3

Aspects of the theory of

OSCILLATORY SYSTEMS

Introduction. When mechanical systems oscillate it is, generally speaking, because they are trapped near a local (if not global) point of minimal potential energy. In one dimension, Taylor expansion of a potential $U(y)$ about a point a gives

$$U(y) = U(a) + U'(a)(y-a) + \tfrac{1}{2}U''(a)(y-a)^2 + \cdots$$

which can be written

$$U(a+x) = e^{x\frac{\partial}{\partial a}}U(a)$$

where the "excursion variable" $x \equiv y - a$. If a is an extremal point then $U'(a) = 0$, and since we can, without physical consequence, assume $U(a) = 0$, we have

$$U(a+x) = \tfrac{1}{2}U''(a)x^2 + \tfrac{1}{3!}U'''(a)x^3 \cdots$$

which is minimal or maximal at a according as $U''(a) \gtrless 0$, and which

$$\approx \tfrac{1}{2}U''(a)x^2 \quad \text{for sufficiently small excursions}$$

We will be interested mainly in multivariable analogs of familiar situation just described. To obtain the Taylor expansion of (say) $U(y^1, y^2)$ about the point $\{a^1, a^2\}$ we write

$$\begin{aligned} U(a^1+x^1, a^2+x^2) &= e^{x^1\partial_1 + x^2\partial_2}\,U(a^1,a^2) \\ &= U(a) + \{x^1U_1(a) + x^2U_2(a)\} \\ &\quad + \tfrac{1}{2}\{x^1x^1U_{11}(a) + x^1x^2U_{12}(a) + x^2x^1U_{21}(a) + x^2x^2U_{22}(a)\} \\ &\quad + \tfrac{1}{3!}\{x^1x^1x^1U_{111}(a) + 3x^1x^1x^2U_{112}(a) + 3x^1x^2x^2U_{122}(a) + x^2x^2x^2U_{222}(a)\} + \cdots \end{aligned}$$

which in the near neighborhood of an extremal point becomes

$$U(a+x) \approx \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} U_{11}(a) & U_{12}(a) \\ U_{21}(a) & U_{22}(a) \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} + \cdots$$

$$\equiv \boldsymbol{x}^{\mathsf{T}} \mathbb{U} \boldsymbol{x} + \cdots \tag{1.1}$$

where
$$\mathbb{U} \equiv \| \partial^2 U(a) / \partial a^i \partial a^j \| \tag{1.2}$$

is obviously symmetric.

Equations (1) clearly retain their meaning and validity in the n-dimensional case, and I proceed in the assumption that our mechanical system *does* have n degrees of freedom. Write $\boldsymbol{x}^{\mathsf{T}} \mathbb{U} \boldsymbol{x} = \boldsymbol{x}^{\mathsf{T}} \mathbb{R}^{-1} \mathbb{R} \mathbb{U} \mathbb{R}^{-1} \mathbb{R} \boldsymbol{x}$ and require of $\mathbb{R}$ that it be a rotation matrix ($\mathbb{R}^{-1} = \mathbb{R}^{\mathsf{T}}$), so that we can write

$$\boldsymbol{x}^{\mathsf{T}} \mathbb{U} \boldsymbol{x} = \boldsymbol{X}^{\mathsf{T}} \mathbb{R} \mathbb{U} \mathbb{R}^{-1} \boldsymbol{X} \quad \text{with} \quad \boldsymbol{X} \equiv \mathbb{R} \boldsymbol{x}$$

Require, moreover, that $\mathbb{R}$ be the "diagonalizer" of $\mathbb{U}$:

$$\mathbb{K} \equiv \mathbb{R} \mathbb{U} \mathbb{R}^{-1} = \begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & k_n \end{pmatrix}$$

It is clear that $\mathbb{K}$ and $\mathbb{R} \mathbb{U} \mathbb{R}^{-1}$ have *identical* eigenvalues, and that the eigenvalues of $\mathbb{K}$ are just the numbers $\{k_1, k_2, \dots, k_n\}$. And we know that the eigenvalues of any real symmetric matrix are necessarily real (because the eigenvalues of any Hermitian matrix are, which is easy to prove). So we have

$$\boldsymbol{x}^{\mathsf{T}} \mathbb{U} \boldsymbol{x} = X^1 k_1 X^1 + X^2 k_2 X^2 + \cdots + X^n k_n X^n$$

of which $\boldsymbol{X} = \boldsymbol{0}$ marks the location of

- a local minimum only if all k_i are positive;
- a saddle point if the k_i are of assorted signs;
- a local maximum if all k_i are negative.

The intrusion here of ideas borrowed from linear algebra is no accident, for quadratic potentials give rise to linear equations of motion, and the resulting theory will be dominated by a **principle of superposition**. As is all of quantum mechanics, and most of wave theory (in all of its manifestations), but which is true only of this "low energy corner" of classical mechanics. It is a circumstance that carries with it access to powerful analytical resources that—though irrelevant to the great bulk of the problems presented by classical mechanics (which is, in general, a highly non-linear subject)—are the "name of the game" in (for example) quantum mechanics and electrical/mechanical engineering.

My intent here will be to review some aspects of "linear classical mechanics" that, though for the most part entirely commonplace, tend to be omitted from introductory accounts of the subject.

1. Lagrangian theory of a damped oscillator. Let $L(\dot{x}, x, t)$ be the Lagrangian of a system with one degree of freedom. To describe the motion of such a system we have

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{\partial^2 L}{\partial \dot{x}\partial \dot{x}}\ddot{x} + \frac{\partial^2 L}{\partial \dot{x}\partial x}\dot{x} + \frac{\partial^2 L}{\partial \dot{x}\partial t} - \frac{\partial L}{\partial x} = 0$$

which in the case $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$ becomes the familiar oscillator equation

$$m\ddot{x} + kx = 0$$

To model the motion of a *damped* oscillator one usually writes

$$m\ddot{x} + 2b\dot{x} + kx = 0 \tag{2}$$

We ask: *What modified Lagrangian leads to the preceding equation*? On has the **Helmholtz conditions**,[1] according to which system of n coupled differential equations

$$G_\nu(\ddot{q}, \dot{q}, q, t) = 0 \quad : \quad \nu = 1, 2, \ldots, n$$

can be obtained from a Lagrangian if and only if these $\frac{1}{2}n(3n-1)$ conditions are satisfied:

$$\left.\begin{aligned} \frac{\partial G_\nu}{\partial \ddot{q}_\lambda} - \frac{\partial G_\lambda}{\partial \ddot{q}_\nu} &= 0 \\ \frac{\partial G_\nu}{\partial \dot{q}_\lambda} + \frac{\partial G_\lambda}{\partial \dot{q}_\nu} &= \frac{d}{dt}\Big[\frac{\partial G_\nu}{\partial \ddot{q}_\lambda} + \frac{\partial G_\lambda}{\partial \ddot{q}_\nu}\Big] \\ \frac{\partial G_\nu}{\partial q_\lambda} - \frac{\partial G_\lambda}{\partial q_\nu} &= \tfrac{1}{2}\frac{d}{dt}\Big[\frac{\partial G_\nu}{\partial \dot{q}_\lambda} - \frac{\partial G_\lambda}{\partial \dot{q}_\nu}\Big] \end{aligned}\right\} \tag{3.1}$$

In the case $n = 1$ only one of these conditions is lively, and it reads

$$\frac{\partial G}{\partial \dot{q}} = \frac{d}{dt}\frac{\partial G}{\partial \ddot{q}} \tag{3.2}$$

which in the case of interest becomes $2b = \frac{d}{dt}m = 0$. It would appear that damping cannot be brought within the compass of Lagrangian mechanics.

However... suppose we were, in place of (2), to write the clearly equivalent equation

$$f(t)\cdot(m\ddot{x} + 2b\dot{x} + kx) = 0 \tag{4}$$

The Helmholtz condition (3.2) then reads $2bf = \frac{d}{dt}(mf)$, the implication being that if in (4) we set $f(t) = e^{(2b/m)t}$ then the resulting equation *is* derivable

[1] See CLASSICAL MECHANICS (1983/84), page 119.

from a Lagrangian. And indeed: from

$$L(\dot{x}, x, t) = e^{(2b/m)t} \cdot \left\{ \tfrac{1}{2} m\dot{x}^2 - \tfrac{1}{2} kx^2 \right\} \tag{5}$$

we are led to

$$e^{(2b/m)t} \cdot (m\ddot{x} + 2b\dot{x} + kx) = 0$$

The appearance of (5) makes it natural to introduce a new variable

$$X \equiv e^{(b/m)t} x \tag{6}$$

Then $$e^{(b/m)t}\dot{x} = \dot{X} - \tfrac{b}{m} X$$

and the Lagrangian becomes

$$L(\dot{X}, X) = \tfrac{1}{2} m\dot{X}^2 - \tfrac{1}{2} \tfrac{mk - b^2}{m} X^2 - \tfrac{1}{2} b \cdot \tfrac{d}{dt} X^2$$

where the final term can be abandoned ("gauged away"). When that is done one has

$$= \tfrac{1}{2} m\dot{X}^2 - \tfrac{1}{2} k' X^2 \tag{7}$$

with $k' \equiv k - (b^2/m)$. A typical solution of the damped oscillator equation (2) is

$$x(t) = x_0\, e^{-(b/m)t} \sin \omega' t$$

with $\omega' \equiv \sqrt{k'/m}$. The curve traced in phase space by $\{x(t), p(t) \equiv \dot{x}(t)/m\}$ is shown in the following figure:

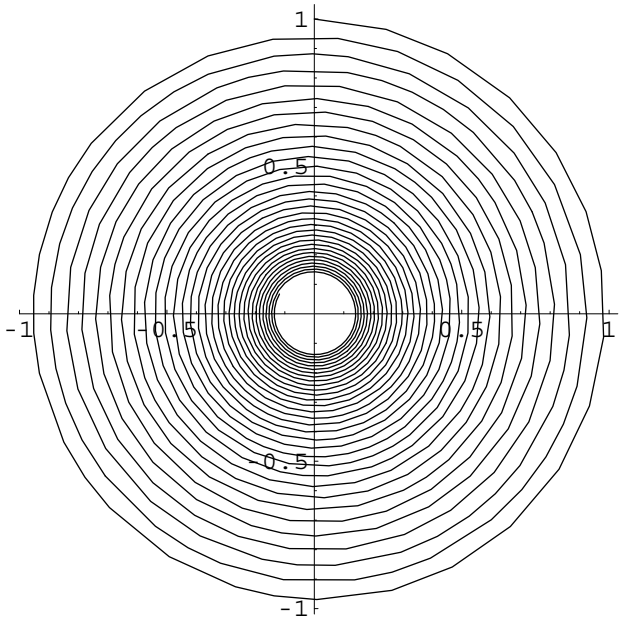


FIGURE 1: *Death in phase space of a damped harmonic oscillator.*

The variables $\{X, P\}$ inflate in such a way as to hold death at bay. I have seen people use the "inflation trick" as the basis of attempts to construct a "quantum theory of RLC circuits," in which the point of departure is provided by an equation

$$L\ddot{Q} + R\dot{Q} + C^{-1}Q = 0$$

that is structurally identical to (2).

2. Quadratic Lagrangians. In Lagrangian mechanics, linearity of the equations of motion presumes quadraticity of the Lagrangian. The most general such Lagrangian, if $\dot{\boldsymbol{x}}$ and $\boldsymbol{x}$ are all one has to work with, can be written

$$L(\dot{\boldsymbol{x}}, \boldsymbol{x}) = \tfrac{1}{2}\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{M}\dot{\boldsymbol{x}} + \dot{\boldsymbol{x}}^{\mathsf{T}}(\mathbb{S} + \mathbb{A})\boldsymbol{x} - \tfrac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\boldsymbol{x}$$

where $\mathbb{M}$ and $\mathbb{K}$ can, without loss of generality, be assumed to be symmetric (antisymmetric parts, if present, would fail to survive the summation process), and where we understand $\mathbb{S}$ and $\mathbb{A}$ to be the symmetric/antisymmetric parts of a matirx with presently unspecified symmetry. Actually

$$\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{S}\boldsymbol{x} = \tfrac{1}{2}\tfrac{d}{dt}\boldsymbol{x}^{\mathsf{T}}\mathbb{S}\boldsymbol{x}$$

so the $\mathbb{S}$-term can be gauged away, leaving us with

$$L(\dot{\boldsymbol{x}}, \boldsymbol{x}) = \tfrac{1}{2}\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{M}\dot{\boldsymbol{x}} + \dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{A}\boldsymbol{x} - \tfrac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\boldsymbol{x} \tag{8}$$

The resulting equations of motion read

$$\tfrac{d}{dt}\big\{\mathbb{M}\dot{\boldsymbol{x}} + \mathbb{A}\boldsymbol{x}\big\} - \big\{-\mathbb{A}\dot{\boldsymbol{x}} - \mathbb{K}\boldsymbol{x}\big\} = \boldsymbol{0}$$

or finally

$$\mathbb{M}\ddot{\boldsymbol{x}} + 2\mathbb{A}\dot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \boldsymbol{0} \tag{9}$$

Note that it is from the Lagrangian origin of these equations that the matrices $\mathbb{M}$, $\mathbb{A}$ and $\mathbb{K}$ have acquired their enforced symmetry properties: if we were willing to do without the support apparatus provided by Lagrangian formalism then we could relax those requirements.[2]

The momentum conjugate to $\boldsymbol{x}$ is

$$\boldsymbol{p} = \mathbb{M}\dot{\boldsymbol{x}} + \mathbb{A}\boldsymbol{x} \tag{10}$$

so the Hamiltonian $H(\boldsymbol{x}, \boldsymbol{p}) = \dot{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{p} - L(\dot{\boldsymbol{x}}, \boldsymbol{x})$ is found after some straightforward calculation to be given by

$$H(\boldsymbol{x}, \boldsymbol{p}) = \tfrac{1}{2}(\boldsymbol{p} - \mathbb{A}\boldsymbol{x})^{\mathsf{T}}\mathbb{M}^{-1}(\boldsymbol{p} - \mathbb{A}\boldsymbol{x}) + \tfrac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\boldsymbol{x} \tag{11}$$

Equations (10) and (11) have a distinctly "electrodynamic" look about them.

[2] Readers may recall that our discussion of the theory of celts culminated (Chapter 2, page 80) in equations of precisely the form (9).

3. Elimination of the gyroscopic term. I have first to establish the possibility of writing

$$\mathbb{M} = \mathbb{N}^2$$

We expect the symmetric matrix $\mathbb{M}$ to have "mass-like" (therefore positive) eigenvalues m_i. Let $\mathbb{D}$ denote the rotational "diagonalizer" of $\mathbb{M}$:

$$\mathbb{D}\mathbb{M}\mathbb{D}^{\intercal} = \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & m_n \end{pmatrix}$$

Then

$$\begin{aligned} \mathbb{M} &= \mathbb{D}^{\intercal} \begin{pmatrix} \sqrt{m_1} & 0 & \dots & 0 \\ 0 & \sqrt{m_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sqrt{m_n} \end{pmatrix} \mathbb{D} \cdot \mathbb{D}^{\intercal} \begin{pmatrix} \sqrt{m_1} & 0 & \dots & 0 \\ 0 & \sqrt{m_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sqrt{m_n} \end{pmatrix} \mathbb{D} \\ &= \mathbb{N} \cdot \mathbb{N} \end{aligned}$$

where the matrix $\mathbb{N}$ is symmetric, and can be assumed to be real. Since we can assign signs independently to the radicals there are a total of 2^n such "square roots of $\mathbb{M}$."

