_s)_t = -c_g(k_s)\,(k_s)_x = -\dfrac{x}{t}\,(k_s)_x$, so that (3.13) results.

Example 3.2 Special case solution for (3.12–3.14). Take $k =$ constant and $\omega = \Omega(k)$, which trivially satisfies (3.12). Then (3.14) reduces to $\boldsymbol{A_t + c_g(k)\, A_x = 0}$ — the amplitude equation for a carrier wave with fixed frequency and slowly varying amplitude. ♣

Remark 3.1 The general form behind (3.12–3.14). Because $t \gg 1$, k_s and ω_s (which are functions of x/t), change slowly over bounded regions in space-time. In fact, let $\epsilon =$ "typical size" of $1/t$. Then, with $\chi = \epsilon\, x$ and $\tau = \epsilon\, t$, we can write:† $k = k(\chi,\, \tau)$, $\omega = \omega(\chi,\, \tau)$, $A = \sqrt{\epsilon}\tilde{A}(\chi,\, \tau)$, and $\theta = \frac{1}{\epsilon}\,\Theta(\chi,\, \tau)$, where $\Theta = k_s\,\chi - \omega_s\,\tau$. Elsewhere we show that solutions of the form $u = A(\chi,\, \tau)\, e^{i\,\theta}$, with $\theta = (1/\epsilon)\,\Theta(\chi,\, \tau)$, satisfy the equations in (3.12–3.14). ♣

† In fact, k and ω are functions of χ/τ.

3.2 Stationary Phase

In this subsection we *consider the behavior of integrals of the form*

$$I(\mu) = \int_{-\infty}^{\infty} f(k)\, e^{i\,\mu\,\phi(k)}\, \mathrm{d}k, \quad \text{where } \phi \text{ is real valued,} \tag{3.16}$$

in the limit $\mu \gg 1$ — the main result is in equation (3.25). We assume that the functions $f = f(k)$ and $\phi = \phi(k)$ are smooth enough, with f and f' integrable.

Here we present a simple version of the theory. For a concise, clear, and complete, presentation of this topic see:

[5] Let $0 < \epsilon \ll 1$ be a typical size for $1/t$. Then, in terms of $X = \epsilon\, x$ and $T = \epsilon\, t$, we have: (i) $x/t = X/T$, so that $k_s = k_s(X,\, T)$, (ii) $\theta = \Theta/\epsilon$, where $\Theta = \Theta(X,\, T) = k_s\, X - \omega_s\, T$, and (iii) $A = \sqrt{\epsilon}\,\alpha(X,\, T)$.

E. T. Copson, *Asymptotic Expansions,* Cambridge University Press, 1965. The book also includes related topics, important for wave behavior, such as: Watson's lemma, steepest descents, and turning points.

Motivation/Intuition.

In the situation under consideration, the exponential in (3.16) is highly oscillatory, and produces a large amount of cancellation in the integral. Specifically, the contributions from near k (any k) and that from $k_* = k + \dfrac{\pi}{\mu\,\phi'(k)}$ are approximately opposite in sign:

$$f(k_*)\,e^{i\,\mu\,\phi(k_*)} = -f(k)\,e^{i\,\mu\,\phi(k)} + O\left(\frac{1}{\mu\,|\phi'(k)|}\right), \tag{3.17}$$

and so cancel (approximately). In fact:

Consider the situation where ϕ' is bounded away from zero everywhere. (3.18)

Thus, WLOG,[6] **assume that $\phi' > q =$ constant > 0.**[7] Then change variables, and rewrite the integral above in the form

$$I(\mu) = \int_{-\infty}^{\infty} F(\phi)\,e^{i\,\mu\,\phi}\,\mathrm{d}\phi, \tag{3.19}$$

where $F(\phi) = f(K)/\phi'(K)$, and $K = K(\phi)$ is the function inverse to $\phi = \phi(k)$. Then it follows that:

$$I = \int_{-\infty}^{\infty} F(\phi)\,\mathrm{d}\frac{e^{i\,\mu\,\phi}}{i\,\mu} = -\frac{1}{i\,\mu}\int_{-\infty}^{\infty} F'(\phi)\,e^{i\,\mu\,\phi}\,\mathrm{d}\phi = \cdots = \left(-\frac{1}{i\,\mu}\right)^n \int_{-\infty}^{\infty} F^{(n)}(\phi)\,e^{i\,\mu\,\phi}\,\mathrm{d}\phi, \tag{3.20}$$

which applies as long as F is n times differentiable with integrable derivatives. Thus

For the case in (3.18), $I = O(\mu^{-n})$, where n is determined by the degree of smoothness and integrability of F. For smooth functions, I is smaller than any power: $I = O(\mu^{-n})$. In fact, for analytic functions it can be shown that I is exponentially small in μ. (3.21)

The arguments above indicate that:

The dominant contributions to the integral in (3.16) arise from the neighborhoods of the stationary points of the phase ϕ — i.e.: the points where $\phi' = 0$. (3.22)

Single and simple stationary point.

Thus let us **assume that the phase has only one stationary point;** that is: a value $k = k_s$ at which $\boldsymbol{\phi'(k_s) = 0}$. In addition, **assume that $\boldsymbol{\phi''_s = \phi''(k_s) \neq 0}$.** Following (3.22) we pick a small neighborhood of k_s, $|k - k_s| \leq \delta \ll 1$, and write

$$I(\mu) \sim \int_{k_s-\delta}^{k_s+\delta} f(k)\,e^{i\,\mu\,\phi(k)}\,\mathrm{d}k \sim f(k_s)\,e^{i\,\mu\,\phi(k_s)} \int_{k_s-\delta}^{k_s+\delta} e^{i\,\frac{1}{2}\,\mu\,\phi''_s\,(k-k_s)^2}\,\mathrm{d}k, \tag{3.23}$$

where, because the interval is small, we have approximated f and ϕ as follows:

$f(k) = f(k_s) + O(|k - k_s|) = f(k_s) + O(\delta)$, and

[6] WLOG means: Without Loss Of Generality.

[7] Note that this implies $\phi \to \pm\infty$ as $k \to \pm\infty$.

$\mu\,\phi(k) = \mu\,\phi(k_s) + \dfrac{1}{2}\,\mu\,\phi_s''\,(k-k_s)^2 + O(\mu\,\delta^3)$ — see remark 3.2.

Furthermore, let ..... $\boldsymbol{\nu = \mathbf{sign}(\phi_s'')}$ and $\boldsymbol{\alpha = \sqrt{\dfrac{2}{\mu\,|\phi_s''|}}}$, (3.23b)

and change variables $k = k_s + \alpha\,s$. Then

$$I(\mu) \sim f(k_s)\,\alpha\,e^{i\,\mu\,\phi(k_s)} \int_{-\delta/\alpha}^{\delta/\alpha} e^{i\,\nu\,s^2}\,\mathrm{d}s. \tag{3.24}$$

Select $r = \delta/\alpha \gg 1$, see remark 3.2, so we can use remark 3.3 and conclude that: **For a unique stationary point at which ϕ'' does not vanish**

$$\boldsymbol{I(\mu) \sim f(k_s)\,\sqrt{\dfrac{2\,\pi}{\mu\,|\phi_s''|}}\,\exp\left(i\,\mu\,\phi(k_s) + i\,\nu\,\dfrac{\pi}{4}\right).} \tag{3.25}$$

If there are several stationary points at which ϕ'' does not vanish, their contributions should be added.[8]

Remark 3.2 What size should δ have in (3.23)? Clearly, δ cannot be too small, otherwise the selected interval will not be large enough to capture the full contribution from the stationary point. On the other hand, it cannot be too large either — otherwise the approximations made for f and $\mu\,\phi$ are not valid. These approximations require $\delta \ll 1$ and $\mu\,\delta^3 \ll 1$, resp. Furthermore, to use remark 3.3, $\delta \gg \alpha$ is needed (where, from (3.23b) $\boldsymbol{\alpha \sim 1/\sqrt{\mu}}$). All these conditions are satisfied if

$$\boldsymbol{\mu^{-1/2} \ll \delta \ll \mu^{-1/3}.}$$

As a final check, we need to show that neglecting the integration beyond $k \pm \delta$ in (3.23) is justified if δ is selected as above. This can be done, but the calculations are too cumbersome to include here (see the book by Copson mentioned earlier for a complete justification of the method). ♣

Remark 3.3 Here we show that, ... for $\boldsymbol{r \gg 1}$, $\boldsymbol{J = \displaystyle\int_{-r}^{r} e^{i\,s^2}\,\mathrm{d}s \sim \sqrt{\pi}\,e^{i\,\frac{\pi}{4}}}$.

Consider the closed path Γ in the complex z-plane going: (i) From $z=-r$ to $z=r$ along the real axis; (ii) From $z=r$ to $z = r\,e^{i\,\pi/4}$, counterclockwise, along the circle of radius r; (iii) From $z = r\,e^{i\,\pi/4}$ to $z = -r\,e^{i\,\pi/4}$ along the $\mathrm{Re}(z) = \mathrm{Im}(z)$ line; and (iv) From $z = -r\,e^{i\,\pi/4}$ to $z=-r$, clockwise, along the circle of radius r. Since $e^{i\,z^2}$ is analytic, its integral along this path vanishes. Furthermore, the integral of $e^{i\,z^2}$ over the circular portions of the path are $O(r^{-1})$ — as shown below. It follows that we can rotate the path of integration for J by $\pi/4$, incurring an $O(r^{-1})$ error. That is

$$J + O(r^{-1}) = \int_{\Lambda} e^{i\,z^2}\,\mathrm{d}z = e^{i\,\frac{\pi}{4}} \int_{-r}^{r} e^{-s^2}\,\mathrm{d}s = e^{i\,\frac{\pi}{4}} \int_{-\infty}^{\infty} e^{-s^2}\,\mathrm{d}s + O(r^{-1}\,e^{-r^2}) = e^{i\,\frac{\pi}{4}}\sqrt{\pi} + O(r^{-1}\,e^{-r^2}),$$

where Λ is the contour given by $z = e^{i\,\pi/4}\,s$, $-r < s < r$, and we have used that $\displaystyle\int_r^{\infty} e^{-s^2}\mathrm{d}s = O(r^{-1}\,e^{-r^2})$ — which follows by integration by parts. This proves the desired result.

We conclude the argument by showing that the integral of $e^{i\,z^2}$ over the circular portions of the path Γ are $O(r^{-1})$. Thus, consider the path Γ_c given by $z = r\,e^{i\,\psi}$, $0 < \psi < \pi/4$ (the argument is the same for the other circular portion), and let $\boldsymbol{M = \int_{\Gamma_c} e^{i\,z^2}\mathrm{d}z = i\,r\,\int_0^{\pi/4} e^{i\,z^2}\,\mathrm{d}\psi}$. Then

$$|M| \le r\int_0^{\pi/4} e^{-r^2\,\sin(2\,\psi)}\,\mathrm{d}\psi \le r\int_0^{\pi/4} e^{-\frac{4}{\pi}\,r^2\,\psi}\,\mathrm{d}\psi \le r\int_0^{\infty} e^{-\frac{4}{\pi}\,r^2\,\psi}\,\mathrm{d}\psi = \frac{\pi}{4\,r},$$

which proves the point. ♣

[8] In this case the integral for I has to be split into several pieces of the form in (3.23).

Remark 3.4 For integrals over finite intervals

$$I(\mu) = \int_a^b f(k)\, e^{i\,\mu\,\phi(k)}\, \mathrm{d}k, \tag{3.26}$$

where either a or b (or both) are finite, contributions from the boundaries (as well as those from stationary points, if any) have to be accounted for. We will not consider this situation here. ♣

3.3 Turning points: transitions from waves to no-wave

The purpose here is to investigate *what happens when* $\Omega''(k_s)$ *vanishes in* (3.7); that is: (3.6) *fails*. Then (3.7) yields an infinite amplitude for the wave, which is (at the very least) suspicious — in fact, wrong, as we show below in (3.28). So, let us look at this carefully:

1. The first clue comes from the answer to the question: *how does the pre-factor* $1/\sqrt{t\,|\Omega''(k_s)|}$ *in the amplitude originate in* (3.7)? To answer this note that: in the stationary phase method we (i) Expand $k\,x - \omega\,t = k_s\,x - \omega_s\,t + (x - c_g(k_s)\,t)(k - k_s) - \frac{1}{2}\,\Omega_s''\,(k-k_s)^2\,t + O((k-k_s)^3\,t)$, where k_s is defined by $x - c_g(k_s)\,t = 0$. (ii) Neglect the higher order term, which reduces the evaluation of the Fourier Transform solution to performing the integral $I = \int e^{-i\frac{1}{2}\,\Omega_s''(k-k_s)^2\,t}\,\mathrm{d}k = \int e^{-i\frac{1}{2}\,\Omega_s''\kappa^2\,t}\,\mathrm{d}\kappa$. (iii) Scale κ to reduce I to the pre-factor above times a constant $= \int e^{\pm i\,\frac{1}{2}\,\mu^2}\mathrm{d}\mu$.

2. When $\Omega_s'' = 0$, we need to consider the next order in the expansion of the phase. Then the same process above reduces the Fourier Transform solution to the integral $I = \int e^{-i\frac{1}{2}\,\Omega_s'''\kappa^3\,t}\,\mathrm{d}\kappa$. Scaling now produces the pre-factor $\boldsymbol{(t\,\Omega'''(k_s))^{-1/3}}$. (3.27)
Thus, when moving at a speed for which $c_g(k_s) = x/t$ leads to $\Omega_s'' = 0$, an *observer sees the wave amplitude decaying like* $t^{-1/3}$, *instead of* $t^{-1/2}$. *The wave amplitude is much larger (by a factor of* $t^{1/6}$*), but not infinity.* (3.28)

3. The condition $c_g(k_s) = x/t$ with $\Omega_s'' \neq 0$ corresponds to: k_s is a local maximum/minimum of the phase as a function of k, with $V = x/t$ fixed. † *What does* $\Omega_s'' = 0$ *correspond to?* We claim that, generically, *this corresponds to two stationary points (a local maximum and a local minimum) that merge and disappear as* V *varies* (see figure 3.2) (3.29)

† This follows because $\frac{\mathrm{d}\theta}{\mathrm{d}k} = x - c_g\,t$ and $\frac{\mathrm{d}^2\theta}{\mathrm{d}k^2} = -\Omega''(k)\,t$.

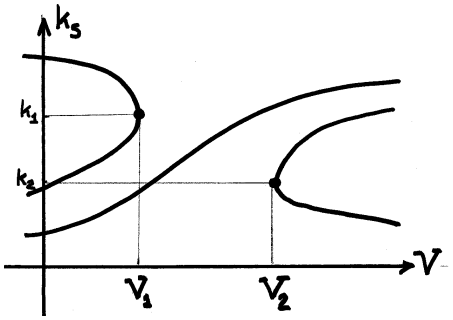


Figure 3.2: Stationary wave-numbers as V changes
Illustration of how the solutions to $c_g(k_s) = V$, generally, change as functions of V. (i) There are several solution branches. (ii) The number of branches is not constant. (iii) At points where $\Omega''(k_s) \neq 0$, the branches are smooth functions of V (inverse function theorem). (iv) Points for which $\Omega''(k_s) = 0$, but $\Omega'''(k_s) \neq 0$, correspond to branch mergers: two branches on one side and none on the other. In the picture this happens at $(V_1,\, k_1)$ and $(V_2,\, k_2)$.

Let us prove this:

3.1 From the inverse function theorem, a solution branch to $c_g(k_s) = V$ ($k_s = k_s(V)$) is smooth when $\Omega''(k_s) \neq 0$.

3.2 Assume a k_0 such that $\Omega''(k_0) = 0$, but $q = \frac{1}{2}\Omega'''(k_0) \neq 0$, and let us investigate what happens for V near $V_0 = c_g(k_0)$. Substituting $k_s = k_0 + \delta k_s$ and $V = V_0 + \delta V$ into the equation $c_g(k_s) = V$ yields, at leading order, $q\,\delta k_s^2 = \delta V$. Thus there are two solutions for $q\,\delta V > 0$, and none if $q\,\delta V < 0$. Near the critical point we can write $V = V(k_s)$, where V is smooth and has a local maximum (or minimum) at $k_s = k_0$. **QED**.

3.3 Using item **3.2** to compute $\Omega''(k_s)$ for V slightly below V_1 in figure 3.2 (where $q<0$), it is easy to see that the upper branch is a local minimum and the middle branch is a local maximum. It follows that the lower branch is a local minimum.

Given the above, an important question is:

The boundary between waves and shadow; two stationary points coalesce: How does the transition from two waves with decay rate $t^{-1/2}$, to a single one with decay rate $t^{-1/3}$, to no waves, happen? (3.30)

We answer this below.

Let $u=u(x,\,t)$ be the (single branch †) solution to a 1-D scalar dispersive equation

$$u=\int_{-\infty}^{\infty} U(k)\,e^{i\,(k\,x-\omega\,t)}\,\mathrm{d}k,\quad \omega=\Omega(k). \tag{3.31}$$

† The general solution consist of a sum of integrals like this, one per branch of the dispersion relation.

Consider the far field limit of u, in the region where two stationary points coalesce. That is:
the **limit $t\gg 1$, with $x/t\approx V_0=c_g(k_0)$, where $\Omega''(k_0)=0$ and $\mu=-\frac{1}{2}\,\Omega''(k_0)\neq 0$.** (3.32)

As in (3.3–3.9), we argue that the main contribution to the integral arises from the neighborhood of the stationary points. However, because the stationary points coalesce for x/t as above, we cannot expand near each of them separately — as in the process leading to (3.7). Instead *we expand the phase $\boldsymbol{\theta=k\,x-\omega\,t}$ in an approximation valid for a (small region) that includes the two stationary points.* Specifically, let: $\boldsymbol{k=k_0+\kappa}$, where $\boldsymbol{\kappa}$ **is small**, and expand:

$$\begin{aligned}\omega &= \omega_0+V_0\,\kappa-\frac{1}{3}\,\mu\,\kappa^3+O(\kappa^4), && \text{so that}\\ \theta &= k_0\,x-\omega_0\,t+(x-V_0\,t)\,\kappa+\frac{1}{3}\,\mu\,\kappa^3 t+O(\kappa^4\,t).\end{aligned} \tag{3.33}$$

For definiteness **we assume $\mu>0$, so that two stationary points exist (near k_0) for $V<V_0$, and none exists for $V>V_0$.** The case $\mu<0$ is similar.

Remark 3.5 Since $c_g=V_0-\mu\,\kappa^2+O(\kappa^3)$, the stationary points (defined by $c_g=V=x/t$) occur for $V<V_0$, and satisfy $k_s\approx k_0\pm\sqrt{(V_0-V)/\mu}$. On the other hand, for $\mu<0$ they occur for $V>V_0$ ♣

Substituting (3.33) into (3.31) leads to

$$u\sim U(k_0)\,e^{i\,(k_0\,x-\omega_0\,t)}\int_{-\infty}^{\infty} e^{i\,\left(\kappa\,(x-V_0\,t)+\frac{1}{3}\,\mu\,\kappa^3\,t\right)}\,\mathrm{d}\kappa. \tag{3.34}$$

We now change variables to $s=(\mu\,t)^{1/3}\,\kappa$, so that

$$\boldsymbol{u\sim\frac{1}{(\mu\,t)^{1/3}}U(k_0)\,e^{i\,(k_0\,x-\omega_0\,t)}\int_{-\infty}^{\infty} e^{i\,\left(s\,\zeta+\frac{1}{3}\,s^3\right)}\,\mathrm{d}s=\frac{2\,\pi}{(\mu\,t)^{1/3}}U(k_0)\,e^{i\,(k_0\,x-\omega_0\,t)}\,\mathrm{Ai}(\zeta),} \tag{3.35}$$

where $\boldsymbol{\zeta=(x-V_0\,t)/(\mu\,t)^{1/3}}$ and Ai is the *Airy* function — a Bessel function of order $\frac{1}{3}$; see §3.3.2. It can be shown that **this approximation is valid for $|\frac{x}{t}-V_0|\ll 1$, and $t\gg 1$. Note that (3.35) only accounts for the contribution from the two stationary points near k_0. If there are other stationary points (away from k_0), their contributions must be added to the approximation to u.**

Using the asymptotic behavior of the Airy function for $\zeta \gg 1$, we see that (3.35) yields

$$u \sim \frac{\sqrt{\pi}}{(\mu\, t)^{1/3}\, \zeta^{1/4}} U(k_0)\, e^{i\,(k_0\, x - \omega_0\, t) - \frac{2}{3}\,\zeta^{3/2}}, \tag{3.36}$$

valid for $t^{1/3} \ll x - V_0\, t \ll t$ (shadow zone). † Thus, when the two stationary points coalesce and disappear, the contribution from the region becomes exponentially small.

† Note that $(\mu\, t)^{1/3}\, \zeta^{1/4} = (x - V_0\, t)^{1/4} (\mu\, t)^{1/4}$.

In §3.3.1 we show how (3.35) transitions into the wave region that exists for $x/t < V_0$.

3.3.1 Matching of (3.35) with (3.7)

Let k_0 be as in (3.32), with $\mu > 0$. Consider the situation in (3.3–3.8). Assume $t^{1/3} \ll V_0\, t - x \ll t$ (this is the same as: $t^{-2/3} \ll V_0 - V \ll 1$ or $1 \ll -\zeta \ll t^{2/3}$).

Then (see remark 3.5)
$$k_s \approx k_0 \pm \sqrt{\frac{V_0 - V}{\mu}} = k_0 \pm \frac{\sqrt{-\zeta}}{(\mu\, t)^{1/3}}. \tag{3.37}$$

Using this with (3.33) then yields:
$$\Omega''(k_s) \approx -2\,\mu\,\kappa_s = \mp \frac{2\,\mu\,\sqrt{-\zeta}}{(\mu\, t)^{1/3}}, \tag{3.38}$$

and †
$$\theta + \nu_s\, \frac{\pi}{4} \approx \theta_0 \mp \frac{2}{3} |\zeta|^{3/2} \pm \frac{\pi}{4}, \tag{3.39}$$

where $\theta_0 = k_0\, x - \omega_0\, t$.

† Note that (3.33) yields $\theta \approx \theta_0 + (\mu\, t)^{1/3} \zeta\, \kappa_s + \frac{1}{3}\,\mu\, t\, \kappa_s^3$, with $\kappa_s = \pm\sqrt{|\zeta|}/(\mu\, t)^{1/3}$.

Using the above in (3.7) reduces the expression there to

$$\begin{aligned} u &\sim \sum_{\pm} \frac{\sqrt{\pi}}{(\mu\, t)^{1/3} |\zeta|^{1/4}}\, U(k_0)\, e^{i\,\theta_0}\, e^{\mp i\,\left(\frac{2}{3}\,|\zeta|^{3/2} - \frac{\pi}{4}\right)} \\ &= \frac{2\,\sqrt{\pi}}{(\mu\, t)^{1/3} |\zeta|^{1/4}}\, U(k_0)\, e^{i\,\theta_0}\, \cos\left(\frac{2}{3}\,|\zeta|^{3/2} - \frac{\pi}{4}\right). \end{aligned} \tag{3.40}$$

Since $\cos(y - \pi/4) = \sin(y + \pi/4)$, **this is the same expression that (3.35) yields when (3.44) is used.**

3.3.2 The Airy function

The Airy function can be defined by the integral

$$Ai(\zeta) = \frac{1}{2\,\pi} \int_{-\infty}^{\infty} e^{i(z\,\zeta + \frac{1}{3}\, z^3)}\, \mathrm{d}z. \tag{3.41}$$

However, this integral converges conditionally, and only for ζ real. However, for $0 < \arg(z) < \pi/3$ and $2\,\pi/3 < \arg(z) < \pi$, $e^{i\, z^3/3}$ decays exponentially fast as $|z| \to \infty$. Thus we can change the contour of integration to Γ in figure 3.3, defined by: $\arg(z) = 5\,\pi/6$ for $\mathrm{Re}(z) < 0$, and $\arg(z) = \pi/6$ for $\mathrm{Re}(z) > 0$.

$$Ai(\zeta) = \frac{1}{2\,\pi} \int_{\Gamma} e^{i(z\,\zeta + \frac{1}{3}\, z^3)}\, \mathrm{d}z. \tag{3.42}$$

Note that, on Γ, $e^{i\, z^3/3} = e^{-|z|^3/3}$. Thus (3.42) defines Ai as an entire function of ζ. For ζ real, Ai switches from oscillatory on the left, to exponentially decaying on the right — see figure 3.3. In fact, it can be

shown that

$$Ai(\zeta) \sim \frac{1}{2\sqrt{\pi}\,\zeta^{1/4}}\, e^{-\frac{2}{3}\zeta^{3/2}} \qquad \text{for } 1 \ll +\zeta. \tag{3.43}$$

$$Ai(\zeta) \sim \frac{1}{\sqrt{\pi}\,|\zeta|^{1/4}}\, \sin\left(\frac{2}{3}\,|\zeta|^{3/2} + \frac{\pi}{4}\right) \qquad \text{for } 1 \ll -\zeta. \tag{3.44}$$

These formulas can be obtained by either (i) The steepest descent method on (3.42); or (ii) Matching of asymptotic expansions for the solutions of (3.45) in the complex plane — matching in the argument of ζ, for $|\zeta|$ large.

