6

ONE-DIMENSIONAL WAVES

Introduction. We turn now from the oscillation of discrete systems (crystals, molecules, coupled oscillators) to an examination of the vibratory motion of *spatially distributed* systems—systems which cannot (like a particle) be said at time t to be "at the point $\boldsymbol{x}(t)$" or (like a molecule) to have constituent parts at the points $\{\boldsymbol{x}_1(t), \boldsymbol{x}_2(t), \ldots, \boldsymbol{x}_n(t)\}$ but which (like a string, or an air column, or a membrane) possess a *continuum* of "constituent parts" that are *spread out in space*.

We recognize at the outset that some/most of the systems with which we will concern outselves do resolve into interactive particles ("atoms") when examined in sufficiently fine detail: such systems can (very usefully!) be considered to be continuous only in macroscopic approximation. Other systems —most notably the electromagnetic field, whose vibrations we call "light"— do *not* resolve into "particles" when examined closely, but appear to retain their distributed character at every scale, however fine.[1] Remarkably, the essential structural features of the theory that will emerge are for the most part insensitive to whether the physical system to which they are being applied is "really" smooth or only seems so, insensitive to whether the system has a truly infinite or only a very large number of degrees of freedom.

Think (non-atomistically) of a fluid. Such a system can swirl, spash, do a lot of things that, though they fall within the perview of "fluid dynamics," lie beyond the bounds of the present discussion: we will be concerned only with the *concerted vibratory* motions that fluids and other distributed systems support when they sufficiently near a point of stable equilibrium. We will, in short, be

[1] The effect of quantum mechanical considerations, when they come into play, is not so much to introduce "electromagnetic discreteness" as it is to bring about a profound adjustment in the way we think/speak about electromagnetic phenomena.

concerned with the kinematics and dynamics of *waves*.

To gain certain graphic advantages, and to keep the mathematics as uncluttered as possible, we will restrict our attention to waves that are one-dimensional in the sense that the motion of the constituent parts of the system that supports the wave is along a line. The acoustic vibration of an air column provides, in leading approximation,[2] a good example. So does the propagation of an electrical signal along a wire. But such systems are a bit difficult to picture. It is for that reason that we will speak frequently of "strings," though the vibrational excursions of a marked point on a guitar string lie mainly in the 2-dimensional plane that is locally normal to the string. Our idealized strings will—at least initially—be assumed to have one of those degrees of freedom "frozen out."

1. Derivation of the wave equation. It is from our knowledge of how particles respond to impressed forces (Newton's $F = m\ddot{x}$) that we will extract a description of how—in the simplest instance—waves propagate in a one-dimensional medium. Our strategy (see FIGURE 1) will be to set up the coupled equations that describe the vibration of a "one-dimensional crystal," then to imagine a *sequence of such crystals* (lattices) in which the atoms are made progressively

- less massive, but
- more numerous and closer together

in such a way as to keep constant the mass per unit length. Notational adjustments will then permit us to carry this process (which I call "refinement of the lattice") to the continuous limit.

Working from the figure we have

$$\left.\begin{aligned} m\ddot{\varphi}_1 &= -k(\phantom{-\varphi_{n-1}} + 2\varphi_1 - \varphi_2\phantom{_{+1}}) \\ m\ddot{\varphi}_2 &= -k(-\varphi_1 \phantom{_{n-1}} + 2\varphi_2 - \varphi_3\phantom{_{+1}}) \\ &\vdots \\ m\ddot{\varphi}_n &= -k(-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) \qquad n = 2, 3, \ldots, N-1 \\ &\vdots \\ m\ddot{\varphi}_N &= -k(-\varphi_{N-1} + 2\varphi_N \phantom{- \varphi_{n+1}}) \end{aligned}\right\} \tag{1}$$

where $\varphi_n(t)$ answers the question: How far (at time t) has the n^{th} particle been displaced to the right of its equilibrium position? The reason for this notational adjustment will become apparent almost immediately.

We stipulate that total length ℓ and total mass M are constants of the lattice refinement process. At the N^{th} state of the process we therefore have

$$\begin{aligned} \text{particle mass } m &= M/N \\ \text{particle separation } a &= \ell/(N+1) \approx \ell/N \end{aligned}$$

[2] In an organ pipe a certain amount of vortex swirling does go on.

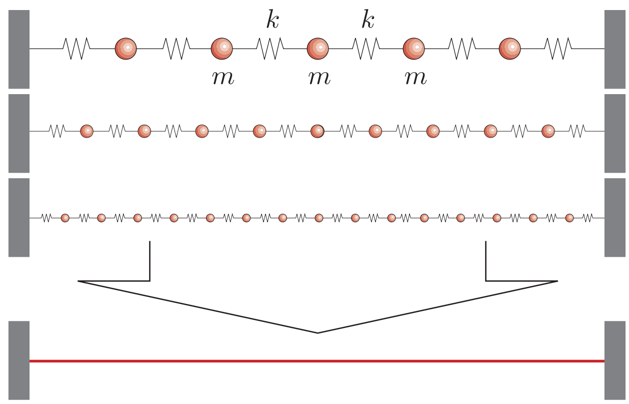


FIGURE 1: *Sequence of one-dimensional crystals that in the limit serves to model a continuous string or rod. It is assumed that at each step of the refinement process all masses are the same, and that so also are the springs that connect nearest neighbors to each other. The total mass and total length of the system are assumed to be held constant during the refinement.*

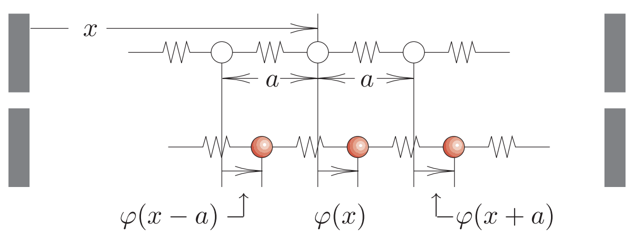


FIGURE 2: *Significance of the notations used for passage to the continuous limit.*

Introducing the (invariant) *linear mass density* $\mu = M/\ell$ we have

$$m = (M/\ell)(\ell/N) \approx \mu a$$

If it is our intention to let $N \uparrow \infty$ then we must <u>abandon ordinal indexing of the lattice particles</u>, for the p^{th} particle will in the limit lie infinitely close to the left end of the lattice, no matter how large is the value assigned to p. We give up discrete indexing in favor of "continuous indexing," writing

$$\varphi(x,t) = \begin{cases} \text{instantaneous displacement of the element} \\ \text{which when the system is at rest lives at } x \end{cases}$$

In that notation the equations of motion (1) become[3]

$$\mu a\,\frac{\partial^2}{\partial t^2}\varphi(x,t) = k\big\{\varphi(x-a,t) - 2\varphi(x,t) + \varphi(x+a,t)\big\}$$

which, after division by a, can be written

$$\mu\,\frac{\partial^2}{\partial t^2}\varphi(x,t) = ka\cdot\frac{\dfrac{\varphi(x+a,t)-\varphi(x,t)}{a} - \dfrac{\varphi(x,t)-\varphi(x-a,t)}{a}}{a}$$

If we assume that, during the course of the lattice refinement process, the *springs get stiffer as they get shorter* in such a way as to produce[4]

$$\lim_{a\downarrow 0} k(a)a = \text{constant, call it } \kappa$$

then in the limit $a\downarrow 0$ (which is to say: in the limit $N\uparrow\infty$) we have

$$\mu\,\frac{\partial^2}{\partial t^2}\varphi(x,t) = \kappa\,\frac{\partial^2}{\partial x^2}\varphi(x,t) \tag{2.1}$$

Dimensionally $[\mu] = ML^{-1}$ and $[\kappa] = [ka] = [\text{force}] = MLT^{-2}$ so

$$[\kappa/\mu] = L^2T^{-2} = [(\text{velocity})^2]$$

which is to say:

$$u = \sqrt{\kappa/\mu}\ \text{ is a velocity natural to the string/rod}$$

---

[3] This is true for the particles that live at the interior points

$$x_2 = 2a,\ x_3 = 3a,\ \ldots,\ x_{N-1} = (N-1)a$$

The particles that live at $x_1 = a$ and $x_N = Na$ satisfy eccentric equations of motion because each is "deprived of a neighbor," but need not concern us because in the limit $a\downarrow 0$ each condenses to an *endpoint* of the string/rod, and those (see again FIGURE 1) we presently consider to be fixed.

[4] But this is exactly how springs *do* behave! For springs k_1 and k_2 connected in series produce

$$k_{\text{effective}} = \left[\frac{1}{k_1} + \frac{1}{k_2}\right]^{-1}$$

while springs connected in parallel produce

$$k_{\text{effective}} = k_1 + k_2$$

Springs combine like capacitors! Cut a spring in half: each half is twice as stiff as the original spring. Cut a spring into N identical fragments: each fragment is N times as stiff. And in the limit $N\uparrow\infty$ each fragment becomes *infinitely* stiff!

and (2) can be written in any of the following equivalent ways:

$$\left.\begin{aligned}\Big\{\frac{1}{u^2}\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial x^2}\Big\}\varphi(x,t)&=0\\ \Big\{\frac{1}{u^2}\partial_t^2-\partial_x^2\Big\}\varphi(x,t)&=0\\ \Box\varphi&=0 \quad : \quad \Box \text{ is the “wave operator”}\\ \frac{1}{u^2}\varphi_{tt}-\varphi_{xx}&=0\end{aligned}\right\}\tag{2.2}$$

In (2)—however written—the large linear system of coupled ordinary differential equations has been replaced by a <u>single linear partial differential equation</u>, the so-called **wave equation**.[5]

It was Newton's $m\ddot{x} = F$ that led us to the wave equation (2.1), and is useful to notice that the equation retains a very Newtonian structure. On the left we have

$$\mu\frac{\partial^2}{\partial t^2}\varphi(x,t) = (\text{mass density})\cdot\begin{cases}\text{instantaneous acceleration of the}\\ \text{string element that lives at } x\end{cases}$$

while on the right we have

$$\begin{aligned}\text{force density} &\propto \frac{\varphi(x+dx)+\varphi(x-dx)}{2}-\varphi(x)\\ &=\begin{cases}\text{amount by which the value of } \varphi \text{ at } x \text{ is exceeded by}\\ \text{the average of its values at nearest-neighbor points}\end{cases}\end{aligned}$$

It becomes plausible in this light to conjecture that the force density of a "stiff string" would be sensitive also to the values assumed by φ at next-nearest-neighboring points, or perhaps at points even more remote.

2. Mechanical properties of simple waves. The total energy of the vibrating crystal shown in FIGURE 1 (and renotated in FIGURE 2) is

$$\begin{aligned}E &= \tfrac{1}{2}\sum_n m\dot{\varphi}_n^2+\tfrac{1}{2}\sum_n k(\varphi_{n+1}-\varphi_n)^2\\ &= \sum_{x\text{-addresses}}\Big\{\tfrac{1}{2}\mu\Big(\frac{\partial\varphi(x)}{\partial t}\Big)^2+\tfrac{1}{2}\kappa\Big(\frac{\varphi(x+a)-\varphi(a)}{a}\Big)^2\Big\}a\end{aligned}$$

---

[5] In fact there exist *many* "wave equations," variants of (2), but they bear a variety of specialized names: when one says "wave equation" one is universally understood to refer either to (2) or to its higher-dimensional generalization

$$\Big\{\frac{1}{u^2}\partial_t^2-\nabla^2\Big\}\varphi(x,t)=0$$

which in the continuous limit becomes[6]

$$E = \int_0^\ell \mathcal{E}(\varphi, \partial\varphi)\, dx \tag{3.1}$$

$$\begin{aligned}\mathcal{E}(\varphi, \partial\varphi) &= \tfrac{1}{2}\mu\Big(\frac{\partial\varphi(x)}{\partial t}\Big)^2 + \tfrac{1}{2}\kappa\Big(\frac{\partial\varphi(x)}{\partial x}\Big)^2\\ &= \tfrac{1}{2}\mu(\partial_t\varphi)^2 + \tfrac{1}{2}\kappa(\partial_x\varphi)^2\\ &= \textbf{energy density}\end{aligned} \tag{3.2}$$

Notice that $\mathcal{E}(\varphi, \partial\varphi)$ is everywhere non-negative, and that $[\mathcal{E}] = [\text{energy/length}]$. The mechanical energy of a vibrating string/rod resides nowhere in particular, but is *distributed along its length*. $\mathcal{E}dx$ describes how much of it resides at time t in the neighborhood dx of x.

From (3) it follows that

$$\begin{aligned}\dot{E} &= \int_0^\ell \frac{\partial\mathcal{E}}{\partial t}\, dx\\ \frac{\partial\mathcal{E}}{\partial t} &= \mu\varphi_{tt}\varphi_t + \kappa\varphi_x\varphi_{xt}\\ &= \kappa\varphi_{xx}\varphi_t + \kappa\varphi_x\varphi_{xt} \quad \text{by the wave equation}\\ &= \partial_x[\kappa\varphi_x\varphi_t]\end{aligned} \tag{4}$$

We conclude that E-variations—if any—must be end effects

$$\dot{E} = \kappa\varphi_x\varphi_t\Big|_0^\ell \tag{5}$$

and that imposition of any of several natural boundary conditions[7] would serve to render such effects impossible. Under such conditions **energy sloshes about, but is conserved in the aggregate**. That image sets in motion the following train of thought: let

$$\begin{aligned}\mathcal{F}(x,t) &= \textbf{energy flux}\\ &= \begin{cases}\text{temporal rate at which energy flows}\\ \text{to the right past the inspection point } x\end{cases}\end{aligned}$$

Fixing our attention now upon some infinitesimal string element, we have—as

[6] Here $\partial\varphi$ stands collectively for all the first partials of φ.

[7] We might, for example, clamp the ends of the the string/rod:

$$\varphi(0,t) = \varphi(\ell,t) = 0 \quad : \quad \text{all t}$$

This would enforce $\varphi_t(0,t) = \varphi_t(\ell,t) = 0$, and thus ensure $\dot{E} = 0$.

a *local* statement of energy conservation—

$$\begin{aligned}\frac{\partial}{\partial t}(\mathcal{E}dx) &= (\text{flux in on the left}) - (\text{flux out on the right})\\ &= \mathcal{F}(x,t) - \mathcal{F}(x+dx,t)\end{aligned}$$

which gives

$$\partial_t\mathcal{E} + \partial_x\mathcal{F} = 0 \tag{5}$$

Comparison with (4) leads to the conclusion that

$$\mathcal{F}(x,t) = -\kappa\varphi_x(x,t)\varphi_t(x,t) \tag{6}$$

Turning our attention now from energy to the *momentum* of a vibrating string/rod, we have

$$P = \sum_n m\dot{\varphi}_n = \sum_{x\text{-addresses}} \mu\,\frac{\partial\varphi(x)}{\partial t}a$$

$$\downarrow$$

$$= \int_0^\ell \mathcal{P}(\varphi,\partial\varphi)dx \tag{7.1}$$

$$\mathcal{P}(\varphi,\partial\varphi) = \mu\varphi_t \tag{7.2}$$

$$= \textbf{momentum density}$$

It follows that

$$\dot{P} = \int_0^\ell \mu\varphi_{tt}dx = \int_0^\ell \kappa\varphi_{xx}dx = \kappa\varphi_x\Big|_0^\ell$$

and that

$$\partial_t\mathcal{P} + \partial_x\mathcal{G} = 0 \tag{8}$$

where

$$\mathcal{G}(x,t) = -\kappa\varphi_x \tag{9}$$

$$= \textbf{momentum flux}$$

When one turns (as below) from the simple wave equation to the simpler of its many variants one finds that the energy/momentum densities and fluxes are described by expressions that change from case to case, but that conservation laws of the form (5/8)

3. Frequently encountered variants of the simple wave equation. FIGURE 3 shows a "crystal" in which each element is harmonically coupled not only not only to its nearest neighbors but also to its equilibrium position, about which it would oscillated even if decoupled from its neighbors. Refinement of such a system leads to a wave equation of the form

$$(\Box + \varkappa^2)\varphi = 0 \tag{10}$$

where $\varkappa$ is a constant with the physical dimension $[\varkappa] = 1/\text{length}$. This is known as the **Klein-Gordon equation**, and is fundamental to the relativistic quantum

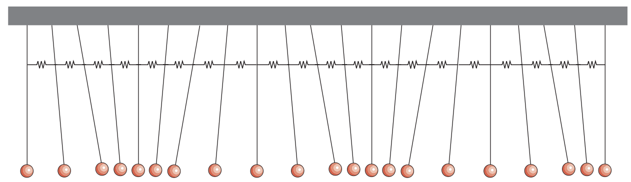


FIGURE 3: *This system of coupled pendula serves to model a crystal in which each element is coupled not only to its nearest neighbors, but also to its equilibrium site.*

theory of massive particles, in which context

$$\varkappa = mc/\hbar = \frac{1}{\text{Compton length}}$$

In the massless limit $m \downarrow 0$ one recovers the wave equation (2).

If one introduces *damping* into the system shown in FIGURE 1 then after refinement one is left with a partial differential equation of the form

$$\left\{\frac{1}{u^2}\partial_t^2 + 2(\gamma/u)\,\partial_t - \partial_x^2\right\}\varphi(x,t) = 0 \tag{11}$$

The damped version of the system shown in FIGURE 3 leads similarly to

$$\left\{\frac{1}{u^2}\partial_t^2 + 2(\gamma/u)\,\partial_t - \partial_x^2 + \varkappa^2\right\}\varphi(x,t) = 0$$

which is often written

$$\left\{\partial_t^2 + 2u\gamma\,\partial_t - u^2\partial_x^2 + (u\varkappa)^2\right\}\varphi(x,t) = 0 \tag{12}$$

and is known as the **telegrapher's equation**, for the interesting reason which I now explain.

