UNIVERSITY OF WEST BOHEMIA

Applied and Computational Mechanics 1 (2007) 613 - 620

Applied and Computational Mechanics

Comparison of optimization methods for human vocal tract resonance properties tuning

V. Radolf [a,*]

[a] *Institute of Thermomechanics, Czech Academy of Sciences, Dolejškova 5, 182 00 Praha , Czech Republic*

Received 7 September 2007; received in revised form 8 October 2007

Abstract

The paper deals with the two various optimization processes finding such geometrical form of acoustical cavities which leads to excitation of predefined acoustic resonance. A computing times and accuracy of a solutions are compared. The attention is focused both on the first two formants that are important for vowel production, and on a domain between the third and the fifth formant. This frequency domain is important for voice timbre, namely for singing voice. The problem is solved by the help of transfer matrix method using conic acoustic elements. The results should help to have a physical background for voice rehabilitation, for teaching of opera singers at musical faculties and for better understanding of biomechanics of voice production.

Keywords: biomechanics, acoustics, optimization, formant tuning, vocal tract, transfer matrix method

1. Introduction

Theme of the paper is focused on optimization possibilities of geometrical form of human vocal tract. An optimization process is designed to find such configuration of acoustic cavities, respecting real physiological limits, which leads to excitation of predefined acoustic resonance. Especially frequency domain between the third and the fifth formant is important for voice timbre, namely for singing voice ('the singer's formant') [5]. The opera singers are able to reach these resonances, but without real physical image.

2. The mathematical model

2.1. Direct numerical method

The transfer matrix method using conic elements was used. The base of this method is wave equation of an acoustic duct with variable cross section *A(x)* and viscous losses (specific acoustic resistance r_s) [2]

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{1}{A} \cdot \frac{\partial A}{\partial x} \cdot \frac{\partial \phi}{\partial x} - \frac{1}{c_0^2} \cdot \left(\frac{\partial^2 \phi}{\partial t^2} + \frac{r_S}{\rho} \cdot \frac{\partial \phi}{\partial t} \right) = 0 \,. \tag{1}$$

*Corresponding author. Tel.: +420 266 053 783, e-mail: radolf@it.cas.cz.

The vocal tract was made of 23 elements as a model of oral and epilaryngeal cavities from vocal folds to mouth for Czech vowel /a/.

The first optimization procedure was programmed in Matlab on the basis of polytop method [1]. This is a general procedure of numerical searching for minimum of a goal function of several variables. A goal function was designed on the basis of required eigenfrequencies (formants). Let's call this method 'direct tuning' in the following text.

2.2. Energy sensitivity method

The second optimization procedure was programmed in Matlab too. According to [4] sensitivity of a particular eigenfrequency to a change in cross-sectional area can be written as

$$S_{i,\nu} = \frac{Ek_{i,\nu} - Ep_{i,\nu}}{Et_\nu}, \tag{2}$$

where $Ek_{i,\nu}$ and $Ep_{i,\nu}$ means the kinetic and potential energy in the i-th cross-section for the ν-th eigenfrequency F_ν . Et_ν means the total energy in all cross-sections. Meaning of the sensitivity function is obvious from the relation

$$S_{i,\nu} = \frac{\Delta F_\nu / F_\nu}{\Delta A_i / A_i}. \tag{3}$$

Nearly according to [4], using the sensitivity function, an iterative process was used for an i-th cross-sectional area computation

$${}^{k+1}A_i = {}^{k}A_i \cdot \left(1 + \sum_\nu {}^{k}z_\nu \cdot {}^{k}S_{i,\nu}\right), \tag{4}$$

where ${}^{k}z_\nu$ is a function of the difference between desired and instantaneous ν-th eigenfrequency. Let's call this method 'energy tuning' in the following text.

The advantage of energy sensitivity execution according to (2) is that the computation time is 120 – 150 times shorter than the computation time of direct execution for particular cross-section A_i according to (3). Sensitivity functions for i = 1-24 and ν = 1-5 computed by direct method for $\Delta A_i : A_i = 0.05$ and by energy metod are displayed in fig. 1 – fig. 5.

