Formants from the Wave Equation

and

Recording Speech During MRI

Pertti Palo <jpalo (ät) math.hut.fi>,

Antti Hannukainen, Ville Havu, Teemu Lukkari and Jarmo Malinen

Institute of Mathematics

HELSINKI UNIVERSITY OF TECHNOLOGY

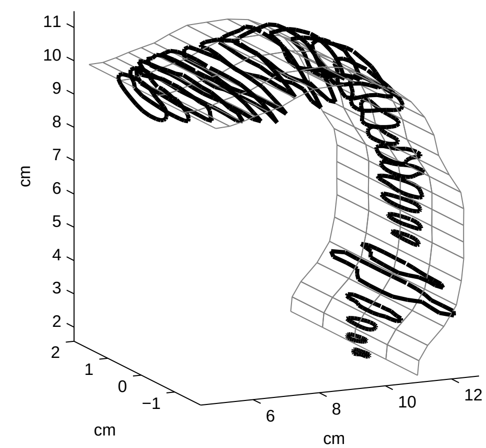
Introduction: Who are we?

- Antti Hannukainen: Programming, FEM
- Ville Havu: Mathematical analysis, numerics, FEM
- Teemu Lukkari: Mathematical analysis, acoustics, signal processing
- Jarmo Malinen: Mathematical analysis, measurements
- Pertti Palo: Phonetics, programming, measurements, acoustics, PR . . .

Background

$$\begin{cases} \Phi_{tt} = c^2 \Delta \Phi & \text{for } (\mathbf{r}, t) \in \Omega \times \mathbb{R}, \\ \Phi = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_1 \times \mathbb{R}, \\ \frac{\partial \Phi}{\partial \nu} = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_2 \times \mathbb{R}, \text{ and} \\ \Phi_t + c\frac{\partial \Phi}{\partial \nu} = 2\sqrt{\frac{c}{\rho_0}} u & \text{for } (\mathbf{r}, t) \in \Gamma_3 \times \mathbb{R}, \end{cases}$$

(a)



(b)

Figure 1: (a) The wave equation model and (b) a sample vowel geometry

- The main goal is to simulate vowels based on a wave equation model.
- We need accurate anatomic data and simultaneously recorded sound to validate the simulation results.

Acoustical model in more detail

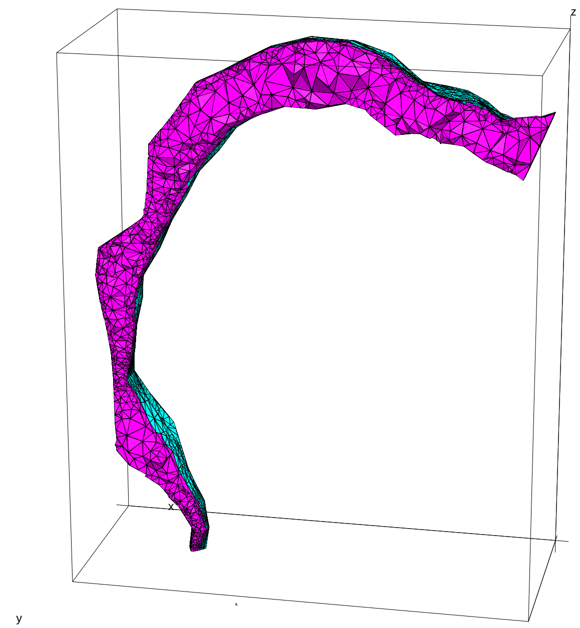
By using the velocity potential Φ the perturbation pressure can be expressed as $p' = \rho_0 \Phi_t$, where Φ is a function, which is related to the particle velocity by $\boldsymbol{v} = -\nabla \Phi$.

Solve Φ, for a given input signal u:

$$\begin{cases} \Phi_{tt} = c^2 \Delta \Phi & \text{for } (\mathbf{r}, t) \in \Omega \times \mathbb{R}, \\ \Phi = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_1 \times \mathbb{R}, \\ \frac{\partial \Phi}{\partial \nu} = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_2 \times \mathbb{R}, \text{ and} \\ \Phi_t + c \frac{\partial \Phi}{\partial \nu} = 2\sqrt{\frac{c}{\rho_0}} u & \text{for } (\mathbf{r}, t) \in \Gamma_3 \times \mathbb{R}, \end{cases} \tag{1}$$

where $u = u(\mathbf{r}, t)$ is a power input signal at the glottis end (per unit area), c the speed of sound within the VT, ν the outer normal of $\partial\Omega$, and $\frac{\partial \Phi}{\partial \nu} = \nu \cdot \nabla \Phi$.

Mathematics: Finite Element Method in our project



- Our initial mesh had about 64000 tetrahedral elements.
- Computation of the resonance (Helmholtz) problem took about half an hour.
- With the time dependent case, we do not aim to produce real time synthesis, but do expect to get relatively close.

