MASTER

The strike note of bells

Eggen, J.H.

Award date:
1986

THE STRIKE NOTE OF BELLS.

J.H EGGEN

Eindhoven, Februari 1986

Final report of the work carried out in the "Hearing and Speech" group of the Institute for Perception Research under the supervision of dr.A.J.M. Houtsma.

SUMMARY

The physical sound spectrum of a bell contains a great number of eigenfrequencies (partials), which are not harmonically related. Despite this inharmonicity, one can usually perceive a clear pitch when a bell is struck. In a mistuned bell this strike note pitch often doesn't coincide with a frequency present in the spectrum of the sound. The strike note pitch therefore is called a subjective tone. In this project the subjective pitch of the strike note is measured as a function of systematic frequency shift in selected partials. A digital filtering technique is used to manipulate the sound of a real bell. The octave, the twelfth and the upper octave partials turn out to be important with respect to the strike note pitch.

SAMENVATTING

Het fysische geluidsspectrum van een klok bevat een groot aantal eigenfrequenties, die onderling geen harmonische relatie vertonen. Ondanks deze inharmoniciteit, kan men gewoonlijk een duidelijke toonhoogte toekennen aan het klokkengeluid. In het geval van een ontstemde klok is de toonhoogte van deze slagtoon niet altijd gelijk aan een van de frequenties uit het fysische geluidsspectrum. De slagtoon wordt daarom een subjectieve toon genoemd. In dit onderzoek is de subjectieve toonhoogte van een klok gemeten, als functie van een systematische frequentie verschuiving van de afzonderlijke partialen. Het geluid van een echte klok werd bewerkt met een digitale filtertechniek. Het octaaf, de duodeciem en het dubbel oktaaf bleken belangrijke partialen te zijn, met betrekking tot de slagtoon.

CONTENTS

SUMMARY 1

CONTENTS 2

1 INTRODUCTION 3

2 HUMAN PITCH PERCEPTION 6

2.1 Historical review 6

2.2 Theoretical models 10

2.2.1 The optimum processor theory (Goldstein) 10

2.2.2 The virtual pitch theory (Terhardt) 14

2.3 Pitch meters based on human pitch perception 17

2.3.1 The DWS-pitch meter 17

2.3.2 The virtual pitch extractor (Terhardt) 20

3 THE ACOUSTICS OF BELLS 24

3.1 Physical aspects 24

3.2 Psychophysical aspects 28

4 ANALYSIS AND SYNTHESIS OF BELL STIMULI 31

4.1 Analysis 31

4.2 The synthesis of bell stimuli 33

5 EXPERIMENT 40

5.1 Apparatus 40

5.2 Procedure 43

5.3 Results 43

6 DISCUSSION AND CONCLUSIONS 52

7 ACKNOWLDGMENTS 54

8 REFERENCES 60

APPENDIX A1 : Formulas Virtual Pitch Theory A-1

APPENDIX A2 : Results of the individual subjects A-3

APPENDIX A3 : Computer program TVP-meter A-27

1 INTRODUCTION

Pitch is one of several subjective qualities, often used to describe the perceived sensation of a physical sound stimulus. Although these subjective attributes are related to physical parameters of the sound stimulus, they cannot be measured directly. The pitch percept of a sound stimulus makes it possible to order sounds on a musical scale. In the case of a sine wave, the pitch is closely related to the frequency, although sound level may affect the perceived pitch. For complex sounds, pitch also depends on frequencies and amplitudes of other partials in the spectrum.

According to Ohm's acoustical law a certain pitch could only be heard if the spectrum of the sound contained power at the corresponding frequency. Seebeck (1841) and Schouten (1940) showed that a periodic sound in which the fundamental frequency was very weak or even absent, was still peceived as having a pitch corresponding to the fundamental frequency. Ohm's law was contradicted ! This very interesting pitch phenomenon of the "missing fundamental", clearly showed that the pitch percept was a subjective attribute of sound.

All modern theories of pitch perception of complex tones involve a central neural processor (Goldstein,1973; Terhardt,1974). This central processor uses signals derived form peripherally resolved stimulus components as input. A great amount of experimental research has been done on the pitch of complex tones comprising two or three successive or nonsuccessive harmonics (Schouten,Ritsma and Cardozo,1962; Houtsma and Goldstein,1972; Goldstein et al.,1978; Houtsma,1979). Modern pitch theories are able to account for most of these data in a quantative sense. Much less research was carried out on large and spectrally dense complex sounds.

A bell is a typical example of a very complex sound. The physical sound spectrum of a bell contains a great number of characteristic eigenfrequencies, which are not harmonically related (van Heuven,1949; Slaymaker and Meeker,1954; Lehr,1965). Despite this inharmonicity, one can perceive a clear pitch when a bell is struck. Early investigations showed that this strike note was a subjective tone (Jones,1930; Meyer and Klaes,1933; Arts, 1938). In the case of a properly tuned bell, the strike note pitch coincides with the frequency of the second partial. The strike note pitch of a mistuned bell often doesn't coincide with the second partial. The most succesful explanation of the origin of the strike note was given by Schouten and 't Hart (1965). They concluded that the strike note pitch was a residue pitch. The strong partials, named the octave, the twelfth and the double octave with the frequency ratios 2:3:4, were perceived together as a tonal residue with a pitch corresponding to the fundamental frequency of this harmonic series.