Now multiply $\mathbb{N}^{-1}$ into (9) to obtain

$$\mathbb{N}\ddot{\boldsymbol{x}} + 2\mathbb{N}^{-1}\mathbb{A}\mathbb{N}^{-1} \cdot \mathbb{N}\dot{\boldsymbol{x}} + \mathbb{N}^{-1}\mathbb{K}\mathbb{N}^{-1} \cdot \mathbb{N}\boldsymbol{x} = \boldsymbol{0}$$

which we will agree to notate

$$\ddot{\boldsymbol{x}} + 2\mathbb{A}\dot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \boldsymbol{0} \tag{12.1}$$

Notice now that $\left(\frac{d}{dt} + \mathbb{A}\right)^2 \boldsymbol{x} = \ddot{\boldsymbol{x}} + 2\mathbb{A}\dot{\boldsymbol{x}} + \mathbb{A}^2$ so (12) can be written

$$\left(\tfrac{d}{dt} + \mathbb{A}\right)^2 \boldsymbol{x} + \hat{\mathbb{K}}\boldsymbol{x} = \boldsymbol{0} \quad \text{with} \quad \hat{\mathbb{K}} \equiv \mathbb{K} - \mathbb{A}^2 \tag{12.2}$$

We had occasion to remark already at (141.2) in Chapter 2 that if $\mathbb{W}$ is a t-dependent rotation matrix and

$$\mathbb{W}\boldsymbol{x} = \boldsymbol{x}$$

then

$$\mathbb{W}\ddot{\boldsymbol{x}} = \left(\tfrac{d}{dt} + \mathbb{A}\right)^2 \boldsymbol{x}$$

where $\mathbb{A} \equiv -\dot{\mathbb{W}}\mathbb{W}^{-1}$. The implication is that (12) can be written

$$\ddot{\boldsymbol{x}} + \hat{\mathbb{K}}\boldsymbol{x} = \boldsymbol{0} \quad \text{with} \quad \hat{\mathbb{K}} \equiv \mathbb{W}^{\intercal}\hat{\mathbb{K}}\mathbb{W} \tag{13}$$

in which connection we note especially the symmetry of $\hat{\mathbb{K}}$. We have managed to "rotate away" the "gyroscopic $\mathbb{A}$-term" that was a conspicuous feature of (9), and was still present in (12.1). Had that term been *absent* from (9) the procedure that led to (12.1) would retain its utility, but the steps that led on to (13) would collapse into pointless triviality.

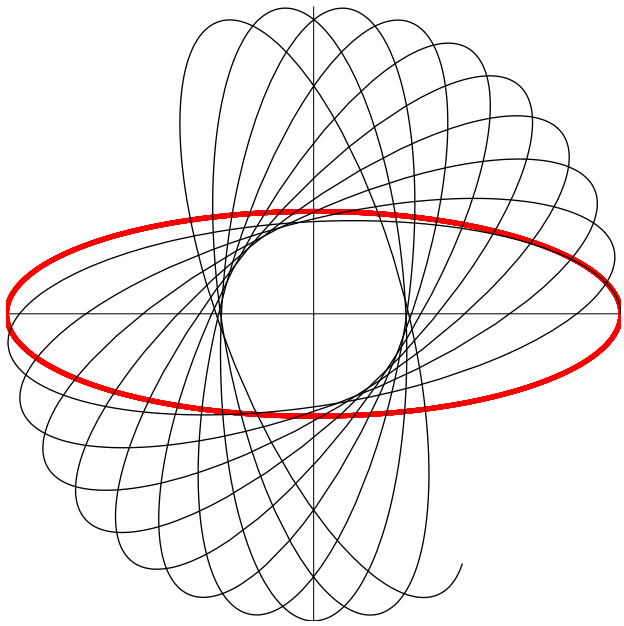


FIGURE 2: *Demonstration of the effect of the gyroscopic term. The red curve resulted from setting*

$$\mathbb{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{A} = \mathbb{O} \quad \textit{and} \quad \mathbb{K} = \begin{pmatrix} 3^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$

with $x(0) = 1$, $y(0) = 0$, $\dot{x}(0) = 0$, $\dot{y}(0) = 1$. *The precessing figure resulted from introducing a gyroscopic term with*

$$\mathbb{A} = \tfrac{1}{5} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

That the gyroscopic effect can be rotated away is not at all surprising on this evidence, but is not an obvious implication of some of the wild figures produced at other points in parameter space.

"Rotational elimination of the gyroscopic term" is reminiscent of the "inflationary elimination of the damping term" that we encountered on page 4. Indeed, the two terms display an obvious kinship, in that both depend linearly on velocity, both act to depress the "spring potential."[3] But there are, at the same time, some important differences: the gyroscopic term is necessarily absent from 1-dimensional theory (there is no such thing as a "non-trivial antisymmetric 1×1 matrix"), and the damping term necessarily absent from

[3] At (7) we had $k \mapsto k' \equiv k - (b^2/m)$; at (12.2) we had $\mathbb{K} \mapsto \hat{\mathbb{K}} \equiv \mathbb{K} - \mathbb{A}^2$.

Lagrangian theory (unless, perhaps, one succeeds in playing a generalization of the game played on page 4). The most important distinction has to do, however, with the energetics of the situation:

In most physical (meaning all mechanical) applications of the present theory we expect

- $\frac{1}{2}\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{M}\dot{\boldsymbol{x}}$ to describe the *kinetic energy* of the system of vibrating masses
- $\frac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\,\boldsymbol{x}$ to describe the *potential energy* stored in the "springs"

but it is not clear what (if any) energetic interpretation we should attach to the gyroscopic term $\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{A}\boldsymbol{x}$ in the Lagrangian (8).[4] That *something* is conserved is known, however, to follow already from the time-translation invariance[5] of the system (Noether's theorem), and the identity of that "something" was known already to Jacobi: it is

$$J = \sum_i x^i \frac{\partial L}{\partial \dot{x}^i} - L = \tfrac{1}{2}\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{M}\dot{\boldsymbol{x}} + \tfrac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\boldsymbol{x} \tag{14}$$

Note the absence of any reference to $\mathbb{A}$! By computation we verify that J is in fact conserved:

$$\begin{aligned}\tfrac{d}{dt}J &= \dot{\boldsymbol{x}}^{\mathsf{T}}(\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x}) \\ &= -2\dot{\boldsymbol{x}}\,\mathbb{A}\,\dot{\boldsymbol{x}} \quad \text{by the equations of motion (9)} \\ &= 0 \qquad\quad \text{because } \mathbb{A} \text{ is antisymmetric}\end{aligned}$$

Had a (non-Lagrangian) term $\mathbb{S}\dot{\boldsymbol{x}}$ been introduced into the equations of motion we would have obtained

$$\tfrac{d}{dt}J = -2\dot{\boldsymbol{x}}\,\mathbb{S}\,\dot{\boldsymbol{x}}$$

Evidently the effect of such a term would be to destroy J-conservation.

Looking back from our present vantage point to the celt equations encountered on page 80, we would expect to be able to show that, while the $\mathbb{A}$-term lends rotational directionality to the motion of $\boldsymbol{n}$, it is the $\mathbb{S}$-term that in some parts of parameter space accounts for the instability (inverse dissipation).

In 1948 B. D. H. Tellegen, the celebrated circuit theorist, described[6] a "new passive circuit element" (additional to the resistor, capacitor, inductor and ideal transformer) and presented FIGURE 3 to describe the action of such a device. He also described how such a device might be realized in practice. In what appear to me to be all essential respects (except those having to do with practical

[4] In non-mechanical applications of the theory (applications to circuit theory, for example) even the kinetic/potential interpretations of the other two terms become physically dubious.

[5] Which would be lost if we allowed $\mathbb{M}$, $\mathbb{A}$ or $\mathbb{K}$ to be time-dependent.

[6] "The gyrator. A new electric circuit element," Philips Laboratory Research Reports **3**, 81 (1948). The paper is reprinted in M. E. Valkenburg (editor), *Circuit Theory: Foundations and Classical Contributions* (1974).

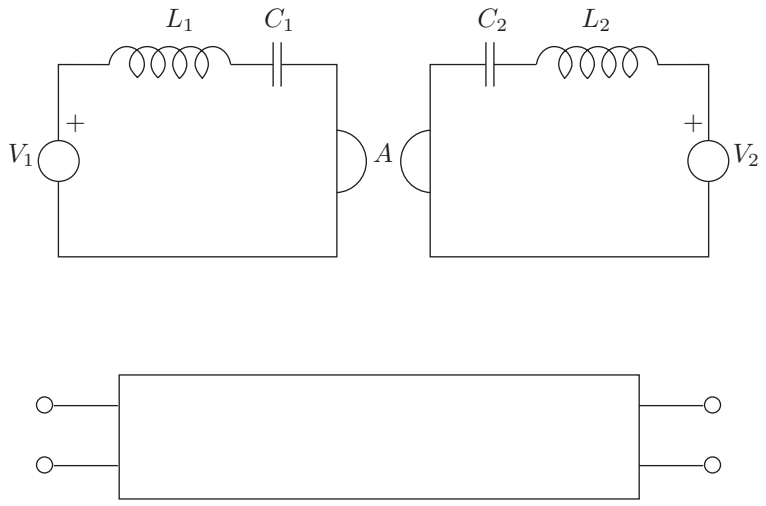


FIGURE 3: *Above*: *Tellegen's schematic representation of two circuits that have been coupled by means of a "gyrator." To describe the action of the device he writes*

$$L_1\ddot{Q}_1 - A\dot{Q}_1 + C_1^{-1}Q_1 = V_1$$
$$L_2\ddot{Q}_2 + A\dot{Q}_2 + C_2^{-1}Q_2 = V_2$$

Below: *A figure intended to emphasize that we are talking here about a 4-port circuit. Gyroscopic coupling—like mutual inductance—has nothing to contribute to the theory of 2-port devices. But whereas mutual inductance—when applicable—is represented by* symmetric *matrix* (*off-diagonal elements of the electrical analog of the "mass matrix"*), *gyroscopic coupling is represented by an* antisymmetric *matrix, and therefore stands in violation of what circuit theorists call the "reciprocity relation."*

realization), Tellegen had been anticipated by A. Bloch,[7] who—though he emphasizes the mechanical applications of the subject—also writes at length about their electrical analogs. An elaborate account of the subject—which, though addressed mainly to the realization problem, also treats its history and theory in great detail (one finds here illustrations of gyroscopes spinning at the ends of pendular strings)—was published in 1952 by C. L. Hogan.[8]

[7] "A new approach to the dynamics of systems with gyroscopic coupling terms," Phil. Mag. **35**, 315 (1943). Bloch's paper is actually cited by Tellegen, who mentions that the mechanical theory had been sketched almost a century earlier in Part I, §345VI of W. Thomson & P. W. Tait's monumental *Treatise on Natural Philosophy* (1873).

[8] "The ferromagnetic Faraday effect at microwave frequencies and its applications," Bell System Technical Journal **31**, 1 (1952).

4. Solution strategies: one degree of freedom. The differential equation

$$m\ddot{x} + 2b\dot{x} + kx = F(t)$$

—which describes the motion of a forced damped harmonic oscillator, and can (in the case $b = 0$) be considered to have derived from

$$L(\dot{x}, x, t) = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}kx^2 + xF(t) \tag{15}$$

—is simple enough that it can be solved in a great many ways. My emphasis here will be on methods that either extend naturally to the theory of coupled systems of oscillators, or that are especially instructive in some other way. Division by m gives

$$\ddot{x} + 2\beta\dot{x} + \omega^2 x = f(t) \tag{16}$$

which when specialized in various ways supplies the equations with which we will actually be working.

Introduce the notation

$$D \equiv \tfrac{d}{dt}$$

and observe that (16) can be written

$$P(D)x(t) = f(t)$$

where $P(D)$ is the *linear differential operator* that can in the present instance be described

$$\begin{aligned} P(D) &= D^2 + 2\beta D + \omega^2 \\ &= \big(D + \beta + i\sqrt{\omega^2 - \beta^2}\big)\big(D + \beta - i\sqrt{\omega^2 - \beta^2}\big) \\ &\equiv A_+(D) \cdot A_-(D) \end{aligned}$$

Turn off the forcing function and observe that

$$\text{if} \quad A_\pm(D)z(t) = 0 \quad \text{then certainly} \quad P(D)z(t) = 0$$

But

$$\begin{aligned} A_+(D)z(t) = 0 \quad &\Longrightarrow \quad z(t) = z_+ \cdot e^{-(\beta + i\sqrt{\omega^2 - \beta^2})\,t} \\ A_-(D)z(t) = 0 \quad &\Longrightarrow \quad z(t) = z_- \cdot e^{-(\beta - i\sqrt{\omega^2 - \beta^2})\,t} \end{aligned}$$

Linearity permits us to assert now that

$$z(t) = z_+ \cdot e^{-(\beta + i\sqrt{\omega^2 - \beta^2})\,t} + z_- \cdot e^{-(\beta - i\sqrt{\omega^2 - \beta^2})\,t}$$

comprises the *general* solution of $P(D)z(t) = 0$. And because $P(D)$ is a *real* linear operator we know that $\Re[z(t)]$ and $\Im[z(t)]$ are *by themselves* solutions of the unforced damped oscillator equation. We are brought thus to the familiar

conclusion that the general solution of

$$\ddot{x} + 2\beta\dot{x} + \omega^2 x = 0$$

can be described[9]

$$x(t) = e^{-\beta t}\Big\{A\cos\big[\sqrt{\omega^2-\beta^2}\,t\big] + B\sin\big[\sqrt{\omega^2-\beta^2}\,t\big]\Big\} \tag{17}$$

In the presence of forcing it is tempting to write

$$x(t) = [P(D)]^{-1} f(t) \tag{18}$$

which would provide a solution if only we could assign meaning to the operator on the right. To that end we notice that for all $g(t)$

$$De^{at}g(t) = e^{at}(D+a)g(t)$$

so we have the "shift rule"

$$D + a = e^{-at}De^{at}$$

Evidently

$$(D+a)^n = e^{-at}D^n e^{at} \quad : \quad n = 0, 1, 2, \ldots$$

and by natural formal extension

$$(D+a)^{-1} = e^{-at}\int_0^t d\tau\, e^{a\tau} \tag{19}$$

(which has at least this to recommend it: it entails $(D+a)(D+a)^{-1} = 1$). Bring (19) to this more detailed statement of (18)

$$x(t) = (D+\beta+i\Omega)^{-1}(D+\beta-i\Omega)^{-1}f(t)$$
$$\Omega \equiv \sqrt{\omega^2-\beta^2} \tag{20}$$

---

[9] The assumption here is that the unforced oscillator is either undamped ($\beta = 0$) or underdamped ($\beta^2 < \omega^2$). To deal with overdamped cases ($\beta^2 > \omega^2$) we might use $\cos(i\theta) = \cosh(\theta)$ and $\sin(i\theta) = i\sinh(\theta)$, but the critically damped case ($\beta^2 = \omega^2$) requires special treatment. For in that case the operators $A_\pm(D)$ become coincident: we have

$$P(D) = (D+\beta)^2$$

The function $e^{-\beta t}$ is a solution of $(D+\beta)x(t) = 0$, whence of $P(D)x(t) = 0$. But the function $te^{-\beta t}$, while also a solution of $P(D)x(t) = 0$, is *not* a solution of $(D+\beta)x(t) = 0$. I do not belabor these points because they are irrelevant to theories that take an orthodox Lagrangian as their point of departure.

