CHAPTER 2

The Equations of Motion for Extensible Strings

1. Introduction

The main purpose of this chapter is to give a derivation, which is mathematically precise, physically natural, and conceptually simple, of the quasilinear system of partial differential equations governing the large motion of nonlinearly elastic and viscoelastic strings. This derivation, just like all our subsequent derivations of equations governing the behavior of rods, shells, and 3-dimensional bodies, is broken down into the description of (i) the kinematics of deformation, (ii) fundamental mechanical laws (such as the generalization of Newton's Second Law to continua), and (iii) material properties by means of constitutive equations. This scheme separates the treatment of geometry and mechanics in steps (i) and (ii), which are regarded as universally valid, from the treatment of constitutive equations, which vary with the material. Since this derivation serves as a model for all subsequent derivations, we examine each aspect of it with great care. We pay special attention to the Principle of Virtual Power and the equivalent Impulse-Momentum Law, which are physically and mathematically important generalizations of the governing equations of motion and which play essential roles in the treatments of initial and boundary conditions, jump conditions, variational formulations, and approximation methods. In this chapter we begin the study of simple concrete problems, deferring to Chaps. 3 and 6 the treatment of more challenging problems.

The exact equations for the large planar motion of a string were derived by Euler (1751) in 1744 and those for the large spatial motion by Lagrange (1762). By some unfortunate analog of Gresham's law, the simple and elegant derivation of Euler (1771), which is based on Euler's (1752) straightforward combination of geometry with mechanical principles, has been driven out of circulation and supplanted with baser derivations, relying on ad hoc geometrical and mechanical assumptions. (Evidence for this statement can be found in numerous introductory texts on partial differential equations and on mathematical physics. Rare exceptions to this unhappy tradition are the texts of Bouligand (1954) and Weinberger (1965).) A goal of this chapter is to show that it is easy to derive the equations correctly, much easier than following many modern expositions, which ask the reader to emulate the Red Queen by believing six impossible things before breakfast.

The correct derivation is simple because Euler made it so. Modern authors should be faulted not merely for doing poorly what Euler did well, but also for failing to copy from the master. A typical ad hoc assumption found in the textbook literature is that the motion of each material point is confined to the plane through its equilibrium position perpendicular to the line joining the ends of the string. In Sec. 7 we show that scarcely any elastic strings can execute such a motion. Most derivations suppress the role of material properties and even the extensibility of the string by assuming that the tension is approximately constant for all small motions. Were it exactly constant, then no segment of a uniform string could change its length, and if the ends of such a string were held at a separation equal to the length of the string, then the string could not move. (One author of a research monograph on 1-dimensional wave propagation derived the wave equation governing the motion of an inextensible string. Realizing that an inextensible string with its ends separated by its natural length could not move, however pretty its governing equations, he assumed that one end of the string was joined to a fixed point by a spring.) One can make sense out of such assumptions as those of purely transverse motion and of the constancy of tension by deriving them as consequences of a systematic perturbation scheme applied to the exact equations, as we do in Sec. 8.

Parts of Secs. 1–4, 6, 8 of this chapter are adapted from Antman (1980b) with the kind permission of the Mathematical Association of America.

2. The Classical Equations of Motion

In this section we derive the classical form of the equations for the large motion of strings of various materials. A *classical* solution of these equations has the defining property that all its derivatives appearing in the equations are continuous on the interiors of their domains of definition. To effect our derivation, we accordingly impose corresponding regularity restrictions on the geometrical and mechanical variables. Since it is well known on both physical and mathematical grounds that solutions of these equations need not be classical, we undertake in Secs. 3 and 4 a more profound study of their derivation, which dispenses with simplified regularity assumptions.

Kinematics of deformation. Let $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ be a fixed right-handed orthonormal basis for the Euclidean 3-space $\mathbb{E}^3$. A *configuration of a string* is defined to be a curve in $\mathbb{E}^3$. A *string* itself is defined to be a set of elements called *material points* (or *particles*) having the geometrical property that it can occupy curves in $\mathbb{E}^3$ and having the mechanical property that it is 'perfectly flexible'. The definition of perfect flexibility is given below.

We refrain from requiring that the configurations of a string be simple (nonintersecting) curves for several practical reasons: (i) Adjoining the global requirement that configurations be simple curves to the local requirement that configurations satisfy a system of differential equations can lead to severe analytical difficulties. (ii) If two different parts of a string come into contact, then the nature of the resulting mechanical interaction must be carefully specified. (iii) A configuration with self-intersections may serve as a particularly convenient model for a configuration in which distinct parts of a string are close, but fail to touch. (iv) It is possible to show that configurations corresponding to solutions of certain problems must be simple (see, e.g., Chap. 3).

We distinguish a configuration $s \mapsto s\boldsymbol{k}$, in which the string lies along an interval in the $\boldsymbol{k}$-direction, as the *reference configuration*. We identify each material point in the string by its coordinate s in this reference configuration. If the domain of definition of the reference configuration is a bounded interval, then, without loss of generality, we scale the length variable s to lie in the unit interval $[0, 1]$. If this domain is semi-infinite or doubly infinite, then we respectively scale s to lie in $[0, \infty)$ or $(-\infty, \infty)$. In our ensuing development of the theory, we just treat the case in which this domain is $[0, 1]$; adjustments for the other two cases are straightforward. (If the string is a closed loop, we could take a circle as its reference configuration, but there is no need to do this because the reference configuration need not be one that can be continuously deformed from topologically admissible configurations; the main purpose of the reference configuration is to name material points.)

For a string undergoing some motion, let $\boldsymbol{r}(s, t)$ denote the position of the material point (with coordinate) s at time t. For the purpose of studying initial-boundary-value problems, we take the domain of $\boldsymbol{r}$ to be $[0, 1] \times [0, \infty)$. The function $\boldsymbol{r}(\cdot, t)$ defines the *configuration* of the string at time t. In this section we adopt the convention that every function of s and t, such as $\boldsymbol{r}$, whose values are exhibited here is ipso facto assumed to be continuous on the interior of its domain. (We critically examine this assumption in the next two sections.) The vector $\boldsymbol{r}_s(s, t)$ is tangent to the curve $\boldsymbol{r}(\cdot, t)$ at $\boldsymbol{r}(s, t)$. (By our convention, $\boldsymbol{r}_s$ is assumed to be continuous on $(0, 1) \times (0, \infty)$.) Note that we do not parametrize the curve $\boldsymbol{r}(\cdot, t)$ with its arc length. The parameter s, which identifies material points, is far more convenient on mathematical and physical grounds.

The length of the material segment (s_1, s_2) in the configuration at time t is the integral $\int_{s_1}^{s_2} |\boldsymbol{r}_s(s, t)|\, ds$. The *stretch* $\nu(s, t)$ of the string at (s, t) is

$$\nu(s, t) := |\boldsymbol{r}_s(s, t)|. \tag{2.1}$$

(It is the local ratio at s of the deformed to reference length, i.e., it is the limit of $\int_{s_1}^{s_2} |\boldsymbol{r}_s(s, t)|\, ds/(s_2 - s_1)$ as the material segment (s_1, s_2) shrinks down to the material point s.) An attribute of a ‘regular’ motion is that this length ratio never be reduced to zero:

$$\nu(s, t) > 0 \quad \forall\ (s, t) \in [0, 1] \times [0, \infty). \tag{2.2}$$

Provided that the reference configuration is natural, which means that there is zero contact force acting across every material point in this configuration (see the discussion of mechanics below), the string is said to be *elongated* where $\nu(s, t) > 1$, and to be *compressed* where $\nu(s, t) < 1$. (The difficulty one encounters in compressing a real string is a consequence of an instability due to its great flexibility.)

To be specific, we assume that the ends $s = 0$ and $s = 1$ of the string are fixed at the points $\boldsymbol{o}$ and $L\boldsymbol{k}$ where L is a given positive number. In the optimistic spirit that led us to assume that $\boldsymbol{r}$ is continuous on $(0, 1)\times(0, \infty)$, we further suppose that $\boldsymbol{r}(\cdot, t)$ is continuous on $[0, 1]$ for all $t > 0$. In this case, our prescription of $\boldsymbol{r}$ at $s = 0$ and at $s = 1$ leads to boundary conditions expressed by the following pointwise limits:

$$\lim_{s\searrow 0} \boldsymbol{r}(s, t) = \boldsymbol{o}, \quad \lim_{s\nearrow 1} \boldsymbol{r}(s, t) = L\boldsymbol{k} \quad \text{for} \quad t > 0, \tag{2.3a}$$

which imply that $\boldsymbol{r}(\cdot, t)$ is continuous up to the ends of its interval of definition. These conditions are conventionally denoted by

$$\boldsymbol{r}(0, t) = \boldsymbol{o}, \quad \boldsymbol{r}(1, t) = L\boldsymbol{k}. \tag{2.3b}$$

We assume that the string is released from configuration $s \mapsto \boldsymbol{u}(s)$ with velocity field $s \mapsto \boldsymbol{v}(s)$ at time $t = 0$. If $\boldsymbol{r}_t(s, \cdot)$ is assumed to be continuous on $[0, \infty)$ for each $s \in (0, 1)$, then these initial conditions have the pointwise interpretations

$$\lim_{t \searrow 0} \boldsymbol{r}(s,t) = \boldsymbol{u}(s), \quad \lim_{t \searrow 0} \boldsymbol{r}_t(s,t) = \boldsymbol{v}(s) \quad \text{for} \quad s \in (0,1), \tag{2.4a}$$

which are conventionally written as

$$\boldsymbol{r}(s,0) = \boldsymbol{u}(s), \quad \boldsymbol{r}_t(s,0) = \boldsymbol{v}(s). \tag{2.4b}$$

The requirement that the data given on the boundary of $[0,1] \times [0, \infty)$ by (2.3) and (2.4) be continuous, so that $\boldsymbol{r}_t$ could be continuous on its domain, is expressed by the *compatibility conditions*

$$\boldsymbol{u}(0) = \boldsymbol{o}, \quad \boldsymbol{u}(1) = L\boldsymbol{k}, \quad \boldsymbol{v}(0) = \boldsymbol{o}, \quad \boldsymbol{v}(1) = \boldsymbol{o}. \tag{2.5}$$

Mechanics. Let $0 < a < b < 1$. We assume that the forces acting on (the material of) (a, b) in configuration $\boldsymbol{r}(\cdot, t)$ consist of a *contact force* $\boldsymbol{n}^+(b, t)$ exerted on (a, b) by $[b, 1]$, a *contact force* $-\boldsymbol{n}^-(a, t)$ exerted on (a, b) by $[0, a]$, and a *body force* exerted on (a, b) by all other agents. We assume that the body force has the form $\int_a^b \boldsymbol{f}(s,t)\, ds$. The contact force $\boldsymbol{n}^+(b, t)$ has the defining property that it is the same as the force exerted on (c, b) by $[b, d]$ for each c and d satisfying $0 < c < b < d < 1$. Analogous remarks apply to $-\boldsymbol{n}^-$. Thus $\boldsymbol{n}^\pm(\cdot, t)$ are defined on an interval $(0, 1)$ of real numbers, as indicated (and not on a collection of pairs of disjoint intervals). We shall see that the distinction between open and closed sets in the definitions of contact forces will evaporate (for the problems we treat; this distinction can play a critical role when the string is in contact with another body). The minus sign before $\boldsymbol{n}^-(a, t)$ is introduced for mathematical convenience. (It corresponds to the sign convention of structural mechanics.)

Let $(\rho A)(s)$ denote the mass density per unit length at s in the reference configuration. This rather clumsy notation, using two symbols for one function, is employed because it is traditional and because it suggests that the density per unit reference length at s in a real 3-dimensional string is the integral of the density per unit reference volume, traditionally denoted by ρ, over the cross section at s with area $A(s)$. It is important to note, however, that the notion of a cross-sectional area never arises in our idealized model of a string. We assume that ρA is everywhere positive on $(0, 1)$ and that it is bounded on $[0, 1]$.

The integrand $\boldsymbol{f}(s, t)$ of the body force is the *body force per unit reference length* at s, t. The most common example of the body force on a segment is the weight of the segment, in which case $\boldsymbol{f}(s,t) = -(\rho A)(s) g \boldsymbol{e}$ where g is the acceleration of gravity and $\boldsymbol{e}$ is the unit vector pointing in the vertical direction. $\boldsymbol{f}(s, t)$ could depend on $\boldsymbol{r}$ in quite complicated ways. For example, $\boldsymbol{f}$ could have the composite form

$$\boldsymbol{f}(s,t) = \boldsymbol{g}\big(\boldsymbol{r}(s,t), \boldsymbol{r}_t(s,t), s, t\big) \tag{2.6}$$

where $\boldsymbol{g}$ is a prescribed function, which describes the effects of the environment. The dependence of $\boldsymbol{g}$ on the velocity $\boldsymbol{r}_t$ could account for air resistance and its dependence upon the position $\boldsymbol{r}$ could account for variable gravitational attraction.

The requirement that at typical time t the resultant force on the typical material segment $(a, s) \subset (0, 1)$ equal the time derivative of the *linear momentum* $\int_a^s (\rho A)(\xi)\boldsymbol{r}_t(\xi, t)\, d\xi$ of that segment yields the following integral form of the *equation of motion*

$$\begin{aligned} &\boldsymbol{n}^+(s,t) - \boldsymbol{n}^-(a,t) + \int_a^s \boldsymbol{f}(\xi,t)\, d\xi \\ &\qquad = \frac{d}{dt}\int_a^s (\rho A)(\xi)\boldsymbol{r}_t(\xi,t)\, d\xi = \int_a^s (\rho A)(\xi)\boldsymbol{r}_{tt}(\xi,t)\, d\xi. \end{aligned} \tag{2.7}$$

This equation is to hold for all $(a, s) \subset (0, 1)$ and all $t > 0$.

The continuity of $\boldsymbol{n}^+$ implies that $\boldsymbol{n}^+(a, t) = \lim_{s\to a} \boldsymbol{n}^+(s, t)$. Since $\boldsymbol{f}$ and $\boldsymbol{r}_{tt}$ are continuous, we let $s \to a$ in (2.7) to obtain

$$\boldsymbol{n}^+(a,t) = \boldsymbol{n}^-(a,t) \quad \forall\ a \in (0,1). \tag{2.8}$$

Since the superscripts $\pm$ on $\boldsymbol{n}$ are thus superfluous, we drop them. We differentiate (2.7) with respect to s to obtain the *classical form of the equations of motion*:

$$\boldsymbol{n}_s(s,t) + \boldsymbol{f}(s,t) = (\rho A)(s)\, \boldsymbol{r}_{tt}(s,t) \quad \text{for} \quad s \in (0,1),\ t > 0. \tag{2.9}$$

Students of mechanics know that the motion of bodies is governed not only by a linear momentum principle like (2.7), but also by an angular momentum principle. We shall shortly explain how the assumption of perfect flexibility together with two additional assumptions ensure that the angular momentum principle is identically satisfied. Under these conditions, (2.9) represents the culmination of the basic mechanical principles for strings.

Constitutive equations. We describe those material properties of a string that are relevant to mechanics by specifying how the contact force $\boldsymbol{n}$ is related to the change of shape suffered by the string in every motion $\boldsymbol{r}$. Such a specification, called a *constitutive relation*, must distinguish the material response of a rubber band, a steel band, a cotton thread, and a filament of chewing gum. The system consisting of (2.9) and the constitutive equation is formally determinate: It has as many equations as unknowns.

A defining property of a string is its perfect flexibility, which is expressed mathematically by the requirement that $\boldsymbol{n}(s,t)$ be tangent to the curve $\boldsymbol{r}(\cdot, t)$ at $\boldsymbol{r}(s,t)$ for each s, t:

$$\boldsymbol{r}_s(s,t) \times \boldsymbol{n}(s,t) = \boldsymbol{o} \quad \forall\ s, t \tag{2.10a}$$

or, equivalently, that there exist a scalar-valued function N such that

$$\boldsymbol{n}(s,t) = N(s,t)\frac{\boldsymbol{r}_s(s,t)}{|\boldsymbol{r}_s(s,t)|}. \tag{2.10b}$$

(Note that (2.2) ensures that $\boldsymbol{r}_s(s,t) \neq \boldsymbol{o}$ for each s, t.) Why (2.10) should express perfect flexibility is not obvious from the information at hand. One motivation for this condition could come from experiment. The best motivation for this tangency condition comes from outside our self-consistent theory of strings, namely, from the theory of rods, which is developed in Chaps. 4 and 8. The motion of a rod is governed by (2.9) and a companion equation expressing the equality of the resultant torque on any segment of the rod with the time derivative of the angular momentum for that segment. In the degenerate case that the rod offers no resistance to bending, has no angular momentum, and is not subjected to a body couple, this second equation reduces to (2.10a) (and the rod theory reduces to the string theory).

The force (component) $N(s,t)$ is the *tension* at (s,t). It may be of either sign. Where N is positive it is said to be *tensile* and the string is said to be *under tension*; where N is negative it is said to be *compressive* and the string is said to be *under compression*. (This terminology is typical of the inhospitability of the English language to algebraic concepts.)

From primitive experiments, we might conclude that the tension $N(s,t)$ at (s,t) in a rubber band depends only on the stretch $\nu(s,t)$ at (s,t) and on the material point s. Such experiments might not suggest that this tension depends on the rate at which the deformation is occurring, on the past history of the deformation, or on the temperature. Thus we might be led to assume that the string is *elastic*, i.e., that there is a constitutive function $(0,\infty) \times [0,1] \ni (\nu, s) \mapsto \hat{N}(\nu, s) \in \mathbb{R}$ such that

$$N(s,t) = \hat{N}\big(\nu(s,t), s\big). \tag{2.11}$$

Note that (2.11) does not allow $N(s,t)$ to depend upon $\boldsymbol{r}(s,t)$ through $\hat{N}$. Were there such a dependence, then we could change the material properties of the string simply by translating it from one position to another. (In this case, it would be impossible to use springs to measure the acceleration of gravity at different places, as Hooke did, by measuring the elongation produced in a given spring by the suspension of a given mass.) Similarly, (2.11) does not allow $N(s,t)$ to depend upon all of $\boldsymbol{r}_s(s,t)$, but only on its magnitude, the stretch $\nu(s,t)$. A dependence on $\boldsymbol{r}_s(s,t)$ would mean that we could change the material response of the string by merely changing its orientation. Finally, (2.11) does not allow $N(s,t)$ to depend explicitly on absolute time t (i.e., $\hat{N}$ has no slot for the argument t alone). At first sight, this omission seems like an unwarranted restriction of generality, because a real rubber band becomes more brittle with the passage of time. But a careful consideration of this question suggests that the degradation of a rubber band depends on the time elapsed since its manufacture, rather than on the absolute time. Were the constitutive function to depend explicitly on t, then the outcome of an experiment performed today on a material manufactured yesterday would differ from the outcome of the same experiment performed tomorrow on the same material manufactured today. This dependence on time lapse can be generalized by allowing $N(s,t)$ to depend on the past history of the deformation at (s, t). We shall soon show how to account for this dependence. In using (2.11) one chooses to ignore such effects. That the material response should be unaffected by rigid motions and by time translations is called the *Principle of Frame-Indifference* (or the *Principle of Objectivity*).