Recent investigations of bells were carried out by Terhardt (1984) and Greenhough (1976). Terhardt compared the perceived pitch of a number of bells to the pitch predicted by his virtual pitch theory (Terhardt, 1982). In the case of Terhardt's experiment, like most experiments on bells, the stimuli couldn't be manipulated systematically. Greenhough was the first one to examine the relative importance of the most prominent partials of a bell systematically. He used synthetic bell sounds, generated on a digital computer by means of additive synthesis. Instead of the generally accepted dominant role of the octave with respect to the strike note pitch, Greenhough found the double octave to be an even more important partial.

In this project we will measure the effect on the strike note pitch caused by a frequency shift of the most prominent partials of a particular bell. Instead of synthetic bell sounds, our research will deal with real bell sounds. A specific partial is removed from the original bell sound using a IIR digital bandreject filter. This partial is replaced by a sinusoid with an amplitude function that approximates the amplitude envelope of the original partial. The frequency of the new partial is varied systematically. The very good preservation of the sound quality is an important feature of this digital filtering technique. This technique will therefore be used to find out whether the findings of Greenhough will hold for real bell sounds. We will also examine the effect of a pitch shift of the strike note in the case two or three partials are shifted in frequency at the same time.

Chapter 2 will discuss the theory of human pitch perception of complex tones. Two pitch extraction algorithms will also be described. The physics and psychophysics of bells are discussed in chapter 3. Chapter 4 provides an outline of the digital signal processing applied in this project. The experimental procedures and results are described in chapter 5. In chapter 6 the results are discussed and some conclusions are drawn.

2 HUMAN PITCH PERCEPTION

Section 2.1 gives a brief historical review of theoretical and experimental developments in the area of human pitch perception. Two modern pitch perception theories, known as the "optimum processor theory" (Goldstein, 1973) and the "virtual pitch theory" (Terhardt,1974) are discussed in section 2.2 . The DWS-pitch meter based on Goldstein's theory and the virtual pitch extractor of Terhardt are described in section 2.3 .

2.1 Historical review

The first mathematical theory on pitch perception was developed by G.S. Ohm (1843). According to Ohm's acoustical law, the ear performs a fourier analysis on the acoustical signal. Pitches corresponding to certain frequencies can be heard only if the sound signal contains power at those frequencies. Psychophysical experiments on sirens carried out by Seebeck, however, did not support Ohm's acoustical law (Seebeck,1841). The signal's component at the fundamental frequency was far too weak to explain the strong sensation of a pitch at the fundamental. Ohm called this phenomenon an auditory illusion, whereas Seebeck called it a relevant perceptual fact. The idea of Ohm, the ear being a frequency analyzer, was adapted by Helmholtz (1877). Helmholtz explained the phenomenon of the "missing fundamental", demonstrated by the siren experiments of Seebeck, with the concept of distortion products generated in the ear. The fundamental was supplied by the ear of the listener as a difference tone.

According to Helmholtz, the ear's basilar membrane can be regarded as a frequency analyzer. The basilar membrane consists of a great number of transverse fibers. When a sound stimulus excites the basilar membrane, certain fibers will resonate at frequencies determined by their length, mass and tension. Bekesy's investigations (1960) showed that the concept of individual transverse fibers was not correct. The individual fibers couldn't resonate freely, but rather the basilar membrane as a whole showed places of maximum vibration. (fig. 2.1.b.) Due to a pressure stimulus, a moving wave is travelling along the basilar membrane. The envelope of the travelling wave shows a maximum, which place along the membrane is a function of frequency. (fig.2.1.c)

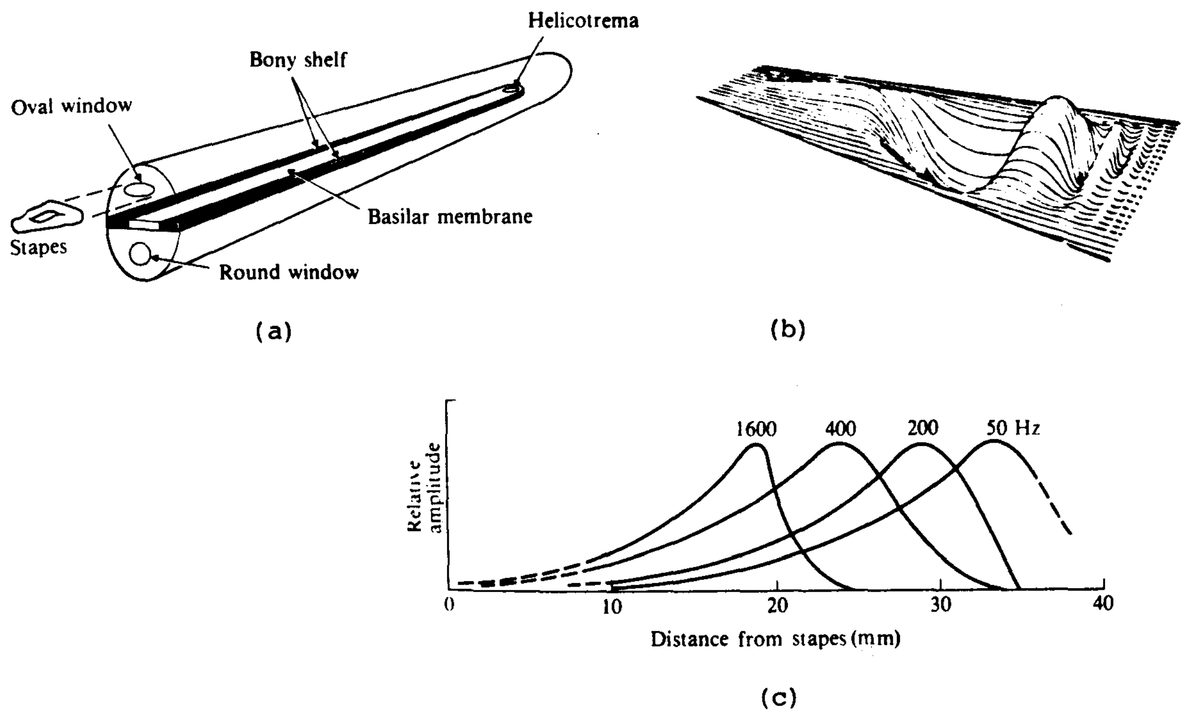


figure 2.1.: a. Schematic diagram of uncoiled cochlea, b. Snapshot of traveling wave along the basilar membrane, c. basilar membrane displacement amplitude as a function of distance from stapes for several frequencies (Rossing,1982).