The electrical properties of *transmission lines* (paired wires intended to carry electrical signals) are not localized—as are the elements of most electrical circuits—but are *distributed*: one speaks of the "resistance per unit length," the "inductance/capacitance per unit length," the "leakage conductance per unit length" and obtains equations that describe the behavior of such lines by working out the properties of such discrete circuits as are shown in FIGURE 4 and proceeding to the continuous limit in the now-familiar way. It is the "leaky lossy system" that leads to the telegrapher's equation, with

$$\begin{aligned} u^2 &= \frac{1}{LC} \\ 2u\gamma &= \frac{RC+GL}{LC} \qquad \text{where} \quad G = 1/r \\ (u\varkappa)^2 &= \frac{RG}{LC} \end{aligned}$$

So cunningly complicated are biological systems that—surprisingly to a physicist

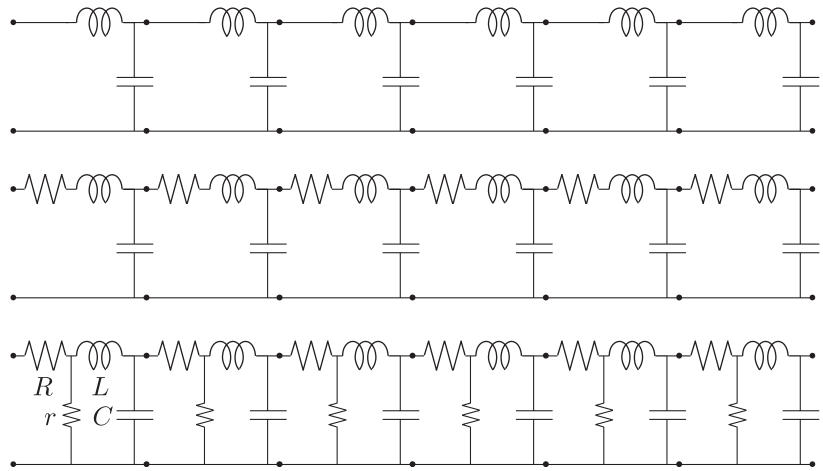


FIGURE 4: *Top figure*: *discrete circuit serves in the continuous limit to model a* **lossless** *transmission line. Middle figure*: *model of a* **lossy** *transmission line. Bottom figure*: *model of a* **leaky lossy** *transmission line.*

—the theory of transmission lines is only distantly relevant to the theory of signal transmission along axions.

> **HISTORICAL NOTE:** It was Jean le Rond d'Alembert (1717–1783)[8] who first took up study of the dynamics of vibrating strings (a subject some phenomenological aspects of which had also been of interest also to Pythagoras!)—whence of the wave equation—though the subject soon attracted the attention also of Euler, Daniel Bernoulli and Lagrange. It was this physically motivated work that launched the mathematical theory of partial differential equations.
>
> About a century passed before Oliver Heaviside and Lord Kelvin were motivated by the invention (∼1840) of telegraphy, and more particularly by the effort to lay a cable across the Atlantic (first accomplished in 1858), to devise and to study the solutions of the telegrapher's equation.
>
> Serious interest in the Klein-Gordon equation had to await the invention (1926) of the non-relativistic Schrödinger equation, which motivated physicists to try to devise a variant of that equation that conformed to the principle of relativity.

[8] d'Alembert was an interesting character. For basic information see, for example, http://en.wikipedia.org/wiki/Jean_le_ Rond_d'Alembert.

4. Solution of the wave equation. It should be noted at the outset that the wave function $\varphi(x,t)$ enters linearly into the wave equation (2), as also into each of the variants encountered in §3, the immediate implication being that

$$\begin{aligned}&\text{if } \varphi_1(x,t) \text{ and } \varphi_2(x,t) \text{ are solutions then}\\&\text{so also is } c_1\varphi_1(x,t)+c_2\varphi_2(x,t)\text{: all } c_1, c_2\end{aligned} \tag{13.1}$$

From linearity and reality it follows moreover that

$$\begin{aligned}&\text{one can always construe real solutions } \varphi(x,t) \text{ to}\\&\text{be the real parts of complex solutions } \psi(x,t)\end{aligned} \tag{13.2}$$

The "solution problem" can be approached in many ways, of which I will sketch several: each draws heavily upon (13.1), and most draw also upon (13.2). I remark in passing that the **free particle Schrödinger equation**—which reads

$$\psi_{xx} + i(2m/\hbar)\psi_t = 0$$

—is of such a form that (13.1) pertains but (13.2) does not.

FACTORIZATION METHOD We begin with a method that is wonderfully swift and elegant, but quite limited in its applicability: in its simplest form it pertains only to the wave equation (2), and works only in the one-dimensional case.[9] The method proceeds from the elementary observation that

$$\Box \equiv \tfrac{1}{u^2}\partial_t^2 - \partial_x^2 = \big(\tfrac{1}{u}\partial_t + \partial_x\big)\big(\tfrac{1}{u}\partial_t - \partial_x\big) \tag{14}$$

Clearly, if φ is killed by either of the operators $\frac{1}{u}\partial_t \pm \partial_x$ is will assuredly be killed by $\Box$. But

$$\begin{aligned}\big(\tfrac{1}{u}\partial_t + \partial_x\big)\varphi(x,t) = 0 \quad &\Longleftrightarrow \quad \varphi(x,t) = f(x-ut) \ :\ \text{any } f(\bullet)\\ \big(\tfrac{1}{u}\partial_t - \partial_x\big)\varphi(x,t) = 0 \quad &\Longleftrightarrow \quad \varphi(x,t) = g(x+ut) \ :\ \text{any } g(\bullet)\end{aligned}$$

so $\Box\varphi = 0$ is satisfied by all functions of the form

$$\varphi(x,t) = f(x-ut) + g(x+ut) \quad : \quad f(\bullet) \text{ and } g(\bullet) \text{ arbitrary} \tag{15}$$

and (less obviously) all solutions $\varphi(x,t)$ of $\Box\varphi = 0$ admit of such representation. Which is pretty remarkable when you consider that

- $f(x-ut)$ decribes a waveform sliding rigidly to the right with speed u;
- $g(x+ut)$ decribes a waveform sliding rigidly to the left with speed u.

[9] It was P. A. M. Dirac's relativity-motivated attempt to extend the method to three dimensions that led to the invention of the **Dirac equation**, which provides the foundation for so much of modern physics.

I reproduce now a pretty argument that was original to d'Alembert himself. Suppose, by way of introduction, that we possessed all the solutions $x(t)$ of a mechanical equation $m\ddot{x} = F(x)$. It would be standard practice for us to select the particular solution of interest by specifying initial data $x_0 = x(0)$ and $v_0 = \dot{x}(0)$. For example: in the ballistic case $m\ddot{x} = -mg$ we would have $x(t) = x_0 + v_0 t - \frac{1}{2}gt^2$. Proceeding now in that same spirit, suppose that the **initial wave data**

$$\varphi(x,0) \quad \text{and} \quad \varphi_t(x,0)$$

has been specified. Working from (15), we have

$$f(x) + g(x) = \varphi(x,0) \tag{16.1}$$

and $-uf'(x) + ug'(x) = \varphi_t(x,0)$, which after integration becomes

$$-f(x) + g(x) = \tfrac{1}{u}\int^x \varphi_t(y,0)\,dy \quad : \quad \text{lower limit arbitrary} \tag{16.2}$$

From (16) we obtain

$$f(x) = \tfrac{1}{2}\Big\{\varphi(x,0) - \tfrac{1}{u}\int^x \varphi_t(y,0)\,dy\Big\}$$
$$g(x) = \tfrac{1}{2}\Big\{\varphi(x,0) + \tfrac{1}{u}\int^x \varphi_t(y,0)\,dy\Big\}$$

whence

$$\varphi(x,t) = \tfrac{1}{2}\big\{\varphi(x-ut,0) + \varphi(x+ut,0)\big\} + \tfrac{1}{2u}\int_{x-ut}^{x+ut} \varphi_t(y,0)\,dy \tag{17}$$

the significance of which is developed in FIGURES 5.

Often we have interest in waves $\varphi(x,t)$ that conform to prescribed **boundary conditions**. Suppose, for example, that a vibrating string has been *clamped* at $x = 0$ and $x = a$, or that an air column vibrates within a pipe of length a that is closed at both ends. We would then insist that

$$\varphi(0,t) = \varphi(a,t) = 0 \quad : \quad \text{all } t \tag{18}$$

The first condition supplies (by (15)) $f(-ut)+g(ut) = 0$ (all t). The implication is that $g(z) = -f(-z)$, and therefore that

$$\varphi(x,t) = f(x-ut) - f(-x-ut) \tag{19}$$

The second condition therefore reads $f(a-ut) - f(-a-ut) = 0$, of which the implication is that $f(z) = f(z+2a)$: $f(\bullet)$ must be *periodic*, with period $2a$. It was first evident (by about 1806) to Joseph Fourier[10] (1768–1830) that the

[10] Consult http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Fourier.html for biographical information.

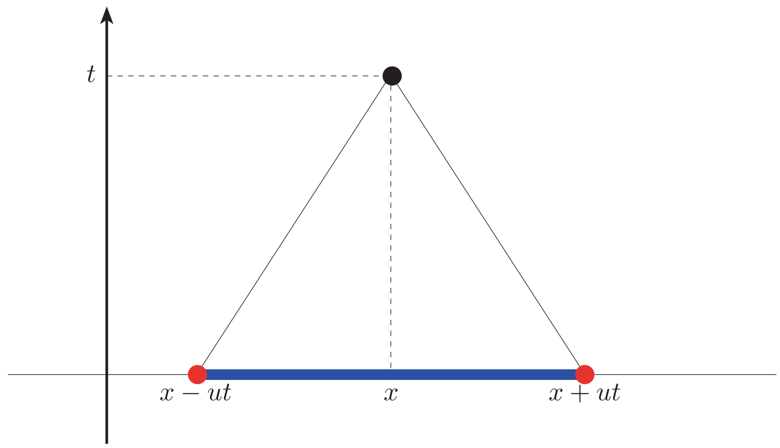


FIGURE 5A: *The • points, according to d'Alembert's equation* (17)*, contribute* $\varphi(x,0)$*-data to the valuation of* $\varphi(x,t)$*, while the* ▬ *points contribute* $\varphi_t(x,0)$*-data.*

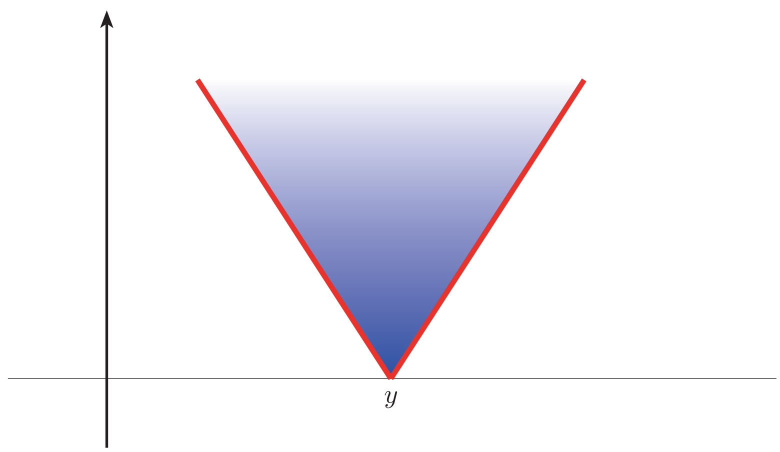


FIGURE 5B: *Complementary interpretation of d'Alembert's equation* (17)*: at spacetime points that lie on the* **red boundary** *of the wedge that extends forward from* $(y,0)$ *the wave function* $\varphi(x,t)$ *is sensitive to the initial datum* $\varphi(y,0)$*, while at points in the* **blue interior of the wedge** $\varphi(x,t)$ *is sensitive to* $\varphi_t(y,0)$*.*

most general such function can be developed

$$f(z) = \sum_n \Big\{ A_n \cos\big(n\pi \tfrac{z}{a}\big) + B_n \sin\big(n\pi \tfrac{z}{a}\big) \Big\}$$

Returning with this information to (19), we have

$$\begin{aligned}
\varphi(x,t) = \sum_n \Big\{ A_n \Big[&\cos\big(n\pi \frac{x-ut}{a}\big) - \cos\big(n\pi \frac{-x-ut}{a}\big) \Big] \\
+ B_n \Big[&\sin\big(n\pi \frac{x-ut}{a}\big) - \sin\big(n\pi \frac{-x-ut}{a}\big) \Big]
\end{aligned}$$

which according to *Mathematica* can be rewritten

$$\varphi(x,t) = \sum_{n=1}^{\infty} 2\big[A_n \sin \omega_n t + B_n \cos \omega_n t\big] \sin k_n x \tag{20.1}$$

with

$$k_n \equiv n\pi\tfrac{1}{a} \quad \text{and} \quad \omega_n \equiv n\pi\tfrac{u}{a} = uk_n \tag{20.2}$$

But (20)—see FIGURE 6—is physics familiar to every first-year student; its qualitative essentials were familiar already to (and exerted a powerful influence upon) Pythagoras, and lie at the tonal base of all music, both Western and Eastern.

If $\varphi(x,t)$ is clamped at $x = 0$ but *unclamped* at $x = a$ (a circumstance not feasible for stringed instruments, but natural to the design of most wind instruments) then we require of $\varphi(x,t)$ that

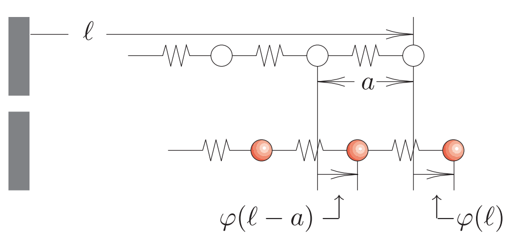
$$\left.\begin{aligned} \varphi(0,t) &= 0 \\ \varphi_x(a,t) &= 0 \end{aligned}\right\} \quad : \quad \text{all t} \tag{21}$$

—the latter condition arising by (9) from the physical requirement that the momentum flux at the free end be constantly zero.[11] The former condition leads as before to (19): $\varphi(x,t) = f(x - ut) - f(-x - ut)$, from which we obtain

$$\varphi_x(x,t) = f\,'(x - ut) + f\,'(-x - ut)$$

The second condition therefore supplies $f\,'(z) = -f\,'(z+2a) = (-)^2 f\,'(z+4a)$, from which we conclude that $f(\bullet)$ itself must be periodic with period $4a$.

[11] Or from the following line of argument: Reverting to the notation of §1 (at present a identifies an endpoint finitely removed from the origin, but in §1 ℓ played that role, while a referred to the "interatomic spacing"), look to the



dynamics of the particle at the unclamped end of a lattice. Immediately

$$\mu a \,\frac{\partial^2}{\partial t^2}\varphi(\ell,t) = k\big\{\varphi(\ell,t) - \varphi(\ell - a,t)\big\}$$

Divide by a, proceed to the limit $a \downarrow 0$ and note that—because $k = \kappa/a$ becomes infinite, a sensible result will be achieved if and only if $\varphi_x(\ell,t) = 0$ for all t. "The last two atoms move in synchrony, as though welded together."

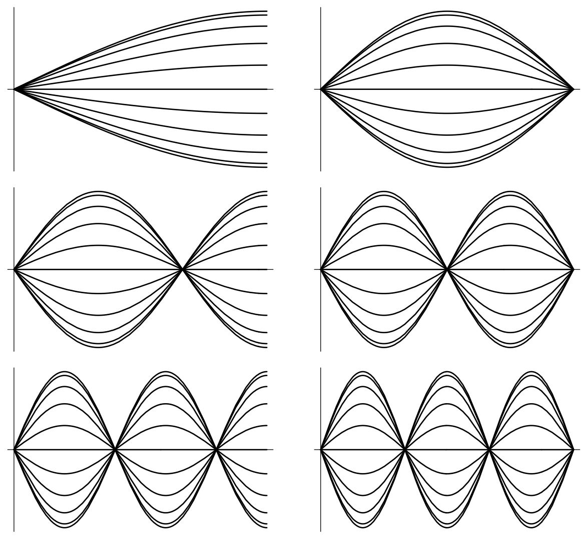


FIGURE 6: *Lowest-lying waveforms for a clamped/unclamped string* (*left column*) *and a clamped/clamped string* (*right column*). *Not only do the waveforms interdigitate: so do the vibrational frequencies, which according to* (20.2) *and* (22) *stand in the progression*

$$\begin{array}{ll} \omega_1 = \frac{1}{2}\cdot\pi\frac{u}{a} & \omega_2 = 1\cdot\pi\frac{u}{a} \\ \omega_3 = \frac{3}{2}\cdot\pi\frac{u}{a} & \omega_4 = 2\cdot\pi\frac{u}{a} \\ \omega_5 = \frac{5}{2}\cdot\pi\frac{u}{a} & \omega_6 = 3\cdot\pi\frac{u}{a} \\ \vdots & \vdots \end{array}$$

According to Fourier we can expect therefore to have

$$f(z) = \sum_n \Big\{ A_n \cos\big(n\pi\tfrac{z}{2a}\big) + B_n \sin\big(n\pi\tfrac{z}{2a}\big)\Big\}$$

so by (19)

$$\begin{aligned} \varphi(x,t) = \sum_n \Big\{ A_n\Big[\cos\big(n\pi\frac{x-ut}{2a}\big) - \cos\big(n\pi\frac{-x-ut}{2a}\big)\Big] \\ + B_n\Big[\sin\big(n\pi\frac{x-ut}{2a}\big) - \sin\big(n\pi\frac{-x-ut}{2a}\big)\Big] \end{aligned}$$

from which we are led again to (20.1), the difference being that now

$$k_n \equiv n\pi\tfrac{1}{2a} \quad\text{and}\quad \omega_n \equiv n\pi\tfrac{u}{2a} = uk_n \tag{22}$$

We recover *precisely* (20.1) when n is even (which places a node at $x = a$, where we want an anitnode), so in (22) it must be understood that $n = 1, 3, 5, \ldots$ The waveforms and spectra in the clamped/clamped and clamped/unclamped cases (for organ pipes the distinction would be between closed/closed and closed/open) *interdigitate*, as shown in FIGURE 6.