Let us sketch how the use of this principle leads to a systematic method for reducing a constitutive equation in a general form such as

$$N(s,t) = N_0\big(\boldsymbol{r}(s,t), \boldsymbol{r}_s(s,t), s, t\big) \tag{2.12a}$$

to a very restricted form such as (2.11). (In Chaps. 8 and 12, we give major generalizations of this procedure.) A motion differing from $\boldsymbol{r}$ by a rigid motion has values of the form $\boldsymbol{c}(t) + \boldsymbol{Q}(t) \cdot \boldsymbol{r}(s,t)$ where $\boldsymbol{c}$ is an arbitrary vector-valued function and where $\boldsymbol{Q}$ is an arbitrary proper-orthogonal tensor-valued function. (A full discussion of these tensors is given in Chap. 11.) Then N_0 is invariant under rigid motions and time translations if and only if

$$N_0\big(\boldsymbol{r}, \boldsymbol{r}_s, s, t\big) = N_0\big(\boldsymbol{c}(t) + \boldsymbol{Q}(t)\cdot\boldsymbol{r}, \boldsymbol{Q}(t)\cdot\boldsymbol{r}_s, s, t+a\big) \tag{2.12b}$$

for all vector-valued functions $\boldsymbol{c}$, for all proper-orthogonal tensor-valued functions $\boldsymbol{Q}$, and for all real numbers a. First we take $\boldsymbol{c} = \boldsymbol{o}$, $\boldsymbol{Q} = \boldsymbol{I}$. Then (2.12b) implies that N_0 is independent of its last argument t. Next we take $\boldsymbol{Q} = \boldsymbol{I}$ and let $\boldsymbol{c}$ be arbitrary. Then (2.12b) implies that N_0 is independent of its first argument $\boldsymbol{r}$. Finally we let $\boldsymbol{Q}$ be arbitrary. We write $\boldsymbol{r}_s = \nu\boldsymbol{e}$ where $\boldsymbol{e}$ is a unit vector. Then (2.12b) reduces to

$$N_0\big(\nu\boldsymbol{e}, s\big) = N_0\big(\nu\boldsymbol{Q}(t)\cdot\boldsymbol{e}, s\big). \tag{2.12c}$$

We regard the N_0 of (2.12c) as a function of the three arguments $\nu \in (0, \infty)$, the unit vector $\boldsymbol{e}$, and s. But (2.12c) says that (2.12c) is unaffected by the replacement of $\boldsymbol{e}$ with any unit vector, so that N_0 must be independent of $\boldsymbol{e}$, i.e., (2.12a) must have the form (2.11).

There is no physical principle preventing the constitutive function from depending in a frame-indifferent way on higher s-derivatives of $\boldsymbol{r}$. Such a dependence arises in certain more refined models for strings that account for thickness changes. For example, to obtain a refined model for a rubber band, one might wish to exploit the fact that rubber is nearly incompressible, so that the volume of any piece of rubber is essentially constant. Within a theory of strings, this constraint can be modelled by taking the thickness to be determined by the stretch, with the consequence that higher derivatives enter the constitutive equations and the inertia terms. (See the discussion in Sec. **16**.12.) Similar effects arise in string models for compressible materials (cf. Sec. **8**.9. These can be interpreted as describing an internal surface tension, which seems to be of limited physical importance except for problems of shock structure and phase changes where its role can be critical. See Carr, Gurtin, & Slemrod (1984), Hagan & Slemrod (1983)), and the references cited in item (iv) Sec. **14**.16.

Anyone who rapidly deforms a rubber band feels an appreciable increase in temperature θ. One can also observe that the mechanical response of the band is influenced by its temperature. To account for these effects we may replace (2.11) with the mechanical constitutive equation for a *thermoelastic* string:

$$N(s,t) = \hat{N}_{00}\big(\nu(s,t), \theta(s,t), s\big). \tag{2.13}$$

When this equation is used, the equation of motion must be supplemented with the energy equation, and the new variables entering the energy equation must be related by constitutive equations.

The motion of a rubber band fixed at its ends and subject to zero body force is seen to die down in a short time, even if the motion occurs in a vacuum. The chief source of this decay is internal friction, which is intimately associated with thermal effects. The simplest model for this friction, which ignores thermal effects, is obtained by assuming that the tension $N(s,t)$ depends on the stretch $\nu(s,t)$, the rate of stretch $\nu_t(s,t)$, and the material point s; that is, there is a function $(0,\infty) \times \mathbb{R} \times [0,1] \ni (\nu, \dot{\nu}, s) \mapsto \hat{N}_1(\nu, \dot{\nu}, s) \in \mathbb{R}$ such that

$$N(s,t) = \hat{N}_1\big(\nu(s,t), \nu_t(s,t), s\big). \tag{2.14}$$

(Note that in general $\nu_t \equiv |\boldsymbol{r}_s|_t$ is not equal to $|\boldsymbol{r}_{st}|$. In the argument $\dot{\nu}$ of $\hat{N}_1$, the superposed dot has no operational significance: $\dot{\nu}$ is just a symbol for a real variable, in whose slot, however, the time derivative ν_t appears in (2.14).) When (2.14) holds, the string may be called *viscoelastic of strain-rate type with complexity 1*. (Some authors refer to such materials as being of *rate type*, while others refer to them as being of *differential type*, reserving *rate type* for an entirely different class.) It is clear that (2.14) ensures that the material response is unaffected by rigid motions and translations of time:

2.15. Exercise. Prove that a frame-indifferent version of the constitutive equation $N(s,t) = \hat{N}_1(\boldsymbol{r}_s(s,t), \boldsymbol{r}_{st}(s,t), s)$ must have the form (2.14).

The form of (2.14) suggests the generalization in which $N(s,t)$ depends upon the first k t-derivatives of $\nu(s,t)$ and on s. (Such a string is termed *viscoelastic of strain-rate type with complexity k*.) This generalization is but a special case of that in which $N(s,t)$ depends upon the past history of $\nu(s,\cdot)$ and upon s. To express the constitutive equation for such a material, we define the *history $\nu^t(s,\cdot)$ of $\nu(s,\cdot)$ up to time t* on $[0,\infty)$ by

$$\nu^t(s,\tau) := \nu(s, t-\tau) \quad \text{for} \quad \tau \geq 0. \tag{2.16a}$$

Then the most general constitutive equation of the class we are considering has the form

$$N(s,t) = \hat{N}_\infty\big(\nu^t(s,\cdot), s\big). \tag{2.16b}$$

The domain of $\hat{N}_\infty(\cdot, s)$ is a class of positive-valued functions. A material described by (2.16) (that does not degenerate to (2.11)) and that is dissipative may be called *viscoelastic*. This term is rather imprecise; in modern continuum mechanics it is occasionally used as the negation of *elastic* and is thus synonymous with *inelastic*.

Note that (2.14) reduces to (2.11) where the string is in equilibrium. Similarly, if the string with constitutive equation (2.16b) has been in equilibrium for all time before t (or, more generally, for all such times $t-\tau$ for which $\nu(s, t-\tau)$ influences $\hat{N}_\infty$), then (2.16b) also reduces to (2.11). Thus "the equilibrium response of all strings (in a purely mechanical theory) is elastic." We shall pay scant attention to constitutive equations of the form (2.16b) more general than (2.14). There is a fairly new and challenging mathematical theory for such materials with nonlinear constitutive equations; see Renardy, Hrusa, & Nohel (1987).

A string is said to be *uniform* if ρA is constant and if its constitutive function $\hat{N}$, $\hat{N}_1, \dots$ does not depend explicitly on s. A real (3-dimensional) string fails to be uniform when its material properties vary along its length or, more commonly, when its cross section varies along its length. If only the latter occurs, we can denote the cross-sectional area at s by $A(s)$. Then $(\rho A)(s)$ reduces to $\rho\, A(s)$ where ρ is the given constant mass density per reference volume. In this case, the constitutive function $\hat{N}$ might well have the form $\hat{N}(\nu, s) = A(s)\bar{N}(\nu)$, etc.

Not every choice of the constitutive functions $\hat{N}$, etc., is physically reasonable: We do not expect a string to shorten when we pull on it and we do not expect friction to speed up its motion. We can ensure that an increase in tension accompany an increase in stretch for an elastic string by assuming that $\nu \mapsto \hat{N}(\nu, s)$ is (*strictly*) *increasing*, i.e., $\hat{N}(\nu_2, s) > \hat{N}(\nu_1, s)$ if and only if $\nu_2 > \nu_1$. This condition can be expressed more symmetrically by

$$(2.17a) \qquad [\hat{N}(\nu_2, s) - \hat{N}(\nu_1, s)][\nu_2 - \nu_1] > 0 \quad \text{if and only if} \quad \nu_2 \neq \nu_1.$$

Our statement that (2.17a) is physically reasonable does not imply that constitutive functions violating (2.17a) are unreasonable. Indeed, models satisfying (2.17a) except for ν in a small interval have been used to describe instabilities associated with phase transitions (see Ericksen (1975, 1977b), James (1979, 1980), Magnus & Poston (1979), and Carr, Gurtin, & Slemrod (1984) and the references cited in item (iv) of Sec. **14**.16).

A stronger condition, which is physically reasonable but not essential for many problems, is that $\nu \mapsto \hat{N}(\nu, s)$ be *uniformly increasing*, i.e., that there be a positive number c such that

$$(2.17b) \qquad [\hat{N}(\nu_2, s) - \hat{N}(\nu_1, s)][\nu_2 - \nu_1] > c[\nu_2 - \nu_1]^2.$$

If $\hat{N}(\cdot, s)$ is differentiable, then (2.17b) is equivalent to

$$(2.17c) \qquad \hat{N}_\nu \geq c \quad \text{everywhere.}$$

If $\hat{N}(\cdot, s)$ is differentiable, then a condition intermediate to (2.17a) and (2.17c) is that

$$(2.17d) \qquad \hat{N}_\nu > 0 \quad \text{everywhere.}$$

Conditions (2.17a,b) could be equivalently expressed as inequalities for difference quotients, but such inequalities do not naturally generalize to the case (treated extensively in later chapters) in which $\hat{N}$ is replaced with a vector-valued function. Note that there is not a perfect correspondence between our conditions on differences and those on derivatives.

One can impose hypotheses on $\hat{N}$ short of differentiability that ensure that $\hat{\nu}$ has properties somewhat better than mere continuity (and weaker than (2.17b): Suppose that $\hat{N}$ is continuous and further that there is a function f on $[0,\infty)$ with $x \mapsto f(x)/x$ strictly increasing from 0 to ∞ such that

$$[\hat{N}(\nu_1, s) - \hat{N}(\nu_2, s)](\nu_1 - \nu_2) \geq f(|\nu_1 - \nu_2|). \tag{2.17e}$$

This condition strengthens (2.17a).

Since $\nu \mapsto \hat{N}_1(\nu, 0, s)$ describes elastic response, we could require it to satisfy (2.17a). A stronger, though reasonable, restriction on $\hat{N}_1$ is that:

$$\nu \mapsto \hat{N}_1(\nu, \dot{\nu}, s) \quad \text{is strictly increasing.} \tag{2.18}$$

Similar restrictions could be placed on other constitutive functions.

The discussion of armchair experiments in the preceding paragraph is intentionally superficial. If we pull on a real string, we prescribe either its total length or the tensile forces at its ends. But in pulling the string we may produce a stretch that varies from point to point; the integral of the stretch is the total actual length. In typical experiments, one measures the tensile force at the ends when the total length is prescribed, and one measures the total length when the tensile force at the ends is prescribed. These experimental measurements of global quantities correspond to information coming from the solution of a boundary-value problem. It is in general a very difficult matter to determine the constitutive function, which has a local significance and which determines the governing equations, from a family of solutions.

For an elastic string the requirements that an infinite tensile force must accompany an infinite stretch and that an infinite compressive force must accompany a total compression to zero stretch are embodied in

$$\hat{N}(\nu, s) \to \infty \quad \text{as} \quad \nu \to \infty, \qquad \hat{N}(\nu, s) \to -\infty \quad \text{as} \quad \nu \to 0. \tag{2.19a,b}$$

The reference configuration is *natural* if the tension vanishes in it. Thus for elastic strings this property is ensured by the constitutive restriction

$$\hat{N}(1, s) = 0. \tag{2.20}$$

It is easy to express assumptions corresponding to those of this paragraph for other materials.

That (2.14) describes a material with a true internal friction, i.e., a material for which energy is dissipated in every motion, is ensured by the requirement that

$$[\hat{N}_1(\nu, \dot{\nu}, s) - \hat{N}_1(\nu, 0, s)]\, \dot{\nu} > 0 \quad \text{for} \quad \dot{\nu} \neq 0. \tag{2.21}$$

A proof that (2.21) ensures that (2.14) is 'dissipative' is given in Ex. 2.29. A stronger restriction, which ensures that the frictional force increases with the rate of stretch, is that

$$\dot{\nu} \mapsto \hat{N}_1(\nu, \dot{\nu}, s) \quad \text{is strictly increasing.} \tag{2.22a}$$

Clearly, (2.22a) implies (2.21). The function $\hat{N}_1(\nu, \cdot, s)$ can be classified just as in (2.17). Condition (2.22a) is mathematically far more tractable than (2.21), but much of modern analysis requires the yet stronger condition

$$\dot{\nu} \mapsto \hat{N}_1(\nu, \dot{\nu}, s) \quad \text{is uniformly increasing.} \tag{2.22b}$$

There are a variety of mathematically useful consequences of the constitutive restrictions we have imposed. In particular, hypothesis (2.19) and the continuity of $\hat{N}$ enable us to deduce from the Intermediate-Value Theorem that for each given $s \in [0,1]$ and $N \in \mathbb{R}$ there is a ν satisfying $\hat{N}(\nu, s) = N$. Hypothesis (2.17a) implies that this solution is unique. We denote it by $\hat{\nu}(N, s)$. Thus $\hat{\nu}(\cdot, t)$ is the inverse of $\hat{N}(\cdot, t)$, and (2.11) is equivalent to

$$\nu(s,t) = \hat{\nu}\big(N(s,t), s\big). \tag{2.23}$$

If $\hat{N}$ is continuously differentiable and satisfies the stronger hypothesis (2.17d), then the classical Local Implicit-Function Theorem implies that $\hat{\nu}$ is continuously differentiable because $\hat{N}$ is. These results constitute a simple example of a *global implicit function theorem.* We shall employ a variety of generalizations of it throughout this book.

Let g be the inverse of $x \mapsto f(x)/x$ where f is given in (2.17e). Then (2.17e) immediately implies that

$$|\hat{\nu}(N_1, s) - \hat{\nu}(N_2, s)| \le g\left(|N_1 - N_2|\right),$$

which implies that $\hat{\nu}$ is continuous and gives a modulus of continuity for it.

We substitute (2.11) or (2.14) into (2.10b) and then substitute the resulting expression into (2.9). We thus obtain a quasilinear system of partial differential equations for the components of $\boldsymbol{r}$. The full *initial-boundary-value problem for elastic strings* consists of (2.3), (2.4), (2.9), (2.10b), and (2.11). That for the viscoelastic string of strain-rate type is obtained by replacing (2.11) with (2.14). If we use (2.16b), then in place of a partial differential equation we obtain a partial functional-differential equation, for which we must supplement the initial conditions (2.4) by specifying the history of $\boldsymbol{r}$ up to time 0.

It proves mathematically convenient to recast these initial-boundary-value problems in an entirely different form, called the weak form of the equations by mathematicians and the Principle of Virtual Power (or the Principle of Virtual Work) by physicists and engineers. The traditional derivation of this formulation from (2.9) is particularly simple: We introduce the class of functions $\boldsymbol{y} \in C^1([0,1] \times [0,\infty))$ such that $\boldsymbol{y}(0,t) = \boldsymbol{o} = \boldsymbol{y}(1,t)$ (for all $t \ge 0$) and such that $\boldsymbol{y}(s,t) = \boldsymbol{o}$ for all t sufficiently large. These functions are termed *test functions* by mathematicians and *virtual velocities* (or *virtual displacements*) by physicists and engineers. We take the dot product of (2.9) with a test function $\boldsymbol{y}$ and integrate the resulting expression by parts over $[0,1]\times[0,\infty)$. Using (2.4) and the properties of $\boldsymbol{y}$ we obtain

$$\begin{aligned}(2.24)\qquad &\int_0^\infty\int_0^1 [\boldsymbol{n}(s,t)\cdot\boldsymbol{y}_s(s,t)-\boldsymbol{f}(s,t)\cdot\boldsymbol{y}(s,t)]\,ds\,dt\\ &=\int_0^\infty\int_0^1 (\rho A)(s)[\boldsymbol{r}_t(s,t)-\boldsymbol{v}(s)]\cdot\boldsymbol{y}_t(s,t)\,ds\,dt \quad \text{for all test functions } \boldsymbol{y}.\end{aligned}$$

Equation (2.24) expresses a version of the *Principle of Virtual Power* for any material. We can substitute our constitutive equations into it to get a version of this principle suitable for specific materials.

Under the smoothness assumptions in force in this section, we have shown that (2.7) and (2.4) imply (2.24). An equally simple procedure (relying on the Fundamental Lemma of the Calculus of Variations) shows that the converse is true.

