

ОБРАБОТКА АКУСТИЧЕСКИХ СИГНАЛОВ.
КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

УДК 612.85

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ АКУСТИЧЕСКОГО РЕЧЕОБРАЗОВАНИЯ
С ПОДВИЖНЫМИ СТЕНКАМИ РЕЧЕВОГО ТРАКТА

© 2016 г. Н. А. Любимов, Е. В. Захаров

Московский государственный университет им. М.В. Ломоносова
119991 ГСП-1, Москва, Ленинские горы

E-mail: lubimov.nicolas@gmail.com

Поступила в редакцию 29.01.2015 г.

Рассматривается математическая модель речеобразования, описывающая распространение акустических колебаний в речевом тракте с подвижными стенками. Функция волнового поля удовлетворяет уравнению Гельмгольца с граничными условиями 3-го рода (импедансного типа). Режим импеданса соответствует трехпараметрической модели колебания маятника. Экспериментальные исследования демонстрируют нелинейный характер влияния подвижности стенок тракта на спектральную огибающую речевого сигнала.

Ключевые слова: модель речеобразования, уравнение Гельмгольца, импеданс.

DOI: 10.7868/S032079191602009X

1. ВВЕДЕНИЕ

Акустическая теория речеобразования имеет довольно долгую историю, начиная с наиболее ранних работ Гельмгольца [1], а также трудов Фанта [2] и Рабинера [3]. Подробный анализ ранних математических моделей дан в [4].

Целостное описание явлений распространения звуковых волн в речевом тракте человека является весьма сложным процессом, поскольку должны быть учтены многие эффекты, влияющие на характер порождаемого на выходе звукового поля. К таким эффектам относятся: переменная площадь поперечного сечения тракта, характер колебания голосовых связок, излучение при открытом ротовом отверстии, подвижность и проницаемость стенок речевого тракта. Простейшие и наиболее часто используемые на практике модели пользуются одномерным приближением для описания функции звукового давления или скорости потока частиц. К таким моделям можно отнести, например, уравнение Вебстера относительно волновой функции давления $p(z, t)$, где $0 \leq z \leq L$ – пространственная координата распространения плоской волны вдоль речевого тракта длины L , $t > 0$ – время и c – скорость звука:

$$\frac{\partial^2 p}{\partial t^2} = \frac{c^2}{A(z)} \frac{\partial}{\partial z} \left(A(z) \frac{\partial p}{\partial z} \right).$$

Колебания голосовых связок и излучение ротового отверстия задаются граничными режимами. Функция $A(z)$ представляет собой переменную площадь поперечного сечения цилиндра, внутри которого распространяются волны. Уравнение

Вебстера хорошо изучено и имеет аналитические решения для некоторых частных случаев формы речевого тракта [5, 6]. В более общем случае, например, когда функция $A(z)$ задана таблично, задача разрешается численно. Часто используемыми численными методами решения считаются методы, основанные на так называемой *схеме Келли–Локбаума* [9]. Данная схема использует представление Даламбера искомого решения в виде бегущих волн, рассматриваемого на дискретной пространственно-временной сетке:

$$p(z, t) = p(z - ct) + p(z + ct) \xrightarrow{z=i\Delta z, t=n\Delta t} \bar{p}_i(n) + \bar{p}_i(n).$$

Стрелками обозначены волны, бегущие в соответствующих направлениях. Данная схема обобщается для двумерного и трехмерного случаев моделирования речеобразования, в том числе в неоднородных средах распространения волн [10, 11]. Основными проблемами при работе с такими многомерными *цифровыми волноводами* (digital waveguide) являются жесткие ограничения, налагаемые на сетку пространственно-временной дискретизации, и сопутствующее им возрастание вычислительной сложности при переходе к многомерным случаям [12].

В других работах [13, 14] рассматриваются задачи, изначально учитывающие многомерность структуры речевого тракта, а следовательно, убирающие предположение о плоскости распространения волнового фронта. В работе [13] предложена двумерная модель, где роль области излучения играет полубесконечный цилиндр. Авторы сравнивают предлагаемую модель с описанным выше уравнением Вебстера. В обоих случаях для получения решения рассматривается численный ме-

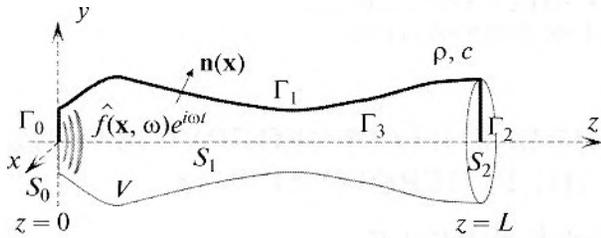


Рис. 1. Схематичное изображение речевого тракта человека от голосовой щели ($z = 0$) до ротового отверстия и области губ ($z = L$).

тод конечных элементов. Важным заключением, сделанным в данной работе, является экспериментально подтвержденный факт расхождения решения уравнения Вебстера с решением рассматриваемой двумерной задачи в области высоких частот. Авторы поясняют это тем, что для больших волновых чисел при заданной геометрии речевого тракта возникают поперечные моды, амплитуды которых нельзя полагать пренебрежимо малыми. Более сложная трехмерная модель речевого тракта рассмотрена в работе Ханнукайна и Луккари [14]. Акустическое поле в ограниченной области $\chi \in \Omega$ описывается уравнением колебаний

$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u$ относительно потенциала скоростей $u = u(\mathbf{x}, t)$. Особое внимание уделено граничному условию в области голосовых связок, которое задается в виде $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial n} = 2\sqrt{\frac{c}{\rho}}u$, где ρ – плотность среды.

В описанных выше моделях остается неучтенным характер влияния стенок речевого тракта, определяемый потерями энергии и подвижностью стенки под действием звукового давления. В обоих случаях вводится допущение о непроницаемости стенок, выражаемое в виде граничного условия в форме Неймана $\frac{\partial u}{\partial n} = 0$. В работе [7] продемонстрировано влияние нежестких стенок на изменение скорости звука. Также подвижность стенок речевого тракта была учтена в моделях генерации звука на примере автоколебательных систем, описанных в работе [8]. Приведенные уравнения задают процесс распространения волн в плоской трубе с открытым концом, где действие голосовых связок моделируется за счет мембраны в одном случае и поперечно колеблющихся пластин в другом случае.

В данной работе решается более общая задача, с использованием трехмерной модели речеобразования, подобной [14], однако учитывающей подвижность всей границы речевого тракта, соответствующей как голосовым связкам, так и непосредственно стенкам речевого тракта. При этом, помимо гладкости, не накладывается никаких других ограничений на форму данной границы. Коле-

бания границы также носят достаточно общий характер, описываемый импедансными граничными условиями.

2. ПОСТАНОВКА ЗАДАЧИ

В качестве основных компонентов рассматриваются следующие органы речевого тракта:

- голосовые связки и прилегающая к ним область;
- боковые стенки речевого тракта и язык, образованные мягкими тканями человеческого тела;
- ротовое отверстие, которое может быть либо открыто (как в случае произнесения гласных звуков), либо закрыто (для некоторых взрывных согласных, таких как “б”);
- полость речевого тракта (гортань), заполненная воздухом.