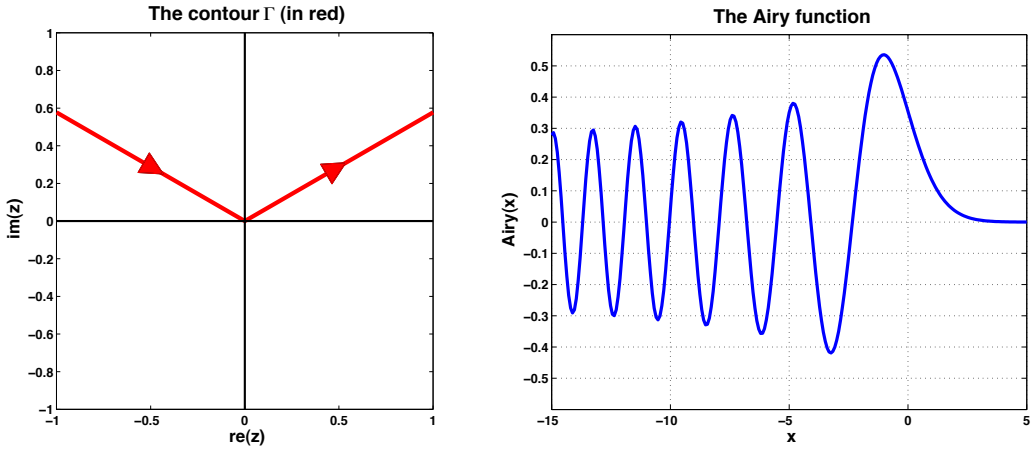


Figure 3.3: The Airy function. Left: integration contour. Right: plot of the function for a real argument.

Finally, note that $\displaystyle\int_\Gamma \frac{\mathrm{d}}{\mathrm{d}z}\left(e^{i(z\,\zeta+\frac{1}{3}\,z^3)}\right)\mathrm{d}z = 0$. Expanding the derivative, we see that this means that

$$\frac{\mathrm{d}^2 Ai}{\mathrm{d}\zeta^2} = \zeta\, Ai. \tag{3.45}$$

This is the Airy equation.

Remark 3.6 The Airy function appears in most wave-to-no-wave transitions, such as, for example: (i) Optics, where it describes what happens near **caustics**. Caustics are "bright" because of the same effect described in item **2** of this subsection. The difference is that t large is replaced by "ratio of curvature of the wave-fronts to wave-length" is large. (ii) **Tunneling** in quantum mechanics. Here the issue is that the transition from waves to no-waves is not sharp: there is an exponential tail that penetrates into the no-waves region. (iii) The front edge of the wave pattern produced by a ship [Kelvin's ship wave pattern].

4 Modulation theory

In this section we consider the behavior of "locally" monochromatic waves with slowly varying parameters (modulation theory). We present several approaches: (i) Multiple scales asymptotic expansions; (ii) Fourier Transform; and (iii) Average Lagrangian.

4.1 Modulation theory for 1-D scalar, single branch dispersive

Consider a wave-field in 1-D such that, at every point in space-time there is a clearly identifiable wave-length λ, wave-frequency f, and wave-amplitude A. However, these are not constant, but change from point to point. This makes sense if the wave-field is nearly mono-chromatic, with the wave properties changing slowly in space and time — where **slowly means** that significant changes occur over distances and time intervals much larger than the wave-length and/or the wave-period. Thus, in **a-dimensional units** where the **wave-length** and the **wave-period** are **used to a-dimensionalize** space and time, we postulate a situation where the wave-number, wave-frequency, and wave-amplitude satisfy

$$\boldsymbol{k = k(X,\, T),} \quad \boldsymbol{\omega = \omega(X,\, T),} \quad \text{and} \quad \boldsymbol{A = A(X,\, T),} \tag{4.1}$$

with $\boldsymbol{X = \epsilon\, x}$, $\boldsymbol{T = \epsilon\, t}$, and $\boldsymbol{0 < \epsilon \ll 1}$.

Note that ϵ is the ratio of the typical wave-length to the typical distance over which significant changes occur, which we have assumed is of the same order as ratio of the typical wave-period to the typical time over which significant changes occur. As long as the wave changes propagate at a speed $= O(\lambda\, f)$, this is a reasonable assumption, since the variations in space and in time are related by this speed. Since $\lambda\, f$ is the phase speed, and wave changes propagate at the group speed, it should be clear that:

The scaling in (4.1) is not appropriate for situations where $c_g/c_p \neq O(1)$.

What about the wave-phase? While k, ω, and A, are directly observable (at least $|A|$ is), the wave-phase θ is not. In fact, it is not even clear that there should be one. However, *if there is a phase,* at any fixed time it should increase by $2\,\pi\,\delta$ when x changes by $\lambda\,\delta$ — where $0 < \delta \ll 1$ is small enough that we can consider λ constant in the interval $[x,\, x + \lambda\,\delta]$. This leads to

$$\theta(x_1,\, t) - \theta(x_2,\, t) = \int_{x_1}^{x_2} k\, \mathrm{d}x \quad \text{for} \quad |x_2 - x_1| \ll 1 \quad \Longleftrightarrow \quad \theta_x = k. \tag{4.2}$$

A similar argument shows that it should be

$$\theta(x,\, t_1) - \theta(x,\, t_2) = -\int_{t_1}^{t_2} \omega\, \mathrm{d}t \quad \text{for} \quad |t_2 - t_1| \ll 1 \quad \Longleftrightarrow \quad \theta_t = -\omega. \tag{4.3}$$

Of course, these two equations are consistent if and only if the waves are conserved

$$\boldsymbol{k_t + \omega_x = 0,} \tag{4.4}$$

which is a consequence of the assumptions made in the first paragraph of this subsection. Hence, **there is a phase**, and we can write

$$\boldsymbol{\theta = \frac{1}{\epsilon}\,\Theta(X,\, T),} \quad \text{with} \quad \boldsymbol{k = \theta_x = \Theta_X} \quad \text{and} \quad \boldsymbol{\omega = -\theta_t = -\Theta_T.} \tag{4.5}$$

Note that

1. The phase is not a "slow" function of x and/or t, since it must change by $2\,\pi$ over one wave-length or one wave-period.
2. Equations (4.2–4.3) leave an undetermined free "constant" that can be added to θ. This free constant can, in fact, be a slow function of x and t (since the argument leading to (4.2–4.3) neglects second order effects — i.e.: $O(\delta^2)$ quantities). We can incorporate this quantity into the phase of the complex wave-amplitude A — see (4.6) below.

3. While the phase of A is a well defined mathematical object, it is generally not something available from observations, as it requires rather precise measurements of the wave field.

Consider now a scalar, single branch dispersive, wave system in 1-D. Then the arguments above indicate that we should investigate solutions of the form

$$\boldsymbol{u = A\, e^{i\,\theta}}, \quad \text{with } \boldsymbol{\theta = \frac{1}{\epsilon}\,\Theta(X,\,T)}, \;\; \boldsymbol{A = A(X,\,T)}, \;\; \boldsymbol{k = \theta_x}, \;\text{ and } \;\boldsymbol{\omega = -\theta_t}, \tag{4.6}$$

where $\boldsymbol{X = \epsilon\, x}$, $\boldsymbol{T = \epsilon\, t}$, and $\boldsymbol{0 < \epsilon \ll 1}$. Note that

4. Linear combinations of solutions of the form (4.6) are allowed. These are, of course, not single frequency.

5. Generalizations of the approach below to non-scalar systems, and/or multiple branch dispersive systems, and/or several dimensions, are easy to do. We restrict the presentation here to the simplest context, to keep the main ideas un-encumbered by technical issues.

The governing equation can be written in the form

$$\boldsymbol{i\,u_t = \Omega(D)\,u}, \quad \text{where } \boldsymbol{D = -i\,\frac{\partial}{\partial x}}, \;\; \boldsymbol{\omega = \Omega(k)} \tag{4.7}$$

is the dispersion relation, and Ω is a real valued analytic function.

There are many subtleties involved in defining what exactly $\Omega(D)$ means. We will not go into them. However, notice that: if $\Omega = P/Q$ is a quotient of polynomials, then (4.7) has a clear meaning: $i\,Q(D)\,u_t = P(D)\,u$.

Now:

$$\begin{aligned}
D\,u &= (k\,A + DA)\;e^{i\,\theta},\\
D^2\,u &= \left(k^2\,A + (Dk)\,A + 2\,k\,DA + O\left(\epsilon^2\right)\right)\;e^{i\,\theta},\\
D^3\,u &= \left(k^3\,A + 3\,k\,(Dk)\,A + 3\,k^2\,DA + O\left(\epsilon^2\right)\right)\;e^{i\,\theta},\\
\ldots &= \ldots\\
D^n\,u &= \left(k^n\,A + \frac{n\,(n-1)}{2}\,k^{n-2}\,(Dk)\,A + n\,k^{n-1}\,DA + O\left(\epsilon^2\right)\right)\;e^{i\,\theta}.
\end{aligned}$$

where $DA = -i\,\epsilon\,A_X$ and $Dk = -i\,\epsilon\,k_X$. Thus [9]

$$\boldsymbol{\Omega(D)\,u = \left(\Omega(k)\,A + \frac{1}{2}\,\Omega''(k)\,(Dk)\,A + c_g(k)\,DA + O\left(\epsilon^2\right)\right)\;e^{i\,\theta}.} \tag{4.8}$$

In addition

$$i\,u_t = (\omega\,A + i\,\epsilon\,A_T)\;e^{i\,\theta}. \tag{4.9}$$

Substituting (4.8–4.9) into (4.7), and collecting equal orders of ϵ yields

$$\boldsymbol{\omega = \Omega(k)}, \quad \text{and} \quad \boldsymbol{A_T + c_g(k)\,A_X + \frac{1}{2}\,(c_g(k))_X\,A = 0}, \tag{4.10}$$

with

$$\boldsymbol{k_T + \omega_X = 0} \qquad \boldsymbol{(\text{thus } \; k_T + c_g(k)\,k_X = 0)}, \tag{4.11}$$

following from $k = \Theta_X$ and $\omega = -\Theta_T = \Omega(k)$. These equations imply that

$$\boldsymbol{\left(A^2\right)_T + \left(c_g(k)\,A^2\right)_x = 0} \quad \text{and} \quad \boldsymbol{\left(|A|^2\right)_T + \left(c_g(k)\,|A|^2\right)_x = 0.} \tag{4.12}$$

Note that (4.10–4.12) *are the same as* (3.12–3.14) (far field limit).

[9] Strictly speaking, this shows that (4.8) applies when $\boldsymbol{\Omega}$ is a polynomial. More work is required for the general case.

Remark 4.1 A special case of these equations occurs when $k = k_0 =$ constant, and $\omega = \omega_0 = \Omega(k_0)$. Then

$$u = A(X,\, T)\, e^{i\,\theta} \quad \text{and} \quad \theta = k_0\, x - \omega_0\, t + \theta_0, \quad \text{with} \quad A_T + c_g(k_0)\, A_X = 0.$$

This corresponds to an amplitude modulated single frequency carrier wave (e.g.: AM radio signals). In the non-dispersive case, where $c_g =$ constant, it is also possible to have $A =$ constant, with a variable wave-number satisfying $k_T + c_g\, k_X = 0$ (e.g.: FM radio signals). ♣

4.1.1 Characteristic form and solution to the IVP (Initial Value Problem)

Assume known initial values for the modulation equations: $\boldsymbol{k(X,\, 0) = k_0(X)}$, $\boldsymbol{A(X,\, 0) = A_0(X)}$, with $\boldsymbol{\omega(X,\, 0) = \Omega(k_0)}$. Introduce now the characteristic curves in space-time, $\boldsymbol{X = X(T,\, \zeta)}$, defined by

$$\frac{dX}{dT} = c_g(k), \quad \text{with} \quad X(0,\, \zeta) = \zeta. \tag{4.13}$$

Then, since $k_T + c_g(k)\, k_X = 0$, along each curve

$$\frac{dk}{dT} = 0, \qquad \text{so that} \quad k \equiv k_0(\zeta). \tag{4.14}$$

$$\frac{dA}{dT} = -\frac{1}{2}\,(c_g(k))_X\, A, \quad \text{with} \quad A(0,\, \zeta) = A_0(\zeta). \tag{4.15}$$

Multiply $k_T + c_g\, k_X = 0$ by c_g', to obtain $(c_g)_T + c_g\, (c_g)_X = 0$.
Take ∂_X of this equation, and get $\alpha_T + c_g\, \alpha_X = -\alpha^2$, where $\boldsymbol{\alpha = (c_g(k))_X}$.
Hence

$$\frac{d\alpha}{dT} = -\alpha^2, \quad \text{so that} \quad \alpha = \frac{\alpha_0(\zeta)}{1 + T\, \alpha_0(\zeta)}, \quad \text{where} \quad \boldsymbol{\alpha_0(X) = (c_g(k_0(X)))_X}. \tag{4.16}$$

Therefore, from (4.15), $A = \frac{1}{\sqrt{1+\alpha_0(\zeta)\, T}}\, A_0(\zeta)$, and we can write the full solution in the form

$$X = c_g(k_0(\zeta))\, T + \zeta, \tag{4.17}$$

$$k = k_0(\zeta), \tag{4.18}$$

$$\omega = \Omega(k_0(\zeta)), \tag{4.19}$$

$$A = \frac{1}{\sqrt{1 + \alpha_0(\zeta)\, T}}\, A_0(\zeta), \tag{4.20}$$

$$\Theta = k\, X - \omega\, T + \Theta_0(\zeta) - k\, \zeta, \tag{4.21}$$

where the last equation follows from $\dfrac{D\Theta}{dT} = k\, c_g(k) - \omega$. Note that

6. *Equations (4.17–4.21) provide a solution of the equations in implicit form.* For every characteristic curve, labeled by the parameter ζ (the value of X where the curve intersects the X-axis at $T = 0$), the solution is given as a function of time T. To get the solution as a function of X and T, equation (4.17) must be used to find ζ as a function $\zeta = \zeta(X,\, T)$.

7. *The solution is valid as long as* $1 + \alpha_0(\zeta)\, T$ *remains positive.* Note that $X_\zeta = 1 + \alpha_0(\zeta)\, T$ (keeping T constant), so that this condition is precisely what is needed to be able to solve for $\zeta = \zeta(X,\, T)$. Beyond this point, some waves that started at different positions at $T = 0$ cross paths (because they have different group speeds). At the moment they cross, the hypothesis implicit in (4.6) no longer

holds,[10] and the calculations above break down. An example of this situation, where crossing of the wave paths (focusing) occurs in some region is studied in § 3.3. There we characterize the nature of the solutions in these regions.

4.1.2 The $T \gg 1$ limit of the IVP solution

Assume that $\alpha_0(\zeta) > 0$ in (4.17–4.21), and consider the $T \gg 1$ limit. In this limit $X \sim c_g(k_0(\zeta))\, T$, and the characteristics appear to arise from the single point $X = T = 0$. Hence ζ is no longer a good parameter for them. Instead, we use the value of k each one carries to label them. In particular, we think now of A_0 as a function of k, $A_0 = A_0(k)$, and assume that we also can write k_0' in terms of k, $k_0'(\zeta) = h(k)$ — so that $\alpha_0 = \Omega''(k)\, h(k)$. Thus we can write

$$X = c_g(k)\, T, \quad A = \sqrt{\frac{2\,\pi}{|\Omega''(k)|\, t}}\, U(k), \quad \text{and} \quad \Theta = k\, X - \omega\, T, \quad \text{with} \quad U = \frac{A_0}{\sqrt{2\,\pi\,\epsilon\,|h|}}. \tag{4.22}$$

This is, basically, the same result obtained using Fourier Transforms in §3.1 (far field limit). That is, equation (3.10) — except for the term $\nu_s\, \dfrac{\pi}{4}$ in θ (which can be absorbed into U).

4.1.3 The Fourier Transform approach

Here we consider a Fourier Transform approach to problems with slowly varying initial conditions. Specifically, to situations where

$$u(x,0) = A_0(X)\, e^{i\,\theta_0}, \quad \text{with} \;\; \theta_0 = \frac{1}{\epsilon}\,\Theta_0(X), \;\; X = \epsilon\, x, \quad \text{and} \quad 0 < \epsilon \ll 1. \tag{4.23}$$

The objective is to get some insight as to what happens beyond the points at which $1 + \alpha_0(\zeta)\, T$ vanishes in (4.17–4.21) — see (4.28).

Remark 4.2 Why do things as early in this subsection, when Fourier Analysis provides, basically, the same answers with less work? In fact, better answers, since it allows the parameters in the approximation to be related to the initial data without ambiguity, and does not run into problems such as what to do once $1 + \alpha_0(\zeta)\, T$ vanishes in (4.17–4.21). The **answer is that the earlier approach can be generalized to situations where Fourier Analysis cannot be used,** such as non-homogeneous situations (variable coefficients) or when weak nonlinear effects matter. ♣

Before proceeding, we must ascertain to what does (4.23) correspond in the Fourier Transform side. Thus consider the Fourier Transform

$$U(k) = \frac{1}{2\,\pi}\int_{-\infty}^{\infty} A_0(X)\, e^{i\,\theta_0}\, e^{-i\,k\,x}\, \mathrm{d}x = \frac{1}{2\,\pi\,\epsilon}\int_{-\infty}^{\infty} A_0(X)\, \exp\left(i\,\frac{1}{\epsilon}\,\phi\right)\, \mathrm{d}X, \tag{4.24}$$

where $\phi = \Theta_0(X) - k\, X$. The method of stationary phase (see §3.2), then yields

$$U(k) \sim \sum_s \frac{1}{\sqrt{2\,\pi\,\epsilon\,|k_0'(X_s)|}}\, A_0(X_s)\, \exp\left(\frac{i}{\epsilon}\,\phi_s(k) + i\,\nu_s\,\frac{\pi}{4}\right), \tag{4.25}$$

[10] Because a whole bunch of (initially well separated) wave-numbers pile up in a small region, the slowly varying wave property breaks down. This well occur when there is a region in the initial conditions where the group speed is decreasing, which will cause the waves in this region to focus.

where $\phi_s = \Theta_0(X_s) - k\,X_s$, $k_0 = \Theta_0'$, $\nu_s = \text{sign}(k_0'(X_s))$, and the sum is over the stationary points X_s, defined as the solutions to $k_0(X_s) = k$. Of course, this expression breaks down at the values of k where a stationary point with $k_0'(X_s) = 0$ occurs.

Motivated by (4.25), we now look at the asymptotic behavior of solutions to (4.7) of the form

$$u = \frac{1}{\sqrt{2\,\pi\,\epsilon}} \int_{-\infty}^{\infty} U(k)\, \exp\left(i\,(k\,x - \omega\,t) - \frac{i}{\epsilon}\,\psi(k)\right) \mathrm{d}k = \frac{1}{\sqrt{2\,\pi\,\epsilon}} \int_{-\infty}^{\infty} U(k)\, \exp\left(i\,\frac{\phi}{\epsilon}\right) \mathrm{d}k, \tag{4.26}$$

where $\phi = k\,X - \omega\,T - \psi(k)$, $\omega = \Omega(k)$, and $T = \epsilon\,t$. Then, using the method of stationary phase,

$$u \sim \sum_s \frac{1}{\sqrt{|\phi_s''|}}\, U(k_s)\, \exp\left(i\,\nu_s\,\frac{\pi}{4}\right)\, e^{i\,\theta}, \tag{4.27}$$

where: $\phi_s'' = -\Omega''(k_s)\,T - \psi''(k_s)$, $\nu_s = \text{sign}(\phi_s'')$, $\theta = \dfrac{1}{\epsilon}\,\Theta$, $\Theta = k_s\,X - \omega_s\,T - \psi(k_s)$, $\omega_s = \Omega(k_s)$, and the sum is over the stationary points $k_s = k_s(X,\,T)$, defined by $X = c_g(k_s)\,T + \psi'(k_s)$.

Equation (4.27) is the same as (4.17–4.21) provided that we identify $k_0 = k_0(\zeta)$ as the inverse function of $\psi'(k)$, so that $\zeta = \psi'(k_0)$. Then define: $\Theta_0 = k_s\,\zeta - \psi(k_s)$ and

$$A_0 = \frac{1}{\sqrt{|\psi''(k_s)|}}\, U\, \exp\left(i\,\nu_s\,\frac{\pi}{4}\right). \quad \text{Thus } \phi_s'' = -\psi''(k_s)\left(1 + \frac{\Omega''(k_s)}{\psi''(k_s)}\,T\right) = -\psi''(k_s)\,(1 + \alpha_0\,T)$$

follows upon identifying k_s with k_0, because $1 = \psi''(k_0)\,(k_0)_\zeta$.

The moral of this is that, **if we change (4.20) to**

$$A = \frac{1}{\sqrt{|1 + \alpha_0(\zeta)\,T|}}\, A_0(\zeta)\, \exp\left(-i\,\text{sign}\left(1 + \alpha_0(\zeta)\,T\right)\,\frac{\pi}{2}\right), \tag{4.28}$$

then the solution remains valid even when $1 + \boldsymbol{\alpha_0(\zeta)\,T} < 0$ — though it fails for $1 + \alpha_0(\zeta)\,T \approx 0$.

4.2 Modulation in the n-D case

We now consider the analog of (4.6) for a (scalar) dispersive equation $\quad \boldsymbol{i\,u_t = \Omega(-i\,\nabla_x)\,u}$, (4.29)
where: $\boldsymbol{\nabla_x}$ is the gradient operator in n-D, and $\boldsymbol{\omega = \Omega(\vec{k})}$ is the dispersion relation (Ω = real valued analytic function). That is: we look for solutions of the form

$$\boldsymbol{u = A\,e^{i\,\theta}}, \quad \text{with } \boldsymbol{\theta = \frac{1}{\epsilon}\,\Theta(\vec{X},\,T)}, \quad \boldsymbol{A = A(\vec{X},\,T)}, \quad \boldsymbol{\vec{k} = \nabla_x\theta}, \quad \text{and } \boldsymbol{\omega = -\theta_t}, \tag{4.30}$$

where $\boldsymbol{\vec{X} = \epsilon\,\vec{x}}$, $\boldsymbol{T = \epsilon\,t}$, and $\boldsymbol{0 < \epsilon \ll 1}$. Using the same approach as in §4.1, it is easy to see that this leads † to the equations

$$\boldsymbol{\theta_t + \Omega(\nabla_x\theta) = 0}, \quad \text{and} \quad \boldsymbol{(A^2)_t + \text{div}\,(\vec{c}_g(k)\,A^2) = 0}, \tag{4.31}$$

where $\boldsymbol{\vec{c}_g = \nabla_k\Omega}$ (here $\boldsymbol{\nabla_k}$ indicates the gradient operator with respect to $\boldsymbol{\vec{k}}$).

† It is left as an exercise to the reader to fill in the details.

Note that:

1. In terms of $\boldsymbol{\vec{k}}$ and $\boldsymbol{\omega}$, the first equation in (4.31) can be written in the form:

$$\boldsymbol{\vec{k}_t + \nabla_x\omega = 0} \text{ and } \boldsymbol{\text{curl}\,\vec{k} = 0}, \text{ with } \boldsymbol{\omega = \Omega(\vec{k})}. \tag{4.32}$$

In particular:
$$\boldsymbol{\vec{k}_t + (\vec{c}_g \cdot \nabla_x)\vec{k} = 0}. \tag{4.33}$$

2. Write $\boldsymbol{A = \rho\, e^{i\,\phi}}$ in polar form. Then $$(\boldsymbol{\rho^2})_t + \text{div}\,(\vec{\boldsymbol{c}}_{\boldsymbol{g}}\,\boldsymbol{\rho^2}) = \boldsymbol{0}, \tag{4.34}$$ as well as $\boldsymbol{\phi_t + \vec{c}_g \cdot \nabla_x\, \phi = 0}$.

3. For any function $\boldsymbol{\gamma = \gamma(\vec{k})}$, $$(\boldsymbol{\gamma\rho^2})_t + \text{div}\,(\vec{\boldsymbol{c}}_{\boldsymbol{g}}\,\boldsymbol{\gamma\rho^2}) = \boldsymbol{0}. \tag{4.35}$$ In particular, since the energy density has the form $\boldsymbol{\mathcal{E} = e(\vec{k})\,\rho^2}$, this yields the conservation of energy, with energy flux $\boldsymbol{\mathcal{F} = \vec{c}_g \mathcal{E}}$.

4.2.1 Characteristic form of the equations

Introduce the wave rays, the curves in space time defined by the equation $$\frac{\mathbf{d}\vec{\boldsymbol{x}}}{\mathbf{d}\boldsymbol{t}} = \vec{\boldsymbol{c}}_{\boldsymbol{g}}(\vec{\boldsymbol{k}}). \tag{4.36}$$ It is then easy to see that, along the rays:

4. $\dfrac{\mathbf{d}\vec{\boldsymbol{k}}}{\mathbf{d}\boldsymbol{t}} = \boldsymbol{0}$. Hence *the rays are straight lines, along which* $\vec{\boldsymbol{k}}$ *and* $\boldsymbol{\omega = \Omega(\vec{k})}$ *are constant.*

$$\vec{\boldsymbol{x}} = \vec{\boldsymbol{c}}_{\boldsymbol{g}}(\vec{\boldsymbol{k}})\,\boldsymbol{t} + \vec{\boldsymbol{\zeta}}, \quad \text{with} \quad \vec{\boldsymbol{k}} = \vec{\boldsymbol{k}}_{\boldsymbol{0}}(\vec{\boldsymbol{\zeta}}), \quad \text{and} \quad \boldsymbol{\omega = \omega_0(\vec{\zeta})} \tag{4.37}$$

where the subindex $\boldsymbol{0}$ indicate the values at $\boldsymbol{t = 0}$, where $\vec{\boldsymbol{x}} = \vec{\boldsymbol{\zeta}}$. Note that we use $\vec{\boldsymbol{\zeta}}$ to parameterize the rays: one ray for every initial data point.

