# A Potential-Function Parameterisation of Vowel Acoustics

Barbara J. Forbes* and E. R. Pike
*D.E.M.E., The Open University, Walton Hall, Milton Keynes MK7 6AA, UK; Email: forbes@phonologica.com

## Introduction

A physico-mathematical framework for the compact parameterisation of vowels is described in terms of the ‘acoustical Klein-Gordon’ equation, specifically by just two limiting cases on the wave-mechanical potential function, $U(x)$. It is shown that as few as 6 binary parameters, referring to the presence or absence of a perturbing potential function at the eigenfunction turning points, are sufficient to generate the formant shifts necessary for the simulation of a 25-vowel space.

## 1. Theory

Recent work by the authors[1,2] has been examining the reduced[3] form of the Webster equation for possible applications in speech parameterisation. We have noted that the second-order time derivative invokes an analogy with the Klein-Gordon equation of quantum mechanics, rather than with the Schrödinger. This ‘acoustical Klein-Gordon’ equation (eq. 1)

$$\frac{\partial^2\Psi(x,t)}{\partial t^2} = c^2\left\{\frac{\partial^2\Psi(x,t)}{\partial x^2} - U(x)\Psi(x,t)\right\}, \qquad \text{eq. 1}$$

is parameterised by the ‘potential function’, $U(x)$, which is defined as

$$U(x) = \frac{d^2S(x)^{1/2}/dx^2}{S(x)^{1/2}}. \qquad \text{eq. 2}$$

Under the simplifying assumption of circular symmetry, the potential function is found to be directly proportional to the tract curvature. It therefore takes a positive value in regions of expansion, for example, and a negative value in a constriction. For piecewise-constant potentials, such that $U(x)=U_0$, these potential-function types may be referred to as rectangular ‘barriers’ and ‘wells’, respectively, in analogy with the wave-mechanical literature. Straight tube and conical sections have a potential function of zero.

By rearranging eq. 2 as a second-order homogeneous differential equation, it can be shown that an infinite ‘equivalence set’ of area function solutions, of arbitrary radius and slope at the glottis, can be chosen to correspond to any given potential function. The resonance characteristics of the infinite set are identical to those of the single potential function, and so it can be concluded that the latter forms a more compact basis for the parameterisation of the tract geometry. Fig. 1 gives examples from the equivalence set of area functions corresponding to a barrier of 1mm width and $10^5$ m$^{-2}$ height. The tract is terminated in an infinite conical baffle of varying angle at the throat.

For piecewise-constant potentials, the time-independent eigenfunction solutions, $\psi(x,t)$, are in terms of a dispersive wavenumber, $k'$, where

$$k' = \sqrt{k^2 - U_0}. \qquad \text{eq. 3}$$

for $k$ the free-space wavenumber. Two limiting cases arise.

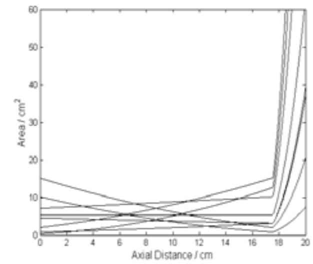


**Fig. 1: Examples from the equivalence set of a wave-mechanical barrier of width 1 mm at *x*=17.5 cm.**

### *1.1 The barrier potential*

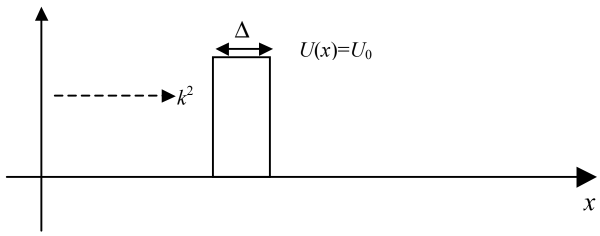


**Fig. 2. Schematic rectangular ‘barrier’ potential function.**

As $U_0$ approaches $k^2$, the dispersive wavenumber approaches zero and the phase velocity rises towards infinity. The limiting case $k^2<U_0$ describes a wave that decays exponentially, or is ‘evanescent’, over the barrier width, Δ. Reflections from the barrier edges are found to lead to strong cavity resonances[1,2]. For a typical maximum vocal-tract cross-section of around 15 cm$^2$, giving higher-mode cuton at around 5 kHz, this condition can be achieved over the complete plane-wave range by setting a barrier height of the order $10^4$ m$^{-2}$.

