Normal modes in vocal cord tissues

Ingo R. Titze

Department of Physics, College of Petroleum and Minerals, Dhahran, Saudi Arabia

William J. Strong

Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602

The human vocal cord is treated as an elastic material capable of propagating compressional, shear, and surface waves. A mathematical formulation for commonly observed vibrational modes is developed. This includes horizontal as well as vertical tissue displacements. Current vocal cord models are discussed and evaluated in terms of their ability to describe adequately the vibrational degrees of freedom of the cords.

Subject Classification: 70.20; 20.40.

INTRODUCTION

The anatomy and physiology of phonation have received considerable attention for many years; so has the acoustics of speech production. Unfortunately, the attempts to establish rigorous support for glottal behavior in terms of the basic vibrational modes of the vocal cord tissues have been qualitative and descriptive, to a large extent. As a result, the functional models of the larynx, and specifically those of the vocal cords, have consisted of mechanical oscillators with one, or at best a few, degrees of freedom. There are many unanswered fundamental questions, such as: What are the normal modes of vibration of the vocal cords? Which modes are excited by the glottal air stream? What is the nature of the energy-loss mechanism within the vocal cord tissues? What is the effective mass of the cords for different vocal adjustments? This paper deals only with the first of these questions. The others, closely connected with the first, will be considered in a succeeding paper. An attempt is made here to treat the vocal cords as an elastic medium capable of absorption, transmission, and reflection of acoustic energy. Based on this point of view, the energy coupling between the glottal air stream and the vocal cord medium is bilateral. The vocal cord modulates the glottal air stream, and the air stream excites acoustic waves in the tissues. As one might expect, the mechanical properties of muscular, connective, and membranous tissues of the vocal folds play an important part in this analysis. In fact, it is the uncertainties of these properties under a variety of stresses which account for the most serious limitations presently experienced in the quantitative aspects of this study.

I. REVIEW OF TISSUE ANATOMY

As a viscoelastic medium with well established boundaries, the human vocal cords are capable of natural vibration and energy dissipation. A variety of soft tissues, including muscle fibers, sarcolemma (elastic connecting tissue), ligaments, collagen, mucus, and membranes, constitute the medium in which energy can be propagated and dissipated. Cartilage, stationary muscle, and the glottis constitute the boundary, where reflection and coupling of energy to surrounding media

take place. Unfortunately, both the medium and the boundary are quite irregular, the former being both non-homogeneous and anisotropic, the latter being neither rectangular nor cylindrical. Any quantitative mathematical treatment, therefore, requires some simplifying assumptions.

The primary anisotropy results from the muscle fibers. There is general agreement among anatomist (Wustrow, 1952; Behringer, 1955; van den Berg, 1958; van den Berg and Moll, 1955; von Leden, 1961) that there are two main bundles of fibers, the thyrovocalis and the thyromuscularis. These are referred to as simply vocalis and muscularis in this paper. Both sets of these muscle fibers extend longitudinally from the thyroid cartilage to various portions of the arytenoid cartilage. Their functions, however, are somewhat different (von Leden, 1961). The vocalis helps to regulate the tension of the vocal cord, whereas the muscularis, whose fibers are oriented somewhat obliquely to those of the vocalis, is primarily responsible for gross adjustments in vocal cord thickness. Both sets of muscle fibers play a part in the control of vocal register (van den Berg, 1960).

For the purpose of this treatment, we assume an idealized rectangular configuration of the vocal cord, as shown in the cross section in Fig. 1. Various tissues of the vocal cords are indicated in layers, proceeding from the glottis toward the thyroid cartilage, which constitutes the lateral boundary. The glottal edge of the vocal cord consists of the vocal ligament, an elastomer which is probably the most solid-like member in the overall vocal cord structure. It can be strained about 30% and has been known to support longitudinal tensions up to 2.2 kg in cadavers (Sonninen, 1973).

There has been much dispute about the manner in which the ligament and vocalis are connected. Some authors (Goertler, 1950; Wustrow, 1952) have claimed that muscle fibers extend into the ligament, thus creating a very firm coupling, while other authors (Behringer, 1955; van den Berg and Moll, 1955; von Leden, 1961) insist that only connective tissue, primarily in the form of collagen, exists between ligament and muscle (see Fig. 1). In recent years the latter point of view has become more acceptable, and we adhere to it here. Since
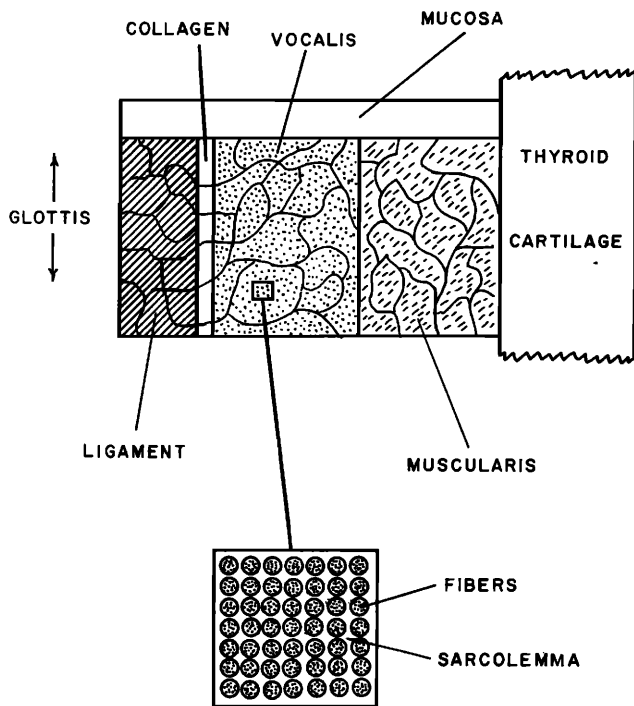
FIG. 1. Simplified sketch representing the internal tissue structure of the left vocal cord in coronal cross section. A rectangular geometry has been assumed.

collagen is a gel under normal temperature conditions, it would appear that the coupling between ligament and vocalis is rather loose. The situation is complicated, however, by the fact that elastic strands extend from the ligament and conus elasticus deep into the muscular tissues, creating, in effect, a "web" or "honey comb" structure, i.e., a strong, elastic framework. Both collagen and muscle fibers are therefore partitioned, allowing for only moderate amounts of shear in any given direction.