6. $\dfrac{\mathbf{d}\boldsymbol{\theta}}{\mathbf{d}\boldsymbol{t}} = \vec{\boldsymbol{k}} \cdot \vec{\boldsymbol{c}}_{\boldsymbol{g}}(\vec{\boldsymbol{k}}) - \boldsymbol{\omega}$. Thus

$$\boldsymbol{\theta = \theta_0(\vec{\zeta}) + (\vec{k} \cdot \vec{c}_g(\vec{k}) - \omega)\, t = \vec{k} \cdot \vec{x} - \omega\, t + \theta_0(\vec{\zeta}) - \vec{k} \cdot \vec{\zeta}}, \tag{4.38}$$

where we have used $\vec{\boldsymbol{x}} = \vec{\boldsymbol{c}}_{\boldsymbol{g}}(\vec{\boldsymbol{k}})\,\boldsymbol{t} + \vec{\boldsymbol{\zeta}}$ to obtain the second equality.

Ray tubes and the energy. Let $\boldsymbol{\mathcal{R}_0}$ be a region in space (say, at time $t = 0$), with boundary $\boldsymbol{\partial\mathcal{R}_0}$. Now consider the set of all the rays that go through $\mathcal{R}_0$. These rays span a *ray tube* in space-time, with boundary made by all the rays that go through $\partial\mathcal{R}_0$. Because the energy flow is along the group speed, see (4.35), no energy enters or leaves the ray tube. Hence $$\int_{\boldsymbol{\mathcal{R}(t)}} |\boldsymbol{A}|^{\boldsymbol{2}}\mathbf{d}\boldsymbol{V} = \text{constant}, \tag{4.39}$$ where $\mathrm{d}V = \mathrm{d}x_1 \ldots \mathrm{d}x_n$ and $\mathcal{R}(t)$ is the intersection of the ray tube with space at time t. From this we see that *the average energy is inversely proportional to the cross sectional area of the ray tube;* in particular, when a ray tube collapses, this predicts that the wave amplitude goes to infinity. †

† This is the same issue that arises in the 1-D case — see: item **7**, (4.20), and (4.28) in §4.1. When this happens the approximation in (4.30) is no longer valid and modulation theory (as stated here) breaks down. The wave amplitude does not become infinity, though it does become much larger than elsewhere. In §3.3 we present an example resolving a situation of this type, with an expansion for the waves near the region where a lower order approximation predicts an infinity.

Because of the second equation in (4.31), it is also the case that $$\int_{\boldsymbol{\mathcal{R}(t)}} \boldsymbol{A^2}\mathbf{d}\boldsymbol{V} = \text{constant}. \tag{4.40}$$ By considering an "infinitesimal" ray tube enveloping an arbitrary ray, this can be used to compute the evolution of A along rays. However, there is a "better" way, which we use next to show that

$$\boldsymbol{A} = \left(\prod_{\boldsymbol{j=1}}^{\boldsymbol{n}} \frac{\boldsymbol{1}}{\sqrt{\boldsymbol{1 + \lambda_j\, t}}}\right) \boldsymbol{A_0(\vec{\zeta})} \tag{4.41}$$

where the $\boldsymbol{\lambda_j}$ are the eigenvalues of the matrix $\boldsymbol{\mathcal{A}_0}$ with entries $(\mathcal{A}_0)_{pq} = \partial_{x_q}(c_g^0)_p$ evaluated at $\vec{\zeta}$, with $(c_g^0)_p$ the components of $\vec{c}_g(\vec{k}_0(\vec{x}))$ — the group speed as a function of $\vec{x}$ at time $t = 0$. Note that: **(i)** *This reduces to* (4.20) *when* $n = 1$. **(ii)** If any of the eigenvalues is negative, this predicts an infinity at a critical time (ray tube collapse). **(iii)** If any eigenvalue is complex, then the square root must be interpreted as a principal value. In this case the complex conjugate eigenvalue also arises, with the sum of the two inverse square roots staying real.

Proof of (4.41). We write the equation for $\boldsymbol{A}$ in the form

$$\boldsymbol{A_t + (\vec{c}_g \cdot \nabla_x)\, A + \frac{1}{2}\,\mathrm{div}(\vec{c}_g)\, A = 0.} \tag{4.42}$$

Define the matrix $\boldsymbol{\mathcal{A}}$ by $\mathcal{A}_{pq} = \partial_{x_q}(c_g)_p$, so that $\mathrm{div}(\vec{\boldsymbol{c}}_{\boldsymbol{g}}) = \mathrm{Tr}(\boldsymbol{\mathcal{A}})$. Then, from the chain rule, $\boldsymbol{\mathcal{A} = \mathcal{J}\,\mathcal{K}}$, where $\mathcal{J}$ and $\mathcal{K}$ are the symmetric matrices: $\boldsymbol{\mathcal{J} = \{\partial_{k_n}\partial_{k_m}\Omega\}}$ and $\boldsymbol{\mathcal{K} = \{\partial_{x_n}k_m\} = \{\partial_{x_n}\partial_{x_m}\theta\}}$.
Since J is a function of $\vec{k}$, it is constant along the rays. On the other hand, applying ∂_{x_ℓ} $(1 \le \ell \le n)$ to (4.33) yields the equation $\frac{\mathrm{d}\mathcal{K}}{\mathrm{d}t} + \boldsymbol{\mathcal{K}\,J\,\mathcal{K} = 0}$. Hence, upon multiplying by $\mathcal{J}$ we obtain

$$\boldsymbol{\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}t} + \mathcal{A}^2 = 0 \quad\Longrightarrow\quad \mathcal{A} = (1 + \mathcal{A}_0\, t)^{-1}\,\mathcal{A}_0.} \tag{4.43}$$

Hence $\mathrm{div}(\vec{\boldsymbol{c}}_{\boldsymbol{g}}) = \mathrm{Tr}\left((\mathbf{1} + \boldsymbol{\mathcal{A}_0\, t})^{-1}\,\boldsymbol{\mathcal{A}_0}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Tr}\left(\log(1 + \mathcal{A}_0\, t)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\log\det(1 + \mathcal{A}_0\, t) = \frac{\mathrm{d}}{\mathrm{d}t}\log\prod(1 + \lambda_j\, t)$.
Using this in (4.42) yields

$$\boldsymbol{\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}t} + \frac{1}{2}\left(\sum_{j=1}^{n}\frac{\lambda_j}{1 + \lambda_j\, t}\right)\mathcal{A} = 0.} \tag{4.44}$$

The solution to this equation is (4.41). **QED**

4.2.2 Example: the classical limit of Quantum Mechanics

The Schrödinger equation (in 1-D) for a particle of mass m in a potential $\tilde{V} = \tilde{V}(\tilde{x})$ is

$$i\,\hbar\,\psi_{\tilde{t}} = -\frac{\hbar^2}{2\,m}\,\psi_{\tilde{x}\tilde{x}} + \tilde{V}\,\psi, \quad \hbar = \frac{h}{2\,\pi} = \text{reduced Planck's constant.}\ \dagger \tag{4.45}$$

† $h = 6.6260700410^{-34}$ m^2 kg/s

Selecting a-dimensional units $t = \tilde{t}/T$, $x = \tilde{x}/L$, and $V = T^2\,\tilde{V}/(m\,L^2)$, this takes the form

$$\boldsymbol{i\,\epsilon\,\psi_t = -\frac{\epsilon^2}{2}\,\psi_{xx} + V\,\psi, \quad \text{with} \quad \epsilon = \frac{\hbar\,T}{m\,L^2}.} \tag{4.46}$$

When L and T are "classical" scales, $\mathbf{0 < \epsilon \ll 1}$,
and we look for solutions of the form † $$\boldsymbol{\psi = A(x,\, t)\, e^{i\,\theta}}, \tag{4.47}$$
where $\boldsymbol{\theta = \frac{1}{\epsilon}\,\Theta(x,\, t)}$.

† Why? Because on the quantum scales (characterized by ϵ) the potential "looks" constant, so that the elementary solutions are plane wave exponentials.

This leads to the equations

$$\boldsymbol{\omega = \frac{1}{2}\,k^2 + V, \quad k_t + \omega_x = 0, \quad \text{and} \quad (A^2)_t + (k\,A^2)_x = 0,} \tag{4.48}$$

where $\boldsymbol{\omega = -\Theta_t}$, and $\boldsymbol{k = \Theta_x}$ — note that the group speed here is $\boldsymbol{c_g = k}$.

Proof. $\epsilon\,\psi_t = (-i\,\omega\,A + \epsilon\,A)\,e^{i\,\theta}$ and $\epsilon^2\,\psi_{xx} = \left(-k^2\,A + i\,\epsilon\,(k_x\,A + 2\,k\,A_x) + \epsilon^2\,A_{xx}\right)\,e^{i\,\theta}$. The first equation follows from the $O(1)$ terms, the last from the $O(\epsilon)$ terms, and the middle one from the definition of ω and k.

These equations imply the "conservation of probability" $$(|\boldsymbol{A^2}|)_t + (\boldsymbol{k}\,|\boldsymbol{A^2}|)_x = \boldsymbol{0}, \tag{4.49}$$ which flows at the group speed.

Introduce now the rays, defined by $\frac{\mathrm{d}x}{\mathrm{d}t} = k$. Then, from $0 = k_t + \omega_x = k_t + k\,k_x + V_x$, it follows that

$$\frac{\mathbf{d}\boldsymbol{x}}{\mathbf{d}\boldsymbol{t}} = \boldsymbol{k} \quad \text{and} \quad \frac{\mathbf{d}\boldsymbol{k}}{\mathbf{d}\boldsymbol{t}} = -\boldsymbol{V_x}. \tag{4.50}$$

Or, in dimensional variables

$$\frac{\mathbf{d}\tilde{\boldsymbol{x}}}{\mathbf{d}\tilde{\boldsymbol{t}}} = \frac{\mathbf{1}}{\boldsymbol{m}}\,\tilde{\boldsymbol{p}} \quad \text{and} \quad \frac{\mathbf{d}\tilde{\boldsymbol{p}}}{\mathbf{d}\tilde{\boldsymbol{t}}} = -\tilde{\boldsymbol{V}}_{\tilde{\boldsymbol{x}}}, \quad \text{with} \quad \boldsymbol{p} = \frac{\boldsymbol{m}\,\boldsymbol{L}}{\boldsymbol{T}}\,\boldsymbol{k}. \tag{4.51}$$

These are the classical mechanics equations for a particle in a potential, where $\tilde{p}$ is the momentum. Note also that, in the limit $\epsilon \to 0$, if the probability is concentrated at a point, the probability moves at the particle speed. Thus classical mechanics is fully recovered in the $\epsilon \to 0$. Of course, if ϵ is merely small, then over times $O(1/\epsilon)$ dispersion kicks in, and an initially sharply confined probability disperses. †

† For $L = 1$ meter, $T = 1$ second, and $m = 1$ kg, $\epsilon = 6.63 \times 10^{-34}$. So you would have to wait $\approx 4.8 \times 10^{25}$ years for this to happen. Don't expect the baseball ball to disperse during a game.

More generally, consider a generic Schrödinger equation of the (a-dimensional) form

$$\boldsymbol{i\,\epsilon\,\psi_t = H(\vec{x},\, -i\,\epsilon\,\nabla)\,\psi = \mathcal{H}\,\psi}, \tag{4.52}$$

for some Hamiltonian H (see remark 4.3). Then the same process as above leads to the "Hamilton-Jacobi" equation

$$\boldsymbol{\Theta_t + H(\vec{x},\, \nabla\,\Theta) = 0} \tag{4.53}$$

for for the phase. Now introduce $\vec{\boldsymbol{k}} = \boldsymbol{\nabla\,\Theta}$, and the rays $\frac{\mathrm{d}x_p}{\mathrm{d}t} = H_{k_p}$, $1 \le p \le$ dimension. Then (4.53) yields

$$\frac{\mathbf{d}\vec{\boldsymbol{x}}}{\mathbf{d}\boldsymbol{t}} = \boldsymbol{\nabla_k H} \quad \text{and} \quad \frac{\mathbf{d}\vec{\boldsymbol{k}}}{\mathbf{d}\boldsymbol{t}} = -\boldsymbol{\nabla_x H}, \tag{4.54}$$

where ∇_k and ∇_x are the gradients with respect to $\vec{k}$ and $\vec{x}$. Now, **(4.54) is a "classical" Hamiltonian system with Hamiltonian $\boldsymbol{H = H(\vec{x},\, \vec{k})}$.**

Remark 4.3 The reverse path: from classical to quantum. Knowing the classical Hamiltonian is not enough to define the operator on the right in (4.52). The reason is commutativity. For example, a term $x\,k$ could translate into $-i\,\epsilon\,x\,\partial_x$, or $-i\,\epsilon\,\partial_x\,x$, or a mixture of both; all of them different as operators. However, $\mathcal{H}$ must be self-adjoint, † and in many situations this is enough to pick a unique path $H \to \mathcal{H}$. ♣

† $\mathcal{H}$ self-adjoint is needed to have conservation of probability — the integral of $|\psi|^2$ must be constant. It also guarantees that the eigenvalues of $\mathcal{H}$, the energy levels, are real. Incidentally, this is also what is needed to make the system dispersive.

4.3 Average Lagrangian

For the Average Lagrangian approach, see the book: *Linear and Nonlinear Waves*, by G. B. Whitham, Wiley-Interscience, New York, 1974.

5 Loose topics related to modulation theory

This subsection includes notes for various topics related to Modulation Theory — see §4.

5.1 Conservation of waves and group speed

For linear waves the *concept of group speed is very generic and it is a direct consequence of the existence of a phase,* for which only the notion of a slowly varying (periodic) plane is needed. At a more general level (no linearity required), *the existence of a phase is also directly linked to various "conservation of waves" principles.*

To understand this imagine a system with periodic plane waves. That is, solutions such that

1. Their only space-time dependence is through a phase $\boldsymbol{\theta = \vec{k}\cdot\vec{x} - \omega\, t}$, where $\vec{\boldsymbol{k}}$ is the wave-vector and $\boldsymbol{\omega}$ is the wave-frequency;
2. They are periodic in $\boldsymbol{\theta}$, of period $\boldsymbol{2\,\pi}$ (normalizing the period to $\boldsymbol{2\,\pi}$ determines the size of $\vec{\boldsymbol{k}}$, else only its direction is determined).

These solutions will generally depend on various parameters
(e.g.: amplitude), and involve a restriction (the *dispersion relation*) $\boldsymbol{\omega = \Omega(\vec{k},}$ parameters$\boldsymbol{)}$. (5.1)

Remark 5.1 Note that there is no linearity assumption here. Periodic plane waves do not require linear equations to exist, hence (5.1–5.5) applies to both the linear and nonlinear cases. On the other hand, *periodic plane waves require constant coefficients* [space and time homogeneity]. However, it is possible to extend the approach to non-homogeneous systems as long as the wave-length and wave-period of the waves involved is much shorter than the scales over which the system changes. Nevertheless, **here we will restrict our attention to the constant coefficients case.** ♣

Assuming now that the periodic plane waves have continuous dependence on their parameters (i.e.: $\vec{\boldsymbol{k}}$ and the other parameters can take values in some open sets) we can now postulate a solution to the system that is "locally" a periodic plane wave. That is:

> *In any neigborhood in space-time of size comparable with the wave-length and wave-period, the solution is well approximated by some periodic plane wave. But on larger scales the wave-parameters smoothly change with $\vec{\boldsymbol{x}}$ and $\boldsymbol{t}$.* (5.2)

Given this, we can *define the phase* $\boldsymbol{\theta = \theta(\vec{x},\, t)}$, (5.3)
as follows: in any neighborhood (where $\vec{\boldsymbol{k}}$ and $\boldsymbol{\omega}$ are well defined) it should be
$\boldsymbol{\nabla\theta = \vec{k}}$ and $\boldsymbol{\theta_t = -\omega}$ — in this fashion the Taylor expansion for $\boldsymbol{\theta}$ near any point agrees (to leading order) with the phase for the locally valid plane wave. †

† *Can we do this?* Given that the description in (5.2) is not mathematically rigorous, a "proof" is not possible. In the end we must **assume** that the solution described in (5.2) has an associated phase, as described by (5.3). We can only offer arguments indicating that this assumption is reasonable, such as:
(i) The Taylor expansion argument below (5.3); and
(ii) The following **intuitive geometrical argument:** Consider a wave-crest for the solution. Along such a wave-crest θ is some constant (we are free to choose this constant). The wave-crest cannot abruptly disappear, as this would violate the last sentence in (5.3). At any point along the wave-crest we can use the local periodic solution to select the "next" wave crest (and the value θ should have there). Following

this second wave-crest allows now us to define θ in the region between the first and second wave-crest. We can then repeat this for wave-crest after wave-crest to slowly, but surely, get θ defined "everywhere".

Now, given that $\boldsymbol{k = \nabla\,\theta}$, $\boldsymbol{\omega = -\theta_t}$, $\boldsymbol{\theta_{x_i t} = \theta_{t x_i}}$, and $\boldsymbol{\theta_{x_i x_j} = \theta_{x_j x_i}}$,
we can write the following equations: $$\boldsymbol{\vec{k}_t + \nabla\,\omega = 0}, \qquad (5.4)$$
and $$\boldsymbol{\mathrm{curl}(\vec{k}) = 0}. \qquad (5.5)$$
Of course: **(5.1) applies as well.** Note that

- For any j, in the restricted space-time plane $\boldsymbol{(x_j,\, t)}$, $\boldsymbol{\frac{1}{2\,\pi}\,k_j}$ is the number of waves per unit length, while $\boldsymbol{\frac{1}{2\,\pi}\,\omega}$ is the corresponding flux. Thus the j^{th} equation in (5.4) expresses the conservation of waves in this plane. Because of (5.2), no new waves can appear suddenly. Hence the waves are conserved.
- We can also look at the waves, at any given time, in any of the planes $\boldsymbol{(x_i,\, x_j)}$. Then, along each of the directions, $\boldsymbol{\frac{1}{2\,\pi}\,k_n}$ is the number of waves per unit length. Again, we can argue that (5.2) means that no waves can be created or destroyed on any given patch, which leads to the conservation equation $\boldsymbol{(k_i)_{x_j} - (k_j)_{x_i}}$ — i.e.: the components of (5.5).

Thus **(5.4–5.5) are equivalent to the conservation of waves.** (5.6)
However, see §5.1.2.

Assume now a linear system. Then $\boldsymbol{\Omega}$ is a function of $\vec{k}$ only, and the equations can be re-written in the form
$$\boldsymbol{\vec{k}_t + (\vec{c}_g \cdot \nabla)\vec{k} = 0} \quad \text{and} \quad \boldsymbol{\mathrm{curl}(\vec{k}) = 0}, \qquad (5.7)$$
where $\boldsymbol{\omega = \Omega(\vec{k})}$ and $\boldsymbol{\vec{c}_g = \nabla\Omega}$.

Proof. The j^{th} component of $\boldsymbol{\nabla\omega}$ in (5.4) is: $\partial_{x_j}\omega = \sum_n (\partial_{k_n}\Omega)(\partial_{x_j} k_n) = \sum_n (\partial_{k_n}\Omega)(\partial_{x_n} k_j) = (\vec{c}_g \cdot \nabla) k_j$. ♣

Note that the general solution to (5.7)
can be written in the implicit form $$\boldsymbol{\vec{k} = \vec{k}_0(\vec{x} - \vec{c}_g(\vec{k})\,t)}, \qquad (5.8)$$
where $\boldsymbol{\vec{k}_0}$ is the initial data for $\vec{k}$, assumed to satisfy $\boldsymbol{\mathrm{curl}(\vec{k}_0)}$.

Proof. Note that (5.4) implies $(\mathrm{curl}(\vec{k}))_t = \mathbf{0}$. Hence, if (5.5) is satisfied initially, it holds for all time. Introduce now the "characteristic" curves, defined by $\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \vec{c}_g$. Then the first equation in (5.7) is equivalent to: $\vec{k}$ is constant along characteristics. But then so is $\vec{c}_g$, and we can write $\vec{x} = \boldsymbol{c_g(\vec{k})\,t} + \vec{x}_0$ for the characteristic that starts at $\vec{x}_0$ — along which $\vec{k} = \vec{k}_0(\vec{x}_0)$. Since $\vec{x}_0 = \vec{x} - \vec{c}_g(\vec{k})\,t$, this proves the result. ♣

This ends the argument showing that the statements in the first paragraph of this subsubsection apply. However, one may ask: **What happens with the group speed when nonlinearity is present?**

5.1.1 Group speed and nonlinearity

The presentation here is an adaptation taken from §14.2 of the book *Linear and Nonlinear Waves* by G. B. Whitham.

When nonlinearity is present, $\boldsymbol{\omega}$ is not just a function of $\vec{k}$, but the wave amplitude as well. † For simplicity, let us examine the *1-D case in the weakly nonlinear regime*
(small, but finite amplitude waves). Then we can expand ‡ $$\boldsymbol{\Omega = \Omega_0(k) + \epsilon^2 a^2 \Omega_2(k) + \ldots}, \qquad (5.9)$$
where $\boldsymbol{0 < \epsilon \ll 1}$ is a measure of the nonlinearity in the

system, $a = O(1)$ is the (scaled by ϵ) wave amplitude, and Ω_0 is the linear dispersion function.

† It may depend on other parameters as well (e.g.: mean value), but here we look at the simplest scenario.

‡ Ω does not depend on the phase of the amplitude (which can be absorbed into θ). Furthermore, if the dependence is smooth, it has to be via the square of the amplitude.

Now, while (5.4) remains valid, it is no longer an equation for the wave-number alone, as it involves the amplitude as well. Hence, in order to analyze what happens, we need this second equation. Unfortunately there is no way (at least, not one I know) to obtain this second equation by a generic argument such as the one leading to (5.4–5.5). On the other hand, for the weakly nonlinear regime we can do something fairly generic. For consider the linear equations, which can be written in the form

$$\boldsymbol{Y_t + \mathcal{A}_0\, Y_x = 0}, \quad \text{where} \quad \begin{pmatrix} c_g^0 & 0 \\ \Omega_0'' a^2 & c_g^0 \end{pmatrix}, \quad \boldsymbol{Y} = \begin{pmatrix} k \\ a^2 \end{pmatrix}, \tag{5.10}$$

and $c_g^0 = \Omega_0'$ is the linear group speed – the zero in the top right entry of $\mathcal{A}_0$ is why in this case the evolution of the wave-number is independent of the wave-amplitude.

When weakly nonlinear effects are added, $\mathcal{A}_0$ is replaced by an expansion $\mathcal{A} = \mathcal{A}_0 + \epsilon^2 \mathcal{A}_2 + O(\epsilon^2)$. In fact, from (5.4) and (5.9) we can write

$$\mathcal{A} = \begin{pmatrix} c_g^0 + \epsilon^2 a^2 \Omega_2' + \dots & \epsilon^2 \Omega_2 + \dots \\ \Omega_0''\, a^2 + \epsilon^2\, \gamma_{21} + \dots & c_g^0 + \epsilon^2\, \gamma_{22} + \dots \end{pmatrix}, \tag{5.11}$$

where γ_{21} and γ_{22} are some functions of k and a^2. The eigenvalues of $\mathcal{A}$ are then given by

Case $\Omega_2\, \Omega_0'' > 0$. Then $$\lambda = c_g^0 \pm \epsilon\, a \sqrt{\Omega_2\, \Omega_0''} + O(\epsilon^2). \tag{5.12}$$
Thus the eigenvalues are real. However, unlike the linear case **there are now two distinct characteristic speeds.** The group speed splits into two speeds, and changes in the wave parameters are carried by these two speeds.

Case $\Omega_2\, \Omega_0'' < 0$. Then $$\lambda = c_g^0 \pm i\, \epsilon\, a \sqrt{\Omega_2\, \Omega_0''} + O(\epsilon^2). \tag{5.13}$$
Thus the eigenvalues are complex. This means that the system in (5.10) is not well-posed. Small perturbations can grow arbitrarily fast, with no limit in the growth rate as the frequency gets larger. What this means is that **the scenario postulated in (5.2) is not possible for a system where the nonlinear corrections produce $\Omega_2\, \Omega_0'' < 0$.** In this case a uniform wave-train evolves away from this configuration on scales comparable with the wave-period and wave-length, and (typically) splits into non-periodic separate units (e.g.: breathers/oscillons). This situation is known as a **Modulational instability.**

5.1.2 Violation of conservation

There are actually many situations where (5.6) is violated by waves patterns, ‡ with waves abruptly vanishing or appearing. For example: at singular rays in Geometrical Optics — see Figure 7.6. For another example: ripples in sand, formed on both beaches and dunes by the action of the wind, often exhibit this phenomena. † I cannot display pictures for these examples without violating copyright, but you can find very many by searching the web with the code names: "spiral waves", "sand wave patterns", "underwater sand ripples", "Patterns in bedforms", etc.

In all these examples conservation fails because some wave-crests merge or split, or sometimes just end at a tip. These situations cannot be described by equations such as (5.4–5.5), and present many challenges.

But, since wave patterns are usually organized around its defects, they are important and the subject of much research.

‡ More generally, (5.2). Wave fronts can develop kinks and other singularities that do not violate conservation, but still destroy the *smooth and slow change* assumption. Example: at caustics.

† Of course, this only works if the sand is left undisturbed — i.e.: no people or dogs walking around it. Similar ripples arise underwater in calm tropical beaches with sandy bottoms.