### *1.2 The well potential*

Helmholz resonances can be expected due to acoustical wells, particularly in the limit as $U_0 \rightarrow -\infty$ and the phase velocity tends towards zero. However, it can be shown that trapped ‘bound-state’ solutions in terms of $-k^2$ exist within the well, that increase to infinity with time[4] and violate principles of energy conservation. The presence of bound-state solutions are an indication that the standard physical assumptions made in the derivation of the Webster equation, particularly that of non-viscosity, must fail in the context of severe constrictions. They may, indeed, be modified by adopting a lossy (complex) wavenumber, which permits tunnelling of the trapped modes out of the well. However, it is known from the quantum-mechanical literature that bound states may be entirely avoided if the well potential is followed by an infinitely-high barrier (Fig. 3) in a ‘well-barrier pair’ configuration, given the condition that

$$|U_0|\,\Delta^2 \le \frac{\pi^2}{4}. \qquad \text{eq. 4}$$

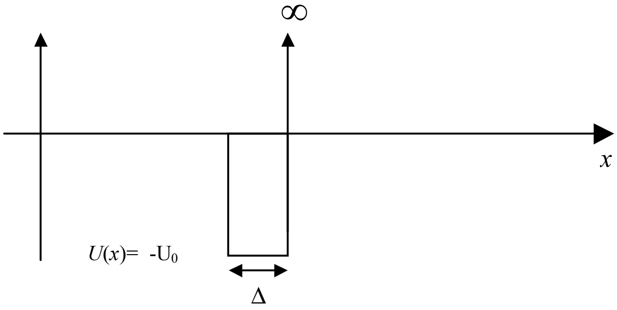


**Fig. 3. Schematic rectangular 'well-barrier pair' potential function.**

The theoretical limit on acoustical well parameters in the nonviscous approximation is, thus, that $|U_0|\Delta^2=\pi^2/4$.

The aim of the research was to investigate whether the formant shifts appropriate to a full vowel space could be generated from just these two limiting cases on potential functions. A time-independent perturbation analysis was conducted[5], and the turning points of the first three eigenfunctions, with the tract midpoint at $x=l/2$, were identified as positions where a perturbing potential function would produce the maximum effect on resonance. It can be noted that, in principle, the tract length $l$ is an arbitrary, renormalisable, quantity.

## 2. Numerical Simulation

The potential functions of section 1 were discretised along the tract according to the perturbation theory. The tract length was set to $l$=17.5 cm, for which the unperturbed quarter-wavelength eigenvalues fall at 500, 1,500, 2,500 ... Hz. The acoustical Klein-Gordon equation (eq. 1) was solved by finite differences for a boundary condition at the glottis set by the noninteractive Klatt 3-parameter polynomial model[1] ($f_0$ = 100 Hz) and for a wave-mechanical approximation to the infinite-baffle radiation impedance[6].

## 3. Results

In terms of both sound quality and spectra, it was found that certain combinations of potential functions produced aural effects that could be identified as vowel-like. Table 1 summarises these 25 potential-function 'strings' and their appropriate IPA labels. The notation is such that a '1' in a column means a perturbing potential function is present, '0' that there is no perturbation. The formant characteristics of the strings were plotted following the method of Ladefoged: fig. 4 shows that they span a classic vowel quadrilateral. A particularly interesting feature of the results is the characterisation of the main phonological groups by the presence or absence of specific 'bits' in the table. For example, the 1st bit characterises the mid and low vowels, the 2nd the low, and 3rd and 5th, the front; and the 6th the rounded vowels,.