A dividing line between the vocalis and muscularis is indicated in Fig. 1. This is to accommodate the possibility that the two main bundles of fibers can be activated independently to some degree. Should this be the case (a notion which has not received much debate), a sizable contraction of the muscularis would cause rigidity in the lateral portions of the vocal cord, thus moving the effective boundary of the cord medially. This concept has been used to explain the phenomenon of "damping," i.e., fixing part of the cord, while another portion remains free to vibrate (Rubin and Hirt, 1960; von Leden, 1961).

The mucous membrane mantles the inner vocal cord structure of the glottal and superior surfaces. During phonation, however, this loose structure is initially blown upward within the glottis, thereafter residing primarily on top of the ligament and vocalis. Thus, the upper layer of the vocal cord resembles a viscous liquid (mucus) with exceptionally high surface tension (the mucous membrane).

In summary, the five regions identified in Fig. 1 provide a reasonable representation of the piecewise homo-

geneous tissue structure. The complexity of this structure can only be fully appreciated when one recognizes that, over the wide range of possible vocal adjustments, both the regional configuration, and the tissue properties within each region, are subject to change. Thus, the degree of muscular activity, externally imposed tissue strains, and even temperature, can cause significant modification in the arrangement of the tissue members and their mechanical properties. A precise description of the normal modes of this system must therefore be limited to a particular vocal adjustment. For example, if the normal chest (modal) register is under consideration, the muscular portions of the vocal cord constitute the bulk of the system. A single-layer structure would then suffice, ligamental and mucosal portions being added as perturbations to the primarily muscular system. For the falsetto register, on the other hand, the muscular portions can be omitted entirely, the ligament can be treated as a thin bar (or a thick string), and the mucosa as a secondary "string," this being much more tightly coupled to the ligament than in the lower register. We are concerned primarily with the single-layer configuration in this report.

II. WAVE PROPAGATION IN VOCAL CORD TISSUES

In order to approximate the natural responses of the vocal cord, let us begin by deriving the wave equation for a "miniscopically homogeneous" fibrous substance. In other words, we assume that the vocal cord consists of longitudinal muscle fibers, as illustrated in the enlarged square in Fig. 1. The muscle fibers are surrounded and separated, however, by thin layers of connective tissue called sarcolemma. This substance is primarily elastic and accounts for the bulk of the restoring force under compression and shear. We allow the tissue properties to vary as a function of vocal cord depth, i.e., laterally. Figure 2 illustrates a rectangular
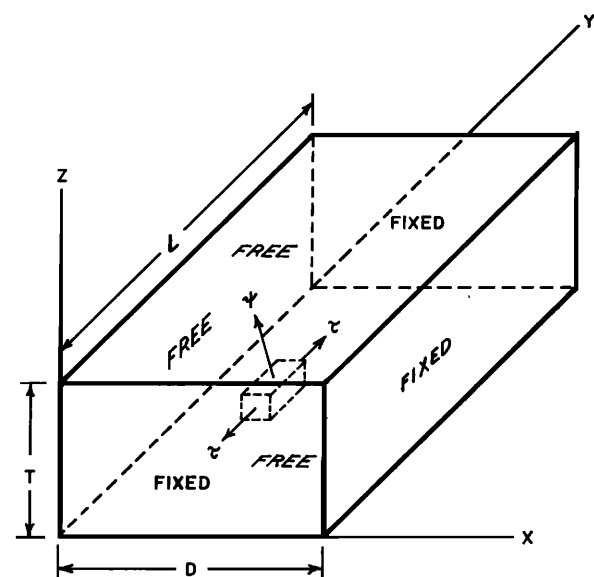

FIG. 2. Differential element of tissue, with displacement vector $\vec{\psi}$ and longitudinal stress τ, within a rectangular parallelepiped representing the vocal cord. Transverse elastic moduli are not indicated. Note boundary conditions.

parallelepiped in Cartesian coordinates. Some bound-
ary conditions are indicated, but these are not relevant
at this point, and are discussed later. The origin of the
coordinate system is centered at the vocal processes,
i.e., posteriorly and on the center line of the glottis.
The anterior end of the cord is therefore at $y = L$, the
thickness T is measured along the z axis. We assume
that any portion of the vocal cord can be displaced ar-
bitrarily by a vector $\vec{\Psi}$, which has components ξ in the
x direction and ζ in the z direction. Thus we may write

$$\vec{\Psi}(x,y,z,t) = \xi(x,y,z,t)\hat{x} + \zeta(x,y,z,t)\hat{z} . \tag{1}$$

Now consider a differential element of tissue, i.e., a
small group of fibers, represented by the smaller par-
allelepiped in Fig. 2. If the longitudinal stress is $\tau(x)$,
the Young's modulus $\epsilon(x)$, the shear modulus $\eta(x)$, and
the tissue density $\rho(x)$, the equations of motion for the
differential element are

$$\Delta x \Delta z \tau [(\partial \xi/\partial y)_y - (\partial \xi/\partial y)_{y+\Delta y}]$$

$$+ \Delta y \Delta z [(\epsilon \partial \xi/\partial x)_x - (\epsilon \partial \xi/\partial x)_{x+\Delta x}]$$

$$+ \Delta x \Delta y \eta [(\partial \xi/\partial z)_z - (\partial \xi/\partial z)_{z+\Delta z}]$$

$$+ \Delta x \Delta y \Delta z \rho (\partial^2 \xi/\partial t^2) = 0, \tag{2}$$

$$\Delta x \Delta z \tau [(\partial \zeta/\partial y)_y - (\partial \zeta/\partial y)_{y+\Delta y}]$$

$$+ \Delta y \Delta x \epsilon [(\partial \zeta/\partial z)_z - (\partial \zeta/\partial z)_{z+\Delta z}]$$

$$+ \Delta y \Delta z [(\eta \partial \zeta/\partial x)_x - (\eta \partial \zeta/\partial x)_{x+\Delta x}]$$

$$+ \Delta x \Delta y \Delta z \rho (\partial^2 \zeta/\partial t^2) = 0. \tag{3}$$

Each bracketed term represents a differential force ex-
erted on the tissue element owing to stresses on opposite
parallel planes. The stresses are related to the strains
linearly through Hooke's law. Although the overall
stress–strain curve for human tissue is nonlinear, there
are many phonatory adjustments in speech and singing
which result in small amplitude vibration, and thus
linearization is valid. Whenever the ligament is stretched,
for example, because of external cricothyroid contrac-
tion, the vibrational amplitude is relative small. Small
amplitudes are also experienced when the blowing pres-
sure is small, or when adduction is incomplete.