В данном приближении исключены некоторые другие внутренние органы, играющие роль в речевом образовании. Действие этих органов неявно включено в граничные условия рассматриваемой задачи. Например, легкие и подвязочная область учитываются при описании звукового поля на голосовой щели. Носовая полость не учитывается, поскольку она играет существенную роль только для сонорных согласных (“н”, “м”). В остальных случаях небная занавеска закрыта, и таким образом замыкает поверхность речевого тракта. Действие губ можно также приближенно включить в описание поля на границе ротового отверстия.

На схематичном рис. 1 изображена область, участвующая в процессе речеобразования в рамках текущей модели. В соответствии с вышеперечисленными компонентами речевого тракта, эта область задается следующим образом:

- S_0 – граница со стороны голосовой щели;
- S_1 – граница, определяемая стенками речевого тракта;
- S_2 – виртуальная граница, описывающая ротовое отверстие;
- V – внутренняя область речевого тракта.

Полная поверхность речевого тракта образована сочленением частичных границ $S = S_0 \cup S_1 \cup S_2$. Вектор нормали $\mathbf{n} = \mathbf{n}(\mathbf{x})$ является внешним по отношению к области V .

Звуковые колебания порождаются потоком воздуха, выходящим из легких и проходящим через голосовую щель. Воздушный поток модулируется колебанием голосовых складок, который имеет квазипериодический характер. С некоторой точностью приближения можно говорить, что в области S_0 присутствует постоянный источник возбуждения вида $\hat{f}(\omega)e^{i\omega t}$, где $\hat{f}(\omega) \in C$ – спектральный профиль возбуждения. Конкретный вид амплитудной огибающей источника будет рассмотрен ниже в разделе 4.

В данной модели свойства поверхности речевого тракта задаются через импеданс, связывающий давление и нормальную составляющую скорости потока:

$$\hat{z}(\mathbf{x}, \omega) = \frac{\hat{p}(\mathbf{x}, \omega)}{\hat{\mathbf{v}}(\mathbf{x}, \omega) \cdot \mathbf{n}(\mathbf{x})}. \quad (1)$$

Нежесткость стенок речевого тракта приводит к тому, что они начинают колебаться под действием звукового давления на поверхности стенок. Предполагается, что данный процесс является локальным, т.е. каждое достаточно малое сечение поверхности стенок совершает колебания независимо от соседних. Иными словами, имеет место модель *локально реактивной поверхности* (locally reacting surface). Обзор подобных моделей можно найти, например, в [15]. Для текущей модели выбрано уравнение движения маятника с затуханием:

$$m \frac{d^2 \zeta}{dt^2} + R \frac{d\zeta}{dt} + K \zeta(\mathbf{x}, t) = p(\mathbf{x}, t). \quad (2)$$

Модель, формулируемая относительно движения элемента границы $\zeta = \zeta(\mathbf{x}, t)$, является трехпараметрической с параметрами: m – погонной массой элемента, R – коэффициентом сопротивления и K – коэффициентом упругости элемента стенки речевого тракта. В частотной области уравнение (2) приводит к выражению для импеданса стенок тракта следующего вида:

$$\hat{z}_b(\omega) = R + i \left(m\omega - \frac{K}{\omega} \right). \quad (3)$$

Отражение волн на открытой границе S_2 (ротовое отверстие) отличается от внутренней поверхности речевого тракта. Уравнение для соответствующего импеданса описывается через низкочастотную аппроксимацию излучения открытой трубы с бесконечным плоским фланцем [16]:

$$\hat{z}_a(\omega) = \rho c \left(\frac{(kr_L)^2}{2} + ikr\delta \right), \quad (4)$$

где $\delta = 0.8216$ – так называемая *поправка длины трубы на открытый конец* (open end length correction), а r_L – радиус ротового отверстия. Несложно показать, что при малом значении радиуса r_L по сравнению с радиусом головы формула (4) эквивалентна более общему представлению, рассмотренному в работе [17].

Непосредственно сама постановка задачи формулируется относительно потенциала скоростей в спектральной области $u(\mathbf{x}, \omega)$, который связан с давлением и скоростью потока соотношениями

$$\hat{\mathbf{v}}(\mathbf{x}, \omega) = \nabla \hat{u}(\mathbf{x}, \omega), \quad (5)$$

$$\hat{p}(\mathbf{x}, \omega) = i\omega \rho \hat{u}(\mathbf{x}, \omega). \quad (6)$$

Имеет место внутренняя задача Гельмгольца с граничными условиями 3-го рода:

$$\Delta u(\mathbf{x}, \omega) + k^2 u(\mathbf{x}, \omega) = 0, \quad \mathbf{x} \in V, \quad (7)$$

$$\frac{\partial u}{\partial n}(\mathbf{x}, \omega) + \alpha(\mathbf{x}, \omega) u(\mathbf{x}, \omega) = f(\mathbf{x}, \omega), \quad \mathbf{x} \in S, \quad (8)$$

где функция $\alpha(\mathbf{x}, \omega)$ определена как

$$\alpha(\mathbf{x}, \omega) = \begin{cases} 0, & \mathbf{x} \in S_0, \\ ik/\hat{z}_b(\omega), & \mathbf{x} \in S_1, \\ ik/\hat{z}_a(\omega), & \mathbf{x} \in S_2, \end{cases} \quad (9)$$

а функция $f(\mathbf{x}, \omega)$ – как

$$f(\mathbf{x}, \omega) = \begin{cases} \hat{f}(\omega), & \mathbf{x} \in S_0, \\ 0, & \mathbf{x} \in S_1 \cup S_2. \end{cases} \quad (10)$$

Относительно каждой частоты ω постановка задачи не меняется, поэтому в дальнейшем эту переменную можно исключить, полагая $u(\mathbf{x}) = u(\mathbf{x}, \omega)$.

Как известно, справедливо интегральное представление решения для $\mathbf{x} \in V$ [19]:

$$u(\mathbf{x}) = \int_S \frac{\partial u}{\partial n_y}(\mathbf{y}) G(\mathbf{x}, \mathbf{y}) - u(\mathbf{y}) \frac{\partial G}{\partial n_y}(\mathbf{x}, \mathbf{y}) dy \quad (11)$$

с использованием фундаментального решения в виде функции Грина $G(\mathbf{x}, \mathbf{y}) = \frac{e^{ikR(\mathbf{x}, \mathbf{y})}}{4\pi R(\mathbf{x}, \mathbf{y})}$, где $R(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ – евклидово расстояние между некоторыми точками \mathbf{x} и \mathbf{y} . Исходная постановка задачи (7) может быть сведена к интегральному уравнению на границе $\mathbf{x} \in S$ [20]:

$$\begin{aligned} \Omega(\mathbf{x})u(\mathbf{x}) + \int_S \left(\frac{\partial G}{\partial n_y}(\mathbf{x}, \mathbf{y}) + \alpha(\mathbf{y})G(\mathbf{x}, \mathbf{y}) \right) u(\mathbf{y}) ds_y = \\ = \int_S G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) ds_y \end{aligned} \quad (12)$$

и функцией $\Omega(\mathbf{x})$ для внутренней задачи

$$\Omega(\mathbf{x}) = - \int_S \frac{\partial}{\partial n_y} \left(\frac{1}{4\pi R(\mathbf{x}, \mathbf{y})} \right) dy, \quad (13)$$

равной величине телесного угла, образованного касательными к поверхности в точке \mathbf{x} . В частности, для регулярной поверхности $\Omega(\mathbf{x}) = \frac{1}{2}$.