A difference between the curves, particularly obvious in fig. 4 and fig. 5, is not large. It is important that the character of both curves is the same.

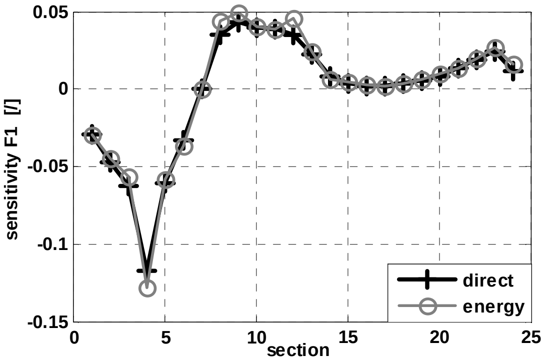

Fig. 1. Sensitivity function $S_{1\div 24,1}$.

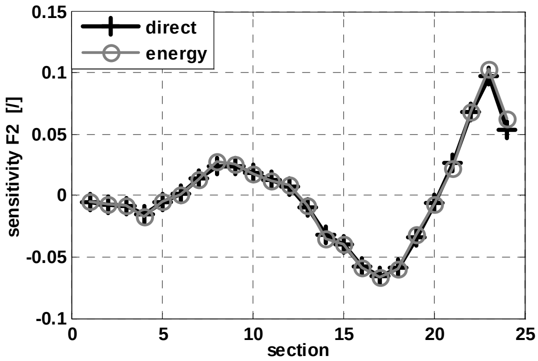

Fig. 2. Sensitivity function $S_{1\div 24,2}$.

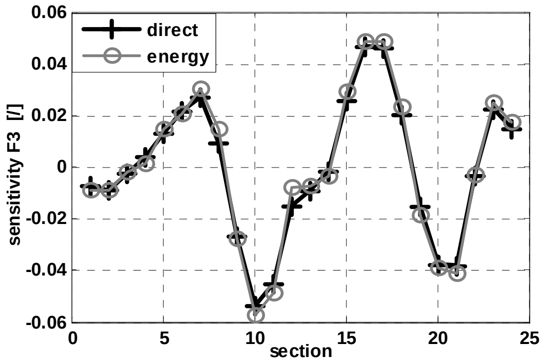

Fig. 3. Sensitivity function $S_{1\div24,\,3}$.

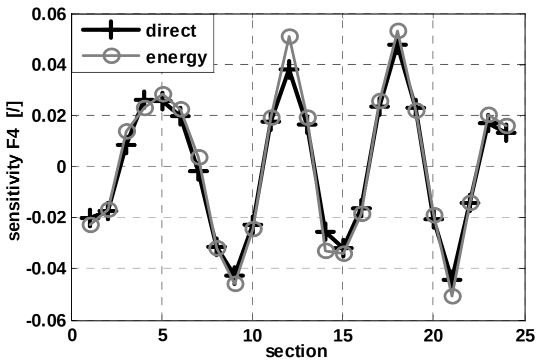

Fig. 4. Sensitivity function $S_{1\div24,\,4}$.

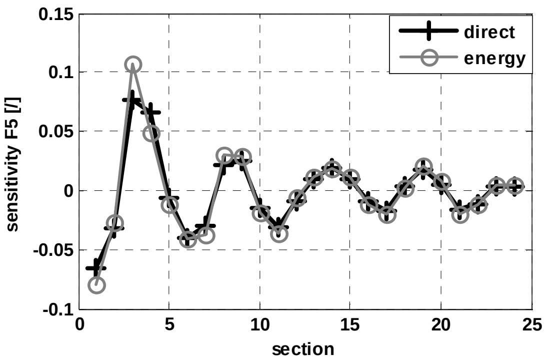

Fig. 5. Sensitivity function $S_{1\div24,\,5}$.

3. Optimization examples – formant tuning

3.1. First example – tuning of F_1 and F_2

There were all 24 cross-sections variables within the range of ± 50 % of original values in these examples. The speed of sound and the density of the air were considered c_o = 353 ms^{-1} ; ρ = 1.2 kgm^{-3}, viscous losses were neglected r_s = 0 kgm^{-3}s^{-1}. The values of the original vocal tract formants $F_1 - F_5$ were 644.3, 1186.2, 3101.8, 3944.1, 4967.5 Hz.