---

6 Energy for Dispersive Systems

This section has notes for various topics related to conservation of energy for dispersive systems.

6.1 Conservation of energy for 1-D first order dispersive equations

Here we consider linear, constant coefficients, first order in time, 1-D dispersive equations. These are very simple equations for which a complete and thorough analysis is possible. We consider equations with both real and complex valued solutions. The main results are:

1. These equations have an infinite number of conservation laws, $\boldsymbol{e_t + f_x = 0}$, which can be written "explicitly" — see §6.1.5. For every such law, the corresponding "wave averages" satisfy $\boldsymbol{\mathcal{F} = c_g\, \mathcal{E}}$, where $\boldsymbol{c_g}$ is the group speed — see (6.5–6.7), (6.24–6.27), and remark 6.4.
2. These equations have an associated variational principle, see §6.1.2.
3. These equations have an associated Hamiltonian form, see §6.1.3.

Below we start by considering the simples case (a polynomial dispersion relation) in §6.1.1–§6.1.3, and then show how to extend/generalize the results to the general case in §6.1.4.

6.1.1 Introduction and the simplest case (polynomial dispersion relation)

Linear, first order, 1-D dispersive equations are characterized by the fact that their solutions are linear combinations of exponentials of the form $\boldsymbol{u = e^{i\,(k\,x - \omega\,t)}}$, where the dispersion relation $\boldsymbol{\omega = \Omega(k)}$ has only one branch. Thus they are associated with first order in time, scalar, dispersive pde. Let us **consider the case when $\boldsymbol{\Omega(k) = \sum_0^d a_n\, k^n}$ is a polynomial** — since Ω is real valued for k real, **the coefficients $\boldsymbol{a_n}$ are all real.** The associated pde is

$$\boldsymbol{u_t = -i \sum_0^d a_n (-i\,\partial_x)^n\, u = \mathcal{L}_\Omega\, u.} \tag{6.1}$$

Note that

- **a.** $\boldsymbol{\mathcal{L}_\Omega}$ is a skew-adjoint operator in $\boldsymbol{L^2([-\infty,\, \infty])}$, or with respect to periodic B.C. (6.2)
- **b.** $\boldsymbol{\mathcal{L}_\Omega}$ is a real (resp. pure imaginary) operator if $a_n = 0$ for n even (resp. odd). (6.3)
- **c.** $\int |u|^2\, dx$ is time independent **(conservation of "energy").** (6.4)

Proof. $(\int |u|^2 dx)_t = (\langle u,\, u\,\rangle)_t = \langle u_t,\, u\,\rangle + \langle u,\, u_t\,\rangle = \langle \mathcal{L}_\Omega\, u,\, u\,\rangle + \langle u,\, \mathcal{L}_\Omega\, u\,\rangle$, as follows from (6.2). Here $\langle \cdot,\, \cdot\,\rangle$ is the standard L^2 inner product. †

† $\boldsymbol{\langle f,\, g\,\rangle = \int f(x)\, g^*(x)\, dx}$, **where the asterisk indicates complex conjugation.**

d. An alternative proof for (6.4) follows from the fact that **$|u|^2$ is a conserved density**. (6.5)

Proof:

$$\begin{aligned}(|u|^2)_t &= \sum_{n=0}^{d}(-i)^{n+1}a_n\left(u^*\,(\partial_x^n u)+(-1)^{n+1}(\partial_x^n u^*)\,u\right)\\ &= \partial_x\left(\sum_{n=0}^{d}(-i)^{n+1}a_n\sum_{m=0}^{n-1}(-1)^m(\partial_x^m u^*)(\partial_x^{n-m-1}u)\right).\end{aligned} \tag{6.6}$$

e. The **wave average "energy" flux $\mathcal{F}$** is related to the **density $\mathcal{E}=|A|^2$** by: $\mathcal{F}=c_g\,\mathcal{E}$. (6.7)

Proof. Substitute $u=A\,e^{i\,(k\,x-\omega\,t)}$, with k and ω real, into the density and flux of (6.6).
Then collect equal powers of k, and note that $(-i)^m(i)^{n-m-1}=-(-1)^m\,i^{n+1}$.

About the "energy". What we call the energy above need not be the "physical" energy in any particular context where the equation occurs (see remark 6.4). The important point is that (linear) dispersive equations have quadratic quantities that are conserved — the one here is an example; the Hamiltonians in §6.1.3 are another. For any such quantity the analog of the relationship in (6.7) applies. This is a consequence of Modulation theory (developed elsewhere in these notes). For a direct proof of this see §6.1.4. ♣

6.1.2 Variational principles (VarPr)

f. Consider the system of equations made up by (6.1) and the equation $v_t^*=\mathcal{L}_\Omega\,v^*$. (6.8)

This **system follows from the VarPr $\delta J=0$** where with u and v fixed at the start and end time.

$$J=\int\left(\frac{1}{2}\left(\langle u_t,\,v^*\rangle-\langle u,\,v_t^*\rangle\right)-\langle\mathcal{L}_\Omega\,u,\,v^*\rangle\right)\mathrm{d}t, \tag{6.9}$$

Proof. The result follows from $\delta J=\int\left(\langle u_t-\mathcal{L}_\Omega\,u,\,\delta v^*\rangle-\langle\delta u,\,v_t^*-\mathcal{L}_\Omega\,v^*\rangle\right)\mathrm{d}t$, where we use (6.2).

Since (6.8) is the equation for $v=u^*$, we *avoid introducing an extra equation for the variational principle* by using

$$\begin{aligned}J &= \int\left(\tfrac{1}{2}\left(\langle u_t,\,u\rangle-\langle u,\,u_t\rangle\right)-\langle\mathcal{L}_\Omega\,u,\,u\rangle\right)\mathrm{d}t\\ &= i\,\mathrm{Im}\left(\int\left(\langle u_t,\,u\rangle-\langle\mathcal{L}_\Omega\,u,\,u\rangle\right)\mathrm{d}t\right)\mathrm{d}t.\end{aligned} \tag{6.10}$$

However, **this works only for complex variations:** that is δu is allowed to be complex. ‡

‡ In particular, note that J vanishes when $\mathcal{L}_\Omega$ is a real operator and u is real.

To check this, let us calculate δJ for J as in (6.10). This leads to

$$\begin{aligned}\frac{1}{2\,i}\,\delta J &= \mathrm{Im}\left(\int\langle u_t-\mathcal{L}_\Omega\,u,\,\delta u\rangle\mathrm{d}t\right)\\ &= \int\left(\langle b_t-\mathcal{A}\,b-\mathcal{B}\,a,\,\delta a\rangle-\langle a_t-\mathcal{A}\,a+\mathcal{B}\,b,\,\delta b\rangle\right)\mathrm{d}t,\end{aligned} \tag{6.11}$$

where $u=a+i\,b$ and $\mathcal{L}_\Omega=\mathcal{A}+i\,\mathcal{B}$ (real and imaginary parts). Then $\frac{\delta J}{\delta a}=0$ and $\frac{\delta J}{\delta b}=0$ is equivalent to (6.1), but both equations are needed.

g. When $a_n=0$ for n even, and **$\mathcal{L}_\Omega$ is a real operator,** it is *convenient to have a variational principle that involves only real valued solutions.* To do this let $u=\phi_x$, and write the equation in the form

$$\phi_{xt}=\mathcal{L}_S\,\phi,\quad\text{with}\quad\mathcal{L}_S=\mathcal{L}_\Omega\,\partial_x\ \textbf{(real and symmetric).} \tag{6.12}$$

Then $$J = \frac{1}{2}\int (\langle \phi_t,\, \phi_x\,\rangle + \langle \mathcal{L}_S\,\phi,\, \phi\,\rangle)\,\mathrm{d}t \qquad (6.13)$$
works for ϕ **real valued.**

6.1.3 Hamiltonian form

h. Equation (6.1) is Hamiltonian, † $u_t = -2\,i\,\frac{\delta H}{\delta u^*}$, with Hamiltonian: $$H = \frac{i}{2}\,\langle \mathcal{L}_\Omega\,u,\, u\,\rangle. \qquad (6.14)$$
Note that, because of (6.2) and the definition of the scalar product, H is a real valued functional of u and u^*. † See remark 6.1.

Proof. We have: $-2\,i\,\delta H = \langle \mathcal{L}_\Omega\,\delta u,\, u\,\rangle + \langle \mathcal{L}_\Omega\,u,\, \delta u\,\rangle = -\langle \delta u,\, \mathcal{L}_\Omega\,u\,\rangle + \langle \mathcal{L}_\Omega\,u,\, \delta u\,\rangle$.
Thus, using the definition of the inner product, $-2\,i\,\frac{\delta H}{\delta u^*} = \mathcal{L}_\Omega\,u$. **QED**

i. When $a_n = 0$ for n even, and $\mathcal{L}_\Omega$ **is a real operator,** it is *convenient to have a Hamiltonian form involving only real valued solutions.* Hence write $$\mathcal{L}_\Omega = \partial_x\,\mathcal{M}, \qquad (6.15)$$
where $\mathcal{M}$ is **real and symmetric**. Then $$u_t = \partial_x \frac{\delta H}{\delta u}, \text{ where } H = \frac{1}{2}\,\langle \mathcal{M}\,u,\, u\,\rangle, \qquad (6.16)$$
where u is real valued.
The proof of this is as in prior cases. As for why (6.16) is Hamiltonian, see remark 6.2.

Remark 6.1 Complex form of the Hamiltonian equations. The standard form for a Hamiltonian system is $\dot q_n = H_{p_n}$ and $\dot p_n = -H_{q_n}$, where $H = H(\vec q,\, \vec p)$ is some (real valued) function. Let $z_n = q_n + i\,p_n$, so that $H = H(\vec z,\, \vec z\,^*)$. Then, using the chain rule:
$\dot q_n = i\,H_{z_n} - i\,H_{z_n^*}$ and $\dot p_n = -H_{z_n} - H_{z_n^*}$. Thus $$\frac{\mathrm{d}\vec z}{\mathrm{d}t} = -2\,i\,\nabla^* H, \qquad (6.17)$$
where ∇^* denotes the gradient with respect to $\vec z\,^*$. ♣

Remark 6.2 Generalized Hamiltonian equation. Consider the set of all the square integrable real valued functions, periodic with period $2\,\pi/\mu > 0$, zero mean, and scalar product $\langle f,\, g\,\rangle = \frac{\mu}{2\,\pi}\int f(x)\,g(x)\,\mathrm{d}x$ — the integral being over one period. Let now $H = H[u]$ be some functional defined in this set, and consider the equation $$u_t = \partial_x \tfrac{\delta H}{\delta u}, \qquad (6.18)$$
which preserves the mean. Equation (6.18) *is a classical Hamiltonian system.*

Proof. Write $u = \sum_{n=1}^{\infty} \sqrt{2\,n\,\mu}\,(p_n\,\cos(n\,\mu\,x) + q_n\,\sin(n\,\mu\,x))$. Then $\partial_{p_n} H = \sqrt{2\,n\,\mu}\,\langle \frac{\delta H}{\delta u},\, \cos(n\,\mu\,x)\,\rangle$ and $\partial_{q_n} H = \sqrt{2\,n\,\mu}\,\langle \frac{\delta H}{\delta u},\, \sin(n\,\mu\,x)\,\rangle$. On the other hand, taking the scalar product of (6.18) with the cosines yields $\frac{1}{2}\,\sqrt{2\,n\,\mu}\,\dot p_n = \langle \partial_x \frac{\delta H}{\delta u},\, \cos(n\,\mu\,x)\,\rangle = -n\,\mu\,\langle \frac{\delta H}{\delta u},\, \sin(n\,\mu\,x)\,\rangle$. That is $\dot p_n = -\partial_{q_n} H$. Similarly, the scalar products with the sines yields $\dot q_n = \partial_{p_n} H$. **QED** ♣

6.1.4 Non-polynomial dispersion relation

Finally: most of the results in this section extend easily to the case where Ω is "arbitrary". Say, for example, a rational function $\Omega(k) = p(k)/q(k)$, with p and q polynomials, real valued for k real, and $q \neq 0$ on the real axis. Then the equation is $$q(-i\,\partial_x)\,u_t = -i\,p(-i\,\partial_x)\,u. \qquad (6.19)$$
The results that extend easily are those that follow from:
(1) $\mathcal{L}_\Omega = -i\,p(-i\,\partial_x)/q(-i\,\partial_x)$ is skew-adjoint; (2) $\mathcal{L}_\Omega\,\partial_x$ is real and symmetric; and
(3) $\mathcal{L}_\Omega = \partial_x\,\mathcal{M}$, where $\mathcal{M}$ is real and symmetric;

all of which remain valid. The **exceptions are (6.5–6.7),** because (6.6) involves a detailed manipulation of the equation which does not extend to the general case. **Below we include a general proof of (6.5–6.7).**

Equation (6.1) can be written in the general form $$u_t = -i \int \Omega(k)\, U(k,\, t)\, e^{i\,k\,x}\, \mathrm{d}k, \tag{6.20}$$
where U **is the Fourier Transform of** u. Then

$$(|u^2|)_t = -i \iint (\Omega(k) - \Omega(\ell))\; U(k,\, t)\, U^*(\ell,\, t)\, e^{i\,(k-\ell)\,x}\, \mathrm{d}k\, \mathrm{d}\ell. \tag{6.21}$$

Now define $$Q(k,\, \ell) = \frac{\Omega(k) - \Omega(\ell)}{k - \ell}, \quad \text{with} \quad Q(k,\, k) = c_g(k), \tag{6.22}$$
so that Q has no singularities. Then

$$(|u^2|)_t = -\partial_x \iint Q(k,\, \ell)\, U(k,\, t)\, U^*(\ell,\, t)\, e^{i\,(k-\ell)\,x}\, \mathrm{d}k\, \mathrm{d}\ell, \tag{6.23}$$

which generalizes (6.6). In particular: $u = A\, e^{i\,(k_0\,x - \omega_0\,t)}$ corresponds to $U = A\, e^{-i\,\omega_0\,t}\, \delta(k - k_0)$. Using this with (6.22–6.23) yields (6.7) again $$\mathcal{F} = c_g\, \mathcal{E}, \tag{6.24}$$
where $\mathcal{E} = |A|^2 =$ **energy density**, and $\mathcal{F} =$ **energy flux** (see remark 6.4).

As pointed out below (6.7), $|u|^2$ need not be the actual "physical" energy. Thus, consider **an "arbitrary" conservation law for (6.20)** $$e_t + f_x = 0. \tag{6.25}$$
That is, (6.25) is satisfied by any solution to (6.20), where e **and** f **are linear combinations of terms** $(\mathcal{D}_1\, u)(\mathcal{D}_2\, u)^*$**, where the** $\mathcal{D}_j$ **are linear, constant coefficients, differential operators in space-time** † — e.g.: $\mathcal{D} = \partial_x^7\, \partial_t^3$.

† Note that e and f need not be real valued, though this would be true for a physical energy.

Hence
$$e = \iint E(k,\, \ell)\, U(k,\, t)\, U^*(\ell,\, t)\, e^{i\,(k-\ell)\,x}\, \mathrm{d}k\, \mathrm{d}\ell \quad \text{and} \quad f = \iint F(k,\, \ell)\, U(k,\, t)\, U^*(\ell,\, t)\, e^{i\,(k-\ell)\,x}\, \mathrm{d}k\, \mathrm{d}\ell, \tag{6.26}$$
for some functions E and F.
Note that (6.26) is more general than the assumptions on e and f above.
Substituting (6.26) into (6.25), and using that (6.25) should apply to any solution to (6.20), we obtain
Hence $U = A\, e^{-i\,\omega_0\,t}\, \delta(k - k_0)$ in (6.26) yields (6.24).
$$F(k,\, \ell) = \frac{\Omega(k) - \Omega(\ell)}{k - \ell}\, E(k,\, \ell). \tag{6.27}$$

Remark 6.3 Nonlocality. The flux in (6.23) is, generally, nonlocal. The same applies to (6.27); even if E corresponds to a local density (i.e.: it is a polynomial), F is nonlocal. Not surprising, given that (6.20) is nonlocal. However, *when* Ω *is a rational function, it is possible to write things in an "almost" local fashion.* Then
$$Q = \frac{p(k)\, q(\ell) - p(\ell)\, q(k)}{k - \ell}\, \frac{1}{q(k)\, q(\ell)}, \tag{6.28}$$
where the first factor is actually a polynomial. Hence, if we introduce $$v = \textstyle\int \frac{U}{q}\, e^{i\,k\,x}\, \mathrm{d}k, \tag{6.29}$$
which is related to u by the ode $$q(-i\,\partial_x)\, v = u, \tag{6.30}$$
the flux in (6.23) can be written in terms of derivatives of v — if E in (6.26) is a polynomial, the same trick works for F in (6.27). See example 6.1. ♣

Example 6.1 Linear regularized KdV (Korteweg-de Vries) equation. $$u_t + u_x - u_{xxt} = 0. \tag{6.31}$$
We can also write the equation as the system $$u_t + v_x = 0 \quad \text{and} \quad u = v - v_{xx}. \tag{6.32}$$
From this it immediately follows that $$(u^2)_t + (v^2 - v_x^2)_x = 0. \tag{6.33}$$

Note: we consider real valued solutions; see remark 6.4.

Another conservation law for this equation is $$(u^2 + u_x^2)_t + (u^2 - 2\,u\,u_{xt})_x = 0. \quad (6.34)$$

The equation follows from the Lagrangian $$L = \tfrac{1}{2}\left(\phi_x\,\phi_t + \phi_x^2 + \phi_{xx}\,\phi_{xt}\right), \quad (6.35)$$

where $u = \phi_x$.

We can also write the Hamiltonian form $$u = \partial_x \tfrac{\delta H}{\delta u}, \text{ where } H = -\tfrac{1}{2}\int v\,u\,\mathrm{d}x, \quad (6.36)$$

and $u = v - v_{xx}$. ♣

Remark 6.4 Averages for energy and fluxes. In (6.6) the density and flux are such that, when $u = A\,e^{i\,(k\,x-\omega\,t)}$ is substituted into it, the result is a constant. Thus $\mathcal{E}$ and $\mathcal{F}$ in (6.7) *do not involve any "average" at all.* The same is true for the calculations in (6.24–6.27). This makes sense for complex valued solutions. However, for situations where one is interested in real valued solutions (i.e.: **Ω is an odd function of k**), the appropriate wave form to use is $u = a\,\cos(k\,x - \omega\,t + \theta_0) = a\,\cos\theta$, not a complex exponential $A\,e^{i\,\theta}$, and the process of obtaining $\mathcal{E}$ and $\mathcal{F}$ involves an actual average over θ.

Below we show how to obtain (6.24) for real equations with real valued solutions.

Note that (6.26–6.27) still applies, since for u real valued $u^2 = u\,u^*$ (similar for other quadratic terms). Then u is real valued $\iff$ $$U^*(k,\,t) = U(-k,\,t), \quad (6.37)$$

while e and f real valued means $$E^*(k,\,\ell) = E(-k,\,-\ell) \text{ and } F^*(k,\,\ell) = F(-k,\,-\ell). \quad (6.38)$$

Substitute $U = A\,e^{-i\,\omega_0\,t}\,\delta(k-k_0) + A^*\,e^{i\,\omega_0\,t}\,\delta(k+k_0)$, $A = \frac{1}{2}\,a\,e^{i\,\theta_0}$ (a and θ_0 real constants), into (6.26). Then

$$e = \tfrac{1}{4}\left(E(k_0,\,k_0) + E(-k_0,\,-k_0) + E(k_0,\,-k_0)\,e^{i\,2\,\theta} + E(-k_0,\,k_0)\,e^{-i\,2\,\theta}\right)a^2, \quad (6.39)$$

and

$$f = \tfrac{1}{4}\left(F(k_0,\,k_0) + F(-k_0,\,-k_0) + F(k_0,\,-k_0)\,e^{i\,2\,\theta} + F(-k_0,\,k_0)\,e^{-i\,2\,\theta}\right)a^2. \quad (6.40)$$

Averaging this leads to $$\mathcal{E} = \tfrac{1}{2}\mathrm{Re}(E(k_0,\,k_0))\,a^2 \text{ and } \mathcal{F} = c_g(k_0)\,\mathcal{E}. \quad (6.41)$$

Note that $\mathcal{F} = c_g(k_0)\,\mathcal{E}$ applies **even if** e and f are not real valued and (6.38) fails. ♣

6.1.5 Infinite number of conservation laws

Note that (6.25–6.27) provides a way to write an infinite number of conservation laws: *For "any" function $E = E(k,\,\ell)$, define: e and f using* (6.26)*, with F given by* (6.27)*. Then* (6.25) *applies.* **QED**

Furthermore, note that

- **A.** If E is a polynomial, the e is local (can be written as a finite linear combination of quadratic terms involving u and its derivatives). Then f is also local if Ω is a polynomial, and can be written in an "almost" local fashion using the process in remark 6.3
- **B.** If Ω is odd (so that the equation admits real valued solutions) and $E^*(k,\,\ell) = E(-k,\,-\ell)$, then e and f are real valued when u is.

Of course, for any given equation, none of these laws (with a few exceptions) will have no physical meaning.

7 Geometrical Optics

In this section we present a brief introduction to Geometrical Optics, which has many points in common with the theory in §4 (modulation). The main ideas behind these (and other) theories were developed separately, and independently, in various fields (ater waves, optics, quantum mechanics, plasma physics, acoustics, seismology, etc.), and it is instructive to see them in more than one incarnation.

7.1 The Eikonal equation

In this subsection we derive the Eikonal equation by two different approaches, and study its solutions.

7.1.1 Wave front propagation with prescribed normal velocity

Consider a wave-front in n-D (mathematically: a hyper-surface) which propagates normal to itself at a known velocity $c > 0$ — see figure 7.1. The wave-front could be an idealization for the leading edge of

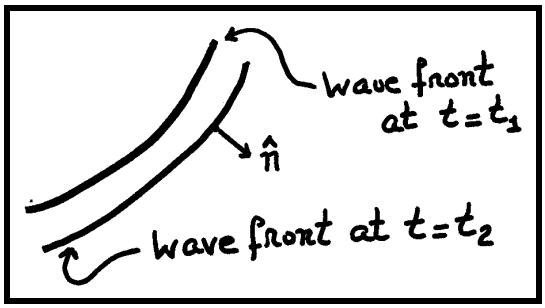


Figure 7.1: Wave front normal propagation
Picture illustrating the evolution of a wave front propagating normal to itself at some prescribed (known) velocity: $c > 0$. The picture here is for the 2-D case, but the same principle applies to any number of dimensions.

a forest fire propagating into the un-burnt forest, or the wave-fronts for light or a high frequency sound wave, or a sonic boom generated by a jet plane, or the shock wave generated by a super-nova explosion, etc. These cases differ by how the velocity c is determined. Here **we will only consider the situations where $c = c(\vec{x})$ is a reasonable approximation** (front speed almost always depends on more factors than just location, but this is a good approximation in many cases).

As long as the wave-fronts do not self-intersect, we can describe them
via a phase function ϕ, defined by: $$\boldsymbol{\phi(\vec{x}) = t} \quad \textbf{if and only if } \boldsymbol{\vec{x}} \in \textbf{wave-front at time } \boldsymbol{t}. \tag{7.1}$$
Thus it should be $\boldsymbol{\phi(\vec{x} + c\,\hat{n}\,dt) = \phi(\vec{x}) + dt}$, where $\hat{\boldsymbol{n}}$ is the **unit normal** to the wave-front at $\vec{\boldsymbol{x}}$. Expanding the left side of this equation yields

$$c\,\hat{n}\cdot\nabla\phi = 1, \quad \text{i.e.:} \quad \boldsymbol{c^2\,(\nabla\phi)^2 = 1}, \tag{7.2}$$

where the second equation follows by using $\hat{\boldsymbol{n}} = \boldsymbol{\nabla\phi/|\nabla\phi|}$. The equation $\boldsymbol{c^2\,(\nabla\phi)^2 = 1}$ is called the **Eikonal equation**, and can be used to determine the wave-fronts, given that the wave-front at some time $t = 0$ is known. This is done in §7.1.3, where we show how to solve the equation given the surface $\phi = 0$ and the direction of propagation (i.e.: to which "side" of the wave-front should $\hat{n}$ point towards).

7.1.2 High frequency, monochromatic, waves for the wave equation

We show here that the Eikonal equation can be used to describe the behavior of high, single frequency waves in set-ups governed by the wave equation (optics, sound, etc.). Write the equation in a-dimensional units

$$u_{tt} - \mathrm{div}(c^2\,\nabla u) = 0, \tag{7.3}$$

where $c = c(\vec{x}) > 0$ is $O(1)$ with $O(1)$ derivatives. We now look for solutions whose wave-length λ is much smaller than any other length scale in the problem — e.g.: curvature of the wave fronts, or scales over which c varies significantly. Specifically, we seek solutions of the form

$$\boldsymbol{u \sim A(\vec{x})\, e^{i\,\omega\,(t-\phi(x))} = A\, e^{i\,\theta}, \quad \text{where} \quad \omega \gg 1.} \tag{7.4}$$

Note that the position of any wave-front, as a function of time, is given by $\phi = t + \phi_0$, where $\phi = \phi_0$ is the wave-front at time $t = 0$. It follows that

$$\begin{aligned} u_{tt} &\sim -\omega^2\, A\, e^{i\,\theta} \quad \text{and} \\ \operatorname{div}(c^2\,\nabla u) &\sim \left(-\omega^2\, c^2\, A\,(\nabla\phi)^2 - i\,\frac{\omega}{A}\operatorname{div}(c^2\, A^2\,\nabla\phi) + \operatorname{div}(c^2\,\nabla A)\right) e^{i\,\theta}. \end{aligned}$$

Then, collecting equal powers of ω we obtain the

$$\textbf{Eikonal equation } \boldsymbol{c^2\,(\nabla\phi)^2 = 1} \quad \text{and the } \textbf{Transport} \text{ equation: } \boldsymbol{\operatorname{div}(c^2\, A^2\,\nabla\phi) = 0}. \tag{7.5}$$

Note that

$$\boldsymbol{\hat{n} = c\,\nabla\phi} \tag{7.6}$$

is the **normal to the wave-fronts in the direction of propagation.**

Remark 7.1 One can obtain higher order approximations by expanding $A = A_0 + \omega^{-1}\, A_1 + \omega^{-2}\, A_2 + \ldots$, where A_0 solves the transport equation and the A_j provide higher order corrections to the amplitude. However, these corrections do not introduce any new qualitative information. ♣

7.1.3 Rays, ray tubes, and the solution to the Eikonal equation

Given a solution, define the **rays** by

$$\boldsymbol{\frac{d\vec{x}}{ds} = c\,\nabla\phi}. \tag{7.7}$$

Note that, from (7.6), s is the **arclength** along the rays.