2.25. Exercise. Derive (2.24) from (2.9) and (2.4) and then derive (2.9) and (2.4) from (2.24). The Fundamental Lemma of the Calculus of Variations states that if f is integrable on a measurable set $\mathcal{E}$ of $\mathbb{R}^n$ and if $\int_{\mathcal{E}} fg\,dv=0$ for all continuous g, then $f=0$ (a.e.). Here dv is the differential volume of $\mathbb{R}^n$.

Equation (2.9) is immediately integrated to yield (2.7) with $\boldsymbol{n}^+=\boldsymbol{n}^-=\boldsymbol{n}$. Then the integral form (2.7), the classical form (2.9), and the weak form (2.24) of the equations of motion are equivalent under our smoothness assumptions. In Sec. 4 we critically reexamine this equivalence in the absence of such smoothness.

2.26. Exercise. When undergoing a steady whirling motion about the $\boldsymbol{k}$-axis, a string lies in a plane rotating about $\boldsymbol{k}$ with constant angular velocity ω and does not move relative to the rotating plane. Let $\boldsymbol{f}(s,t)=g(s)\boldsymbol{k}$, where g is prescribed. Let (2.3) hold. Find a boundary-value problem for a system of ordinary differential equations, independent of t, governing the steady whirling motion of an elastic string under these conditions. Show that the steady whirling of a viscoelastic string described by (2.14) is governed by the same boundary-value problem. How is this result influenced by the frame-indifference of (2.14)? (Suppose that N were to depend on $\boldsymbol{r}_s$ and $\boldsymbol{r}_{st}$.)

2.27. Exercise. For an elastic string, let $W(\nu,s):=\int_1^\nu \hat{N}(\bar\nu,s)\,d\bar\nu$. Suppose that $\boldsymbol{f}$ has the form $\boldsymbol{f}(s,t)=\boldsymbol{g}(\boldsymbol{r}(s,t),s)$ where $\boldsymbol{g}(\cdot,s)$ is the Fréchet derivative (gradient) of the scalar-valued function $-\omega(\cdot,s)$, i.e., $\boldsymbol{g}(\boldsymbol{r},s)=-\omega_{\boldsymbol{r}}(\boldsymbol{r},s)$, where ω is prescribed. (Thus $\boldsymbol{f}$ is conservative.) W is the *stored-energy* or *strain-energy function* for the elastic string and ω is the *potential-energy density function* for the body force $\boldsymbol{f}$. Show that the integration by parts of the dot product of (2.9) with $\boldsymbol{r}_t$ over $[0,1]\times[0,\tau)$ and the use of (2.3) and (2.4) yield the *conservation of energy*:

$$\begin{aligned}(2.28)\qquad &\int_0^1 \left[W\big(\nu(s,\tau),s\big)+\omega\big(\boldsymbol{r}(s,\tau),s\big)+\tfrac12(\rho A)(s)|\boldsymbol{r}_t(s,\tau)|^2\right]\,ds\\ &\qquad=\int_0^1 \left[W\big(|\boldsymbol{u}_s(s)|,s\big)+\omega\big(\boldsymbol{u}(s),s\big)+\tfrac12(\rho A)(s)|\boldsymbol{v}(s)|^2\right]\,ds.\end{aligned}$$

(This process parallels that by which (2.24) is obtained from (2.9) and (2.4).) Show that (2.28) can be obtained directly from (2.24) and (2.3) by choosing $\boldsymbol{y}(s,t)$ in (2.24) to equal $\boldsymbol{r}_t(s,t)\chi(t,\tau,\varepsilon)$ where

$$\chi(t,\tau,\varepsilon):=\begin{cases}1 & \text{for}\quad 0\le t\le\tau,\\ 1+(\tau-t)/\varepsilon & \text{for}\quad \tau\le t\le\tau+\varepsilon,\\ 0 & \text{for}\quad \tau+\varepsilon\le t,\end{cases}$$

and then taking the limit of the resulting version of (2.24) as $\varepsilon \to 0$. See Sec. 10 for further material on energy.

2.29. Exercise. Let (2.14) hold and set $\hat{N}(\nu, s) = \hat{N}_1(\nu, 0, s)$. Define W as in Ex. 2.27. Let $\boldsymbol{f}$ have the conservative form shown in Ex. 2.26. Define the *total energy of the string at time* τ to be the left-hand side of (2.28). Form the dot product of (2.9) with $\boldsymbol{r}_t$, integrate the resulting expression with respect to s over $[0, 1]$, and use (2.3) to obtain an expression for the time derivative of the total energy at time t. This formula gives a precise meaning to the remarks surrounding (2.21).

2.30. Exercise. Formulate the boundary conditions in which the end $s = 1$ is constrained to move along a frictionless continuously differentiable curve in space. Let this curve be given parametrically by $a \mapsto \bar{\boldsymbol{r}}(a)$. (Locate the end at time t with the parameter $a(t)$.) A mechanical boundary condition is also needed.

2.31. Exercise. Formulate a suitable Principle of Virtual Power for the initial-boundary-value problem of this section modified by the replacement of the boundary condition at $s = 1$ with that of Ex. 2.30. The mechanical boundary condition at $s = 1$ should be incorporated into the principle.

The first effective steps toward correctly formulated equations for the vibrating string were made by Taylor (1713) and Joh. Bernoulli (1729). D'Alembert (1743) derived the first explicit partial differential equation for the small motion of a heavy string. The correct equations for the large vibrations of a string in a plane, equivalent to the planar version of (2.9), (2.10b), were derived by Euler (1751) in 1744 by taking the limit of the equations of motion for a finite collection of beads joined by massless elastic springs as the number of beads approaches infinity while their total mass remains fixed. The correct linear equation for the small planar transverse motion of an elastic string, which is just the wave equation, was obtained and beautifully analyzed by d'Alembert (1747). Euler (1752) stated 'Newton's equations of motion' and in his notebooks used them to derive the planar equations of motion for a string in a manner like the one just presented. A clear exposition of this derivation together with a proof that $\boldsymbol{n}^+ = \boldsymbol{n}^-$ was given by Euler (1771). Lagrange (1762) used the bead model to derive the spatial equations of motion for an elastic string. The Principle of Virtual Power in the form commonly used today was laid down by Lagrange (1788). A critical historical appraisal of these pioneering researches is given by Truesdell (1960), upon whose work this paragraph is based.

We note that the quasilinear system (2.9), (2.10b), (2.11) arising from the conceptually simple field of classical continuum mechanics is generally much harder to analyze than semilinear equations of the form $u_{tt} - u_{ss} = f(u, u_s)$, which arise in conceptually difficult fields of modern physics.

3. The Linear Impulse-Momentum Law

The partial differential equations for the longitudinal motion of an elastic string are the same as those for the longitudinal motion of a naturally straight elastic rod (for which compressive states are observed). It has long been known that solutions of these equations can exhibit shocks, i.e., discontinuities in $\boldsymbol{r}_s$ or $\boldsymbol{r}_t$. (See the discussion and references in Chap. 18.) Shocks can also arise in strings with constitutive equations of the form (2.16b) (see Renardy, Hrusa, & Nohel (1987)). On the other hand, Andrews (1980), Andrews & Ball (1982), Antman & Seidman (1996), Dafermos (1969), Greenberg, MacCamy, & Mizel (1968), Kanel' (1969), and MacCamy (1970), among many others, have shown that the longitudinal motions of nonlinearly viscoelastic strings (or rods) for special cases of (2.14) satisfying a uniform version of (2.22) do not exhibit shocks. The burden of these remarks is that the smoothness assumptions made in Sec. 2 are completely unwarranted for nonlinearly elastic strings and for certain kinds of nonlinearly viscoelastic strings.

It is clear that the integral form (2.7) of the equations of motion makes sense under smoothness assumptions weaker than those used to derive (2.9). In this section we study natural generalizations of (2.3), (2.4), (2.7), and (2.8) under such weaker assumptions. In the next section we demonstrate the equivalence of these generalizations with a precisely formulated version of the Principle of Virtual Power.

We formally integrate (2.7) with respect to t over $[0, \tau]$ and take account of (2.4) to obtain the *Linear Impulse-Momentum Law*:

$$\begin{aligned}\int_0^\tau \Big[\boldsymbol{n}^+(s,t) - \boldsymbol{n}^-(a,t)\Big]\,dt + \int_0^\tau \int_a^s \boldsymbol{f}(\xi,t)\,d\xi\,dt \\ = \int_a^s (\rho A)(\xi)[\boldsymbol{r}_t(\xi,\tau) - \boldsymbol{v}(\xi)]\,d\xi,\end{aligned} \tag{3.1}$$

which is to hold for (almost) all a, s, τ. The left-hand side of (3.1) is the *linear impulse of the force system* $\{\boldsymbol{n}^\pm, \boldsymbol{f}\}$ and the right-hand side is the *change in linear momentum* for the material segment (a, s) over the time interval $(0, \tau)$. We regard (3.1) as the natural generalization of the equations of motion (2.7).

We now state virtually the weakest possible conditions on the functions entering (3.1) for its integrals to make sense as Lebesgue integrals and for our boundary and initial conditions to have consistent generalizations. These generalizations are the highlights of the ensuing development, the details of which can be omitted by the reader unfamiliar with real analysis.

We assume that there are numbers σ^- and σ^+ such that

$$0 < \sigma^- \leq (\rho A)(s) \leq \sigma^+ < \infty \quad \forall\, s \in [0, 1]. \tag{3.2}$$

We assume that $\boldsymbol{r}_s$ and $\boldsymbol{r}_t$ are locally integrable on $[0, 1] \times [0, \infty)$, that $\boldsymbol{r}$ satisfies the boundary conditions (2.3) in the sense of *trace* (see Adams (1975), Nečas (1967)), i.e., that

$$\lim_{s \searrow 0} \int_{t_1}^{t_2} \boldsymbol{r}(s,t)\,dt = \boldsymbol{o}, \quad \lim_{s \nearrow 1} \int_{t_1}^{t_2} [\boldsymbol{r}(s,t) - L\boldsymbol{k}]\,dt = \boldsymbol{o} \quad \forall\,(t_1, t_2) \subset [0, \infty), \tag{3.3}$$

that $\boldsymbol{u}$ is integrable on $[0, 1]$, that the first initial condition of (2.4) is assumed in the sense of trace:

$$\lim_{t \searrow 0} \int_a^b (\rho A)(s)[\boldsymbol{r}(s,t) - \boldsymbol{u}(s)]\,ds = \boldsymbol{o} \quad \forall\,[a, b] \subset [0, 1], \tag{3.4}$$

and that $\boldsymbol{v}$ is integrable on $[0, 1]$. Conditions (3.3) and (3.4) are consistent with the local integrability of $\boldsymbol{r}_s$ and $\boldsymbol{r}_t$ (see Adams (1975), Nečas (1967)). We do not prescribe a generalization of the second initial condition of (2.4) because we shall show that it is inherent in (3.1), as the presence there of $\boldsymbol{v}$ suggests. We finally assume that $\boldsymbol{n}^\pm$ and $\boldsymbol{f}$ are locally integrable on $[0, 1] \times [0, \infty)$.

Since we are merely assuming that our variables are integrable over compact subsets of $[0,1]\times[0,\infty)$, we must show that the single integrals in (3.1) make sense: By Fubini's Theorem, the local integrability of $\boldsymbol{n}^+$ implies that for each $\tau\in(0,\infty)$ there is a set $\mathcal{A}^+(\tau)\subset[0,1]$ with Lebesgue measure $|\mathcal{A}^+(\tau)|=1$ such that $\boldsymbol{n}^+(s,\cdot)$ is integrable over $[0,\tau]$ for $s\in\mathcal{A}^+(\tau)$. Moreover, the Lebesgue Differentiation Theorem implies that there is a subset $\mathcal{A}_0^+(\tau)$ of $\mathcal{A}^+(\tau)$ with $|\mathcal{A}_0^+(\tau)|=1$ such that for $s\in\mathcal{A}_0^+(\tau)$, the integral $\int_0^\tau \boldsymbol{n}^+(s,t)\,dt$ has the 'right' value in the sense that it is the limit of its averages over intervals centered at s. The corresponding statements obtained by replacing the superscript '+' by '−' are likewise true. Let $\mathcal{A}(\tau):=\mathcal{A}_0^+(\tau)\cap\mathcal{A}_0^-(\tau)$. (Thus $|\mathcal{A}(\tau)|=1$ for each τ.) Let $\mathcal{B}$ be the set of $t\geq 0$ for which $(\rho A)(\cdot)\boldsymbol{r}_t(\cdot,t)$ is integrable over $[0,1]$ and for which $\int_0^1(\rho A)(s)\boldsymbol{r}_t(s,t)\,ds$ has the 'right' value. (Fubini's Theorem and Lebesgue's Differentiation Theorem imply that $|\mathcal{B}\cap[0,T]|=T$ for all $T\geq 0$.) Thus each term in (3.1) is well-defined for each $\tau\in\mathcal{B}$ and for each a and s in $\mathcal{A}(\tau)$ with $a\leq s$. Hence (3.1) holds a.e.

We now derive some important consequences from (3.1). Since Fubini's Theorem allows us to interchange the order of integration in the double integral, we can represent the the first integral on the left-hand side of (3.1) as an integral over (a,s) of an integrable function of ξ for $\tau\in\mathcal{B}$. Thus for each $\tau\in\mathcal{B}$, the function $s\mapsto\int_0^\tau \boldsymbol{n}^+(s,t)\,dt$ is absolutely continuous, not merely on $\mathcal{A}(\tau)$, but on all of $[0,1]$. Consequently,

$$\int_0^\tau \boldsymbol{n}^+(a,t)\,dt=\lim_{s\to a}\int_0^\tau \boldsymbol{n}^+(s,t)\,dt\qquad\forall\,\tau\in\mathcal{B}. \tag{3.5}$$

Then (3.1) implies that

$$\int_0^\tau \boldsymbol{n}^+(a,t)\,dt=\int_0^\tau \boldsymbol{n}^-(a,t)\,dt\qquad\forall\,\tau\in\mathcal{B}. \tag{3.6}$$

Thus the superscripts '+' and '−' are superfluous even in this more general setting and will accordingly be dropped.

The properties of the Lebesgue integral imply that if $a,s\in\mathcal{A}(T)$, then $a,s\in\mathcal{A}(\tau)$ for all $\tau\in[0,T]$. Let us fix $T>0$. Let $a,s\in\mathcal{A}(T)$. Since left-hand side of (3.1) is an integral over $(0,\tau)$ of an integrable function of t, the right-hand side of (3.1) defines an absolutely continuous function of τ for $a,s\in\mathcal{A}(T)$. Thus

$$\lim_{\tau\searrow 0}\int_a^s(\rho A)(\xi)[\boldsymbol{r}_t(\xi,\tau)-\boldsymbol{v}(\xi)]\,d\xi=\boldsymbol{o}\quad\forall\,a,s\in\mathcal{A}(T). \tag{3.7}$$

This generalization of the second initial condition of (2.4), which has the same form as (3.4), is thus implicit in (3.1).

It is important to note that the generalizations (3.3), (3.4), (3.7) of the boundary and initial conditions (2.3) and (3.3) represent averages of the classical pointwise conditions. As such, the limiting processes they embody correspond precisely to the way they could be tested experimentally.

Our basic smoothness assumption underlying the development of this section is the local integrability of $\boldsymbol{r}_s$, $\boldsymbol{r}_t$, and $\boldsymbol{n}$. Since $\boldsymbol{n}$ is to be given as a constitutive function of the stretch and possibly other kinematic variables, the local integrability of $\boldsymbol{n}$ imposes restrictions on the class of suitable constitutive functions.

In the modern study of shocks, physically realistic solutions $\boldsymbol{r}$ are sought in larger classes of functions, such as functions of bounded variation, which need not have locally integrable derivatives. Thus there is a need for mathematically sound and physically realistic generalizations of the development of this and the next section. (See the refernces cited at the end of Sec. **12**.9.)

3.8. Exercise. Repeat Ex. 2.26, but now obtain the same equations for the steady whirling of the string directly from (3.1) and (3.6). This derivation can easily be performed with complete mathematical rigor.

4. The Equivalence of the Linear Impulse-Momentum Law with the Principle of Virtual Power

In this section we prove that the Linear Impulse-Momentum Law formulated in Sec. 3 is equivalent to a generalized version of the Principle of Virtual Power stated in (2.24). Our proof is completely rigorous and technically simple. Although we couch our presentation in the language of real analysis to ensure complete precision, all the steps have straightforward interpretations in terms of elementary calculus.

The demonstration of equivalence given at the end of Sec. 2, which is universally propounded by mathematicians and physicists alike, pivots on the classical form (2.9) of the equations of motion. But this form is devoid of meaning in the very instances when the Linear Impulse-Momentum Law and the Principle of Virtual Power are essential, i.e., when there need not be classical solutions. In our approach given below, Eq. (2.9) never appears.

Since (2.9) never appears, it is therefore never exposed to abuse. The most dangerous sort of abuse would consist in multiplying (2.9) by a positive-valued function depending on the unknowns appearing in (2.9), thereby converting (2.9) to an equivalent classical form. But its corresponding weak form, obtained by the procedure leading to (2.24), would not be equivalent to (2.24), because the integration by parts would produce additional terms caused by the presence of the multiplicative factor. Consequently, the corresponding jump conditions at discontinuities (see Sec. 5) would have forms we deem wrong because they are incompatible with the jump conditions coming from the generalization of (2.24). This generalization is deemed correct because, as we shall show, it is equivalent to the Linear Impulse-Momentum Law, which we regard as a fundamental principle of mechanics.

Note that the Principle of Virtual Power as stated in (2.24) makes sense when the smoothness restrictions imposed on $\boldsymbol{r}$ and $\boldsymbol{n}$ in Sec. 2 are replaced by the much weaker conditions of Sec. 3. The resulting form of (2.24) can be further extended to apply to all test functions $\boldsymbol{y}$ that have essentially bounded generalized derivatives, that vanish for large t, and that vanish in the sense of trace on the boundaries $s = 0$ and $s = 1$. (These functions form a subspace of the Sobolev space $W^1_\infty([0,1]\times[0,\infty))$.) The smoothness assumptions on the variables entering these formulations are the weakest that allow all the integrals to make sense as Lebesgue integrals. We refer to the resulting version of (2.24) as the *generalized* Principle of Virtual Power.