This view on the working of the auditory system is usually referred to as the place theory of hearing. A vibration in time is converted into a vibration pattern in space. Next, the space pattern excites a spatial pattern of neural activity.

The time (or periodicity) theory of pitch perception represents a completely different way of looking at the auditory system. According to this theory the ear performs a time analysis on the sound stimulus instead of a frequency analysis. Information about the time structure of the sound stimulus is reflected in the time distribution of spikes in the auditory nerve. This information is decoded in the central nervous system (Wever, 1949). According to te volley theory of Wever, the total responses of a group of nerve fibers represent the frequency of a sound wave (see figure 2.2).

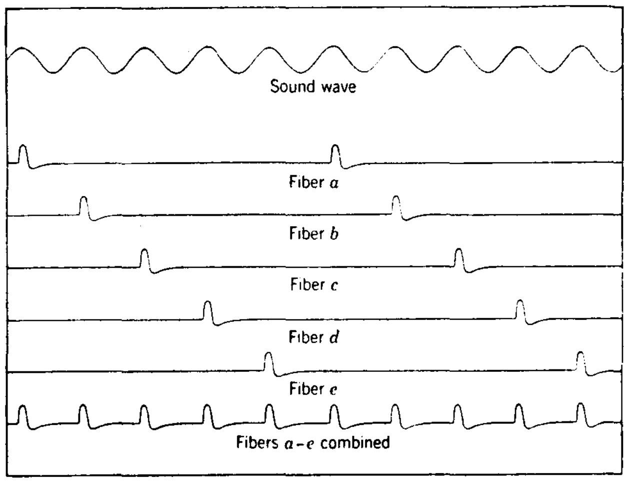


figure 2.2.: The volley principle. Each fiber responds to certain parts of the sound wave, and their total responses represent the full frequency of the wave (Wever, 1949).

Schouten (1940) repeated Seebeck's experiments and showed the inadequacy of the distortion hypothesis of Helmholtz. In order to explain the phenomenon of the "missing fundamental", Schouten formulated his "residue theory" (Schouten,1940). According to this theory the auditory system performs a frequency analysis on the sound stimulus (Place theory) followed by a time analysis (Time theory). The lower frequency components locally excite the basilar membrane. These lower partials are detected as resolved pure tone components. Several of the higher partials will fall within the resonance bandwidth of a given region of the membrane. The unresolved components form the so-called "residue". The time pattern of the residue on the basilar membrane closely resembles that of the stimulus as a whole (fig.2.3), and is encoded in the time distibution of spikes carried by the auditory nerve. Both place and time theory are strictly cochlear based theories. According to the place theory of pitch perception, pitch is mediated by means of spectral clues present in the cochlear output. Time theory assumes temporal clues in the coclear output to represent the pitch information.

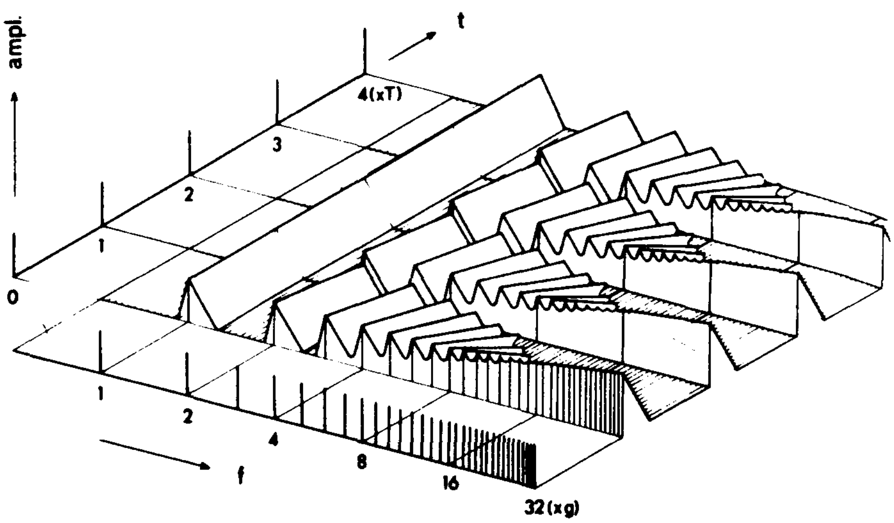


figure 2.3.: Response of the basilar membrane to a series of periodic impulses. linear time scale, logarithmic frequency and amplitude scales. The waveform of the sound stimulus can be seen on the left and its spectrum at the bottom of the figure.For the central part of the figure, the frequency axis can be seen as a position axis, indicating distance along the basilar membrane (Moore,1982).