and obtain

$$x(t) = e^{-(\beta+i\Omega)t} \int_0^t e^{(\beta+i\Omega)\sigma} \left[e^{-(\beta-i\Omega)\sigma} \int_0^\sigma e^{(\beta-i\Omega)\tau} f(\tau)\, d\tau \right] d\sigma$$

which a little manipulation serves to cast in much more attractive form. Write

$$= e^{-(\beta+i\Omega)t} \iint_{\mathcal{R}} e^{(\beta-i\Omega)\tau} f(\tau) e^{2i\Omega\sigma}\, d\tau d\sigma$$

where $\mathcal{R}$ refers to the triangular domain $0 \leqslant \tau \leqslant \sigma \leqslant t$. Reverse the order of integration to obtain

$$\begin{aligned} &= e^{-(\beta+i\Omega)t} \int_0^t e^{(\beta-i\Omega)\tau} f(\tau) \left[\int_\tau^t e^{2i\Omega\sigma} d\sigma \right] d\tau \\ &= \int_0^t e^{-\beta(t-\tau)} \frac{\sin \Omega(t-\tau)}{\Omega} f(\tau)\, d\tau \qquad : \text{ underdamped} \end{aligned} \tag{21}$$

from which follow

$$\begin{aligned} &\downarrow \\ &= \int_0^t e^{-\beta(t-\tau)} (t-\tau) f(\tau)\, d\tau \qquad : \text{ critically damped} \\ &\downarrow \\ &= \int_0^t e^{-\beta(t-\tau)} \frac{\sinh \sqrt{\beta^2-\omega^2}\,(t-\tau)}{\sqrt{\beta^2-\omega^2}} f(\tau)\, d\tau \quad : \text{ overdamped} \end{aligned}$$

However objectionably formal (informal?) has been the argument that led to (21), a quick calculation confirms that it does in fact describe a particular solution of (16), a solution for which

$$x(0) = \dot{x}(0) = 0$$

To that *particular* solution—call it $x_p(t)$—of (16) we need only add $x_0(t)$—the general solution (17) of the unforced equation—to obtain the *general solution of the forced oscillator equation*:

$$x(t) = x_0(t) + x_p(t)$$

In the presence of damping $x_0(t)$ dies in characteristic time $1/\beta$, and with it all memory of the initial data: the motion of x is determined entirely by recent activity of the forcing function $f(t)$. But in the absence of damping the $x_0(t)$ term is not "transcient:" its effects persist indefinitely.

5. A quick look at some properties of causal Green functions. It is clear from the argument that led to (21) that *all* functions of the form

$$x_a(t) \equiv \int_a^t e^{-\beta(t-\tau)} \frac{\sin \Omega(t-\tau)}{\Omega} f(\tau)\, d\tau$$

satisfy the forced oscillator equation (16).[10] It is as a matter merely of formal convenience that we concentrate here on the solution that results from setting $a = -\infty$. That done, we can write

$$x(t) = \int_{-\infty}^{+\infty} G(t-\tau) f(\tau)\, d\tau \tag{22}$$

$$G(t-\tau) \equiv e^{-\beta(t-\tau)} \frac{\sin \Omega(t-\tau)}{\Omega}\, \theta(t-\tau) \tag{23}$$

where $\theta(t-\tau)$ is the Heaviside step function (known to *Mathematica* as the `UnitStep` function):

$$\theta(t-\tau) = \int_{-\infty}^{t} \delta(u-\tau)\, du = \begin{cases} 1 & : \quad \tau < t \\ 0 & : \quad \tau > t \end{cases}$$

The Green function admits of two complementary interpretations:

- $G(t-\tau)$ describes the *weight* with which successive past values $f(\tau)$ of input contribute to the present output $x(t)$: see FIGURE 4
- $G(t-\tau)$ describes the lingering output that results from a *unit kick* at time $\tau < t$ (FIGURES 5), and it is in *this* sense that it provides a realization of Green's original idea.

We can (as was anticipated in the language of the preceding paragraph) look upon (22) as a description of how a damped harmonic oscillator —assembled from mass, spring and dashpot—functions as a device to convert "input" $f(t)$ to "output" $x(t)$. It informs us that the action of the device is

- **linear**: superimposed input produces superimposed output
- **time-independent**: time-translated input produces time-translated output (this because t and τ enter into the Green function only through their difference)
- **causal**: output depends only upon *past* input—this by action of the step function $\theta(t-\tau)$, that renders the future invisible/inconsequential.

We expect similar properties to attach to the Green function of *any* physical system (many-particle population of vibrating particles, electromagnetic field,

[10] Evidently $x_{ab}(t) \equiv x_a(t) - x_b(t)$ is a solution of the *un*forced oscillator equation. It is evident also that in the presence of damping $x_a(t)$ and $x_b(t)$ —assume $a > b$—will become very nearly identical at times $t - a \gg 1/\beta$, when both solutions act upon memory of the same recent past.

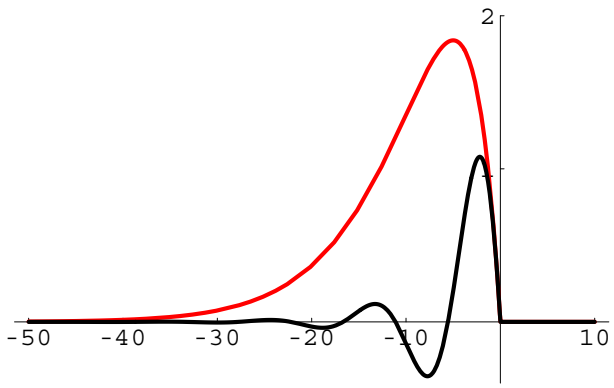


FIGURE 4: *Write $G(t,\tau;\beta,\omega)$ to make the parameter-dependence of the oscillator Green function explicit. Shown here are graphs of $G(0,\tau;0.2,0.2)$ and $G(0,\tau;0.2,0.6)$, illustrating how the Green function serves to weight past values of input to create the present value $x(0)$ of output. The red curve pertains to a critically damped case ($\beta=\omega$), the black curve to an underdamped case ($\beta<\omega$).*

idealized electronic component of the sort that does not have to be connected to an external power source) of time-independent design that acts linearly and causally.

Fourier analysis ("linear algebra in function space") can be used to provide deepened insight into the points at issue, by methods that I will sketch as they pertain specifically to the dampened harmonic oscillator.[11] Write

$$x(t) = \tfrac{1}{\sqrt{2\pi}}\int \xi(\alpha)\, e^{i\alpha t} d\alpha$$

$$f(t) = \tfrac{1}{\sqrt{2\pi}}\int \varphi(\alpha)\, e^{i\alpha t} d\alpha$$

$$G(t) = \tfrac{1}{2\pi}\int \Gamma(\alpha)\, e^{i\alpha t} d\alpha$$

In this notation (22)—with its convolution on the right—is readily brought to the form

$$\xi(\alpha) = \Gamma(\alpha)\varphi(\alpha)$$

The differential equation (16) supplies, on the other hand,

$$\big[-\alpha^2 + 2i\beta\alpha + \omega^2\big]\xi(\alpha) = \varphi(\alpha)$$

of which the solution has become a matter of simple algebra: we have

$$\Gamma(\alpha) = -\frac{1}{[\alpha-(+\Omega+i\beta)][\alpha-(-\Omega+i\beta)]}$$

[11] For more general discussion, see pages 46–51 in CLASSICAL THEORY OF FIELDS (1965/66).

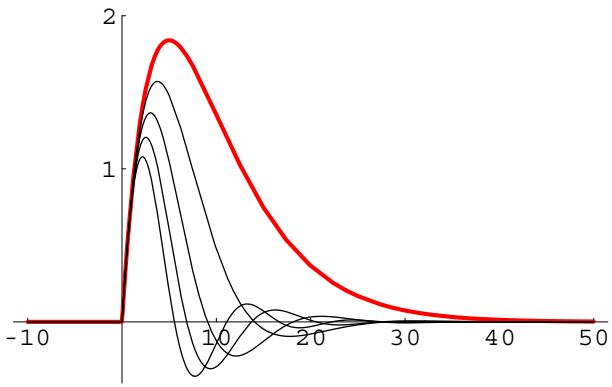


FIGURE 5A: *Graphs of* $G(t,0;0.2,\omega)$ *showing the output* $x(t)$ *that results from a unit kick at time* $\tau = 0$: $f(\tau) = \delta(\tau)$. *The red response is critically damped* $(\omega = \beta = 0.2)$; *the black curves are progressively more underdamped*: $\omega = 0.3, 0.4, 0.5, 0.6$.

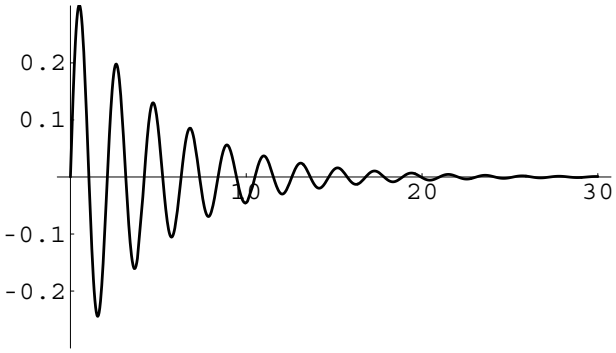


FIGURE 5B: *Graph of* $G(t,0;0.2,3.0)$, *showing response of a more underdamped oscillator to a unit kick at time* $t = 0$.

giving (by inverse Fouriertransformation)

$$x(t) = -\tfrac{1}{\sqrt{2\pi}} \int \frac{1}{[\alpha-(+\Omega+i\beta)][\alpha-(-\Omega+i\beta)]}\, \varphi(\alpha)\, e^{i\alpha\, t}\, d\alpha$$

Introduce

$$\varphi(\alpha) = \tfrac{1}{\sqrt{2\pi}} \int f(\tau)\, e^{i\alpha\, \tau}\, d\tau$$

and, after interchanging the order of integration, obtain

$$x(t) = \int \left\{ -\tfrac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{[\alpha-(+\Omega+i\beta)][\alpha-(-\Omega+i\beta)]} e^{i\alpha(t-\tau)}\, d\alpha \right\} f(\tau)\, d\tau \qquad (24)$$

To evaluate $\{etc.\}$ we (i) complexify α ($\alpha \mapsto a + ib$), (ii) replace the $\int_{-\infty}^{+\infty}$ with a contour integral $\oint_C$ (see FIGURES 6) and have recourse to the calculus of residues: specifically, we draw upon the identity ("method of partial fractions")

$$\frac{1}{(x-A)(x-B)} = \frac{1}{A-B}\Big[\frac{1}{x-A} - \frac{1}{x-B}\Big]$$

and the Cauchy integral theorem

$$f(a) = \tfrac{1}{2\pi i}\oint_C \frac{1}{z-a} f(z)\,dz$$

(the assumption here being that $f(z)$ is analytic on the region $\mathcal{R}$ bounded by $\mathrm{C} = \partial\mathcal{R}$) to write

$$\Big\{etc.\Big\} = \frac{1}{2i\Omega}\left[\tfrac{1}{2\pi i}\oint_C \tfrac{1}{\alpha-(+\Omega+i\beta)} e^{i\alpha(t-\tau)}d\alpha - \tfrac{1}{2\pi i}\oint_C \tfrac{1}{\alpha-(-\Omega+i\beta)} e^{i\alpha(t-\tau)}d\alpha\right]$$

$$= \begin{cases} \dfrac{1}{2i\Omega}\Big[e^{i(+\Omega+i\beta)(t-\tau)} - e^{i(-\Omega+i\beta)(t-\tau)}\Big] & : \quad t > \tau \\ 0 & : \quad t < \tau \end{cases}$$

$$= e^{-\beta(t-\tau)}\frac{\sin\Omega(t-\tau)}{\Omega}\,\theta(t-\tau)$$

... which brings (25) into precise agreement with (22/23).

So Fourier analysis has taught us what we already knew. But it has, as will emerge, done more: it has cast new light on what we knew, drawn attention to the importance of $\Gamma(\alpha)$—the complex-valued Fourier transform of the Green function—and illustrated the truth of the general proposition that *to understand clearly what is going on one must get off the real line onto the complex plane*. Note first that if one assumes harmonic forcing $f(t) = f \cdot e^{i\alpha t}$ and looks for similarly harmonic response, one is led immediately from the equation of motion (16) to the statement

$$x(t) = f \cdot \Gamma(\alpha)\, e^{i\alpha t}$$

where[12] $\Gamma(\alpha)$ can be described

$$\begin{aligned}\Gamma(\alpha) &= \frac{1}{(\omega^2-\alpha^2)+2i\beta\alpha} \\ &= \frac{\omega^2-\alpha^2}{(\omega^2-\alpha^2)^2+4\beta^2\alpha^2} + i\frac{-2\beta\alpha}{(\omega^2-\alpha^2)^2+4\beta^2\alpha^2} \\ &= \Re[\Gamma(\alpha)] + i\Im[\Gamma(\alpha)] \\ &= A(\alpha)e^{-i\delta(\alpha)}\end{aligned}$$

[12] See again (24).

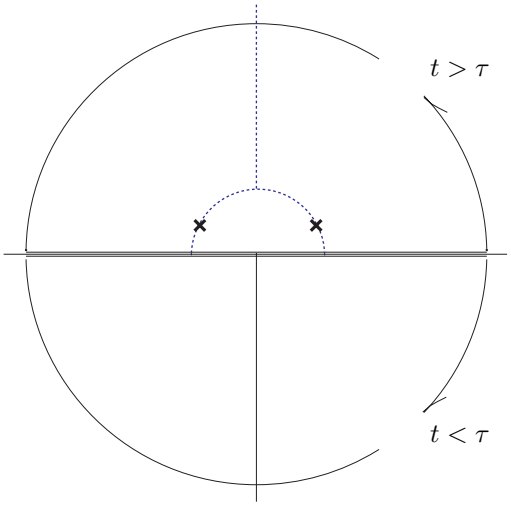


FIGURE 6A: *When $\beta > 0$ two poles lie in the upper halfplane. Contour closure $\circlearrowleft$ or $\circlearrowright$ is forced by the requirement that the factor*

$$e^{i\alpha(t-\tau)} = e^{ia(t-\tau)} \cdot e^{-b(t-\tau)}$$

must kill the integrand on the "return arc at infinity." The contour envelops poles when $t > \tau$, none when $t < \tau$. It is "contour flipping" that produces the step function in (23)*, and damping that makes causality automatic. The dashed blue curve shows the locus of the poles as β is increased with ω held constant. Critical damping occurs at the top of the blue arc.*

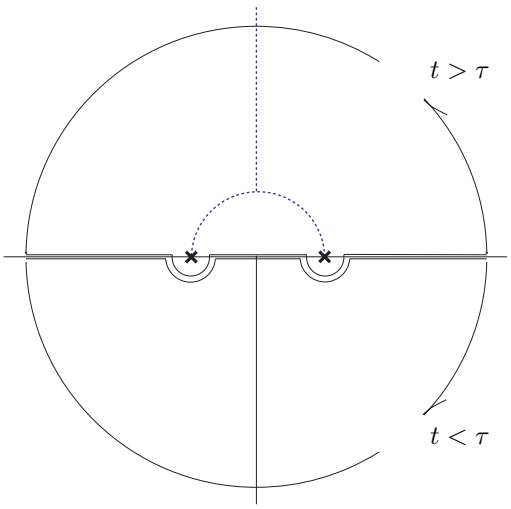


FIGURE 6B: *Causality-preserving management of the contour in the absence of damping.*

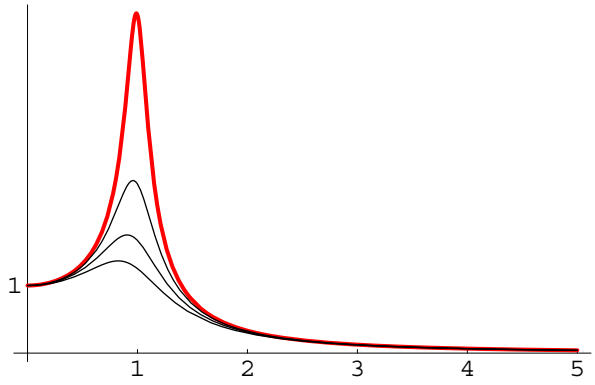


FIGURE 7A: *Amplitude of a harmonically driven damped oscillator, plotted as a function of driving frequency α. The damping parameter β ranges upward from the value* 0.1 (*red curve*) *through the values* 0.2, 0.3, 0.4. *In all cases* $\omega = 1.0$. *Note that the resonant peak shifts slightly downward as β is increased.*

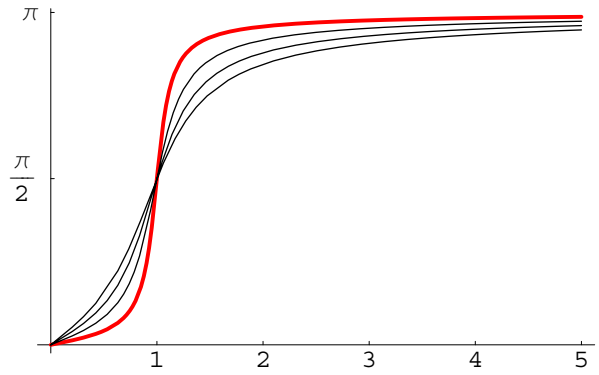


FIGURE 7B: *Relative phases of the harmonic stimulus and response, for the same values of ω and β as were used to construct the preceding figure.*

where evidently

$$A(\alpha) = \frac{1}{\sqrt{(\omega^2 - \alpha^2)^2 + 4\beta^2\alpha^2}}$$

$$\delta(\alpha) = \arctan\left\{\frac{2\beta\alpha}{\omega^2 - \alpha^2}\right\}$$

These familiar functions are plotted in FIGURES 7.