A further justification for the present approach is,
however, not based upon displacement amplitude or
linearity at all. We are trying to establish the normal
modes of vibration, and therewith the number of degrees
of freedom which are necessary to describe adequately
the observed vibrational patterns. It is conceded at
onset that the driven frequencies of vibration in many
cases differ substantially from the normal-mode fre-
quencies. For example, in the typical modal register
of phonation, τ, ϵ, and η may be subject to large fluc-
tuations, i.e., their effective value may change over
different portions of a cycle. During glottal closure,
the effective Young's modulus ϵ increases owing to
direct contact with the opposite vocal cord. Further-
more, the longitudinal stress τ can readily increase
during peak excursions of the vocal ligament. In other
words, since the strain–stress curves of the ligament
and vocalis are highly nonlinear (van den Berg, 1960),
dynamic stresses and static stresses can differ appreci-
ably. These nonlinear features of the tissue properties

have all been incorporated in precise dynamic modeling
of the vocal cords (Ishizaka and Flanagan, 1972; Titze,
1973,), but can be ignored in the present investigation
of the normal modes.

Proceeding with the analysis, Eqs. 2 and 3 can be
cast into more familiar elastic wave equations by divid-
ing by $\Delta x \Delta y \Delta z$ and allowing the differential lengths to
approach zero:

$$\tau \frac{\partial^2 \xi}{\partial y^2} + \epsilon \frac{\partial^2 \xi}{\partial x^2} + \eta \frac{\partial^2 \xi}{\partial z^2} + \frac{\partial \epsilon}{\partial x} \frac{\partial \xi}{\partial x} - \rho \frac{\partial^2 \xi}{\partial t^2} = 0, \tag{4}$$

$$\tau \frac{\partial^2 \zeta}{\partial y^2} + \epsilon \frac{\partial^2 \zeta}{\partial z^2} + \eta \frac{\partial^2 \zeta}{\partial x^2} + \frac{\partial \eta}{\partial x} \frac{\partial \zeta}{\partial x} - \rho \frac{\partial^2 \zeta}{\partial t^2} = 0. \tag{5}$$

Henceforth, the solutions to Eq. 4 are entitled x modes
and the solutions to Eq. 5 z modes. Thus, the x modes
describe horizontal motion of the vocal cords (trans-
verse to the glottal air flow) and the z modes describe
vertical motion of the vocal cords (in the direction of
the air flow).

III. DISCUSSION OF TISSUE PARAMETERS

Equations 4 and 5 are of course similar to the classi-
cal wave equations for small vibrations in linear iso-
tropic elastic media. There are some notable differ-
ences, however, The transverse shear is τ with re-
spect to the y axis and η with respect to the x and z
axes. This, of course, reflects the anisotropy of the
tissue, and the fact that large external and internal
stresses can be applied in the direction of the tissue
fibers. It has already been pointed out that the vocal
ligament can be subjected to enormous stresses in ca-
davers. But what about the stresses which the human
himself can exert during normal conditions of phonation?
Most of the information available at present is in the
form of electromyographic data. Although emg record-
ings of muscular activities are very useful for qualita-
tive description of muscular behavior, the authors have
so far been unsuccessful in transforming "level of activ-
ity" into the mechanical equivalents of dynes/cm². One
attempt at such a transformation has been carried out
by Hast (1966) in an experiment on dog larynxes. In
this study the recurrent nerve was artificially stimulated
by electronic pulses at varying frequencies. The section
of the thyroid cartilage containing the anterior attach-
ment of the thyrovocalis muscle was carefully cut out
and attached to a force transducer. In this manner the
level of muscular activity could be correlated with the
isometric tension. The results indicate that the over-
all muscular tension ranges between 100 and 300 g
when the muscle is stimulated, and between 0 and 100
g when the muscle is passive. Total elongation (strain)
and frequency of stimulation govern the exact values
within those ranges. Unfortunately, the value of the
cross-sectional area of the muscle is not given. Thus,
in order to determine the average longitudinal stress τ,
we estimate a muscular cross section of about 0.25
cm². Then

$$\tau = (200 \text{ g})(980 \text{ cm/sec}^2)/(0.25 \text{cm}^2) \approx 10^6 \text{ dynes/cm}^2$$

for the average active stress, and approximately one
fourth that for the average passive stress. It is shown
later, by an indirect method using normal-mode fre-

quencies, that the above value derived from measurements on dogs is not unrealistic for human subjects.

Now let us focus our attention on the remaining elastic constants ϵ and η. It might at first appear as though ϵ were a bulk compressibility, as it would be if Eqs. 4 and 5 were to represent wave motion in infinite, or semi-infinite, elastic media. Such a bulk compressibility would be extremely high compared to any other elastic constants of the tissue—in the order of 10^{10} dynes/cm^2 (Oestreicher, 1951). It is inconceivable that compressional waves of such high velocities would play a role of any significance in vocal cord vibration, since frequencies are low and the extent of the material is very limited. On the other hand, since the vocal cord depth and thickness are so small, a compression of one of these can take place at the expense of an extension of the other, still maintaining overall bulk incompressibility, or near incompressibility. The amount of transverse extension is, of course, measured by Poisson's ratio of the tissue. The circular geometry of the muscle fibers and surrounding sarcolemma turns into a slightly oval geometry. It might be expected that such a deformation is accompanied by a sizable loss of energy. Viscosity of compression and shear will be treated in a succeeding paper.

Before attempting to give an estimate of the Young's modulus ϵ, it is beneficial to discuss the shear modulus η, for it has been determined for some tissues. Oestreicher (1951) gives a value of 2.5×10^4 dynes/cm^2 for tissues near the body surface. This constant was derived from measurements of the mechanical impendance at the surface of the body for a wide range of frequencies Since the vocal ligament is composed of elastic fibers with considerably higher shear modulus than the softer muscle and fatty tissues, and since elastic strands penetrate throughout the entire vocal cord, it is probably more correct to assume an average η of about 10^5 dynes/cm^2.