Results: Computed formants in F2-F1 plane

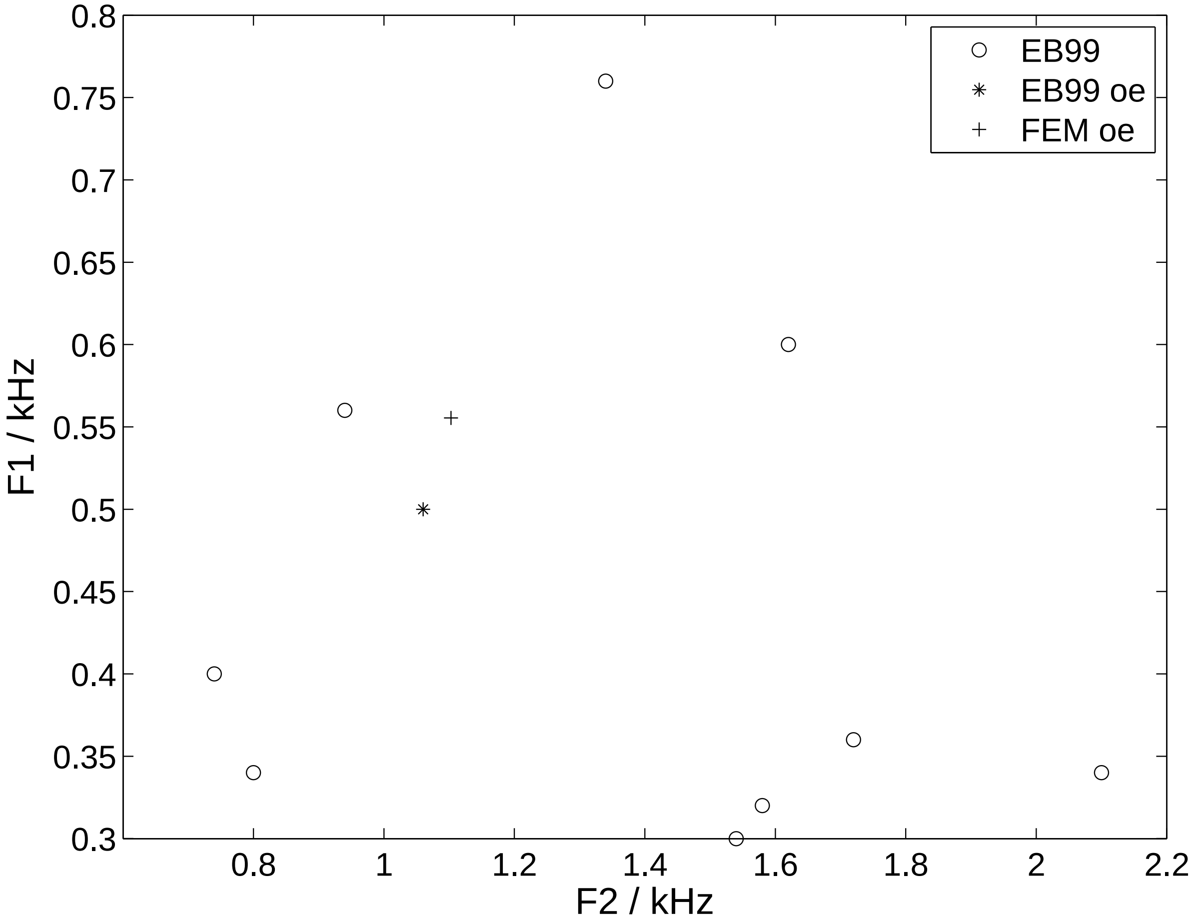


Figure 2: Computed formants and Olov Engwall's measured formants for long vowels in F2-F1 plane.

Results: Pressure distributions

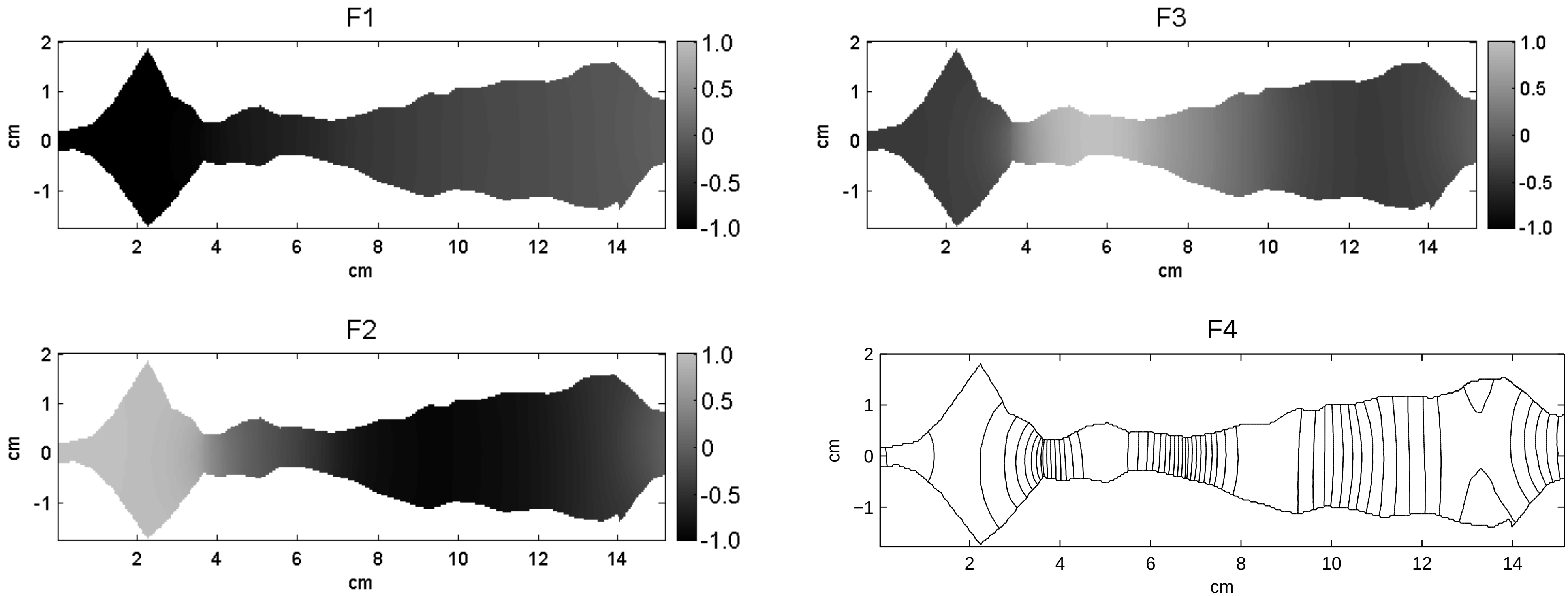


Figure 3: Four approximate eigenfunctions corresponding to the lowest eigenvalues ie. pressure distributions for formants 1-4. Glottis is on the left and mouth on the right.