The aim was to tune first and second formants to frequencies 880 Hz and 1320 Hz. These values are the second and the third harmonic components of a tone with fundamental frequency 440 Hz (standard musical pitch). This is a possibility how to amplify the sung tone that gets near upper limit of tenorists.

In the first example only the first and second formants were tuned without reference to location of the others. The goal function for direct method was then

$$f_{goal} = (F_1 - 880)^2 + (F_2 - 1320)^2 . \tag{5}$$

Energy metod doesn't need any goal function, it has just prescribed F_1 = 880, F_2 = 1320 and $\nu = 1.2$.

Geometry of the original vocal tract and 'direct method' modified one is shown in fig. 6. Geometry of the vocal tract modified by 'energy method' is shown in fig. 7. Evidently 'energy method' gives smoother shape without sharp edges.

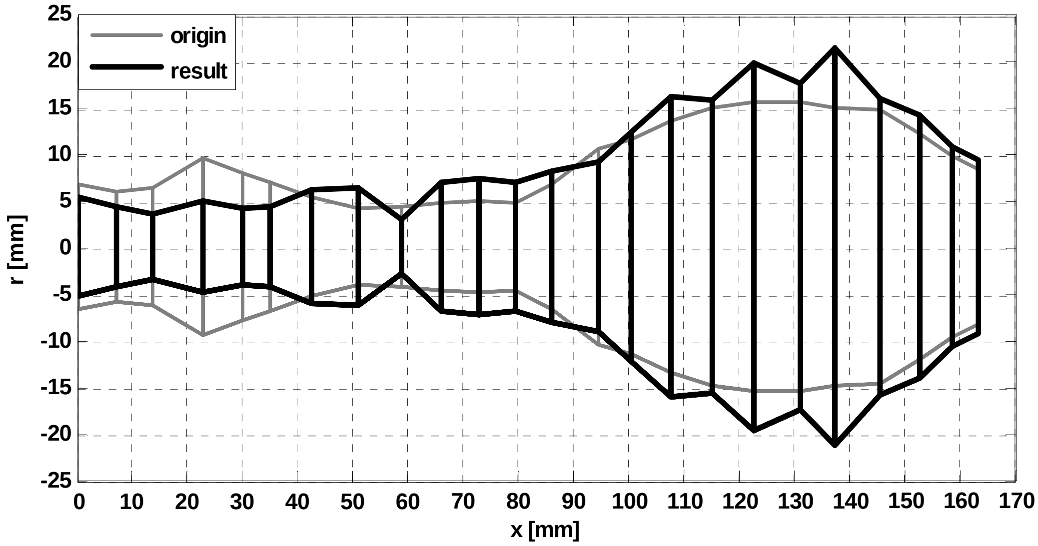

Fig. 6. Geometry of vocal tract, 'direct tuning' of F1 and F2.