SOLUTION BY SEPARATION OF VARIABLES The simple wave equation

$$\Box\varphi = 0$$

presents a context that is insufficiently rich to expose many points of principle. We look now therefore to the Klein-Gordon equation

$$(\Box + \varkappa^2)\varphi = 0$$

which gives back the wave equation in the limit $\varkappa^2 \downarrow 0$.

Assume φ to possess the factored form $\varphi(x, t) = F(x)\cdot G(t)$. Then

$$(\Box + \varkappa^2)\varphi = F(x)\cdot \tfrac{1}{u^2}G''(t) - F(x)''\cdot G(t) + \varkappa^2 F(x)\cdot G(t) = 0$$

Divide by $F(x)\cdot G(t)$—possible at spacetime points where φ does not vanish—and get

$$\frac{G''(t)}{G(t)} = u^2\Big\{\frac{F''(x)}{F(x)} - \varkappa^2\Big\}$$

The only way a function of t can be identically equal to a function of x is for them to be separately equal to some constant:

$$\frac{G''(t)}{G(t)} = \alpha \tag{23.1}$$

$$u^2\Big\{\frac{F''(x)}{F(x)} - \varkappa^2\Big\} = \alpha \tag{23.2}$$

To avoid solutions of $G''(t) = \alpha G(t)$ that blow up asymptotically we insist that α be negative, which we emphasize by writing $\alpha = -\omega^2$. Then(23.1) gives

$$G(t) = G_0 \cos\big(\omega t + \delta_G\big)$$

and (23.2) becomes

$$\frac{F''(x)}{F(x)} = \underbrace{-(\omega/u)^2 + \varkappa^2}_{\substack{\text{must again be negative (call it } -k^2) \\ \text{to avoid asytmptotic blow-up}}}$$

which gives

$$F(x) = F_0 \cos\big(kx + \delta_F\big)$$

Thus are we led to particular solutions of the form

$$\begin{aligned}\varphi(x,t;\delta_F,\delta_G) &= \Phi_0\cos\big(kx+\delta_F\big)\cdot\cos\big(\omega t+\delta_G\big) \\ &= \tfrac{1}{2}\Phi_0\big[\cos\big(kx-\omega t+\delta_-\big)+\cos\big(kx+\omega t+\delta_+\big)\big]\end{aligned} \tag{24.1}$$

with $\Phi_0 = F_0G_0$, $\delta_\pm = \delta_F \pm \delta_G$. These **standing-wave solutions** can be taken in fairly obvious linear combination to construct right/left **running-wave solutions**

$$A\cos\big(kx-\omega t+\alpha\big) \quad \text{and} \quad B\cos\big(kx+\omega t+\beta\big) \tag{24.2}$$

of the K-G equation (10). In all cases, specification of k serves *via*

$$\omega^2 = u^2\big(k^2+\varkappa^2\big) \tag{25}$$

to determine the value of ω, to within a sign. In the limit $\varkappa^2 \downarrow 0$ we recover standing/running-wave solutions of the simple wave equation (2). The separation of variables method works but provides no particular advantage in the one-dimensional case discussed above, but in higher-dimensional cases it is often the method of choice.

$\boxed{\text{SOLUTION BY ANSATZ}}$ Assume

$$\varphi(x,t) = \text{real part of } Ze^{i(kx+\omega t)}$$

From $\big(\Box+\varkappa^2\big)e^{i(kx+\omega t)} = \big(-\frac{1}{u^2}\omega^2+k^2+\varkappa^2\big)e^{i(kx+\omega t)}$ we learn that $e^{i(kx+\omega t)}$ will satisfy the K-G equation if and only if ω and k satisfy the "dispersion relation" (25); *i.e.*, if and only if

$$\omega = \pm\omega(k) \quad \text{with} \quad \omega(k) \equiv u\sqrt{k^2+\varkappa^2} \tag{26}$$

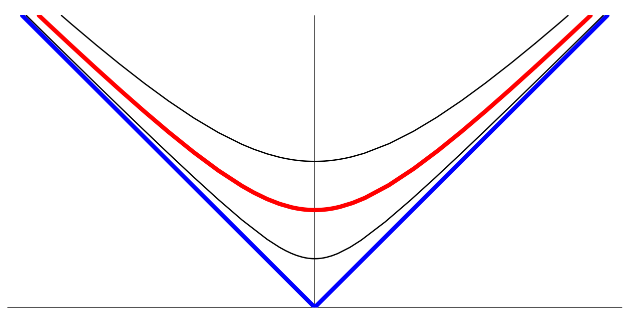


FIGURE 7: *Graph of $\omega(k)$. ω/u runs $\uparrow$, k runs $\rightarrow$. The hyperbolic curve intersects the ω/u-axis at $\varkappa$. In the figure $\varkappa$ runs through the values $\varkappa = 0, \frac{1}{2}, 1, \frac{3}{2}$.*

PROBLEM 1 : Assume the K-G system to be clamped at $x = 0$ and $x = a$. What then are the allowed values of ω?

Appealing now to the *linearity*, we conclude that every $\varphi(x,t)$ of the form

$$\varphi(x,t) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Big\{ \varPhi_+(k)\, e^{i[+\omega(k)t + kx]} + \varPhi_-(k)\, e^{i[-\omega(k)t + kx]} \Big\}\, dk \qquad (27.1)$$

is a solution of the K-G equation, and speculate (correctly, but are not yet in position to prove) that *every* solution admits of such representation. The expression on the right side of (27) will be real-valued if (and, as it turns out, only if)

$$[\varPhi_\pm(k)]^* = \varPhi_\mp(-k) \qquad (27.2)$$

It is to extract the juice from (27), and to open many other doors, that I enter now upon a mathematical digression:

ELEMENTS OF FOURIER ANALYSIS

Consider the set $\mathcal{F}$ of all (sufficiently nice) functions $f(x)$ with period a:

$$f(x+a) = f(x) \quad : \quad \text{all } x$$

Contained within $\mathcal{F}$ are, in particular, the elementary functions

$$\left.\begin{aligned} C_0(x\,;a) &= \sqrt{1/a} \\ C_n(x\,;a) &= \sqrt{2/a}\,\cos[2\pi n x/a] \quad : \quad n = 1,2,3,\ldots \\ S_n(x\,;a) &= \sqrt{2/a}\,\sin[2\pi n x/a] \quad : \quad n = 1,2,3,\ldots \end{aligned}\right\} \qquad (28)$$

which are (ask *Mathematica*) *orthonormal* in the sense[12]

$$\left.\begin{aligned} \int_0^a C_m(x)C_n(x)\,dx = \int_0^a S_m(x)S_n(x)\,dx &= \delta_{mn} \\ \int_0^a C_m(x)S_n(x)\,dx &= 0 \quad : \quad \text{all cases} \end{aligned}\right\} \qquad (29)$$

Every function of a single variable—whether periodic or not—can be written

$$f(x) = f_{\text{even}}(x) + f_{\text{odd}}(x)$$

with

$$\begin{aligned} f_{\text{even}}(x) &= \tfrac{1}{2}\big\{f(x) + f(-x)\big\} = +f_{\text{even}}(-x) \\ f_{\text{odd}}\,(x) &= \tfrac{1}{2}\big\{f(x) - f(-x)\big\} = -f_{\text{odd}}\,(-x) \end{aligned}$$

Functions of the form

$$f(x) = \sum_{n=0} f_n C_n(x) \qquad (30.1)$$

$$f_n = \int_0^a f(y)C_n(y)\,dy \qquad (30.2)$$

[12] Compare Chapter 1, page 9.

are manifestly even. It was Fourier who first asserted (correctly, as it turned out, given a suitable **theory of convergence**, the development of which required several decades) that *every* nice $f \in \mathcal{F}_{\text{even}}$ admits of such representation.[13] Similarly, for every nice $f \in \mathcal{F}_{\text{odd}}$ we have

$$f(x) = \sum_{n=1} f_n S_n(x) \tag{31.1}$$

$$f_n = \int_0^a f(y) S_n(y)\, dy \tag{31.2}$$

EXAMPLE: Let $f(x)$ be the even function that is defined by periodic continuation $f(x+a) = f(x)$ of

$$f(x) = (2x/a)^2 \quad : \quad -\tfrac{1}{2}a < x < +\tfrac{1}{2}a \tag{32.1}$$

Working from (30) we then have

$$f(x) = \tfrac{1}{3} + \sum_{n=1}^{\infty} (-)^n \frac{4\cos[2\pi n x/a]}{n^2\pi^2} \tag{32.2}$$

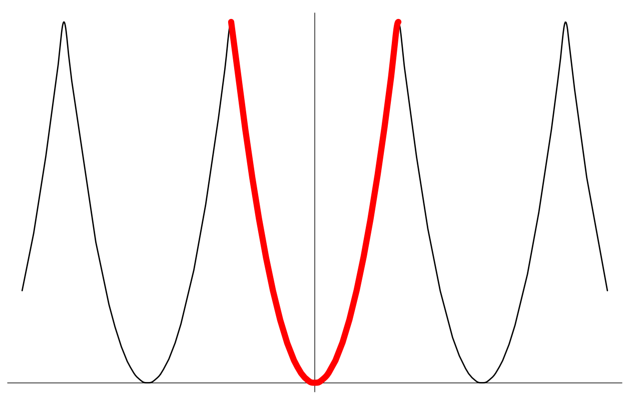


FIGURE 8: *Figure obtained from* (32.2) *in the case* $a = 2$*, with the sum truncated at* $n = 20$*. The central peaks stand at* $x = \pm 1$*, where* $f(\pm 1) = 1$*.*

EXAMPLE: Writing $\theta(x)$ to denote (see again Chapter 3, page 24) what *Mathematica* calls the `UnitStep` function, we write

$$\begin{aligned} g(x) = -\theta(x+3) + 2\theta(x+2) - 2\theta(x+1) + \theta(x) \\ - 2\theta(x-1) + 2\theta(x-2) - 2\theta(x-3) + \cdots \end{aligned} \tag{33.1}$$

---

[13] We see now why $\boldsymbol{a} = \sum_i (\boldsymbol{a}, \boldsymbol{e}_i)\boldsymbol{e}_i$ (Chapter 1, page 8) is called "Fourier's identity."

to describe the the central portion of a **square wave**:

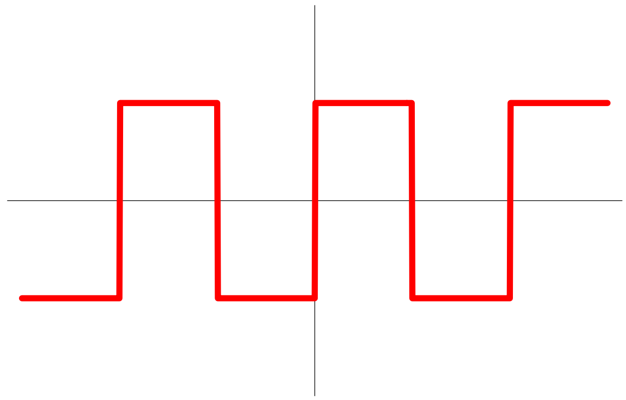


FIGURE 9: *Graph of the square wave $g(x)$ defined at* (33)*. This is an odd function of period $a=2$, bounded above and below by ±1.*

From

$$g_n = \int_0^2 g(y)S_n(y;2)\,dy = \begin{cases} 0 & : \quad n \text{ even} \\ 4/n\pi & : \quad n \text{ odd} \end{cases}$$

we obtain

$$g(x) = \tfrac{4}{\pi}\sum_{k=1}^{\infty}\tfrac{1}{2k-1}\sin[(2k-1)\pi x] \tag{33.2}$$

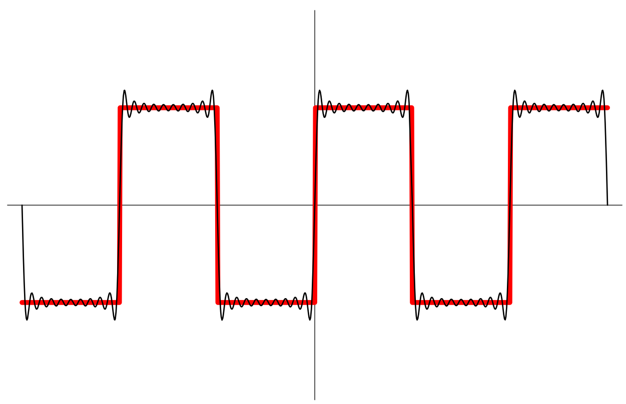


FIGURE 10: *Here superimposed upon the preceding figure is the graph that results from* (33.2) *when the sum is truncated at $k=10$. The confusion at the points of discontinuity is typical, and is known as* **Gibbs' phenomenon***.*

Translation $x \longmapsto x + \frac{1}{2}$ turns the odd square wave into an even squarewave, and turns the sine-terms into cosine-terms of alternating sign:

$$h(x) \equiv g(x + \tfrac{1}{2}) = \tfrac{4}{\pi} \sum_{k=1}^{\infty} (-)^{k-1} \tfrac{1}{2k-1} \cos[(2k-1)\pi x] \tag{33.3}$$

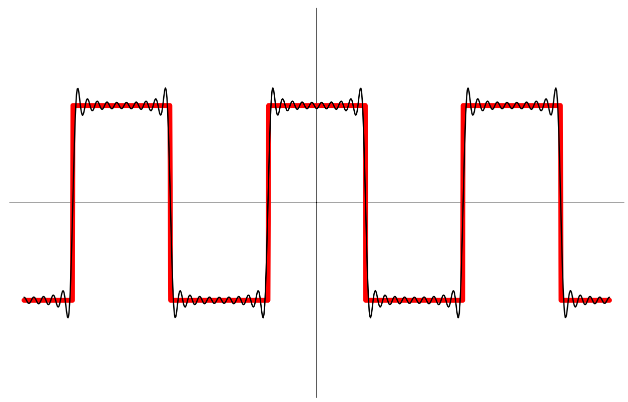


FIGURE 11: *Graph of the first* 10 *terms of the cosine series* (33.3), *superimposed upon a graph of the shifted square wave.*

PROBLEM 2: Let $f(x)$ be the odd **sawtooth** function that is defined by periodic continuation $f(x+1) = f(x)$ of

$$f(x) = 2x \quad : \quad -\tfrac{1}{2} < x < +\tfrac{1}{2}$$

Your objective is to display $f(x)$ as a Fourier sine series

$$f(x) = \sum_{n=1}^{\infty} f_n S_n(x;1)$$
$$f_n = \int_0^1 f(y) S_n(y;1)\,dy$$

To that end, compute $f_1, f_2, f_3, f_4, f_5, f_6, \ldots$ until you recognize the pattern, then plot (on $\{x, -2, 2\}$) the finite series you get by truncating the sum at 5, 10, 30. You will notice that Gibbs' phenomenon is again evident.

PROBLEM 3: You will now use the resources of *Mathematica* to accomplish the same objective. First plot `Round[x]`, $\{x, -2, 5\}$ to gain an understanding of what the command `Round[ ]` does. Then

```
Plot[2(x - Round[x]), {x, -2, 2}, AspectRatio→Automatic];
```

Now turn on the Fourier series package:

```
<<FourierSeries`
```

Now command

```
FourierSinCoefficient[2(x - Round[x]), x, n]
```

with $n = 1, 2, 3, 4, 5, 6, \ldots$ and compare your results with the f_n obtained "by hand" in the preceding exercise. Finally, command

```
FourierTrigSeries[2(x - Round[x]), x, 10]
```

and compare your result with that obtained previously. You might find it instructive at this point to look up `Fourier Series Package` in the *Mathematica* Documentation Center.

PROBLEM 4: Compare the results achieved by the commands

```
FourierTrigSeries[Cos[2πx]^8, x, 8] // Expand
```

and

```
Cos[2πx]^8 // TrigReduce // Expand
```

Finally, extract this same result from

$$\cos[2\pi x]^8 = \Big(\frac{e^{i2\pi x} + e^{-i2\pi x}}{2}\Big)^8$$

At the final step in the calculation use `ExpToTrig`.