Furthermore, along a ray

$$\boldsymbol{\frac{d\phi}{ds} = (c\,\nabla\phi)\cdot\nabla\phi = \frac{1}{c}}. \tag{7.8}$$

Hence, since $\phi = t +$ constant defines the wave fronts:

$$\boldsymbol{\frac{ds}{dt} = c}. \tag{7.9}$$

That is: the fronts move along the rays at speed c — note that the results in §7.1.1 show that this is equivalent to the Eikonal equation.

By taking the gradient of the equation, written in the form $(\nabla\phi)^2 = 1/c^2$, it is easy to see that along the rays †

$$\boldsymbol{\frac{d\nabla\phi}{ds} = \nabla\frac{1}{c}}. \tag{7.10}$$

† The gradient yields $2\,[(\nabla\phi)\cdot\nabla]\,\nabla\phi = \nabla c^{-2} = 2\,c^{-1}\,\nabla c^{-1}$.

This last equation, together with (7.7), shows that **the rays are straight lines if $c =$ constant.**

Remark 7.2 Solve the Eikonal using rays. Equations (7.7–7.8) and (7.10) provide an ode system, which allows to solve for $(\vec{x},\, \nabla\phi,\, \phi)$ along a ray (if these quantities are known somewhere). In particular, if a wave-front is known (as well as which way it moves), we can use it to initialize the ode on each ray as it crosses the known front ((7.6) provides $\nabla\phi$). Then, by solving the ode for each ray, the solution to the Eikonal equation follows, parameterized by s and the point $\vec{\zeta}$ along the known front where the ray starts. This solution remains valid as long as the rays do not cross (ray tubes do not collapse) — see below. ♣

Now write the complex wave amplitude in **polar form $A = \rho\, e^{i\,\psi}$**. Then it is easy to see that the transport equation in (7.5) is equivalent to

$$\frac{d\psi}{ds} = 0 \quad \text{and} \quad \boldsymbol{\operatorname{div}(c^2\,\rho^2\,\nabla\phi) = 0}. \tag{7.11}$$

The first equation simply says that the phase of A is constant along rays. The second is explained below.

Remark 7.3 Ray tubes and energy. Given a set Ω_1 on a wave-front at some time $t = t_1$, we define the **ray tube through Ω_1** as the set of all the rays that cross the wave-front through Ω_1 — see figure 7.2. Then $\Omega = \Omega(t)$ is the intersection of the ray tube with the wave-front at time t.

Using the second equation in (7.11) and the divergence theorem (Gauss) we arrive at (7.12), which states that **the wave energy flows along the ray tubes** (see remark 7.4). In particular, **when a ray tube collapses,** the equations predict that: **the wave-amplitude must have an infinity. However, the equations are no longer valid when this happens** — see remark 7.5. ♣

$$\int_{\Omega(t)} c\,\rho^2\,\mathrm{d}S = \textbf{constant}, \qquad (7.12)$$

$\mathrm{d}S$ = area element on the wave-front.

Note about (7.12). When using Gauss theorem, note that $c^2\,\rho^2\,\nabla\phi$ is parallel to the walls of the ray tube, which thus give no contribution. On the other hand, on the two ends of the ray tube (Ω_1 and $\Omega(t)$), the normal vector is given by (7.6).

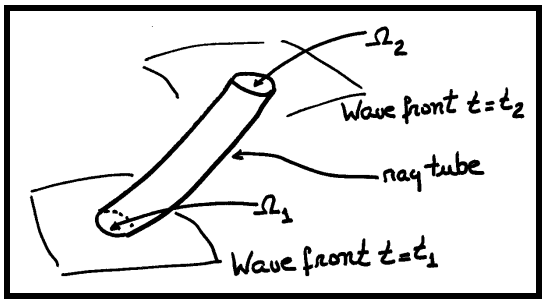


Figure 7.2: Ray tubes and energy

The picture illustrates a ray tube, comprised of all the rays that intersect the wave front on some set Ω_1 at some time $t = t_1$. Then $\Omega(t)$ is the intersection of the ray tube with the wave front at time t. The wave energy flows along ray tubes; thus the energy in the set Ω_1 at time $t = t_1$ is the same as the energy in the set Ω_2 at time $t = t_2$.

Remark 7.4 Energy density on a wave-front. The conservation of energy equation, $\boldsymbol{E_t + \mathrm{div}(\vec{F}) = 0}$, associated with (7.3) has: $\boldsymbol{E = \frac{1}{2}\,u_t^2 + \frac{1}{2}\,c^2\,(\nabla u)^2}$ and $\boldsymbol{\vec{F} = -c^2\,u_t\,\nabla u}$. Thus, for a solution like the one in (7.4), the average (over one period) leading order energy fluxes and densities are: †

$$\boldsymbol{\mathcal{E} = \tfrac{1}{2}\,\omega^2\,\rho^2} \text{ and } \boldsymbol{\vec{\mathcal{F}} = \tfrac{1}{2}\,\omega^2\,c^2\,\rho^2\,\nabla\phi}. \qquad \text{(r7.4a)}$$

However, $\boldsymbol{\mathcal{E}}$ is energy per unit volume. To obtain the energy density per wave-front unit area, we multiply $\boldsymbol{\mathcal{E}}$ by the wave-length $\boldsymbol{\lambda = \frac{2\,\pi\,c}{\omega}}$.

$$\boldsymbol{\mathcal{E}_{wf} = \pi\,\omega\,c\,\rho^2}. \qquad \text{(r7.4b)}$$

This is the energy density in (7.12) ♣

† To compute energies and fluxes a real valued solution is needed. Thus use $\mathrm{Re}(u)$ in the calculations. Note: the second equation in (7.11) is $\mathbf{0} = \mathrm{div}(\vec{\mathcal{F}}) = \boldsymbol{\mathcal{E}_t} + \mathrm{div}(\vec{\mathcal{F}})$, because $\mathcal{E}$ is time independent.

Remark 7.5 Ray tube collapse and failure of the equations. When a ray tube collapses, because the wave-fronts move along the rays at speed c (with the rays normal to the wave-front), † the wave-fronts focus along the collapse direction, and develop infinite curvature. This then violates the hypothesis above (7.4), and it becomes no longer true that the wave-length is much smaller than any other length scale in the problem. As a consequence, **the approximations involved cease to be valid, and the expansion fails.**

† This follows from (7.7–7.9).

Furthermore, note that the map $(\vec{\zeta},\, s) \to \vec{x}$ (see remark 7.2) is invertible before ray tube collapse, so that the *"solution by rays" in remark* 7.2 yields the solution $\phi = \phi(\vec{x})$ to the Eikonal equation. *Beyond this point the solution becomes multiple valued,* as several rays may go through a given point in space. ‡ ♣

‡ Fortunately, this does not mean that Geometrical Optics becomes useless beyond this point. In fact, the multiple values simply mean that the solution is there a linear combination of several solutions of the form in (7.4), one for each value of ϕ. The key question is then: but what amplitude should one use, since at the collapse A does not just become multiple

valued; it blows up. The answer comes through expansions valid in the regions where the collapse occurs, which provide *connection formulas* relating the amplitudes along the rays (pre-collapse to post-collapse) — the situation is somewhat similar to the one studied in §3.3. We will see some examples of how to deal with situations where Geometrical Optics fails later in these notes. ♣

7.1.4 Fermat's principle

Note that eliminating $\nabla\phi$ from (7.7) and (7.10) yields

$$\frac{\mathbf{d}}{\mathbf{d}s}\left(\frac{1}{c}\frac{\mathbf{d}\vec{x}}{\mathbf{d}s}\right)=\boldsymbol{\nabla}\left(\frac{1}{c}\right). \tag{7.13}$$

This equation is consistent with $s=$ arclength, provided it is initialized with $\left|\frac{\mathbf{d}\vec{x}}{\mathbf{d}s}\right|=\mathbf{1}$.

Proof: Let $\vec{k}=\frac{\mathbf{d}\vec{x}}{\mathbf{d}s}$. Then, using that $\frac{\mathbf{d}}{\mathbf{d}s}\frac{1}{c}=\vec{k}\cdot\boldsymbol{\nabla}\frac{1}{c}$, we obtain from the equation $\frac{\mathbf{d}k^2}{\mathbf{d}s}=\mathbf{2}\,c\,(\mathbf{1}-k^2)\,\vec{k}\cdot\boldsymbol{\nabla}c^{-1}$. Hence if $k^2=1$ somewhere, $k^2=1$ everywhere. **QED**

Consider now two points along a ray, $\vec{a}$ and $\vec{b}$. Then the time it takes a wave-front to travel from $\vec{a}$ to $\vec{b}$ is given by

$$T=\int_{\vec{a}}^{\vec{b}}\mathbf{d}t=\int_{\vec{a}}^{\vec{b}}\frac{1}{c}\,\mathbf{d}s. \tag{7.14}$$

Claim: **The rays are the stationary paths for the travel time T.** (7.15)

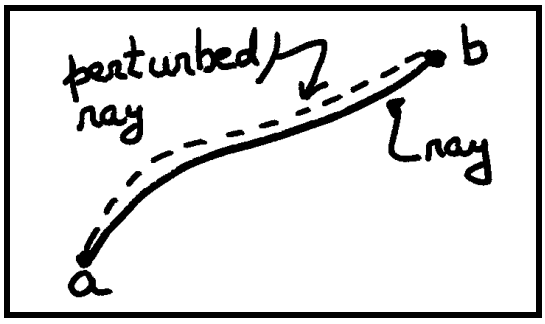


Figure 7.3: Fermat's principle: extremizing ray

Fermat's principle states that rays are the extrema of the optical path. The picture on the left illustrates a ray connecting the points a and b, and a perturbation to the ray. For any such perturbation the optical path change is higher order than the size of the perturbation (if the perturbation is small enough).

Proof. We begin by writing T for an arbitrary path connecting $\vec{a}$ to $\vec{b}$,

$$T[\vec{x}]=\int_0^1\frac{v}{c}\,\mathbf{d}\tau, \tag{7.15a}$$

where $\vec{x}=\vec{x}(\tau)$, $v=\left|\frac{\mathbf{d}\vec{x}}{\mathbf{d}\tau}\right|>0$ † (hence $\mathbf{d}s=v\,\mathbf{d}\tau$), $\vec{a}=\vec{x}(0)$, and $\vec{b}=\vec{x}(1)$.

† For the rays $\left|\frac{\mathbf{d}\vec{x}}{\mathbf{d}s}\right|=1$. Thus we only consider parameterizations with $v>0$.

This form of writing the travel time has the advantage that, when considering perturbations to the path, we do not have to worry about parameterizing the perturbed path by arc-length.

Consider now a perturbed path, $\vec{x}_0+\vec{x}_1$, where $\vec{x}_1$ vanishes at $\tau=0,\,1$ — see figure 7.3. Assuming that $\vec{x}_1$ is small, we expand: [11]

$$v=\sqrt{v_0^2+2\,\vec{v}_0\cdot\vec{v}_1+v_1^2}=v_0+v_0^{-1}\vec{v}_0\cdot\vec{v}_1+\text{HOT},$$

where $\vec{v}_j$ is the τ-derivative of $\vec{x}_j$ and $v_j=|\vec{v}_j|$. Similarly:

$$\alpha=\alpha_0+(\boldsymbol{\nabla}\alpha)_0\cdot\vec{x}_1+\text{HOT},$$

where $\alpha=1/c$, and the subscript zero indicates evaluation at $\vec{x}_0$. It follows that

$$\begin{aligned}T[\vec{x}_0+\vec{x}_1] &= T[\vec{x}_0]+\int_0^1\alpha_0\,v_0^{-1}\,\vec{v}_0\cdot\vec{v}_1\,\mathrm{d}\tau+\int_0^1 v_0\,(\nabla\alpha)_0\cdot\vec{x}_1\,\mathrm{d}\tau+\text{HOT}\\ &= T[\vec{x}_0]+\int_0^1\left(v_0\,(\nabla\alpha)_0-\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\alpha_0\,v_0^{-1}\,\vec{v}_0\right)\right)\cdot\vec{x}_1\,\mathrm{d}\tau+\text{HOT},\end{aligned} \tag{7.15b}$$

where we did an integration by parts in the first integral. However,

$\vec{x}_0$ is an extrema (or stationary) path if and only if: the

[11] HOT = Higher Order Terms.

linear term in (7.15b) vanishes for all choices of $\vec{x}_1$. That is:

$$v_0\,(\nabla\alpha)_0 - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\alpha_0\, v_0^{-1}\,\vec{v}_0\right) = \mathbf{0}. \quad (7.15c)$$

But $v_0\,\mathrm{d}\tau = \mathrm{d}s$. Hence (7.13) and (7.15c) are the same. **QED**

Remark 7.6 The rays are extrema, not minima. In many places you will find Fermat's principle stated as *rays are local minima of the travel time* (or even just minima). **This is, generally, not correct** — i.e.: some rays are a minimum, or a local minimum, but they do not have to be. Below a simple example to illustrate this. ♣

Example 7.1 Mirror laws and example where a ray is not a minimum *(in 2-D).*[12]

Below we **examine when the rays from mirror reflections correspond to local minimums, or maximus, of the travel time T.** Let a mirror be at $y = m(x) > 0$, such that $m'(0) = 0$ and $c \equiv 1$ below the mirror — see the left panel in figure 7.4, and **Note #1**. We will consider rays that go from $P_1 = (x,\, y) = (-1,\, 0)$, to the mirror, and back to $P_2 = (x,\, y) = (1,\, 0)$. However, **let us first get some intuition.**

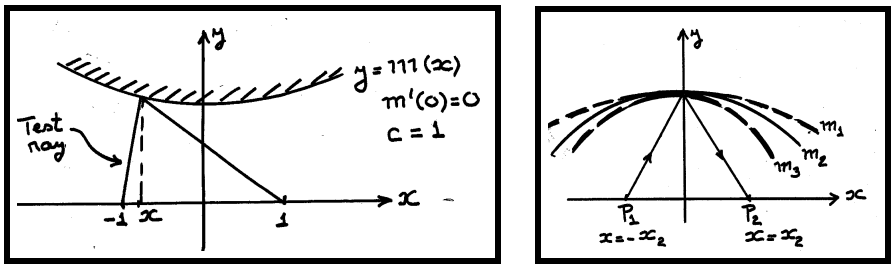


Figure 7.4: Examples of rays that do not minimize the optical time.

Intuitive argument. Imagine that the mirror is an ellipse, with foci at P_1 and P_2 — say, $m = m_2$ in the right panel of figure 7.4. Then T is constant for straight lines from P_1 to the mirror at $(x,\, m(x))$, and back to P_2 (call such a line L_x). Hence all L_x are rays — see **Note #2**. (7.16)

Let us now concentrate L_0. If the mirror is slightly changed, so that the curvature at $x = 0$ is larger (e.g.: $m = m_2$), or smaller (e.g.: $m = m_1$), L_0 remains a ray (because it satisfies "angle of incidence = angle of reflection" — see **Note #3**). However, it should be obvious that (a-b) below apply. (7.17)

(a) If $m = m_3$, L_x is shorter than L_0 for $x \neq 0$ small. That is: T is a local maximum for L_0.

(b) If $m = m_1$, L_x is longer than L_0 for $x \neq 0$ small. That is: T is a local minimum for L_0. See **Note #4**.

That is, roughly: *the ellipse is the "boundary shape" separating local maximums from local minimums.*

Note #1. Because $c = 1$, the rays are straight lines within the media, and travel time $= T =$ ray length.

Note #2. The statement in (7.16) is the well known fact that, if we place a light source at one of the foci of an elliptical mirror, the light focus back on the other. This is the basis for lithotripsy, with sound instead of light.

Note #3. That **mirror reflection satisfies "angle of incidence = angle of reflection"** also follows from Fermat's principle — see item **3** in the comments after the example.

Note #4: The direct path from P_1 to P_2 is always shorter. Thus, even when the mirror reflected path is a local minimum, it is not a global minimum.

Mathematics. In general (figure 7.4 left panel) the travel time for L_x is

$$T = \sqrt{(1+x)^2 + m^2(x)} + \sqrt{(1-x)^2 + m^2(x)}. \quad (7.18)$$

Using the fact that $m'(0) = 0$, it is easy to see that $T'(0) = 0$, i.e.: L_0 is a ray (see **Note #5**). This corresponds to a local minimum (resp.

[12] I would like to thank **Alexander Zaslavsky**, who found a serious mathematical error in a prior version of this example.

maximum) if $T''(0) > 0$ (resp. $T''(0) < 0$).
A straightforward calculation yields: $$T''(0) = 2\,\alpha\,((1+m_0^2)(1+m_0 m''(0)) - 1), \quad (7.19)$$
where $\alpha = (1+m_0^2)^{-1.5}$ and $m_0 = m(0)$
— see **Note #6**. Hence, a **strict local minimum** occurs for $$m''(0) > -m_0/(1+m_0^2) = \kappa_c, \quad (7.20)$$
while $m''(0) < \kappa_c$ yields a **strict local maximum.** When $m''(0) = \kappa_c$ higher order derivatives are needed to decide. It is then also **possible to have neither a minimum, nor a maximum;** e.g.: following the intuitive argument, have m track m_1 for $x < 0$. and m_3 for $x > 0$.

Note #5. Lines other than L_0 may be rays, depending on the shape of the mirror. The ellipse is the extreme case where all L_x satisfy Fermat's principle.

Note #6. The calculation leading to (7.19) is a bit cumbersome. The following approach is less error prone: (i) Write $T = \sqrt{g_+} + \sqrt{g_-}$, where $g_\pm = (1 \pm x)^2 + m^2$. (b) Calculate the derivatives in terms of $g_\pm$. (c) When done substitute the values for the derivatives of $g_\pm$ at $x = 0$.

Consistency. The equation for the ellipse with foci at $(\pm 1, 0)$, going through $(0,\, m_0)$, is $$\frac{y^2}{m_0^2} + \frac{x^2}{1+m_0^2} = 1. \quad (7.21)$$
It is easy to check that $y''(0) = \kappa_c$. Hence the two approaches give the same answer.

The situation in this example is generic. Once you have curved mirrors, curved interfaces, of generic variable c, you cannot longer expect the rays to be minimums for T. ♣

Further comments and issues related to remark 7.6 and example 7.1.

1. This example involves a mirror. Can this be done without mirrors? *Yes:* replace the mirror by an interface, where the media above the interface has $c > 1$. Then if $m(0)$ is small enough, total reflection of rays from $(-1,\, 0)$ occurs at the interface near $(0,\, m(0))$. At this point the same argument applies. If having a discontinuity in c bothers you, smear it a little; this will not destroy the effect.
2. One could argue that maybe there is some "extra" fact that prevents extrema which are not local minima from being rays. However, as long as we are dealing with the wave equation (7.3), the key equation is (7.13), and Fermat's principle follows from it. And (7.13) allows rays that do not minimize T in (7.14), not even locally. Of course, if one changes the physical setting, "anything" can happen. For example, we mentioned in §7.1.1 that the Eikonal equation can be used to (approximately) describe the propagation of a forest fire. In this case, once a front goes through some region, another front cannot go by through the same region. In particular: the notion of "reflection" used in the example above ceases to make sense; the only ray allowed would be the one going directly from $(-1,\, 0)$ to $(1,\, 0)$. Maybe in this case the rays must be local minima (or flat out minima) of T; **I do not know.**
3. **Proof of the mirror reflection law.** For any ray reflecting of a mirror, one can set a local coordinate frame where, modulo scale factors, the problem looks as in the example: The ray goes from $(-1,\, 0)$ to $(0,\, m(0))$ to $(1,\, 0)$. Since the ray must be an extrema for T, it follows that $m'(0) = 0$ — i.e.: the mirror reflection law.

7.1.5 The 2-D case; trapping and Snell's law

Here we consider the 2-D case, where $c = c(x,\, y)$. Let $n = 1/c$ by the **index of refraction,** then the equations for the rays (i.e.: (7.13)) are
$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}x}{\mathrm{d}s}\right) = n_x \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}y}{\mathrm{d}s}\right) = n_y, \quad (7.22)$$
where s is the arclength. Introduce the ray angle θ (angle that the tangent to the ray makes with the positive x-axis) so that $\frac{\mathrm{d}x}{\mathrm{d}s} = \cos\theta$ and $\frac{\mathrm{d}y}{\mathrm{d}s} = \sin\theta$. Then
$$n\,\frac{\mathrm{d}\theta}{\mathrm{d}s} = (\nabla n)\cdot\hat{n}, \quad (7.23)$$

where $\hat{\boldsymbol{n}} = (-\sin\boldsymbol{\theta}, \cos\boldsymbol{\theta})$ is the unit normal to the ray (see figure 7.5).

Proof. Note that, with $\hat{\boldsymbol{t}} = (\cos\boldsymbol{\theta}, \sin\boldsymbol{\theta})$, (7.22) is equivalent to:

$$(\boldsymbol{\nabla} n)\cdot\hat{\boldsymbol{t}}\,\cos\boldsymbol{\theta} - \boldsymbol{n}\sin\boldsymbol{\theta}\,\tfrac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}\boldsymbol{s}} = \boldsymbol{n_x} \quad \text{and} \quad (\boldsymbol{\nabla} n)\cdot\hat{\boldsymbol{t}}\,\sin\boldsymbol{\theta} + \boldsymbol{n}\cos\boldsymbol{\theta}\,\tfrac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}\boldsymbol{s}} = \boldsymbol{n_y}.$$

Multiplying the first equation here by $\cos\boldsymbol{\theta}$, the second by $\sin\boldsymbol{\theta}$, and adding, gives an identity. On the other hand, multiplying by $-\sin\boldsymbol{\theta}$, $\cos\boldsymbol{\theta}$, and adding, gives (7.23). Hence (7.22) and (7.23) are equivalent. **QED**

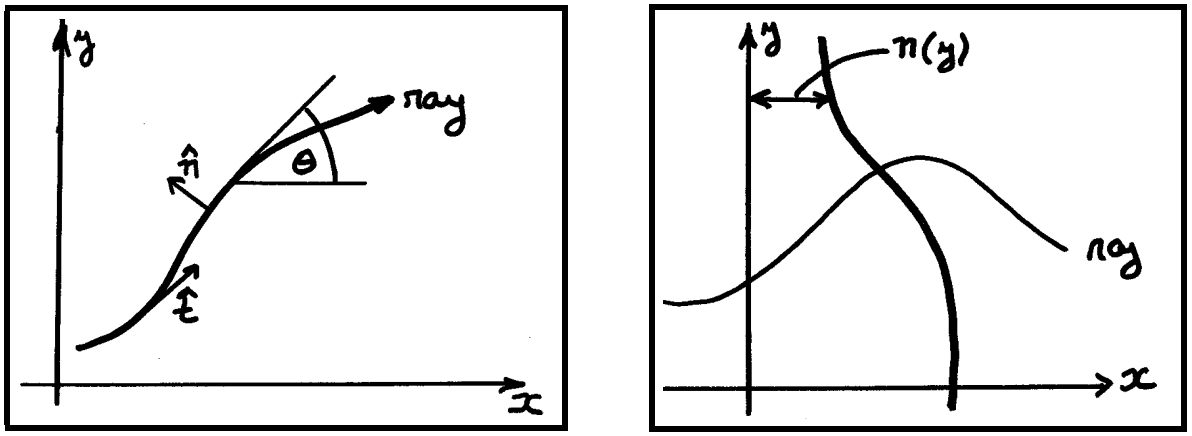


Figure 7.5: Ray angle and ray trapping. **Left:** The ray angle $\boldsymbol{\theta}$, as well as the unit tangent to the ray $\hat{\boldsymbol{t}} = (\cos\boldsymbol{\theta}, \sin\boldsymbol{\theta})$ and the unit normal $\hat{\boldsymbol{n}} = (-\sin\boldsymbol{\theta}, \cos\boldsymbol{\theta})$. **Right:** Trapping by a wave-speed channel.

Consider now a layered media, where $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{y})$.