The deformation of the fiber—sarcolemma structure under shear is really no different from that already discussed for lateral or vertical "compressibility." Circular ministructures are converted into slightly oval shapes, only the orientation is different. We conclude, therefore, that ϵ and η must be of the same order of magnitude. In summary, then, the values for the bulk tissue parameters in Eqs. 4 and 5 are assumed to be

$\tau = 10^5 - 10^7$ dynes/cm^2 (average 10^6 dynes/cm^2),

$\epsilon \approx 10^5$ dynes/cm^2 (average),

$\eta \approx 10^5$ dynes/cm^2 (average),

$\rho = 1.1$ g/cm^3 (approximately water).

Obviously, a great deal of refinement is necessary before one can treat the zonal structure of Fig. 1 in great detail.

IV. NATURAL MODES OF THE VOCAL CORDS

We shall derive the normal modes for two simplified cases of the overall configuration indicated in Fig. 1. First, let us assume that the entire vocal cord consists of a single layer of muscle with uniform properties throughout. Standard techniques of separation of variables are employed in the solution of Eqs. 4 and 5. For sinusoidal variation in time, the solutions can be written as

$$\zeta(x,y,z,t) = A e^{i\omega_x t} \cos\frac{(2n_x-1)\pi x}{2D} \sin\frac{n_y \pi y}{L} \cos\frac{n_z \pi z}{T}, \quad (6)$$

$$\zeta(x,y,z,t) = B e^{i\omega_z t} \cos\frac{(2n_x-1)\pi x}{2D} \sin\frac{n_y \pi y}{L} \cos\frac{n_z \pi z}{T}, \quad (7)$$

where the free—fixed, fixed—fixed, and free—free boundary conditions indicated in Fig. 2 have been imposed. n_x, n_y, n_z are integers, A and B arbitrary constants, and ω_x and ω_z are the radian frequencies of vibration. When Eqs. 6 and 7 are substituted into the wave equations 4 and 5, the normal-mode frequencies are obtained as

$$f_x = \frac{1}{2}\left(\frac{\tau}{\rho}\right)^{1/2}\left[\frac{n_y^2}{L^2} + \frac{(2n_x-1)^2}{4D^2}\frac{\epsilon}{\tau} + \frac{n_z^2}{T^2}\frac{\eta}{\tau}\right]^{1/2} \quad x \text{ modes}, \quad (8)$$

$$f_z = \frac{1}{2}\left(\frac{\tau}{\rho}\right)^{1/2}\left[\frac{n_y^2}{L^2} + \frac{(2n_x-1)^2}{4D^2}\frac{\eta}{\tau} + \frac{n_z^2}{T^2}\frac{\epsilon}{\tau}\right]^{1/2} \quad z \text{ modes}. \quad (9)$$

Note, that if ϵ and η are equal, as they have been assumed here, the x and z modes are degenerate. Furthermore, if both ϵ and η are much less than τ, conventional string modes result. Figure 3 illustrates tissue displacements for some common x and z modes. Discussion of the eigenfrequencies themselves is entertained later.

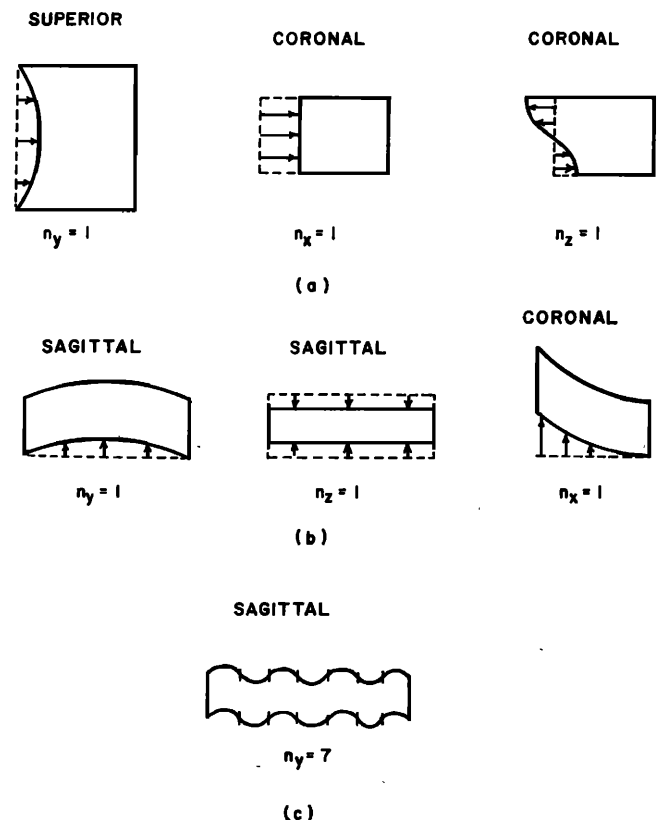In most cases of phonation, only the lowest-order modes are excited by the glottal air stream. Thus, the

FIG. 3. Tissue displacements for some vibrational modes which have been observed. (a) Low-order x modes, (b) low-order z modes, and (c) higher-order z mode.

$n_y = 1$ x mode [Fig. 3(a), left] reflects the often observed "oval" glottis, whereas the more nearly rectangular glottis corresponds to $n_x = 1$ [Fig. 3(a), center]. Whenever there are sudden adjustments, such as "voice breaks," or whenever the vocal cords are "pulse excited," as is the case during very forceful glottal closure or opening, there is a strong possibility that higher modes of the cords are indeed excited. This has been demonstrated by Rubin and Hirt (1960) via high-speed cinematography. Six ridges of mucus, collected on the superior surface of the vocal ligaments, are evidence of the fact that standing waves are present within the ligament and mucous membrane. This indicates that higher z modes, at least up to $n_y = 7$, are possible in some subjects under some conditions. In the particular case reported by Rubin and Hirt, the vocalist was a female singing a high C (approximately 1000 Hz). Normal-mode frequencies up to 6000 Hz were excited temporarily within the vocal cord tissues. Figure 3(c) illustrates displacements of the ligaments for the z mode just described. These displacements are, of course, highly exaggerated. Any mucus residing on the surface of the vocal cord collects at the locations designated by small vertical lines.