From the real periodic functions introduced at (28) led us now assemble the complex-valued periodic functions

$$\left.\begin{aligned} E_0(x\,;a) &\equiv C_0(x\,;a) &&= \sqrt{1/a} \\ E_{\pm n}(x\,;a) &\equiv \frac{C_n(x\,;a) \pm iS_n(x\,;a)}{\sqrt{2}} &&= \sqrt{1/a}\,e^{\pm i\,2\pi n x/a} \end{aligned}\right\} \tag{34}$$

which, by (29), are orthonormal in the sense that[14]

$$\int_{-\frac{1}{2}a}^{+\frac{1}{2}a} \bar{E}_\mu(x\,;a)E_\nu(x\,;a)\,dx = \delta_{\mu\nu} \tag{35}$$

The implication is that periodic functions $f(x)$, whether complex-valued or not,

[14] Note that *the leading factor in the integrand has been conjugated*, as signaled by the $^-$. And that the subscripts now range on $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

can be developed

$$f(x)=\sum_{n=-\infty}^{\infty} f_n E_n(x\,;a) \tag{36.1}$$
$$f_n=\int_{-\frac{1}{2}a}^{+\frac{1}{2}a} f(y)\bar{E}_n(y\,;a)\,dy \tag{36.2}$$

and, incidentally, that

$$\int_{-\frac{1}{2}a}^{+\frac{1}{2}a} \bar{f}(x)g(x)\,dx=\sum_{n=-\infty}^{\infty}\bar{f}_n g_n \tag{38}$$

PROBLEM 5: Plot `Abs[Sin[`πx`]` on the interval $0<x<5$. Develop that function as a Fourier series. The result you have obtained suggests pretty conclusively that

$$|\sin(\pi x)|=\frac{2}{\pi}-\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[2n\pi x]}{4n^2-1}$$

Setting first $x=0$, then $x=\frac{1}{2}$, deduce values of

$$\sum_{n=1}^{\infty}\frac{1}{4n^2-1}\quad\text{and}\quad\sum_{n=1}^{\infty}(-)^n\frac{1}{4n^2-1}$$

Add those results to obtain an evaluation of

$$\sum_{n=1}^{\infty}\frac{1}{16n^2-1}$$

The methods of Fourier analysis can be extended to **aperiodic functions** if, by inspired trickery, we consider such functions to be "periodic" on $\left[-\frac{1}{2}a,+\frac{1}{2}a\right]$ with $a\uparrow\infty$. At (36) we had

$$f(x)=\sum_{n=-\infty}^{\infty} f_n\sqrt{1/a}\,e^{i\,2\pi nx/a} \tag{36.1}$$
$$f_n=\int_{-\frac{1}{2}a}^{+\frac{1}{2}a} f(y)\sqrt{1/a}\,e^{-i\,2\pi ny/a}\,dy \tag{36.2}$$

which can be combined to give (here $b = \frac{1}{2}a$)

$$f(x) = \sum_{n=-\infty}^{\infty} \tfrac{1}{2b}\int_{-b}^{+b} f(y)\, e^{i\pi(x-y)n/b}\, dy$$

As $b \uparrow \infty$ we bring into play a device familiar from the "lattice refinement procedure" encountered in §1: we set $n/b = \ell$, $\Delta n/b = 1/b = d\ell$ and obtain

$$\downarrow$$

$$= \tfrac{1}{2}\int_{-\infty}^{+\infty} d\ell \int_{-\infty}^{+\infty} f(y)\, e^{i\pi\ell(x-y)}\, dy$$

which with a change of variables $\ell \longmapsto k = \pi\ell$ becomes the **Fourier Integral theorem**

$$= \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} dk \int_{-\infty}^{+\infty} f(y)\, e^{ik(x-y)}\, dy \tag{37}$$

The Fourier integral theorem—which is susceptible to direct and more convincing proof[15]—can be formulated in several useful alternative ways. If we write

$$g(k) \equiv \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(y)\, e^{-iky}\, dy \;\;:\;\; \textbf{Fourier transform} \text{ of } f(x) \tag{38.1}$$

then (37) becomes

$$f(x) \equiv \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(k)\, e^{+ikx}\, dk \tag{38.2}$$

EXAMPLE: The functions

$$f(x\,; a) \equiv \frac{\texttt{Sign}[a-x] + \texttt{Sign}[a+x]}{4a}$$

are "box functions." Each is centered at the origin and envelops unit area; $f(x, a)$ gets <u>wider/narrower</u> according as the positive parameter a gets larger/smaller (see the following figure). The command

$$\texttt{FourierTransform}[f(x\,;a), x, k]$$

supplies

$$g(k\,; a) = \frac{\sin ak}{ak\sqrt{2\pi}}$$

which gets <u>narrower/wider</u> according as a gets larger/smaller. That

[15] See, for example, §6.1 on pages 77–79 of R. Courant & D. Hilbert, *Methods of Mathematical Physics*: *Volume One* (English translation 1953).

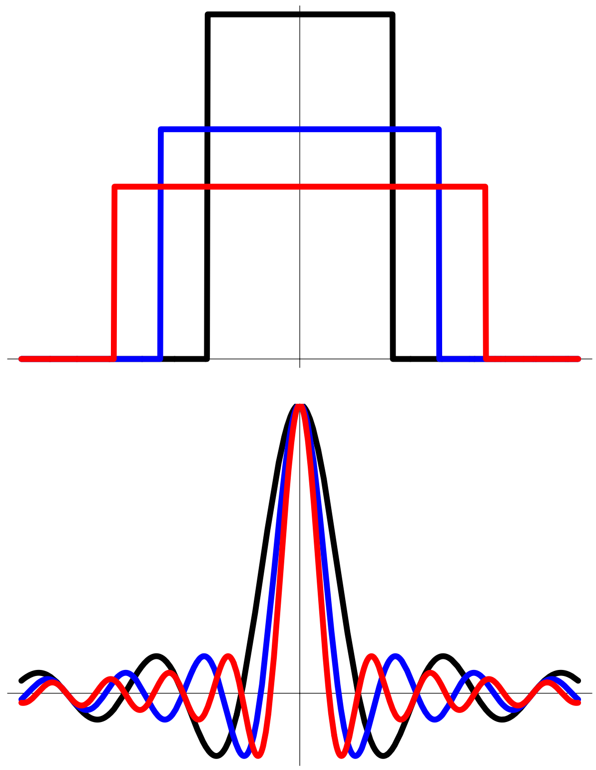


FIGURE 12: *Shown above*: *three box functions $f(x;a)$ that progress* **black/ blue / red** *as a progresses through the values $1, \frac{3}{2}, 2$. Shown below*: *the Fourier transforms of those functions.*

contrariwise tendency is universally characteristic of the relation between functions and their Fourier transforms, and in a quantum mechanical context accounts for the **Heisenberg uncertainty principle**. For discussion of some typical engineering applications (mainly to image processing) of the Fourier transform see, for example,

http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html

http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Contents.html

http://www.engr.trinity.edu/~paul/fourier/fourier/node3.html

If, in (37), we reverse the order of integration (*i.e.*, if we do what we can "while waiting for $f(y)$ to be specified") we obtain

$$f(x) = \int_{-\infty}^{+\infty} f(y) \underbrace{\left\{ \tfrac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x-y)} dk \right\}}_{\text{provides a representation of } \delta(y-x)} dy \tag{39}$$

and it is to provide the pretty interpretation of this result that I digress now to review the

ESSENTIALS OF THE THEORY OF COMPLEX VECTOR SPACES

A vector space $\mathcal{V}_n$ becomes a complex vector space when considered to be closed under multiplication by complex scalars:

$$\boldsymbol{a} \in \mathcal{V}_n \Longrightarrow \lambda\boldsymbol{a} \in \mathcal{V}_n \quad : \quad \lambda \text{ complex}$$

$\mathcal{V}_n$ becomes a *complex inner product space* when endowed with a (complex) number-valued "inner product" $(\boldsymbol{a}, \boldsymbol{b})$ possessing the following properties:

- LINEARITY: $(\boldsymbol{a}, \lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2) = \lambda_1(\boldsymbol{a}, \boldsymbol{b}_1) + \lambda_2(\boldsymbol{a}, \boldsymbol{b}_2)$
- HERMITIAN SYMMETRY: $(\boldsymbol{a}, \boldsymbol{b}) = \overline{(\boldsymbol{b}, \boldsymbol{a})}$

  NOTE: I use $\overline{}$ or * to signify complex conjugation, which ever seems to work best in the context at hand.

- POSITIVE DEFINITENESS: $(\boldsymbol{a}, \boldsymbol{a})$ is automatically real (since equal to its own conjugate). We require that $(\boldsymbol{a}, \boldsymbol{a}) \geqslant 0$ for all $\boldsymbol{a} \in \mathcal{V}_n$, with equality if and only if $\boldsymbol{a} = \boldsymbol{0}$.

Develop $\boldsymbol{a}$ and $\boldsymbol{b}$ with respect to an arbitrary basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$:

$$\boldsymbol{a} = \textstyle\sum_i a^i \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{b} = \sum_j b^j \boldsymbol{e}_j$$

Then $(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i,j} \bar{a}^i (\boldsymbol{e}_i, \boldsymbol{e}_j) b^j$, the specific meaning of which hinges upon our having assigned values to the basic inner products

$$h_{ij} = (\boldsymbol{e}_i, \boldsymbol{e}_j) = \bar{h}_{ji}$$

If we interpret the h_{ij} to be elements of an $n \times n$ matrix, then that matrix $\mathbb{H} = \|h_{ij}\|$ must (if we are to achieve $(\boldsymbol{a}, \boldsymbol{b}) = \overline{(\boldsymbol{b}, \boldsymbol{a})}$) be "**hermitian**," meaning <u>equal to its own conjugate transpose</u>: $\mathbb{H} = \mathbb{H}^{\mathsf{t}}$, where $^{\mathsf{t}}$ signifies $^{*\mathsf{T}}$. Arguments quite similar to those presented on pages 27–28 in Chapter1 establish that

- the eigenvalues of hermitian matrices are invariably *real*;
- eigenvectors associated with distinct eigenvalues are invariably *orthogonal*.

Positive definiteness of $(\boldsymbol{a}, \boldsymbol{a})$ requires that all the eigenvalues of $\mathbb{H}$ be positive.

It is usually advantageous to work with bases that are *orthonormal*: $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$, which in matrix notation reads

$$\boldsymbol{e}_i^{\mathsf{t}} \boldsymbol{e}_j = \delta_{ij} \tag{40.1}$$

Fourier's identity (see again page 8 in Chapter 1) $\boldsymbol{a} = \sum_k (\boldsymbol{a}, \boldsymbol{e}_k)\boldsymbol{e}_k$

can in matrix notation be written $\boldsymbol{a} = \sum_k \boldsymbol{e}_k \boldsymbol{e}_k^{\mathsf{t}} \boldsymbol{a}$ which, since valid for *all* $\boldsymbol{a} \in \mathcal{V}_n$, amounts to a statement of the "*completeness relation*"

$$\sum_k \boldsymbol{e}_k \boldsymbol{e}_k^{\mathsf{t}} = \mathbb{I} \tag{40.2}$$

With those ideas in mind, we construe "nice functions" $f(x)$ to be elements of an **infinite-dimensional complex vector space** $\mathcal{V}_\infty$, and the functions

$$e(x\,;k) = \tfrac{1}{\sqrt{2\pi}}\, e^{ikx} \quad : \quad -\infty < k < +\infty \tag{41}$$

to be the continuously indexed (!) elements of a basis in $\mathcal{V}_\infty$. At (39) we had

$$\tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ik(x-y)} dk = \delta(x-y) \tag{42}$$

which when written

$$\int_{-\infty}^{+\infty} e(x\,;k)\,\overline{e(y\,;k)}\,dk = \delta(y-x) \tag{43.1}$$

becomes a **continuous analog of the completeness relation**. The Fourier integral theorem (37/39), viewed in this light, emerges as the infinite-dimensional analog of "Fourier's identity." From (42) it follows also, by notatonal adjustment, that

$$\tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ix(-r+s)} dx = \delta(r-s)$$

which when written

$$\int_{-\infty}^{+\infty} \overline{e(x\,;r)}\,e(x\,;s)\,dx = \delta(y-x) \tag{43.2}$$

becomes a **continuous analog of the orthonormality condition**.

The point I have sought to establish (in remarks that began on page 17, and that might on another occasion be pursued in book-length detail) is that *the architecture of Fourier analysis is identical to that of the theory of vector spaces* (complex inner product spaces), of which it provides a natural extension (and the invention of which it, in point of historical fact, preceded, and served in part to motivate).

Returning now to (27.1), we have

$$\varphi(x,0) = \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \Big\{\varPhi_+(k) + \varPhi_-(k)\Big\}e^{ikx}\,dk$$

$$\varphi_t(x,0) = \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} i\omega(k)\Big\{\varPhi_+(k) - \varPhi_-(k)\Big\}e^{ikx}\,dk$$

By inverse Fourier transformation[16]

$$\varPhi_+(k) + \varPhi_-(k) = \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \varphi\,(y,0)e^{-iky}\,dy$$

$$\varPhi_+(k) - \varPhi_-(k) = \tfrac{1}{i\omega(k)}\tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \varphi_t(y,0)e^{-iky}\,dy$$

whence

$$\varPhi_\pm(k) = \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \tfrac{1}{2}\Big[\varphi(y,0) \pm \tfrac{1}{i\omega(k)}\varphi_t(y,0)\Big]e^{-iky}dy \tag{44}$$

This equation describes $\varPhi_\pm(k)$ in terms of the initial data $\varphi(x,0)$, and $\varphi_t(x,0)$, which we are free to prescribe as we will. If we insert (44) into (27.1) and reverse the order of integration we obtain a result that can be notated[17]

$$\varphi(x,t) = \int_{-\infty}^{+\infty} \Big\{\varphi(y,0)G_t(x-y,t) + \varphi_t(y,0)G(x-y,t)\Big\}\,dy \tag{45}$$

with

$$G(x-y,t) \equiv \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{i[+\omega(k)t+k(x-y)]} - e^{i[-\omega(k)t+k(x-y)]}}{2i\omega(k)}\,dk \tag{46.1}$$

$$= \tfrac{1}{\pi u}\int_0^{\infty} \frac{\sin\sqrt{k^2+\varkappa^2}\,ut}{\sqrt{k^2+\varkappa^2}}\cos k(x-y)\,dk \tag{46.2}$$

These results are less complicated than they look. It is, for example, clear from (46.1) that G—whence also G_t—is itself a solution of the K-G equation

$$\big(\tfrac{1}{u^2}\partial_t^2 - \partial_x^2 + \varkappa^2\big)G(x-y,t) = \big(\tfrac{1}{u^2}\partial_t^2 - \partial_x^2 + \varkappa^2\big)G_t(x-y,t) = 0$$

---

[16] Though we will have many other occasions to dreaw upon that fundamental material, this is the step that motivated the mathematical digression that began on page 17.

[17] This is structurally reminiscent of the statement that for a simple oscillator

$$x(t) = x(0)\,G_t(t) + \dot{x}(0)\,G(t) \quad\text{with}\quad G(t) \equiv \frac{\sin\omega t}{\omega}$$

and that it is a solution with these distinguishing initial-value properties:

$$G_t(x-y,0) = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ik(x-y)}\,dk = \delta(x-y)$$
$$G(x-y,0) = G_{tt}(x-y,0) = 0$$

Returning with this information to (45), we see that $\varphi(x,t)$ is in fact a solution of the K-G equation that conforms to the prescribed initial conditions. According to (46.2), $G(x-y,t)$ is the cosine transform of $\big[\pi u\sqrt{k^2+\varkappa^2}\,\big]^{-1}\sin\sqrt{k^2+\varkappa^2}\,ut$. Consulting the tables,[18] we find that

$$G(x-y,t) = \begin{cases} \pm\,\frac{1}{2u}J_0\big(\varkappa\sqrt{(ut)^2-(x-y)^2}\,\big) \text{ if } (ut)^2-(x-y)^2 \geqslant 0 \\[2ex] 0 \text{ otherwise, } i.e.\text{, if } y \text{ lies farther from } x \text{ than a signal,} \\ \text{propagating at speed } u\text{, could travel in time } t \end{cases} \tag{46.3}$$

where the sign is $\pm$ according as $t>0$ or $t<0$., and where $J_0(\bullet)$ refers to the 0^{th}-order **Bessel function**. To describe $G_t(x-y,t)$ in similar detail we need only to know that $J_0'(z) = -J_1(z)$.

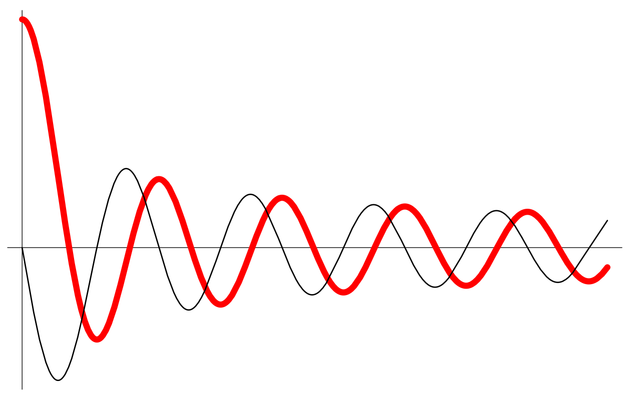


FIGURE 13: *Graphs of* $J_0(z)$*—shown in* **red***—and of* $J_0'(z) = -J_1(z)$.

If one introduces (46.3) into (45) and proceeds to the limit $\varkappa\downarrow 0$ it is not difficult to recover d'Alembert's equation (17), though I will not belabor the details. Here as there, only initial data interior to the wedge (cone) that extends backward from (x,t) contributes to the value of $\varphi(x,t)$ (see again FIGURE 5A), but when $\varkappa \neq 0$ the functions G and G_t serve to *modulate the weighting of the contributory initial data.*

The result summarized in (45/46) is physically quite striking and, though analytically non-trivial, was obtained by an argument that is fundamentally quite simple—an argument that hinges essentially upon the principle of

[18] See Erdélyi *et al*, *Tables of Integral Transforms*, Volume I, page 26, **30** (19*xx*). *Mathematica* appears on this occasion to be of no assistance.

superposition, and that is fairly typical of the application of Fourier analysis to physical problems.