The Fourier analytic line of argument served to trace "causality" to the circumstance that

$$\Gamma(\alpha) \equiv \int_{-\infty}^{+\infty} G(\tau) e^{-i\alpha\tau} d\tau = \int_{0}^{\infty} G(\tau) e^{-i\alpha\tau} d\tau$$

is analytic (*i.e.*, has no poles) on the lower half of the complex α plane. From this fact it follows (by an argument I am about to sketch) that[13]

$$\left.\begin{aligned} \Re[\Gamma(a)] &= -\tfrac{1}{\pi}\mathbf{P}\int_{-\infty}^{+\infty} \frac{1}{x-a}\Im[\Gamma(x)]\,dx \\ \Im[\Gamma(a)] &= +\tfrac{1}{\pi}\mathbf{P}\int_{-\infty}^{+\infty} \frac{1}{x-a}\Re[\Gamma(x)]\,dx \end{aligned}\right\} \tag{26}$$

according to which *either of the functions* $\Re[\Gamma(\alpha)]$ *and* $\Im[\Gamma(\alpha)]$ *serves to determine the other*. In circumstances where both of those functions can be independently measured in the laboratory one might check to see whether that data is consistent with the **Kramers-Krönig relations**[14] (26), and thus determine whether the "microsystem inside the black box" is or is not causal.[15] Now from motivation to the details:

Looking to the first of the following figures, and proceeding in the presumption that $\Gamma(\alpha)$

- is analytic on the lower halfplane and
- vanishes fast enough at the "edge" of the lower halfplane

[13] Here the **P** signals that one is to take the "principal part" of the singular integral, a concept that will be defined very shortly. Equations (26) announce that the real and imaginary parts of $\Gamma(\alpha)$ are "Hilbert transforms" of one another: see Chapter 15 in A. Erdélyi *et al*, *Tables of Integral Transforms* (1954) for a long list of examples of such transforms.

[14] This subject derives from work done by H. A. Kramers in 1927, to which R. Krönig made contributions in 1942.

[15] Such questions became lively in the 1960s, when difficulties encountered in theoretical elementary particle physics led people to question whether the microworld *is* causal. A population of relations analogous to (26)—known collectively as "dispersion relations"—was developed so that "causality question" could be addressed experimentally. See J. S. Toll, "Causality and the dispersion relation: logical foundations," Phys. Rev. **104**, 1760 (1956) for a lucid review of the subject and its early history. The term "dispersion relation" derives, by the way, from the circumstance that one early application served to relate the real to the imaginary (absorptive) part of the complex index of refraction $n(\omega)$, and it is from the frequency-dependence of phase velocity that the phenomenon of optical dispersion derives: see L. Mandel & E. Wolf, *Optical Coherence & Quantum Optics* (1995), §3.1.1.

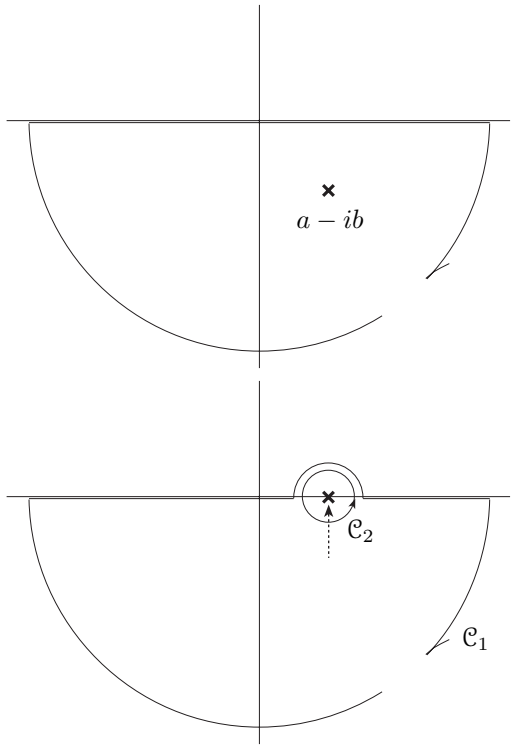


FIGURE 8: *In the upper figure a curve "bounds the lower halfplane," on which we have planted a pole at $\alpha = a - ib$. In the lower figure we have "turned b off" and placed a second contour—of reversed (which is to say: standard) sense $\circlearrowleft$—about the pole's repositioned location. We will take the radius ϵ of $\mathcal{C}_2$ down to zero.*

we have

$$\begin{aligned}\Gamma(a-ib) &= -\tfrac{1}{2\pi i}\oint_{\mathcal{C}_1}\frac{1}{z-(a-ib)}\Gamma(z)\,dz \\ &= -\tfrac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{1}{x-(a-ib)}\Gamma(x)\,dx\end{aligned}$$

Proceeding now to the limit $b \downarrow 0$ (lower figure) we have

$$\begin{aligned}\Gamma(a) &= -\tfrac{1}{2\pi i}\left\{\left[\int_{-\infty}^{a-\epsilon}+\int_{a+\epsilon}^{+\infty}\right]\frac{1}{x-a}\Gamma(x)\,dx - \tfrac{1}{2}\oint_{\mathcal{C}_2}\frac{1}{z-a}\Gamma(z)\,dz\right\} \\ &= -\tfrac{1}{2\pi i}\mathbf{P}\int_{-\infty}^{+\infty}\frac{1}{x-a}\Gamma(x)\,dx + \tfrac{1}{2}\Gamma(a)\end{aligned}$$

where the "principal part" of the improper integral is (with $\lim_{\epsilon\to 0}$ understood)

defined in the preceding line.[16] So we have

$$\Gamma(a) = -\tfrac{1}{\pi i}\mathbf{P}\int_{-\infty}^{+\infty} \frac{1}{x-a}\Gamma(x)\,dx$$

of which (26) presents the real/imaginary parts.

We digress to assure ourselves that the Kramers-Krönig relations (26) pertain accurately to the specific case of a damped harmonic oscillator.To that end we first sharpen our notation, writing

$$R(x;\omega,\beta) \equiv \Re[\Gamma(x)] = \frac{\omega^2 - x^2}{(\omega^2 - x^2)^2 + 4\beta^2 x^2}$$
$$S(x;\omega,\beta) \equiv \Im[\Gamma(x)] = \frac{-2\beta x}{(\omega^2 - x^2)^2 + 4\beta^2 x^2}$$

Experiments show that these functions are pretty wild when plotted after test values have been assigned to a, ω and β. And it appears that *Mathematica* cannot handle the symbolic integrals that appear on the right sides of (26). But the *numerical* integrals give it no trouble at all: we find, for example, that

$$R(2.5;1,0.2) = -0.183807$$
$$-\tfrac{1}{\pi}\Big[\texttt{NIntegrate[}\frac{S(x,1,0.2)}{x-2.5}\texttt{,}\ \{x,-\infty,2.4999\}\texttt{]}$$
$$+\texttt{NIntegrate[}\frac{S(x,1,0.2)}{x-2.5}\texttt{,}\ \{x,2.5001,+\infty\}\texttt{]}\Big] = -0.183804$$

and that inversely

$$S(2.5;1,0.2) = -0.035011$$
$$+\tfrac{1}{\pi}\Big[\texttt{NIntegrate[}\frac{R(x,1,0.2)}{x-2.5}\texttt{,}\ \{x,-\infty,2.4999\}\texttt{]}$$
$$+\texttt{NIntegrate[}\frac{R(x,1,0.2)}{x-2.5}\texttt{,}\ \{x,2.5001,+\infty\}\texttt{]}\Big] = -0.035021$$

These results are encouraging, particularly since they are duplicated when a, ω and β are randomly assigned other values, and are improved when ϵ (here $\epsilon = 0.0001$) is made smaller. But these are the results of *mathematical* experimentation. What *physical* data might we feed into (26) to test the causality of the system?

The steady response (response after start-up transcients have died) of a damped oscillator to harmonic stimulation $f(t) = f \cdot e^{i\alpha t}$ can be described[17]

$$\begin{aligned} x(t) &= \Re[\Gamma(\alpha) f e^{i\alpha t}] \\ &= f A(\alpha)\cos[\alpha t - \delta(a)] \end{aligned}$$

[16] See also (for example) E. T. Whittaker & G. N. Watson, *Modern Analysis* (1902), pages 75 and 117.

[17] See again the bottom of page 16.

so the energy

$$\begin{aligned} E(t) = \tfrac{1}{2}m\big[\dot{x}^2(t) + \omega^2 x^2(t)\big] &= \tfrac{1}{2}mf^2A^2\big\{\alpha^2\sin^2(\alpha t-\delta) + \omega^2\cos^2(\alpha t-\delta)\big\} \\ &= \tfrac{1}{2}mf^2A^2\big\{\alpha^2 + (\omega^2-\alpha^2)\cos^2(\alpha t-\delta)\big\} \end{aligned}$$

surges periodically, and when averaged over a period becomes

$$\begin{aligned} \langle E\rangle &\equiv \tfrac{1}{\tau}\int_0^\tau E(t)\,dt \\ &= \tfrac{1}{2}mf^2A^2\cdot\tfrac{1}{2}(\omega^2+\alpha^2) \end{aligned}$$

Therefore

$$\begin{aligned} E(t) - \langle E\rangle &= \tfrac{1}{2}mf^2A^2(\omega^2-\alpha^2)\big\{\cos^2(\alpha t-\delta) - \tfrac{1}{2}\big\} \\ &= \tfrac{1}{4}mf^2\Re[\Gamma(\alpha)]\cos 2(\alpha t-\delta) \end{aligned}$$

from which it follows that

$$\sigma_{_E} \equiv \sqrt{\langle[E(t)-\langle E\rangle]^2\rangle}$$

—which quantifies the extent to which $E(t)$ ripples about its mean value—can be described

$$\begin{aligned} &= \tfrac{1}{4}mf^2\Re[\Gamma(\alpha)]\sqrt{\langle\cos^2 2(\alpha t-\delta)\rangle} \\ &= \tfrac{1}{\sqrt{32}}mf^2\cdot\Re[\Gamma(\alpha)] \end{aligned} \tag{27}$$

This result carries us forward, but not quite far enough. For direct measurement of $\sigma_{_E}$ would appear to require that we be "inside the black box," which we consider to be disallowed.

Observe now that

$$\begin{aligned} \dot{E}(t) &= -\tfrac{1}{2}mf^2A^2\big\{(\omega^2-\alpha^2)\alpha\sin 2(\alpha t-\delta)\big\} \\ &= -\tfrac{1}{2}mf^2\,\Re[\Gamma(\alpha)]\cdot\alpha\sin 2(\alpha t-\delta) \end{aligned}$$

is related to the **instantaneous energy injection rate**

$$\begin{aligned} \mathcal{J}(t) &= \dot{x}\cdot mf(t) \\ &= -fA(\alpha)\alpha\sin(\alpha t-\delta)\cdot mf\cos\alpha t \end{aligned}$$

and the **instantaneous energy dissipation rate**

$$\begin{aligned} \mathcal{D}(t) &= \dot{x}\cdot 2m\beta\dot{x} \\ &= 2m\beta\,[-fA(\alpha)\alpha\sin(\alpha t-\delta)]^2 \end{aligned}$$

through

$$\begin{aligned} \mathcal{J}(t) - \mathcal{D}(t) &= \dot{x}\cdot(mf(t) - 2m\beta\dot{x}) \\ &= \dot{x}\cdot(m\ddot{x}+kx) \\ &= \dot{E}(t) \end{aligned}$$

The **averaged rate of energy injection** is therefore

$$\begin{aligned}\langle \mathcal{J} \rangle \equiv \tfrac{1}{\tau}\int_0^\tau \mathcal{J}(t)\,dt &= -mf^2A(\alpha)\alpha \cdot \tfrac{1}{\tau}\int_0^\tau \sin(\alpha t - \delta)\cos\alpha t\,dt \\ &= \tfrac{1}{\tau}\pi\, mf^2A(\alpha)\sin\Big(\arctan\frac{2\beta\alpha}{\omega^2-\alpha^2}\Big) \\ &= \tfrac{1}{\tau}\pi\, mf^2A(\alpha)\frac{2\beta\alpha}{\sqrt{(\omega^2-\beta^2)^2+4\beta^2\alpha^2}} \\ &= -\tfrac{1}{2}mf^2\alpha\cdot\Im[\Gamma(\alpha)] \end{aligned} \tag{28}$$

and we are not surprised to discover, by an almost identical calculation, that $\langle \mathcal{J} \rangle = \langle \mathcal{D} \rangle$: the oscillator does, on average, neither accumulate nor lose energy when harmonically driven. Notice that energy injection and dissipation rates $\mathcal{J}(t)$ and $\mathcal{D}(t)$ are measurable *from outside the box*, and that when taken together they permit one to infer the value of

$$E(t) = \int_0^t \big[\mathcal{J}(s) - \mathcal{D}(s)\big]\,ds + E(0)$$

whence of $\sigma_{_E}$ (which, we note, is insensitive to the (unmeasured) value of $E(0)$).