Rubin and Hirt also illustrate motion picture sequences for a typical "voice break," i.e., a sudden adjustment from falsetto to full voice, or vice versa. By careful observation of their frame-by-frame sequence, one can note higher-order x modes during the transition. Higher-order x modes were also observed by Rubin and Hirt along the glottal edge. The authors called the phenomenon "scalloping," but it is not nearly as convincing as the "ridging" phenomenon described earlier.

Of particular interest has been the phenomenon called "vertical phasing," i.e., the movement of different vertical portions of the glottal edge of the vocal cord in opposite directions at a given time (Farnsworth, 1940; Flanagan and Landgraf, 1968; Ishizaka and Flanagan, 1972; Titze, 1973; Titze, 1974). What is described is usually a combination of two phenomena which are somewhat independent. If the phasing occurs primarily within the vocalis muscle, we are talking about the $n_z = 1$ shear mode illustrated in Fig. 3(a), right side. The shear modulus η governs this mode. It is a particularly likely mode to be excited by the air stream when the vocal folds are thick and lax (Gupta et al., 1973). During glottal closure the bottom portion of the vocal fold is forced laterally by the subglottal pressure, whereas during the open portion of the cycle the Bernoulli pressure tends to "suck" the same portion medially. The top portion follows out of phase. In this mode of vibration the upper and lower portions have very nearly equal masses, as dictated by the cosine term in Eq. 7.

There is another kind of vertical phasing which has been described by Hirano (1968), among others. This involves the relative motions between the mucosa and the ligament vocalis, and occurs whenever the vocal cord is unstretched. A surface wave is seen to propagate laterally from the glottis toward the vocal cord boundary, as illustrated in Fig. 4. Due to the high surface tension of the mucosa (which is, of course, the
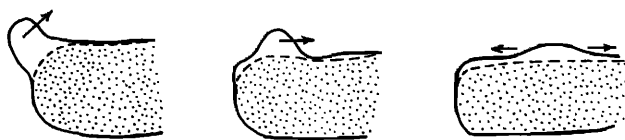
FIG. 4. Surface wave in the mucosal portions of the vocal cord.

tension of the mucous membrane), the surface wave is readily dispersed, but occasionally gets reflected from the boundary and travels back toward the glottis (Farnsworth, 1940). It is important to realize that a surface wave of this type originates from an initial upward motion of the mucosa. When the ligaments separate after closure of the glottis, the mucosal portions have a tendency to remain adducted for a short period of time. This allows the full subglottal pressure to be applied to part of the mucosa. The result is a sudden upward displacement, followed by lateral propagation and dispersion.

In certain types of phonation the z modes (vertical motion of the cord) may be more significant than the x modes. This is the case, for example, when the vocal processes are firm in place, but the vocal cord length is equal to, or even less than, the neutral length. The center of the vibrating portion of the cord is then blown upward like a swinging door before it proceeds laterally. Low-order z modes, corresponding to $n_y = 1$, $n_z = 1$, and $n_x = 1$ are depicted in Fig. 3(b). The net effect of a combination of the $n_y = 1$ z mode and the $n_y = 1$ x mode is a rotary motion of the vocal cord edge. This kind of motion is particularly likely to occur when the vocal cord edge is thin, because the area of attachment of the ligament to the vocalis is reduced, causing the ligament to move quite independently from the rest of the cord. It is conveivable that under some conditions the lateral motion may become very minimal, or cease altogether. This might occur when there is incomplete adduction and high tension in the ligaments. The air flow through the narrow glottis is then unmodulated, but the vocal tract is still excited because of the vertical motion of the cord. One might call this a "membrane" mode, because, aside from the slit in the center, the vocal cords are functioning like a membrane stretched across a circular tube.

V. DISCUSSION OF EIGENFREQUENCIES

Now let us briefly discuss the eigenfrequencies themselves, shown in Fig. 5. Only 18 of the lower frequencies are depicted. These correspond to all combinations of the integer values:

$$n_x = 1, 2, 3, \quad n_y = 1, 2, 3, \quad n_z = 0, 1.$$

For the sake of simplicity, the degenerate case was chosen ($\epsilon = \eta = 10^5$ dynes/cm^2). Unequal values of ϵ and η would produce an additional set of 18 frequencies, corresponding to the z modes. These would be distributed evenly within the region of the existing eigenfrequencies.
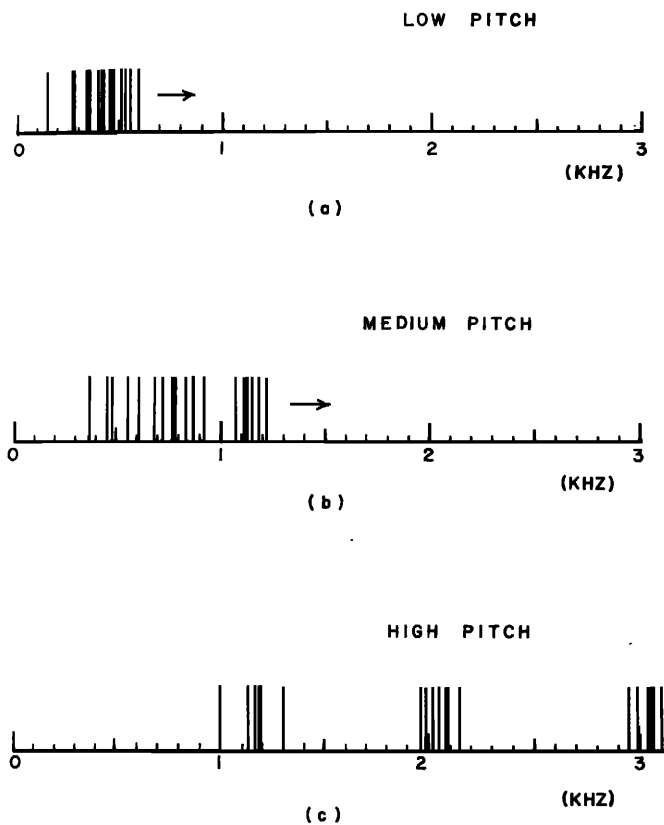
The parameter of primary interest is the longitudinal