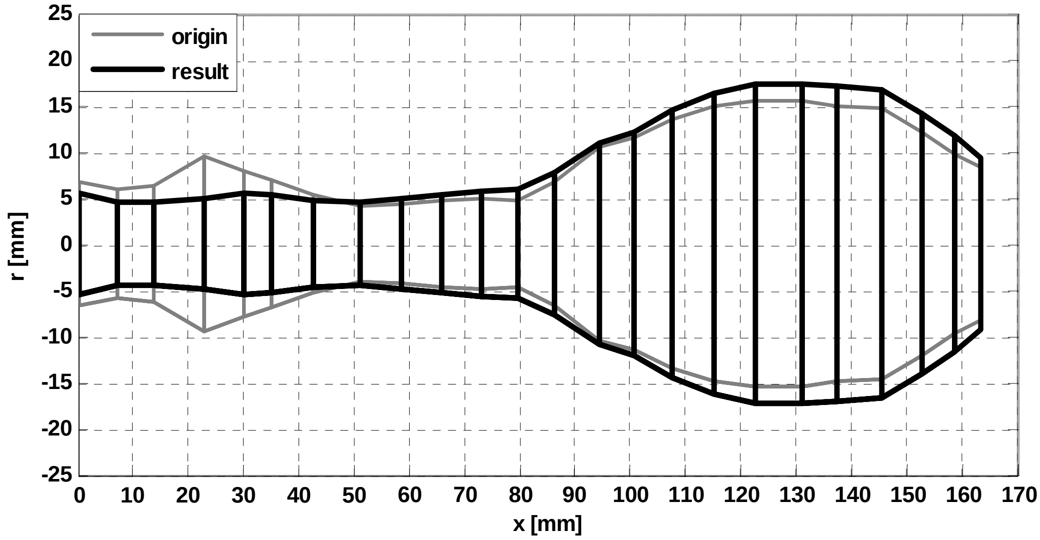

Fig. 7. Geometry of vocal tract, 'energy tuning' of F1 and F2.

In tab. 1. modified formants, number of iterations and computing time are displayed. Formants F_1 and F_2 were successfully tuned to desired frequencies by the both methods, while the others changed in a different way. Nevertheless the difference between second and third formant reduced in both cases, hence the reduction (negative part of transfer function) between 1500 – 2800 Hz disappeared. Transfer function of the original and modified vocal tract model is shown in fig. 8. for 'direct method' and in fig. 9. for 'energy method'.

Method	**Formants F_1 – F_5 [Hz]**	**Number of iterations**	**Computing time [s]**
direct	880.0, 1320.0, 2669.1, 3520.6, 4432.4	290	298
energy	880.0, 1320.0, 2894.8, 3747.4, 4740.0	136	5

Tab. 1. Results of the first example by two different methods.

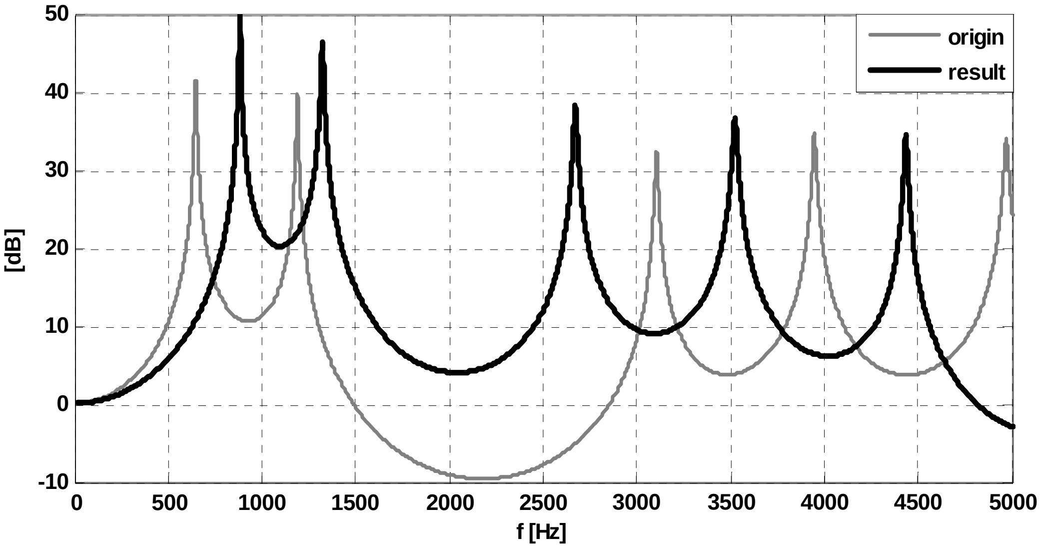

Fig. 8. Transfer function of vocal tract, 'direct tuning' of F1 and F2.