Which brings us to the point of this discussion. By (*i*) monitoring the instantaneous rates of energy injection and dissipation when the system is harmonically driven at various frequencies α, and by (*ii*) feeding that data into (27/28), one can assign "experimentally determined values" to the real and imaginary parts of the Fourier transform $\Gamma(\alpha)$ of the Green function. And by (*iii*) feeding that information into the Kramers-Krönig relations (26) one becomes able to announce (with a confidence limited only by measurement error) whether the linear passive system inside the box—whatever it is—is or is not causal. This procedure involves activity in the frequency-domain, and might be recommended when it is not feasible, by activity in the time-domain, to test *directly* whether it is true (FIGURES 5) that "output never precedes input."

An engineer is asked to design a device that achieves

$$\text{input signal } S_{\text{in}}(\alpha) \longrightarrow \text{output signal } S_{\text{out}}(\alpha) = \Gamma(\alpha)S_{\text{out}}(\alpha)$$

To realize prescribed operating characteristics he can "sprinkle poles as he likes on the upper half of the complex α plane," but is prevented by causality from allowing poles to dribble onto the lower halfplane.

6. Simultaneous diagonalization & its limitations. Equations of the form

$$\mathbb{M}\ddot{\boldsymbol{x}} + 2(\mathbb{S}+\mathbb{A})\dot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \boldsymbol{F}(t) \tag{29}$$

would be very easy to solve if all the matrices in question were *diagonal*, for the system (29) would then simply present n copies of the one-dimensional problem,

the solution of which we have already discussed at length.[18] If we knew how to simultaneously diagonalize a population of real symmetric matrices we would have an obvious solution strategy ready at hand, and it that mathematical issue to which we now turn our attention.

Recall how one undertakes to diagonalize a *solitary* symmetric matrix $\mathbb{M}$. One computes the eigenvalues $\{m_1, m_2, \ldots, m_n\}$ and normalized eigenvectors $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n\}$, and from the latter assembles

$$\mathbb{R} \equiv \begin{pmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \ldots & \boldsymbol{r}_n \\ \vert & \vert & & \vert \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

From $\mathbb{M}\boldsymbol{r}_i = m_i \boldsymbol{r}_i$ it then follows that

$$\mathbb{M}\mathbb{R} \equiv \begin{pmatrix} m_1\boldsymbol{r}_1 & m_2\boldsymbol{r}_2 & \ldots & m_n\boldsymbol{r}_n \\ \vert & \vert & & \vert \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

Therefore

$$\mathbb{R}^{\intercal}\mathbb{M}\mathbb{R} = \begin{pmatrix} \boldsymbol{r}_1 \longrightarrow \\ \boldsymbol{r}_2 \longrightarrow \\ \vdots \\ \boldsymbol{r}_n \longrightarrow \end{pmatrix} \cdot \begin{pmatrix} m_1\boldsymbol{r}_1 & m_2\boldsymbol{r}_2 & \ldots & m_n\boldsymbol{r}_n \\ \vert & \vert & & \vert \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

which by the known orthogonality of the (already normalized) eigenvectors becomes

$$= \begin{pmatrix} m_1 & 0 & \ldots & 0 \\ 0 & m_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & m_n \end{pmatrix}$$

Were we to omit the $\mathbb{M}$-factor we would have

$$\mathbb{R}^{\intercal}\mathbb{R} = \begin{pmatrix} \boldsymbol{r}_1\cdot\boldsymbol{r}_1 & \boldsymbol{r}_1\cdot\boldsymbol{r}_2 & \ldots & \boldsymbol{r}_1\cdot\boldsymbol{r}_n \\ \boldsymbol{r}_2\cdot\boldsymbol{r}_1 & \boldsymbol{r}_2\cdot\boldsymbol{r}_2 & \ldots & \boldsymbol{r}_2\cdot\boldsymbol{r}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{r}_n\cdot\boldsymbol{r}_1 & \boldsymbol{r}_n\cdot\boldsymbol{r}_2 & \ldots & \boldsymbol{r}_n\cdot\boldsymbol{r}_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

which shows $\mathbb{R}$ to be a rotation matrix.

[18] We recognize, of course, that the $\mathbb{S}$-term could never have arisen from a time-independent Lagrangian, and that it is too much to ask for a "non-trivially diagonal antisymmetric matrix"!

Now introduce the "contraction (or dilation) matrix"

$$\mathbb{C} = \begin{pmatrix} 1/\sqrt{m_1} & 0 & \dots & 0 \\ 0 & 1/\sqrt{m_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1/\sqrt{m_n} \end{pmatrix} = \mathbb{C}^{\intercal}$$

and obtain

$$\mathbb{M}' \equiv \mathbb{C}^{\intercal}\mathbb{R}^{\intercal}\mathbb{M}\,\mathbb{R}\,\mathbb{C} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

All we can say about $\mathbb{K}' \equiv \mathbb{C}^{\intercal}\mathbb{R}^{\intercal}\mathbb{K}\,\mathbb{R}\,\mathbb{C}$, however, is that it is real and symmetric. But that is enough. Let $\left\{\omega_1^2, \omega_2^2 \dots, \omega_n^2\right\}$ and $\left\{\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_n\right\}$ denote the eigenvalues and orthonormal eigenvectors of $\mathbb{K}'$, and from the latter assemble the rotation matrix

$$\mathbb{Q} \equiv \begin{pmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \\ \vert & \vert & & \vert \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

Arguing as before, we have

$$\mathbb{K}'' \equiv \mathbb{Q}^{\intercal}\mathbb{K}'\mathbb{Q} = \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \omega_n^2 \end{pmatrix} \tag{30}$$

while—and this is the point—$\mathbb{M}'$ is so simple that it *responds by invariance* to this last transformation:

$$\mathbb{M}'' \equiv \mathbb{Q}^{\intercal}\mathbb{M}'\mathbb{Q} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \mathbb{M}'$$

The upshot: if $\mathbb{W} \equiv \mathbb{R}\mathbb{C}\mathbb{Q}$ then

$$\left.\begin{aligned} \mathbb{W}^{\intercal}\mathbb{M}\,\mathbb{W} &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \\ \mathbb{W}^{\intercal}\mathbb{K}\,\mathbb{W} &= \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \omega_n^2 \end{pmatrix} \end{aligned}\right\} \tag{31}$$

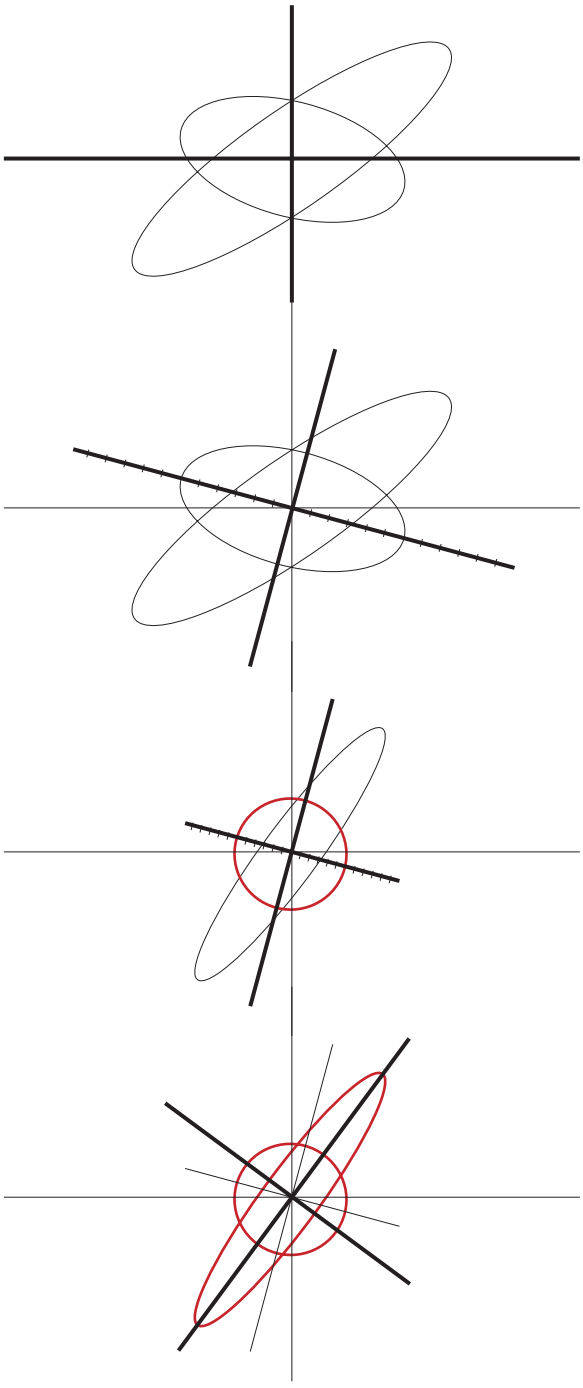


FIGURE 9: *Geometrical representation of the three-step procedure for simultaneously diagonalizing two quadratic forms* (*symmetric matrices*).

The sequence of steps that led to this result is illustrated in FIGURE 9, where in place of $\mathbb{M}$ itself we look to the associated "quadratic form" $\boldsymbol{x}^{\mathsf{T}}\mathbb{M}\boldsymbol{x}$, and to portray the latter we look actually to the graph of $\boldsymbol{x}^{\mathsf{T}}\mathbb{M}\boldsymbol{x} = 1$, which might (depending upon the spectrum of $\mathbb{M}$) be any conic section, but—simply to keep the diagram on the page—has been assumed to be elliptical.

It is clear from the figure that with the diagonalization of two quadratic forms we have exhausted the method's potential, for as a first step toward diagonalization of a third form we would have to stabilize the first two by rescaling *both* to circles (hyperspheres), which is clearly impossible. It should be noted that, though $\mathbb{R}$ and $\mathbb{Q}$ are dimensionless rotation matrices, $\mathbb{C}$ (at least in the intended mechanical application) has the physical dimension of $1/\sqrt{\text{mass}}$. So also, therefore, does $\mathbb{W}$. Which means that, while $\mathbb{W}^{\mathsf{T}}\mathbb{M}\mathbb{W}$ is dimensionless,

$$[\,\mathbb{W}^{\mathsf{T}}\mathbb{K}\mathbb{W}\,] = [k/m] = (\text{frequency})^2$$

This accounts for the ω^2-notation introduced at (30). Note also that $\mathbb{W}$ is itself *not a rotation matrix*: when wrapped around a symmetric matrix it yields a matrix with altered eigenvalues.

7. Normal modes. Given coupled linear equations of the form (29), we construct $\mathbb{W}$ and with its aid write

$$\mathbb{W}^{\mathsf{T}}\mathbb{M}\ddot{\boldsymbol{x}} + 2\mathbb{W}^{\mathsf{T}}(\mathbb{S}+\mathbb{A})\dot{\boldsymbol{x}} + \mathbb{W}^{\mathsf{T}}\mathbb{K}\boldsymbol{x} = \mathbb{W}^{\mathsf{T}}\boldsymbol{F}(t)$$

whence

$$\ddot{\boldsymbol{X}} + 2\mathbb{W}^{\mathsf{T}}(\mathbb{S}+\mathbb{A})\mathbb{W}\cdot\dot{\boldsymbol{X}} + \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \omega_n^2 \end{pmatrix}\boldsymbol{X} = \boldsymbol{f}(t)$$

with $\boldsymbol{X} \equiv \mathbb{W}^{-1}\boldsymbol{x}$ and $\boldsymbol{f} \equiv \mathbb{W}^{\mathsf{T}}\boldsymbol{F}$. The matrix $\mathbb{W}^{\mathsf{T}}\mathbb{S}\mathbb{W}$ is symmetric, but we have no resources left with which to achieve its diagonalization: since it is anyway non-Lagrangian I will at this point abandon the $\mathbb{S}$-term.[19]

[19] This arbitrary decision is a bit high-handed, for if we were doing circuit theory we would be dismissing all resistive effects. An easily workable theory would result if we assumed less drastically that

$$\mathbb{W}^{\mathsf{T}}\mathbb{S}\mathbb{W} = 2\begin{pmatrix} \beta_1 & 0 & \dots & 0 \\ 0 & \beta_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \beta_n \end{pmatrix}$$

Or we could agree to allow $\mathbb{W}^{\mathsf{T}}\mathbb{S}\mathbb{W}$ to be an arbitrary symmetric matrix and, rolling up our sleeves, agree to work with equations that are irreducibly coupled.

The matrix $\mathbb{W}^{\intercal}\mathbb{A}\mathbb{W}$ is antisymmetric, and could be removed by the rotational technique described in §3; I am content therefore to abandon also the $\mathbb{A}$-term. We are left with an n-fold replication of the one-dimensional theory of forced undamped oscillators. Immediately

$$\boldsymbol{X}(t) = \begin{pmatrix} X_1 \sin(\omega_1 t + \delta_1) + \int G_1(t-\tau) f_1(\tau) d\tau \\ X_2 \sin(\omega_2 t + \delta_2) + \int G_2(t-\tau) f_2(\tau) d\tau \\ \vdots \\ X_n \sin(\omega_n t + \delta_n) + \int G_n(t-\tau) f_n(\tau) d\tau \end{pmatrix}$$

where by (23)

$$G_k(t-\tau) = \tfrac{1}{\omega_k} \sin[\omega_k(t-\tau)] \cdot \theta(t-\tau)$$

The motion of the physical variables $\boldsymbol{x}$ can therefore be described

$$\boldsymbol{x}(t) = \mathbb{W}\boldsymbol{X}(t)$$

which—in component-wise detail (and in the absence of forcing)—reads

$$x_i(t) = \sum_{j=1}^{n} W_{ij} X_j \sin(\omega_j t + \delta_j)$$

Suppose that all X_j were to vanish except the k^{th}. We would than have a "monochromatic" solution

$$\boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}_{k^{\text{th}}\ \text{mode}} = X_k \begin{pmatrix} W_{1k} \\ W_{2k} \\ \vdots \\ W_{nk} \end{pmatrix} \sin(\omega_k t + \delta_k)$$

$$\equiv X_k \boldsymbol{x}_k \sin(\omega_k t + \delta_k)$$

of the equations of motion—a solution in which all variables x_i are *vibrating in synchrony*, with angular frequency ω_k.

Notice that the vectors $\boldsymbol{x}_k$ are generally not "normal" if the familiar sense

$$\boldsymbol{x}_j^{\intercal}\boldsymbol{x}_k = 0 \ : \ j \neq k \qquad \text{NOT GENERALLY TRUE}$$

But at (31) we had $\mathbb{W}^{\intercal}\mathbb{M}\mathbb{W} = \mathbb{I}$ which can be notated $\boldsymbol{x}_j^{\intercal}\mathbb{M}\boldsymbol{x}_k = \delta_{jk}$, and on page 6 learned to write

$$\mathbb{M} = \mathbb{N}\cdot\mathbb{N} = \mathbb{N}^{\intercal}\cdot\mathbb{N}$$

giving

$$(\mathbb{N}\mathbb{W})^{\intercal}(\mathbb{N}\mathbb{W}) = \mathbb{I}$$

So what is invariably true is that the vectors $\boldsymbol{e}_j \equiv \mathbb{N}\boldsymbol{x}_j$ are orthonormal:

$$\boldsymbol{e}_j^{\mathsf{T}}\boldsymbol{e}_k = \delta_{jk}$$

I present now a slight variant of the argument that led us above to the "normal mode" concept. Starting from the unforced equations

$$\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \mathbb{O}$$

we look for solutions of the (monochromatic) form

$$\boldsymbol{x}(t) = \boldsymbol{x} \cdot e^{i\omega t}$$

Immediately

$$(\,\mathbb{K} - \omega^2\mathbb{M})\boldsymbol{x} = \boldsymbol{0}$$

which—if we make use once again of $\mathbb{M} = \mathbb{N}^{\mathsf{T}}\cdot\mathbb{N}$ and agree to write $\boldsymbol{e} = \mathbb{N}\boldsymbol{x}$ —becomes $(\mathbb{K}\mathbb{N}^{-1} - \omega^2\,\mathbb{N})\boldsymbol{e} = \boldsymbol{0}$ or

$$(\hat{\mathbb{K}} - \omega^2\,\mathbb{I})\boldsymbol{e} = \boldsymbol{0} \quad \text{with} \quad \hat{\mathbb{K}} \equiv \mathbb{N}^{-1}\mathbb{K}\mathbb{N}^{-1} \text{ (real symmetric)}$$

We have encountered here a straightforward eigenvalue problem: ω^2 must be one of the roots $\{\omega_1^2, \omega_2^2, \ldots, \omega_n^2\}$ of

$$\det(\hat{\mathbb{K}} - \omega^2\,\mathbb{I}) = 0$$

and the solution vectors $\boldsymbol{e}$ must (to within scalar factors) be the associated orthonormal eigenvectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. We arrive thus at monochromatic solutions

$$\boldsymbol{x}_k(t) = \mathbb{N}^{-1}\boldsymbol{e}_k \cdot e^{i\omega_k t}$$

and general (polychromatic) solutions that are weighted linear combinations of the (real parts) of those.