LOW PITCH

(a)

MEDIUM PITCH

(b)

HIGH PITCH

(c)

FIG. 5. Normal-mode frequencies of the vocal cord for three different adjustments of longitudinal stress. (a) $\tau = 10^5$ dynes/cm^2, (b) $\tau = 10^6$ dynes/cm^2, and (c) $\tau = 10^7$ dynes/cm^2. In all cases $\epsilon = \eta = 10^5$ dynes/cm^2.

stress τ. In an attempt to illustrate the eigenfrequencies for three distinctly different vocal adjustments, the following conditions were chosen:

(1) $L = 1.2$ cm, $D = 1.0$ cm, $T = 0.7$ cm,

$\tau = 10^5$ dynes/cm^2,

τDT = longitudinal tension

= 70 000 dynes (approximately 70 g),

lowest eigenfrequency = 154 Hz;

(2) $L = 1.4$ cm, $D = 0.8$ cm, $T = 0.5$ cm,

$\tau = 10^6$ dynes/cm^2,

τDT = longitudinal tension

= 400 000 dynes (approximately 400 g),

lowest eigenfrequency = 371 Hz;

(3) $L = 1.6$ cm, $D = 0.6$ cm, $T = 0.3$ cm,

$\tau = 10^7$ dynes/cm^2,

τDT = longitudinal tension

= 18 000 000 dynes (approximately 1.8 kg),

lowest eigenfrequency = 1000 Hz.

It must be noted that the longitudinal tension computed in part (3) above is unrealistic. Although neither the dimensions of the vocal cord nor the value of τ is un-

realistic, it must be borne in mind that the vocal cord supports high tensions primarily in the ligament, i.e., in the glottal edge of the vocal cord, rather than throughout its entire depth. An estimate of the cross-sectional area of the vocal ligament can be made from pictures given by Hirano (1968) and van den Berg (1958). Such an estimate is in the order of 0.05 cm^2, depending on the subject. This is less than one third of the product of D and T indicated in part (3) above. With this consideration, the longitudinal tension required to produce a 1000-Hz fundamental frequency would be 500 g instead of 1.8 kg.

It was stated earlier that it is possible to utilize the normal-mode analysis to estimate the ligament stress required to produce the extremely high pitches achieved by trained vocalists. This is now evident. Assuming that the ligament supports virtually all of the longitudinal tension, that ϵ and η are much smaller than τ, and that the fundamental frequency of the glottal air stream (F_0) corresponds to the lowest-order x mode, Eq. 8 yields

$$\tau = 4L^2 f^2 \rho \ . \tag{10}$$

For a female vocalist singing high C (≈ 1000 Hz), this corresponds to the value of 10^7 dynes/cm^2 in part (3) above ($L = 1.6$, $\rho = 1.0$ g/cm^3). This is about five times larger than the average stress and about three times larger than the maximum stress computed earlier from the measurements made by Hast (1966) on live dogs. Recall, however, that an uncertainty in the cross-sectional area of the muscle existed in that computation. For this reason it might be a little more accurate to compare the total longitudinal tension, which has already been suggested to be 500 g for this case. This compares favorably with the 300-g maximum tension noted by Hast on dogs.

The foregoing discussion points out the limitations of the single-zone structure considered here. The change of elastic properties from ligament to vocalis can be accounted for by allowing two zones. If the ligament extends from $x = 0$ to $x = d$ and the vocalis from $x = d$ to $x = D$, the input impedance at the glottis can be determined by a two-step impedance transformation as follows:

$$Z(d) = -i(k_{x2}\epsilon_2\omega) \cot k_{x2}(D-d), \tag{11}$$

$$Z(0) = (k_{x1}\epsilon_1\omega)\frac{-i(k_{x2}\epsilon_2/k_{x1}\epsilon_1)\cot k_{x2}(D-d) + i\tan k_{x1}d}{1 + (k_{x2}\epsilon_2/k_{x1}\epsilon_1)\cot k_{x2}(D-d)\tan k_{x1}d} \ . \tag{12}$$

In the above, k_{x1}, k_{x2}, ϵ_1, and ϵ_2 represent the propagation constants (wavenumbers) and Young's moduli of the ligament and vocalis, respectively. The propagation constants are determined by the tissue properties and the the overall boundary conditions:

$$k_{x1} = [\omega_x^2 \rho/\epsilon_1 - (n_y\pi/L)^2\tau_1/\epsilon_1 - (n_z\pi/T)^2\eta_1/\epsilon_1]^{1/2}, \tag{13}$$

$$k_{x2} = [\omega_x^2 \rho/\epsilon_2 - (n_y\pi/L)^2\tau_2/\epsilon_2 - (n_z\pi/T)^2\eta_2/\epsilon_2]^{1/2}. \tag{14}$$

The normal modes of the system are then obtained by letting the input impedance, defined by Eq. 12, vanish:

$$(k_{x2}\epsilon_2)/(k_{x1}\epsilon_1)\cot k_{x2}(D-d) = \tan k_{x1}d \ . \tag{15}$$

An exact solution of ω_x from Eq. 13, 14, and 15 requires numerical methods. Qualitatively, the results are not

much different from the simpler single-zone model.
Since the tissue parameters are not known with much
precision at this point, it is not fruitful to continue a
quantitative analysis of multiple-zone models.

Returning once more to the eigenfrequencies in Fig.
5, note that at low pitches the spectrum is almost con-
tinuous, especially in view of the fact that any amount of
damping results in a broadening of each of the lines. It
is therefore highly probable that the fundamental fre-
quency is determined in part by the aerodynamic forces.
This substantiates the myoelastic—aerodynamic theory
of voice production for low pitches. For extremely high
pitches, however, the natural frequencies congregate
in clusters around the harmonic string modes. It is
therefore concluded that at high pitches F_0 very nearly
corresponds to the natural frequency of vibration of the
cords, allowing for only slight modification owed to
aerodynamic forces.

Although the vocal cord length, depth, and thickness
have been systematically adjusted to reduce the effec-
tive mass of the vocal cord at higher pitches, no attempt
is made here to categorize these adjustments by reg-
ister. Since register is intimately connected with the
efficiency of conversion of steady flow into time-varying
flow, the shape of the glottis is more crucial than the
shape of the vocal cord. Thus, in order to accommodate
registration, a vocal cord model must not only have a
number of longitudinal and vertical degrees of freedom
in the tissue but also provisions for adjusting the static
(prephonatory) position of the vocal cord. Discussion
of vocal cord models and their adequacy is entertained
presently.