Experiments of Houtsma and Goldstein (1972) on the musical pitch of two-tone complexes showed the inadequacy of these cochlear based theories. Complex tones comprising two successive harmonics were presented dichotically to the listener (one component to the left ear, the other to the right ear). Despite this dichotic presentation of the stimuli and the absence of unresolved components, the subjects were able to perceive a residue pitch. These experiments led to the hypothesis that the pitch of complex sounds was formed by a central neural processor which operates on signals derived from peripherally resolved stimulus components. Modern pitch perception theories, particularly the "optimum processor theory" (Goldstein,1973) and the "virtual pitch theory" (Terhardt,1974) are based on this hypothesis. These two theories of pitch perception will be discussed in more detail in the next section.

2.2 Theoretical models

2.2.1 The optimum processor theory (Goldstein)

The optimum processor theory (Goldstein,1973) was developed to account for experimental data on musical pitch perception of two-tone complexes (Houtsma and Goldstein,1972). A more general formulation of the theory was given by Gerson and Goldstein (1978). Duifhuis,Willems and Sluyter (1982) developed a practical pitch meter based on Goldstein's theory of pitch perception. This pitch meter could be implemented on a digital computer (see section 2.3.1). Recently, a further generalization of Goldstein's theory was given by Beerends and Houtsma (1986). They showed that the optimum processor theory could account for pitch phenomena involving simultaneous complex tones.

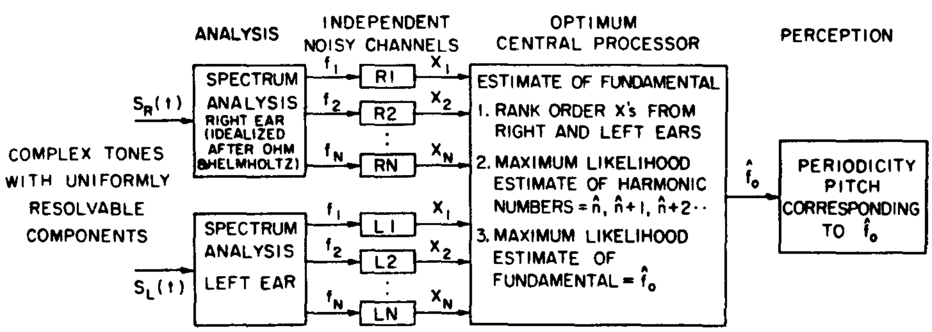


figure 2.3.: Schematic diagram of the optimum processor theory for the central formation of periodicity pitch of complex tones (Goldstein, 1973).

The four stages of processing of the optimum processor theory are shown in figure 2.3 . The first stage describes the spectral analysis performed by the auditory system. Both ears are represented as a spectrum analyzer. The output of this stage consists of a set of individually resolved spectral components. Amplitude and phase information of the components are ignored. Next, the frequency information of the resolved components is conveyed stochastically to the optimum central processor. The optimum processor maximizes the correlation between the estimated frequencies and a harmonic template. The fourth stage represents the final pitch perception of the fundamental frequency corresponding to the best matching template. Next a mathematical formulation of the optimum central processor theory will be given.

The frequencies of the individually resolved stimulus components $\{f_i\}$ are presented to the central processor as samples $\{x_i\}$ of Gaussian distributions G.

$$G(f_k,\sigma_k) = (\sigma_k\sqrt{2\pi})^{-1}.\exp[\ -(x_k-f_k)^2/2\sigma_k^2\ ] \qquad (2.1)$$

f_k is the mean of the Gaussian distribution. The central processor assumes that f_k equals $n_k.f_0$. σ_k is the standard deviation of the Gaussian distribution. This standard deviation is known to the processor and is a function of frequency only. Houtsma (1979) gives the following idealized standard deviation function :

$$\begin{aligned} &\sigma(f)=0.3125\ f^{1/2} \quad &,f\leqslant 2500\ \text{Hz.} \\ &\sigma(f)=10^{-9}\ f^3 \quad &,f>2500\ \text{Hz.} \end{aligned} \qquad (2.2)$$

(σ and f in Hz.)

The central processor makes an optimum estimate of the fundamental $\hat{f}_0$ and the harmonic numbers $\hat{n}_k$, which maximize the likelihood function L (Van Trees,1968).

$$L = \prod_k(\sigma_k\sqrt{2\pi})^{-1}.\exp[\ -(x_k-n_kf_0)^2/2\sigma_k^2\ ] \qquad (2.3)$$

Equation (2.3) is replaced by it´s natural logarithm (monotic transformation) :

$$\Lambda = -\sum_k \ln(\sigma_k\sqrt{2\pi}) - 1/2\sum_k(x_k-n_kf_0)^2/2\sigma_k^2 \qquad (2.4)$$

The following assumptions on σ_k are made :

1. $\sigma(f)$ is proportional to f so that $\sigma(f_k)=K(f_k).f_k=K(f_k).n_k.f_0$
2. $\sigma(f_k)$ is constant in the immediate vicinity of f_k ;$\sigma(f_k)=\sigma(n_k f_0)\approx\sigma(x_k)$

Because the first term on the right hand of eq.(2.4) is constant, maximizing (2.4) is equivalent to a least-squares procedure, where

$$\varepsilon^2 = \sum_k (x_k - n_k f_0)^2 / 2\sigma_k^2 \qquad (2.5)$$

is minimized and the σ_k values are known.($\sigma(n_k f_0)=\sigma(x_k)$)

Given the optimum values of n_k ($\hat{n}_k$) the solution for the pitch $\hat{f}_0$ follows from :

$$\left.\frac{\partial \varepsilon^2}{\partial f_0}\right|_{f_0=\hat{f}_0} = 0 \qquad (2.6)$$

$$\hat{f}_0 = \sum_k \hat{n}_k x_k/\sigma_k^2 \Big/ \sum_k \hat{n}_k^2/\sigma_k^2 \qquad (2.7)$$

In order to estimate the pitch $\hat{f}_0$, the set of integers $\{\hat{n}_k\}$ must be estimated first. By substituting eq.(2.7) into eq.(2.5) the ε^2-function can be computed for all possible combinations of n_k. The processor determines the set $\{\hat{n}_k\}$ which minimizes ε^2. This set is used in eq.(2.7) to estimate the pitch $\hat{f}_0$.