In the first of the methods described above, simultaneous diagonalization (reduction of the many-variable problem to multiple disjoint copies of the single-variable problem) was the controlling idea: it led *via* instances of the eigenvalue problem to a superposition of monochromatic solutions. In the second method, monochromoticity was the controlling idea, and its execution entailed solution of the same set of eigenvalue problems. There is not much to recommend one method over the other, though in most applications most people seem to prefer the latter. A procedure that is computationally swifter than either will be described near the end of the next section.

8. EXAMPLE: Small oscillations of a hanging chain. We study the system shown in FIGURE 10. Immediately

$$\begin{aligned}
U &= mg\ell\Big\{4(1-\cos x_4)+3(1-\cos x_3)+2(1-\cos x_2)+(1-\cos x_1)\Big\}\\
&\approx \tfrac{1}{2}mg\ell\Big\{x_1^2+2x_2^2+3x_3^2+4x_4^2\Big\}\\
T &\approx \tfrac{1}{2}m\ell^2\Big\{\quad \dot{x}_4^2\\
&\qquad\quad +(\dot{x}_4+\dot{x}_3)^2\\
&\qquad\quad +(\dot{x}_4+\dot{x}_3+\dot{x}_2)^2\\
&\qquad\quad +(\dot{x}_4+\dot{x}_3+\dot{x}_2+\dot{x}_1)^2\Big\}
\end{aligned}$$

so $\mathbb{M}$ and $\mathbb{K}$ have the values spelled out below:[20]

$$\mathbb{M} = m\ell^2\begin{pmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{pmatrix}$$

$$\mathbb{K} = mg\ell\begin{pmatrix}1&0&0&0\\0&2&0&0\\0&0&3&0\\0&0&0&4\end{pmatrix}$$

Here as in all pendular problems the m-factors drop away, and to facilitate numerical work we will adopt units in which $g=\ell=1$. To the command `Eigenvalues[`$\mathbb{M}$`]` *Mathematica* supplies

$$\begin{aligned}m_1&=0.28312\\m_2&=0.42602\\m_3&=1.00000\\m_4&=8.29086\end{aligned}$$

while `Eigenvectors[`$\mathbb{M}$`]` which after normalization are displayed as the successive columns in the matrix[21]

$$\mathbb{R} = \begin{pmatrix}-0.42853&+0.65654&-0.57735&+0.22801\\+0.65654&-0.22801&-0.57735&+0.42853\\-0.57735&-0.57735&0.00000&+0.57735\\+0.22801&+0.42853&+0.57735&+0.22801\end{pmatrix}$$

[20] Usually it is $\mathbb{M}$ that is simple (often already diagonal) and $\mathbb{K}$ that is relatively complicated. Here the serial construction of the system has had the effect of reversing that situation.

[21] I display only leading digits in the 16-place numbers that *Mathematica* carries in its memory, and abandon complex terms of the order $i10^{-15}$ that sometimes show up as artifacts of its polynomial root extraction technique.

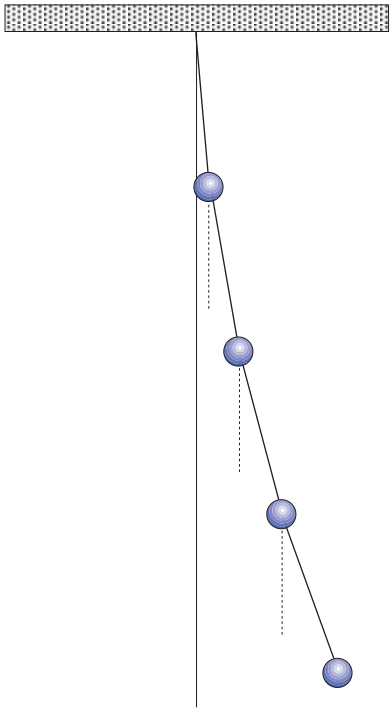


FIGURE 10: *A short "hanging chain" system. The balls will be assumed all to have mass m, the connecting threads all to have length ℓ. The deviation angles—progressing upward—will be given the not-very-angular-looking names x_1, x_2, x_3, x_4.*

Computation confirms that $\mathbb{R}^{\mathsf{T}}\mathbb{R} = \mathbb{I}$ and that $\mathbb{R}^{\mathsf{T}}\mathbb{M}\mathbb{R}$ is diagonal with m_k's strung along the diagonal (in ascending order). Introduce

$$\mathbb{C} = \begin{pmatrix} 1.87939 & 0 & 0 & 0 \\ 0 & 1.53209 & 0 & 0 \\ 0 & 0 & 1.00000 & 0 \\ 0 & 0 & 0 & 0.34730 \end{pmatrix}$$

Verify that $\mathbb{C}^{\mathsf{T}}\mathbb{R}^{\mathsf{T}}\mathbb{M}\,\mathbb{R}\mathbb{C} = \mathbb{I}$ and construct the symmetric matrix

$$\mathbb{K}' \equiv \mathbb{C}^{\mathsf{T}}\mathbb{R}^{\mathsf{T}}\mathbb{K}\,\mathbb{R}\mathbb{C} = \begin{pmatrix} 7.96020 & 2.33258 & 0.02984 & 0.04163 \\ 2.33258 & 5.32733 & 1.33885 & 0.04238 \\ 0.02984 & 1.33885 & 2.33333 & 0.30901 \\ 0.04163 & 0.04238 & 0.30901 & 0.37914 \end{pmatrix}$$

Compute the eigenvalues of $\mathbb{K}'$

$$\begin{aligned}\omega_1^2 &= 0.32255\\ \omega_2^2 &= 1.74576\\ \omega_3^2 &= 4.53662\\ \omega_4^2 &= 9.39507\end{aligned}$$

and from the normalized eigenvectors of $\mathbb{K}'$, entered as columns in sequence, assemble

$$\mathbb{Q} = \begin{pmatrix} -0.01996 & -0.15546 & +0.50650 & -0.84788\\ +0.05020 & +0.42865 & -0.73706 & -0.05008\\ -0.18395 & -0.86989 & -0.44602 & -0.10262\\ +0.98145 & -0.18813 & -0.03559 & -0.00988\end{pmatrix}$$

Observe that $\mathbb{Q}^{\intercal}\mathbb{Q} = \mathbb{I}$ and that $\mathbb{Q}^{\intercal}\mathbb{K}'\mathbb{Q}$ is diagonal with ω_k^2's strung along the diagonal (in ascending order). Construct

$$\mathbb{W} \equiv \mathbb{R}\mathbb{C}\mathbb{Q} = \begin{pmatrix} +0.25050 & +1.04369 & -0.89462 & +0.21818\\ +0.21010 & +0.13267 & +1.13466 & -0.80673\\ +0.17405 & -0.24821 & +0.09525 & +1.37805\\ +0.14198 & -0.33032 & -0.53248 & -0.76628\end{pmatrix}$$

and verify that $\mathbb{W}^{\intercal}\mathbb{M}\mathbb{W} = \mathbb{I}$ and that $\mathbb{W}^{\intercal}\mathbb{K}\mathbb{W}$ is diagonal with ω_k^2's strung along the diagonal (in ascending order). The simultaneous diagonalization is now complete: the natural frequencies of the system have been recognized to be

$$\begin{aligned}\omega_1 &= \sqrt{\omega_1^2} = 0.56793\\ \omega_2 &= \sqrt{\omega_2^2} = 1.32127\\ \omega_3 &= \sqrt{\omega_3^2} = 2.12993\\ \omega_4 &= \sqrt{\omega_4^2} = 3.06514\end{aligned}$$

and the modal vectors $\boldsymbol{x}_k$ are the columns of $\mathbb{W}$, read in sequence.

There is, however, a much swifter way to proceed. Into $\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \mathbb{O}$ introduce the assumption $\boldsymbol{x}(t) = \boldsymbol{x}\cdot e^{i\omega t}$ and obtain $(\mathbb{K} - \omega^2\mathbb{M})\boldsymbol{x} = \boldsymbol{0}$. Multiply from the left by $\mathbb{M}^{-1}$ and obtain[22]

$$(\mathbb{M}^{-1}\mathbb{K} - \omega^2\,\mathbb{I})\boldsymbol{x} = \boldsymbol{0}$$

Recognize this to be *as it stands* a classic eigenvalue/vector problem, and proceed undetered by the observation that $\mathbb{M}^{-1}\mathbb{K}$ has no obviously attractive properties. *Mathematica* leads one *immediately* to eigenvalues

$$\begin{aligned}\omega_1^2 &= 0.32255\\ \omega_2^2 &= 1.74576\\ \omega_3^2 &= 4.53662\\ \omega_4^2 &= 9.39507\end{aligned}$$

[22] Compare this with how the same equation was processed on page 29.

and to (non-orthogonal) eigenvectors

$$\boldsymbol{x}_1=\begin{pmatrix}+1.76429\\+1.47976\\+1.22582\\+1.00000\end{pmatrix},\ \boldsymbol{x}_2=\begin{pmatrix}-3.15966\\-0.40165\\+0.75141\\+1.00000\end{pmatrix},\ \boldsymbol{x}_3=\begin{pmatrix}+1.68009\\-2.13088\\-0.17887\\+1.00000\end{pmatrix},\ \boldsymbol{x}_4=\begin{pmatrix}-0.28472\\+1.05277\\-1.79836\\+1.00000\end{pmatrix}$$

which are readily seen to differ only by obvious scalar multipliers from the corresponding columns in $\mathbb{W}$.

9. Classical molecules. We will—with the assistance of *Mathematica*—use the simple method just described to examine properties of a graded sequence of "classical molecules." We look first to the

LINEAR TRIATOMIC MOLECULE show in FIGURE 11. The kinetic energy is $\frac{1}{2}\dot{\boldsymbol{x}}^{\mathsf{T}}\mathbb{M}\dot{\boldsymbol{x}}$, the potential energy is $\frac{1}{2}\boldsymbol{x}^{\mathsf{T}}\mathbb{K}\boldsymbol{x}=\frac{1}{2}m\omega^2\big[(x_2-x_1)^2+(x_3-x_2)^2\big]$ with

$$\mathbb{M}=\begin{pmatrix}m&0&0\\0&M&0\\0&0&m\end{pmatrix},\qquad \mathbb{K}=m\omega^2\begin{pmatrix}1&-1&0\\-1&2&-1\\0&-1&1\end{pmatrix}$$

so the equations of motion read

$$\mathbb{K}\boldsymbol{x}+\mathbb{M}\ddot{\boldsymbol{x}}=\boldsymbol{0}$$

and if we assume the atoms to move in harmonic synchrony $\boldsymbol{x}(t)=\boldsymbol{x}\,e^{i\omega t}$ becomes

$$(\mathbb{M}^{-1}\mathbb{K}-\omega^2\mathbb{I})\boldsymbol{x}=\boldsymbol{0}$$

The (squared) natural frequencies are

$$\omega_0^2=0,\quad \omega_1^2=\omega^2,\quad \omega_2^2=\big(1+2\tfrac{m}{M}\big)\omega^2$$

and the associated modal vectors are[23]

$$\boldsymbol{x}_0=\begin{pmatrix}1\\1\\1\end{pmatrix}a,\quad \boldsymbol{x}_1=\begin{pmatrix}+1\\0\\-1\end{pmatrix}a,\qquad \boldsymbol{x}_2=\begin{pmatrix}+1\\-2\frac{m}{M}\\+1\end{pmatrix}a$$

of which

- $\boldsymbol{x}_0$ describes non-oscillatory uniform translation of the center of mass,
- $\boldsymbol{x}_1$ describes a "breathing" mode (no center of mass motion),
- $\boldsymbol{x}_2$ describes a "sloshing" mode (no center of mass motion).

Notice that—whatever the (necessarily positive) value of the mass ratio m/M—sloshing is always faster than breathing.

[23] In the following equations a is a constant of arbitrary value but the enforced dimensionality of "length." It is most naturally identified with the characteristic interatomic separation, the "size of the molecule."

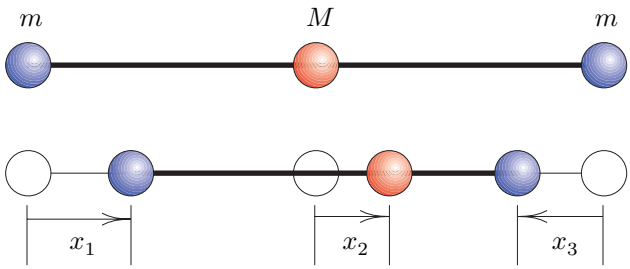
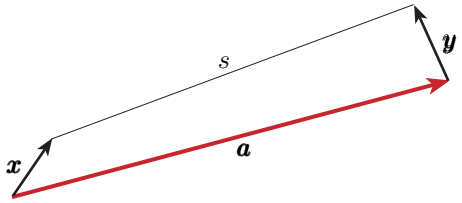


FIGURE 11: *"Linear triatomic molecule," with a central atom of different mass. Excursion variables* $\{x_1, x_2, x_3\}$ *are positive or negative according as they refer to displacements to right or left.*

PLANAR TRIATOMIC MOLECULE This system (FIGURE 12) has six degrees of freedom, and might be expected therefore to have six modes of vibration. The $\mathbb{M}$ matrix is trivial in this instance. To facilitate description of the potential I digress to develop this little

LEMMA: In the following figure $\boldsymbol{a}$ represents a relaxed spring of length a. Infinitesimal displacements $\boldsymbol{x}$ and $\boldsymbol{y}$ of its respective ends



cause the spring to have altered length

$$\begin{aligned} s &= \sqrt{(\boldsymbol{a}+\boldsymbol{y}-\boldsymbol{x})\cdot(\boldsymbol{a}+\boldsymbol{y}-\boldsymbol{x})} \\ &= a\sqrt{1+\frac{2\boldsymbol{a}\cdot(\boldsymbol{y}-\boldsymbol{x})+(\boldsymbol{y}-\boldsymbol{x})\cdot(\boldsymbol{y}-\boldsymbol{x})}{a}} \\ &= a+\frac{\boldsymbol{a}\cdot(\boldsymbol{y}-\boldsymbol{x})+(\boldsymbol{y}-\boldsymbol{x})\cdot(\boldsymbol{y}-\boldsymbol{x})}{a}-\dots \end{aligned}$$

The potential energy stored in the distorted spring is

$$U = \tfrac{1}{2}k(s-a)^2$$

which in leading order has become

$$\approx \tfrac{1}{2}k\big[\hat{\boldsymbol{a}}\cdot(\boldsymbol{y}-\boldsymbol{x})\big]^2 \tag{32}$$

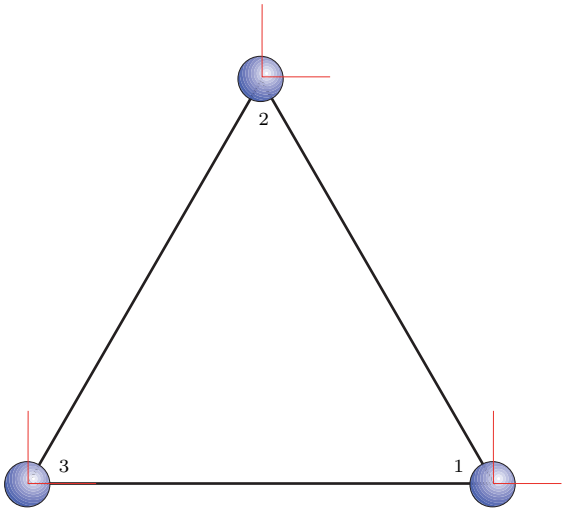


FIGURE 12: *"Planar triatomic molecule," assembled from atoms of identical mass. Attached to each resting atom is a frame used to assign coordinates to its excited excursions. We will write* $\boldsymbol{a}_{ij}$ *to denote the vector that at rest extends from* m_i *to* m_j*: evidently*

$$\hat{\boldsymbol{a}}_{12} = \tfrac{1}{2}\begin{pmatrix} -1 \\ +\sqrt{3} \end{pmatrix}, \quad \hat{\boldsymbol{a}}_{23} = \tfrac{1}{2}\begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix}, \quad \hat{\boldsymbol{a}}_{31} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

We will write

$$\boldsymbol{z}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$$

to describe m_i*'s displacement from its rest position, and write*

$$\boldsymbol{x} \equiv \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

to assign canonical order to the variables we use will to describe the momentary shape of the excited molecule.