VI. LUMPED MASS-SPRING MODELS

Although many analog models of the vocal cords have
been devised over the last century, the digital versions
of the last five years have been the significant ones for
quantitative analysis. Of these, the single-mass model
by Flanagan and Landgraf (1968), the two-mass model
by Ishizaka and Flanagan (1972), and the 16-mass model
by Titze (1973, 1974) are discussed briefly. The Flan-
agan and Landgraf model has a single horizontal degree
of freedom. Although it has been successful in demon-
strating self-oscillation, some gross features of regis-
ter, and plausible relationships between subglottal pres-
sure and fundamental frequency, there are some serious
limitations. The restricted range of oscillation, for ex-
ample, resulting from interactions between the vocal
cords and the vocal tract, has been discussed in detail
by Ishizaka and Flanagan (1972). Aside from this, how-
ever, it is evident from present considerations that a
single-mass model cannot realistically account for shear
in the tissues. Furthermore, a single spring represent-
ing the lateral restoring force does not adequately dif-
ferentiate between the longitudinally applied stress τ and
the lateral Young's modulus ϵ.

The two-mass model by Ishizaka and Flanagan does
incorporate shear, but is equally inadequate with re-
spect to longitudinal details. Figure 6(a) is a sketch of
the model. The spring constants k_1 and k_2 are effective
spring constants which take into consideration both lon-

gitudinal tension and Young's modulus, whereas k_c ac-
counts for the shear between the masses. It has been
pointed out earlier that, unless the upper mass repre-
sents different tissue (such as the mucosa), the two
masses must be of equal size to satisfy the second co-
sine term in Eq. 7. The so-called "typical" conditions
by Ishizaka and Flanagan consist of unequal masses,
suggesting modeling of both types of tissue. Since the
model has no vertical degrees of freedom, however, it
is still inadequate for complete description of the dy-
namics of the mucosa. Titze (1973) has included the
vertical degrees of freedom in his 16-mass representa-
tion of the vocal cords, and has been successful in sim-
ulating the motions of the mucosa. A sketch of the
model proposed by Titze is shown in Fig. 6(b). How-
ever, since there are still only two masses in cross-
section, both types of vertical phasing cannot be achieved
simultaneously. It is therefore concluded that at least
three masses in cross section are necessary to account
for the shear in muscle tissues as well as the mucosal
coverings.

Gupta *et al.* (1973) have recently argued that, based
upon nonsteady air-flow analysis in the subglottal region,
a one-mass vibrator can oscillate only in a limited fre-
quency range, i.e., from 250 to 500 Hz, whereas the
two-mass vibrator can cover the 50—450-Hz range.
Based upon the present analysis, the authors of this
paper agree that additional degrees of freedom increase
the oscillatory range of the vocal cords, but it seems
unlikely that these additional degrees of freedom *lower*
the range. It is a well-known fact that the lowest mode
of any vibrational system is the one in which all dis-
placements are in phase. The fact that vertical phasing
occurs at low pitches and not at high ones results from
two factors. First, the vocal cord is thicker at low
pitches, resulting in a larger surface over which a pres-
sure gradient can exist, and second, the ligament is
lax, allowing for a higher degree of shear between var-
ious tissues near the glottal edge. If we were to assume
a hypothetical case in which the longitudinal tension re-
mains uniform throughout the depth of the cord as the
pitch is raised, and the thickness likewise remains con-
stant, a transition from $n_z = 0$ to $n_z = 1$ would appear to
raise the pitch, not lower it, as Gupta *et al.* suggest.
There are, however, some small pitch-lowering phe-
nomena associated with vertical phasing. Viscous losses
in the vocalis and various connective tissues do lower
the pitch slightly (Titze, 1975), especially the viscous
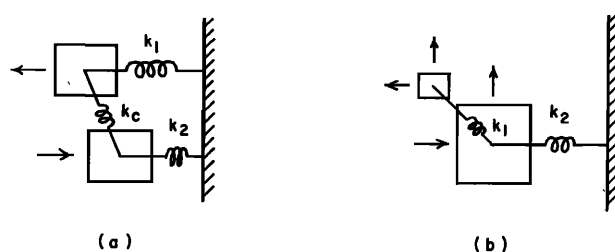drag associated with the motion of the mucosa. Since

(a)   (b)

FIG. 6. Vocal cord models which capture some features of the
natural responses of the vocal cords. (a) The Ishizaka—Flana-
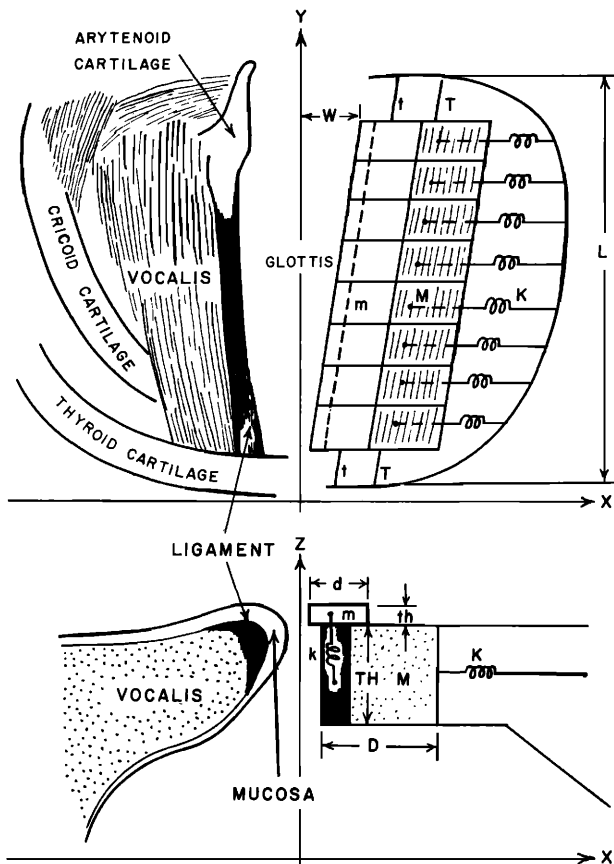gan model, and (b) the Titze model.