The stochastic transformation of the frequency information of the resolved stimulus components ($\{f_i\}$ -> $\{x_i\}$) causes errors in the estimated set of harmonics (see fig.2.3). These errors give rise to ambiguities in the pitch perception.

In the original presentation of the optimum processor theory (Goldstein, 1973) the central processor presumes that the stimulus consists of a set of successive harmonics. This successive harmonic constraint was removed by Gerson and Goldstein (1978). According to them, the central processor estimates periodicity pitch by optimizing the match between the aurally measured frequencies and a general harmonic template over a certain predetermined pitch range.

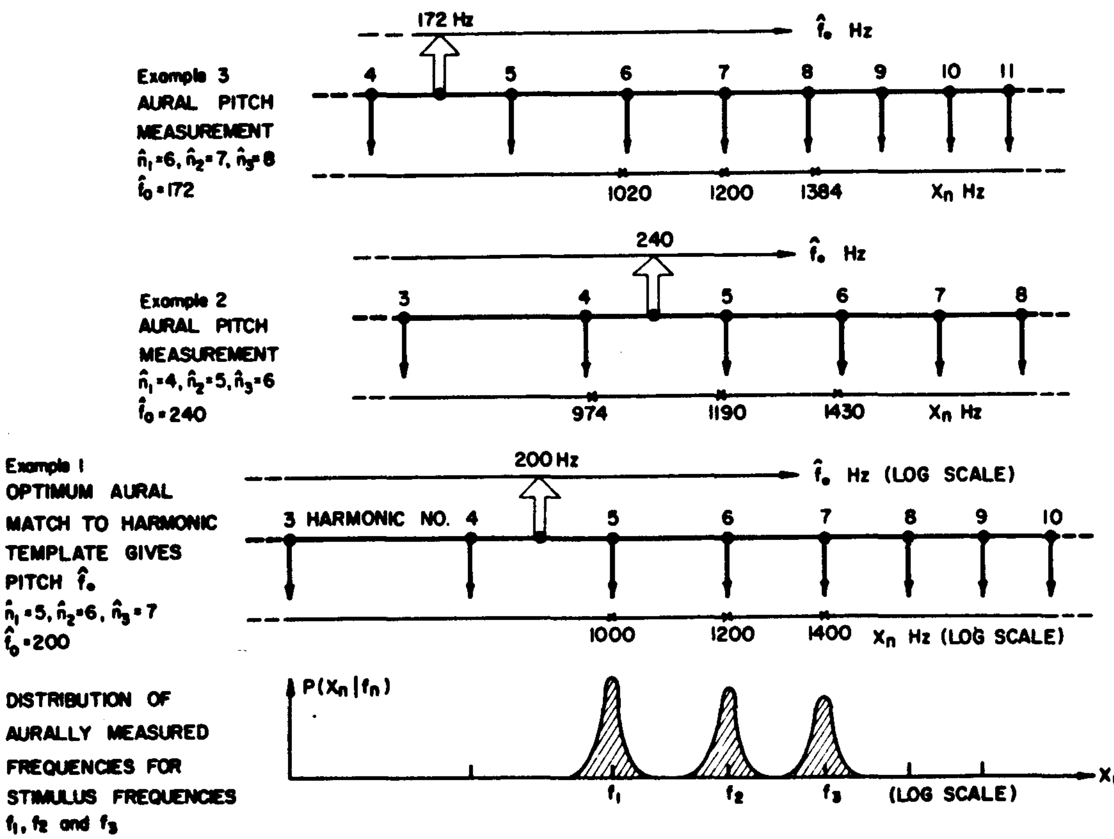


figure 2.4.: Periodicity pitch as a harmonic frequency pattern recognition. Bottom: Probability distribution of aural estimates of frequencies of component tones in a complex (three) tone stimulus comprising 5th,6th and 7th harmonics of 200 Hz. Examples 1-3 illustrate best matches of harmonic frequency template to three different sets of component frequency estimates. Large errors in f_0 occur when the harmonics are misaligned (e.g. 2 and 3) (Goldstein et al.,1978).

Goldstein derived an analytical solution for two-tone complexes comprising successive harmonics (Goldstein,1973). In order to predict the results of experiments on pitch perception comprising more than two harmonics, Gerson and Goldstein (1978) used a computer simulation technique.By making some essential changes Duifhuis,Willems and Sluyter (1982) managed to implement a pitch meter based on Goldstein's theory (see section 2.3.1).

2.2.2 The virtual pitch theory (Terhardt)

Terhardt developed his virtual pitch theory in order to explain two different modes, usually involved in the pitch perception of complex tones (Terhardt,1974). The principles of the virtual pitch theory are shown in figure 2.5 .