5. Motion of waves and wave packets. First some points of terminology: the K-G wave $e^{i[\omega(k)t-kx]}$ contains but a single (k-dependent) frequency, and is said on that account to be "monochromatic." The "phase" of the wave is defined by the equation

$$e^{i[\omega(k)t-kx]} = e^{i(\textsf{\textbf{phase}})}$$

The point where the phase assumes a specified constant value moves along with a **phase velocity** u which by

$$\tfrac{d}{dt}(\text{phase}) = \omega(k) - k\dot{x} = 0$$

has the value

$$v = \frac{\omega(k)}{k} \tag{47}$$

The (angular) frequency ω has dimension $[\omega] = T^{-1}$ while the "wave number" k has dimension $[k] = L^{-1}$, so it is indeed the case that v is a "velocity": $[v] = L/T$. For monochromatic K-G waves

$$v = \frac{u\sqrt{k^2+\varkappa^2}}{k} \tag{48}$$

Evidently

- $v > u$ for all k if $\varkappa \neq 0$;
- $v \downarrow u$ as $k \uparrow \infty$ if $\varkappa \neq 0$;
- $v = u$ for all k if $\varkappa = 0$.

Look now to the <u>superposition of two</u> right-running K-G waves

$$\varphi(x,t) = \cos\big[ut\sqrt{k^2+\kappa^2} - kx\big] + \cos\big[ut\sqrt{(k+\Delta k)^2+\kappa^2} - (k+\Delta k)x\big]$$

with nearly identical wave numbers. The `TrigFactor` command supplies

$$\begin{aligned} &= 2\cos\Big[\tfrac{1}{2}ut\Big(\sqrt{(k+\Delta k)^2+\kappa^2} - \sqrt{k^2+\kappa^2}\Big) - \tfrac{1}{2}(\Delta k)x\Big] \\ &\quad\cdot\cos\Big[\tfrac{1}{2}ut\Big(\sqrt{(k+\Delta k)^2+\kappa^2} + \sqrt{k^2+\kappa^2}\Big)\Big] - \tfrac{1}{2}(\Delta k)x - kx\Big] \end{aligned}$$

Expand the expressions [*etc.*] in powers of Δk, abandon all but the leading terms and obtain

$$\varphi(x,t) \approx 2\cos\Big[\tfrac{1}{2}\Big(\frac{uk}{\sqrt{k^2+\kappa^2}}\,t - x\Big)\Delta k\Big]\cdot\cos\Big[ut\sqrt{k^2+\kappa^2} - kx\Big] \tag{49}$$

The second factor slides rigidly to the right with the phase velocity

$$v_{\text{phase}} = \frac{u\sqrt{k^2+\varkappa^2}}{k} = \omega/k$$

while the leading factor slides rigidly to the right with what we will acquire reason to call the **group velocity**

$$v_{\text{group}} = \frac{uk}{\sqrt{k^2+\varkappa^2}} = u^2k/\omega \tag{50}$$

Evidently

- $v_{\text{group}} < u$ for all k if $\varkappa \neq 0$;
- $v_{\text{group}} \uparrow u$ as $k \uparrow \infty$ if $\varkappa \neq 0$;
- $v_{\text{group}} = u = v_{\text{phase}}$ for all k if $\varkappa = 0$.

Moreover

$$v_{\text{group}} \cdot v_{\text{phase}} = u^2 \quad : \quad \text{all } k \tag{51}$$

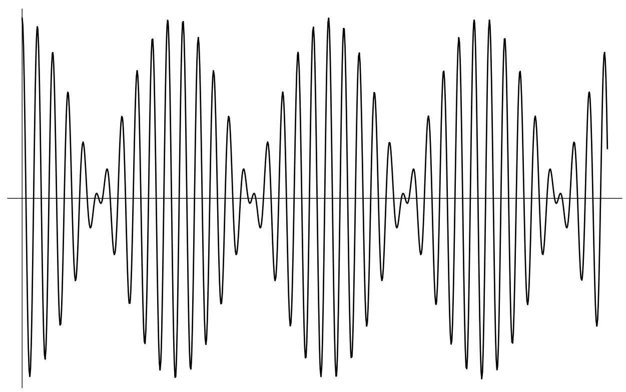


FIGURE 14: *Frame from a filmstrip based upon* (49)*, which refers in effect to the "beat" phenomenon in a dispersive medium* (*and more specifically to the beats produced by two K-G waves of nearly the same frequency*)*. The envelope advances with speed v_{group}, while the wavetrain rushes forward with speed $v_{\text{phase}} > v_{\text{group}}$.*

Two-wave superposition is, however, too simple to expose some of the most salient aspects of the situation. Suppose we are sitting at the origin and see the passing wave to have the form (see FIGURE 15) of a **slowly modulated carrier wave** of frequency ω_0:

$$\varphi(0,t) = A(t) \cdot \cos \omega_0 t$$
$$A(t) = \int_0^{\Omega} a(\nu) \cos \nu t \, d\nu \quad : \quad \Omega \ll \omega_0$$

Here $a(\nu)$ is the **spectrum of the signal**, and it is the condition $\Omega \ll \omega_0$ that expresses the stipulation that the modulation is "slow;" *i.e.*, that the frequencies present in the signal are all much lower than the frequency of the carrier. Immediately

$$\begin{aligned}
\varphi(0,t) &= \int_0^{\Omega} \tfrac{1}{2} a(\nu) \Big[\cos[(\omega_0 - \nu)t] + \cos[(\omega_0 + \nu)t] \Big] d\nu \\
&= \int_{\omega_0 - \Omega}^{\omega_0} \tfrac{1}{2} a(\omega_0 - \omega) \cos \omega t \, d\omega + \int_{\omega_0}^{\omega_0 + \Omega} \tfrac{1}{2} a(\omega - \omega_0) \cos \omega t \, d\omega \\
&= \int_{\omega_0 - \Omega}^{\omega_0 + \Omega} \Phi(\omega) \cos \omega t \, d\omega
\end{aligned} \tag{52}$$

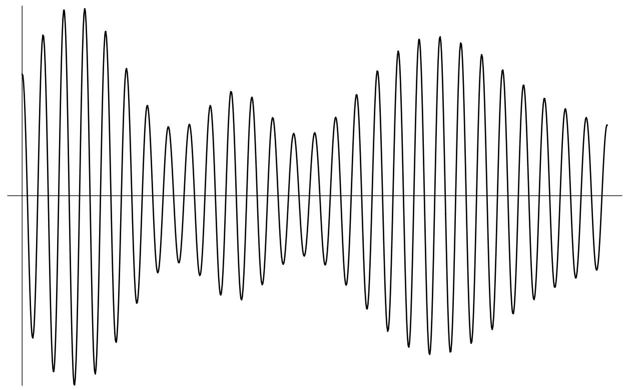


FIGURE 15: *A modulated carrier wave.*

where $\varPhi(\omega)$ is assembled from back-to-back copies of $a(\nu)$, centered on ω_0: $\varPhi(\omega_0 \pm \nu) = \frac{1}{2}a(\nu)$.

EXAMPLE: Consider the case in which $\underline{a(\nu)\text{ is constant}}$ (call the constant $= A/\Omega$) for $0 \leqslant \omega \leqslant \Omega$, and vanishes elsewhere. Then

$$A(t) = (A/\Omega)\int_0^\Omega \cos\nu t\,d\nu = A\frac{\sin\Omega t}{\Omega t}$$

and

$$\varphi(0,t) = A\frac{\sin\Omega t}{\Omega t}\cdot\cos\omega_0 t = \tfrac{1}{2}(A/\Omega)\int_{\omega_0-\Omega}^{\omega_0+\Omega}\cos\omega t\,d\omega \tag{53}$$

An observer stationed downstream at x would write

$$\varphi(x,t) = \tfrac{1}{2}(A/\Omega)\int_{\omega_0-\Omega}^{\omega_0+\Omega}\cos\big[\omega t - k(\omega)x\big]\,d\omega \tag{54}$$

$$k(\omega) = \sqrt{(\omega/u)^2 - \varkappa^2}$$

The integral, even in this simplest of cases, is intractable, but can be made to yield useful information if approached circumspectly. Expanding in powers of $(\omega - \omega_0)$—a variable which will remain "small" if the bandwidth is narrow—we have

$$\omega t - k(\omega)x = P_0 + P_1(\omega-\omega_0) + P_2(\omega-\omega_0)^2 + \cdots \tag{55}$$

with

$$\begin{aligned}
P_0 &= \omega_0 t - k_0 x\\
P_1 &= t - x/v_g\\
P_2 &= +\tfrac{1}{2}\Big(\frac{1}{v_g^2 k_0} - \frac{1}{u^2 k_0}\Big)x\\
P_3 &= -\tfrac{1}{2}\Big(\frac{1}{v_g^2 k_0} - \frac{1}{u^2 k_0}\Big)\frac{1}{v_g k_0}x\\
&\vdots
\end{aligned}$$

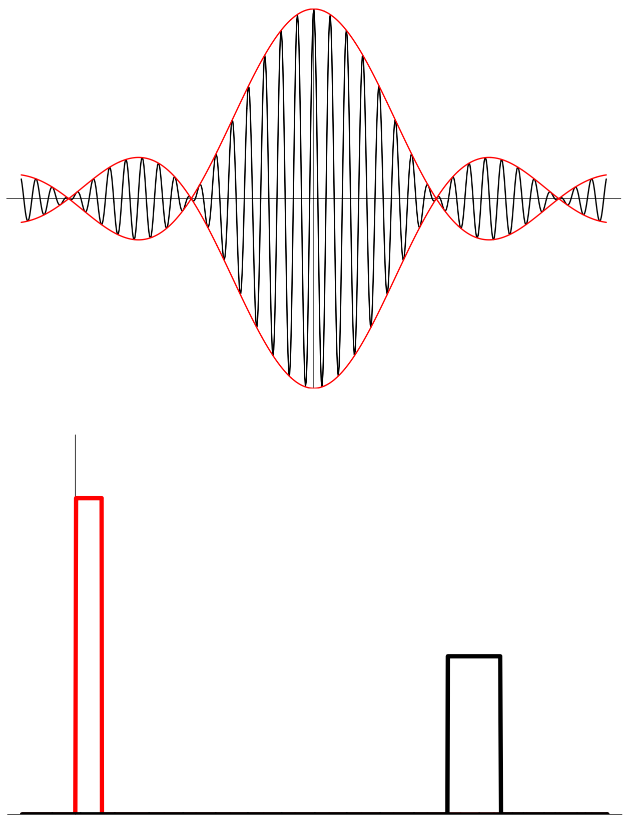


FIGURE 16: *Shown above*: *the modulated carrier—or "signal"—encountered at* (52). *Shown below*: *the Fourier transform* (*in* **red**) *of the "modulation function"* $A(t)$, *and* (*in* **black**) *the Fourier transform of the modulated carrier. The former has frequencies distributed between* 0 *and* Ω, *the latter has frequencies distributed between* $\omega_0 - \Omega$ *and* $\omega_0 + \Omega$; *has, that is to say, a* **bandwidth** *of* 2Ω.

where $k_0 = k(\omega_0)$ and $v_g = u^2 k_0/\omega_0$. If we abandon terms of order $n \geqslant 2$ we find

$$\begin{aligned}\varphi(x,t) &\approx \tfrac{1}{2}(A/\Omega)\int_{\omega_0-\Omega}^{\omega_0+\Omega}\cos\big[P_0 + P_1(\omega-\omega_0)\big]\,d\omega\\ &= A\,\frac{\sin\Omega P_1}{\Omega P_1}\cos P_0\\ &= A\,\frac{\sin\Omega(t-x/v_g)}{\Omega(t-x/v_g)}\cdot\cos\omega_0(t-x/v_p)\end{aligned}\tag{56}$$

where $v_p = \omega_0/k_0$ is the phase velocity of the carrier wave. In this approximation the the envelope, or "signal," is seen to glide rigidly along at the group velocity v_g, and to be written onto a carrier that also moves rigidly, but at the higher velocity v_p. In the limit $\varkappa \downarrow 0$ the velocities v_g and v_p become mutually equal to the wave velocity u, and the whole construction (carrier + modulation) moves rigidly. The downstream observer sees a *delayed but undistorted copy* of the signal we detected at the $x = 0$.

If in (55) we retain the second-order term but abandon terms of order $n \geqslant 3$, writing

$$\varphi(x,t) \approx \tfrac{1}{2}(A/\Omega)\int_{\omega_0-\Omega}^{\omega_0+\Omega} \cos\big[P_0 + P_1(\omega-\omega_0) + P_2(\omega-\omega_0)^2\,\big]\,d\omega$$

we find the integral still to be analytically tractable, but to involve the functions[19]

$$\begin{aligned}\texttt{FresnelC[}z\texttt{]} &= C(z) \equiv \int_0^z \cos\big(\tfrac{\pi}{2}u^2\big)du\\ \texttt{FresnelS[}z\texttt{]} &= S(z) \equiv \int_0^z \sin\big(\tfrac{\pi}{2}u^2\big)du\end{aligned}$$

It would serve no useful purpose to spell out the intricate details of the result to which the integral leads, since those can be reproduced at the pleasure of the reader who follows the these steps:

- Define

$$\begin{aligned}k[\omega\_,\varkappa\_,u\_] &:= \sqrt{(\omega/u)^2 - \varkappa^2}\\ vg[\omega\_,\varkappa\_,u\_] &:= u^2 k[\omega,\varkappa,u]/\omega\\ P_0[\omega\_,\varkappa\_,u\_,x\_,t\_] &:= \omega t - k[\omega,\varkappa,u]x\\ P_1[\omega\_,\varkappa\_,u\_,x\_,t\_] &:= t - \frac{x}{vg[\omega,\varkappa,u]}\\ P_2[\omega\_,\varkappa\_,u\_,x\_,t\_] &:= -\Big(\frac{1}{vg[\omega,\varkappa,u]} - \frac{1}{u^2}\Big)\frac{x}{2k[\omega,\varkappa,u]}\end{aligned}$$

- Return with this information to the integrand.
- Assign numeric values to $\omega = \omega_0$, $\varkappa$, u and x.
- Integrate, and plot the result.

To produce FIGURE 17 I set $\omega_0 = 15$ and $u = \Omega = 1$ (the same values as were used to construct FIGURE 16). Additionally, I set $\varkappa = 10$ and placed my signal detectors at $x = 20, 60, 100$. When the

[19] See Abramowitz & Stegun, §§7.3 & 7.4, pages 300–304. The Fresnel integrals are famous for their occurance in optical diffraction theory, but they are encountered also in many contexts that have nothing to do with diffraction. Readers are encouraged to take a moment to `Plot` $C(z)$ and $S(z)$.

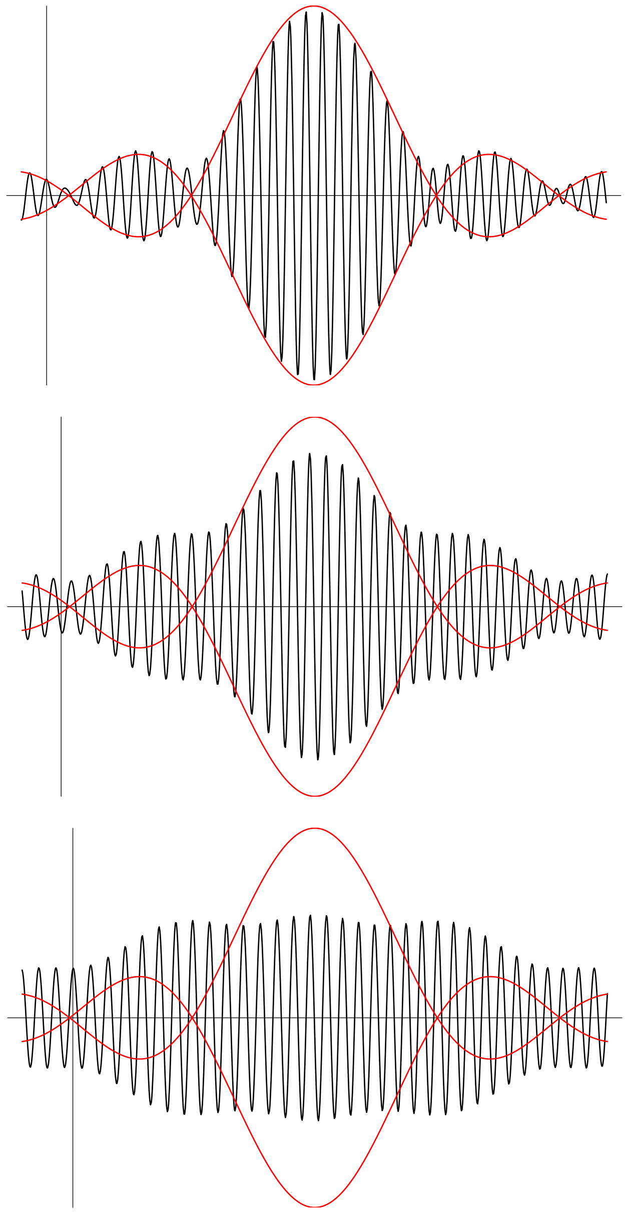


FIGURE 17: *Evolved descendents* $\varphi(20,t)$, $\varphi(60,t)$ *and* $\varphi(100,t)$ *of the signal* $\varphi(0,t)$ *shown in* FIGURE 16. *The top signal is centered at* $t_c = 26.85$, *the middle figure at* $t_c = 80.50$, *the bottom figure at* $t_c = 134.2$. *In all cases (as also in* FIGURE 16), *time runs from* $t_c - 7.5$ *to* $t_c + 7.5$. *The envelope of the signal* $\varphi(0,t)$ *has in each case been superimposed to make clear the effect of progressive* **signal dispersion**.

parameters are set to those values we expect the carrier wave to advance with speed

$$v_{\text{phase}} = vp[15, 10, 1] = 3/\sqrt{5} = 1.34164$$

and the signal to advance with speed

$$v_{\text{group}} = vg[15, 10, 1] = \sqrt{5}/3 = 0.74536$$

and on the latter basis expect to have

$$\begin{aligned} t_c[20] &= \frac{20}{0.74536} = 26.83 \ \ : \ \ \text{compare the observed } 26.85 \\ t_c[60] &= \frac{60}{0.74536} = 80.50 \ \ : \ \ \text{compare the observed } 80.50 \\ t_c[100] &= \frac{100}{0.74536} = 134.2 \ \ : \ \ \text{compare the observed } 134.2 \end{aligned}$$

It is the $\varkappa$-factor which is responsible for the dispersion of K-G wave packets. In the case $\varkappa = 0$ one has $v_g = u$, with the consequence (see again page 31) that $P_n = 0$ for $n \geqslant 2$: all the waves that constitute a signal advance at the same speed: the signal has become a locked, nondispersive unit.