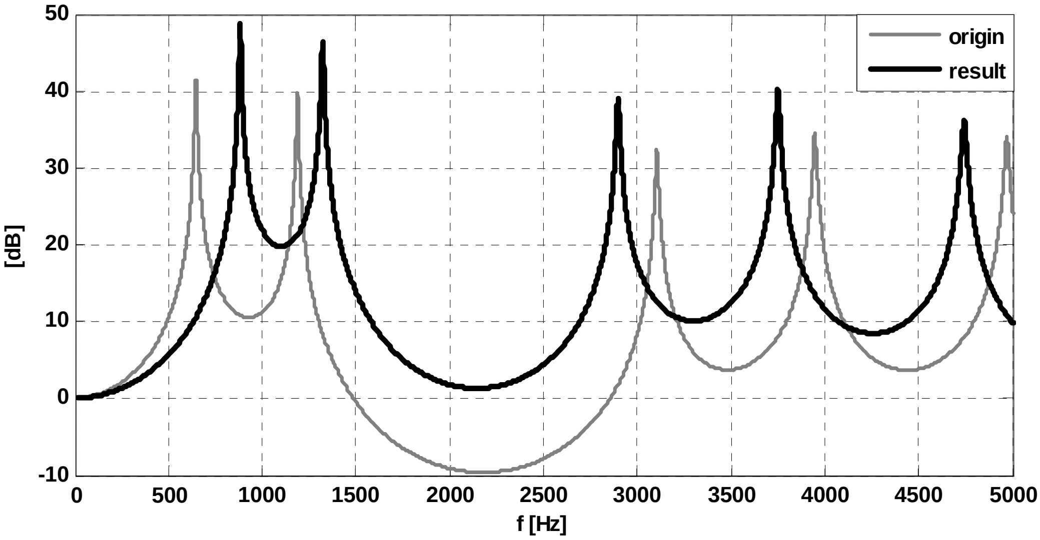

Fig. 9. Transfer function of vocal tract, 'energy tuning' of F1 and F2.

Both methods are essentially different therefore number of iterations in tab. 1. can't be compared. But the computing time is suitable quantity to confront. The 'energy method' is 60 times faster in this example.

3.2. Second example – tuning of F_1 - F_5

In the second example the first and second formants were tuned to frequencies 880 Hz and 1320 Hz again, but the others were required to keep their original values. The goal function for direct method was then

$$f_{goal} = (F_1 - 880)^2 + (F_2 - 1320)^2 + (F_3 - 3101.8)^2 + (F_4 - 3944.1)^2 + (F_5 - 4967.5)^2 . \quad (6)$$

Geometry of the original vocal tract and 'direct method' modified one is shown in fig. 10. Geometry of the vocal tract modified by 'energy method' is shown in fig. 11. The 'energy method' gives smoother shape than the 'direct method' again.

Transfer function of the original and modified vocal tract model is shown in fig. 12. for 'direct method' and in fig. 13. for 'energy method'. Both modified transfer functions are nearly the same.

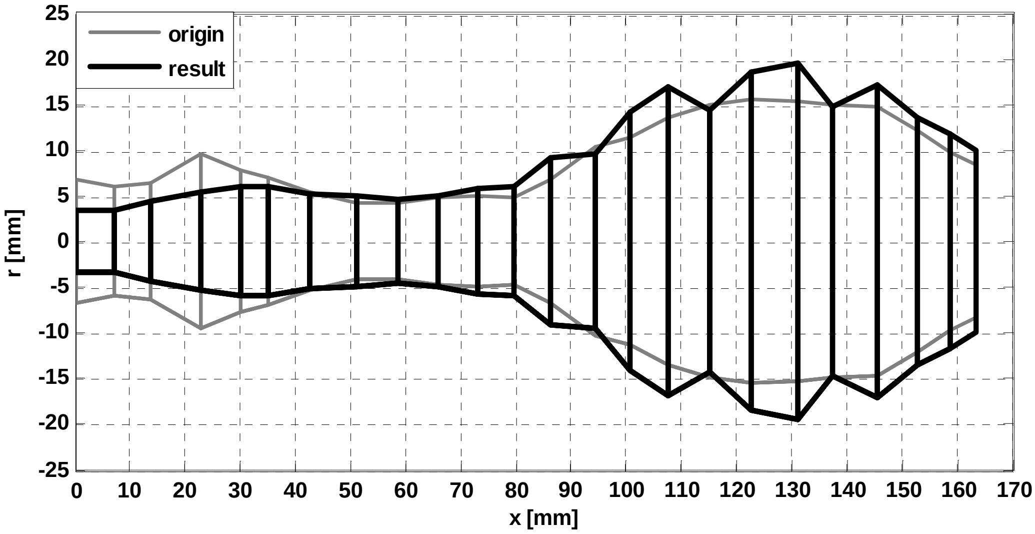

Fig. 10. Geometry of vocal tract, 'direct tuning' of F1-F5.