В операторном виде интегральное уравнение можно записать в виде

$$\mathbb{I}u + \mathbb{A}[\alpha u] + \mathbb{B}u = \mathbb{A}[f]. \quad (14)$$

Существование и единственность решения для интегрального уравнения подобного типа основывается на альтернативе Фредгольма. Однако в данном случае ее нельзя непосредственно применять в силу того, что функции α и f не принадлежат классу непрерывных на границе функций $C(S)$. Разрывы функций импеданса и источника при переходе границ голосовые складки (S_0) – стенки тракта (S_1) – ротовое отверстие (S_2) можно устранить при помощи, например, линейной интерполяции. В этом случае оператор $\mathbb{I} + \mathbb{A}\alpha + \mathbb{B}$ является компактным, и его ядро содержит только нулевое значение $\ker(\mathbb{I} + \mathbb{A}\alpha + \mathbb{B}) = \ker(\mathbb{I} + \mathbb{B}) = \{0\}$ почти для всех волновых чисел k , за исключением

тех, где нарушается единственность решения однородной внутренней задачи Неймана [20]. На практике соответствующие им частоты $\omega = kc$ можно исключить из рассмотрения, ограничившись поиском решения в прилегающей области конкретной резонансной частоты.

3. МЕТОД РЕШЕНИЯ ИНТЕГРАЛЬНОГО УРАВНЕНИЯ

Для поиска решения задачи (12) удобно перейти к цилиндрической системе координат:

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, \quad 0 \leq r \leq r_L, \\ \theta &= \arctg \frac{y}{x}, \quad 0 \leq \theta < 2\pi, \\ z &= z, \quad 0 \leq z \leq L, \end{aligned}$$

в терминах которой расстояние между точками $\mathbf{x} = (r_x, \theta_x, z_x)$ и $\mathbf{y} = (r_y, \theta_y, z_y)$ записывается в виде

$$R(\mathbf{x}, \mathbf{y}) = \sqrt{r_x^2 + r_y^2 - 2r_x r_y \cos(\theta_x - \theta_y) + (z_x - z_y)^2}.$$

Вводится условие об осесимметричности модели речевого тракта, где осью симметрии является прямая $r = 0, -\infty < z < +\infty$. Данное условие позволяет рассмотреть поверхность S в виде тела вращения с образующим контуром $\gamma \in \Gamma: r = r(\gamma), z = z(\gamma)$, включающем также ось симметрии $r = 0$ (см. жирную линию на рис. 1). Аналогично разбиению поверхности, контур также образован сочленением трех частей кривых, задающих разные области речевого тракта, вдобавок вводится четвертая часть кривой, отвечающая оси симметрии: $\Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$. На данной кривой действует дополнительное граничное условие симметрии $\frac{\partial u}{\partial n} = 0$, которое, однако, не учитывается при построении интегрального уравнения, что может быть продемонстрировано непосредственным дифференцированием и постановкой значения $r = 0$.

Будем искать решение в виде

$$u(r, \theta, z) = \sum_{m=0}^{\infty} u_m(r, z) \cos m\theta.$$

Относительно неизвестных коэффициентов $u_m = u(r_x, z_x)$ в точке $\theta_x = 0$, с учетом гладкой поверхности вращения, интегральное уравнение (12) записывается в явном виде:

$$\begin{aligned} \frac{1}{2} u_m + \int_{\Gamma_0} B_m u_m J d\gamma_y + \int_{\Gamma_1} \left(\frac{ik}{\hat{z}_b} A_m + B_m \right) u_m J d\gamma_y + \\ + \int_{\Gamma_2} \left(\frac{ik}{\hat{z}_a} A_m + B_m \right) u_m J d\gamma_y = f \int_{\Gamma_0} B_m J d\gamma_y \end{aligned} \quad (15)$$

с функциями

$$A_m = A_m(r_x, z_x, r_y, z_y) = \int_0^{2\pi} G(\mathbf{x}, \mathbf{y}) \cos m\theta_y d\theta_y, \quad (16)$$

$$\begin{aligned} B_m = B_m(r_x, z_x, r_y, z_y) = \\ = \int_0^{2\pi} \frac{\partial G}{\partial r_y} \frac{dz}{d\gamma_y} - \frac{\partial G}{\partial z_y} \frac{dr}{d\gamma_y} \cos m\theta_y d\theta_y \end{aligned} \quad (17)$$

и якобианом $J = J(\gamma) = r_y \left\| \frac{dr}{d\gamma}, \frac{dz}{d\gamma} \right\| = r_y \sqrt{\left(\frac{dr}{d\gamma_y} \right)^2 + \left(\frac{dz}{d\gamma_y} \right)^2}$. Интеграл по границе Γ_3 обращается в 0 и поэтому не вносит вклад в решение задачи.

Далее рассматривается численный метод, сводящий интегральное уравнение (15) к системе линейных алгебраических уравнений. Необходимо также учесть, что подынтегральные значения интегралов (16), (17) обращаются в бесконечность в точке $\mathbf{x} = \mathbf{y}$.

3.1. Метод Галеркина

Для дискретизации задачи (15) используется метод Галеркина с линейными элементами. Кривая Γ аппроксимируется кусочно-линейными функциями, связывающими N узлов сетки $u_i = u(r_i, z_i), i = 1, 2, \dots, N$ так, что значение для некоторой точки $r_i \leq r(\gamma) \leq r_{i+1}, z_i \leq z(\gamma) \leq z_{i+1}$ точки представимо в виде

$$u(\gamma) = u_i(1 - \gamma) + u_{i+1}\gamma, \quad (18)$$

$$r(\gamma) = r_i(1 - \gamma) + r_{i+1}\gamma, \quad z(\gamma) = z_i(1 - \gamma) + z_{i+1}\gamma,$$

где γ локально параметризует кривую, соединяющую точки (r_i, z_i) и (r_{i+1}, z_{i+1}) и изменяется в пределах $0 \leq \gamma \leq 1$. Якобиан в этом случае становится равен $J(\gamma) = r(\gamma) \sqrt{(r_{i+1} - r_i)^2 + (z_{i+1} - z_i)^2} = r(\gamma) \sqrt{\Delta r_i^2 + \Delta z_i^2} = r(\gamma) \Delta_i$, а внешняя нормаль в произвольной точке внутри отрезка равна $\mathbf{n} = \frac{1}{\Delta_i} (\Delta z_i, -\Delta r_i)$.