The literature seems not to supply many examples of signal dispersion that allow of development in closed-form analytical detail, but even in their absence the physical essence of the phenomenon in question is seems clear: certainly

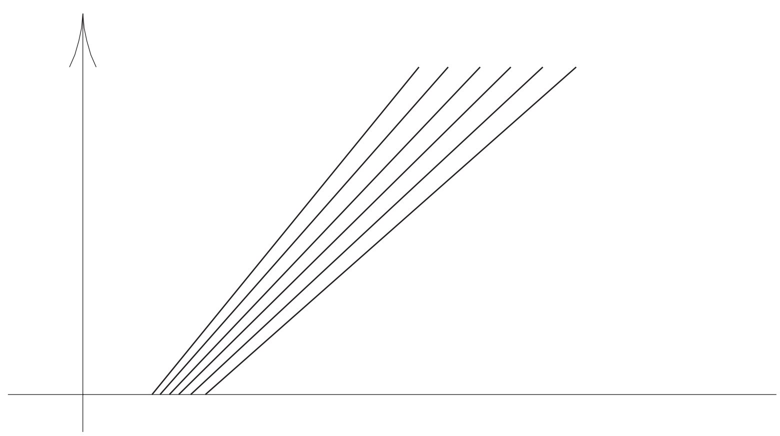


FIGURE 18: *A spatially localized packet of particles is launched with a narrow assortment of initial velocities. As time passes the packed becomes progressively more disperse.*

clear is the mechanical situation shown in the figure, and it is with such an image in mind that one might be tempted to understand the dispersion of wave packets. It would, however, be overly simplistic to attribute the effect in question to the fact that a wave packet has been assembled by superposition of monochromatic waves that move with a narrow variety of phase velocities—this for the subtle reason that I now discuss.

Suppose it to be the case that $e^{i[\omega t - kx]}$ satisfies the linear wave equation of momentary interest to us—now *not* assumed to be the K-G equation—if and only if ω and k satisfy the **dispersion relation**

$$\omega = \omega(k) \quad \text{or inversely} \quad k = k(\omega)$$

To construct a right-running wave packet we write

$$\varphi(x,t) = \int \varPhi(\omega)e^{i[\omega t - k(\omega)x]}d\omega$$

where $\varPhi(\omega)$ vanishes except on a narrow frequency band centered about the carrier frequency ω_0. Developing $\omega t - k(\omega)x$ in powers of $(\omega-\omega_0)$, we find

$$\begin{aligned}\omega t - k(\omega)x = [\omega_0 t - k(\omega_0)x] + [t - xk'(\omega_0)](\omega-\omega_0) &- \tfrac{1}{2}xk''(\omega_0)(\omega-\omega_0)^2\\ &- \tfrac{1}{6}xk'''(\omega_0)(\omega-\omega_0)^3\\ &\;\vdots\end{aligned}$$

so if the red terms (terms of orders $n \geqslant 2$) could be abandoned we would have

$$\varphi(x,t) \approx \underbrace{\int \varPhi(\omega)e^{i[(\omega-\omega_0)(t - x/v_g)]}d\omega}_{A(t-x/v_g)} \cdot e^{i[\omega_0(t-x/v_p)]}$$

The second factor describes the carrier wave: it advances rigidly with speed

$$v_{\text{phase}} = \frac{\omega_0}{k(\omega_0)} \tag{57.1}$$

The first factor modulates the carrier (gives shape to the wave packet, comprises the "signal"): it also advances rigidly, but with speed

$$v_{\text{group}} = \frac{1}{k'(\omega_0)} = \frac{d\omega(k)}{dk}\Big|_{k_0} \tag{57.2}$$

The waves within the contributory frequency band, whether they move with the same or different phase velocities, conspire to make the signal move, but not to make it disperse. It is the population of red terms—and most importantly the term

$$-\tfrac{1}{2}xk''(\omega_0)(\omega-\omega_0)^2$$

—that is responsible for dispersion. A quick calculation supplies

$$-k''(\omega_0) = \Big[\frac{1}{v_g(\omega)}\Big]^2 \frac{dv_g(\omega)}{d\omega}\Big|_{\omega=\omega_0} \tag{58}$$

Evidently it is not variation (within the frequency band) of the phase velocity but *variation of the group velocity* that accounts for dispersion. In those subject areas where wave physics and particle physics have reason to converse (quantum mechanics provides an example) it is typically the case that

$$\text{group velocity} \;\longleftrightarrow\; \text{particle velocity}$$

If read in that light, FIGURE 18 *does* serve to describe the essence of the dispersion phenomenon.

Basic to the phenomenological description of the passage of monochromatic light waves through transparent media is the frequency-dependent[20] **index of refraction**

$$n(\nu) \equiv \frac{c}{v_{\text{phase}}(\nu)} \quad : \quad c = \text{velocity of light in vacuum}$$

From (wavelength)(frequency) $= v_{\text{phase}}$ it follows that

$$\lambda_{\text{medium}} = \tfrac{1}{n}\lambda_{\text{vacuum}}$$

and this, by $k = 2\pi/\lambda$ and $\omega = 2\pi\nu$, entails

$$\begin{aligned} k_{\text{medium}} &= n{\cdot}k_{\text{vacuum}} \\ &= \tfrac{1}{c}\,\omega\, n(\omega) \end{aligned}$$

For most commonly-encountered transparent materials (air, glass) the index of refraction is (at optical frequencies) a *rising* function of frequency. This circumstance accounts, by Snell's law (see the figure), for the fact that blue

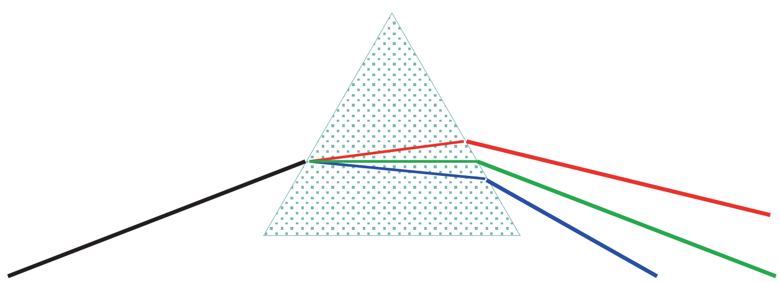


FIGURE 19: *"Dispersion" in the sense used to describe a property of prisms. It is because $n(\nu)$ is a rising function of frequency that $\theta_{\text{blue}} > \theta_{\text{red}}$.*

light is more strongly bent than red when passed through a prism. "Dispersion" in the prismatic sense can be traced to a property of $k'(\omega)$, and is said to be "normal" or "anomalous" according as $k'(\omega)$ is positive or negative. On the

[20] Here I will, in deference to well-estaablished tradition, write $\nu = \frac{1}{2\pi}\omega$ to denote literal frequency.

other hand, "dispersion" in the sense discussed previously has been traced to a property of $k''(\omega)$. In common materials $k''(\omega) < 0$, so people sometimes (confusingly!) say of dispersion in that sense that it is "normal" or "anomalous" according as $k''(\omega)$ is negative or positive.[21]

EXAMPLE: Consider the dispersion relation

$$\omega^2 = u^2(k^2 + \alpha k^4)$$

that is implicit in the 4^{th}-order linear field equation

$$\tfrac{1}{u^2}\varphi_{tt} = \varphi_{xx} - \alpha\varphi_{xxxx}$$

When k is very large (λ very small) we have $\omega \sim \sqrt{\alpha}\,uk^2$: the string has become "stiff" (becomes agitated at short wavelengths). But if k is small (in the sense $\alpha k^2 \ll 1$) then

$$\omega = uk\sqrt{1+\alpha k^2} = uk(1 + \tfrac{1}{2}\alpha k^2 + \cdots)$$

Calculus supplies

$$k'(\omega) = \Big(\frac{d\omega}{dk}\Big)^{-1} \quad\text{and}\quad k''(\omega) = -\Big(\frac{d\omega}{dk}\Big)^{-3}\frac{d^2\omega}{dk^2}$$

on which basis we compute

$$k''(\omega) = -\frac{24\alpha k}{u^2(2+3\alpha k^2)^3}$$

We conclude that the dispersive properties of such a system ("dispersion" understood here to refer to wave packet deformation) are "normal" or "anomalous" according as α is positive or negative. Notice also that we have

$$\begin{aligned} v_{\text{phase}} &= \frac{\omega}{k} = u(1 + \tfrac{1}{2}\alpha k^2) \\ v_{\text{gauge}} &= \frac{d\omega}{dk} = u(1 + \tfrac{3}{2}\alpha k^2) \\ &= v_{\text{phase}} + u\alpha k^2 \end{aligned}$$

so when $\alpha > 0$ we encounter the curious situation

$$v_{\text{group}} > v_{\text{phase}} > u$$

[21] For an instance of such usage, see Iain Main, *Vibrations and Waves in Physics* (3^{rd} edition 1993), page 219. Derivation of $n(\omega)$ from physical first principles is a highly non-trivial matter. For a good introduction to the subject, and to related topics,see F. S. Crawford, Jr., *Waves*: *Berkeley Physics Course, Volume 3* (1968), §4.3, pages 176–191.

6. The telegrapher's equation: The Klein-Gordon equation (10) can be written

$$\big\{\partial_t^2 - u^2\partial_x^2 + (u\varkappa)^2\big\}\varphi(x,t) = 0$$

Bearing in mind the adjustment $\ddot{x} + \omega^2 x = 0 \longmapsto \ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$ that in the theory of simple oscillators serves to model dissipation,[22] we look now to a variant of the K-G equation

$$\big\{\partial_t^2 + 2u\gamma\,\partial_t - u^2\partial_x^2 + (u\varkappa)^2\big\}\varphi(x,t) = 0 \tag{59}$$

that has, as I have already indicated (see again page 8), an ancient and honorable history: (58) is the "telegrapher's equation," first studied by William Thomson (Lord Kelvin) in 1855.[23]

Looking for solutions of the form $e^{i(\omega t - kx)}$ we obtain

$$-\omega^2 + i\,2u\gamma\omega + u^2(k^2 + \varkappa^2) = 0$$

giving

$$\omega = iu\gamma \pm \underbrace{u\sqrt{k^2 + \varkappa^2 - \gamma^2}}_{\omega(k)} \tag{60}$$

Thus are we led to write

$$\varphi = e^{-u\gamma\,t}\cdot e^{i[\omega(k)\,t \pm k\,t]}$$

and to distinguish several cases:

$\boxed{\text{CASE } \varkappa_0^2 \equiv \varkappa^2 - \gamma^2 > 0}$ All that we have learned from study of the K-G equation can be retained virtually without change, the only differences being that (i) the role formerly played by $\varkappa^2$ is now assigned to $\varkappa_0^2$, and (ii) all K-G functions have now acquired $e^{-u\gamma t}$-factors which are—and this is the remarkable point—k-independent: all φ fields die, with characteristic times given universally (*i.e.*, at all frequencies) by $\tau = 1/u\gamma$. Wave packets go about their familiar dispersive business, but die before they get very far.

$\boxed{\text{CASE } \varkappa_0^2 \equiv \varkappa^2 - \gamma^2 = 0}$ The dispersion relation (60) now assumes the form

$$\omega(k) = uk$$

characteristic of a simple string: **all dispersive effects are extinguished**, and the general solution of (59) becomes

$$\varphi(x,t) = e^{-u\gamma\,t}\cdot\big\{f(x - ut) + g(x + ut)\big\}$$

---

[22] See again equations (2) and (14) in Chapter 3.

[23] For a detailed account of the history of the telegrapher's equation see http://www.du.edu/~jcalvert/tech/cable.htm#Intr. For an account of the basic technical details see http://en.wikipedia.org/wiki/Telegrapher's_equations.

REMARK: In notation appropriate to the theory of transmission lines

$$4u^2(\gamma^2 - \varkappa^2) = \Big(\frac{RC + GL}{LC}\Big)^2 - 4\frac{RG}{LC} = \Big(\frac{RC - GL}{LC}\Big)^2 = 0$$

so the condition $\varkappa_0^2 = 0$ amounts simply to the requirement that

$$RC = GL$$

On lines that conform to this condition **signals propagate without distortion**, though they do attenuate.

$\boxed{\text{CASE } \varkappa^2 - \gamma^2 \equiv -k_0^2 < 0}$ In this highly damped circumstance

$$\omega(k) = u\sqrt{k^2 - k_0^2} \text{ is real or imaginary according as } k^2 \gtrless k_0^2$$

The implication is that Fourier components with $k^2 < k_0^2$ are *superattenuated* (*i.e.*, die faster than normal, at k-dependent rates). Such strings act as **high-pass filters** (but not very good ones, for even at pass frequencies they are dispersive and lossy).

7. Symptoms of underlying discreteness: It was by "refinement of the lattice" (FIGURE 1) that we were led from the Newtonian dynamics of vibrating "crystals" to an account of the vibrations supported by idealized "strings." Systems of the former type have finitely many degrees of freedom, are described by functions $\{\varphi_1(t), \varphi_2(t), \ldots \varphi_N(t)\}$ that satisfy coupled ordinary differential equations of the form (compare (1))

$$\ddot{\boldsymbol{\varphi}} + \mathbb{\Omega}^2\boldsymbol{\varphi} = \mathbf{0} \tag{61.1}$$

$$\mathbb{\Omega}^2 \equiv \omega_0^2 \begin{pmatrix} 2 & -1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & \ldots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & -1 & 2 \end{pmatrix} \tag{61.2}$$

with $\omega_0^2 = k/m$. Systems of the latter type are, on the other hand, described by a single field function $\varphi(x, t)$ that satisfies a solitary partial differential equation

$$\big\{\partial_t^2 - u^2\partial_x^2\big\}\varphi(x, t) = 0 \tag{62}$$

Several considerations recommend (62)—and equations like it—to our attention:

- The computational problems latent in (62) are—for large N—much more tractable than those latent in (61).

- If a string is the physical object of interest it would be retrograde to discard the theory of idealized strings in favor of a theory of many (equally idealized) particles.
- From higher-dimensional generalizations of (62) it becomes possible to gain direct insight into such phenomena as polarization, interference, diffraction —insight which it would be virtually impossible to extract from many-particle physics, even if it be granted that wave-conducting media are ultimately granular/particulate. Waves are *cooperative phenomena*: wave physics permits/invites one to forget the dancers, concentrate on the dance.

However. . . the wave equation (62) speaks of idealized systems that are able to support waves of arbitrarily high frequency, arbitrarily short wavelength. This the electromagnetic field (for example) *is* able to do (though at very high frequencies we have to abandon the classical conception of electromagnetism). But it would clearly be absurd to contemplate soundwaves on a crystal with

$$\text{wavelength} \ll \text{mean interatomic separation}$$

and unreasonable to suppose that (62) speaks meaningfully about waves on a guitar string with

$$\text{wavelength} \ll \text{string diameter}$$

Clearly, there are contexts in which the high-frequency implications of wave theory can lead us astray, contexts in which "refinement of the lattice" leads us away from the physical facts into a kind of fantasy land. I explore this issue by study of a concrete example.

Equations (61) and (62) are both *linear*. In both instances, therefore, it becomes natural to look for modes in which all parts of the system vibrate in synchrony, and from those to assemble more general solutions by superposition. To implement the idea as it relates to (61), we understand $\boldsymbol{\varphi}$ to be the real part of $\boldsymbol{\psi}$, assume $\boldsymbol{\psi}(t) = \psi \cdot e^{i\omega t} \cdot \boldsymbol{\Phi}$ (with ψ an arbitrary complex number) and are led to the eigenvalue problem

$$(\mathbb{\Omega}^2 - \omega^2\,\mathbb{I})\,\boldsymbol{\Phi} = \mathbf{0}$$

Looking specifically to the case $N = 25$, we use `Eigenvalues[]` and `ListPlot[]` to obtain the data displayed on the following page. The commands

$$\texttt{NullSpace[}\,\mathbb{\Omega}^2 - \omega_n^2\,\mathbb{I}\,\texttt{]} \quad : \quad n = 1, 2, \ldots, 25$$

produce the normalized eigenvectors $\boldsymbol{\Phi}_n$—the first few of which are displayed graphically on page 43.