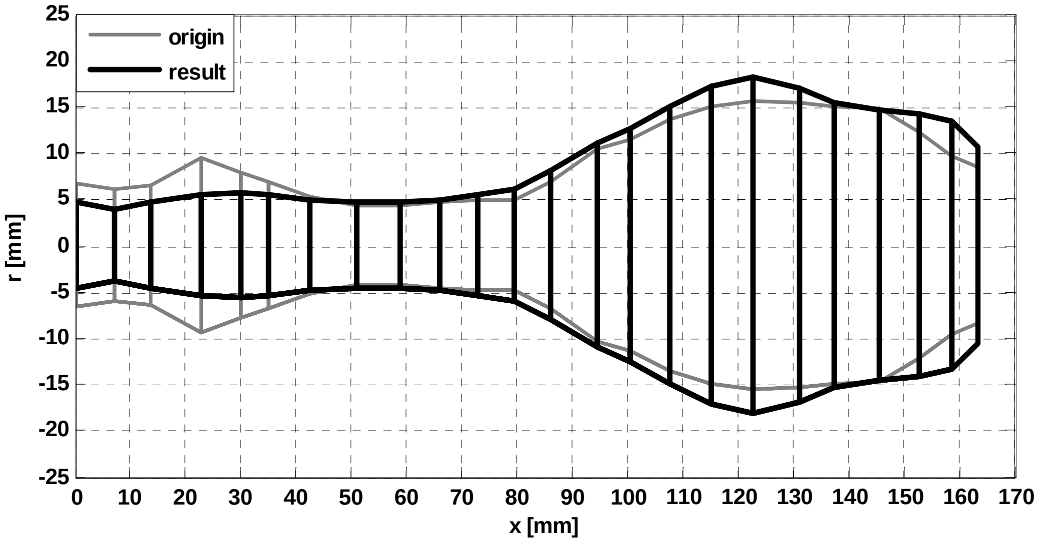

Fig. 11. Geometry of vocal tract, 'energy tuning' of F1-F5.

Resultant values of modified formants, number of iterations and computing time are displayed in tab. 2. All the formants were successfully tuned to desired frequencies by the both methods. The 'energy method' is more than 200 times faster in this case.

Method	**Formants $F_1 - F_5$ [Hz]**	**Number of iterations**	**Computing time [s]**
direct	880.0, 1320.0, 3101.8, 3944.1, 4967.5	970	2171
energy	880.0, 1320.0, 3101.8, 3944.1, 4967.5	125	10

Tab. 2. Results of the second example by two different methods.

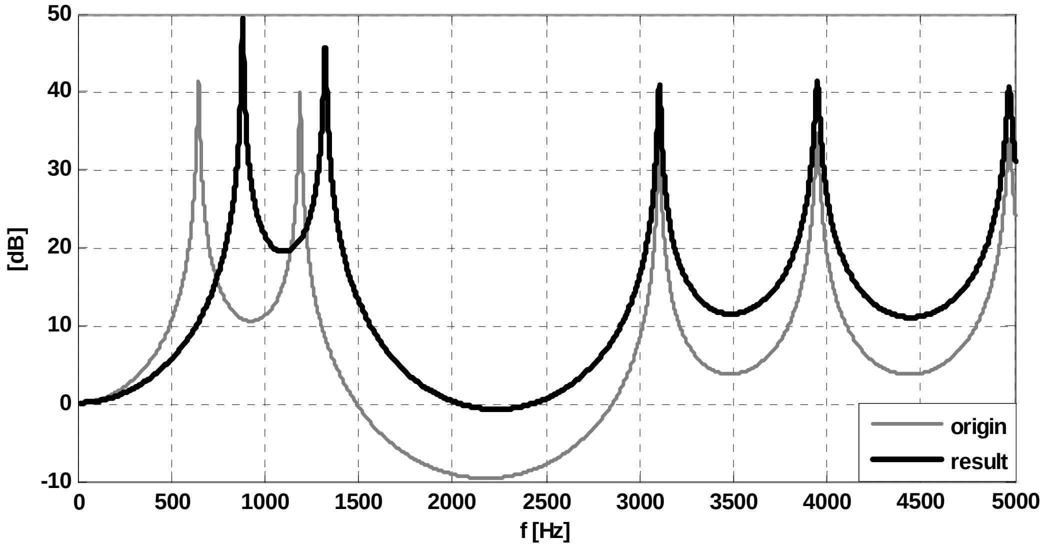

Fig. 12. Transfer function of vocal tract, 'direct tuning' of F1-F5.