Then, since $\boldsymbol{0} = \boldsymbol{n_x} = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{s}}(\boldsymbol{n}\,\cos\boldsymbol{\theta})$, we have

$$\boldsymbol{n(y)}\cos\boldsymbol{\theta} = \textbf{constant}, \tag{7.24}$$

which is **Snell's law.** Thus

$$\left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}\boldsymbol{s}}\right)^{\boldsymbol{2}} = \sin^{\boldsymbol{2}}\boldsymbol{\theta} = \boldsymbol{1} - \boldsymbol{\kappa}^{\boldsymbol{2}}\,\boldsymbol{c}^{\boldsymbol{2}}(\boldsymbol{y}), \tag{7.25}$$

where $\boldsymbol{\kappa}$ is the constant in (7.24). Hence **near local minima of c rays can get trapped, with y oscillating periodically between the solutions to $\kappa^2 c^2(y) = 1$.** †

† Note that κ depends on the ray. Hence not all rays get trapped (if κ is small enough the ray can escape).

7.1.6 Solution of the transport equation using rays (constant c)

Finally, here we derive the solution to the transport equation (second in (7.5)) along rays, for the case when $\boldsymbol{c}$ is a constant. This is done by producing an ode for the Hessian of the phase, which can then be solved to obtain a formula analogous to (4.41).

WLOG, assume $\boldsymbol{c} = \boldsymbol{1}$. Then, along the rays, the transport equation in (7.5) takes the form

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{s}} + \frac{\boldsymbol{1}}{\boldsymbol{2}}\,\boldsymbol{\Delta\phi}\,\boldsymbol{A} = \boldsymbol{0}. \tag{7.26}$$

Next, apply to the Eikonal equation $(\boldsymbol{\nabla\phi}) = \boldsymbol{1}$ the operator $\boldsymbol{\partial^2_{x_p x_q}}$, all $\boldsymbol{p}, \boldsymbol{q}$. This yields

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{s}}\boldsymbol{\Phi} + \boldsymbol{\Phi}^{\boldsymbol{2}} = \boldsymbol{0}, \tag{7.27}$$

where $\boldsymbol{\Phi}$ is the matrix with entries $\boldsymbol{\phi_{x_p x_q}}$. Note that (7.10) can be written in the form $\boldsymbol{\Phi}\,\boldsymbol{\nabla\phi} = \boldsymbol{0}$. Thus

one eigenvalue of Φ is zero, the other(s) are the principal curvatures of the wave-fronts. (7.28)

The solution to (7.27) is $\boldsymbol{\Phi} = (\boldsymbol{1} + \boldsymbol{s}\,\boldsymbol{\Phi_0})^{-\boldsymbol{1}}\,\boldsymbol{\Phi_0}$, where $\boldsymbol{\Phi_0}$ is the value of $\boldsymbol{\Phi}$ on the ray at the initial

wave-front. † Then, since $\boldsymbol{\Delta\,\phi = \mathrm{Tr}(\Phi)}$, the solution to (7.26) is (same steps as those used to get (4.41))

$$A = \left(\prod_j \frac{1}{\sqrt{1+\lambda_j\, s}}\right) A_0, \tag{7.29}$$

where the λ_j are the non-zero eigenvalues of Φ_0, and A_0 is the amplitude on the ray at the initial wave-front.

† **Can we obtain Φ_0 from just knowledge of the initial wave-front? Yes: all this information can be extracted from the Eikonal equation.** For example: we know that ϕ is constant on the wave-front; we also know the derivative normal to the wave-front of ϕ; we also know the derivative normal to the wave-front of $\nabla\phi$ (i.e.: (7.10)); etc. The gory details are left to the reader.

Remark 7.7 What happens when c is not constant. In this case we can still write ode that determine $\boldsymbol{A}$ along the rays. But, of course, we cannot solve them explicitly (as above). Nevertheless, for a smooth $\boldsymbol{c}$, the solutions above are indicative of the ray dynamics in any small region (where $\boldsymbol{c}$ is nearly constant). The general equations are:

$$\frac{\mathbf{d}}{\mathbf{d}s}(c\,A) + \frac{1}{2}\,(\Delta\phi)\,c^2\,A = 0 \quad \text{and} \quad \frac{\mathbf{d}}{\mathbf{d}s}\Phi + c\,\Phi^2 = \mathcal{N}.$$

Here Φ is as above and $\mathcal{N}$ is the matrix with entries $\mathcal{N}_{pq} = n_{x_p x_q} + c\, n_{x_p}\, n_{x_q}$, where $n = 1/c$ is the index of refraction. ♣

7.2 Singular rays (an example)

The Eikonal equation can "spontaneously" develop singularities, where the approximation $\boldsymbol{\lambda \ll L}$ no longer applies ‡ — hence the Eikonal equation may no longer be valid. Here we consider a particular example of this, involving a singular ray, and show how to resolve the singularity.

‡ Here λ is the wave-length and L is a measure of the other lengths in the problem; e.g.: curvature of the wave-fronts, etc. See below (7.3).

For simplicity let us consider a 2-D situation with a constant wave-speed, which thus can be **normalized to $c \equiv 1$.** Then the equation is

$$u_{tt} = \Delta\, u = u_{xx} + u_{yy}. \tag{7.30}$$

The Geometrical Optics approximation is then:

$$u \sim A\, e^{i\,\omega\,(t-\phi)},\ \omega \gg 1, \quad \text{with} \quad (\nabla\phi)^2 = 1 \text{ and } \mathrm{div}(A^2\,\nabla\phi) = 0. \tag{7.31}$$

The second equation can also be written in the form $((\nabla\phi)\cdot\nabla)\,A^2 + (\Delta\phi)\,A^2 = 0$, which is useful for the ray formulation below.

The **ray formulation** is where $\vec{x} = (x,\, y)$.

$$\frac{\mathbf{d}}{\mathbf{d}s}\vec{x} = \nabla\phi, \quad \frac{\mathbf{d}}{\mathbf{d}s}\nabla\phi = 0, \quad \frac{\mathbf{d}}{\mathbf{d}s}\phi = 1, \quad \text{and} \quad \frac{\mathbf{d}}{\mathbf{d}s}A + \frac{1}{2}\,\Delta\phi\,A = 0, \tag{7.32}$$

Remark 7.8 In principle we can solve (7.31) by solving the ode system (7.32) for every ray starting at a wave-front, say $\phi = 0$. † This yields $\vec{x} = \vec{x}(s,\,\zeta)$, $\phi = \phi(s,\,\zeta)$, and $A = A(s,\,\zeta)$ — with ζ = parameter on the initial wave-front. Then invert $\vec{x} = \vec{x}(s,\,\zeta)$ to obtain (ζ, s) as a function of $\vec{x}$. Finally: note the role of $\Delta\phi$ (wave-front *curvature*) in driving the amplitude evolution along rays. [† This requires solving for $\Delta\phi$ as well (see §7.1.6).] ♣

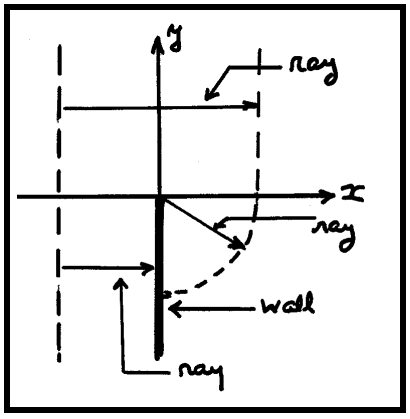


Figure 7.6: Wave-fronts hit an absorbing half wall

Consider a situation with a (very thin) absorbing wall along the negative y-axis, with a plane wave arriving from the left, parallel to the y-axis. As the wave-fronts (dashed lines) arrive to the wall, the lower portion is absorbed and only the top, $\boldsymbol{y > 0}$, goes through. However, the wave equation (7.30) does not admit an abrupt cut-off in the solution, so that the wave-fronts must continue below $\boldsymbol{y < 0}$ for $\boldsymbol{x > 0}$. The result is that they bend **(diffraction),** as in the picture. Here we construct an Eikonal-based solution for this situation. The **singular ray** (positive x-axis) requires special treatment.

7.2.1 Example description: absorbing half wall parallel to the fronts (diffraction)

The example we will consider is described in Figure 7.6.

Remark 7.9 What does $\omega \gg 1$ mean here? The basis for writing $\omega \gg 1$ in (7.31) is the selection of non-dimensional variables for (7.30) based on space, L, and time, $T = L/c$, scales such that L is much larger than the wavelength, ℓ. Then ω is given by the ratio L/ℓ. Furthermore, L is supposed to characterize the distances over which the waves depart from plane and uniform, and it is generally determined by some feature of the problem itself: a scale imposed by either the initial or the boundary conditions, or by the scale over which the wave speed c varies, or some other similar factor.

However, in this example there are *only two dimensional constants:* the wave speed c, and the wavelength ℓ. So, *what determines L?* The answer is: what region we look at. Specifically, we *want to find out what the solution looks like on the far-field,* at distances L from the origin (where the wall ends). The point is that, in this limit, the solution is "simpler", and can be approximated by theories such as Geometrical Optics — as we will show here.

Note, though, that there is a region, $\mathcal{U}$, of size $O(\ell)$ near the origin where the full wave equation is needed, and no simplifications are possible. The *existence of $\mathcal{U}$ creates a puzzle:* as shown in Figure 7.6 (see §7.2.2 as well), the Geometrical Optics' rays in the 4^{th} quadrant emanate from this region. Hence, without determining the solution in $\mathcal{U}$, how can we determine it in the 4^{th} quadrant? The reason is geometry: in the far-field $\mathcal{U}$ looks like a point, so no detailed information of the solution there is needed — see §7.2.3, in particular the "Matching with quadrant #4" process (7.48–7.50). ♣

7.2.2 Solution to the Eikonal and Transport equations

We propose the following **solution for the Eikonal equation**

$$\boldsymbol{\phi = x} \quad \text{for} \quad \boldsymbol{y \geq 0} \quad \text{or} \quad \boldsymbol{y < 0, x < 0}, \qquad \text{and} \qquad \boldsymbol{\phi = r} \quad \text{for} \quad \boldsymbol{y \leq 0, x > 0}, \tag{7.33}$$

where $\boldsymbol{r = \sqrt{x^2 + y^2}}$.

Remark 7.10 Note that *this defines $\boldsymbol{\phi}$ in the domain of interest (plane minus negative y-axis), with both $\boldsymbol{\phi}$ and $\boldsymbol{\nabla\phi}$ continuous, hence it is a perfectly "reasonable" solution to* $\boldsymbol{(\nabla\phi)^2 = 1}$. But the curvature is discontinuous across the positive x-axis (the *singular ray*), and this causes trouble with the transport equation (see below). In addition, $\nabla\phi$ does not have a well defined limit at $r \to 0$, which means that *the*

rays in the region $(x > 0,\, y < 0)$ *are not connected to the incident fronts* — since they all start at the origin, forming an *expansion fan of rays*. Below we explain how to get around these problems. ♣

Solution to the transport equation. Everywhere but in the 4^{th} quadrant, the rays follow from $\frac{\mathrm{d}x}{\mathrm{d}s} = 1$ and $\frac{\mathrm{d}y}{\mathrm{d}s} = 0$. Hence they are given by $\boldsymbol{x = s}$ and $\boldsymbol{y =}$ constant. Thus the equation for A is $\frac{\mathrm{d}A}{\mathrm{d}x} = 0$. Since **the amplitude is uniform for the incident wave,** we have

$$\boldsymbol{A \equiv 1} \tag{7.34}$$

everywhere but in the $4^{\mathbf{th}}$ quadrant.

In the 4^{th} quadrant the rays are the lines $\boldsymbol{r = s > 0}$ and $\boldsymbol{\theta =}$ const., where $-\pi/2 < \theta < 0$ is the polar angle. Furthermore

$$\frac{\mathbf{d}}{\mathbf{d}r}\boldsymbol{A} + \frac{\mathbf{1}}{\mathbf{2}\,\boldsymbol{r}}\,\boldsymbol{A} = \mathbf{0}, \tag{7.35}$$

where $\boldsymbol{\Delta\phi = 1/r}$ is the curvature. ‡

‡ Generally the curvature is given by $\boldsymbol{\kappa} = \mathrm{div}(\boldsymbol{c}\,\boldsymbol{\nabla\phi})$, because $\boldsymbol{c}\,\boldsymbol{\nabla\phi}$ is the unit normal to the wave front. (7.36)

It follows that

$$\boldsymbol{A} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{r}}}\boldsymbol{\alpha}(\boldsymbol{\omega},\, \boldsymbol{\theta}), \tag{7.37}$$

for some function $\boldsymbol{\alpha}$ that we need to determine, and **does not follow** from the theory as so far presented. † However, the situation is worse than merely not knowing what $\boldsymbol{\alpha}$ is:

The values of $\boldsymbol{A}$ cannot match across the singular ray $\boldsymbol{x > 0,\, y = 0}$. (7.38)

This is because (i) for $y = +0$, $A \neq 0$ is constant; while (ii) for $y = -0$ it is proportional to $1/\sqrt{r}$. There is also the issue of what happens as $r \to 0$; clearly the wave amplitude should not be infinity there.

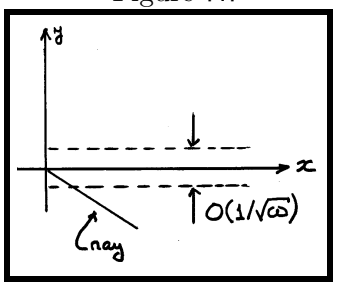
† We expect $\boldsymbol{\alpha \to 0}$ as $\boldsymbol{\omega \uparrow \infty}$. The amount of energy "turning" the corner should vanish as ray theory becomes more accurate.

Issues to resolve:

- **i1.** Determine $\boldsymbol{\alpha}$.
- **i2.** Fix the inconsistency in $\boldsymbol{A}$ across the singular ray.
- **i3.** Resolve the infinity at $\boldsymbol{r = 0}$.

These issues will be resolved in §7.2.3 via a *singular ray expansion*, valid in a $\boldsymbol{O(1/\sqrt{\omega})}$ wide band along the singular ray. Note that then all the rays in the 4^{th} quadrant start inside the band, and are initialized there.

Figure 7.7



Remark 7.11 A final point: the resolution of the difficulty in (7.38) leads to α having a singularity as $\theta \uparrow 0$. A singularity that is not actually reached by the solution because it occurs within the region covered by the singular expansion in §7.2.3, which is not singular there. ♣

7.2.3 Singular ray expansion

Motivation. Near the singular ray the Eikonal wave-fronts are almost plane, but not quite. In fact, if we look at some generic wave-front (given by $\boldsymbol{\phi = t > 0}$), we can write:

$$x = t \;\text{ for } y > 0 \quad \text{and} \quad x = t - \frac{1}{2\,t}\,y^2 + O(y^4/t^3) \;\text{ for }\; 0 < -y \ll 1. \tag{7.39}$$

Thus, in order to capture the curvature change across the singular ray, $O(y^2)$ corrections should be as important as the leading order $x \approx t$. To reinforce this point, consider the solution provided by (7.33) and (7.37) in the 4^{th} quadrant and expand $\boldsymbol{r = x + \frac{1}{2\,x}\,y^2 + O(y^4/x^3)}$ for $\boldsymbol{0 < -y \ll 1}$.

$$\boldsymbol{u} \sim \frac{\boldsymbol{\alpha}}{\sqrt{\boldsymbol{r}}}\,\boldsymbol{e}^{\boldsymbol{i\,\omega\,(t-r)}}, \tag{7.40}$$

This yields
$$u \sim \frac{\alpha}{\sqrt{x}}\, e^{-i\,\omega\, y^2/(2\,x)}\, e^{i\,\omega\,(t-x)} = \mathcal{A}(\sqrt{\omega}\, y,\, x,\, y)\, e^{i\,\omega\,(t-x)}, \quad (7.41)$$
for some $\mathcal{A}$, where we have used that $\theta = \theta(x,\, y)$ to incorporate α into $\mathcal{A}$. This form is the basis for what follows next.

Expansion. We propose **the following expansion for the solution to (7.30) near the singular ray:**
$$u \sim B(z,\, x,\, y)\, e^{i\,\omega\,(t-x)}, \quad \text{where} \quad z = \sqrt{\omega}\, y \qquad \text{(see remark 7.16)}. \quad (7.42)$$

Remark 7.12 A generic singular ray is a ray that separates two regions such that (i) ϕ is smooth in each region; (ii) ϕ and $\nabla\phi$ are continuous across the ray; but (iii) the wave-fronts curvature changes across the ray. In this generic case, the analog of (7.42) replaces y by the coordinate transversal to the ray, and x by the coordinate along the ray. Note also that the form in (7.41) generalizes to each side of the ray if one places $x = 0$ at the center of curvature of the rays on the corresponding side. ♣

Substituting (7.42) into (7.30), the leading order yields
$$-2\,i\,B_x + B_{zz} = 0, \quad (7.43)$$
where y appears as a parameter [13] only. Now we use this equation/approximation to cover the gap between the expansion for $y > 0$, (7.34), and the expansion for $y < 0$, (7.37). To do this we solve an **initial value problem** for it, with data given at $x = 0$ (where the solution is determined by the incoming wave from the left). That is
$$B \equiv 1 \;\text{ for }\; z > 0 \quad \text{and} \quad B \equiv 0 \;\text{ for }\; z < 0. \quad (7.44)$$

Why these initial values? For any $z \neq 0$, let $x \downarrow 0$. Then, from (7.42) and Figure 7.6, (7.44) follows. In particular: on the $(x,\, z)$-scale the unresolved region near the origin (see remark 7.9) is still a point — since $\sqrt{\omega} \ll \omega$ for ω large.

Remark 7.13 *Important:* there is the implicit assumption here that what happens at the "point" $z = x = 0$ has no influence on the solution on the scale where (7.43) applies. This seems reasonable, but we cannot justify it here. However, the Fourier analysis in §7.3 confirms this assumption. ♣

Side note: with asymptotic approximations, quite often "reasonable" assumptions that cannot be justified are made. In the absence of a rigorous mathematical justification [which, even when they exist, can be cumbersome], the best strategy is to show consistency. Show that all the pieces click together and produce a solution without obvious flaws. If possible, checking that the higher order corrections actually stay small is also helpful.

Because (7.43–7.44) is invariant under $z \to \nu\, z$ and $x \to \nu^2\, x$ ($\nu > 0$ an arbitrary constant), the problem reduces to the ode
$$B = B(\zeta), \quad \zeta = \frac{z}{\sqrt{x}}, \quad \text{and} \quad i\,\zeta\,\frac{\mathrm{d}B}{\mathrm{d}\zeta} + \frac{\mathrm{d}^2 B}{\mathrm{d}\zeta^2} = 0, \quad (7.45)$$
with $B(\infty) = 1$ and $B(-\infty) = 0$ (to satisfy the initial conditions).
The ode can be integrated once, to $\frac{\mathrm{d}B}{\mathrm{d}\zeta} = c_i\, e^{-i\,\zeta^2/2}$,
where c_i is a constant. Thus
$$B = 1 - c_i \int_\zeta^\infty e^{-i\,s^2/2}\, \mathrm{d}s, \quad (7.46)$$
where we have already implemented $B(\infty) = 1$.
The other condition, $B(-\infty) = 0$, yields
$$c_i = \left(\int_{-\infty}^{\infty} e^{-i\,s^2/2}\, \mathrm{d}s\right)^{-1} = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,\pi/4}. \quad (7.47)$$

Remark 7.14 Integrals such as the one in (7.46) are conditionally convergent only. However, they can be made to converge fast by appropriately rotating the path of integration into the complex plane. ♣

[13] This is the same equation as in the Paraxial approximation, §8.1. Not surprising: same scaling (but different motivation).

Matching with quadrant #1. Let $\sqrt{1/\omega} \ll y \ll 1$ ($\zeta \gg 1$ and $y > 0$ small), with $0 < x = O(1)$. Then Geometrical Optics gives $u \sim e^{i\,\omega\,(t-x)}$ — see (7.33–7.34), with matches with what (7.42) produces.

Matching with quadrant #4. Let $\sqrt{1/\omega} \ll -y \ll 1$ ($-\zeta \gg 1$ and $y < 0$ small), with $0 < x = O(1)$. Then Geometrical Optics gives (7.40–7.41). On the other hand, using that ζ is large and negative,

$$B = c_i \int_{-\infty}^{\zeta} e^{-i\,s^2/2}\,\mathrm{d}s = i\,\frac{c_i}{\zeta}\,e^{-i\,\zeta^2/2} + i\,c_i \int_{-\infty}^{\zeta} e^{-i\,s^2/2}\,\frac{\mathrm{d}s}{s^2} = -\frac{1}{\sqrt{2\,\pi}\,\zeta}\,e^{-i\,\pi/4}\,e^{-i\,\zeta^2/2} + O(\zeta^{-3}) \tag{7.48}$$

where (i) The first equality follows from the definition of c_i; (ii) The second equality follows from integration by parts using $e^{-i\,s^2/2} = -\left(e^{-i\,s^2/2}\right)' \frac{1}{i\,s}$. (iii) The third equality follows by substituting the value for c_i. Hence

$$B\,e^{i\,\omega\,(t-x)} \sim -\frac{1}{\sqrt{2\,\pi}\,\zeta}\,e^{-i\,\pi/4}\,e^{-i\,\zeta^2/2}\,e^{i\,\omega\,(t-x)} = -\frac{e^{-i\,\pi/4}}{\sqrt{2\,\pi\,\omega\,x}}\,\frac{x}{y}\,e^{-i\,\omega\,y^2/(2\,x)}\,e^{i\,\omega\,(t-x)}. \tag{7.49}$$

This agrees with (7.41) provided that we take †

$$\boldsymbol{\alpha = -\frac{e^{-i\,\pi/4}}{\sqrt{2\,\pi\,\omega}}\,\frac{x}{y}}. \tag{7.50}$$

As anticipated in remark 7.11, α is singular as $y \uparrow 0$ for $x > 0$ fixed.
However, (7.50) is valid for $\sqrt{1/\omega} \ll -y$ only. For y smaller one should use the non-singular (7.46).
† Note that $x/y = \mathrm{cotan}(\theta)$. Hence α has the form in (7.37).

Remark 7.15 Above we use the expansion in (7.42–7.44) to fill in the singular ray region (where Geometrical Optics fails), as well as connect/match the solutions provided by Geometrical Optical for the regions $y > 0$ and $y < 0$. In this process we determine what the function α, left open by Geometrical Optics, should be. However, the Geometrical Optics rays in the region $y < 0$ start at the origin and move away from the line $y = 0$. Given this, *question: how can a matching done near $y = 0$ determine the solution in the whole whole 4-th quadrant?* The reason is that most of the rays in the 4-th quadrant enter it by traversing the $1/\sqrt{\omega}$ wide band near $y = 0$ where the singular ray expansion is valid. The only questionable rays are those that go almost glancing to the wall, which enter the quadrant too close (within $O(1/\omega)$) to the origin, where the validity of (7.42–7.44) is questionable. Even though the result above agrees with the waves vanishing near the wall, and introduces no inconsistencies, its validity is dubious. Further checks are needed (which I have not done). ♣

Remark 7.16 Geometrical significance of the scaling in (7.42). In standard Geometrical Optics (as well as Modulation Theory in other sections of these notes), we consider locally plane waves, with variations on scales slow relative to the wave period and wave-length. In the case of this section, the expansion includes a scale, $\omega\,\phi$ (which "detects" the wave-length $1/\omega$), and a scale for the "slow" changes, $\vec{x}$. This leads to a situation where *the curvature of the wave-fronts plays no role at leading order* (i.e.: Eikonal equation, or Dispersion Relation for the Modulation Theory case). Curvature appears at the next order (the transport equation) but it is a one-way interaction only; when something begins to "go wrong" with the wave amplitude, there is no path by which the phase to correct and adapt. *In order for curvature effects to affect the phase, a scale that "detects" curvature must be added.* Since curvature is basically a second derivative of the phase in the directions parallel to the wave-front, this means that a scale $1/\sqrt{\omega}$ along the wave fronts is needed. This is precisely what (7.42) does. However, *the same idea works for many other contexts where wave-front curvature effects matter* (for some examples see §7.4). ♣

7.3 Singular ray example by Fourier Transforms

Here we consider the following problem: Solve $\boldsymbol{\Delta\,\psi + \omega^2\,\psi = 0}$, (7.51)
in the region $\boldsymbol{x > 0}$ and $\boldsymbol{-\infty < y < \infty}$, with $\boldsymbol{\psi(0,\, y) = 1}$ for $\boldsymbol{y > 0}$
and $\boldsymbol{\psi(0,\, y) = 0}$ for $\boldsymbol{y < 0}$, and the *radiation condition:* $\boldsymbol{\psi\, e^{i\,\omega\, t}}$ solves the wave equation with no incoming waves from $\boldsymbol{x = \infty}$. We will also assume $\boldsymbol{\omega \gg 1}$.

Remark 7.17 With $\boldsymbol{u = \psi\, e^{i\,\omega\, t}}$, this is the same problem that was solved (approximately) in §7.2, using Geometrical Optics. Here we construct the exact solution using integral methods, and then use the stationary phase method to double check the approach introduced in §7.2. † ♣

† One may ask: why is the approach in §7.2 needed if integral methods can be used to write the exact solution? The reason is that the integral methods only work for special geometries.