To implement the same idea as it relates to (62), we assume $\varphi(x,t)$ to have the separated form

$$\varphi(x,t) = \psi \cdot e^{i\omega t} \cdot \Phi(x)$$

and require of $\Phi(x)$ that $\Phi(0) = \Phi(\ell) = 0$, where (if a is the lattice constant—the inter-particle separation at rest) $\ell = (N+1)a$ is the overall length of the

n	ω_n/ω_0	n	ω_n/ω_0
1	0.12075	13	1.41421
2	0.24107	14	1.49702
3	0.36051	15	1.57437
4	0.47863	16	1.64597
5	0.59500	17	1.71156
6	0.70921	18	1.77091
7	0.82082	19	1.82380
8	0.92944	20	1.87003
9	1.03468	21	1.90944
10	1.13613	22	1.95188
11	1.23344	23	1.96724
12	1.32625	24	1.98542
		25	1.99635

TABLE 1: *List of the numbers ω_n obtained by taking the square roots of the eigenvalues of Ω^2 in the case $n = 25$.*

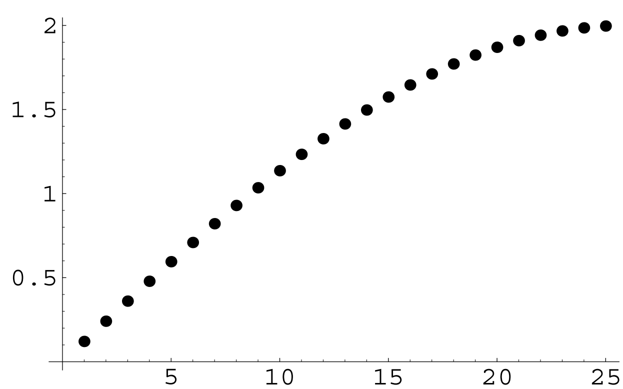


FIGURE 20: `ListPlot` *of the data tabulated above, showing the modal frequencies $\{\omega_1, \omega_2, \ldots, \omega_{25}\}$ of our 25-atom crystal.*

crystal. From the wave equation (62) we then obtain

$$\varPhi''(x) = -(\omega/u)^2\,\varPhi(x) \tag{63}$$

and to achieve conformity with the boundary conditions are forced to set

$$\varPhi_n(x) = A_n \cdot \sin[n\pi x/\ell] \quad : \quad n = 1, 2, \ldots \tag{64}$$

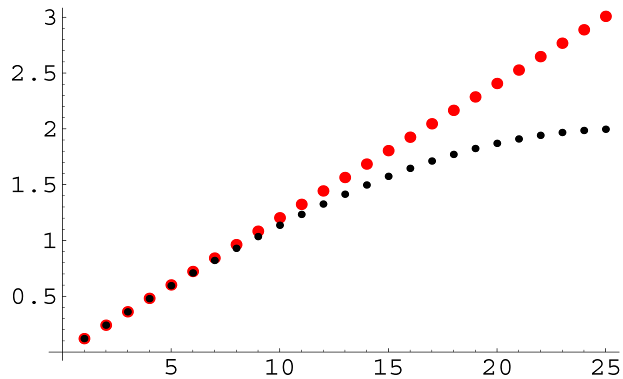


FIGURE 21: *Shown here in red are the modal frequencies* (65) *obtained in the wave-theoretic approximation to the exact dynamics of our 25-atom crystal. The exact modal frequencies, displayed already in* FIGURE 20, *are for purposes of comparison repeated here in black. Wave theory systematically over-estimates the higher frequencies, and supplies an infinitude of numbers* ω_n $(n > 25)$ *that are entirely spurious* (*possess no counterparts in physical reality*).

Normalization—in the sense $\int_0^\ell \varPhi_n^2(x)\,dx = 1$—requires that we set $A_n = \sqrt{2/\ell}$ (all n). It now follows that $\omega_n/u = n\pi/\ell$; *i.e.*, that

$$\omega_n = n\omega_0 \quad \text{with} \quad \omega_0 = \pi u/\ell \tag{65}$$

In the case at hand $\omega_0 = \pi u/26a = \omega_{n+1} - \omega_n$ and it is to achieve agreement (at the lowest frequencies) with the data presented in TABLE 1 that we set $u/a = 0.995745$. The exact modal frequencies and their wave-theoretic approximants are compared in FIGURE 21: it is evident that the agreement is good at low frequencies, but at high frequencies wave theory gives results that are consistently too high.

Looking to what the discrete and continuous theories have to say about *mode shape*, we find (FIGURE 22) that at low frequencies the agreement is excellent. But at high frequencies the two theories are not at all in agreement, which is to say: wave theory becomes profoundly misleading. Looking to the fastest mode, we find (FIGURE 23) that

$$\varPhi_{25}(x) = \sqrt{\tfrac{2}{26}}\sin\left[25\tfrac{\pi}{26}x\right]$$

captures the sign alternations that are a conspicuous feature of $\boldsymbol{\varPhi}_{25}$ but misses altogether the overall modulation, which is also a conspicuous feature. Much better agreement is achieved by

$$\tilde{\varPhi}_{25}(x) = \sin\left[1\tfrac{\pi}{26}x\right]\cdot\sqrt{\tfrac{2}{26}}\sin\left[25\tfrac{\pi}{26}x\right]$$

but $\tilde{\varPhi}_{25}(x)$ does not satisfy (63), and the *ad hoc* adjustment $\varPhi_{25}(x) \mapsto \tilde{\varPhi}_{25}(x)$ lacks wave-theoretic justification.

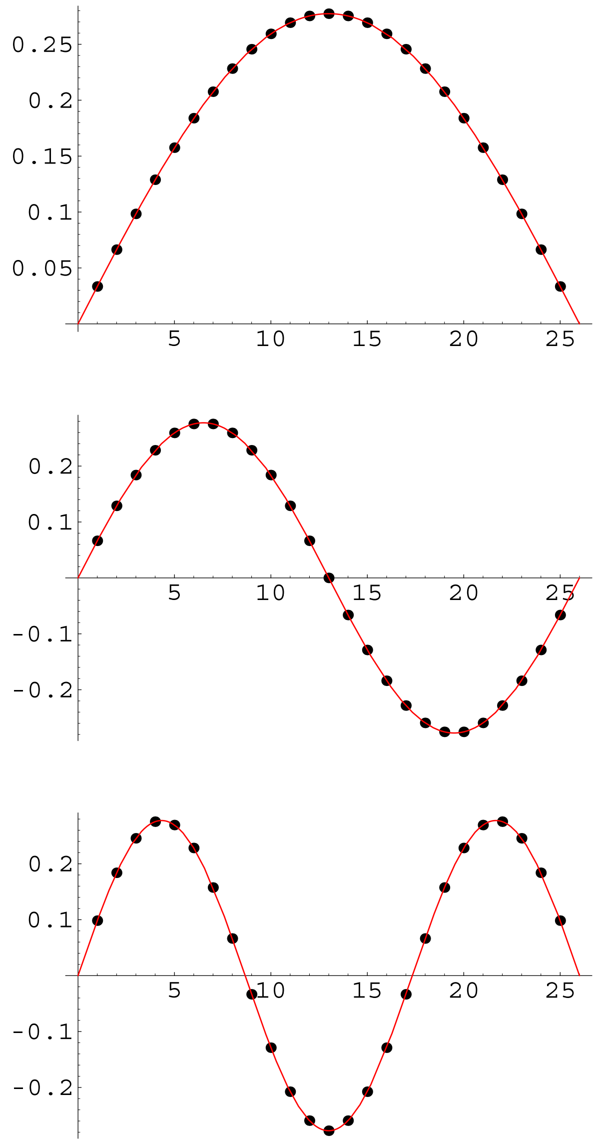


FIGURE 22: **Mode shape at low frequencies**. *Black dots describe the elements of—reading from top to bottom—the eigenvectors* $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$, *obtained by* `NullSpace` *commands. The red curves are graphs of the eigenfuctions*

$$\Phi_n(x) = \sqrt{\tfrac{2}{26}}\sin\left[n\tfrac{\pi}{26}x\right] \quad : \quad n = 1, 2, 3$$

The agreement is seen to be excellent.

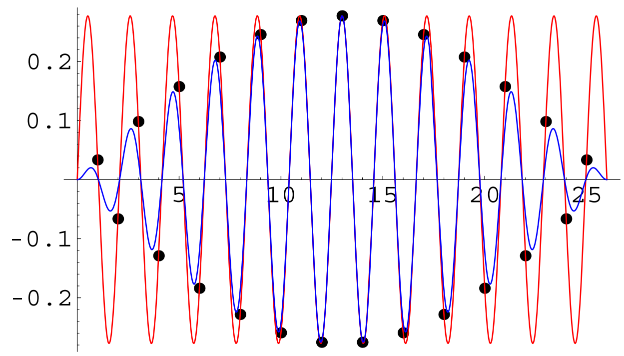


FIGURE 23: *Black dots describe the elements of* $\boldsymbol{\Phi}_{25}$*: note the strict sign alternation. The eigenfunction*

$$\Phi_{25}(x) = \sqrt{\tfrac{2}{26}}\sin\left[25\tfrac{\pi}{26}x\right]$$

captures the sign alternation, but misses the conspicuous modulation. The latter is captured pretty well by a function

$$\tilde{\Phi}_{25}(x) = \sin\left[1\tfrac{\pi}{26}x\right]\cdot\sqrt{\tfrac{2}{26}}\sin\left[25\tfrac{\pi}{26}x\right]$$

that, however, lacks theoretical foundation/justification.

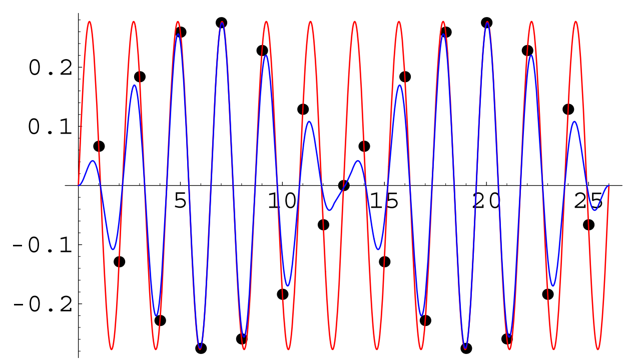


FIGURE 24: *Black dots describe the elements of* $\boldsymbol{\Phi}_{24}$*. The function*

$$\Phi_{24}(x) = \sqrt{\tfrac{2}{26}}\sin\left[24\tfrac{\pi}{26}x\right]$$

does not fit the data at all well. The ad hoc *function*

$$\tilde{\Phi}_{24}(x) = \texttt{Abs[}\sin\left[2\tfrac{\pi}{26}x\right]\texttt{]}\cdot\sqrt{\tfrac{2}{26}}\sin\left[24\tfrac{\pi}{26}x\right]$$

does a good deal better.

The cautionary moral is very simple: wave theory, when used as a computational scheme to approximate the vibrational physics of discrete systems, must—like all approximations—be used with care, for it supplies information that becomes increasingly misleading as decreasing wavelength becomes comparable to the characteristic "grain size" of the discrete system.

8. From "waves without supporting media" to special relativity: That cautionary advice can, however, be disregarded when wave theory is used to describe the vibrations of media—if any there be—that are *really* continuous, that at every degree of resolution fail to reveal any hint of "graininess." Maxwell's discovery[24] that light is an electromagnetic wave led physicists (1870–1900) to inquire into the properties of the medium—called the "æther"—that supported those waves. It was clear from the outset that the æther was made of uncommon stuff: the fact that c is so relatively large suggested that the æther must be very stiff, yet at the same time quite tenuous, for all efforts to detect motion *through* the æther (Fizeau (1853), Michaelson–Morley (1881–1887), Trouton–Noble (1903)) had met with failure. Attempts to account for that failure were resourceful but not very convincing (had the feel of attempts to explain why a ghost was unseen). It was theoretical desperation that led H. A. Lorentz (∼1903) to advance—tentatively—the radical suggestion that electromagnetic waves might get along very will *without* the support of a medium! Like a smile without a face. My objective here will be to describe how it comes about that Lorentz' suggestion leads directly to the invention of special relativity.[25]

To pose the issue, let it be supposed (see FIGURE 25) that we—who call ourselves O to emphasize our inertiality—contemplate a taut string at rest. We remark that *elastic longitudinal vibrations* of the string are described by the wave equation

$$\Big[\Big(\frac{\partial}{\partial x}\Big)^2-\frac{1}{u^2}\Big(\frac{\partial}{\partial t}\Big)^2\Big]\varphi(x,t)=0 \tag{66}$$

and that the general solution of (1) can be represented

$$\begin{aligned}\varphi(x,t)&=f(x-ut)+g(x+ut)\\&=\text{rigidly right-running}+\text{rigidly left-running}\end{aligned} \tag{67}$$

where the wave function $\varphi(x,t)$ refers physically to the *instantaneous local displacement* of the element of string which resides normally at x.

---

[24] *"I have also a paper afloat, containing an electromagnetic theory of light, which, until I am convinced to the contrary, I hold to be great guns."*

—Letter to Charles Cay (a professor of mathematics, and Maxwell's cousin), dated 5 January 1865.

[25] This material has been adapted from material I developed for presentation to my Physics 100 students in March 1979 to commemorate Einstein's 100[th] birthday. For a more fulsome discussion see "How Einstein might have been led to relativity already in 1895" (August 1999), which is available on the Courses Server.

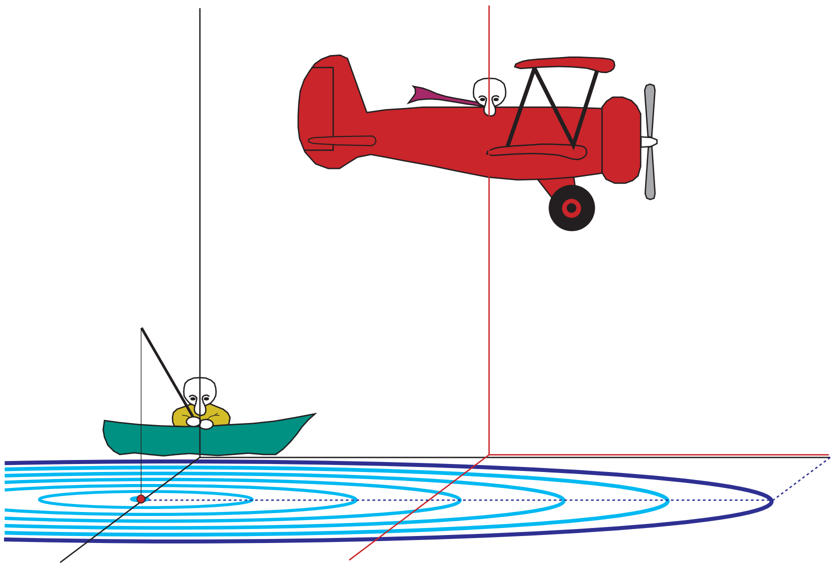


FIGURE 25: *3-dimensional representation of the 1-dimensional issue addressed in the text. We—seated in the boat—are at rest with respect to the water, and see the flying observer to pass by with speed v. In the figure I have attempted to suggest—in anticipation of things to come—that the biplane is very slow: $v < u$, where u denotes the speed (with respect both to us and to the stagnant sea) of wave propagation.*

A second inertial observer O is seen to pass by with speed v. The question arises: How does O render the physics to which we alluded when we wrote (66); *how does the wave equation transform*? Which we take to mean: How does (66) respond to Galilean transformations

$$\left.\begin{aligned} t &= t \\ x &= x + vt \end{aligned}\right\} \tag{68}$$

The answer falls into our lap when we realize that to describe the “displacement field” O writes

$$\begin{aligned} \varphi(x,t) &= f(x + vt - ut) + g(x + vt + ut) \\ &= f(x - [u-v]t) + g(x + [u+v]t) \end{aligned} \tag{69}$$

and that such a function cannot possibly satisfy a wave equation of type (66); evidently (69) is in fact a solution of

$$\Big[\frac{\partial}{\partial x} + \frac{1}{u-v}\frac{\partial}{\partial t}\Big]\Big[\frac{\partial}{\partial x} - \frac{1}{u+v}\frac{\partial}{\partial t}\Big]\varphi(x,t) = 0 \tag{70}$$

which gives back (66) only in the trivial case $v = 0$. Expansion of the differential operator that stands of the left side of (70) gives

$$\underbrace{\Big(\frac{\partial}{\partial x}\Big)^2 + \frac{2v}{u^2 - v^2}\frac{\partial}{\partial x}\frac{\partial}{\partial t} - \frac{1}{u^2 - v^2}\Big(\frac{\partial}{\partial t}\Big)^2}$$

which reduces to

$$= \Big(\frac{\partial}{\partial x}\Big)^2 - \frac{1}{u^2}\Big(\frac{\partial}{\partial t}\Big)^2 \quad \text{at } v = 0$$

and—interestingly—*becomes singular as* $v^2 \uparrow u^2$.

The circumstance that (66) and (70) are—though they refer to the same physics—*structurally distinct* bothers neither O nor O, for both realize that *they stand in asymmetric relationships to the medium*: O is at rest with respect to the string; O in motion with respect to the string.

Enter: Young Einstein I quote now from an English translation of Einstein's "Autobiographical Notes:"[26]

> *"... I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. ... After ten years of reflection such a principle resulted from a paradox upon which I had already hit at the age of 16: if I pursue a beam of light with velocity c... I should observe such a beam as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations. ... It seemed to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer... at rest."*

Suppose we had reason—as Lorentz a few years later did have reason—to posit the existence in Nature of a class of waves which move *without benefit of a "supporting medium."* We lose then all grounds on which to tolerate any asymmetry in the relationship of O to O, at least insofar as concerns the physics of such waves.

If O sees such a wave[27] $\varphi(x,t) = f(x - ct) + g(x + ct)$ to satisfy

$$\Big[\Big(\frac{\partial}{\partial x}\Big)^2 - \frac{1}{c^2}\Big(\frac{\partial}{\partial t}\Big)^2\Big]\varphi(x,t) = 0$$

[26] P. A. Schilpp (editor), *ALBERT EINSTEIN: Philosopher–Scientist* (1951), page 53.

[27] The notational adjustment $u \to c$ is intended to emphasize that we have now in mind waves which are "special" (in the sense just described), though not necessarily electromagnetic; put therefore out of mind the thought that c refers to the "velocity of light."

then O must see that wave to satisfy

$$\Big[\Big(\frac{\partial}{\partial x}\Big)^2 - \frac{1}{c^2}\Big(\frac{\partial}{\partial t}\Big)^2\Big]\varphi(x,t) = 0$$

How can such "form invariance" be achieved?