Intermission

That was our status last spring.

So what have we been up to during the last year or so?

Sound measurements: What would we like to get?

- The fundamental frequency F0, . . .
- F1, F2, F3 and, if possible, F4 . . .
- . . . and their bandwidths . . .
- . . . before, after and during the MR imaging sequence.
- Access to clean speech signal in real time.

Sound measurements: What’s the problem then?

- No metal allowed inside the MRI main coil.
- No magnetic material allowed inside the MRI room.
- All electronics in the MRI room have to be RF-shielded.
- Strong acoustic noise (over 90 dB SPL) present during the imaging sequence.

What did we decide to do?

The recording system is based on three main design principles:

1. using air as signal medium when unavoidable,

2. using real-time analog electronics for first stages of signal processing, and

3. using DSP for post-processing.

System measurement setup

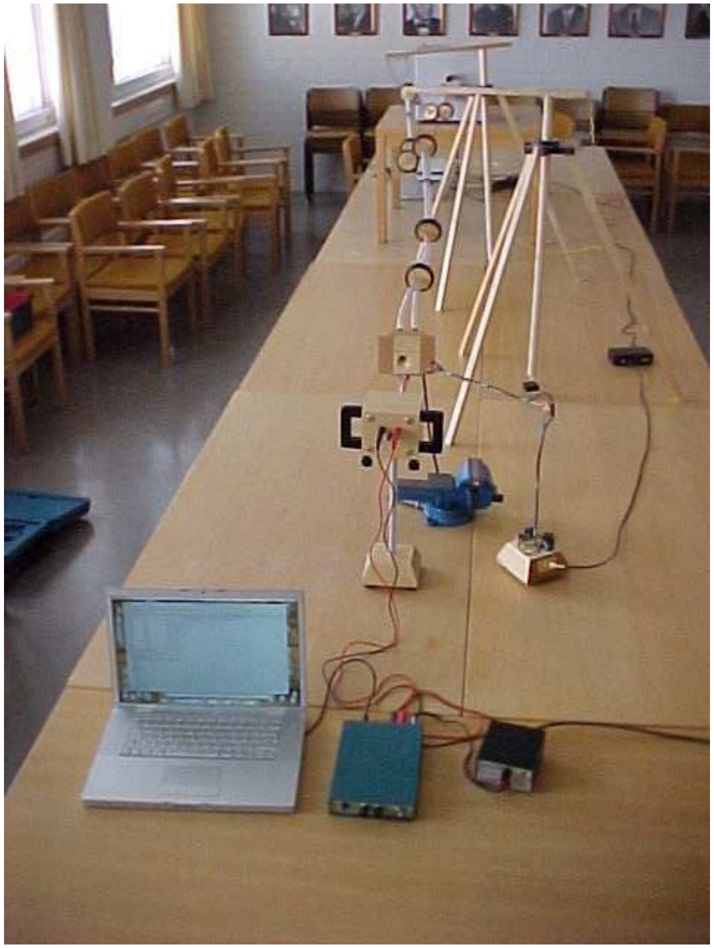


Figure 4: The setup for acoustic field measurements

Point like sound source

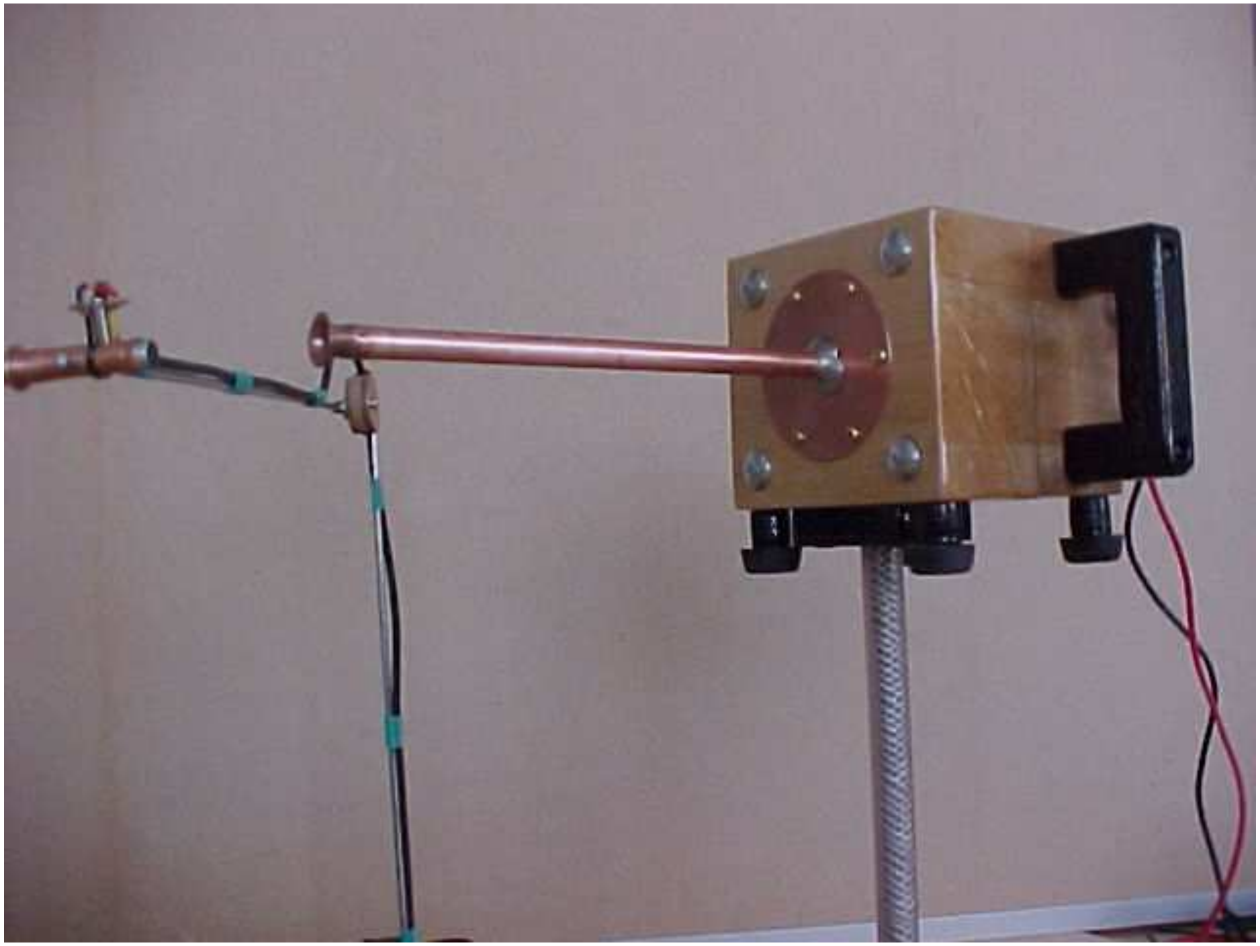
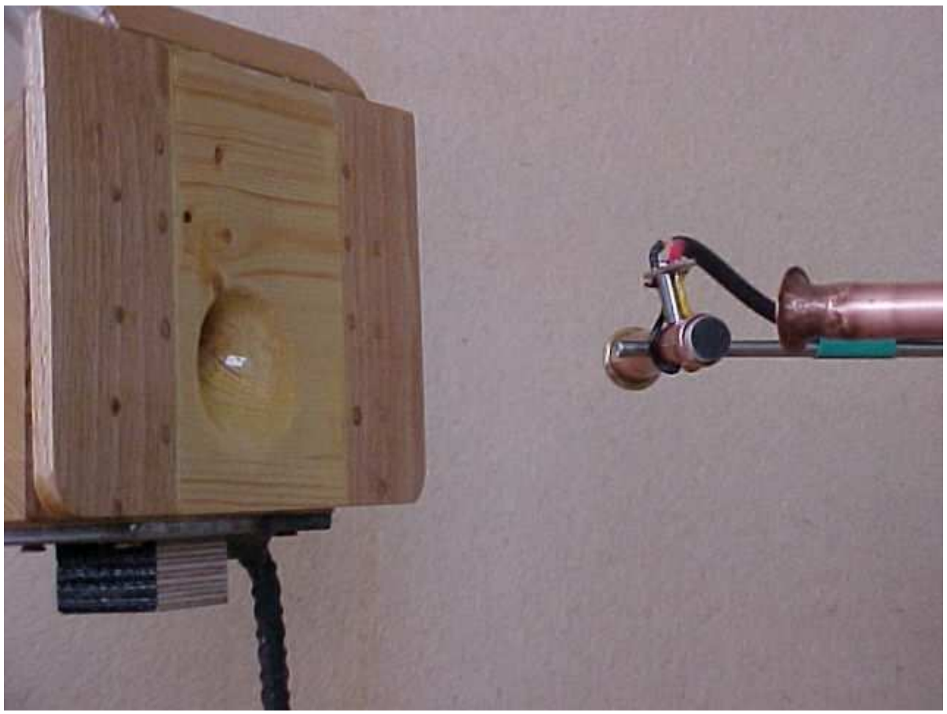


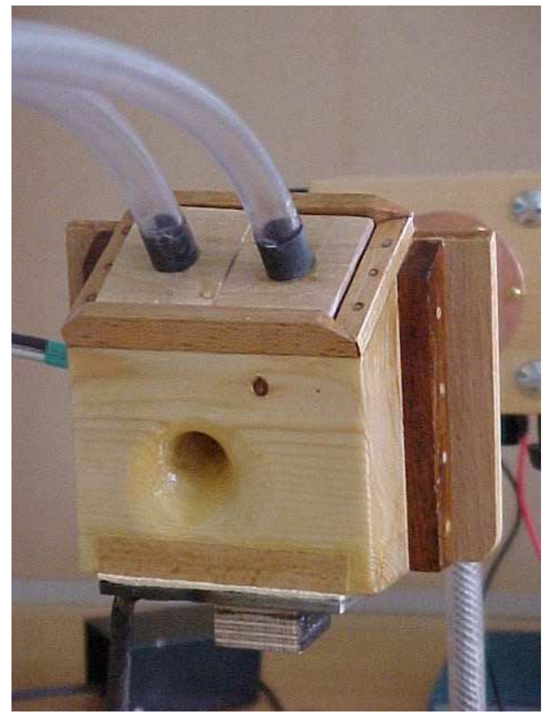
Figure 5: This sound source will be used to measure the frequency response of the sound recording setup as well as the directionality of the sound collector

Sound collector

There is a two channel sound collector in our system. One channel is for noise and the other for the contaminated speech.