We now derive this principle from the Linear Impulse-Momentum Law under the assumptions of Sec. 3. Let ϕ be a polygonal (piecewise affine) function of s with support in $(a, b) \subset [0, 1]$ and let ψ be a polygonal function of t with support in $[0, \tau)$. (The *support* of a function is the closure of the set on which it is not zero.) Note that the support of ψ is contained in a half-closed interval. Let $\boldsymbol{e}$ be a fixed but arbitrary constant unit vector.

Then (3.1) implies that

$$\begin{aligned}(4.1)\qquad \int_0^\tau \int_a^b \phi_s(s)\psi_t(t) \bigg\{ &\int_0^t \boldsymbol{e} \cdot [\boldsymbol{n}(s,\bar{t}\,) - \boldsymbol{n}(a,\bar{t}\,)]\, d\bar{t} \\ &+ \int_0^t \int_a^s \boldsymbol{e} \cdot \boldsymbol{f}(\bar{s},\bar{t}\,)\, d\bar{s}\, d\bar{t} \bigg\}\, ds\, dt \\ = \int_0^\tau \int_a^b \phi_s(s)\psi_t(t) &\int_a^s (\rho A)(\bar{s})\boldsymbol{e} \cdot [\boldsymbol{r}_t(\bar{s},t) - \boldsymbol{v}(\bar{s})]\, d\bar{s}\, ds\, dt.\end{aligned}$$

Since ψ and ϕ are absolutely continuous, we can integrate the triple integral on the left-hand side of (4.1) by parts with respect to t, we can integrate the quadruple integral on the left-hand side of (4.1) by parts with respect to t and s, and we can integrate the right-hand side of (4.1) by parts with respect to s. Since $\psi(\tau) = 0$, $\phi(a) = 0 = \phi(b)$, we thereby convert (4.1) to

$$\begin{aligned}(4.2)\qquad \int_0^\tau \int_a^b \phi_s(s)\psi(t)\boldsymbol{e} \cdot \boldsymbol{n}(s,t)\, ds\, dt &- \int_0^\tau \int_a^b \phi(s)\psi(t)\boldsymbol{e} \cdot \boldsymbol{f}(s,t)\, ds\, dt \\ &= \int_0^\tau \int_a^b \phi(s)\psi_t(t)\boldsymbol{e} \cdot (\rho A)(s)[\boldsymbol{r}_t(s,t) - \boldsymbol{v}(s)]\, ds\, dt.\end{aligned}$$

Let us set

$$\boldsymbol{y}(s,t) = \phi(s)\psi(t)\boldsymbol{e}. \tag{4.3}$$

Since this $\boldsymbol{y}$ has support in $(a,b) \times [0,\tau)$, we can write (4.2) in the form (2.24) for all $\boldsymbol{y}$'s of the form (4.3). More generally, (2.24) holds for all $\boldsymbol{y}$'s in the space that is the completion in the norm of $W^1_\infty([0,1] \times [0,\infty))$ of finite linear combinations of functions of the form (4.3). (Some properties of this space are discussed by Antman & Osborn (1979).)

Equation (2.24) for this large class of $\boldsymbol{y}$'s expresses the *generalized Principle of Virtual Power* or the *Weak Form* of (2.9), (2.3), and (2.4). We henceforth omit the adjective *generalized*. (If we allow $\boldsymbol{n}$ and $\boldsymbol{f}$ to be smoother, we can allow the $\boldsymbol{y}$'s to be rougher.) The *Weak Form of the Initial-Boundary Value Problem* for elastic strings is obtained by inserting (2.10b) and (2.11) into (2.24) and appending (3.3) and (3.4). Analogous definitions hold for other materials.

Without making unwarranted smoothness assumptions, we have thus shown that the Linear Impulse-Momentum Law implies the Principle of Virtual Power. Conversely, we can likewise recover (3.1) (without the superscripts '$\pm$') from (2.24) by taking ε to be a small positive number, taking $\boldsymbol{y}$ to have the form (4.3) (which reduces (2.24) to (2.42)), taking ϕ and ψ to have the forms

$$\phi(\bar{s}) = \begin{cases} 0 & \text{for} \quad 0 \le \bar{s} \le a, \\ \frac{\bar{s}-a}{\varepsilon} & \text{for} \quad a \le \bar{s} \le a+\varepsilon, \\ 1 & \text{for} \quad a+\varepsilon \le \bar{s} \le s-\varepsilon, \\ \frac{s-\bar{s}}{\varepsilon} & \text{for} \quad s-\varepsilon \le \bar{s} \le s, \\ 0 & \text{for} \quad s \le \bar{s} \le 1, \end{cases} \tag{4.4a}$$

$$\psi(t) = \begin{cases} 1 & \text{for} \quad 0 \le t \le \tau, \\ 1 - \frac{t-\tau}{\varepsilon} & \text{for} \quad \tau \le t \le \tau + \varepsilon, \\ 0 & \text{for} \quad \tau + \varepsilon \le t, \end{cases} \tag{4.4b}$$

and then letting $\varepsilon \to 0$. (The functions of (4.4) should be sketched. If H is the Heaviside function, then ϕ is a Lipschitz continuous approximation to $\bar{s} \mapsto H(\bar{s}-a) - H(\bar{s}-s)$ and ψ is a Lipschitz continuous approximation to $t \mapsto 1 - H(t-\tau)$.) In this process, we must evaluate the typical expression

$$\lim_{\varepsilon \to 0} \int_0^{\tau+\varepsilon} \frac{1}{\varepsilon} \int_a^{a+\varepsilon} \boldsymbol{n}(s,t) \cdot \boldsymbol{e} \psi(t) \, ds \, dt, \tag{4.5}$$

which Fubini's Theorem and (4.4b) allow us to rewrite as

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_a^{a+\varepsilon} \int_0^{\tau+\varepsilon} \boldsymbol{n}(s,t) \cdot \boldsymbol{e} \, dt \, ds - \lim_{\varepsilon \to 0} \left\{ \varepsilon \left[\frac{1}{\varepsilon^2} \int_a^{a+\varepsilon} \int_\tau^{\tau+\varepsilon} \boldsymbol{n}(s,t) \cdot \boldsymbol{e} \frac{t-\tau}{\varepsilon} \, dt \, ds \right] \right\}. \tag{4.6}$$

The Lebesgue Differentiation Theorem implies that the first term in (4.6) is

$$\int_0^\tau \boldsymbol{n}(a,t) \cdot \boldsymbol{e} \, dt \tag{4.7}$$

for almost all a in $(0,1)$ and that the supremum of the absolute value of the bracketed expression in the second term of (4.6) is finite for almost all a in $(0,1)$ and τ in $(0,\infty)$. (Note that $|(t-\tau)/\varepsilon| \le 1$ for $t \in [\tau, \tau+\varepsilon]$.) Thus (4.5) equals (4.7). The other terms are treated similarly. The arbitrariness of $\boldsymbol{e}$ allows it to be cancelled in the final expression. Thus (2.24) implies (3.1) and these two principles are equivalent.

The Principle of Virtual Power can be used to exclude certain naive solutions of differential equations as unphysical. We illustrate this property with a differential equation simpler than that for a string. Consider the boundary-value problem

$$u''(s) + \pi^2 u(s) = 0 \quad \text{on} \quad (-1,1), \quad u(\pm 1) = 0. \tag{4.8}$$

The continuous function u^* defined by $u^*(s) = |\sin \pi s|$ is in $W_1^1(-1,1)$, satisfies the boundary conditions, and satisfies the differential equation everywhere except at 0. Other than its failure to be a classical solution of the boundary-value problem, there is nothing intrinsically wrong with u^* from a purely mathematical standpoint. Now suppose that (4.8) is regarded as a symbolic representation for the weak problem

$$\int_{-1}^1 [u'v' - \pi^2 uv] \, ds = 0 \quad \forall\, v \in C^1[-1,1] \quad \text{with} \quad v(\pm 1) = 0, \quad u(\pm 1) = 0, \tag{4.9}$$

solutions of which are sought in $W_1^1(-1,1)$. We presume that (4.9) embodies a Principle of Virtual Power, representing a description of the underlying physics more fundamental than that given by (4.8). It is easy to show that u^* does not satisfy (4.9) and can therefore be excluded as unrealistic. (Just substitute u^* into (4.9) and integrate the resulting system by parts on $[-1,0]$ and $[0,1]$ obtaining $v(0)=0$. Since there are v's that do not vanish at the origin, u^* is not a solution of (4.9).) Indeed, by using methods like those of (4.4)–(4.6) or of Sec. 6, we can show that every (weak) solution of (4.9) is a classical solution of (4.8). In the next section we show how the Principle of Virtual Power enables us to classify precisely those kinds of jumps that are compatible with it.

In much of modern mathematical literature, the classical form of an equation is regarded as merely an abbreviation for the weak form. Since weak formulations of equivalent classical formulations need not be equivalent, this convention should be used with care. The weak form is also sometimes termed the *variational form*, an expression we never employ because it connotes far more generality than the notion of variational structure introduced in Sec. 10.

If there are concentrated or impulsive forces applied to the string, then $\boldsymbol{f}$ would not be locally integrable, and the development of these last two sections would not be valid. Distribution theory, which was designed to handle linear equations with such forces, has recently been extended to handle nonlinear equations (see Colombeau (1990), Rosinger (1987)). But it is not evident how to obtain (3.6) in such a more general setting. In Sec. 6 we comment further on this question for a degenerately simple static problem.

5. Jump Conditions

We now show how the Principle of Virtual Power yields jump conditions that weak solutions must satisfy at their discontinuities.

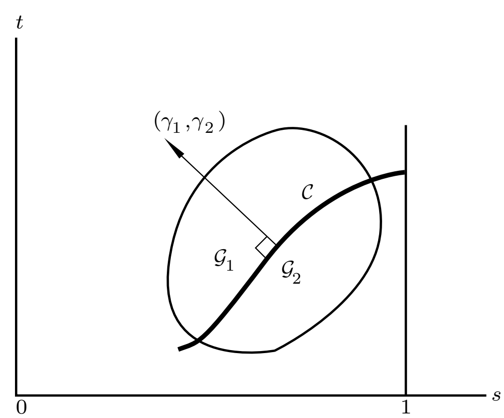

Figure 5.1. The neighborhood of a curve of discontinuity.

Let $\mathcal{C} \in [0,1]\times[0,\infty)$ be (the image of) a simple curve. We assume that $\mathcal{C}$ is so smooth that it possesses a unit normal (γ_1, γ_2) at almost every point.

(It suffices for $\mathcal{C}$ to be uniformly Lipschitz-continuous, i.e., that there be a finite number of open sets covering $\mathcal{C}$ such that in each such set $\mathcal{E}$ there is a coordinate system with respect to which $\mathcal{C} \cap \mathcal{E}$ can be described as the graph of a Lipschitz continuous function.) Suppose that there are two disjoint, simply-connected open sets $\mathcal{G}_1$ and $\mathcal{G}_2$ such that $\emptyset \neq \partial\mathcal{G}_1 \cap \partial\mathcal{G}_2 \subset \mathcal{C}$ (see Fig. 5.1), that (2.9) holds in the classical sense in $\mathcal{G}_1$ and in $\mathcal{G}_2$, and suppose that there are integrable functions $\boldsymbol{n}^1, \boldsymbol{n}^2, \boldsymbol{r}_t^1, \boldsymbol{r}_t^2$ on $\mathcal{C}$ such that

$$\boldsymbol{n} \to \boldsymbol{n}^\alpha, \ \boldsymbol{r}_t \to \boldsymbol{r}_t^\alpha \quad \text{(in the sense of trace)} \qquad \text{as} \quad \mathcal{G}_\alpha \ni (s,t) \to \mathcal{C}, \quad \alpha = 1, 2. \tag{5.2}$$

Set

$$[\![\boldsymbol{n}]\!] := \boldsymbol{n}^2 - \boldsymbol{n}^1, \quad [\![\boldsymbol{r}_t]\!] := \boldsymbol{r}_t^2 - \boldsymbol{r}_t^1 \quad \text{on } \mathcal{C}. \tag{5.3}$$

$[\![\boldsymbol{n}]\!]$ is called the *jump* in $\boldsymbol{n}$ across $\mathcal{C}$.

If $\boldsymbol{y}$ is taken to have support in $\mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{C}$, then (2.24) reduces to

$$\int_{\mathcal{G}_1 \cup \mathcal{G}_2} [\boldsymbol{n} \cdot \boldsymbol{y}_s - \boldsymbol{f} \cdot \boldsymbol{y} - \rho A(\boldsymbol{r}_t - \boldsymbol{v}) \cdot \boldsymbol{y}_t] \, ds \, dt = 0 \tag{5.4}$$

for all such $\boldsymbol{y}$'s. We separately integrate (5.4) by parts over $\mathcal{G}_1$ and $\mathcal{G}_2$ (by means of the divergence theorem), noting that (2.9) is satisfied in each region and that $\boldsymbol{y}$ vanishes on $\partial\mathcal{G}_1 \setminus \mathcal{C}$ and on $\partial\mathcal{G}_2 \setminus \mathcal{C}$. We obtain

$$\int_{\mathcal{C}} \boldsymbol{y} \cdot \{[\![\boldsymbol{n}]\!]\gamma_1 - \rho A [\![\boldsymbol{r}_t]\!]\gamma_2\} \, d\lambda = 0 \tag{5.5}$$

for all such $\boldsymbol{y}$'s. Here $d\lambda$ is the differential arc length along $\mathcal{C}$. Since $\boldsymbol{y}$ is arbitrary on $\mathcal{C}$, Eq. (5.5) implies that

$$[\![\boldsymbol{n}]\!]\gamma_1 - \rho A [\![\boldsymbol{r}_t]\!]\gamma_2 = \boldsymbol{o} \quad \text{a.e. on } \mathcal{C}. \tag{5.6}$$

These are the *Rankine-Hugoniot jump conditions* for (2.24). A curve in the (s,t)-plane across which there are jumps in $\boldsymbol{n}$ or $\boldsymbol{r}_t$ is called a *shock* (*path*). A solution suffering such a jump is said to have (or be) a *shock*. Suppose that the shock path has the equation $s = \sigma(t)$. Then $\sigma'(t)$ is the *shock speed* at $(\sigma(t), t)$. Equation (5.6) thus has the form

$$[\![\boldsymbol{n}]\!] + \rho A \sigma' [\![\boldsymbol{r}_t]\!] = \boldsymbol{o}. \tag{5.7}$$

The foregoing analysis leading to (5.6) is formal to the extent that solutions are presumed classical except on isolated curves. For further information on jump conditions and shocks, see Chap. 18 and the references cited there.

6. The Existence of a Straight Equilibrium State

When none of the variables appearing in the Linear Impulse-Momentum Law depends on the time, it reduces to the static form of (2.7), (2.8):

$$\boldsymbol{n}(s) - \boldsymbol{n}(a) + \int_a^s \boldsymbol{f}(\xi)\, d\xi = \boldsymbol{o} \tag{6.1}$$

for (almost) all a and s in $[0,1]$. If $\boldsymbol{f}$ is Lebesgue-integrable, then (6.1) implies that $\boldsymbol{n}$ is absolutely continuous and has a derivative almost everywhere. Thus the classical equilibrium equation

$$\boldsymbol{n}'(s) + \boldsymbol{f}(s) = \boldsymbol{o} \tag{6.2}$$

holds a.e. When (2.3a) holds, the Principle of Virtual Power (2.24) reduces to

$$\int_0^1 [\boldsymbol{n}(s) \cdot \boldsymbol{y}'(s) - \boldsymbol{f}(s) \cdot \boldsymbol{y}(s)]\, ds = 0 \tag{6.3}$$

for all sufficiently smooth $\boldsymbol{y}$ that vanish at 0 and 1.

Note that in equilibrium the constitutive equation (2.14) reduces to that for an elastic string, namely (2.11). (Indeed, if the string has been in equilibrium for its entire past history, then the general constitutive equation (2.16b) itself reduces to (2.11). This observation must be interpreted with care, because a string described by (2.16b) can creep under the action of an equilibrated system of forces not varying with time.) We accordingly limit our attention to elastic strings, described by (2.10b) and (2.11):

$$\boldsymbol{n}(s) = \hat{N}\big(\nu(s), s\big)\frac{\boldsymbol{r}'(s)}{\nu(s)}. \tag{6.4}$$

We assume that $\hat{N}$ is continuously differentiable, although much can be done with $\hat{N}$'s that are merely continuous. (See the remarks surrounding (2.17e).)

For integrable $\boldsymbol{f}$ we now study the boundary-value problem of finding a function $\boldsymbol{r}$, whose (distributional) derivative $\boldsymbol{r}'$ is integrable, that satisfies the system (6.1), (6.4), (2.3), which we record as

$$\hat{N}\big(|\boldsymbol{r}'(\xi)|, \xi\big)\frac{\boldsymbol{r}'(\xi)}{|\boldsymbol{r}'(\xi)|}\bigg|_a^s + \int_a^s \boldsymbol{f}(\xi)\, d\xi = \boldsymbol{o}, \tag{6.5}$$

$$\boldsymbol{r}(0) = \boldsymbol{o}, \quad \boldsymbol{r}(1) = L\boldsymbol{k}. \tag{6.6}$$

(Conditions (6.6) interpreted as (2.3a) make sense because $\boldsymbol{r}$ is the indefinite integral of the integrable function $\boldsymbol{r}'$ and is accordingly absolutely continuous.)