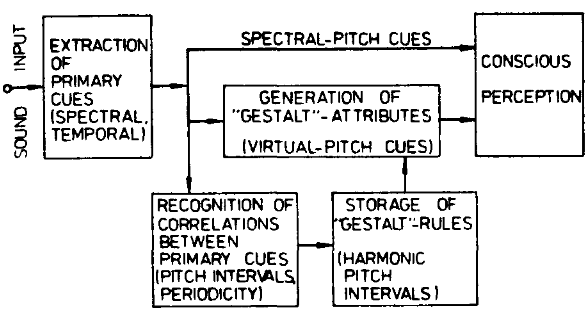


figure 2.5.: Schematic diagram of the concepts involved in the virtual pitch theory (Terhardt, 1974).

In the analytic mode, a particular pure tone component of a complex sound can be heard as a spectral pitch. In the synthetic mode, the complex sound is percieved as a single entity, having a certain virtual pitch. ("Gestalt" perception) Figure 2.6 shows a model for spectral- and virtual-pitch perception.

The perception of spectral pitches (analytic listen-mode) is described by the spectral pitch pattern. This SP-pattern contains the frequencies and the weights of all possible spectral pitch cues. Only spectral components of the sound stimulus that are resolved by the auditory system can serve as spectral pitch cues. The frequency of a cue depends on the sound pressure level of the component and on the masking effect of adjacent partials. The weight of the cues is also dependent on sound pressure level and masking effects. The cues are also weighted according to the principle of spectral dominance. Numerical expressions of the above mentioned effects are discussed in section 2.3.2 .

The virtual pitch pattern (VP-pattern) describes the synthetic listen-mode. In order to understand the mechanism of virtual pitch formation, the learning matrix should be discussed in more detail (see figure 2.6).

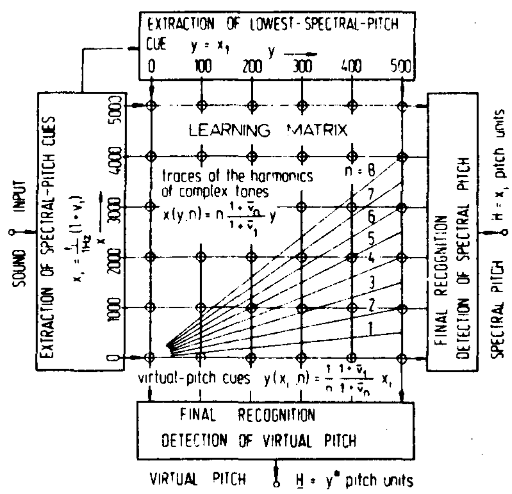


figure 2.6.: Model of spectral- and virtual-pitch perception. The learning matrix consists of a great number of columns and rows, from which only six are drawn (Terhardt, 1974).

The spectral pitch cues x_i are transmitted from left to right along the rows of the matrix.In the learning phase, the lowest spectral pitches are fed into the vertical lines of the matrix. If there is a correlation between the lowest spectral cue y_i and another spectral cue x_i, the resistance R_{ij} in the coincidence detection mechanism at the crosspoint (x_i, y_i) is lowered (see figure 2.7).

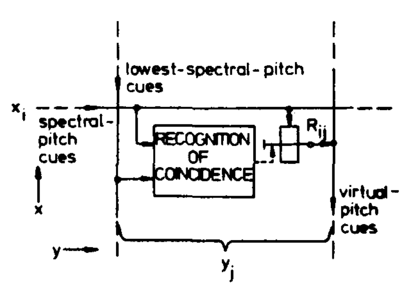


figure 2.7.: Structure of crosspoint $(x_i;y_i)$ of rows and columns of the learning matrix. One column with position y_i consists of two transmission lines, the left one transmits lowest spectral pitch cues into the matrix, the right one transmits the virtual pitch cues which are produced in the matrix to the final recognition system of virtual pitch. Repeated coincidence of cues at (x_i, y_i) causes reduction of the resistance R_{ij} (Terhardt, 1974).

In his early live, a human being acquires the ability to identify voiced speech sounds. When the learning matrix is stimulated by speech sounds, which are characterized by a strong correlation between the individual harmonics, the resistances R_{ij} are systematically lowered at those cross-points (x_i, y_i) where:

$$x_i = n.y_i \qquad , n=1,2,......,M. \qquad (2.8)$$

Pitch shifts of x_i and y_i introduced in the extraction process of spectral pitch cues are ignored in this formula. In the recognition or "post learning" phase a spectral pitch cue x_i produces a set of M virtual pitch cues, at the positions:

$$y_i = (1/n).x_i \qquad , n=1,2,......,M. \qquad (2.9)$$

The virtual pitch cues are subharmonics of the spectral pitch cues. If the complex stimulus provides R spectral pitch cues there will be RxM virtual pitch cues. The significance of a particular virtual pitch cue is expressed by its weight. Terhardt (1979,1982) gives four criteria which determine the weight of a virtual pitch.

1. Weight increases with the number of components that provide the same or nearly the same virtual pitch cue.
2. Weight increases with increasing spectral pitch weight.
3. Weight decreases with increasing subharmonic number.
4. Weight increases with increasing accuracy of the near coincidences.

These criteria are combined into one coincidence coefficent. This coefficient is used by an algorithm which extracts the virtual pitch pattern of a complex sound. The algorithm is discussed in section 2.3.2 .

The weights of the spectral and virtual pitches describe the prominance of the pitches within each group and not between the two groups. The virtual pitch theory, like the central processor theory of Goldstein, accounts for most experimental data on the perception of pitch of complex tones (Terhardt et al,1982; Houtsma,1979).