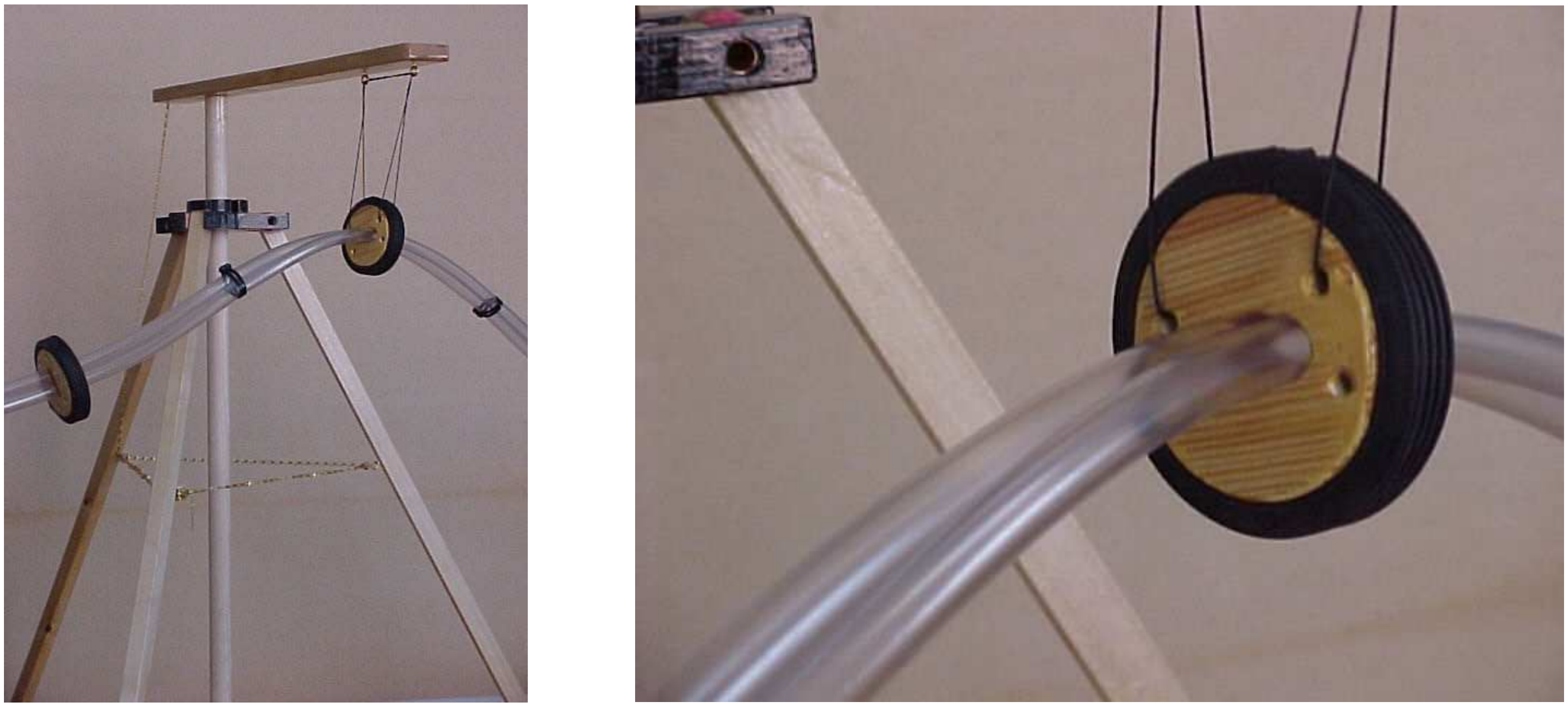
(a)



(b)