Рассматривая вместе все узлы сетки, интегральное уравнение (14) в компактной записи можно свести к системе линейных алгебраических уравнений:

$$\begin{aligned} \frac{1}{2} u_i + \sum_{j=1}^N \mathbf{A}_{ij} [\alpha u] + \\ + \sum_{j=1}^N \mathbf{B}_{ij} [u] = \sum_{j=1}^N \mathbf{A}[f], \quad i = 1, 2, \dots, N, \end{aligned} \quad (19)$$

где значения операторов могут быть найдены путем подстановки выражения неизвестной функции (18) в j -м узле в соответствующие интегралы:

$$A_{ij}[\psi] = \Delta_j \int_0^1 A(r_i, z_i, r(\gamma), z(\gamma)) \psi(r(\gamma), z(\gamma)) r(\gamma) d\gamma, \quad (20)$$

$$B_{ij}[\psi] = \Delta_j \int_0^1 B(r_i, z_i, r(\gamma), z(\gamma)) \psi(r(\gamma), z(\gamma)) r(\gamma) d\gamma. \quad (21)$$

3.2. Эллиптические интегралы

Подынтегральные выражения в (16) и (17) содержат функции, непрерывные всюду, кроме точек $R = R(x, y) = 0$, где функция обращается в бесконечность и, следовательно, не интегрируема через квадратурные формулы. Чтобы избавиться от этого эффекта, предлагается выделить сингулярную часть интегралов, заменив их табличным значением известной функции, представимой в виде эллиптических интегралов [21]. Раскладывая непрерывную функцию e^{ikR} в окрестности точки $R = 0$ в ряд Тейлора

$$e^{ikR} = 1 + ikR + \frac{(ikR)^2}{2!} + \frac{(ikR)^3}{3!} + \dots$$

и разделив экспоненту на R , при $R \rightarrow 0$ получается следующее выражение:

$$\frac{e^{ikR}}{R} = \frac{1}{R} + ik + \frac{(ik)^2 R}{2!} + \frac{(ik)^3 R^2}{3!} + \dots$$

Ряд состоит из членов, непрерывных в точке $R = 0$, за исключением первого члена $\frac{1}{R}$. В случае нормальной производной будет содержаться член вида $\frac{\partial}{\partial n} \left(\frac{1}{R} \right)$, имеющий, однако, тот же порядок сингулярности, что и в случае отсутствия производной. Части интегралов (16), (17), не содержащие рассмотренную сингулярность, интегрируются численно с помощью квадратурных формул Гаусса [22]. Оставшаяся часть вычисляется аналитически с использованием формул

$$\int_0^{2\pi} \frac{\cos m\theta}{4\pi R} d\theta = \frac{(-1)^m \chi_m(\zeta)}{\pi \sigma}, \quad (22)$$

$$\int_0^{2\pi} \frac{\partial}{\partial n} \left(\frac{\cos m\theta}{4\pi R} \right) d\theta = \frac{(-1)^m \sigma}{\pi} \frac{d\chi_m}{d\zeta} \frac{\partial \zeta}{\partial n} - \frac{\chi_m}{\zeta^2} \frac{\partial \sigma}{\partial n}, \quad (23)$$

где введены следующие обозначения:

$$\chi_m(\zeta) = \sum_{n=0}^m a_{2n} I_{2n}(\zeta), \quad (24)$$

$$\sigma(x, y) = \sqrt{(r_x + r_y)^2 + (z_x - z_y)^2}, \quad (25)$$

$$\zeta(x, y) = \sqrt{\frac{4r_x r_y}{(r_x + r_y)^2 + (z_x - z_y)^2}}. \quad (26)$$

Постоянные a_n представляют собой коэффициенты полинома Чебышева. Функции $I_n(\zeta)$ выражаются через полные эллиптические интегралы

1-го $K(\zeta) = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - \zeta^2 \sin^2 \varphi}}$ и 2-го рода $E(\zeta) = \int_0^{\frac{\pi}{2}} \sqrt{1 - \zeta^2 \sin^2 \varphi} d\varphi$:

$$I_0(\zeta) = K(\zeta), \quad (27)$$

$$I_2(\zeta) = \frac{E(\zeta) - (1 - \zeta^2)K(\zeta)}{\zeta^2}, \quad (28)$$

$$I_n(\zeta) = \frac{(2\zeta^2 - 1)I_{n-2} - (n-3)(1 - \zeta^2)I_{n-4}}{(n-1)\zeta^2}. \quad (29)$$

4. КОМПЬЮТЕРНЫЙ ЭКСПЕРИМЕНТ

Целью экспериментального анализа ставится демонстрация влияния подвижности стенок на спектральный состав передаточной функции речевого тракта, рассчитываемой по формуле

$$\hat{h}(\omega) = \frac{\hat{v}(\mathbf{x}_{out}, \omega) \cdot \mathbf{n}(\mathbf{x}_{out})}{\hat{v}(\mathbf{x}_{in}, \omega) \cdot \mathbf{n}(\mathbf{x}_{in})} = \omega \rho \frac{|u(\mathbf{x}_{out}, \omega)|}{|\hat{z}(\mathbf{x}_{out}) \hat{f}(\omega)|}, \quad (30)$$

где \mathbf{x}_{in} , \mathbf{x}_{out} – входная и выходная точка наблюдения. В цилиндрической системе координат принимается $\mathbf{x}_{out} = (0, 0, L)$.

Различные наборы параметров m (погонная масса), R (коэффициент сопротивления) и K (коэффициент упругости) соответствуют различным анатомическим характеристикам речевого тракта [5]. Вводится допущение, что речевой тракт со всех сторон окружен мягкими мышечными тканями. Это можно считать справедливым для области гортани, языка, щек и губ, но задает погрешность моделирования в области твердого неба. Плотность мышечной ткани человека полагается равной $\rho_w = 1060 \text{ кг/м}^3$. Если принять, что толщина стенок h_w в среднем находится в диапазоне от 0.005 до 0.02 м, то погонная масса варьируется в пределах $5.3 \leq m = \rho_w h_w \leq 21.2 \text{ кг/м}^2$. Согласно оценкам [5, 7], коэффициент сопротивления $R \approx (0.8 \dots 1.2) \times 10^4 \text{ кг/м}^2 \text{ с}$, а коэффициент упругости связан с толщиной стенки через модуль Юнга $K = \frac{E_w}{h_w}$, равный $E_w = 10^4 \dots 10^5 \text{ кг/м}^2$, откуда диапазон изменений для коэффициента упругости $K \approx (5 \dots 20) \times 10^5 \text{ кг/м}^2 \text{ с}^2$.

Ниже изложены результаты трех экспериментов по определению зависимости передаточной функции речевого тракта от изменения трех вышеизложенных параметров модели импеданса стенок. Решения удобно рассматривать при пол-

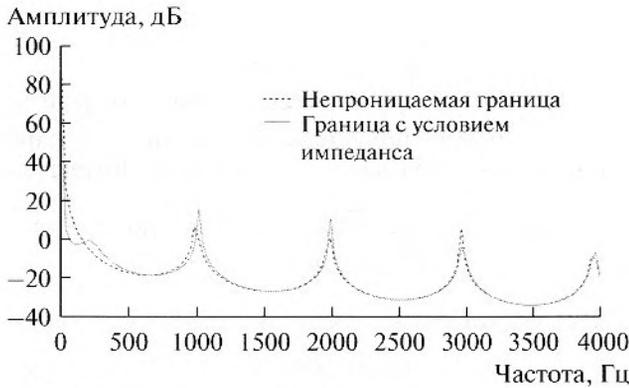


Рис. 2. Передаточная функция решения, полученного для непроницаемых стенок (условия Неймана) и с условием импеданса с коэффициентом сопротивления $R \approx 0.8 \times 10^4$ кг/м² с.