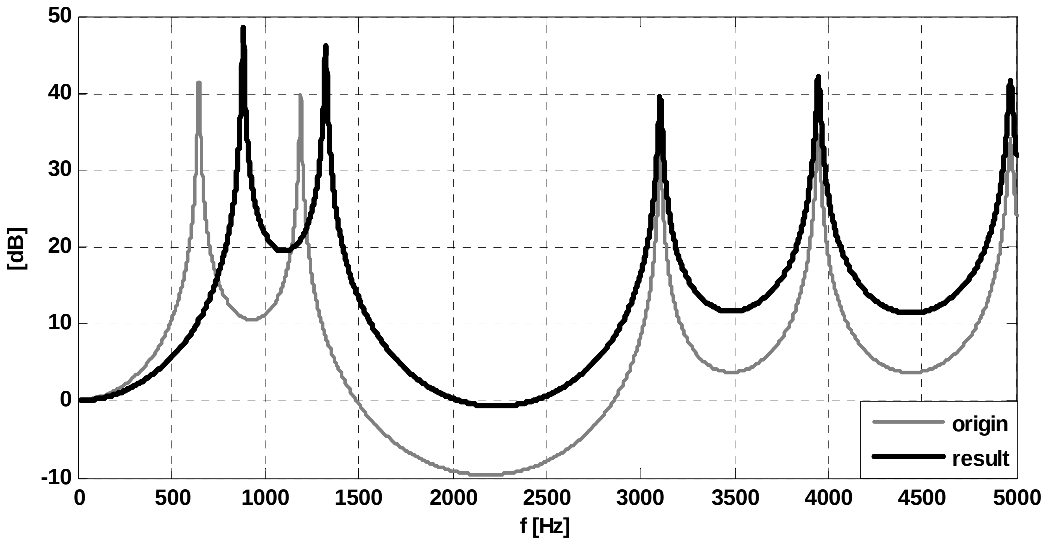

Fig. 13. Transfer function of vocal tract, 'energy tuning' of F1-F5.

4. Conclusion

Two examples of vocal tract formant tuning were carry out. Direct numerical method and energy sensitivity method were compared in both examples.

'Direct method' based on the polytop method [1] generally enables to vary both cross-sectional area and length of each acoustic conic element [3]. In addition it is possible to search for the minimum of a goal function which includes a difference between some formants. This can be important for 'the singer's formant' searching. [5], [3].

'Energy method' according to [4] enables to vary only cross-section areas and its solution just heads for the specified formants. However this method gives smoother shape of vocal tract and moreover its computing time is much shorter in comparison with 'direct method'. This time was 60 times shorter in the first example and more than 200 times shorter in the second example.

Acknowledgements

This work has been supported by the research project AV0Z20760514, by the project of IT CAS PP07-03065 'Acoustic Properties of Cavities and Their Optimization with Respect to Applications at Human Voice Biomechanics'.

References

[1] P. Lederer, Teorie a optimalizace mechanických systémů I, ediční středisko ČVUT, Praha , 1988.
[2] J. Merhaut, Teoretické základy elektroakustiky, Academia, Praha, 1971.
[3] V. Radolf, The possibilities of tuning and optimisation of acoustic resonance properties of the human vocal tract, Book of extended abstracts, Engineering Mechanics 2007, IT CAS Prague, pp. 237-238.
[4] B.H. Story, Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions, J. Acoust. Soc. Am. 119 (2) (2006) 715-718.
[5] I. Titze, Principles of Voice Production, National Center for Voice Speech, Iowa City IA 52242, USA, 2000.