In Chap. 3 we shall study a rich collection of problems for (6.5) and (6.6) in which $\boldsymbol{f}$ depends on $\boldsymbol{r}$. Here we content ourselves with the study of straight equilibrium configurations $s \mapsto \boldsymbol{r}(s) = z(s)\boldsymbol{k}$, in which z is an absolutely continuous, increasing function, when $\boldsymbol{f}$ has the special form

$$\boldsymbol{f}(s) = g(s)\boldsymbol{k}. \tag{6.7}$$

In keeping with the local integrability of $\boldsymbol{f}$ assumed in Secs. 3 and 4, we take g to be Lebesgue-integrable. The requirement that $z'(s) > 0$ for all s ensures (2.2). Under these conditions, the problem (6.5), (6.6) reduces to finding z and a constant K such that

$$\hat{N}(z'(s), s) = G(s) + K \quad \forall\, s \in [0,1], \quad G(s) := -\int_0^s g(\xi)\, d\xi, \tag{6.8}$$

$$z(0) = 0, \quad z(1) = L. \tag{6.9a,b}$$

In view of the equivalence of (2.11) with (2.23), Eq.(6.8) is equivalent to

$$z'(s) = \hat{\nu}\big(G(s) + K, s\big). \tag{6.10}$$

The properties of $\hat{\nu}$ ensure that $z'(s) > 0$ for all s. We integrate (6.10) subject to (6.9a) to obtain

$$z(s) = \int_0^s \hat{\nu}\big(G(\xi) + K, \xi\big)\, d\xi. \tag{6.11}$$

The boundary-value problem (6.8), (6.9) has this z as a solution provided K can be chosen so that (6.11) satisfies (6.9b), i.e., so that

$$\Phi(K) := \int_0^1 \hat{\nu}\big(G(\xi) + K, \xi\big)\, d\xi = L. \tag{6.12}$$

Note that since G is the indefinite integral of the integrable function g, it is absolutely continuous. It follows that the function $\Phi(\cdot)$, just like $\hat{\nu}(\cdot, s)$, strictly increases from 0 to ∞ as its argument increases from $-\infty$ to ∞. Since $L > 0$, we can reproduce the argument justifying the existence of $\hat{\nu}$ to deduce that (6.12) has a unique solution K (depending on G and L). The solution of (6.8), (6.9) is then obtained by substituting this K into (6.11).

Since G and $\hat{\nu}$ are continuous, (6.11) implies that the solution z is continuously differentiable and its derivative is given by (6.10). Let us now suppose that g is continuous. Then G is continuously differentiable. Since $\hat{\nu}$ is continuously differentiable, (6.10) implies that the solution z is twice continuously differentiable. Since (6.10) is equivalent to (6.8), we can accordingly differentiate (6.8) to show that z is a classical solution of the ordinary differential equation

$$\tfrac{d}{ds}\hat{N}\big(z'(s), s\big) + g(s) = 0. \tag{6.13}$$

(The regularity theory of this paragraph is called a *bootstrap* argument.) We summarize our results:

6.14. Theorem. *Let $\hat{N}$ be continuously differentiable on $(0, \infty) \times [0, 1]$, $\hat{N}_\nu(\nu, s) > 0$ for all ν and s, $\hat{N}(\nu, s) \to \infty$ as $\nu \to \infty$, and $\hat{N}(\nu, s) \to -\infty$ as $\nu \to 0$. Let g be Lebesgue-integrable. Then* (6.8), (6.9) *has a unique solution z, which is continuously differentiable and satisfies $z'(s) > 0$ for all s. If g is continuous, then z is twice continuously differentiable and satisfies* (6.13).

Note that (6.10) implies that the stretch z' is constant if the material is uniform, i.e., if $\hat{N}_s = 0$, and if G is constant, i.e., if $g = 0$.

Equation (6.8) and its equivalent, (6.10), make sense if G is merely integrable. In this case, g is defined as the distributional derivative of $-G$. Our analysis goes through with the solution z continuous by (6.11). The only trouble with such a solution lies in its mechanical interpretation: Our proof in Sec. 3 that $\boldsymbol{n}^+ = \boldsymbol{n}^-$ is no longer applicable. Much of the difficulty with this question evaporates if the distribution g were to equal an integrable function a.e. In particular, if G were a Heaviside (i.e., a step) function, then g would be a Dirac delta, and our problem, which would be solvable, would also make mechanical sense.

Note that the unique solution of (6.8), (6.9) may well represent a compressed straight state. This certainly occurs if $g = 0$ and $L < 1$. Such a solution should certainly be unstable under any reasonable physical criterion. (It is unique only among all straight equilibrium states.) This solution is nevertheless worthy of study because its equations are exactly those for the straight equilibrium state of a naturally straight rod, whose bending stiffness allows it to sustain a certain amount of compression without losing stability. A knowledge of the properties of the straight states of a straight rod is necessary for the study of its buckling from that state.

Note that (6.5) and (6.6) may admit straight folded solutions in which z is not increasing. These can have a very complicated structure (see Reeken (1984a) and the treatments of Chaps. 3 and 6). These solutions are not accounted for by Theorem 6.14.

7. Purely Transverse Motions

The ad hoc assumption that the motion of each material point is confined to a plane perpendicular to the line joining the ends of the string, frequently used in textbook derivations of the equations of motion of strings and discussed in Sec. 1, motivates our study in this section of conditions under which such special motions can occur.

J. B. Keller (1959) and B. Fleishman (1959) independently observed that if an elastic string has a constitutive equation of the form

$$\hat{N}(\nu, s) = (EA)(s)\nu \tag{7.1}$$

where EA is a given positive-valued function, then the equations of motion (2.9), (2.10b), and (2.11) reduce to the special form

$$[(EA)(s)\boldsymbol{r}_s(s,t)]_s + \boldsymbol{f}(s,t) = (\rho A)(s)\boldsymbol{r}_{tt}(s,t). \tag{7.2}$$

(The ungainly symbol EA is used because it roughly conforms to traditional engineering notation. See the discussion of the notation ρA in Sec. 2.) If $\boldsymbol{f}$ does not depend on $\boldsymbol{r}$ through a relation such as (2.6), then (7.2) is a system of three uncoupled nonhomogeneous wave equations. In particular, if f satisfies

$$\boldsymbol{k} \cdot \boldsymbol{f}(s,t) = -G'(s) \tag{7.3}$$

and if the initial data satisfy

$$\boldsymbol{k} \cdot \boldsymbol{u} = z, \quad \boldsymbol{k} \cdot \boldsymbol{v} = 0 \tag{7.4}$$

where z is the unique solution of (6.8) and (6.9), then the initial-boundary-value problem consisting of (7.2), (2.3), and (2.4) has a unique solution with $\boldsymbol{k} \cdot \boldsymbol{r} = z$. (The existence and uniqueness of such a solution, under mild conditions on the data, follows from the theory of partial differential equations.) This solution describes a purely transverse motion.

Of course, (7.1) satisfies neither (2.19b) nor (2.20). Keller noted, however, that (7.1) could closely approximate the behavior of certain rubber strings when ν is large. This observation does not imply that the motion of a string satisfying a constitutive equation close to (7.1) is close to the motion given by (7.2), because a small nonlinear coupling can shift energy from one mode to another, as is well known in rigid-body mechanics. In particular, even if (7.3) and (7.4) hold, a string with a constitutive equation close to (7.1) could undergo motions with a significant longitudinal component.

We now address the converse problem of determining what restrictions are imposed on the constitutive functions by the assumption that the string must execute a nontrivial purely transverse motion with

$$\boldsymbol{r}(s,t) \cdot \boldsymbol{k} = z(s), \quad (\boldsymbol{r} \cdot \boldsymbol{i})^2 + (\boldsymbol{r} \cdot \boldsymbol{j})^2 \neq 0 \tag{7.5a,b}$$

for every $\boldsymbol{f}$ satisfying (7.3) and for all initial conditions satisfying (7.4) when z satisfies (6.8) and (6.9) and when ν lies in a certain interval (ν^-, ν^+) in $(0, \infty)$ with $\nu^- \leq \min z'$. The substitution of (7.5a) and (2.10b) into the $\boldsymbol{k}$-component of (2.9) yields

$$\left[\hat{N}\big(\nu(s,t), s\big) \frac{z'(s)}{\nu(s,t)}\right]_s - G'(s) = 0, \tag{7.6}$$

so that

$$\Omega\big(\nu(s,t), s\big) := \frac{\hat{N}\big(\nu(s,t), s\big) z'(s)}{\nu(s,t)} - G(s) = \Omega\big(\nu(0,t), 0\big). \tag{7.7}$$

The theory of initial-value problems for quasilinear partial differential equations, applied to the full system of governing equations, says that in a small neighborhood of the initial time, solutions depend continuously on smooth initial data. Thus smooth initial data satisfy (7.7) for $t = 0$. For any fixed s we can prescribe the initial data $\nu(s, 0)$ and $\nu(0, 0)$ arbitrarily in (ν^-, ν^+). Thus from (7.7) at $t = 0$ we conclude that Ω is a constant function. This constancy of Ω ensures that $\hat{N}(\cdot, s)$, restricted to (ν^-, ν^+), has the form (7.1). We summarize this argument, a modified version of that of J. B. Keller (1959):

7.8. Theorem. *Let* (7.3) *and* (7.4) *hold. If every solution of* (2.9), (2.10b), (2.11), (2.3), (2.4) *for which* $\nu^- < \nu < \nu^+$ *is purely transverse, i.e., satisfies* (7.5), *then* $\hat{N}(\cdot, s)$ *restricted to* (ν^-, ν^+) *has the form* (7.1).

7.9. Problem. Let (7.3) and (7.4) hold. Suppose that Ω is independent of s and that the initial-boundary-value problem admits *a* nontrivial purely transverse motion satisfying (7.5). What restrictions are thereby imposed on $\hat{N}$?

The following exercise, proposed by J. M. Greenberg, also indicates the role played by linear, or more generally, affine constitutive relations.

7.10. Exercise. Consider the free motion of a uniform, nonlinearly elastic string of doubly infinite length. Thus $\boldsymbol{f} = \boldsymbol{o}$, $\hat{N}_s = 0$, $s \in (-\infty, \infty)$. A solution $\boldsymbol{r}$ of the governing equations is called a *travelling wave* iff it has the form

$$\boldsymbol{r}(s,t) = \boldsymbol{p}(s - ct) \tag{7.11}$$

where c is a real number. Show that if there is no nonempty open interval of $(0, \infty)$ on which $\hat{N}$ is affine, then the travelling waves in a string have very special and uninteresting forms. Determine those forms. (In Sec. **9**.3 we shall see that the equations for rods have a very rich collection of travelling waves.)

8. Perturbation Methods and the Linear Wave Equation

In Sec. 6 we proved the existence of a unique straight equilibrium configuration $z\boldsymbol{k}$ for an elastic string when $\boldsymbol{f} = g(s)\boldsymbol{k}$. In this section we study the motion of an elastic string near this equilibrium state by formal perturbation methods. We first outline their application to the initial-boundary-value problem (2.9), (2.10b), (2.11), (2.3), (2.4) and then give a detailed treatment of time-periodic solutions. We discuss the validity of the perturbation methods in the next section.

We begin by studying the initial-boundary-value problem when the data are close to those yielding the straight equilibrium state: Let ε represent a small real parameter and let the data have the form

$$\text{(8.1)}\quad \boldsymbol{u}(s) = z(s)\boldsymbol{k} + \varepsilon\boldsymbol{u}_1(s), \quad \boldsymbol{v}(s) = \varepsilon\boldsymbol{v}_1(s), \quad \boldsymbol{f}(s,t) = g(s)\boldsymbol{k} + \varepsilon\boldsymbol{f}_1(s,t)$$

where z is the solution of the equilibrium problem given in Sec. 6. For well-behaved solutions, the inital data should satisfy the compatibility conditions $\boldsymbol{u}_1(0) = \boldsymbol{o} = \boldsymbol{v}_1(0)$, $\boldsymbol{u}_1(1) = \boldsymbol{o} = \boldsymbol{v}_1(1)$. We suppose that $\hat{N}(\cdot, s)$ is $(p+1)$-times continuously differentiable. We seek formal solutions of the initial-boundary-value problem whose dependence on the parameter ε is specified by a representation of the form

$$\text{(8.2)}\qquad \boldsymbol{r}(s,t,\varepsilon) = z(s)\boldsymbol{k} + \sum_{k=1}^{p} \frac{\varepsilon^k}{k!}\boldsymbol{r}_k(s,t) + o(\varepsilon^p).$$

Since (8.2) implies that

$$\text{(8.3)}\qquad \boldsymbol{r}_k(s,t) = \left.\frac{\partial^k \boldsymbol{r}(s,t,\varepsilon)}{\partial\varepsilon^k}\right|_{\varepsilon=0}, \qquad k = 1, \dots, p,$$

we can find the problem formally satisfied by $\boldsymbol{r}_k$ by substituting $\boldsymbol{r}(s,t,\varepsilon)$ into the equations of the nonlinear problem, differentiating the resulting equations k times with respect to ε, and then setting $\varepsilon = 0$. We find that the equation for $\boldsymbol{r}_k$ is linear and involves $\boldsymbol{r}_1, \dots, \boldsymbol{r}_{k-1}$; thus the equations for $\boldsymbol{r}_1, \dots, \boldsymbol{r}_p$ can be solved successively.

To compute these equations directly in vectorial form, we define

$$\text{(8.4)}\qquad \hat{\boldsymbol{n}}(\boldsymbol{q}, s) := \hat{N}(|\boldsymbol{q}|, s)\boldsymbol{q}|\boldsymbol{q}|^{-1}.$$

Thus (2.9), (2.10b), (2.11) has the form

$$\text{(8.5)}\qquad \hat{\boldsymbol{n}}(\boldsymbol{r}_s, s)_s + \boldsymbol{f} = \rho A \boldsymbol{r}_{tt}.$$

We use the definition of Gâteaux derivative given after (**1**.4.5) to obtain

$$\text{(8.6)}\qquad \hat{\boldsymbol{n}}_{\boldsymbol{q}}(\boldsymbol{q}, s) \cdot \boldsymbol{c} = \hat{N}_\nu(|\boldsymbol{q}|, s)\frac{\boldsymbol{q}\,\boldsymbol{q}\cdot\boldsymbol{c}}{|\boldsymbol{q}|^2} + \frac{\hat{N}(|\boldsymbol{q}|, s)}{|\boldsymbol{q}|}\left[\boldsymbol{c} - \frac{\boldsymbol{q}\,\boldsymbol{q}\cdot\boldsymbol{c}}{|\boldsymbol{q}|^2}\right].$$

(To differentiate $|\boldsymbol{q}|$ with respect to $\boldsymbol{q}$, we write it as $\sqrt{\boldsymbol{q}\cdot\boldsymbol{q}}$ so that $\partial\sqrt{\boldsymbol{q}\cdot\boldsymbol{q}}/\partial\boldsymbol{q} = \boldsymbol{q}/\sqrt{\boldsymbol{q}\cdot\boldsymbol{q}}$.) Thus we find that

$$\partial_\varepsilon \hat{\boldsymbol{n}}(\boldsymbol{r}_s(s,t,\varepsilon),s)|_{\varepsilon=0} = \hat{\boldsymbol{n}}_{\boldsymbol{q}}(z'(s)\boldsymbol{k},s)\cdot\partial_s\boldsymbol{r}_1(s,t). \tag{8.7}$$

Note that $\hat{\boldsymbol{n}}_{\boldsymbol{q}}$ is symmetric.

Differentiating (8.5) once with respect to ε and using (8.6) and (8.7) we reduce the equation for $\boldsymbol{r}_1$ to

$$(\boldsymbol{L}\cdot\boldsymbol{r}_1)(s,t) = \boldsymbol{f}_1(s,t) \tag{8.8a}$$

where the vector-valued partial differential operator $\boldsymbol{L}$ is defined by

$$\begin{aligned}(\boldsymbol{L}\cdot\boldsymbol{r})(s,t) &:= (\rho A)(s)\boldsymbol{r}_{tt}(s,t)\\ &- \frac{\partial}{\partial s}\left\{\frac{N^0(s)}{z'(s)}[\boldsymbol{r}_s(s,t)\cdot\boldsymbol{i}\,\boldsymbol{i} + \boldsymbol{r}_s(s,t)\cdot\boldsymbol{j}\,\boldsymbol{j}] + N^0_\nu(s)\boldsymbol{r}_s(s,t)\cdot\boldsymbol{k}\,\boldsymbol{k}\right\}\end{aligned} \tag{8.8b}$$

with

$$N^0(s) := \hat{N}(z'(s),s), \quad N^0_\nu(s) := \hat{N}_\nu(z'(s),s). \tag{8.9}$$

We use an analogous notation for higher derivatives. Note that the components of (8.8) in the $\boldsymbol{i}$-, $\boldsymbol{j}$-, and $\boldsymbol{k}$-directions uncouple into three scalar wave equations.

$\boldsymbol{r}_1$ must satisfy the boundary conditions

$$\boldsymbol{r}_1(0,t) = \boldsymbol{o}, \quad \boldsymbol{r}_1(1,t) = \boldsymbol{o} \tag{8.10}$$

and the initial conditions

$$\boldsymbol{r}_1(s,0) = \boldsymbol{u}_1(s), \quad \partial_t\boldsymbol{r}_1(s,0) = \boldsymbol{v}_1(s). \tag{8.11}$$

The component $\boldsymbol{r}_1\cdot\boldsymbol{k}$ satisfies the following wave equation obtained by dotting (8.8) with $\boldsymbol{k}$:

$$(\rho A)(s)w_{tt}(s,t) - [N^0_\nu(s)w_s(s,t)]_s = \boldsymbol{f}_1(s,t)\cdot\boldsymbol{k}. \tag{8.12}$$

We can simplify the equations for the other two components of (8.8) by introducing the change of variable

$$\zeta = z(s) \quad \text{or, equivalently,} \quad s = \tilde{s}(\zeta) \tag{8.13}$$

where $\tilde{s}$ is the inverse of z, which exists by virtue of the positivity of z'. We set

$$\tilde{\boldsymbol{r}}_1(\zeta,t) := \boldsymbol{r}_1(\tilde{s}(\zeta),t), \quad \widetilde{\rho A}(\zeta) := \frac{(\rho A)(\tilde{s}(\zeta))}{z'(\tilde{s}(\zeta))}. \tag{8.14}$$

Then $\tilde{\boldsymbol{r}}_1 \cdot \boldsymbol{i}$ and $\tilde{\boldsymbol{r}}_1 \cdot \boldsymbol{j}$ satisfy

$$\widetilde{\rho A}(\zeta)u_{tt}(\zeta,t) - [N^0(\tilde{s}(\zeta))u_\zeta(\zeta,t)]_\zeta = h(\zeta,t) \tag{8.15}$$

where $h(\zeta,t)$ respectively equals $\boldsymbol{f}_1(\tilde{s}(\zeta),t)\cdot\boldsymbol{i}$ and $\boldsymbol{f}_1(\tilde{s}(\zeta),t)\cdot\boldsymbol{j}$. Note that (6.8) and (8.9) imply that

$$N^0(\tilde{s}(\zeta)) = K + G(\tilde{s}(\zeta)) \tag{8.16}$$

where K satisfies (6.12). $\widetilde{\rho A}$ is the mass per unit length in the configuration $z\boldsymbol{k}$. The change of variables (8.13) and (8.14) is tantamount to taking the stretched equilibrium configuration $z\boldsymbol{k}$ as the reference configuration.