2.3 Pitch meters based on human pitch perception

Practical pitch meters based on the modern pitch theories, described in section 2.2, were used in comparing experimental results with theoretical predictions. The DWS-pitch meter was developed at the IPO. During this project an implementation of this meter was available. The virtual pitch extractor, described by Terhardt (1979,1982), was implemented on a VAX 11/780 computer system. Other pitch detection algorithms (see Rabiner et al.,1976) are not discussed in this report, because they tend to pay little attention to perceptual aspects of pitch; pitch is treated as a purely technical issue. The pitch meters that will be discussed in this section operate according to the principles of the human pitch extractor. They are more suited to the task of determining the subjective strike note pitch of a bell (Duifhuis et al.,1982).

2.3.1 The DWS-pitch meter

The DWS-pitch meter consists of a spectral analyser and a harmonic pattern recognizer. The harmonic pattern recognizer estimates the pitch of a sound segment on the basis of the spectral components found by the spectral analyser.

The spectral analyser

The spectral components are determined with the use of a Fast Fourier Transform. Parameters of the FFT are shown in table 2.1 .

Number of points	: 256
Sample frequency	: 5 kHz.
Frequency range	: 0 - 2.5 kHz.
Time window	: 51.2 ms. (40 ms acoustic signal, 11.2 ms silence)
Frequency resolution	: 19.5 Hz.
Window	: Hamming-window

Table 2.1.: Fast Fourier Transform parameters used in DWS-meter.

A spectral sample with amplitude A_i is considered a local maximum when $A_i >$ A_{i-1}, and $A_i > A_{i+1}$. It is checked if the local maximum A_i is above a certain threshold. This fluctuating threshold is 26 dB below the level of the highest spectral peak. A check is also made whether the local maximum isn't masked by a lower spectral component. When A_i is above these thresholds, the amplitude and frequency of the local maximum are determined more precisely by means of a parabolic interpolation. Finally the shape of the peak is checked. The spectral analyser determines only the six lowest components.

The harmonic pattern recognizer

The DWS-meter was mainly designed to estimate the pitch of speech sound. In order to deal with the absence of certain harmonic frequencies in the speech spectrum, as well as with the problem of spurious inharmonic components, Duifhuis et al. introduced their harmonic sieve procedure. The harmonic sieve operates in the frequency domain. The sieve consists of 11 meshes with a bandwidth W(f). The position ℓ of the sieve along the frequency scale is characterized by the value $f_{0\ell}$ of its fundamental frequency. The position of a particular mesh j is given by its harmonic frequency $f_j = j.f_{0\ell}$, with j=1,2,...,11 . The bandwidth of the meshes is proportional to their center frequencies, $W(f)=2\alpha jf_0$. The sieve scans the fundamental frequency range (50 Hz. < $f_{0\ell}$ > 500 Hz.) in discrete steps (index ℓ, ℓ=1,2,...,L). These steps should be smaller than W(f), in order not to miss any valuable information. When α is chosen 5%, the meshes just don't overlap and the fundamental frequency range is stepped most efficient. A component X_i passing through mesh j is given an index $m_{k\ell}=j$, where ℓ indicates the sieve position. A number of sets $\{m_k\}\ell$ is obtained, if the sieve scans the fundamental frequency range. Each set $\{m_k\}\ell$ can be regarded as a vector in a multi-dimensional space. In case of the optimum set $\{\hat{n}_k\}$, the normalized distance between the corresponding vector and a vector which characterizes a reference set (ideal harmonic pattern) is minimized. Minimizing the normalized distance is equivalent to minimizing the quantity C_ℓ:

$$C_\ell = (M_\ell + N_\ell)/K_\ell \qquad (2.10)$$

where : M_ℓ= highest harmonic candidate $m_{k\ell}$,
N_ℓ= total number of components minus those for which $m_{k\ell}>11$,
K_ℓ= total number of components which passed the sieve.

The minimum of C_ℓ over $\ell=1$ to L determines the optimum set $\{\hat{n}_k\}$. The corresponding fundamental frequency is estimated using eq.2.7 of section 2.2.1, with $\sigma(f)=\sigma=$constant :

$$\hat{f}_0 = \sum_k^{N\ell} x_k \hat{n}_k \Big/ \sum_k^{N\ell} \hat{n}_k^2 \qquad (2.11)$$

where $\{x_k\}$ are the detected component frequencies.

In the DWS-meter the fundamental frequency range is scanned in 81 discrete steps, for every 40-ms acoustic segment. Allik et al.(1982) noted that at many positions of the sieve, no components passed through it, because the meshes didn't coincide with the input spectral components. They concluded that there was a lot of useless computation. They proposed an algorithm which fixed the position of the sieve on the basis of an input frequency. Their algorithm, called the common denominator procedure, was more than 20 times faster than the harmonic sieve procedure.

2.3.2 The virtual pitch extractor (Terhardt)

Like the DWS-pitch meter, Terhardt´s virtual pitch extractor (hereafter called the TVP-meter) consists of two elements :

- a spectral analyser which derives the spectral pitch pattern (SP-pattern).
- a (sub)harmonic pattern recognizer, which determines the virtual pitch pattern (VP-pattern).

The spectral analyser

The tonal components are extracted automatically from the FFT spectrum of the sound stimulus. Information about the parameters of the FFT analysis, used by the TVP-meter, is shown in table 2.2.

Number of points	: 800
Sample frequency	: 10 kHz.
Frequency range	: 0 - 5 kHz.
Time window	: 80 ms.
Frequency resolution	: 12.5 Hz.
Window	: Hanning-window

Table 2.2.: Fast Fourier Transform parameters used in TVP-meter.