Figure 6: The sound collector from (a) below with the sound source and (b) above

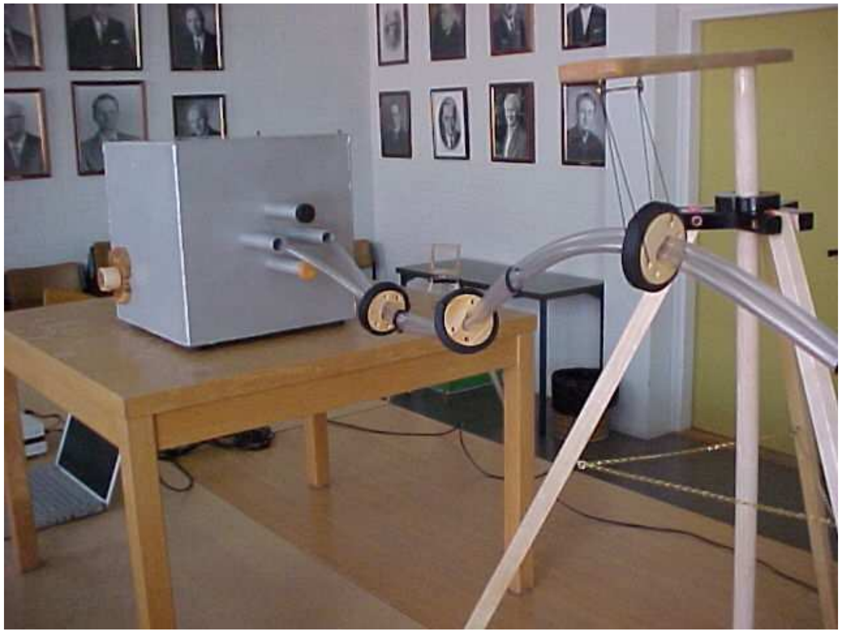
Acoustic wave guides



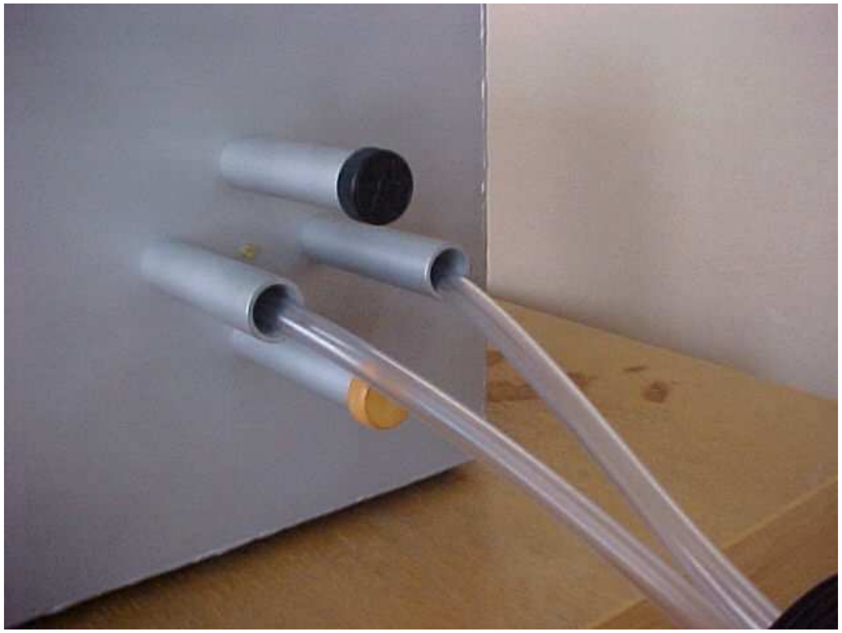
(a) (b)

Figure 7: (a) The acoustic wave guides hanging from a magnet free stative and (b) the suspension in close up

Faraday cage



(a)



(b)

Figure 8: (a) The Faraday cage houses the microphones and (b) the acoustic waveguides enter the cage through electromagnetic waveguides

Microphone array

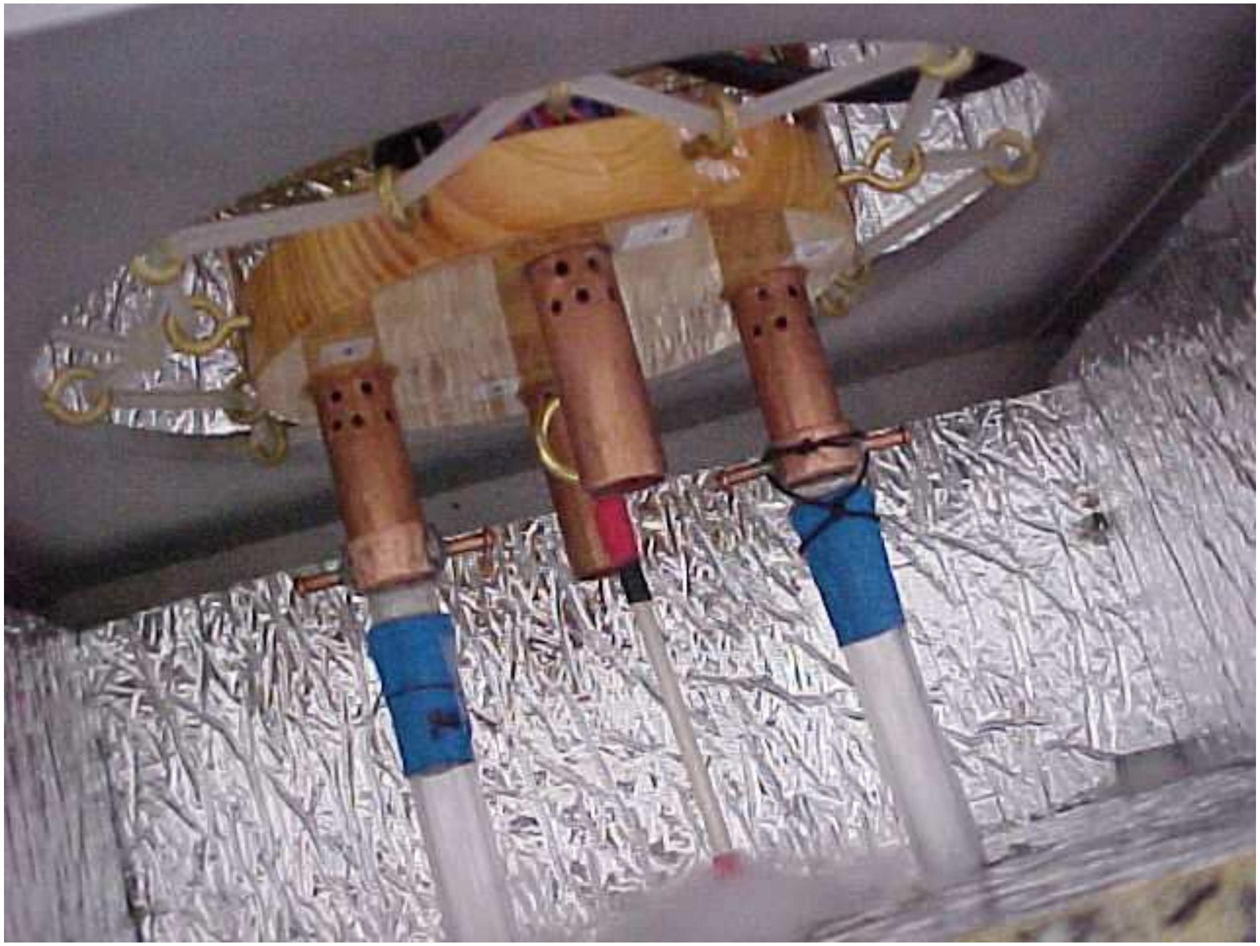


Figure 9: The microphone array consists of four microphones.