Remark 7.18 The variable change $\tilde{x} = a\,x$, $\tilde{y} = a\,y$, and $\tilde{\omega} = \omega/a$, leaves (7.51) unchanged. Hence, *what does* $\omega \gg 1$ *mean?* The answer is given in remark 7.9; we are interested in the behavior of the solution far from the special point at the origin, where "far" means relative to the wave-length. ♣

7.3.1 Solution to the problem

The solution to the problem is

$$\psi = -\frac{\mathbf{1}}{\mathbf{2\,\pi\, i}} \int_{\mathbf{\Gamma}} e^{-i\,\omega\,(k\,y + \ell\, x)}\, \frac{\mathbf{d}k}{k}, \tag{7.52}$$

where $\boldsymbol{\ell = \sqrt{1 - k^2}}$ and $\boldsymbol{\Gamma}$ (see Figure 7.8) are defined as follows:

- **a.** The branch cuts for ℓ are: $k > 1$ and $k < -1$ on the real line, with $\ell > 0$ for $-1 < k < 1$. The signs for $i\,\ell$ along the branches are as indicated in Figure 7.8 (the other side of each branch yields the opposite sign).
- **b.** The contour Γ goes from $k = -\infty$ to $k = \infty$. It tracks the lower side of the branch $k < -1$ on the left half-plane, and the upper side of the branch $k > 1$ on the right half-plane.
  Important: $k = 0$ is *below* Γ. (7.53)
- **c.** (7.52) is absolutely convergent for $x > 0$ (exponential decay of $e^{-i\,\omega\,\ell\, x}$ as $|k| \to \infty$).

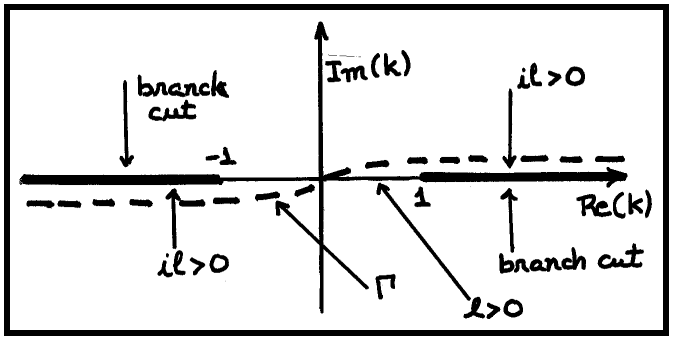


Figure 7.8: Contour Γ
Dashed line: integration contour Γ used in (7.52). The branch cuts for ℓ, and the signs for ℓ and $i\,\ell$ are also indicated.
Important: $k = 0$ is *below* Γ.

Proof that (7.52) is the solution to (7.51).

1. By construction ψ solves the equation, since $k^2 + \ell^2 = 1$.

2. By construction ψ satisfies the radiation condition: (i) The wave components, i.e.: ℓ real, have $\ell > 0$ — thus they correspond to waves moving right. (ii) The non-wave components are evanescent (vanish exponentially as $x \to \infty$).

3. The limit $x \downarrow 0$ of (7.52) is

$$\psi = -\frac{\mathbf{1}}{\mathbf{2\,\pi\,i}} \int_{-\infty}^{\infty} e^{-i\,\omega\,k\,y}\, \frac{\mathbf{d}k}{\boldsymbol{k + i\,0}}, \tag{7.54}$$

where the denominator $k + i\,0$ arises because the path goes above $k = 0$ — see (7.53). Note that **this integral is conditionally convergent only.**

In this expression, if $y < 0$ we can move the path of integration up in the complex plane (integrate from $-\infty + i\,a$ to $\infty + i\,a$, with $a > 0$). This produces a factor $e^{\omega\,a\,y}$ outside the integral. Hence, letting $a \to \infty$ we see that it must be $\boldsymbol{\psi(0,\,y) = 0}$ for $\boldsymbol{y < 0}$. On the other hand, if $y > 0$, we move the path down in the complex plane (after picking up a residue at $k = 0$). The same process then yields $\boldsymbol{\psi(0,\,y) = 1}$ for $\boldsymbol{y > 0}$. **QED**

7.3.2 Alternative form for the solution

An alternative form for (7.52) is

$$\begin{aligned} \psi &= -\frac{\mathbf{1}}{\mathbf{2\,\pi\,i}} \int_{-1}^{1} e^{-i\,\omega\,(k\,y+\ell\,x)}\, \frac{\mathbf{d}k}{\boldsymbol{k + i\,0}} + \frac{\mathbf{1}}{\boldsymbol{\pi}} \int_{1}^{\infty} \frac{\mathbf{sin}(\boldsymbol{\omega\,k\,y})}{\boldsymbol{k}}\, e^{-\omega\,\mu\,x}\, \mathbf{d}k \\ &= \frac{\mathbf{1}}{\mathbf{2}}\, e^{-i\,\omega\,x} + \frac{\mathbf{1}}{\boldsymbol{\pi}} \int_{0}^{1} \frac{\mathbf{sin}(\boldsymbol{\omega\,k\,y})}{\boldsymbol{k}}\, e^{-i\,\omega\,\ell\,x}\, \mathbf{d}k + \frac{\mathbf{1}}{\boldsymbol{\pi}} \int_{1}^{\infty} \frac{\mathbf{sin}(\boldsymbol{\omega\,k\,y})}{\boldsymbol{k}}\, e^{-\omega\,\mu\,x}\, \mathbf{d}k, \end{aligned} \tag{7.55}$$

where $\boldsymbol{\mu = \sqrt{k^2 - 1}}$. **Details:** Equality #1: denominator "$k + i\,0$" in first integral because path goes above $k = 0$; see (7.53). Equality #2: pull out half-residue at $k = 0$ from first integral.

Note that $\boldsymbol{\psi}$ **is not real valued,** even though the equation and the boundary conditions at $x = 0$ are real. The reason is that *the radiation condition breaks the symmetry:* A wave $e^{i\,\omega\,(t - k\,y - \ell\,x)}$, with $\ell = \sqrt{1 - k^2} > 0$ $(-1 < k < 1)$, is allowed; but its complex conjugate is not allowed.

7.3.3 Far field asymptotic behavior

We now consider the $\omega \gg 1$ behavior of (7.52) using the steepest descent method. The stationary points satisfy $\frac{\mathbf{d}}{\mathbf{d}k}(\boldsymbol{k\,y + \ell\,x}) = \mathbf{0}$, that is:

$$\boldsymbol{y - \frac{k}{\ell}\,x = 0}. \tag{7.56}$$

This has two solutions. † One of them is

$$\boldsymbol{k_s = \frac{y}{r}} \quad \text{and} \quad \boldsymbol{\ell_s = \frac{x}{r}}, \quad \text{where} \quad \boldsymbol{r = \sqrt{x^2 + y^2}}. \tag{7.57}$$

The other, $\boldsymbol{k_s = -y/r}$ and $\boldsymbol{\ell_s = -x/r}$, is in the other branch of ℓ.

† From the equation $k = a\,y$ and $\ell = a\,x$, for some constant a. Then $1 = k^2 + \ell^2 = a^2\,r^2$.

First assume $\boldsymbol{y \neq 0}$. Then adjust the contour Γ so that it goes through the stationary point (picking up a residue at $k = 0$ when $y > 0$). Further, adjust Γ so that it crosses the real axis at a $\pi/4$ angle. ‡

‡Since the integrand vanishes exponentially along Γ as $|k| \to \infty$, we can do this (changing the path only in the region

$-1 < k < 1$. The angle adjustment then picks the direction along which the integrand decays the fastest as the path moves away from the stationary point (steepest descent path).

Then we have, for $\boldsymbol{y < 0}$ (and $\boldsymbol{x > 0}$)

$$\begin{aligned} \psi &\sim -\frac{r\, e^{-i\omega r}}{2\pi i y} \int_{k-k_s = e^{i\pi/4} s} \exp\left(i\omega \frac{r^3 (k-k_s)^2}{2x^2} \right) \mathrm{d}k \\ &= -\frac{r\, e^{-i\omega r + i\pi/4}}{2\pi i y} \int_{-\infty}^{\infty} \exp\left(-\omega \frac{r^3 s^2}{2x^2} \right) \mathrm{d}s \\ &= -\frac{\boldsymbol{x}}{\boldsymbol{y}\sqrt{\boldsymbol{2\pi\omega r}}}\, \boldsymbol{e^{-i\omega r - i\pi/4}} \quad (\text{for } \boldsymbol{y<0} \text{ and } \boldsymbol{x>0}). \end{aligned} \tag{7.58}$$

This is exactly that same as (7.40) and (7.50).

On the other hand, for $\boldsymbol{y > 0}$ (and $\boldsymbol{x > 0}$) the same contribution as above arises from the stationary point, † but the residue is larger, hence

$$\boldsymbol{\psi \sim e^{-i\omega x}} \quad (\text{for } \boldsymbol{y>0} \text{ and } \boldsymbol{x>0}). \tag{7.59}$$

This is **the same as (7.34).**

† The circular wave-fronts shown in the 4-th quadrant in Figure 7.6 are actually there in the other quadrants as well, but they are higher order than the main wave. They are produced by the edge of the wall at the origin.

The **stationary phase approach above does not work for $y \approx 0$** (i.e.: near the singular ray). The reason is that in this case k_s is close to the pole at $k = 0$ of the integrand in (7.52). However, the main contribution to the integral still comes from the neighborhood of the stationary point. Thus we expand everything near $k = 0$.

$$\boldsymbol{\ell = 1 - \frac{1}{2} k^2 + O(k^4)} \quad \text{and} \quad \boldsymbol{k y + \ell x = x + k y - \frac{1}{2} k^2 x + O(k^4 x)}$$

Now we use this in (7.52) to approximate the integrand in the portion near $k = 0$, neglecting the rest. Furthermore, we make the change of variables $k = s/\sqrt{x}$, to obtain

$$\psi \sim -\frac{e^{-i\omega x}}{2\pi i} \int_{-\infty}^{\infty} \exp\left(-i\omega \left(s\eta - \frac{1}{2} s^2 + O(\omega s^4/x) \right) \right) \frac{\mathrm{d}s}{s} \tag{7.60}$$

$$= -\frac{e^{-i\omega x}}{2\pi i} \int_{-\infty}^{\infty} \exp\left(-i \left(s\zeta - \frac{1}{2} s^2 + O(s^4/(\omega x)) \right) \right) \frac{\mathrm{d}s}{s}, \tag{7.61}$$

where (i) $\boldsymbol{\eta = y/\sqrt{x}}$ and $\boldsymbol{\zeta = \sqrt{\omega}\,\eta}$; (ii) in the second line the change of variable $s \to s/\sqrt{\omega}$ was implemented; and (iii) *the path of integration goes above the singularity at $s = 0$.*

Detail: These integrals are conditionally convergent only. But they can be made absolutely convergent by appropriately moving the integration contour in the complex plane.

Thus

$$\boldsymbol{\psi \sim B(\zeta)\, e^{-i\omega x}}, \quad \text{where} \quad \boldsymbol{B = -\frac{1}{2\pi i} \int_{-\infty + i0}^{\infty + i0} \exp\left(-i \left(s\zeta - \frac{1}{2} s^2 \right) \right) \frac{\mathrm{d}s}{s}}. \tag{7.62}$$

Note that **this is not valid for $x \approx 0$** because then we cannot neglect the $O(s^4/(\omega x))$ term in the integral.

Now we show that the **B in (7.62) is the same as the one in (7.45–7.47).**

A. From (7.62) ‡

$$\boldsymbol{B_\zeta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-i \left(s\zeta - \frac{1}{2} s^2 \right) \right) \mathrm{d}s = \frac{1}{2\pi}\, e^{-i\zeta^2/2} \int_{-\infty}^{\infty} e^{is^2/2}\, \mathrm{d}s = \frac{1}{\sqrt{2\pi}}\, e^{-i\zeta^2/2 + i\pi/4}}.$$

Hence, see (7.45–7.47), the two B differ by at most a constant.

‡ For the 2-nd equality change variables $s \to s - \zeta$. For the 3-rd equality rotate the integration path by $\pi/4$.

B. For $\zeta < 0$, let $q = e^{i\,\pi/4}$ and rotate the path in (7.62) to $s = \zeta + q\,\tilde{s}$. Then

$B = -\frac{1}{2\,\pi\,i}\, e^{i\,(\pi/4-\zeta^2/2)} \int_{-\infty}^{\infty} e^{-\tilde{s}^2/2}\, \frac{\mathrm{d}\tilde{s}}{\zeta+q\,\tilde{s}}$.

Now, in the integral here, we take $\zeta \to -\infty$ (no singularities are crossed), to obtain $B(-\infty) = 0$. Using now the result in **A**, we see that the two B are equal. **QED**

7.4 Transversal wave modulation

Geometrical Optics is, locally, a 1-D theory. In a small neighborhood of any point in space the wave-fronts are parallel planes. This is the reason that the equations can be reduced to ode along the rays. When the locally 1-D assumption fails, so does Geometrical Optics, and special approximations are needed near these locations — an example of this was given in §7.2.3.

In this subsection we illustrate how to incorporate transversal (multi-D) effects into locally 1-D theories. We do this with a few simple examples. The main idea is the one explained in remark 7.16.

7.4.1 Example: time dependent singular ray expansion

We consider the wave equation $u_{tt} - u_{xx} - u_{yy} = 0$,
and look for solutions of the form
$$u = U(\theta,\, z,\, x,\, y,\, t), \quad (7.63)$$
where $\theta = \frac{x-t}{\epsilon}$, $z = \frac{y}{\sqrt{\epsilon}}$, and $0 < \epsilon \ll 1$. Then

$\epsilon\, u_t = -U_\theta + \epsilon\, U_t$ and $\epsilon^2\, u_{tt} = U_{\theta\theta} - 2\,\epsilon\, U_{\theta t} + \epsilon^2\, U_{tt}$,

$\epsilon\, u_x = +U_\theta + \epsilon\, U_x$ and $\epsilon^2\, u_{xx} = U_{\theta\theta} + 2\,\epsilon\, U_{\theta x} + \epsilon^2\, U_{xx}$,

$\epsilon^2\, u_{yy} = \epsilon\, U_{zz} + 2\,\epsilon^{1.5}\, U_{zy} + \epsilon^2\, U_{yy}$.

Expanding $U = v + \sqrt{\epsilon}\, v_1 + \ldots$, the leading order yields
$$(v_t + v_x)_\theta + \frac{1}{2}\, v_{zz} = 0. \quad (7.64)$$
In particular, note that $v = B(x,\, y,\, z)\, e^{-i\,\theta}$ yields
$$-2\,i\,B_x + B_{zz} = 0, \quad (7.65)$$
which is the singular ray equation (7.43). Outside the context of singular rays this is known as the **parabolic** or **paraxial approximation** — see §8.

Long, weakly non-linear, plane dispersive waves generally satisfy the Korteweg de Vries (KdV) equation $v_t + v\, v_\theta + v_{\theta\theta\theta} = 0$. The same idea can be used to extend this equation, and add 2-D effects. This leads to the Kadomtsev-Petviashvili (KP) equation
$$(v_t + v\, v_\theta + v_{\theta\theta\theta})_\theta + \frac{1}{2}\, v_{zz} = 0. \quad (7.66)$$
Crossing waves: This last equation exhibits nonlinear crossing waves. Such things are observed in the ocean and other contexts. You can find many pictures in the web (search for "Cross sea waves" or 'Square sea waves"). Two examples are:
www.sciencealert.com/this-is-a-cross-sea-you-do-not-want-to-get-caught-in-one
www.paulingraham.com/waves-crossing.html

7.4.2 Example: transonic flow equation

This is a weakly nonlinear version of (7.64), that arises in (for example) flow past thin wings, nonlinear acoustics, weak shock

reflection at glancing angles, etc. $$(v_t + v_x + v\, v_\theta)_\theta + \frac{1}{2}\, v_{zz} = 0. \quad (7.67)$$

8 Paraxial approximation and Gaussian beams

In this section we introduce the Gaussian beam solutions to the wave equation, valid in the context of the paraxial approximation.

8.1 The Paraxial approximation for the wave equation

We want to describe the propagation of waves, governed by the wave equation, with the following characteristics: (i) The waves are nearly monochromatic, with some typical wave-length λ. (ii) The wave variations in the direction transversal to their propagation direction are characterized by a length scale L, such that $0 = \lambda/L = \epsilon \ll 1$. On the other hand, the variations along the propagation direction are on an scale larger than L.

We now select the x-axis to line up with the direction of propagation, and write the equation in adimensional form by using L as the distance scale, and L/c as the time scale — where c is the wave speed. Then the equation becomes

$$u_{tt} - u_{xx} - \Delta_\perp u = 0, \quad (8.1)$$

where $\Delta_\perp = \partial_y^2 + \partial_z^2$. Now we seek a solution of the form $$\boldsymbol{u = A(\chi,\, \tau,\, y,\, z)\, e^{i(x-t)/\epsilon}}, \quad (8.2)$$ where $\boldsymbol{\tau = \epsilon\, t}$ and $\boldsymbol{\chi = \epsilon\, x}$. Substituting, we obtain:

$$2\, i\, (A_\tau + A_\chi) + \Delta_\perp A = \epsilon^2\, (A_{\tau\tau} - A_{\chi\chi}) \quad (8.3)$$

Neglecting the higher order terms yields the **paraxial approximation**

$$\boldsymbol{2\, i\, (A_\tau + A_\chi) + \Delta_\perp A = 0.} \quad (8.4)$$

Next we will exhibit an exact solution for this equation.

8.1.1 The Gaussian beam solution

We look for solutions of (8.4) which are rotationally invariant relative to the axis of propagation. Specifically $$\boldsymbol{A = \beta(\chi,\, \tau)\, \exp\left(-\alpha(\chi,\, \tau)\, r^2\right)}, \quad (8.5)$$ where $\boldsymbol{r^2 = y^2 + z^2}$. Substituting into (8.4) yields

$$i\,(\mathcal{D}\beta - r^2\,\beta\,\mathcal{D}\alpha) - 2\,\alpha\,\beta + 2\,\alpha^2\,\beta\, r^2 = 0, \quad \text{where} \quad \boldsymbol{\mathcal{D} = \partial_\tau + \partial_\chi}. \quad (8.6)$$

Equating the terms independent of r, and proportional to r^2, now gives

$$\boldsymbol{i\,\mathcal{D}\beta = 2\,\alpha\,\beta} \quad \text{and} \quad \boldsymbol{i\,\mathcal{D}\alpha = 2\,\alpha^2.} \quad (8.7)$$

Note that, in fact, these are ode along the characteristics $\frac{\mathrm{d}\chi}{\mathrm{d}\tau} = 1$. The general solution has the form

$$\boldsymbol{\alpha = \frac{i}{f(\chi - \tau) - 2\,\tau} = \frac{i}{g(\chi - \tau) - 2\,\chi}} \quad \text{and} \quad \boldsymbol{\beta = h(\chi - \tau)\,\alpha}, \quad (8.8)$$

where $g(\xi) = 2\,\xi + f(\xi)$ and $h = h(\xi)$ are arbitrary functions. Now we can write

$$g = 2\,\chi_0 + i\,w_0^2, \quad \text{and} \quad h = a_0\,w_0^2\,e^{i\,\psi_0} \tag{8.9}$$

where the variables with subindex zero are (real valued) functions of $\xi = \chi - \tau$. In the **steady case** (no dependence on τ) **all these quantities are constants.** The reason for taking $\mathrm{Im}(g) > 0$ will soon become clear. Define:

$$\zeta = \chi - \chi_0, \quad \zeta_0 = \frac{1}{2}\,w_0^2, \quad w = w_0\sqrt{1 + \frac{\zeta^2}{\zeta_0^2}}, \quad R = \zeta\left(1 + \frac{\zeta_0^2}{\zeta^2}\right), \quad \text{and} \quad \psi = \mathrm{atan}\left(\frac{\zeta}{\zeta_0}\right) - \psi_0. \tag{8.10}$$

Then some straightforward, but cumbersome, calculations show that

$$\alpha = \frac{1}{w^2} - \frac{i}{2\,R} \quad \text{and} \quad \beta = a_0\,\frac{w_0}{w}\,e^{-i\,\psi}. \tag{8.11}$$

Thus

$$A = a_0\,\frac{w_0}{w}\,\exp\left(-\frac{r^2}{w^2} + i\,\frac{r^2}{2\,R} - i\,\psi\right). \tag{8.12}$$

Note that **this solution is contained within a gaussian radius w of the main axis, whose minimum occurs when $\zeta = 0$.** This is why it is called a "Gaussian" beam. Also note that, had we taken $\mathrm{Im}(g) = w_0^2 < 0$, the solution would blows up away from the axis. Note also that $2\,\pi\,R$ plays the role of a radial wave length: the phase of A, say Θ, satisfies $\Theta_r = r/R$.

The variable ψ is called the Gouy's phase. Consider now what happens with ψ, in the steady state case, for ζ small. There $\psi \approx \zeta/\zeta_0 - \psi_0$. The effect of this is to change the main phase in (8.2) from $(x-t)/\epsilon$ to $(x - \epsilon^2\,x/\zeta_0 - t)/\epsilon$. That is: **the phase speed of the wave near its waist is bigger than one.**

8.1.2 Weak dissipative effects

Here we consider the changes that adding a (small) amount of dissipation to the wave equation cause. The dissipation has to be small enough that it does not destroy the paraxial approximation. As an example, consider the equation

$$u_{tt} - 2\,\epsilon^3\,\mu\Delta\,u_t - u_{xx} - \Delta_\perp\,u = 0, \tag{8.13}$$

where $\mu > 0$ is an $O(1)$ constant. It is easy to check that this changes (8.4) to

$$2\,i\,(A_\tau + A_\chi + \mu\,A) + \Delta_\perp\,A = 0. \tag{8.14}$$

Thus let $\tilde{A}$ solve (8.4), and $a = a(\chi,\,\tau)$ solve $a_\tau + a_\chi + \mu\,a = 0$. Then $A = a\,\tilde{A}$ solves (8.14).

We can deal with a larger amount of dissipation by changing the ansatz in (8.2). Thus take

$$u_{tt} - 2\,\epsilon^2\,\mu\Delta\,u_t - u_{xx} - \Delta_\perp\,u = 0. \tag{8.15}$$

Then $u = A(\chi,\,\tau,\,y,\,z)\,e^{i\frac{x-t}{\epsilon} - \mu\,t + q(x-t) + i\epsilon\,p\,t}$, with $p = \frac{1}{2}\,\mu^2 + 2\,\mu\,q$ and $q =$ constant, yields: †

$$2\,i\,(A_\tau + A_\chi) + \Delta_\perp\,A = 0. \tag{8.16}$$

That is: (8.4).

† This involves a bit of messy, but straightforward, algebra.

From this we see that the dissipation has two effects, depending on the type of situation.

1. **Case $q = 0$.** If the wave is initialized with a modulated sinusoidal, $u = A(\chi,\,0,\,y,\,z)\,e^{i\frac{x-t}{\epsilon}}$, then decay in time (and a correction to the frequency) are induced: the factor $e^{-\mu\,t + i\,p\,\tau}$.

2. **Case $q = -\mu$.** If the wave is triggered at a boundary (say $x = 0$) with a sinusoidal behavior in time $u = A(0, \tau, y, z)\, e^{-i\frac{t}{\epsilon}+ip\tau}$ then decay occurs as the wave propagates: the factor $e^{-\mu x}$.

For $-\mu < q < 0$ a mix of the two behaviors is involved. Note that **(i)** It cannot be $q < -\mu$, unless energy is being supplied to the wave from outside. In the case of signaling from the boundary the input amplitude must grow with time. **(ii)** The case $q > 0$ corresponds to a decaying boundary signal, which was much bigger in the past. In this case the wave cannot extend to $x = \infty$, since this corresponds to an arbitrarily large input in the far past.

Remark 8.1 Larger dissipation. Making the dissipation one order larger than in (8.15) results in solutions where the "wave" is dissipated on scales comparable to the wave-length or the period. In this case there is no point in considering modulations of the amplitude over many periods or wave-lengths (as the paraxial approximation involves), since the signal is dissipated in a few periods after triggered (or wave-lengths away from a source producing it). Furthermore, then even the notion of "wave" becomes somewhat suspect.
In fact, it can be shown that (8.14) and (8.16) only make sense for solutions that are "band limited". The reason is that too much short-wave component creates a "residue" that can decay much slower than the main component, and make the solution to (8.14) or (8.16) inaccurate. ♣

8.2 The Paraxial approximation for dispersive equations (with Fourier transforms)

Here we consider rotationally symmetric systems characterized by a dispersion relation $G(\omega, k^2) = 0$. † This means that the solutions to the system can be written as linear combinations of solutions proportional to $e^{i(\vec{k}\cdot\vec{x}-\omega t)}$.

† We require ω to be real whenever $\vec{k} = (k_1, \ldots, k_n)$ is real. However the system need not be strictly dispersive. Thus the matrix $\frac{\partial^2\omega}{\partial k_i\, \partial k_j}$ can be singular, as it is for the wave equation.

Notation: for a vector $\vec{v}$, we use v to indicate its length when this can cause no confusion.

Remark 8.2 Characterization of the paraxial approximation on the Fourier Transform side. Here we show that, in terms of the Fourier Transform, *the paraxial approximation corresponds to solutions with a Fourier Transform concentrated on a small "pancake-like" region, of thickness $O(\epsilon^2)$ and radius $O(\epsilon)$,* where $0 < \epsilon \ll 1$. ‡ Furthermore, the region sits at the end of the main wave-number $\vec{k}_0$, with its flat part normal to $\vec{k}_0$ (equation (8.17) implements this idea). Hence the wave-numbers are preferentially aligned with the main wave-number; we call these solutions: "directional narrow band".
We also show that *the Gaussian beams correspond to the situation where the spectral amplitude in the "pancake" is given by a Gaussian in the direction perpendicular to $\vec{k}_0$.* ♣

‡ This implies a-dimensional units where $\vec{k}_0$ and the corresponding wave-frequency ω_0 are $O(1)$.