Equation (8.12) describes the small longitudinal motion of a string (or of a rod) about its straight stretched equilibrium state. The nonuniformity of the string and the presence of G cause the coefficients of (8.12) to depend on s. Condition (2.17d) ensures that N_ν^0 is positive and that (8.12) is consequently hyperbolic.

Equation (8.15) describes the small transverse vibrations of the string. If $G=0$ and if the string is uniform, then $\widetilde{\rho A}$ is constant. $N^0(\tilde{s}(\zeta))$ is the tension at $\tilde{s}(\zeta)$ in the configuration $z\boldsymbol{k}$. If $G=0$, Eq. (8.16) implies that this tension is constant whether or not the string is uniform. Under the hypotheses (2.17a) and (2.20), Eq. (6.13) implies that for $G=0$ this constant tension is positive if and only if $L>1$. Where this tension (constant or not) is positive, (8.15) is hyperbolic, and where the tension is negative, (8.15) is elliptic. In the latter case, expected physical instabilities are reflected by the ill-posedness of initial-boundary-value problems for (8.15). Analogous statements apply to the full nonlinear system. By endowing the string with resistance to bending and twisting (i.e., by replacing the string theory with a rod theory), we remove this ill-posedness at the cost of enlarging the system. See Chaps. 4 and 8.

8.17. Exercise. Find the linearized equations (satisfied by $\boldsymbol{r}_1$) for the motion of a viscoelastic string satisfying (2.14). Classify the equations as to type.

8.18. Exercise. Suppose that $L>1$ and that $\boldsymbol{v}=\boldsymbol{o}$ and $\boldsymbol{f}=\boldsymbol{o}$. Find equations for $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{r}_3$ for the perturbation solution for (2.9), (2.10b), (2.11), (2.3), (2.4). If $\boldsymbol{u}_1\cdot\boldsymbol{k}=0$, how do these equations illuminate the role of purely transverse motions discussed in Sec. 7?

We now turn to the more interesting problem of determining the properties of free time-periodic motions of an elastic string near the straight equilibrium state. We seek motions satisfying (8.5) with $\boldsymbol{f}(s,t)=g(s)\boldsymbol{k}$, satisfying (2.3), and having an as yet undetermined period $2\pi/\sqrt{\lambda}$ with $\lambda>0$ so that

$$\boldsymbol{r}(s,t+2\pi/\sqrt{\lambda}) = \boldsymbol{r}(s,t). \tag{8.19}$$

Let us set $\bar{t}=\sqrt{\lambda}t$, $\bar{\boldsymbol{r}}(s,\bar{t})=\boldsymbol{r}(s,\bar{t}/\sqrt{\lambda})$, introduce these variables into the governing equations, and then omit the superposed bars. In this case, (8.5) is modified by having λ precede ρA. Equation (8.19) reduces to

$$\boldsymbol{r}(s,t+2\pi) = \boldsymbol{r}(s,t). \tag{8.20}$$

Note that a small parameter ε is not supplied in this problem. For the moment, we may think of it as an amplitude characterizing the departure of time-periodic solutions from the trivial straight equilibrium state.

The following exercise shows that we cannot attack this problem of periodic solutions by blindly following the approach used for the initial-boundary-value problem.

8.21. Exercise. Substitute (8.2) into the problem for time-periodic solutions. Find the frequencies λ for which the problem for $\boldsymbol{r}_1$ has solutions of period 2π in time. Show that if the corrections $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ have the same frequencies, then $\hat{N}$ is subjected to unduly severe restrictions.

We circumvent this difficulty by allowing λ also to depend on ε. The dependence of frequency on the amplitude thereby permitted is a typical physically important manifestation of nonlinearity. We accordingly supplement (8.2) with

$$\lambda(\varepsilon) = \omega^2 + \sum_{k=1}^{p} \frac{\varepsilon^k}{k!}\lambda_k + o(\varepsilon^{p+1}). \tag{8.22}$$

Equations (8.2) and (8.22) give a parametric representation (i.e., a curve) for the configuration and the frequency in a neighborhood of the trivial state. We obtain the equations satisfied by $\boldsymbol{r}_k$ and λ_{k-1} by substituting (8.2) and (8.22) into the governing equations, differentiating them k times with respect to ε, and then setting $\varepsilon = 0$. We find that $\boldsymbol{r}_k$ satisfies the boundary conditions and periodicity conditions

$$\boldsymbol{r}_k(0,t) = \boldsymbol{o} = \boldsymbol{r}_k(1,t), \quad \boldsymbol{r}_k(s, t+2\pi) = \boldsymbol{r}_k(s,t). \tag{8.23a,b}$$

$\boldsymbol{r}_1$ satisfies

$$\boldsymbol{L}(\omega^2)\cdot \boldsymbol{r}_1 = \boldsymbol{o} \tag{8.24}$$

where $\boldsymbol{L}(\lambda)$ is defined by (8.8b) with ρA replaced with $\lambda\rho A$. We assume that system (8.24) is hyperbolic, i.e., we assume that N^0 is everywhere positive. We can solve (8.23), (8.24) for $\boldsymbol{r}_1$ by separation of variables. We find that nontrivial solutions $\boldsymbol{r}_1$ have the form

$$\boldsymbol{r}_1(s,t) = \begin{cases} u_l(s)[\boldsymbol{a}_{lm}\cos mt + \boldsymbol{b}_{lm}\sin mt] & \text{when} \quad \omega^2 = \sigma_l^2/m^2, \\ w_l(s)[\alpha_{lm}\cos mt + \beta_{lm}\sin mt]\boldsymbol{k} & \text{when} \quad \omega^2 = \tau_l^2/m^2, \end{cases} \tag{8.25a,b}$$
$$l = 0,1,2,\dots, \quad m = 1,2,\dots,$$

where the $\{\boldsymbol{a}_{lm}\}$ and the $\{\boldsymbol{b}_{lm}\}$ are arbitrary vectors in $\operatorname{span}\{\boldsymbol{i},\boldsymbol{j}\}$, where the $\{\sigma_l^2\}$ are the eigenvalues and $\{u_l\}$ are the corresponding eigenfunctions of the Sturm-Liouville problem

$$\frac{d}{ds}\left[\frac{N^0(s)u'}{z'(s)}\right] + \sigma^2(\rho A)(s)u = 0, \quad u(0) = 0 = u(1), \tag{8.26}$$

where the $\{\alpha_{lm}\}$ and the $\{\beta_{lm}\}$ are arbitrary real numbers, and where the $\{\tau_l^2\}$ are the eigenvalues and the $\{w_l\}$ are the corresponding eigenfunctions of the Sturm-Liouville problem

$$\frac{d}{ds}\left[N_\nu^0(s)w'\right] + \tau^2(\rho A)(s)w = 0, \quad w(0) = 0 = w(1). \tag{8.27}$$

We normalize $\{u_l\}$ by the requirement that

$$\int_0^1 (\rho A)(s)u_l(s)u_n(s)\,ds = \delta_{ln} := \begin{cases} 1 & \text{if} \quad l = n, \\ 0 & \text{if} \quad l \neq n, \end{cases}, \quad u_l'(0) > 0 \tag{8.28}$$

and adopt the same conditions for $\{w_l\}$. δ_{ln} is the *Kronecker delta*. The positivity everywhere of N^0 and z' ensures that $0 < \sigma_0^2 < \sigma_1^2 < \cdots$ and that $\sigma_l^2 \to \infty$ as $l \to \infty$ by the Sturm-Liouville theory (see Coddington & Levinson (1955) or Ince (1926), e.g.). The positivity everywhere of N_ν^0 ensures that $\{\tau_l^2\}$ has the same properties. The representations of (8.25a,b) respectively correspond to transverse and longitudinal motions. Since (8.23) with $k = 1$ and (8.24) are invariant under translations of time and under rotations about the $\boldsymbol{k}$-axis, we could without loss of generality impose two restrictions the four components of $\boldsymbol{a}_{lm}$ and $\boldsymbol{b}_{lm}$ appearing in (8.25a).

Note that for fixed $\omega^2 = \sigma_l^2/m^2$, there are as many different solutions of the form (8.25) as there are distinct pairs (j, p) of integers, $j = 1, 2, \ldots, p = 0, 1, 2, \ldots$, satisfying

(8.29a,b) $$\frac{\sigma_l^2}{m^2} = \frac{\sigma_j^2}{p^2} \quad \text{or} \quad \frac{\sigma_l^2}{m^2} = \frac{\tau_j^2}{p^2}.$$

(Note that the special condition that $\hat{N}(\cdot, s)$ be linear, discussed in Sec. 7, ensures that $\sigma_l^2 = \tau_l^2$ for all l.) If N^0/z' and ρA are constant, then there are infinitely many pairs (j, p) satisfying (8.29a).

Suppose that $\omega^2 = \sigma_l^2/m^2$ and that there are no pairs of integers (j, p) such that (8.29b) holds. In this case we find that $\boldsymbol{r}_1 \cdot \boldsymbol{k} = 0$. Then the perturbation procedure yields

(8.30) $$\boldsymbol{L}(\sigma_l^2/m^2) \cdot \boldsymbol{r}_2 = 2\lambda_1 \rho A \partial_{tt} \boldsymbol{r}_1 + \partial_s \left\{ \left[\frac{N^0 - z' N_\nu^0}{(z')^2} \right] \partial_s \boldsymbol{r}_1 \cdot \partial_s \boldsymbol{r}_1 \right\} \boldsymbol{k}.$$

Before blindly lurching toward a solution of (8.30), it is useful to take a preliminary step that can greatly simplify the analysis: We take the dot product of (8.30) with $\boldsymbol{r}_1$ of (8.25) and then integrate the resulting expression by parts twice over $[0, 1] \times [0, 2\pi]$. Since $\boldsymbol{r}_1$ satisfies the homogeneous equation and since $\boldsymbol{r}_1$ has no $\boldsymbol{k}$-component, the resulting equation reduces to

(8.31) $$\lambda_1 = 0.$$

Thus the equations for the $\boldsymbol{i}$- and $\boldsymbol{j}$-components of $\boldsymbol{r}_2$ are exactly the same as those for these components of $\boldsymbol{r}_1$. It follows that the contribution of these terms of $\boldsymbol{r}_2$ to (8.2) has exactly the same form as the corresponding components of $\boldsymbol{r}_1$, but with a coefficient of $\varepsilon^2/2$ in place of ε. We accordingly absorb $\boldsymbol{r}_2$ into $\boldsymbol{r}_1$ by taking

(8.32) $$\boldsymbol{r}_2 \cdot \boldsymbol{i} = 0 = \boldsymbol{r}_2 \cdot \boldsymbol{j},$$

In Sec. **5**.6 we describe a more systematic way to get relations like (8.32).

In view of (8.31) and (8.32), problem (8.30) reduces to a nonhomogeneous linear equation for $\boldsymbol{r}_2 \cdot \boldsymbol{k}$:

(8.33) $$\begin{aligned} &\frac{\sigma_l^2}{m^2} \rho A \partial_{tt} \boldsymbol{r}_2 \cdot \boldsymbol{k} - \partial_s [N_\nu^0 \partial_s \boldsymbol{r}_2 \cdot \boldsymbol{k}] \\ &= \tfrac{1}{2} \partial_s \left[\frac{N^0 - z' N_\nu^0}{(z')^2} (u_l')^2 \right] \left[|\boldsymbol{a}_{lm}|^2 + |\boldsymbol{b}_{lm}|^2 + \left(|\boldsymbol{a}_{lm}|^2 - |\boldsymbol{b}_{lm}|^2 \right) \cos 2mt \right. \\ &\qquad \left. + \boldsymbol{a}_{lm} \cdot \boldsymbol{b}_{lm} \sin 2mt \right]. \end{aligned}$$

Since we know that the homogeneous problem for (8.33) has only the trivial solution, we can seek a solution in the form $f(s) + g(s) \cos 2mt + h(s) \sin 2mt$ and obtain boundary-value problems for f, g, and h like (8.27). The solutions of these boundary-value problems can be represented in terms of a Green function associated with the operator of (8.27) or alternatively by an expansion in terms of the eigenfunctions associated with (8.27). We can also represent the solution of (8.33) directly as an eigenfunction expansion with respect to the basis

(8.34) $$\{(s, t) \mapsto \tfrac{1}{\pi} w_q(s) \cos nt,\ \tfrac{1}{\pi} w_q(s) \sin nt, \quad q = 0, 1, \ldots,\ n = 1, 2, \ldots\}.$$

We find the Fourier coefficients of $\boldsymbol{k}\cdot\boldsymbol{r}_2$ by multiplying (8.33) by a member of (8.34) and integrating the resulting equation by parts over $[0,1]\times[0,2\pi]$. (See Stakgold (1998), e.g.) We get

$$\text{(8.35a)}\quad \boldsymbol{k}\cdot\boldsymbol{r}_2(s,t) = \frac{m^2}{2\pi}\sum_{q=0}^{\infty}\mu_{lq}w_q(s)\left[(|\boldsymbol{a}_{lm}|^2-|\boldsymbol{b}_{lm}|^2)\cos 2mt+(\boldsymbol{a}_{lm}\cdot\boldsymbol{b}_{lm})\sin 2mt\right],$$

$$\text{(8.35b)}\quad \left(\tfrac{1}{4}\tau_q^2-\sigma_l^2\right)\mu_{lq} := \int_0^1 \partial_s\left\{\frac{N^0(s)-z'(s)N_\nu^0(s)}{z'(s)^2}u_l'(s)^2\right\}w_l(s)\,ds$$

when (8.25a) holds and when $\tau_q^2\neq 4\sigma_l^2$ for each q. The properties of $\{\tau_q^2\}$ developed in Sturm-Liouville theory ensure that (8.35a) converges. Equation (8.35) shows that the first correction to the purely transverse linear motion is a longitudinal motion.

Using (8.25), (8.31), and (8.32) we find that

$$\text{(8.36)}\quad \boldsymbol{L}(\sigma_l^2/m^2)\cdot\boldsymbol{r}_3 = 3\lambda_2\rho A\partial_{tt}\boldsymbol{r}_1+3\partial_s\left\{(N^0-z'N_\nu^0)\frac{[\partial_s\boldsymbol{r}_1\cdot\partial_s\boldsymbol{r}_1+(\boldsymbol{k}\cdot\partial_s\boldsymbol{r}_2)z']\partial_s\boldsymbol{r}_1}{(z')^3}\right\}.$$

We treat this equation just like (8.30): We dot it with $\boldsymbol{r}_1$ and integrate the resulting equation by parts over $[0,1]\times[0,2\pi]$ to get

$$\text{(8.37)}\quad m^2\pi(|\boldsymbol{a}_{lm}|^2+|\boldsymbol{b}_{lm}|^2)\lambda_2 = \int_0^1\int_0^{2\pi}\partial_s\left\{(N^0-z'N_\nu^0)\frac{[\partial_s\boldsymbol{r}_1\cdot\partial_s\boldsymbol{r}_1+(\boldsymbol{k}\cdot\partial_s\boldsymbol{r}_2)z']\partial_s\boldsymbol{r}_1}{(z')^3}\right\}\cdot\partial_s\boldsymbol{r}_1\,dt\,ds.$$

In view of (8.22) the sign of this expression for λ_2 gives the important physical information of whether the frequency λ increases or decreases with the amplitude ε of the motion. Note how λ_2 depends crucially on the behavior of the constitutive function $\hat{N}$. The procedures we have used in this analysis are quite general.

8.38. Exercise. Obtain an explicit representation for λ_2 when the material is uniform, so that $\hat{N}_s = 0$ and z' and ρA are constant.

A computation analogous to that leading to (8.37) can be carried out for the purely longitudinal motion. But the results are purely formal because it can be shown that no purely longitudinal periodic motion is possible (see Keller & Ting (1966) and Lax (1964)). The solutions must exhibit shocks. For the transverse motions (which have a longitudinal component as we have seen), periodic solutions are possible. (The energy could be shifted about and avoid being concentrated. We give a transparent example of such a phenomenon in Sec. **14**.15.) This possibility of shocks makes it hard to justify the method and to interpret the results. The formal results clearly say something important about the nonlinear system, but it is difficult to give a mathematically precise and physically illuminating explanation of exactly what is being said. In other words, it is not clear what the linear wave equations say about solutions of the nonlinear equations. By introducing a strong dissipative mechanism, corresponding to (2.14) subject to (2.22b), it is likely that we could prevent our equations from having shocks. But this dissipation would prevent periodic solutions unless we introduced periodic forcing. The resulting perturbation scheme would be more complicated, but there is some hope that the approach could be justified. It is physically attractive but notoriously difficult to study the undamped system by taking the limit as the dissipation goes to zero. We comment on related questions at the end of Sec. 11.

Carrier (1945, 1949) used such perturbation methods to study periodic planar vibrations of an elastic string for which $\hat{N}(\cdot,s)$ is taken to be affine, although this restriction is inessential. The work of this section is largely based on Keller & Ting (1966) and J. B. Keller (1968). For other applications of this formalism, see Millman & Keller (1969) and Iooss & Joseph (1990).

9. The Justification of Perturbation Methods

In this section we give precise conditions justifying perturbation methods for static problems. The fundamental mathematical tool for our analysis is the Implicit-Function Theorem in different manifestations. References on the justification of perturbation methods for dynamical problems are given at the end of this section. Our basic result is

9.1. Theorem. *Let z be as in Sec. 6. Let p be a positive integer. If $\hat{N}$ is continuous, if $\hat{N}(\cdot, s) \in C^{p+1}(0, \infty)$, if $\hat{N}_\nu(z'(s), s) \equiv N^0_\nu(s) > 0$ and $\hat{N}(z'(s), s) \equiv N^0(s) > 0$ for each s, if $g \in C^0[0,1]$, and if $\boldsymbol{f}_1 \in C^0[0,1]$, then there is a number $\eta > 0$ such that for $|\varepsilon| < \eta$ the boundary-value problem*

$$\frac{d}{ds}\left[\hat{N}\big(|\boldsymbol{r}'(s)|, s\big)\frac{\boldsymbol{r}'(s)}{|\boldsymbol{r}'(s)|}\right] + g(s)\boldsymbol{k} + \varepsilon \boldsymbol{f}_1(s) = \boldsymbol{o}, \tag{9.2}$$

$$\boldsymbol{r}(0) = \boldsymbol{o}, \quad \boldsymbol{r}(1) = L\boldsymbol{k} \tag{9.3a,b}$$

has a unique solution $\boldsymbol{r}(\cdot, \varepsilon)$ with $\boldsymbol{r}(\cdot, \varepsilon) \in C^2[0,1]$ and $\boldsymbol{r}(s, \cdot) \in C^{p+1}(-\eta, \eta)$. (Thus $\boldsymbol{r}(s, \varepsilon)$ admits an expansion like (8.2).)