De-noising amplifier

- Analog electronics provide real time response.
- Overvoltage and RF shielded inputs
- One speech input channel
- Up to three noise input channels
- Optional low-pass filtering and independent amplifications

Tests: Acoustic wave guides

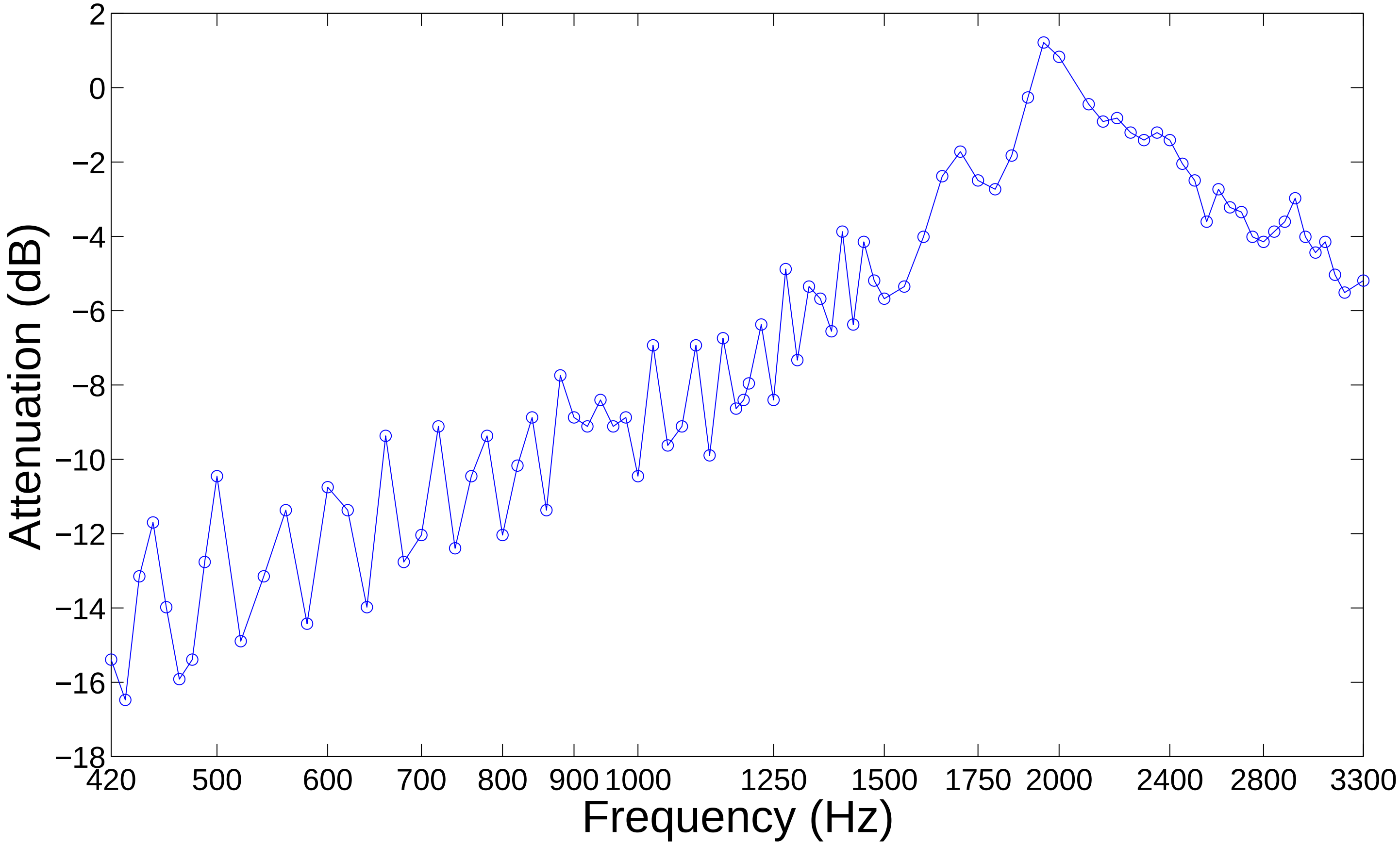


Figure 10: Frequency response of the acoustic wave guides

Tests: Does the noise cancellation work with acoustic components?