ностью закрытом ротовом отверстии, т.е. при выполнении граничного условия 3-го рода на всей границе, поскольку в этом случае вся энергия падающей волны распределена по поверхности речевого тракта.

Для начала зафиксируем реактивную (мнимую) часть импеданса и будем варьировать резистивную компоненту, которая не зависит от частоты сигнала, а определяется только одним параметром R . На рис. 2 проиллюстрированы спектральные профили передаточной функции для случаев жестких и подвижных стенок тракта с низким порогом коэффициента сопротивления $R \approx 0.8 \times 10^4$ кг/м² с. Помимо перераспределения энергии в области резонансных частот, наглядный анализ показывает, что при введении конечного импеданса возникает до-

полнительный резонанс на частоте примерно 200 Гц. Этот резонанс порождается вынужденными колебаниями стенок речевого тракта, и его значение соответствует результатам экспериментов, проведенных в работе [23].

Рассматривая только область низких частот передаточной функции, можно исследовать поведение резонансной области при изменении коэффициента сопротивления. Данный пример продемонстрирован на рис. 3. Поскольку изменение коэффициента сопротивления R влияет на скорость затухания колебаний (2), то при больших значениях наблюдается исчезновение характерного низкочастотного резонанса, как если бы стенки были абсолютно жесткими.

Характер изменения передаточной функции проанализирован также для переменной толщины стенок речевого тракта. Результаты моделирования изображены на рис. 4. Отчетливо видно возникновение пика в области низких частот. Это связано с понижением механического резонанса

колебаний стенок $F_w = \frac{1}{2\pi} \sqrt{\frac{K}{m}}$ в область собственной частоты закрытого речевого тракта 0 Гц. Колебания звукового поля на частоте, равной механическому резонансу колебаний стенок, приводят к тому, что реактивная компонента импеданса обращается в ноль. Стоит также обратить внимание на смещение самих резонансов трубы при изменении массы стенок, что согласуется с исследованиями изменения скорости звука, приведенными в работе [23], а позднее развитыми в работах [5, 7].

Синтез гласных, применяемый в данном эксперименте, состоит из нескольких шагов. Снача-

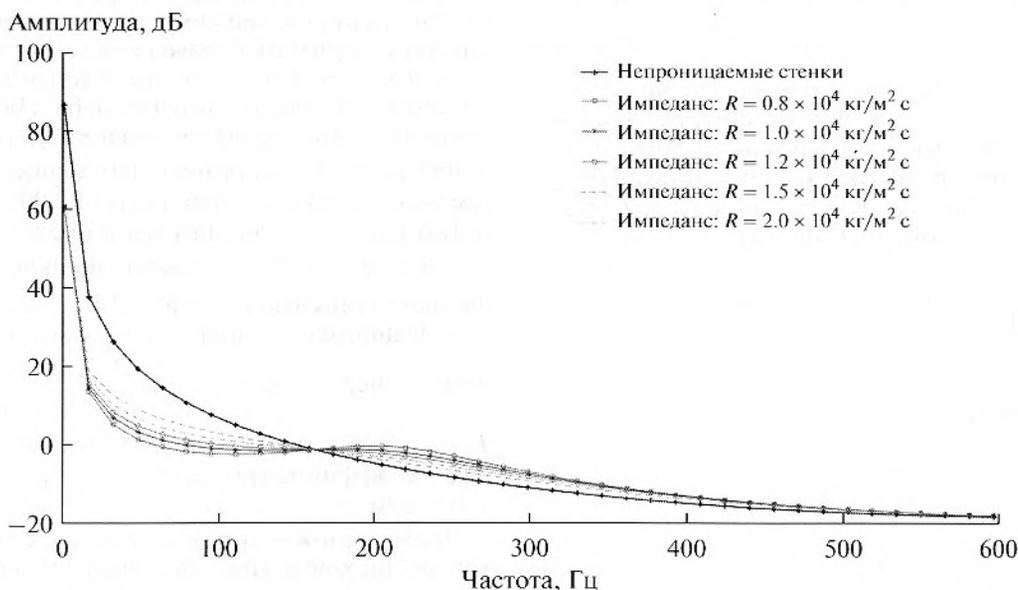


Рис. 3. Передаточная функция решения в области низких частот для различных значений коэффициента сопротивления R .

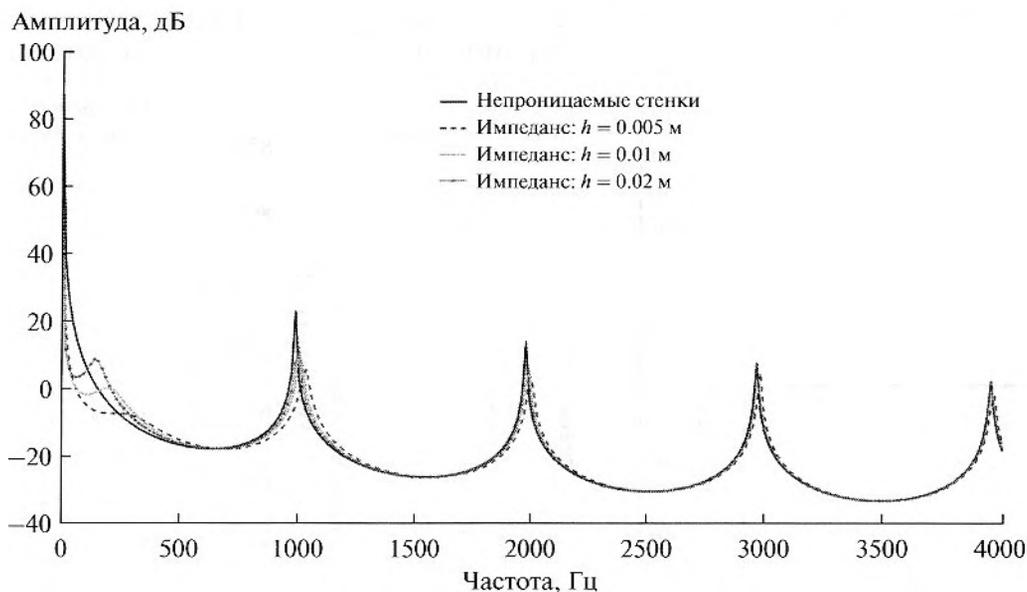


Рис. 4. Передаточная функция решения в области низких частот для различной толщины стенок речевого тракта h .

ла на основе данных о строении тракта вычисляется амплитудно-частотная характеристика речевого тракта (30).