Bringing (32) to management of the molecular data presented in the caption to FIGURE 12, we find

$$\mathbb{M}^{-1}\mathbb{K} = \omega^2 \begin{pmatrix} +\frac{5}{4} & -\frac{1}{4} & -1 & -\frac{\sqrt{3}}{4} & +\frac{\sqrt{3}}{4} & 0 \\ -\frac{1}{4} & +\frac{1}{2} & -\frac{1}{4} & +\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{4} \\ -1 & -\frac{1}{4} & +\frac{5}{4} & 0 & -\frac{\sqrt{3}}{4} & +\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & +\frac{\sqrt{3}}{4} & 0 & +\frac{3}{4} & -\frac{3}{4} & 0 \\ +\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{4} & -\frac{\sqrt{3}}{4} & -\frac{3}{4} & +\frac{3}{2} & -\frac{3}{4} \\ 0 & -\frac{\sqrt{3}}{4} & +\frac{\sqrt{3}}{4} & 0 & -\frac{3}{4} & +\frac{3}{4} \end{pmatrix}$$

of which the eigenvalues are

$$\begin{array}{lcl} \omega_0^2 = 0 & : & \text{3-fold degenerate} \\ \omega_1^2 = \frac{3}{2}\omega^2 & : & \text{2-fold degenerate} \\ \omega_2^2 = 3\omega^2 & : & \text{non-degenerate} \end{array}$$

The associated eigenvectors, as displayed by *Mathematica*, are (to within dimensioned factors a of arbitrary numerical value, which I henceforth omit)

$$\boldsymbol{x}_1 = \begin{pmatrix} 0 \\ \sqrt{3} \\ 0 \\ -1 \\ 0 \\ +1 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ -\sqrt{3} \\ 0 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{x}_4 = \begin{pmatrix} +\frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{3}} \\ +\frac{1}{\sqrt{3}} \\ -1 \\ 0 \\ +1 \end{pmatrix}$$

$$\boldsymbol{x}_5 = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ +\frac{2}{\sqrt{3}} \\ -1 \\ +1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_6 = \begin{pmatrix} -\sqrt{3} \\ 0 \\ +\sqrt{3} \\ +1 \\ -2 \\ +1 \end{pmatrix}$$

These, however, are more informative when taken in certain eigenvalue-sharing

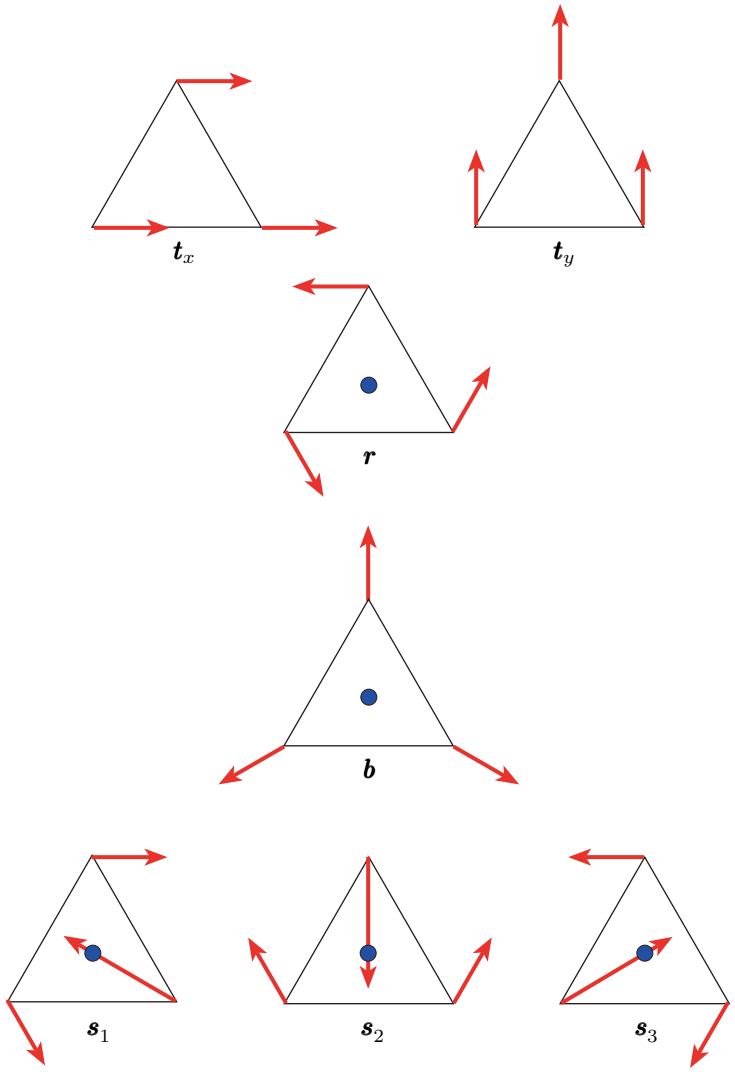


FIGURE 13: *Diagramatic representation of the vibrational modes of "classical ozone." The modes decorated with blue dots • leave invariant the center of mass. The translational modes can be taken in linear combination*

$$\boldsymbol{t} = \cos\theta \cdot \boldsymbol{t}_x + \sin\theta \cdot \boldsymbol{t}_y$$

to achieve translation in any direction. Translation and rotation $\boldsymbol{r}$ *are non-vibrational modes of motion: both have vibrational frequency* $\nu^2 = 0$. *Only two of the scrunch modes are linearly independent*

$$\boldsymbol{s}_1 + \boldsymbol{s}_2 + \boldsymbol{s}_3 = \boldsymbol{0}$$

but all three are necessary to capture the geometric symmetry of the molecule.

linear combinations

$$\begin{aligned}
\boldsymbol{t}_x &\equiv \boldsymbol{x}_3 \\
\boldsymbol{t}_y &\equiv \boldsymbol{x}_1 + \boldsymbol{x}_2 \\
\boldsymbol{r} &\equiv \boldsymbol{x}_3 - \sqrt{3}\boldsymbol{x}_1 \\
\boldsymbol{b} &\equiv \boldsymbol{x}_6 \\
\boldsymbol{s}_1 &\equiv \boldsymbol{x}_5 - 2\boldsymbol{x}_4 \\
\boldsymbol{s}_2 &\equiv \boldsymbol{x}_4 - 2\boldsymbol{x}_5 \\
\boldsymbol{s}_3 &= \boldsymbol{x}_4 + \boldsymbol{x}_5
\end{aligned}$$

that are most clearly described by means of diagrams (FIGURE 13).

In retrospect, we might have anticipated that 0 would appear three times in the spectrum of the molecule (twice for translations, once for rotation), and that it will appear six times in the spectrum of any 3-dimensional molecule (though—why?—only five times in the spectrum of a diatomic molecule). That leaves room for three vibrational modes. We found that one was a breathing mode, two were scrunch modes. We observe that the breathing mode serves by itself to capture the symmetry of the molecule, and that $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3\}$ serve collectively to do so. The physics of the molecule provides a "representation" of the symmetry group (see below: page 38), and if one knows enough group theory one can argue backwards, from the group theory to the physics (or very nearly).[24] It should be noted, however, that if one of the atoms in our triatomic

I	R_1	R_2	D_1	D_2	D_3
R_1	R_2	I	D_3	D_1	D_2
R_2	I	R_1	D_2	D_3	D_1
D_1	D_2	D_3	I	R_1	R_2
D_2	D_3	D_1	R_2	I	R_1
D_3	D_1	D_2	R_1	R_2	I

Group table describing the symmetry structure of the triatomic molecule. Here

R_1 means rotate 120° ↻

R_2 means rotate 240° ↻

D_i means reflect in diagonal through i^{th} vertex

and the table entry identifies the result of performing first the column lable, then the row lable.

[24] See, for example, J. S. Lomont, *Applications of Finite Groups* (1959), Chapter 4, §2, especially pages 117–126; B. Higman, *Applied Group-theoretic and Matrix Methods* (1955), Chapter 10, especially §10.4.

molecule is assigned a different mass $m_2 \mapsto M = m/\mu$ then the spectrum is altered

$$\begin{aligned}
\nu_0^2 &= 0 && : \quad \text{3-fold degenerate}\\
\nu_1^2 &= \omega^2\,\frac{2+\mu}{3} && : \quad \text{non-degenerate}\\
\nu_0^2 &= \omega^2\,\frac{6+3\mp\sqrt{12-12\mu+3\mu^2}}{4} && : \quad \text{non-degenerate}
\end{aligned}$$

but in a smoothly continuous way (ditto the modes), while adjustment of the symmetry group is abrupt and radical: the group described above is replaced by

$$\begin{array}{cc} \mathsf{I} & \mathsf{D}_2 \\ \mathsf{D}_2 & \mathsf{I} \end{array}$$

And most real molecules are notable for their gross *asymmetry* (though they are typically assembled from symmetric components).

The techniques illustrated above as they pertain to "classical ozone" could in principle be used to analyse *any* classical molecule (classical DNA?): all one needs are

- a table of masses m_i
- unit vectors $\hat{\boldsymbol{a}}_{ij}$ that describe the construction of the unexcited molecule
- a table of spring constants k_{ij}.

10. Forced oscillation of many-body systems. We declared an interest in systems of the form

$$\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \boldsymbol{F}(t)$$

already at (29), at the beginning of §6. It was that interest that motivated study of the "simultaneous diagonalization problem," the idea being that by simultaneous diagonalization we might reduce the many-variable problem to many independent copies of the (already solved) single-variable problem. But in subsequent work we looked exclusively to unforced systems

$$\mathbb{M}\ddot{\boldsymbol{x}} + \mathbb{K}\boldsymbol{x} = \boldsymbol{0}$$

and in §9 did so by means of an efficient computational technique (introduced at the end of §8) that proceeds without explicit reference to certain formal niceties that lie near the heart of the simultaneous diagonalization method. I turn now to study of the motion of *stimulated* molecules partly to clarify some formal/ methodological points, partly in recognition of the fact that it is by observing their (quantum mechanical, not classical[25]) response to stimuli that physicists (unlike chemists) have learned most of what they know about the structure of

[25] I remark in passing that if we had in hand a classical molecule assembled from N "atoms" ($3N$ degrees of freedom) we would expect the spectrum to provide typically six 0's (three translational, three rotational modes) and a

molecules. Our basic strategy remains what it has been: to reduce the many-particle problem to multiple copies of the single-particle problem. But the implementation of that strategy requires now more careful attention.

First a word about the recently neglected "formal niceties." It is clear that the equations

$$(\mathbb{K} - \omega^2 \mathbb{M})\boldsymbol{x} = \boldsymbol{0}$$

and

$$(\mathbb{M}^{-1}\mathbb{K} - \omega^2 \mathbb{I})\boldsymbol{x} = \boldsymbol{0}$$

yield identical sets $\{\omega_1^2, \omega_2^2, \ldots, \omega_n^2\}$ and $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ of eigenvalues and eigenvectors, while

$$(\mathbb{K}\mathbb{M}^{-1} - \omega^2 \mathbb{I})\boldsymbol{y} = \boldsymbol{0} \quad \text{with} \quad \boldsymbol{y}_i = \mathbb{M}\boldsymbol{x}_i$$

yields the same set of eigenvalues but (unless $\mathbb{M}$ happens to be diagonal) a *different* population of eigenvectors. Equivalently to any of those equations, we might (as has already been noted) write

$$(\mathbb{N}^{-1}\mathbb{K}\mathbb{N}^{-1} - \omega^2 \mathbb{I})\boldsymbol{e} = \boldsymbol{0} \quad \text{with} \quad \boldsymbol{e}_i = \mathbb{N}\boldsymbol{x}_i \tag{33.1}$$

where the symmetric matrix $\mathbb{N}$ is any one of the 2^n square roots of the symmetric matrix $\mathbb{M}$. From the manifest symmetry of $\hat{\mathbb{K}} \equiv \mathbb{N}^{-1}\mathbb{K}\mathbb{N}^{-1}$ it follows that

- if $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ associate with distinct eigenvalues then automatically $\boldsymbol{e}_i \perp \boldsymbol{e}_j$
- if $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ associate with same eigenvalue then we can ("by hand") *arrange* to have $\boldsymbol{e}_i \perp \boldsymbol{e}_j$, the net implication being that

we can (after normalization) assert quite generally that

$$\boldsymbol{e}_i^{\mathsf{T}}\boldsymbol{e}_j = \delta_{ij} \tag{33.2}$$

(continued from the preceding page) total of $3N - 6$ (possibly degenerate) vibrational frequencies $\{\nu_1, \nu_2, \ldots, \nu_{3N-6}\}$. We might therefore expect the quantized energy spectrum (which by itself provides only an important fragment of a full quantum theory) to have the form

$$E_{n_1, n_2, \ldots, n_{3N-6}} = \hbar(\nu_1 n_1 + \nu_2 n_2 + \cdots + \nu_{3N-6} n_{3N-6}) + E_0$$

(here $E_0 \equiv \frac{3N-6}{2}\hbar$, and the n's range on $\{0, 1, 2, \ldots\}$)—this by "quantizing" each vibrational mode separately. We would, by this reasoning, expect diatomic molecules to have simplest-possible spectra. Conspicuously absent, however, from the primitive theory just sketched is any reference to the spectral complications that result from the quantization of angular momentum. For further discussion see E. B. Wilson, J. C. Decius & P. C. Cross, *Molecular Vibrations: The Theory of Infrared & Raman Vibrational Spectra* (1955) or Chapter 27 in M. Weissbluth, *Atoms & Molecules* (1978).