FIG. 7. A multimass model of the vocal cord allowing for eight longitudinal degrees of freedom.

the shear between the mucosa and the ligament is small, the effect of its motion is primarily one of energy loss, and hence, pitch reduction.

Neither the one-mass model nor the two-mass model are capable of representing horizontal differences along the length of the cord. This presents a number of limitations. The fixed boundary conditions at the anterior and posterior ends of the cord cannot be simulated. Hence, the frequently observed "oval" glottis never exists in the model. Furthermore, to accomodate registration, portions of the cord should be able to be "damped," i.e., fixed, and the prephonatory shape of the glottis should be triangular under some conditions. These features have been incorporated by Titze (1973; 1974) in his 16-mass representation of the vocal cord. Figure 7 illustrates some of the details of the model. Note that the prephonatory shape of the glottis can be controlled by the parameter W, which defines the position of the vocal processes, and hence, the degree of adduction. Longitudinal variations in vocal cord mass, stiffness, and boundary can be incorporated for simulation of localized abnormalities, i.e., lesions and growths. The number of degrees of freedom of this model is, of course, still finite, forcing an upper limit ($n_y = 8$) upon the normal-mode spectrum. An eigenanalysis of the 16-mass system has been carried out by Titze (1973). The results are very similar to those presented in Fig. 5.

VII. SUMMARY

This analysis has focused upon the small amplitude (normal mode) response of the vocal cord. Some assumptions were inherent in the treatment, and should be summarized at this point. A rectangular cross section of the cord was chosen. Most x-ray and tomographic studies indicate that this assumption is reasonable for low pitches. For higher pitches, on the other hand, a wedge-shaped vocal cord is the more common configuration. In addition, the vocal cord boundary should be more nearly circular, allowing for cylindrical modes in the outer portions of the tissue. Finally, in the exact mathematical analysis, uniform tissue properties were assumed over portions of the cord. The effects of the inhomogeneities arising from the ligament and mucosa were analyzed qualitatively. It is expected that, in future considerations, perturbations of this type can be included in a more quantitative fashion. Some vital clues to vocal fold behavior have been offered. In order to construct a complete model of vocal cord vibration, which must take into consideration the aerodynamic driving forces, the nonlinearities, and the loss mechanisms, it is imperative that the representation of the tissue itself be sufficiently complex to account for the possible, and only the possible, degrees of freedom. The elastic properties of the vocal cord have been related to the bulk properties of the tissues, thus allowing for measurements to be made on fragments of the tissues rather than on the whole system including the boundaries. Sonninen et al. (1972) have recently considered some of the stresses and strains—within vocal cord tissues—which might result from vocal abuse, giving rise to temporary or permanent damage. Providing that accurate data on longitudinal stress, Young's modulus, shear modulus, and viscosity of the tissues can be collected, a model based upon the compressional, shear, and viscous modes can predict the maximum strains, the energy loss per cycle, and the vocal efficiency. Finally, the normal modes, governing the transient responses of the vocal cords, must be invoked in the analysis of sudden phonatory attacks, "voice breaks," and other rapid manipulations of vocal constraints.

Behringer, S. (1955), "Die Anordnung der Muskulatur in der Menschlichen Stimmlippe und im Gebiet des Conus Elasticus," Z. Anat. Entwicklung. 118, 324–342.

Farnsworth, D. W. (1940). "High-Speed Motion Pictures of the Human Vocal Cords," Bell Lab. Rev. 18 (7), 203–208.

Flanagan, J. L., and Landgraf, L. (1968). "Self-Oscillating Source for Vocal Tract Synthesizers," IEEE Trans. Audio Electroacoust. AU-16, 57–64.

Goertler, K. (1950). "Die Anordnung, Histologie und Histogenese der Quergeschrieften Muskulatur im Menschlichen Stimmband," A. Anat. 115, 352–401.

Gupta, V., Wilson, T. A., and Beavers, G. S. (1973). "A Model for Vocal Cord Excitation," J. Acoust. Soc. Am. 54, 1607–1617.

Hast, M. H. (1966). "Physiological Mechanism of Phonation: Tension of the Vocal Fold Muscle," Acta Oto-Laryngol. 62, Fascile 4–5, 309–318.

Hirano, M. (1968). "The Vocal Cord During Phonation," Igaku Ayumi 80, No. 10.

Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of Voiced Sounds from a Two-Mass Model of the Vocal Cords,"

Bell Syst. Tech. J. 51, 1233–1267.

Oestreicher (1951). "Field and Impedance of an Oscillating Sphere in a Viscoelastic Medium with an Application to Biophysics," J. Acoust. Soc. Am. 23, No. 6, 707–714.

Rubin, H. J., and Hirt, C. C. (1960). "The Falsetto: A High-Speed Cinematographic Study," Laryngoscope 70, 1305–1324.

Sonninen, A. (1973). Personal correspondence.

Sonninen, A., Damste, P. H., Jol, J., and Fokkens, J. (1972). "On Vocal Strain," Folia Phoniatr. 24, 321–336.

Titze, I. R. (1973), "The Human Vocal Cords: A Mathematical Model. Part I," Phonetica 28, 129–170.

Titze, I. R. (1974), "The Human Vocal Cords: A Mathematical Model. Part II," Phonetica 29, 1–21.

Titze, I. R. (1975), "On Viscous Losses, Effective Mass, and Elasticity of the Human Vocal Cords" (to be published).

van den Berg, J. (1958). "Myoelastic–Aerodynamic Theory of Voice Production," J. Speech Hear. Res. 1, 227–244.

van den Berg, J. (1960), "Vocal Ligaments Versus Registers," Curr. Prob. Phoniatr. Logoped. 1, 19–34.

van den Berg, J., and Moll, J. (1955). "Zur Anatomie des Menschlichen Musculus Vocalis," Z. Anat. Entwicklung. 118, 465–470.

von Leden, H. (1961), "The Mechanism of Phonation," Arch. Otolaryngol. 74, 660–676.

Wustrow, F. (1952), "Bau und Function des Menschlichen Musculus Vocalis," Z. Anat. Entwicklung. 116, 506–522.