Синтез гласных осуществляется на основе модели источник–фильтр (source–filter) [3, 4]. Функцию фильтра выполняет система речевого тракта, а источником звуковых колебаний являются голосовые складки. Для экспериментального анализа была выбрана широко известная модель Лильенкранца–Фанта, или LF-модель [18, 24]:

$$f_{LF}(t) = \begin{cases} E_0 e^{\alpha t} \sin(\omega_g t), & 0 \leq t \leq T_e, \\ -\frac{E_e}{\varepsilon T_a} \left(e^{-\varepsilon(t-T_e)} - e^{-\varepsilon(T_c-T_e)} \right), & T_e \leq t \leq T_c, \\ 0, & T_c \leq T_0, \end{cases} \quad (31)$$

где T_0 – период основного тона голоса, $\omega_g = \frac{\pi}{T_p}$,

параметры T_p, T_e, T_c, T_a задают временной характер колебания складок, а E_0 и E_e определяют амплитуду колебаний. Остальные параметры (α, ε) подбираются исходя из условия непрерывности и требования

$$\int_0^{T_0} a_{LF} dt = 0 \quad (\text{см. детали в работах [24, 25]}).$$

Комплексный спектр модели (31) выписывается в явном виде:

$$\begin{aligned} \hat{f}_{LF}(\omega) = & \frac{E_0}{(\alpha - i\omega)^2 + \omega_g^2} \times \\ & \times \left(e^{(\alpha - i\omega)T_e} \left((\alpha - i\omega) \sin(\omega_g T_e) - \omega_g \cos(\omega_g T_e) \right) + \right. \\ & \left. + \frac{E_e e^{-i\omega T_e}}{\varepsilon T_a i\omega (\varepsilon + i\omega)} \left(\varepsilon (1 - \varepsilon T_a) \left(1 - e^{-i\omega(T_0 - T_e)} \right) - \varepsilon T_a i\omega \right) \right). \end{aligned} \quad (32)$$

Если спектральная характеристика обладает свойством квазистационарности, то речевой сигнал может быть получен методом фильтрации сигнала источника:

$$s(t) = \int h(t - \tau) f_{LF}(\tau) d\tau, \quad (33)$$

где $h(t)$ – передаточная функция (30) во временной области. С учетом того, что известно частотное представление спектральной функции (30), используется более эффективный подход, состоящий в перемножении кратковременного преобразования фурье-сигнала со спектральной функцией-фильтром и последующем восстановлении временных отсчетов при помощи широко известного метода *перекрестного сложения* (overlap-add method) [26].

Целью дальнейшего вычислительного эксперимента является синтез гласных русского языка “а”, “у”, “и” с учетом известной геометрии речевого тракта в момент произношения [27]. Неформальным критерием оценки качества решения являются также аудиальные тесты на предмет разборчивости синтезируемых гласных.

Передаточная функция речевого тракта рассчитывается по формуле (30). Для каждого гласного звука было построено численное решение передаточной функции, состоящее из 256 частотных точек. Ниже на рис. 5–8 продемонстрированы результаты экспериментов: геометрические параметры речевого тракта (слева), спектральная огибающая звука при закрытом речевом тракте (по центру), спектральная огибающая при открытом речевом тракте (справа).

Из графиков наглядно видно, что как в случае закрытого, так и открытого ротового отверстия подвижность стенок речевого тракта вводит не-

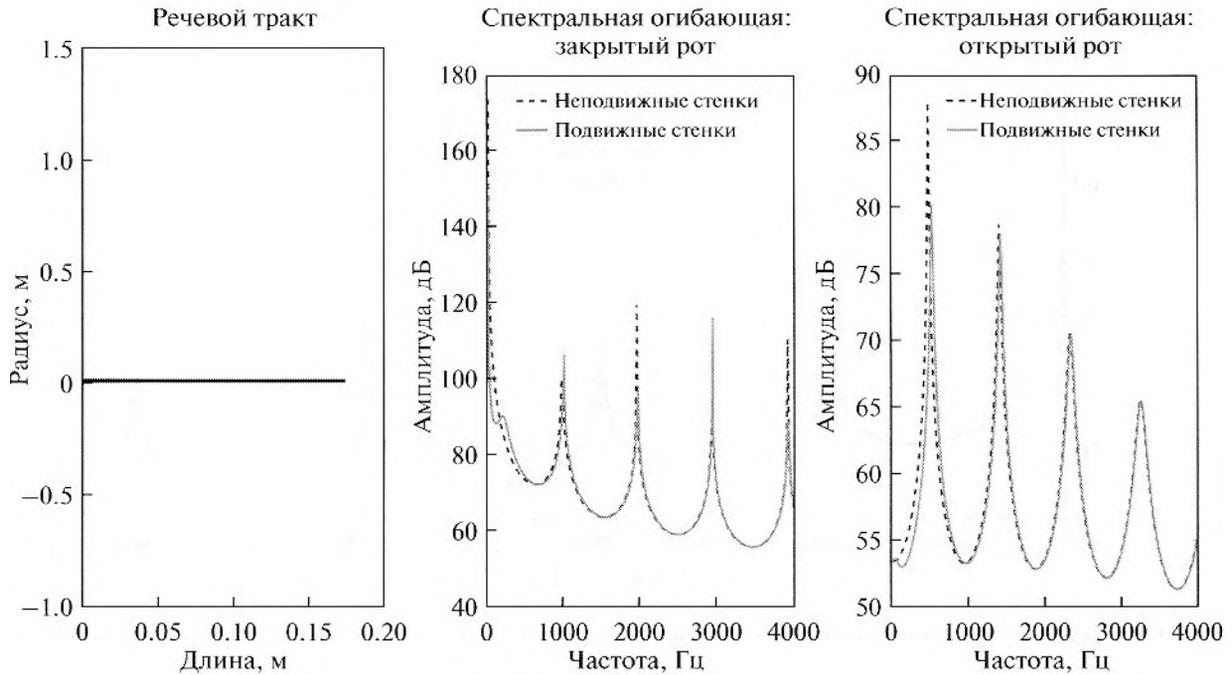


Рис. 5. Плоская труба.