Lorentz transformations We are forced to the conslusion that to achieve

$$f(x-ct)+g(x+ct) \longrightarrow f(x-ct)+g(x+ct) \quad : \quad \text{all } f(\bullet) \text{ and } g(\bullet)$$

we must presume the relation of O to O to be described *not* by

$$\left.\begin{aligned} t &= t \\ x &= x + vt \end{aligned}\right\} \tag{68}$$

but by some *modified transformation equations*

$$\left.\begin{aligned} t &= \mathfrak{T}(x,t;v) \\ x &= \mathfrak{X}(x,t;v) \end{aligned}\right\} \tag{71}$$

and, moreover, that these must have the property that they entail

$$x \pm ct = W_\pm(x \pm ct) \tag{72}$$

For only then will we have

$$\begin{aligned} f(x-ct)+g(x+ct) &= f(W_-(x-ct)) + g(W_+(x+ct)) \\ &\equiv f(x-ct)+g(x+ct) \quad : \quad \text{all } f(\bullet) \text{ and } g(\bullet) \end{aligned}$$

Let us now *assume* (71) to possess the *linear* structure characteristic of the Galilean transformation (68), writing

$$\left.\begin{aligned} t &= Pt + p\,x \\ x &= q\,t + Qx \end{aligned}\right\} \tag{73}$$

where P, p, Q and q depend (in some presently unknown way) upon the kinematic parameter v. From (73) it follows that

$$x \pm ct = (Q \pm cp)x + (q \pm cP)t \tag{74.1}$$

while to achieve compliance with (72) there must exist multipliers $K_\pm(v)$ such that

$$= K_\pm(v)\cdot(x \pm ct) \tag{74.2}$$

From (74) we are led to a quartet of equations

$$\begin{aligned} Q + cp &= K_+ \\ Q - cp &= K_- \\ q + cP &= +cK_+ \\ q - cP &= -cK_- \end{aligned}$$

from which it follows readily that

$$\left.\begin{aligned} P &= \tfrac{c}{2c}(K_+ + K_-) \\ p &= \tfrac{1}{2c}(K_+ - K_-) \\ Q &= \tfrac{1}{2}\,(K_+ + K_-) \\ q &= \tfrac{c}{2}\,(K_+ - K_-) \end{aligned}\right\} \tag{75}$$

But the functions $K_\pm(v)$ and $c(v)$ remain at present unknown.

To make further progress, let us REQUIRE that the transformation which sends

$$O \xrightarrow[v]{\hspace{5em}} O$$

be *symmetric* in the sense that $v \to -v$ achieves its inversion; then

$$\begin{aligned} x \pm ct = K_\pm(v)\cdot(x \pm ct)& \\ &x \pm ct = K_\pm(-v)\cdot(x \pm ct) \end{aligned}$$

supplies the information that

$$K_+(v)\cdot K_+(-v) = K_-(v)\cdot K_-(-v) = 1 \tag{76.1}$$

Note also that time-reversal sends

$$\begin{aligned} x \pm ct &= K_\pm(+v)\cdot(x \pm ct) \\ &\Big\downarrow \text{time-reversal} \\ x \mp ct &= K_\pm(-v)\cdot(x \mp ct) \end{aligned}$$

so if we REQUIRE *time-reversal invariance* we are led to the condition

$$K_\pm(+v) = K_\mp(-v) \tag{76.2}$$

Equations (76) conjointly entail

$$K(v) \equiv K_+(v) = \frac{1}{K_-(v)} \tag{77}$$

which serves to reduce the number of unknown functions.

Consider finally how O describes O's "clock-at-himself," and *vice versa*. Working from (73) and (75), we find that

$$(t, vt) \longleftarrow (t, \quad 0) \quad \text{entails} \quad q/P = v = c\,\frac{K_+ - K_-}{K_+ + K_-} \tag{78.1}$$

$$(t, \quad 0) \longrightarrow (t, -vt) \quad \text{entails} \quad q/Q = v = c\,\frac{K_+ - K_-}{K_+ + K_-} \tag{78.2}$$

which upon comparison give

$$c(v) = c \quad : \quad \text{all } v \tag{79}$$

The striking implication is that

> *If O and O "share the wave equation" then—granted certain natural assumptions—they must necessarily "share the value of c."*

It is on this basis that henceforth we abandon the red c black c distinction. Introducing the dimensionless relative velocity parameter

$$\beta \equiv v/c$$

and returning with (77) to (78), we obtain

$$\beta = \frac{K^2 - 1}{K^2 + 1} \quad \text{whence} \quad K(v) = \sqrt{\frac{1+\beta}{1-\beta}} \tag{80}$$

All the formerly "unknown functions" have now been determined.

Bringing this information to (75), and returning with the results to (73), we obtain

$$\left.\begin{aligned} t &= \tfrac{1}{2}\Big[\sqrt{\tfrac{1+\beta}{1-\beta}} + \sqrt{\tfrac{1-\beta}{1+\beta}}\Big]t + \tfrac{1}{2c}\Big[\sqrt{\tfrac{1+\beta}{1-\beta}} - \sqrt{\tfrac{1-\beta}{1+\beta}}\Big]x = \gamma(\;t + vx/c^2) \\ x &= \tfrac{c}{2}\Big[\sqrt{\tfrac{1+\beta}{1-\beta}} - \sqrt{\tfrac{1-\beta}{1+\beta}}\Big]t + \tfrac{1}{2}\Big[\sqrt{\tfrac{1+\beta}{1-\beta}} + \sqrt{\tfrac{1-\beta}{1+\beta}}\Big]x = \gamma(vt + \;x) \end{aligned}\right\} \tag{81}$$

where

$$\gamma \equiv \tfrac{1}{\sqrt{1-\beta^2}} = 1 + \tfrac{1}{2}\beta^2 + \tfrac{3}{8}\beta^4 + \cdots$$

Transformations $(t, x) \longrightarrow (t, x)$ of the design (81) first presented themselves when Lorentz (∼1904) looked to the transformation-theoretic properties (not of the wave equation but) of Maxwell's equations, and are called **Lorentz transformations**. For $v \ll c$ (equivalently: in the formal limit $c \uparrow \infty$) (81) gives

$$\begin{aligned} t &= \qquad t + \cdots \\ x &= x + vt + \cdots \end{aligned}$$

which exposes the sense in which the Lorentz transformations

- enlarge upon
- reduce to
- contain as approximations

the Galilean transformations (68).

We have achieved (81) by an argument which involves little more than high school algebra—an argument which approaches the masterful simplicity of Einstein's own line of argument (1905). But Einstein's argument—which with its population of idealized trains, lanterns and meter sticks seems to me to read more like "mathematical epistemology" than physics—remains unique in the field; I do not claim to understand it well enough to be able to reproduce it in the classroom, do not find it particularly compelling, and do not know how seriously it today is to be taken (since relativity has been found to pertain at a scale so microscopically fine as to render Einstein's "thought experiments" meaningless). My own argument, on the other hand, springs from a question—*How does this object of interest transform, and what transformations preserve its form?*—which in its innumerable variants has been fruitfully central to mathematical physics for at least 250 years.

Principle of relativity A world in which inertial observers O and O use

$$\left.\begin{aligned} t &= t \\ x &= x + vt \end{aligned}\right\} \tag{68}$$

when comparing mechanical observations, but must use

$$\left.\begin{aligned} t &= \gamma(\ t + vx/c^2) \\ x &= \gamma(vt + \ x) \end{aligned}\right\} \tag{81}$$

when comparing observations pertaining to "mediumless waves" (and must, moreover, be prepared to assign distinct values $c', c'', \ldots$ to c when confronted with distinct systems of such waves) is a world which is *theoretically unhinged* —a world in which the physics books can, in their totality, pertain to the experience of (at most) a single observer. Such a state of affairs would be inconsistent with the spirit of the Copernican revolution.

It follows—on grounds which, whether formal/philosophical/æsthetic, are clearly fundamental—that not more than one of the options spelled out above can figure in a comprehensive physics. How to proceed?

- We might try to stick with (68); we then retain Newtonian dynamics intact, but must give up the notion of a "mediumless wave." This may seem acceptable on its face, but entails that we also abandon Maxwellean electrodynamics and the associated electromagnetic theory of light—theories which conform very well to observation.

- We might, alternatively, adopt some instance of (81)—namely, the instance which results from promoting some c-value to the status of a *universal constant of Nature*. If, in particular, we set

$$c = \text{velocity of light}$$

then we retain Maxwellean electrodynamics intact, but must abandon Newtonian dynamics, which becomes merely the leading low-velocity *approximation* to a "relativistic dynamics." This, clearly, is the more "interesting" way to go (and in Einstein's view—by force of his epistemological argument—the *only* way to go). We are led thus —with Einstein (1905)—to postulate the

> **PRINCIPLE OF RELATIVITY**: Physical formulæ and concepts shall be "admissible" if and only if they are form-invariant with respect to the Lorentz transformations (81).

Note that the Principle of Relativity refers to no specific physical phenomenon; it refers, instead, to the necessary *structure of physical theories in general*. It stands to physics in much the same relationship that the Rules of Syntax (which refer to no specific utterance, but to the design of "well-formed sentences") stand to language, and has much in common with the Principle of Dimensional Homogeneity (see again Chapter 2, page 9).

Invention of "spacetime" Diagrams such as that presented in FIGURE 26 are commonly encountered already in pre-relativistic physics, where they are used to represent kinematic/dynamic events. The point to which I would draw

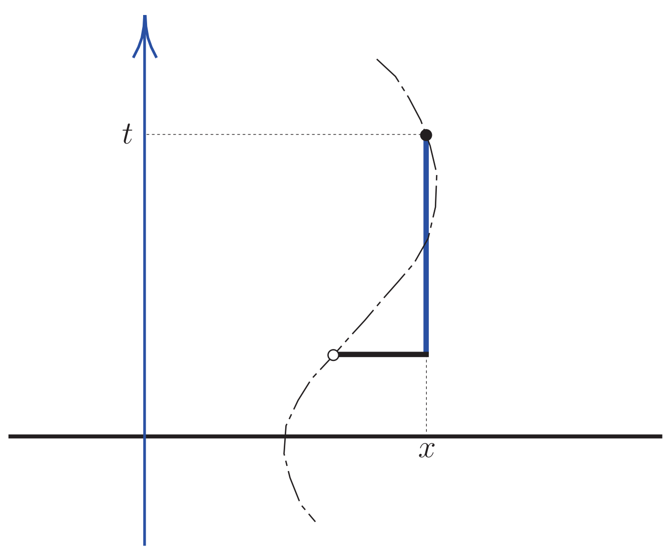


FIGURE 26: *Space-time representation of the motion $x(t)$ of a point (kinematics) or particle (dynamics). The diagram is in effect a movie—a t-parameterized stack of "time-slices." In pre-relativistic physics it makes one kind of good sense to speak of the Euclidean distance between two space points (as measured by a meter stick), and quite a different kind of good sense to speak of the temporal interval separating two time-slices (as measured by a clock), but no sense to speak of "the distance between two events," ∘ and •.*

attention is that such "space-time diagrams" arise by formal fusion of two distinct notions:

- a 1-dimensional "time axis," the points of which are well-ordered, and on which intervals have the physical dimension of TIME;
- an N-dimensional "space," the points of which are (except in the case $\mathrm{N}=1$) *not* well-ordered, and on which intervals have the physical dimension of LENGTH.

Notice now that, upon returning with (77) to (74.2), we have

$$(x \pm ct) = K^{\pm 1}(x \pm ct) \tag{82}$$

from which it follows that

$$(ct)^2 - x^2 = (ct)^2 - x^2 \tag{83}$$

This last equation establishes the sense in which

the expression $(ct)^2 - x^2$ is Lorentz-invariant

just as, on the Euclidean plane,

the expression $x^2 + y^2$ is rotationally invariant

In the latter context, $r^2 \equiv x^2 + y^2$ defines the (squared) "Pythagorean length" of the spatial interval separating the point (x, y) from the origin $(0, 0)$; it is from that definition that the Euclidean plane—now a "metric space"—acquires its distinctive metric properties.

Which brings us to the work of Hermann Minkowski, who had been one of Einstein's teachers (of mathematics) at the ETH in Zurich, and who in 1908 wrote

> *"The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth, space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."*

Minkowski's idea was to let $s^2 \equiv (ct)^2 - x^2$ serve to *define* the (squared) "Minkowskian length" of the spacetime interval separating the "event" (ct, x) from the origin $(0, 0)$. By this simple means (see Figure 27) he achieved

$$\text{space} \otimes \text{time} \longrightarrow \text{spacetime}$$

and managed thus expose the deeper significance of Einstein's accomplishment (which Einstein himself had somehow failed to notice).

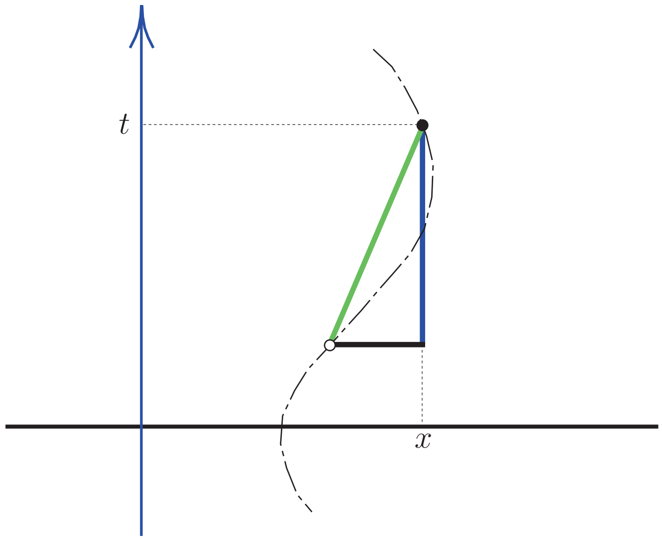


FIGURE 27: *By assigning meaning to the (squared) length of the green interval*

$$s^2 \equiv c^2(t - t_0)^2 - (x - x_0)^2$$

Minkowski invented "spacetime"—the $(3+1)$*-dimensional metric continuum which serves as the blackboard upon which all physics is written, the arena of physical experience. The dashed curve is, in relativistic parlance, called the "worldline" of the point/particle.*

It is to clarify the simple essence of Minkowski's idea that we make at this point a notational adjustment, writing

$$\left.\begin{aligned} &x^0 \text{ in place of } ct\\ &x^1 \text{ in place of } x\end{aligned}\right\}\quad \text{both have the dimensionality of LENGTH}$$

and

$$\begin{aligned} s^2 &= (x^0)^2 - (x^1)^2 \\ &\downarrow \\ &= (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad \text{in 4-dimensional spacetime} \end{aligned} \tag{84}$$

It is, of course, the availability of the dimensioned constant c which makes such an adjustment possible; i.e., which makes it possible to "measure time in centimeters." The right side of (84) is precisely Pythagorean *except for the funny sign*, which makes a world of difference. In this notation (81) becomes

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \mathbb{L}\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \quad \text{with} \quad \mathbb{L} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix}$$

(84) can be written

$$s^2 = \begin{pmatrix} x^0 \\ x^1 \end{pmatrix}^{\mathsf{T}} g \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \quad \text{with} \quad g = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

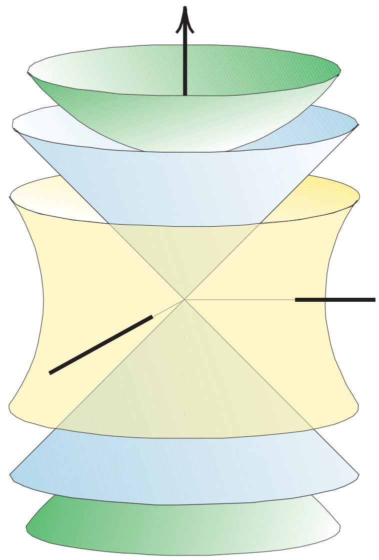


FIGURE 28: *Isometric surfaces in spacetime. An arrow marks the time axis. All points on the green bowl lie at the same distance $s^2 > 0$ from the origin; all points on the blue cone lie at the same distance $s^2 = 0$ from the origin; all points on the yellow girdle lie at the same distance $s^2 > 0$ from the origin. The figure is shared by all inertial observers.*

and the statement (83)—Lorentz-invariance of s^2—becomes a corollary of the fact that Lorentz matrices $\mathbb{L}(\beta)$ satisfy

$$\mathbb{L}^{\mathsf{T}}\boldsymbol{g}\,\mathbb{L} = \boldsymbol{g} \quad : \quad \text{all } \beta$$

This is directly analogous to the statement that $r^2 = \boldsymbol{x}\cdot\boldsymbol{x}$ will be invariant under $\boldsymbol{x} \mapsto \mathbb{R}\boldsymbol{x}$ if and only if $\mathbb{R}^{\mathsf{T}}\mathbb{I}\,\mathbb{R} = \mathbb{I}$.[28]

Since the work of Einstein/Minkowski (1907) it has been recognized that (*i*) "spacetime" is a metric space; (*ii*) its geometry is hyperbolic (see the preceding figure) and (*iii*) appears the same to all inertial observers; (*iv*) the physics that we inscribe on spacetime must necessarily possess the same symmetry-structure as spacetime itself. And it was by enlargement upon this insight that Einstein was led to the invention of general relativity (1916). Quite a lot of physics to extract from the theory of "waves on a string when there is no string"!

[28] All that we commonly consider to be most characteristic of special relativity —breakdown of the concept of non-local simultaneity, the Lorentz contraction, time dilation, even $E = mc^2$—follows by direct implication from the material just sketched. For details see (for example) the material cited in Note 25.