Proof. From (6.5) with $a = 0$ and from (8.4) we get

$$\hat{\boldsymbol{n}}\big(\boldsymbol{r}'(s), s\big) - \hat{\boldsymbol{n}}\big(\boldsymbol{r}'(0), 0\big) - G(s)\boldsymbol{k} + \varepsilon \int_0^s \boldsymbol{f}_1(\xi)\, d\xi = \boldsymbol{o}, \tag{9.4}$$

which can be obtained from the integration of (9.2). From (8.6) we obtain

$$\boldsymbol{c} \cdot \hat{\boldsymbol{n}}_{\boldsymbol{q}}(z'(s)\boldsymbol{k}, s) \cdot \boldsymbol{c} = N^0_\nu(s)(\boldsymbol{k} \cdot \boldsymbol{c})^2 + \frac{N^0(s)}{z'(s)}\left[\boldsymbol{c} \cdot \boldsymbol{c} - (\boldsymbol{k} \cdot \boldsymbol{c})^2\right]. \tag{9.5}$$

Thus $\hat{\boldsymbol{n}}_{\boldsymbol{q}}(z'(s)\boldsymbol{k}, s)$ is positive-definite and therefore nonsingular. The classical Implicit-Function Theorem thus implies that for $\boldsymbol{q}$ near $z'(s)\boldsymbol{k}$, the function $\boldsymbol{q} \mapsto \hat{\boldsymbol{n}}(\boldsymbol{q}, s)$ has an inverse, which we denote by $\boldsymbol{n} \mapsto \boldsymbol{m}(\boldsymbol{n}, s)$. We use it to solve (9.4) for $\boldsymbol{r}'(s)$. We integrate the resulting equation from 0 to s subject to (9.3a) to obtain

$$\boldsymbol{r}(s) = \int_0^s \boldsymbol{m}\left(\hat{\boldsymbol{n}}(\boldsymbol{r}'(0), 0) + G(\xi)\boldsymbol{k} - \varepsilon \int_0^\xi \boldsymbol{f}_1(\sigma)\, d\sigma,\ \xi\right) d\xi. \tag{9.6}$$

The requirement that (9.6) satisfy (9.3b) yields

$$\begin{aligned} &\boldsymbol{l}(\boldsymbol{r}'(0), \varepsilon) \\ &\quad := \int_0^1 \boldsymbol{m}\left(\hat{\boldsymbol{n}}(\boldsymbol{r}'(0), 0) + G(\xi)\boldsymbol{k} - \varepsilon \int_0^\xi \boldsymbol{f}_1(\sigma)\, d\sigma,\ \xi\right) d\xi = L\boldsymbol{k}. \end{aligned} \tag{9.7}$$

If there is a unique solution $\boldsymbol{r}'(0) = \boldsymbol{p}(\varepsilon)$ of this equation, then its substitution for $\boldsymbol{r}'(0)$ in (9.6) yields the solution $\boldsymbol{r}(\cdot, \varepsilon)$ of (9.2), (9.3). We now verify that (9.7) meets the hypotheses of the classical Implicit-Function Theorem: First of all, we must show that

$$\boldsymbol{l}(z'(0)\boldsymbol{k}, 0) = L\boldsymbol{k}. \tag{9.8}$$

But this is equivalent to (6.12) because (8.4) and (6.8) imply that

$$\begin{aligned}\hat{\boldsymbol{n}}\,(z'(0)\boldsymbol{k}, 0) + G(\xi)\boldsymbol{k} &= [N(z'(0), 0) + G(\xi)]\boldsymbol{k} = [K + G(\xi)]\boldsymbol{k} \\ &= \hat{N}(z'(\xi), \xi)\boldsymbol{k} = \hat{\boldsymbol{n}}\,(z'(\xi)\boldsymbol{k}, \xi)\,.\end{aligned} \tag{9.9a}$$

Next, (9.7) implies that

$$\boldsymbol{l}_{\boldsymbol{p}}(\boldsymbol{p}, 0) = \int_0^1 \boldsymbol{m}_{\boldsymbol{n}}\big(\hat{\boldsymbol{n}}(\boldsymbol{p}, 0) + G(\xi)\boldsymbol{k},\, \xi\big) \cdot \hat{\boldsymbol{n}}_{\boldsymbol{q}}(\boldsymbol{p}, 0)\, d\xi. \tag{9.9b}$$

From (9.9a) we find that

$$\boldsymbol{m}_{\boldsymbol{n}}\big(\hat{\boldsymbol{n}}(z'(0)\boldsymbol{k}, 0) + G(\xi)\boldsymbol{k},\, \xi\big) = \boldsymbol{m}_{\boldsymbol{n}}\big(\hat{N}(z'(\xi), \xi)\boldsymbol{k},\, \xi\big), \tag{9.10}$$

so that (9.10) is the inverse of the symmetric positive-definite tensor $\hat{\boldsymbol{n}}_{\boldsymbol{q}}(z'(\xi)\boldsymbol{k}, \xi)$. It follows that (9.10) is positive-definite. Since the product of two symmetric positive-definite tensors is positive-definite (though not necessarily symmetric), we find that

$$\boldsymbol{l}_{\boldsymbol{p}}(z'(0)\boldsymbol{k}, 0) \quad \text{is nonsingular.} \tag{9.11}$$

Conditions (9.8) and (9.11) are the requisite hypotheses for the classical Implicit-Function Theorem, which says that there is a number $\eta > 0$ such that (9.7) has a unique solution $\boldsymbol{p}(\varepsilon)$ for $|\varepsilon| < \eta$ and that $\boldsymbol{p}(\cdot) \in C^{p+1}(-\eta, \eta)$. It then follows from (9.5) that $\boldsymbol{r}(s, \cdot)$ itself is in this space. The regularity of $\boldsymbol{r}(\cdot, \varepsilon)$ can be read off from (9.6). (It is correspondingly enhanced for increased smoothness of g and $\boldsymbol{f}_1$.) $\square$

Note that this theorem is purely local in the sense that it gives information about solutions of the nonlinear problem (9.2), (9.3) only in a neighborhood of a known solution. In contrast, the elementary analysis of Sec. 6 is global. In Chap. 3 we shall give global analyses of equilibrium states of strings under several more interesting force systems.

In this proof we have avoided the use of determinants. They are not suitable for proving (9.11) because $\boldsymbol{l}_{\boldsymbol{p}}(z'(0)\boldsymbol{k}, 0)$ is an integral. If an integrand is a positive-definite tensor everywhere, then its integral is likewise, but if an integrand is merely nonsingular everywhere, then its integral need not be nonsingular.

9.12. Exercise. Prove the last assertion about nonsingular tensors.

The proof of Theorem 9.1 relied on the special nature of (9.2). If $\boldsymbol{f}_1$, say, were to depend upon $\boldsymbol{r}$, then (9.5) would be an integral equation for $\boldsymbol{r}$ and would require a subtler analysis. Procedures for such analyses have been systematized, the most comprehensive methods employing an abstract version of the local Implicit-Function Theorem **20**.1.27 in Banach Space, which is applied in several places in this book. Here we present a related concrete approach, the *Poincaré shooting method,* applicable to systems of ordinary differential equations (more complicated than (9.2)).

Proof of Theorem 9.1 by the Poincaré shooting method. We seek a vector $\boldsymbol{a}$ such that (9.2) subject to the *initial conditions*

$$\boldsymbol{r}(0) = \boldsymbol{o}, \quad \boldsymbol{r}'(0) = \boldsymbol{a} \tag{9.13}$$

has a solution satisfying (9.3b). To apply the basic theory of ordinary differential equations to this problem, it is convenient to write (9.2) as a first-order system in which the derivatives of the unknowns are expressed as functions of the unknowns. This reduction can be effected in two ways: For $\boldsymbol{r}'$ close enough to $z'\boldsymbol{k}$, we can use the tools developed in the above proof of Theorem 9.1 to write (9.2), (9.3) as

$$\boldsymbol{n}' = -g(s)\boldsymbol{k} - \varepsilon\boldsymbol{f}_1(s), \quad \boldsymbol{r}' = \boldsymbol{m}(\boldsymbol{n}, s), \quad \boldsymbol{r}(0) = \boldsymbol{o}, \quad \boldsymbol{n}(0) = \boldsymbol{b} := \hat{\boldsymbol{n}}(\boldsymbol{a}, 0). \tag{9.14}$$

Alternatively, we could carry out the differentiation in (9.2). For $\boldsymbol{r}'$ close enough to $z'\boldsymbol{k}$, we can solve (9.2) for $\boldsymbol{r}''$, obtaining an equation of the form $\boldsymbol{r}'' = \boldsymbol{h}(\boldsymbol{r}', s, \varepsilon)$. We set $\boldsymbol{v} = \boldsymbol{r}'$ and thereby convert this second-order system to the equivalent first-order system

$$\boldsymbol{v}' = \boldsymbol{h}(\boldsymbol{v}, s, \varepsilon), \quad \boldsymbol{r}' = \boldsymbol{v}. \tag{9.15}$$

To be specific, we limit our attention to (9.14). Since the results of Sec. 6 imply that it has a unique solution $\boldsymbol{r} = z\boldsymbol{k}$, $\boldsymbol{n} = \hat{\boldsymbol{n}}(z'(\cdot)\boldsymbol{k}, \cdot))$ for $\varepsilon = 0$ and $\boldsymbol{b} = \hat{\boldsymbol{n}}(z'(0)\boldsymbol{k}, 0))$, the basic theory of ordinary differential equations (see Coddington & Levinson (1955, Chaps. 1,2) or Hale (1969, Chap. 1), e.g.) implies that (9.14) has a unique solution $\boldsymbol{r}(\cdot, \boldsymbol{b}, \varepsilon)$ defined on the whole interval $[0, 1]$ if ε and $\boldsymbol{b}$ are close enough to 0 and $\hat{\boldsymbol{n}}(z'(0)\boldsymbol{k}, 0))$. Moreover, $\boldsymbol{r}(s, \cdot, \cdot)$ is $(p+1)$-times continuously differentiable. $\boldsymbol{r}(\cdot, \boldsymbol{b}, \varepsilon)$ would correspond to a solution of (9.2), (9.3) for small nonzero ε if $\boldsymbol{b}$ can be chosen so that

$$\boldsymbol{r}(1, \boldsymbol{b}, \varepsilon) = L\boldsymbol{k}. \tag{9.16}$$

We know that this system for $\boldsymbol{b}$ has the solution $\boldsymbol{b}_0 := \hat{\boldsymbol{n}}(z'(0)\boldsymbol{k}, 0))$ for $\varepsilon = 0$. The Implicit-Function Theorem then implies that there is a number $\eta > 0$ such that (9.16) has a unique solution $(-\eta, \eta) \ni \varepsilon \mapsto \hat{\boldsymbol{b}}(\varepsilon)$ with $\hat{\boldsymbol{b}} \in C^{p+1}(-\eta, \eta)$ and with $\hat{\boldsymbol{b}}(0) = \boldsymbol{b}_0$ provided that

$$\det \boldsymbol{R}(1) \neq 0, \quad \boldsymbol{R}(s) := \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{b}}(s, \boldsymbol{b}_0, 0). \tag{9.17}$$

The theory of ordinary differential equations implies that the matrix $\boldsymbol{R}$ satisfies the initial-value problem obtained by formally differentiating (9.14) with respect to $\boldsymbol{b}$ and then setting $(\boldsymbol{b}, \varepsilon) = (\boldsymbol{b}_0, 0)$. This process yields

$$\boldsymbol{n}_{\boldsymbol{b}}' = \boldsymbol{O}, \quad \boldsymbol{r}_{\boldsymbol{b}}' = \boldsymbol{m}_{\boldsymbol{n}}\big(\hat{\boldsymbol{n}}(z'(s)\boldsymbol{k}, s)\big) \cdot \boldsymbol{n}_{\boldsymbol{b}}, \quad , \boldsymbol{n}_{\boldsymbol{b}}(0) = \boldsymbol{I}, \quad \boldsymbol{r}_{\boldsymbol{b}}(0) = \boldsymbol{O}, \tag{9.18}$$

whence we obtain

$$\boldsymbol{R}' = \boldsymbol{m}_{\boldsymbol{n}}\big(\hat{\boldsymbol{n}}(z'(s)\boldsymbol{k}, s)\big), \quad \boldsymbol{R}(0) = \boldsymbol{O}. \tag{9.19}$$

We obtain $\boldsymbol{R}(1)$ by integrating (9.19). It is positive-definite because (9.10) is. □

9.20. Problem. Investigate the validity of the perturbation process when N^0 is not everywhere positive.

Chap. 20 contains proofs of a general version of the Implicit-Function Theorem, of the basic existence and uniqueness theorem for initial-value problems for ordinary differential equations in Ex. (**20**.1.22), and of the Poincaré Shooting Method. These proofs are each based upon the Contraction Mapping Principle. Thus both methods discussed above are intimately connected. Methods for justifying perturbation methods for dynamical problems have so far required that the equations have a strong dissipative mechanism. See Koch & Antman (2001), Potier-Ferry (1981,1982), Xu & Marsden (1996).

10. Variational Characterization of the Equations for an Elastic String

If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then the equation

$$f(x) = 0 \tag{10.1}$$

is equivalent to

$$\phi'(x) = 0 \tag{10.2}$$

where

$$\phi(x) = \int_0^x f(\xi)\, d\xi. \tag{10.3}$$

Thus we might be able to study the existence of solutions of (10.1) by showing that ϕ has an extremum (a maximum or a minimum) on $\mathbb{R}$. If ϕ is merely continuous, we can still study the minimization of ϕ, although the corresponding problem (10.1) for f need not be meaningful. (The present situation in 3-dimensional nonlinear elastostatics has precisely this character: Under certain conditions the total energy is known to have a minimizer, but it is not known whether the equilibrium equations, which correspond to the vanishing of the Gâteaux derivative of the energy, have solutions. See Chap. 13.)

If $\mathsf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is continuous, then the system

$$\mathsf{f}(\mathsf{x}) = \mathsf{o} \tag{10.4}$$

may not be equivalent to the vanishing of a gradient

$$\partial\phi(\mathsf{x})/\partial\mathsf{x} = \mathsf{o} \tag{10.5}$$

because there may not be a scalar-valued function ϕ such that $\partial\phi/\partial\mathsf{x} = \mathsf{f}$. If $\mathsf{f} \in C^1(\mathbb{R}^n)$, then a necessary and sufficient condition for the existence of such a ϕ is that:

$$\partial\mathsf{f}/\partial\mathsf{x} \quad \text{is symmetric.} \tag{10.6}$$

In this case, ϕ is defined by the line integral

$$\phi(\mathsf{x}) := \int_{\mathcal{C}} \mathsf{f}(\mathsf{y}) \cdot \, d\mathsf{y} \tag{10.7}$$

where $\mathcal{C}$ is a sufficiently smooth curve joining a fixed point to x. (For a proof, see Sec. **12**.3.) We could then study (10.4) by studying extrema of ϕ.

In this section we show how the equations of motion for an elastic string can be characterized as the vanishing of the Gâteaux differential of a scalar-valued function of the configuration. For this purpose, we must first extend the notion of directional derivative of a real-valued function defined on some part of $\mathbb{R}^n$ to a real-valued function defined on a set of functions, which are to be candidates for solutions of the governing equations.

Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be normed spaces (e.g., spaces of continuous functions; see Chap. 19) and let $\mathcal{A} \subset \mathcal{E}_1$. Let $f[\,\cdot\,] : \mathcal{E}_1 \to \mathcal{E}_2$. (When the argument of a function lies in a function space, we often enclose this argument with brackets instead of parentheses. Examples of such f's are forthcoming.) If

$$\tfrac{d}{d\varepsilon} f[u + \varepsilon y]|_{\varepsilon=0} \tag{10.8}$$

exists for $u \in \mathcal{A}$, $y \in \mathcal{E}_1$, and $\varepsilon \in \mathbb{R} \setminus \{0\}$, then it is called the *Gâteaux differential, directional derivative*, or *first variation of* of f *at* u *in the direction* y. If (10.8) exists for all y in $\mathcal{E}_1$ (for which it is necessary that u be an interior point of $\mathcal{A}$) and if it is a bounded linear operator acting on y (i.e., if (10.8) is linear in y and if the $\mathcal{E}_2$-norm of (10.8) is less than a constant times the $\mathcal{E}_1$-norm of y), then f is said to be *Gâteaux-differentiable* at u. In this case, (10.8) is denoted by $f_u[u] \cdot y$, and $f_u[u]$ is called the *Gâteaux derivative* of f at u. (This terminology is not completely standardized. See Vaĭnberg (1964) for a comprehensive treatment of the interrelationship of various kinds of differentiations.) If $\mathcal{E}_2 = \mathbb{R}$, then f is called a *functional*.

We ask whether the (weak form of the) governing equations for a string can be characterized by the vanishing of the Gâteaux differential of some suitable functional. We show that this can be done for elastic strings under conservative forces. We study the formulation of the equations for time-periodic motions of such strings; the formulation of initial-boundary-value problems by variational methods proves to be somewhat unnatural.