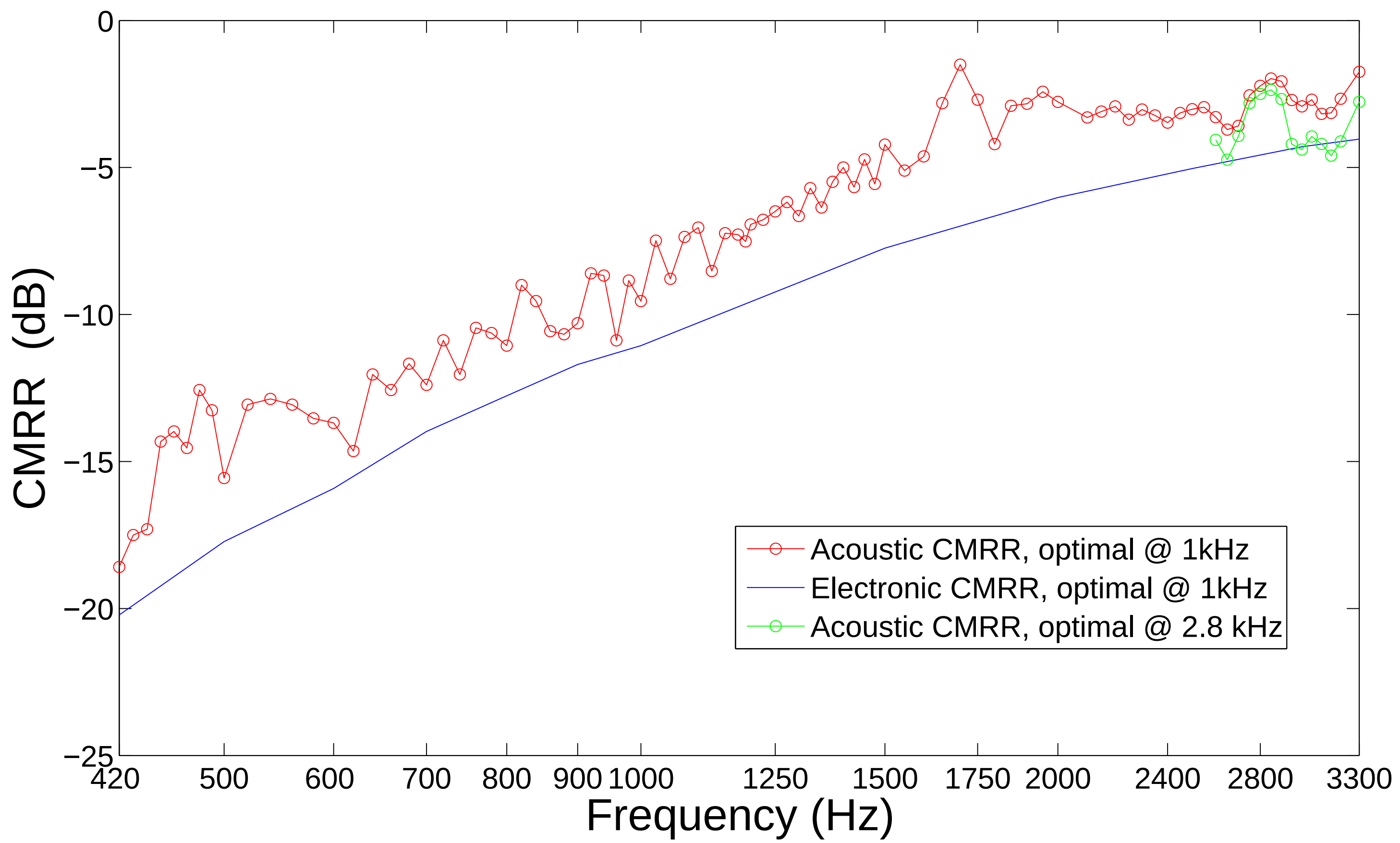
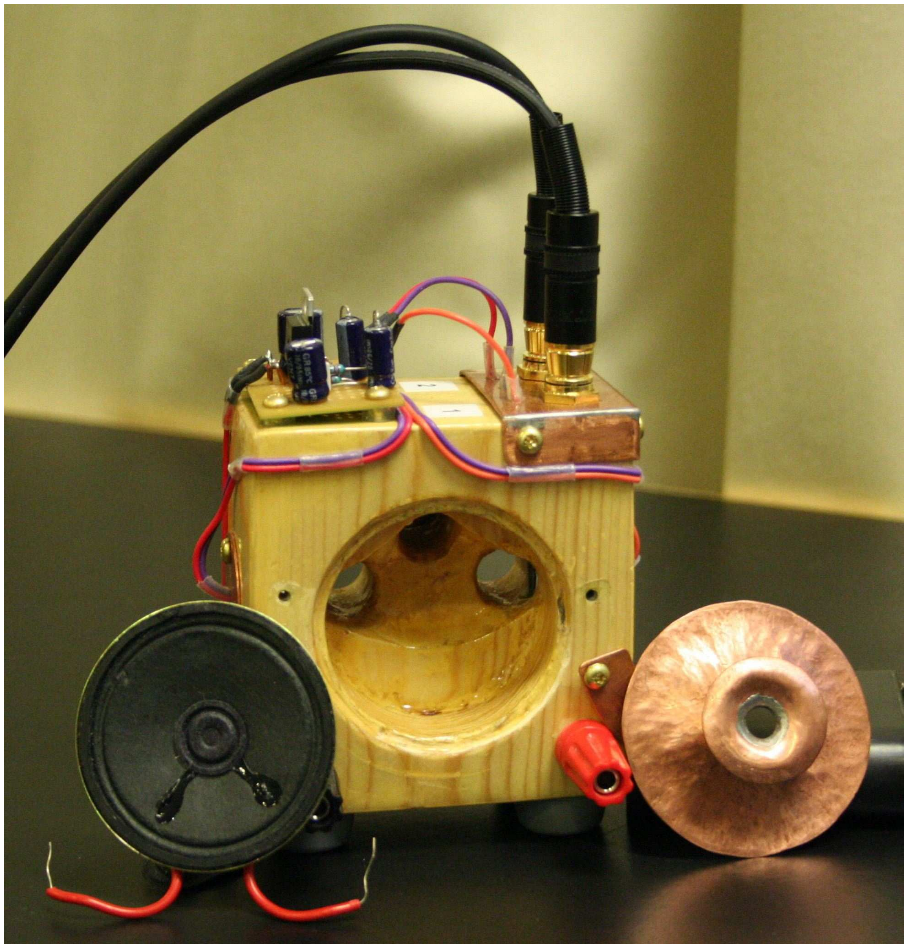


Figure 11: CMRR of the whole system excluding the sound collector

Tests: Two channel signal source



We used a custom built acoustic signal source to obtain the previous data.

Full circle

$$\begin{cases} \Phi_{tt} = c^2 \Delta \Phi & \text{for } (\mathbf{r}, t) \in \Omega \times \mathbb{R}, \\ \Phi = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_1 \times \mathbb{R}, \\ \frac{\partial \Phi}{\partial \nu} = 0 & \text{for } (\mathbf{r}, t) \in \Gamma_2 \times \mathbb{R}, \text{ and} \\ \Phi_t + c\frac{\partial \Phi}{\partial \nu} = 2\sqrt{\frac{c}{\rho_0}} u & \text{for } (\mathbf{r}, t) \in \Gamma_3 \times \mathbb{R}, \end{cases}$$

Thank you.

Questions, please?