Consider now a branch $\omega = \Omega(\vec{k}) = \alpha(k^2)$ of the dispersion relation, and directional narrow band solutions of the form

$$u = \int a\left(\frac{k_1 - k_0}{\epsilon^2}, \frac{\vec{k}_\perp}{\epsilon}\right) e^{i\,(\vec{k}\cdot\vec{x}-\omega\, t)}\, dk_1 \ldots dk_n \tag{8.17}$$

where: (i) $n \geq 2$ is the number of space dimensions. (ii) k_0 is some constant, with corresponding wave-number $\omega_0 = \alpha(k_0^2)$. (iii) We have aligned the coordinate system so that the x_1-axis corresponds to the main direction of propagation: $\vec{k}_0 = (k_0, 0 \ldots, 0)$. (iv) $\vec{k}_\perp = (k_2, \ldots, k_n)$. (v) $0 < \epsilon \ll 1$. (vi) a decays "fast" as its arguments grow in size.
We now write $k_1 = k_0 + \epsilon^2\,\kappa$ and $\vec{k}_\perp = \epsilon\,\vec{\kappa}_\perp$, so that: $k^2 = k_0^2 + \epsilon^2\,(\kappa_\perp^2 + 2\,k_0\,\kappa) + O(\epsilon^4)$. Thus

$$\omega = \omega_0 + \epsilon^2\,\alpha'(k_0^2)(\kappa_\perp^2 + 2\,k_0\,\kappa) + O(\epsilon^4), \tag{8.18}$$

where we note that the group speed satisfies $\vec{c}_g = 2\,\alpha'(k^2)\,\vec{k}$. (8.19)

Substituting this into (8.17), and using the fact that a decays at infinity to neglect the $O(\epsilon^4)$ terms in ω, yields

$$\begin{aligned} u &\sim A(\chi,\,\tau,\,\vec{y})\,e^{i\,(k_0\,x_1-\omega_0\,t)}, \quad \text{where} \\ A &= \int a(\kappa,\,\vec{\kappa}_\perp)\,\exp\left\{i\,\left[\kappa\,\chi+\vec{\kappa}_\perp\cdot\vec{y}-\alpha'(k_0^2)(\kappa_\perp^2+2\,k_0\,\kappa)\,\tau\right]\right\}\mathrm{d}\kappa\,\mathrm{d}\kappa_2\ldots\mathrm{d}\kappa_n, \end{aligned} \tag{8.20}$$

$\chi=\epsilon^2\,x_1$, $\tau=\epsilon^2\,t$, $\vec{y}=\epsilon\,(x_2,\,\ldots,\,x_n)$, and $\vec{\kappa}_\perp=(\kappa_2,\,\ldots,\,\kappa_n)$. Note now that A is the general solution to the **Paraxial approximation** equation:

$$i\,\left(A_\tau+2\,\alpha'(k_0^2)\,k_0\,A_\chi\right)+\alpha'(k_0^2)\,\Delta_\perp A=0, \tag{8.21}$$

where $\Delta_\perp$ denotes the Laplacian in $\vec{y}$. Note that when $\alpha'(k_0^2)=0$ this equation becomes trivial, meaning that we need to go to higher order to obtain a meaningful equation. ***Thus we assume $\alpha'(k_0^2)\neq 0$. We also assume $k_0\neq 0$, so that we the term A_χ is present.***

Remark 8.3 Relationship with §8.1. For $\omega=k$, so that $\alpha(\zeta)=\sqrt{\zeta}$, and $k_0=1$ the equations here reduce to those in §8.1. The only difference is in the non-dimensionalization: here the main wave-length is selected as $O(1)$, while it is $O(\epsilon)$ in §8.1. Thus there is a "conversion" factor of ϵ between the variables there and the ones here. ♣

Next we write (8.21) in various alternative forms:

A. To make role of group speed apparent. Let $\vec{n}=(1,\,0\ldots 0)$ be the unit vector along the main direction of propagation. Then, using (8.19), we have

$$2\,i\,k_0\,\left(A_\tau+(\vec{c}_g)_0\cdot\vec{n}\,A_\chi\right)+(\vec{c}_g)_0\cdot\vec{n}\,\Delta_\perp A=0, \tag{8.22}$$

upon multiplying (8.21) by $2\,k_0$. Next introduce $\zeta=\chi-2\,\alpha'(k_0^2)\,k_0\,\tau=\chi-2\,(\vec{c}_g)_0\cdot\vec{n}\,\tau$.

B. Eliminate one variable. Change variables, and write $A=A(\zeta,\tau,\,\vec{y})$. Then

$$i\,A_\tau+\alpha'(k_0^2)\,\Delta_\perp A=0, \tag{8.23}$$

while (8.20) yields $$A=\int \tilde{a}(\zeta,\,\vec{\kappa}_\perp)\,\exp\left\{i\,\left[\vec{\kappa}_\perp\cdot\vec{y}-\alpha'(k_0^2)\,\kappa_\perp^2\,\tau\right]\right\}\mathrm{d}\kappa_2\ldots\mathrm{d}\kappa_n, \tag{8.24}$$

where $\tilde{a}(\zeta,\,\vec{\kappa}_\perp)=\int a(\kappa,\,\vec{\kappa}_\perp)\,e^{i\,\kappa\,\zeta}\mathrm{d}\kappa$. Note that **here the dependence on ζ is parametric and thus "arbitrary".** The particular case $A=A(\tau,\,\vec{y})$ follows when there is no dependence on ζ.

C. Eliminate one variable. Change variables, and write $A=A(\zeta,\chi,\,\vec{y})$. Then

$$2\,k_0\,i\,A_\chi+\Delta_\perp A=0, \tag{8.25}$$

while (8.20) yields $$A=\int \tilde{a}(\zeta,\,\vec{\kappa}_\perp)\,\exp\left\{i\,\left[\vec{\kappa}_\perp\cdot\vec{y}-\frac{1}{2\,k_0}\,\kappa_\perp^2\,\chi\right]\right\}\mathrm{d}\kappa_2\ldots\mathrm{d}\kappa_n, \tag{8.26}$$

where $\tilde{a}(\zeta,\,\vec{\kappa}_\perp)=\int a(\kappa,\,\vec{\kappa}_\perp)\,\exp\left\{i\,\left(\kappa+\frac{1}{2\,k_0}\kappa_\perp^2\right)\,\zeta\right\}\mathrm{d}\kappa$. Note that **here the dependence on ζ is parametric and thus "arbitrary".** The particular case $A=A(\chi,\,\vec{y})$ follows when there is no dependence on ζ.

8.2.1 Special solutions of the equation — Gaussian beams

Modulo constants, which can be eliminated by a proper scaling of the variables, (8.21) is the same as (8.4). Hence the solution in §8.1.1 applies. Here we show how to obtain this, and other, solutions using the expressions in (8.20), (8.24), and (8.26). Since these are essentially equivalent, we will concentrate on (8.26).

Consider a situation where $\tilde{a}$ is a Gaussian. Specifically: $\tilde{\boldsymbol{a}} = \boldsymbol{p}\, \boldsymbol{e}^{-\boldsymbol{q}\,\boldsymbol{\kappa}_\perp^2}$, (8.27)

where $p = p(\zeta)$ and $q = q(\zeta) > 0$. Then (8.26) becomes

$$\boldsymbol{A} = \boldsymbol{p} \prod_{\mathbf{2}}^{\boldsymbol{n}} \int \boldsymbol{e}^{-\boldsymbol{\sigma}^{\mathbf{2}}\,\boldsymbol{\kappa}_j^{\mathbf{2}} + \boldsymbol{i}\,\boldsymbol{\kappa}_j\,\boldsymbol{y}_j} \mathbf{d}\boldsymbol{\kappa}_j = \frac{\boldsymbol{\pi}^{(\boldsymbol{n}-\mathbf{1})/\mathbf{2}}}{\boldsymbol{\sigma}^{\boldsymbol{n}-\mathbf{1}}} \boldsymbol{p}\, \exp\left(-\frac{\boldsymbol{y}^{\mathbf{2}}}{\mathbf{4}\,\boldsymbol{\sigma}^{\mathbf{2}}}\right), \tag{8.28}$$

where $\boldsymbol{\sigma}$ is the root with $\mathbf{arg}(\boldsymbol{\sigma}) < \boldsymbol{\pi}/\mathbf{4}$ of the equation $\boldsymbol{\sigma}^{\mathbf{2}} = \boldsymbol{q} + \frac{\boldsymbol{i}}{\mathbf{2}\,\boldsymbol{k_0}}\,\boldsymbol{\chi}$, and we have used (8.29). For $n = 3$ this is the same solution in obtained in §8.1.1.

Remark 8.4 Gaussian integrals. Here we show that:

$$\boldsymbol{I} = \int_{-\infty}^{\infty} \boldsymbol{e}^{-\boldsymbol{\sigma}^{\mathbf{2}}\,\boldsymbol{k}^{\mathbf{2}} + \boldsymbol{i}\,\boldsymbol{k}\,\boldsymbol{z}}\, \mathbf{d}\boldsymbol{k} = \frac{\sqrt{\boldsymbol{\pi}}}{\boldsymbol{\sigma}}\, \exp\left(-\frac{\boldsymbol{z}^{\mathbf{2}}}{\mathbf{4}\,\boldsymbol{\sigma}^{\mathbf{2}}}\right). \tag{8.29}$$

where $\boldsymbol{I}$ is defined by the first equality, $\boldsymbol{\sigma}$ and $\boldsymbol{z}$ are complex numbers, and $\mathbf{arg}(\boldsymbol{\sigma}) < \boldsymbol{\pi}/\mathbf{4}$. †

† Then $\mathrm{Re}(\boldsymbol{\sigma}^{\mathbf{2}}) > \mathbf{0}$. Vice versa, if $\mathrm{Re}(\boldsymbol{\alpha}) > \mathbf{0}$ there exist a unique $\boldsymbol{\sigma}$ with $\boldsymbol{\alpha} = \boldsymbol{\sigma}^{\mathbf{2}}$ and $\mathbf{arg}(\boldsymbol{\sigma}) < \boldsymbol{\pi}/\mathbf{4}$.

Proof. For a fixed z, I (as defined by the integral) is an analytic function of σ in the region $\arg(\sigma) < \pi/4$. Thus it is enough to consider the case where σ is real and positive. In this case

$$\boldsymbol{I} = \frac{\mathbf{1}}{\boldsymbol{\sigma}} \int_{-\infty}^{\infty} \boldsymbol{e}^{-\boldsymbol{k}^{\mathbf{2}} + \boldsymbol{i}\,\boldsymbol{k}\,\frac{\boldsymbol{z}}{\boldsymbol{\sigma}}}\, \mathbf{d}\boldsymbol{k} = \frac{\mathbf{1}}{\boldsymbol{\sigma}} \int_{-\infty}^{\infty} \boldsymbol{e}^{-\left(\boldsymbol{k} - \boldsymbol{i}\,\frac{\boldsymbol{z}}{\mathbf{2}\,\boldsymbol{\sigma}}\right)^{\mathbf{2}}}\, \mathbf{d}\boldsymbol{k}\; \boldsymbol{e}^{-\frac{\boldsymbol{z}^{\mathbf{2}}}{\mathbf{4}\,\boldsymbol{\sigma}^{\mathbf{2}}}} = \frac{\mathbf{1}}{\boldsymbol{\sigma}} \int_{-\infty}^{\infty} \boldsymbol{e}^{-\boldsymbol{k}^{\mathbf{2}}}\, \mathbf{d}\boldsymbol{k}\; \boldsymbol{e}^{-\frac{\boldsymbol{z}^{\mathbf{2}}}{\mathbf{4}\,\boldsymbol{\sigma}^{\mathbf{2}}}}.$$

The last equality follows by moving the contour of integration: $k \to k + i\,\frac{z}{2\sigma}$. Then use $\int_{-\infty}^{\infty} e^{-k^2}\, \mathrm{d}k = \sqrt{\pi}$. ♣

Remark 8.5 General solution. The result above can be easily generalized to situations where $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{\zeta}, \vec{\boldsymbol{\kappa}}_\perp)$ is a polynomial in $\vec{\boldsymbol{\kappa}}_\perp$. This follows because (8.29) yields (by differentiation with respect to z) an explicit formula for $\boldsymbol{I} = \int_{-\infty}^{\infty} \boldsymbol{k}^{\boldsymbol{\ell}} \boldsymbol{e}^{-\boldsymbol{\sigma}^{\mathbf{2}}\,\boldsymbol{k}^{\mathbf{2}} + \boldsymbol{i}\,\boldsymbol{k}\,\boldsymbol{z}}\, \mathbf{d}\boldsymbol{k}$ — ℓ a natural number.

Since any square integrable function can be expanded in terms of the orthogonal functions $\{\boldsymbol{H_j}(\boldsymbol{\kappa})\, \boldsymbol{e}^{-\boldsymbol{\kappa}^{\mathbf{2}}/\mathbf{2}}\}$, any solution with a square integrable Fourier Transform can be expanded in terms of the solutions above. Here $\boldsymbol{H_j}(\boldsymbol{x}) = (-\mathbf{1})^{\boldsymbol{j}} \boldsymbol{e}^{\boldsymbol{x}^{\mathbf{2}}} \frac{\mathbf{d}^{\boldsymbol{j}}}{\mathbf{d}\boldsymbol{x}^{\boldsymbol{j}}} \boldsymbol{e}^{-\boldsymbol{x}^{\mathbf{2}}}$ is the j-th Hermite polynomial. ♣

---

9 Moving point sources.

In this section we consider the wave patterns produced by sources moving at a constant speed in a (linear) dispersive media (in open space). In particular, we are interested in the steady state, after all the transients are gone. To simplify matters we *only consider sources whose dimensions can be neglected, so that they can be approximated by a point.*

If the *sources do not actively generate waves at particular frequencies,*[14] then the steady state wave pattern is made by a combination of all the waves that are steady in the fame of reference of the source, with their location in space determined by their group speed. Examples of this can be found below in § 9.1, § 9.2, and § 9.3.

9.1 Moving point source — string on a bed.

Here we consider the example where the waves are governed by the (linear) Klein Gordon equation (string on a bed), so that the problem to be solved is (as usual, *we work in non-dimensional variables*)

$$u_{tt} - u_{xx} + u = \delta(x - V\,t), \quad -\infty < x < \infty, \tag{9.1}$$

where V is a constant — since the equation is left-right symmetric, assume $\boldsymbol{V \geq 0}$. The conditions at $x = V\,t$ are

c1. u is continuous.

c2. u_x and u_t have a simple jump discontinuity, such that (see remark 9.1)

$$-\left(u_t^R - u_t^L\right) V - \left(u_x^R - u_x^L\right) = 1, \tag{9.2}$$

where the super-scripts R and L indicate values immediately to the right and left of $x = V\,t$. For a steady state solution $u = u(\zeta)$, with $\zeta = x - V\,t$, (9.1–9.2) reduce to

$$(V^2 - 1)\,u'' + u = \delta(\zeta) \quad \text{and} \quad (V^2 - 1)\,\left((u')^R - (u')^L\right) = 1, \tag{9.3}$$

where the primes indicate derivatives with respect to ζ.

For solutions proportional to $e^{i\,(k\,x - \omega\,t)}$, the dispersion relation, phase speed, and group speed are

$$\omega^2 = 1 + k^2, \quad c_p = \frac{\omega}{k}, \quad \text{and} \quad c_g = c_p^{-1}, \quad \Longrightarrow \quad 0 < c_g^2 < 1 < c_p^2 \ \text{ for all } \ k \neq 0. \tag{9.4}$$

The steady state solution to (9.1) has to be produced by combining waves which are steady in a frame of reference moving at speed V.

Case $\boldsymbol{V = \sqrt{1 + \mu^2}}$, $\boldsymbol{\mu > 0}$.

Then the steady state solution will combine the waves with $k = \pm 1/\mu$, which are the solutions to the equation $c_p = V$. Both these waves must be to the left of the source (since $c_g < V$). Since $V > 1$, nothing can move faster than the source, so that u must vanish ahead of the source. Thus [15]

$$u = \alpha\,\sin \mathcal{Z} + \beta\,\cos \mathcal{Z}, \quad \text{for} \quad \mathcal{Z} < 0, \tag{9.5}$$

$$u = 0, \qquad\qquad\qquad\quad \text{for} \quad \mathcal{Z} > 0, \tag{9.6}$$

where: (i) α and β are constants; and (ii) $\boldsymbol{\mathcal{Z} = \dfrac{1}{\mu}\,(x - V\,t)}$. Then the conditions **c1** and **c2** yield

$$\beta = 0 \quad \text{and} \quad -1 = (V^2 - 1)\,\frac{\alpha}{\mu} = \mu\,\alpha. \tag{9.7}$$

[14] An example of this occurs if the source is an oscillator, inputing energy into the system at some frequency.

[15] Note that $\omega = \pm\,V/\mu$ for $k = \pm 1/\mu$, and we pick the root that yields $c_p = V$.

Hence

$$\boldsymbol{u = -\frac{1}{\mu}\,\sin \mathcal{Z}, \quad \text{for} \quad \mathcal{Z} < 0, \quad \text{and} \quad u = 0, \quad \text{for} \quad \mathcal{Z} > 0.} \tag{9.8}$$

Case $V = \sqrt{1-\mu^2}$, $0 < \mu \leq 1$.

In this case there are no "waves", and the solution must have the form

$$u = \alpha\, e^{\mathcal{Z}}, \qquad \text{for} \quad \mathcal{Z} < 0, \tag{9.9}$$

$$u = \beta\, e^{-\mathcal{Z}}, \qquad \text{for} \quad \mathcal{Z} > 0, \tag{9.10}$$

where: (i) α and β are constants; and (ii) $\boldsymbol{\mathcal{Z} = \frac{1}{\mu}\,(x - V\,t)}$. Then the conditions **c1** and **c2** yield

$$\alpha = \beta \quad \text{and} \quad 1 = (V^2 - 1)\left(-\frac{\beta}{\mu} - \frac{\alpha}{\mu}\right) = \mu\,(\alpha + \beta). \tag{9.11}$$

Hence

$$\boldsymbol{u = \frac{1}{2\,\mu}\, e^{-|\mathcal{Z}|}.} \tag{9.12}$$

In this case the solution consists of two evanescent waves, with no disturbance far away from the source. In particular, there is no energy exchange between the source and the wave field.[16]

Case $V = 1$.

In this case there are no waves that can propagate away from the source (evanescent or otherwise). Everything stays with the source. The steady solution is then

$$\boldsymbol{u = \delta(x - t),} \tag{9.13}$$

which, of course, only solves the equation in the weak sense of distributions.

Remark 9.1 *To understand the meaning of (9.1–9.2), let us go back to the derivation of the equation using the conservation of transversal momentum. That is, the equation*

$$\frac{d}{dt}\int_a^b \rho\, u_t(x,\,t)\,dx = T\,u_x(b,\,t) - T\,u_x(a,\,t) - \int_a^b k_b\,u(x,\,t)\,dx + \int_a^b S(x,\,t)\,dx, \tag{9.14}$$

where $a < b$ are arbitrary constants, ρ is the mass per unit length of the string, T is the string tension, k_s is the elastic constant of the bed, and S are any additional sources of momentum. In fact, imagine that

> ***S* is produced by a needle point, sliding along the string at constant speed, and applying a constant pressure to it.**

Then we can approximate S by a Delta-function. In non-dimensional variables this yields

$$\frac{d}{dt}\int_a^b u_t(x,\,t)\,dx = u_x(b,\,t) - u_x(a,\,t) - \int_a^b u(x,\,t)\,dx + \int_a^b \delta(x - V\,t))\,dx. \tag{9.15}$$

[16] Other than at the beginning, when the steady state solution is being established — transient period.

In this case u will satisfy $u_{tt} - u_{xx} + u = 0$ on each side of the delta. Consider now a and b such that $a < V\,t < b$. Then we split the integral on the left into the piece from a to $V\,t$, and the piece from $V\,t$ to b. Then

$$\begin{aligned}\frac{d}{dt}\int_a^b u_t(x,\,t)\,dx &= \frac{d}{dt}\left(\int_a^{V\,t} u_t(x,\,t)\,dx + \int_{V\,t}^b u_t(x,\,t)\,dx\right)\\ &= V\,u_t(V\,t-0,t) - V\,u_t(V\,t+0,t) + \int_a^{V\,t} u_{tt}(x,\,t)\,dx + \int_{V\,t}^b u_{tt}(x,\,t)\,dx,\end{aligned}$$

where $V\,t \pm 0$ is used to indicate evaluation on each side of the delta. Furthermore, using the fact that u satisfies the equation on each side of the delta, we can write

$$\int_a^{V\,t} u_{tt}(x,\,t)\,dx = u_x(V\,t-0,\,t) - u_x(a,\,t) - \int_a^{V\,t} u(x,\,t)\,dx$$

and

$$\int_{V\,t}^b u_{tt}(x,\,t)\,dx = u_x(b,\,t) - u_x(V\,t+0,\,t) - \int_{V\,t}^b u(x,\,t)\,dx.$$

Substitute now all of this into (9.15), plus the fact that $\int_a^b \delta(x - V\,t))\,dx = 1$ if $a < V\,t < b$, to obtain

$$V\,u_t(V\,t-0,t) - V\,u_t(V\,t+0,t) + u_x(V\,t-0,t) - u_x(V\,t+0,t) = 1. \tag{9.16}$$

This is the same as (9.2). ♣

9.2 Moving point source — KdV equation.

Here we consider the example where the waves are governed by the (linear) KdV equation, so that the problem to be solved is (as usual, *we work in non-dimensional variables*)

$$u_t + u_{xxx} = \delta(x - V\,t), \quad -\infty < x < \infty, \tag{9.17}$$

where V is a constant. The conditions at $x = V\,t$ are

c1. u is continuous.

c2. u_x is continuous.

c3. u_{xx} has a simple jump discontinuity, with a jump of one across the point $x = V\,t$. That is:

$$\boldsymbol{u^R_{xx} - u^L_{xx} = 1},$$

where the super-scripts R and L indicate values immediately to the right and left of $x = V\,t$.

For solutions proportional to $e^{i\,(k\,x - \omega\,t)}$, the dispersion relation, phase speed, and group speed are

$$\omega = -k^3, \quad c_p = -k^2, \quad \text{and} \quad c_g = -3\,k^2, \quad \Longrightarrow \quad c_g < c_p \ \text{ for all } \ k \neq 0. \tag{9.18}$$

The steady state solution to (9.17) has to be produced by combining waves which are steady in a frame of reference moving at speed V. However, *unlike the situation in* § 9.1, *this also includes the state* $u =$ constant — whose group speed is $c_g = 0$.

The condition for a wave to be steady in the source frame of reference is $\boldsymbol{c_p = V}$. However, the phase speed $c_p = \omega/k$ is not well defined for $k = 0$. In particular, when $\omega(0) = 0$, the constants are solutions which are steady on any frame. Further, there is no problem defining the group speed $c_g = \dfrac{d\omega}{dk}$ for $k = 0$.

Case $V = -\mu^2$, $\mu > 0$.

Then the steady state solution will combine the waves with $k = \pm\mu$, and the constant state. The waves to the left of the source (as $c_g < V$ for them), and a constant state ahead. Thus

$$u = \alpha \sin \mathcal{Z} + \beta \cos \mathcal{Z}, \quad \text{for} \quad \mathcal{Z} < 0, \tag{9.19}$$

$$u = \gamma, \qquad\qquad\qquad \text{for} \quad \mathcal{Z} > 0, \tag{9.20}$$

where: (i) α, β, and γ are constants; and (ii) $\boldsymbol{\mathcal{Z} = \mu\,(x - V\,t) = \mu\,x + \mu^3\,t}$. Then the conditions **c1**, **c2**, and **c3**, yield

$$\beta = \gamma, \quad \mu\,\alpha = 0, \quad \text{and} \quad \mu^2\,\beta = 1. \tag{9.21}$$

Hence

$$\boldsymbol{u = \frac{1}{\mu^2}\,\cos \mathcal{Z}, \quad \text{for} \quad \mathcal{Z} < 0, \quad \text{and} \quad u = \frac{1}{\mu^2}, \quad \text{for} \quad \mathcal{Z} > 0.} \tag{9.22}$$

Case $V = \mu^2$, $\mu > 0$.

In this case there are no "waves", and the solution must have the form [17]

$$u = \alpha\, e^{\mathcal{Z}} + \gamma, \qquad \text{for} \quad \mathcal{Z} < 0, \tag{9.23}$$

$$u = \beta\, e^{-\mathcal{Z}}, \qquad \text{for} \quad \mathcal{Z} > 0, \tag{9.24}$$

where: (i) α, β, and γ are constants; and (ii) $\boldsymbol{\mathcal{Z} = \mu\,(x - V\,t) = \mu\,x - \mu^3\,t}$. Then the conditions **c1**, **c2**, and **c3**, yield

$$\alpha + \gamma = \beta, \quad \mu\,\alpha = -\mu\,\beta, \quad \text{and} \quad \mu^2\,\beta - \mu^2\,\alpha = 1. \tag{9.25}$$

Hence

$$\boldsymbol{u = -\frac{1}{2\,\mu^2}\,e^{\mathcal{Z}} + \frac{1}{\mu^2}, \quad \text{for} \quad \mathcal{Z} < 0, \quad \text{and} \quad u = \frac{1}{2\,\mu^2}\,e^{-\mathcal{Z}}, \quad \text{for} \quad \mathcal{Z} > 0.} \tag{9.26}$$

Case $V = 0$.

There is **no steady state.** The source is now in resonance with the constant state, whose group speed vanishes. *A resonance occurs whenever there is a wave with both phase and group speed equal to those of the source.*

9.3 Moving point source — water waves.

The End.

[17] The constant state is now behind the source, because it has group speed $= 0 < V$.