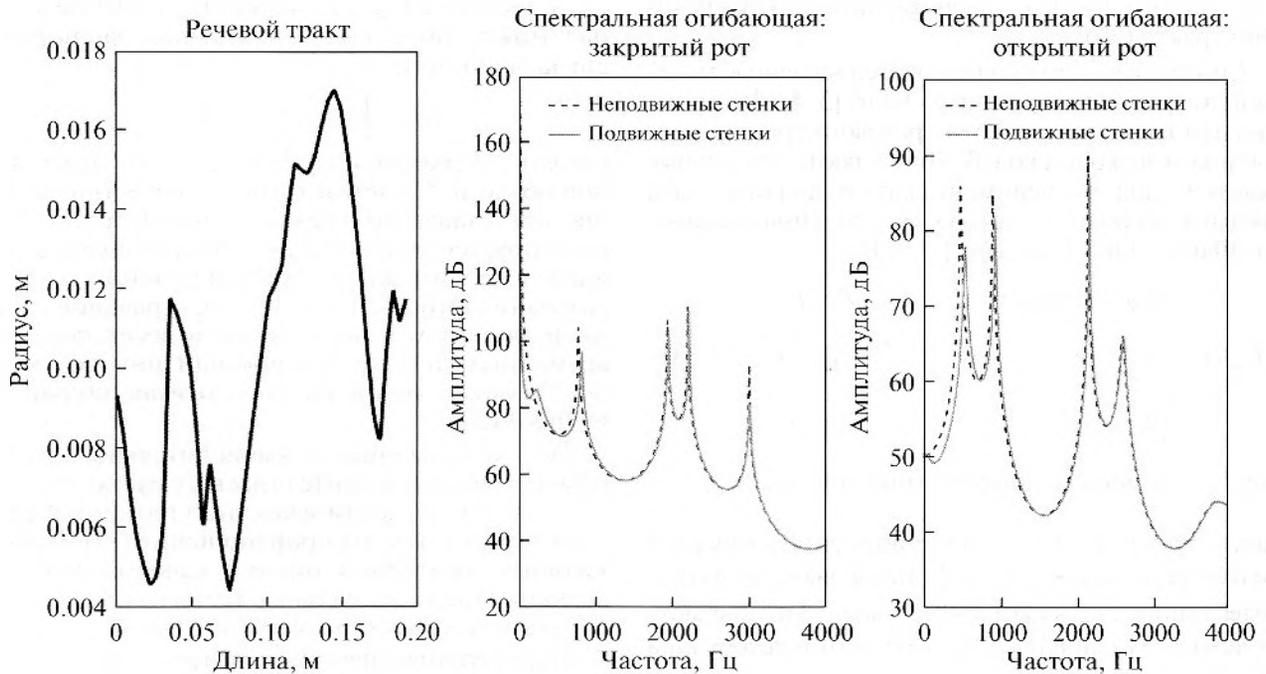


Рис. 6. Фонема "а".

линейные искажения спектральной огибающей сигнала вне зависимости от типа гласных звуков. Основными типами нелинейных преобразования спектра является возникновение резонанса в области низких частот, а также смещение формант, которое также уменьшается с ростом частоты.

Полученные спектральные профили гласных позволяют синтезировать соответствующие им во

времени сигналы. В рамках неформальных аудиальных тестов с несколькими экспериментаторами было выявлено улучшение качества сигнала, построенного с применением описанных в данной работе интегральных уравнений. В качестве референсных результатов предлагалось прослушать аналогичные гласные, синтезированные при помощи модели Келли—Локбаума [9], а также при

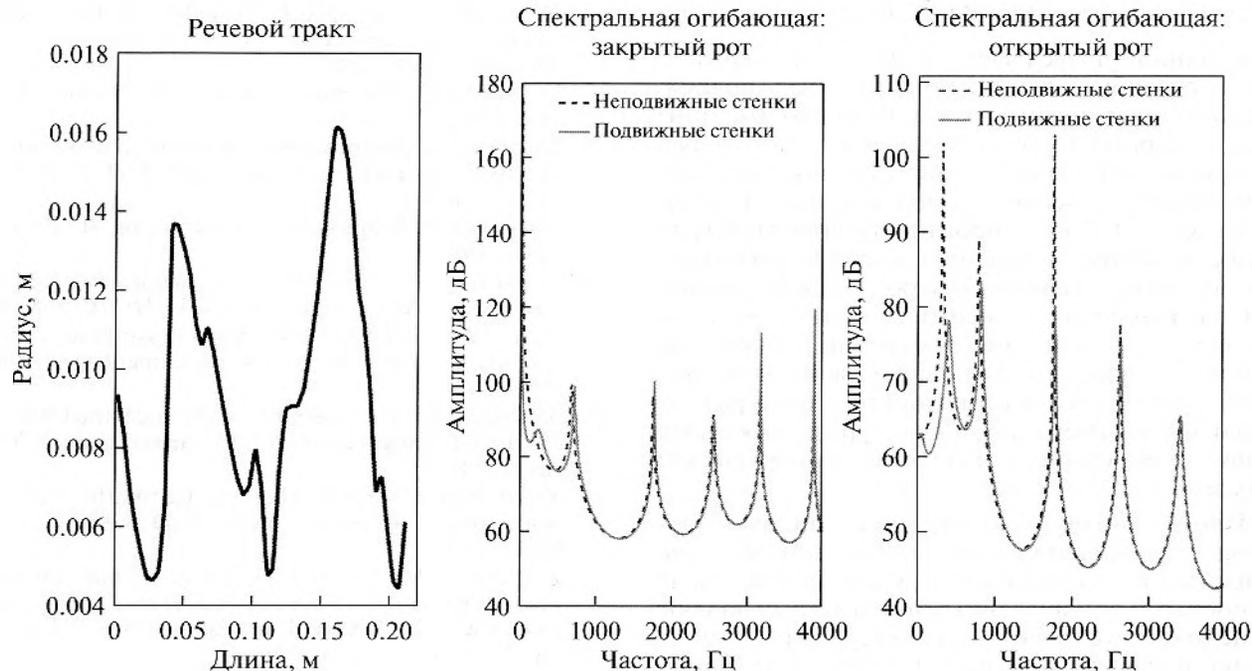


Рис. 7. Фонема "у".

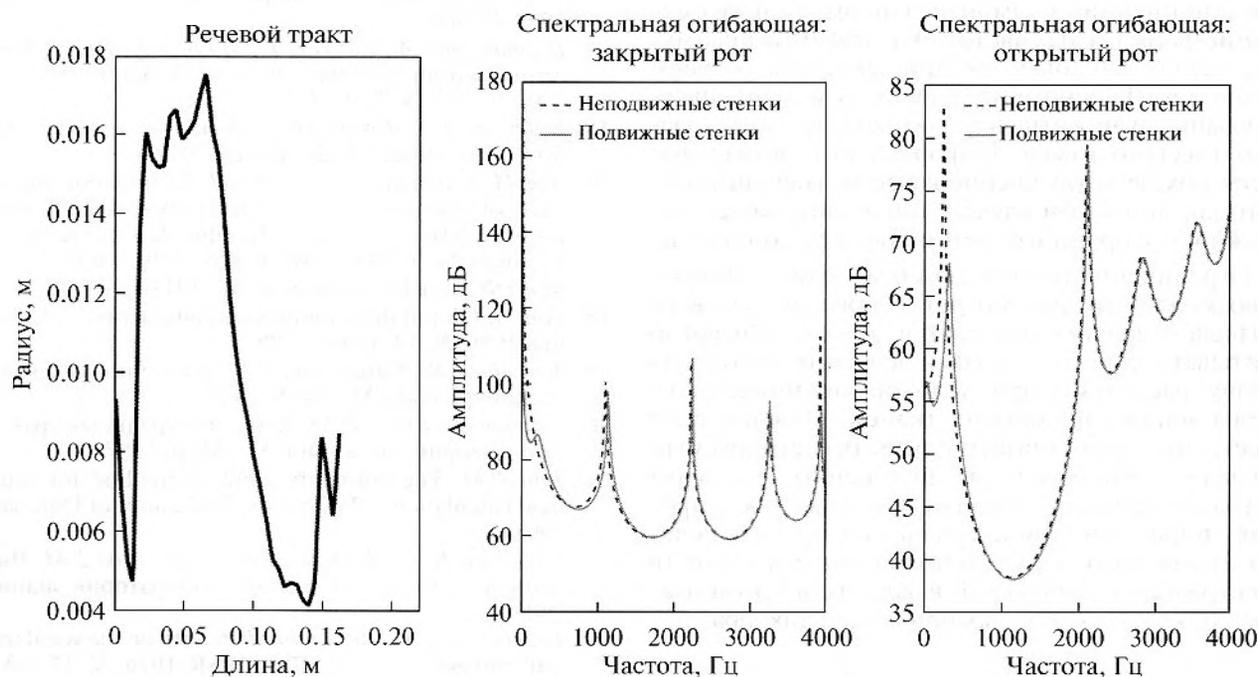


Рис. 8. Фонема "и".