The *kinetic energy* of the string at time t is

$$K[\boldsymbol{r}](t) := \tfrac{1}{2} \int_0^1 (\rho A)(s) |\boldsymbol{r}_t(s,t)|^2 \, ds. \tag{10.9}$$

The *stored-energy* (or *strain-energy*) (*density*) *function* for an elastic string is the function W defined by

$$W(\nu, s) := \int_1^\nu \hat{N}(\beta, s) \, d\beta. \tag{10.10}$$

The (*total*) *stored energy* in the string at time t is

$$\Psi[\boldsymbol{r}](t) := \int_0^1 W\big(\nu(s,t), s\big) \, ds. \tag{10.11}$$

Suppose that $\boldsymbol{f}$ has the form

$$\boldsymbol{f}(s,t) = \boldsymbol{g}(\boldsymbol{r}(s,t), s) \tag{10.12}$$

and that there is a scalar-valued function ω, called the *potential-energy density* of $\boldsymbol{g}$, such that

$$\boldsymbol{g}(\boldsymbol{r}, s) = -\partial\omega(\boldsymbol{r}, s)/\partial\boldsymbol{r} \tag{10.13}$$

(see the remarks following (10.5)). Thus $\boldsymbol{g}$ is conservative. The *potential energy of the body force* $\boldsymbol{g}$ *at time* t is

$$\Omega[\boldsymbol{r}](t) := \int_0^1 \omega\big(\boldsymbol{r}(s,t), s\big)\, ds. \tag{10.14}$$

The *potential-energy functional for the string* is $\Psi + \Omega$.

Let $\mathcal{E}$ consist of all continuously differentiable vector-valued functions $[0,1] \times \mathbb{R} \ni (s,t) \mapsto \boldsymbol{y}(s,t)$ that satisfy

$$\boldsymbol{y}(0,t) = \boldsymbol{o} = \boldsymbol{y}(1,t), \qquad \boldsymbol{y}(s,\cdot) \text{ has period } T. \tag{10.15}$$

The norm on $\mathcal{E}$ can be taken to be

$$\|\,\boldsymbol{y}\| := \max\{|\boldsymbol{y}(s,t)| + |\boldsymbol{y}_s(s,t)| + |\boldsymbol{y}_t(s,t)| \colon (s,t) \in [0,1] \times [0,T]\}. \tag{10.16}$$

We introduce the *Lagrangian functional* Λ by

$$\Lambda[\boldsymbol{r}] := \int_0^T \{K[\boldsymbol{r}](t) - \Psi[\boldsymbol{r}](t) - \Omega[\boldsymbol{r}](t)\}\, dt. \tag{10.17}$$

We study this functional on the class of *admissible* motions

$$\mathcal{A} := \{\boldsymbol{r} \colon \boldsymbol{r}(s,t) = \boldsymbol{y}(s,t) + Ls\boldsymbol{k},\ \boldsymbol{y} \in \mathcal{E},\ |\boldsymbol{r}_s(s,t)| > 0\}. \tag{10.18}$$

For $\boldsymbol{r} \in \mathcal{A}$ and $\boldsymbol{y} \in \mathcal{E}$, we obtain from (10.9)–(10.14) that

$$\begin{aligned}
\Lambda_{\boldsymbol{r}}[\boldsymbol{r}] \cdot \boldsymbol{y} = \frac{d}{d\varepsilon} \int_0^T \int_0^1 \big[&\tfrac{1}{2}(\rho A)(s) |\boldsymbol{r}_t(s,t) + \varepsilon \boldsymbol{y}_t(s,t)|^2 \\
&- W\big(|\boldsymbol{r}_s(s,t) + \varepsilon \boldsymbol{y}_s(s,t)|, s\big) \\
&- \omega\big(\boldsymbol{r}(s,t) + \varepsilon \boldsymbol{y}(s,t), s\big)\big]\, ds\, dt \Big|_{\varepsilon=0} \\
= \int_0^T \int_0^1 \Big[&(\rho A)(s) \boldsymbol{r}_t(s,t) \cdot \boldsymbol{y}_t(s,t) \\
&- \hat{N}\big(|\boldsymbol{r}_s(s,t)|, s\big) \frac{\boldsymbol{r}_s(s,t) \cdot \boldsymbol{y}_s(s,t)}{|\boldsymbol{r}_s(s,t)|} \\
&+ \boldsymbol{g}\big(\boldsymbol{r}(s,t), s\big) \cdot \boldsymbol{y}(s,t) \Big]\, ds\, dt.
\end{aligned} \tag{10.19}$$

The mild difference between the vanishing of (10.19) and the Principle of Virtual Power (2.24), embodied in the presence of $\boldsymbol{v}$ in (2.24), reflects the fact that (2.24) accounts for initial conditions, whereas (10.19) accounts for periodicity conditions. *Hamilton's Principle* for elastic strings under conservative forces states that (the weak form of) the governing equations can be characterized by the vanishing of the Gâteaux differential of the

Lagrangian functional Λ. Any system of equations that can be characterized by the vanishing of the Gâteaux differential of a functional is said to have a *variational structure*; the equations are called the *Euler-Lagrange equations* for that functional.

Hamilton's principle does not require that variables entering it be periodic in time. In fact, in the mechanics of particles and rigid bodies, the configuration is typically required to satisfy boundary conditions at an initial and terminal time. Such conditions are artificial; they are devised so as to yield the governing equations as Euler-Lagrange equations. On the other hand, periodicity conditions define an important class of problems.

In continuum mechanics, Hamilton's principle is applicable only to frictionless systems acted on solely by conservative forces. A criterion telling whether a system of equations admits a natural variational structure is given by Vaĭnberg (1964) and is exploited by Tonti (1969). (Its derivation is just the generalization to function spaces of that for (10.6).) There is also a theory, akin to the theory of holonomicity in classical mechanics, that tells when a system can be transformed into one having a variational principle. The use of such a theory for quasilinear partial differential equations is very dangerous because the requisite transformations may change the weak form of the equations. For physical systems, the altered form may not be physically correct because it does not conform to the Principle of Virtual Power and accordingly does not deliver the correct jump conditions.

For Hamilton's Principle to be useful, it must deliver something more than an alternative derivation of the governing equations with theological overtones. One way for it to be useful would be for it to promote the proof of existence theorems for solutions characterized as extremizers of Λ. Serious technical difficulties have so far prevented this application to the equations of motion of nonlinear elasticity. Hamilton's Principle, however, has recently proved to be very effective in supporting the demonstration of the existence of multiple periodic solutions of systems of ordinary differential equations (see Ekeland (1990), Mawhin & Willem (1989), Rabinowitz et al. (1987), e.g.). The specialization of Hamilton's Principle to static problems, called the *Principle of Minimum Potential Energy*, is very useful for existence theorems and for the interpretation of the stability of equilibrium states, as we shall see in Chaps. 7 and 13. Moreover, Hamiltonian structure has been effectively exploited to derive stability theorems for certain elastic systems (see Simo, Posbergh, & Marsden (1991), e.g.).

11. Discretization

In this section we briefly survey some numerical methods for solving partial differential equations like those for the string. This text is not directly concerned with numerical methods; we examine these questions here because they are intimately related to the Principle of Virtual Power. (They can also be used to produce constructive existence theorems for certain problems.)

We describe a simple method that leads to the formal approximation of the partial differential equations for an elastic string by a system of ordinary differential equations. This procedure, associated with the names of Bubnov, Galerkin, Faedo, and Kantorovich, is sometimes called the method of lines. A special case of it is the semidiscrete finite-element method.

Let $\{s \mapsto \phi_k(s),\ k = 1, 2, \dots\}$ be a given set of functions in $W_p^1(0,1)$ with the properties that $\phi_k(0) = 0 = \phi_k(1)$ and that given an arbitrary function in $W_p^1(0,1)$ and an error, there exists a finite linear combination of the ϕ_k that approximate the given function to within the assigned error in the W_p^1-norm. (The set $\{s \mapsto \sin k\pi s\}$ has these properties. Another such set is defined in (11.9).) We seek to approximate solutions of the initial-boundary-value problem for elastic strings of Sec. 2 by functions $\boldsymbol{r}^K$ of the form

$$\boldsymbol{r}^K(s,t) = Ls\boldsymbol{k} + \sum_{k=1}^{K} \phi_k(s)\boldsymbol{r}_k(t) \tag{11.1}$$

where the functions $\boldsymbol{r}_k$ are to be determined. We approximate the given initial position $\boldsymbol{u}(s)$ and initial velocity $\boldsymbol{v}(s)$ by

$$\boldsymbol{u}^K(s) = Ls\boldsymbol{k} + \sum_{k=1}^{K} \phi_k(s)\boldsymbol{u}_k, \quad \boldsymbol{v}^K(s) = \sum_{k=1}^{K} \phi_k(s)\boldsymbol{v}_k \tag{11.2a,b}$$

where the constant vectors $\{\boldsymbol{u}_k, \boldsymbol{v}_k\}$ are given. In the Principle of Virtual Power (2.24), (2.10b), (2.11) for elastic strings let us replace $\boldsymbol{r}$ and $\boldsymbol{v}$ with $\boldsymbol{r}^K$ and $\boldsymbol{v}^K$ and let us choose

$$\boldsymbol{y}(s,t) = \phi_l(s)\boldsymbol{y}_l(t) \tag{11.3}$$

where $\boldsymbol{y}_l$ is an arbitrary absolutely continuous function that vanishes for large t. (There is no need for $\boldsymbol{y}_l$ to be indexed with l. No summation is intended on the right-hand side of (11.3).) Then this principle reduces to the following weak formulation of the system of ordinary differential equations for $\{\boldsymbol{r}_k\}$:

$$\begin{aligned}\int_0^\infty \int_0^1 \hat{\boldsymbol{n}}\left(L\boldsymbol{k} + \sum_{k=1}^{K} \phi_k'(s)\boldsymbol{r}_k(t),\, s\right)\cdot \boldsymbol{y}_l(t)\phi_l'(s)\,ds\,dt - \int_0^\infty \boldsymbol{f}_l\cdot\boldsymbol{y}_l\,dt \\ = \sum_{k=1}^{K}\langle\phi_k,\phi_l\rangle \int_0^\infty \left(\frac{d\boldsymbol{r}_k}{dt} - \boldsymbol{v}_k\right)\cdot\frac{d\boldsymbol{y}_l}{dt}\,dt\end{aligned} \tag{11.4}$$

for all absolutely continuous $\boldsymbol{y}_l,\ l = 1,\dots,K$, where

$$\boldsymbol{f}_l(t) := \int_0^1 \boldsymbol{f}(s,t)\phi_l(s)\,ds, \quad \langle\phi_k,\phi_l\rangle := \int_0^1 \rho A\phi_k\phi_l\,ds. \tag{11.5}$$

In consonance with (2.4) we require that $\{\boldsymbol{r}_k\}$ satisfy the initial conditions

$$\boldsymbol{r}_k(0) = \boldsymbol{u}_k, \quad \frac{d\boldsymbol{r}_k}{dt}(0) = \boldsymbol{v}_k, \tag{11.6a,b}$$

the second of which is incorporated into (11.4) as we shall see.

11.7. Exercise. Suppose that (11.4) has a continuously differentiable (or more generally an absolutely continuous) solution $\{\boldsymbol{r}_1,\dots,\boldsymbol{r}_K\}$. Take $\boldsymbol{y}_l(t) = \psi(t)\boldsymbol{e}$ where ψ is defined in (4.4b) and and where $\boldsymbol{e}$ is an arbitrary constant vector. Use the method described at the end of Sec. 4 to prove that $\sum_{k=1}^{K}\langle\phi_k,\phi_l\rangle(\frac{d\boldsymbol{r}_k}{dt} - \boldsymbol{v}_k)$ is continuously differentiable (whichever smoothness hypothesis is made about the solution) and accordingly satisfies (11.6b). Show that the solution of (11.4) is thus a classical solution of the system

$$\sum_{k=1}^{K}\langle\phi_k,\phi_l\rangle\frac{d^2\boldsymbol{r}_k}{dt^2} + \int_0^1 \hat{\boldsymbol{n}}\left(L\boldsymbol{k} + \sum_{k=1}^{K}\phi_l'(s)\boldsymbol{r}_k(t),\, s\right)\phi_k'(s)\,ds - \boldsymbol{f}_l = \boldsymbol{o}. \tag{11.8}$$

If we make the very reasonable assumption that the Gram matrix with components $\langle\phi_k,\phi_l\rangle$ is nonsingular, then (11.8) can be put into standard form. In particular, if $\phi_k(s) = \sin k\pi s$ and if ρA is constant, then this matrix as well as the corresponding

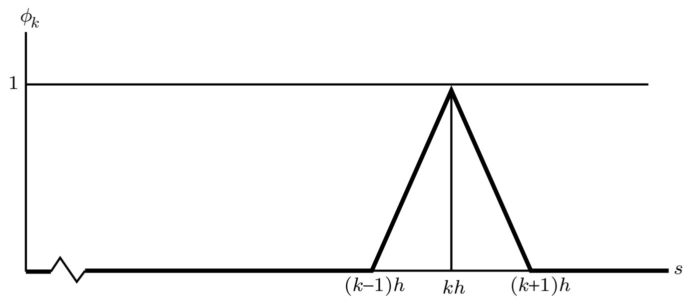

Figure 11.10. The function ϕ_k.

matrix with components $\langle f'_k, \phi'_l \rangle$ is diagonal. For practical computation, this virtue is counterbalanced by the high cost of the numerical evaluation of the integrals in (11.4) and (11.8).

Let us set $h = 1/(K+1)$ and

$$\phi_k(s) = \begin{cases} h^{-1}[s-(k-1)h] & \text{for} \quad (k-1)h \le s \le kh, \\ 1 - h^{-1}(s-kh) & \text{for} \quad kh \le s \le (k+1)h, \\ 0 & \text{elsewhere.} \end{cases} \tag{11.9}$$

(This function is shown in Fig. 11.10.) When (11.9) is used, the matrices whose elements are $\langle \phi_k, \phi_l \rangle$ and $\langle \phi'_k, \phi'_l \rangle$ are tridiagonal. The cost of the numerical evaluation of the integrals in (11.4) and (11.8) is low. (Matrices with components $\langle \phi'_k, \phi'_l \rangle$ arise naturally in the linearization of (11.8) and are associated with the finite-difference approximation of $\boldsymbol{r}_{ss}$.) The choice (11.9) gives the simplest (semi-discrete) finite-element approximation to our nonlinear initial-boundary-value problem. If ρA is constant, say $\rho A = 1$, then

$$\langle \phi_k, \phi_k \rangle = \frac{4h}{6}, \quad \langle \phi_k, \phi_{k+1} \rangle = \frac{h}{6}, \quad \langle \phi_k, \phi_l \rangle = 0 \quad \text{for} \quad l \ne k-1, k, k+1. \tag{11.11}$$

If, however, we were to approximately evaluate these integrals by using the trapezoid rule, then we would find that $\langle \phi_k, \phi_l \rangle = \delta_{kl}$. In this case, the left-hand side of (11.8) would uncouple and the resulting equations could be identified with the equations of motion of K beads joined by massless nonlinearly elastic springs.

11.12. Exercise. Replace $\boldsymbol{r}$ in (10.17) with $\boldsymbol{r}^K$ of (11.1). Show that the vanishing of the Gâteaux derivative of the resulting functional of $\{\boldsymbol{r}_1, \dots, \boldsymbol{r}_K\}$ is equivalent to (11.4).

11.13. Exercise. Using the principles of classical particle mechanics, find the equations of motion of K beads joined in sequence by massless nonlinearly elastic springs, with the first and the Kth bead joined to fixed points by such springs. Compare the resulting equations with (11.8). Formally obtain (2.9)–(2.11) by letting $K \to \infty$ while the total mass of the beads stays constant. (See the discussion at the end of Sec. 2.)

Even though the form of the governing equations for discrete models converges to the form of the governing equations for string models, it does not follow that the solutions of the former converge to solutions of the latter in any physically reasonable sense. Von Neumann (1944) (in a paper filled with valuable insights) advanced the view, now recognized as false, that the solutions for the positions of the beads in the equations of Ex. 11.13, together with their time derivatives and suitable difference quotients should converge respectively to the position, velocity, and strain fields for (2.9)–(2.11). This convergence is valid only where the partial differential equations have classical solutions. Where the velocity and strain suffer jump discontinuities (shocks), the solutions of the discrete problem develop high-frequency oscillations that persist in the limit as $K \to \infty$. Consequently it can be shown that the limiting stress is incorrect. For thorough discussions of this phenomenon and related issues, see Greenberg (1989, 1992), Hou & Lax (1991), and the references cited therein.

Likewise, finite-element discretizations of dynamic problems of nonlinear elasticity may fail to give sharp numerical results because they are not well adapted to capture the shocks such systems may possess. (There is an effort to change this state of affairs.) There are a variety of effective numerical schemes, originally developed for gas dynamics, that can effectively handle shocks. One such scheme, that of Godunov (see Bell, Colella, & Trangenstein (1989), e.g.), may be regarded as a discretization of the Impulse-Momentum Law of Sec. 3 in a way that exploits the characteristics of the underlying hyperbolic system. The trouble with many such schemes is that they have inherent dissipative mechanisms, inspired by those for gas dynamics, that are not invariant under rigid motions and could therefore lead to serious errors in problems for which there are large rotations. (Cf. Antman (1998, 2003a) and Sec. **8**.9.) There is an extensive literature on the finite-element method for equilibrium problems. Among the more mathematical works oriented toward solid mechanics are those of Brenner & Scott (2002), Brezzi & Fortin (1991), Ciarlet (1978), Ciarlet & Lions (1991), Hughes (1987), Johnson (1987), Oden & Carey (1981–1984), and Szabó & Babuška (1991).

Although we do not know in what sense the solution of (11.8) converges to the solution of the partial differential equations for elastic strings, we might be able to resolve this question for viscoelastic strings by using modern analytic techniques associated with the Faedo-Galerkin method. (See Ladyzhenskaya (1985), Ladyženskaja, Solonnikov, & Ural'tseva (1968), Lions (1969), and Zeidler (1990, Vol. IIB).) An analysis along these lines for a quasilinear engineering model of an elastic string was carried out by Dickey (1973). He proved that the solutions of a system like (11.8) converge to the classical solution of the partial differential equations until the advent of shocks. Antman & Seidman (1996) used the Faedo-Galerkin method to treat the longitudinal motion of a viscoelastic rod with a constitutive equation of the form (2.14).

Springer

http://www.springer.com/978-0-387-20880-0

Nonlinear Problems of Elasticity
Antman, S.S.
2005, XVIII, 838 p. 110 illus., Hardcover
ISBN: 978-0-387-20880-0