Which in terms of the $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ vectors reads

$$\begin{aligned} \boldsymbol{x}_i^{\mathsf{T}}\mathbb{M}\boldsymbol{x}_j &= \delta_{ij} \\ \boldsymbol{y}_i^{\mathsf{T}}\mathbb{M}^{-1}\boldsymbol{y}_j &= \delta_{ij} \end{aligned} \tag{34}$$

These vectors are orthonormal *with respect to the mass metric* (or its inverse). Equations (33) identify the language of choice, the framework within which normal modes are literally "$\perp$ modes." It is the language in which we will conduct our business.

Immediately we have the "spectral resolution of $\hat{\mathbb{K}}$:"

$$\hat{\mathbb{K}} = \sum_i \omega_i^2 \mathbb{P}_i \tag{35}$$

where

$$\mathbb{P}_i \equiv \boldsymbol{e}_i \boldsymbol{e}_i^{\mathsf{T}} \quad : \quad i = 1, 2, \ldots n \tag{36}$$

describes a *complete set of orthogonal projection matrices* which necessarily *commute* with one another:

$$\sum_i \mathbb{P}_i = \mathbb{I}, \quad \mathbb{P}_i\mathbb{P}_j = \delta_{ij}\,\mathbb{P}_j = \mathbb{P}_j\mathbb{P}_i \tag{37}$$

The equation of unforced motion has become

$$\Big\{ \sum_i \omega_i^2 \mathbb{P}_i + \partial_t^2 \Big\} \boldsymbol{e}(t) = \boldsymbol{0}$$

of which

$$\boldsymbol{e}(t) = \sum_i \mathcal{X}_i \boldsymbol{e}_i \sin(\omega_i t + \delta_i)$$

is (by $\mathbb{P}_i \boldsymbol{e}_j = \boldsymbol{e}_i \delta_{ij}$) an immediate solution—is, in fact, (since it contains $2N$ independently adjustable constants $\{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n, \delta_1, \delta_2, \ldots \delta_n\}$) the general solution. It follows from $\boldsymbol{e}(t) = \mathbb{N}\boldsymbol{x}(t)$ that $\boldsymbol{x}(t)$ is a solution of $(\mathbb{K}+\mathbb{M}\,\partial_t^2)\boldsymbol{x} = \boldsymbol{0}$.

Multiplying the equation of forced motion on the left by $\mathbb{N}^{-1}$ we obtain

$$\begin{aligned} \mathbb{N}^{-1}(\mathbb{K} + \mathbb{M}\,\partial_t^2)(\mathbb{N}^{-1}\boldsymbol{e}) &= \mathbb{N}^{-1}\boldsymbol{F} \\ &\Downarrow \\ (\hat{\mathbb{K}} + \partial_t^2)\boldsymbol{e}(t) &= \boldsymbol{f}(t) \\ &\Downarrow \\ \Big\{ \sum_i \omega_i^2 \mathbb{P}_i + \partial_t^2 \Big\} \boldsymbol{e}(t) &= \boldsymbol{f}(t) \end{aligned} \tag{38}$$

where the representation of $\boldsymbol{e}(t)$ and $\boldsymbol{f}(t)$ requires an explanatory word. We will write

$$\boldsymbol{e}(t) = \sum_j \mathcal{X}_j(t)\boldsymbol{e}_j \quad \text{and} \quad \boldsymbol{f}(t) = \sum_j \mathcal{F}_j(t)\boldsymbol{e}_j \tag{39}$$

where the $\{\boldsymbol{e}_j\}$ are dimensionless orthonormal *unit vectors* and the $\{\mathcal{X}_j\}$ and $\{\mathcal{F}_j\}$ are dimensioned *coordinates* (that happen in the present instance to be time-dependent)

$$[\,\mathcal{X}_j] = \text{length}\sqrt{\text{mass}} \quad : \quad [\,\mathcal{F}_j] = \frac{\text{length}}{\sqrt{\text{mass}}}$$

Equation (38) can therefore be written

$$\sum_j \left[\ddot{\mathcal{X}}_j + \omega_j^2 \mathcal{X}_j\right] \boldsymbol{e}_j = \sum_j \mathcal{F}_j \boldsymbol{e}_j$$

from which left-multiplication by $\mathbb{P}_k$ serves to project out the k^{th} modal component, giving

$$\ddot{\mathcal{X}}_k(t) + \omega_k^2 \mathcal{X}_k(t) = \mathcal{F}_k(t) \tag{40}$$

Immediately

$$\mathcal{X}_k(t) = \int_{-\infty}^{+\infty} G_k(t-\tau)\, \mathcal{F}_k(\tau)\, d\tau \tag{41}$$

with (see again (23) on page 13)

$$G_i(t-\tau) = e^{-\beta_i(t-\tau)} \frac{\sin \Omega_i (t-\tau)}{\Omega_i}\, \theta(t-\tau) \tag{42}$$
$$\Omega_i = \sqrt{\omega_i^2 - \beta_i^2}$$

into which the β_i's have been introduced to make provision for *modal damping*.[26] We find ourselves in position now to write

$$\boldsymbol{e}(t) = \int_{-\infty}^{+\infty} \mathbb{G}(t-\tau) \boldsymbol{f}(\tau)\, d\tau \tag{43}$$

$$\mathbb{G}(t-\tau) \equiv \sum_i G_i(t-\tau) \mathbb{P}_i \tag{44}$$

which when translated into variables of direct physical significance reads

$$\boldsymbol{x}(t) = \int_{-\infty}^{+\infty} \mathbb{N}^{-1} \mathbb{G}(t-\tau) \mathbb{N}^{-1} \boldsymbol{F}(\tau)\, d\tau \tag{45}$$

---

[26] From this result it follows, by the way, that for zero-frequency modes

$$G_0(t-\tau) = \frac{1 - e^{2\beta_0(t-\tau)}}{2\beta_0}\, \theta(t-\tau)$$

and *Mathematica* assures us that indeed

$$(\partial_t^2 + 2\beta_0 \partial_t)\, G_0(t-\tau) = \delta(t-\tau)$$

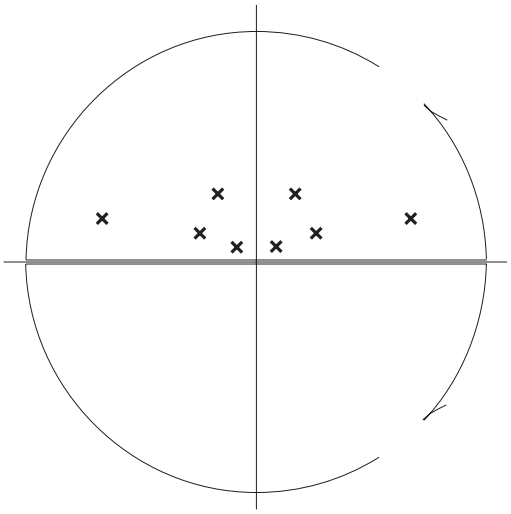


FIGURE 14: *The symmetric matrix $\mathit{\Gamma}(\alpha)$ is a complex-valued function of the complex frequency α. It has poles placed with bilateral symmetry on the upper halfplane, and projection-matrix-valued residues.*

Passing from the time-domain to the frequency-domain by direct analogs

$$\begin{aligned}
\boldsymbol{e}(t) &= \tfrac{1}{\sqrt{2\pi}}\int \boldsymbol{\varepsilon}(\alpha)\, e^{i\alpha t} d\alpha \\
\boldsymbol{f}(t) &= \tfrac{1}{\sqrt{2\pi}}\int \boldsymbol{\varphi}(\alpha)\, e^{i\alpha t} d\alpha \\
\mathbb{G}(t) &= \tfrac{1}{2\pi}\int \mathit{\Gamma}(\alpha)\, e^{i\alpha t} d\alpha
\end{aligned}$$

of equations encountered on page 14, we have

$$\boldsymbol{\varepsilon}(\alpha) = \mathit{\Gamma}(\alpha)\, \boldsymbol{\varphi}(\alpha)$$

with

$$\mathit{\Gamma}(\alpha) = -\sum_k \frac{\mathbb{P}_k}{[\alpha - (+\Omega_k + i\beta_k)][\alpha - (-\Omega_k + i\beta_k)]} \tag{46}$$

There is another way (compare page 16) to understand how it comes about that $\mathit{\Gamma}(\alpha)$ enters as a natural object into the theory. Supposing the system be harmonically stimulated

$$\begin{aligned}
\boldsymbol{F}(t) = \boldsymbol{F}\, e^{i\alpha t} \quad \Rightarrow \quad \boldsymbol{f}(t) &= \boldsymbol{\varphi}\, e^{i\alpha t} \text{ with } \boldsymbol{\varphi} = \mathbb{N}^{-1}\boldsymbol{F} \\
&= e^{i\alpha t} \sum_i \varphi_i(\alpha)\, \boldsymbol{e}_i
\end{aligned} \tag{47.1}$$

(here $\boldsymbol{F}$ and $\boldsymbol{\varphi}$ are constant vectors, as are $\boldsymbol{x}$ and $\boldsymbol{\varepsilon}$ below) we expect to have steady harmonic response:

$$\begin{aligned} \boldsymbol{x}(t) = \boldsymbol{x}\, e^{i\alpha t} \quad \Rightarrow \quad \boldsymbol{e}(t) &= \boldsymbol{\varepsilon}\, e^{i\alpha t} \text{ with } \boldsymbol{\varepsilon} = \mathbb{N}\boldsymbol{x} \\ &= e^{i\alpha t} \sum_i \varepsilon_i(\alpha)\, \boldsymbol{e}_i \end{aligned} \tag{47.2}$$

Bringing those assumptions to[27]

$$\Big\{ \partial_t^2 + \sum_j \big[2\beta_j \partial_t + \omega_j^2 \big] \mathbb{P}_j \Big\}\, \boldsymbol{e}(t) = \boldsymbol{f}(t)$$

and multiplying on the left by $\mathbb{P}_k$, we have

$$\big[-\alpha^2 + 2i\beta_k + \omega_k^2 \big] \varepsilon_k(\alpha) = \varphi_k(\alpha)$$

whence

$$\varepsilon_k(\alpha) = \mathit{\Gamma}_k(\alpha) \varphi_k(\alpha) \tag{48}$$

from which (46) immediately follows, and (41/42) can be recovered by contour integration.

The energy of our vibrating system (crystal) is given by

$$\begin{aligned} E &= \tfrac{1}{2}\, \dot{\boldsymbol{x}}^{\mathsf{T}} \mathbb{M}\, \dot{\boldsymbol{x}} + \tfrac{1}{2}\, \boldsymbol{x}^{\mathsf{T}} \mathbb{K} \boldsymbol{x} \\ &= \tfrac{1}{2}\, \dot{\boldsymbol{e}}^{\mathsf{T}} \dot{\boldsymbol{e}} + \tfrac{1}{2}\, \boldsymbol{e}^{\mathsf{T}} \hat{\mathbb{K}} \boldsymbol{e} \\ &= \sum_k E_k \\ E_k &= \tfrac{1}{2} \big[\, \dot{\mathcal{X}}_k^2 + \omega_k^2 \mathcal{X}_k^2 \,\big] \end{aligned}$$

Proceeding now exactly as on pages 21–23, we express $\mathit{\Gamma}_k(\alpha)$ in polar form

$$\mathit{\Gamma}_k(\alpha) = A_k(\alpha) e^{-i\delta_k(\alpha)}$$

we conclude on the basis of (48) that

$$\begin{aligned} &\text{harmonic stimulus } \boldsymbol{f}(t) = e^{i\alpha t} \sum_i \varphi_i(\alpha)\, \boldsymbol{e}_i \\ &\qquad\qquad \Downarrow \\ &\text{harmonic response } \boldsymbol{e}(t) = \Re\big[e^{i\alpha t} \sum_i \mathit{\Gamma}_i(\alpha) \varphi_i(\alpha) \big] \boldsymbol{e}_i \\ &\qquad\qquad\qquad\qquad = \sum_i A_i(\alpha) \varphi_i(\alpha) \cos[\alpha t - \delta_i(\alpha)]\, \boldsymbol{e}_i \end{aligned}$$

and—proceeding in direct imitation of the argument on pages 22 & 23—find

[27] As the following equation indicates, we have elected to manage "modal dissipation" in the manner suggested in footnote 19 on page 27.

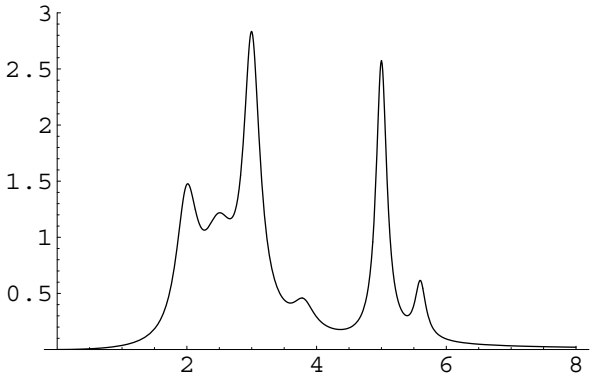


FIGURE 15: *Typical absorption spectrum for an imaginary classical molecule, computed from* (49) *with* $\{\varphi, \omega, \beta\}$*-parameters set to values* $\{1.0, 2.0, 0.2\}$, $\{1.0, 2.5, 0.3\}$, $\{1.5, 3.0, 0.15\}$, $\{0.2, 3.8, 0.2\}$, $\{1.0, 5.0, 0.1\}$ *and* $\{0.2, 5.6, 0.1\}$. *It would be one thing for a molecular spectroscopist to extract that data from the curve, quite another to deduce molecular structure from the data. For look to the economics of the situation*: *more numbers must be specified to describe a structure than to describe a spectrum.*

that under harmonic stimulation the

$$\begin{aligned}
&\text{mean rate } \langle \mathcal{D} \rangle \text{ of energy dissipation} \\
&\qquad = \text{mean rate } \langle \mathcal{J} \rangle \text{ of energy injection} \\
&\qquad = \tfrac{1}{2}\alpha \sum_k \varphi_k^2(\alpha) \frac{2\alpha\beta_k}{(\omega_k^2 - \beta_k^2)^2 + 4\alpha^2\beta_k^2} \\
&\qquad = -\tfrac{1}{2}\alpha \sum_k \varphi_k^2(\alpha) \Im[\Gamma_k(\alpha)] \\
&\qquad = -\tfrac{1}{2}\alpha \boldsymbol{\varphi}^{\intercal}(\alpha) \Im[\mathbb{\Gamma}(\alpha)] \boldsymbol{\varphi}(\alpha) \\
&\qquad = -\tfrac{1}{2}\alpha \boldsymbol{F}^{\intercal}(\alpha) \Im[\mathbb{N}^{-1}\mathbb{\Gamma}(\alpha)\mathbb{N}^{-1}] \boldsymbol{F}(\alpha)
\end{aligned} \tag{49}$$

It is gratifying to observe that according to this formula

$$[\langle \mathcal{D} \rangle] = \frac{[\text{force}/\sqrt{\text{mass}}\,]^2}{\text{frequency}} = \frac{\text{energy}}{\text{time}}$$

Equation (49) pertains in principle to *all* classical vibratory structures. For macroscopic structures the β_k's can be expected to arise typically from a "dashpot effect" (viscosity), but for "classical molecules" it would be more realistic to assume that the stimulation is optical, and that the β_k's refer phenomenologically to a scattering process ("classical Raman scattering").