помощи трехмерной модели речевого тракта с абсолютно жесткими стенками, как это сделано в работе [14]. Было выявлено, что подвижность и проницаемость стенок речевого тракта, определяемая граничной функцией импеданса, существенно влияет на характер восприятия акустического сигнала. Субъективная аудиальная оценка синтеза является, безусловно, неформальным критерием ка-

чества моделей, однако достаточно эффективной для демонстрации возможностей моделирования. Аудиальный тест зачастую применяется в задачах *TTS*-синтеза (Text-To-Speech), а также оценки эффективности сжатия аудиосигнала. Примеры синтезированных аудиосигналов можно найти на интернет-странице <http://research.makseq.com/speech-production>.

5. ЗАКЛЮЧЕНИЕ

В данной статье разработана модель акустического речеобразования, учитывающая подвижность стенок речевого тракта. В предположении о локально-реактивной поверхности тракта рассматривалась трехпараметрическая модель движения элемента поверхности под действием звукового давления. Сформулировано интегральное уравнение на основе внутренней задачи Гельмгольца с граничными условиями 3-го рода. Для численного метода граничных элементов использована схема Галеркина с линейными элементами. Для упрощения метода вводилось условие осевой симметрии, позволяющее свести поверхностное интегральное уравнение к интегралам по контуру, а также представить сингулярные интегралы аналитическим образом.

Проведенный экспериментальный анализ показал, что подвижность стенок вносит существенный вклад в решение акустической задачи. Основными обнаруженными эффектами являются нелинейные искажения спектра передаточной функции речевого тракта: возникновение резонанса в низкочастотной области, отвечающего вынужденным колебаниям стенок тракта с упругим или инерционным импедансом, а также смещение формантных частот. Эти эффекты преимущественно наблюдаются при закрытом ротовом отверстии. На основе предлагаемой модели синтезированы аудиосигналы, соответствующие гласным русского языка. Неформальные аудиальные тесты показали улучшение качества воспринимаемого сигнала в том случае, когда учитывается подвижность и проницаемость стенок речевого тракта.

Ограничения предлагаемой модели сводятся к невозможности рассмотреть свойства речевого сигнала в дальнем поле. Для этого необходимо учитывать влияние головы и решать более общую задачу рассеяния при условии проницаемости среды мягких и костных тканей. Однако, если учесть, что в простейшем случае регистрируемый сигнал в дальнем поле связан линейно с сигналом на границе речевого тракта [7], то даже такая простая аппроксимация позволяет сделать заключение о важности более точного анализа свойств спектральной огибающей в задачах детектирования уникальных речевых признаков диктора.

СПИСОК ЛИТЕРАТУРЫ

1. *Helmholtz H.L.F.* Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. Braunschweig: Vieweg, 1863.
2. *Фант Г.* Акустическая теория речеобразования. М.: Наука, 1964.
3. *Рабинер Л.Р., Шаффер Р.В.* Цифровая обработка речевых сигналов. М.: Радио и связь, 1981.
4. *Morgan N., Gold B.* Speech and audio signal processing: processing and perception of speech and music. New York: John Wiley, 2000.
5. *Сорокин В.Н.* Речевые процессы. М.: Народное образование, 2012.
6. *Макаров И.С.* Аппроксимация речевого тракта коническими рупорами // Акуст. журн. 2009. Т. 55. № 2. С. 256–265.
7. *Сорокин В.Н.* Теория речеобразования. М.: Радио и связь, 1985.
8. *Ланда П.С., Руденко О.В.* О двух механизмах генерации звука // Акуст. журн. 1989. Т. 35. № 5. С. 855–862.
9. *Kelly J.L., Lochbaum C.C.* Speech synthesis / Proc. 4th Int. Congress on Acoustics, Copenhagen. 1962. P. 1–4.
10. *Mullen J.* Physical modelling of the vocal tract with the 2D digital waveguide mesh. PhD thesis, Univ. of York, April 2006.
11. *Speed M.D.A.* Voice synthesis using the three-dimensional digital waveguide mesh. PhD thesis, Univ. of York, 2012.
12. *Karjalainen M., Erku M.* Digital waveguides versus finite difference structures: equivalence and mixed modeling // EURASIP J. on Applied Signal Processing. 2004. № 7. P. 978–989.
13. *Kako T., Kano T.* Numerical simulation of wave propagation phenomena in vocal tract and domain // 11th Int. Conf. on Domain Decomposition Methods, Bergen. 1999. P. 268–273.
14. *Hannukainen A., Lukkari T., Malinen J., Palo P.* Vowel formants from the wave equation // J. Acoust. Soc. Am. 2007. V. 122. № 1. P. EL1.
15. *Brambley E.J.* Review of acoustic linear models with flow // Acoustics, 2012, Nantes, 2012.
16. *Atig M., Dalmont J.-P., Gilbert J.* Termination impedance of open-ended cylindrical tubes at high sound pressure level // Comptes Rendus de l'Académie des Sciences. Series IIB – Mechanics. 2004. № 1.
17. *Морз Ф.* Колебания и звук. М.: ГИТТЛ, 1949.
18. *Fant G.* Glottal flow: models and interaction // J. Phonet. 1986. № 14. P. 393–399.
19. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. М.: МГУ, 1999.
20. *Колтон Д., Кресс Р.* Методы интегральных уравнений в теории рассеяния. М.: Мир, 1987.
21. *Juhl P.M.* The boundary element method for sound field calculations. PhD thesis, Tech. univ. of Denmark, 1993.
22. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. М.: Бином. Лаборатория знаний, 2003.
23. *Fant G., Nord L., Branderud P.* A note on the vocal tract wall impedance // J. STL-QPSR 1976. V. 17. № 4. P. 13–20.
24. *Fant G.* The LF-model revisited. Transformation and frequency domain analysis // STL-QPSR 1995. V. 2. № 3. P. 121–156.
25. *Fant G.* The voice source in connected speech // J. Speech Commun. 1997. P. 125–139.
26. *Smith J.O.* Spectral audio signal processing. CCRMA, Stanford, 2010.
27. *Сорокин В.Н.* Синтез речи. М.: Наука, 1992.