## 4. Conclusions

It can be concluded that a vocal-tract parameterisation in terms of just two limiting cases on potential functions is sufficient to generate a full 25-vowel space and that the results may be of interest in speech recognition and compression. For simulation applications, improvements to the sound quality must be obtained by, for example, incorporating the effects of lossy propagation and a yielding wall, and adopting an interactive glottis model.

| 6-Bit Model: 25-Vowel System. | | | | | | |
|---|---|---|---|---|---|---|
| | Tract Position | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Vowel | $\frac{l}{10}$ | $\frac{2l}{5}$ | $\frac{l}{2}$ | $\frac{2l}{3}$ | $\frac{4l}{5}$ | $l$ |
| [i] | 0 | 0 | 1 | 0 | 1 | 0 |
| [y] | 0 | 0 | 1 | 0 | 1 | 1 |
| [ɨ] | 0 | 0 | 0 | 1 | 0 | 0 |
| [ɯ] | 0 | 0 | 1 | 1 | 0 | 0 |
| [u] | 0 | 0 | 0 | 1 | 0 | 1 |
| [ɪ] | 0 | 0 | 1 | 1 | 1 | 0 |
| [ʏ] | 0 | 0 | 1 | 1 | 1 | 1 |
| [ʊ] | 1 | 0 | 0 | 0 | 0 | 1 |
| [e] | 1 | 0 | 1 | 0 | 1 | 0 |
| [ø] | 1 | 0 | 1 | 0 | 1 | 1 |
| [ɵ] | 0 | 0 | 0 | 0 | 0 | 1 |
| [ɤ] | 1 | 0 | 1 | 1 | 0 | 0 |
| [o] | 1 | 0 | 1 | 1 | 0 | 1 |
| [ə] | 0 | 0 | 0 | 0 | 0 | 0 |
| [ɛ] | 1 | 0 | 1 | 1 | 1 | 0 |
| [œ] | 1 | 0 | 1 | 1 | 1 | 1 |
| [ɜ] | 1 | 0 | 0 | 1 | 0 | 0 |
| [ɞ] | 1 | 0 | 0 | 1 | 0 | 1 |
| [ɔ] | 1 | 1 | 1 | 1 | 0 | 1 |
| [æ] | 1 | 1 | 1 | 1 | 1 | 0 |
| [ɐ] | 1 | 0 | 0 | 0 | 0 | 0 |
| [a] | 1 | 1 | 0 | 0 | 0 | 0 |
| [ɶ] | 1 | 1 | 0 | 1 | 0 | 0 |
| [ɑ] | 1 | 1 | 1 | 1 | 0 | 1 |
| [ɒ] | 1 | 1 | 1 | 1 | 1 | 1 |

**Table 1: 6-bit vocal-tract model.**

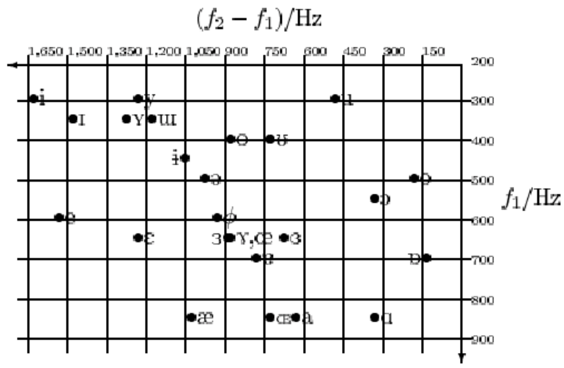


**Fig. 4: First and second formant characteristics of the 6-bit vocal-tract model.**

### *References*


[1] Forbes, B. J. A Potential-Function Analysis of Speech Acoustics. PhD Thesis, King's College London. 2000.

[2] Forbes, B. J., Pike E. R., and Sharp, D. B. The acoustical Klein-Gordon equation: the wave-mechanical step and barrier potential functions. Submitted to JASA.

[3] Benade A. H. and Jansson E. V. On plane and spherical horns with non-uniform flare. I. Theory of radiation, resonance frequencies, and mode conversion. Acustica 31, pp 79-98. 1974.

[4] Berners D.P. Acoustics and Signal Processing Techniques for Physical Modelling of Brass Instruments. PhD Thesis, Department of Electrical Engineering, Stanford University, CA. 1999.

[5] Forbes, B. J and Pike E. R. The acoustical Klein-Gordon equation: a time-independent perturbation analysis. Paper presented at the 144th Meeting of the Acoustical Society of America, Cancun, Mexico. 2002.

[6] Forbes, B. J. Piecewise-constant horn functions: the radiation impedance and end corrections in Schrödinger solutions of theWebster equation. Proc. International Seminar on Musical Acoustics, Perugia, Italy, 517-520. 2001.