In order to be detected as a tonal component, a spectral sample with relative sound pressure level L, has to satisfy the following conditions:

$$1.\ L_{i-1} < L_i \geq L_{i+1} \qquad (2.12)$$
$$2.\ L_i - L_{i+j} \geq 7\ \text{dB} \qquad ;\ j=-3,-2,+2,+3.$$

A more precise component frequency(about $\pm$ 1 Hz. accuracy) is obtained by means of an interpolation formula. Due to masking effects, some components may become inaudible while others are reduced in audibility. Terhardt (1982) takes the difference between the components SPL and that SPL which expresses the masking effect of the other partials, as a measure of the relevance of the component. The so-called sound pressure level excess of the ith tonal component, LX_i is given by :

$$LX_i = L_i - 10\log_{10}\left[\left(\sum_{\substack{k=1 \\ k\neq i}}^{N} 10^{L_{Ek}(f_i)/20}\right)^2 + I_{Ni} + 10^{L_{TH}(f_i)/10} \right] \qquad (2.13)$$

where: L_i = SPL of the ith component,

N = total number of detected tonal components,

$L_{Ek}(f_i)$ = exitation level produced at f_i by the kth component,

I_{Ni} = noise intensity present in the critical band around the ith component,

$L_{TH}(f_i)$ = hearing threshold at frequency f_i .

Numerical expressions for $L_{Ek}(f_i)$, I_{Ni} and L_{TH} are given in appendix A1.

The pitch of an isolated pure tone undergoes a frequency shift when the same tone is presented as a component of a complex tone. It´s also known that the pitch of a pure tone depends on its sound pressure level. The pitch of a spectral cue with frequency f_i is given by :

$$H_i=f_i\ (1+v_i) \qquad (2.14)$$

H_i is called the <u>true</u> spectral pitch, whereas f_i is the <u>nominal</u> spectral pitch. The quantity v_i represents the pitch shift. A formula for the calculation of v_i is shown in appendix A1. The spectral pitches are weighted in order to express their relative prominence. The weight WS_i of a spectral pitch cue depends on its SPL excess LX_i and its frequency f_i :

$$WS_i=[1-\exp(-LX_i/15)][1+ 0.07((f_i/0.7)-(0.7/f_i))^2]^{-1/2} \quad ,LX_i\geqslant 0$$

$$WS_i=0 \quad ,LX_i < 0 \qquad (2.15)$$

LX_i is expressed in dB. and f_i in kHz.

The first factor describes the influence of the SPL excess on the weight. Weight increases with increasing LX and saturates at about 20 dB SPL. The second factor accounts for the frequency dependence of the weight. The final result of the spectral analysis part of the TVP-meter is the spectral pitch pattern. The SP-pattern contains the nominal and true spectral pitches and their corresponding weights. The following stage of the TVP-meter, the subharmonic pattern recognizer, uses the SP-pattern as input.

The subharmonic pattern recognizer

According to the virtual pitch theory, every subharmonic of a spectral pitch cue represents a possible virtual pitch cue (see section 2.2.2). The pitch of the virtual cues is specified by :

$$\underline{H}_{im} = f_i/m \qquad (2.16)$$

where $\underline{H}_{im}$ is the mth subharmonic of the ith spectral pitch.

In case of a virtual pitch, the character $\underline{H}$ is underlined. This formula doesn't account for small pitch shifts, therefore $\underline{H}_{im}$ is called the nominal virtual pitch. An expression for the true virtual pitch is described in appendix A1. There are RxM potential virtual pitches, when the SP-pattern contains R relevant spectral pitches and M is the maximal subharmonic number taken in consideration. Terhardt et al. (1982) formulated an efficient algorithm which evaluates the significance of the many potential virtual pitches.

First, the algorithm sorts the spectral pitch weights WS_i in terms of descending weights. Only those spectral pitches with a weight of at least 70% of the strongest pitch are considered. The coincidence coefficient of the first subharmonic (m=1) of the first spectral pitch with any subharmonic of the second spectral pitch is calculated. The subharmonic number n of the second spectral pitch (f_j) which is most nearest to the subharmonic of the first spectral pitch (f_i) can be determined at once by the following formula :

$$n = Int(m.f_j/f_i + 0.5) \qquad (2.17)$$

Next, a coincidence coefficient C_{ij} is calculated :

$$C_{ij} = (WS_i WS_j/mn)^{1/2}(1-\gamma/\delta), \quad \text{if } \gamma \leqslant \delta,$$
$$C_{ij} = 0, \text{ if } \gamma > \delta \text{ and/or } n > 20. \tag{2.18}$$

where : WS : spectral pitch weight,VP-weight increases with increasing spectral pitch weight.

n,m : subharmonic numbers, VP-weight decreases with increasing subharmonic number

γ: degree of inharmonicity, $\gamma = |(nf_i/mf_j)-1|$, VP-weight decreases with increasing inharmonicity.

δ: width of the coincidence interval. VP-weight increases with increasing accuracy of coincidence.

Hereafter, the coincidence coefficient of the first subharmonic (m=1) of the first spectral pitch with any subharmonic of the third spectral pitch is calculated, and so on. The result is a virtual pitch weight W_{11}, which is the sum of the coincidence coefficients C_{1j} . The same procedure is repeated for the other subharmonics of the first spectral pitch, resulting in the virtual pitch weights W_{1m}, m=2,3,...,12. In general, W_{im} is the virtual pitch weight which corresponds to the mth subharmonic of the ith spectral pitch. The total result of the subharmonic pattern recognizer is the VP-pattern. The VP-pattern contains the virtual pitches and their weights.

3 THE ACOUSTICS OF BELLS

3.1 Physical aspects

A bell is a hollow vibrating object with axial symmetry. The names used to indicate the various parts of a bell are shown in figure 3.1.

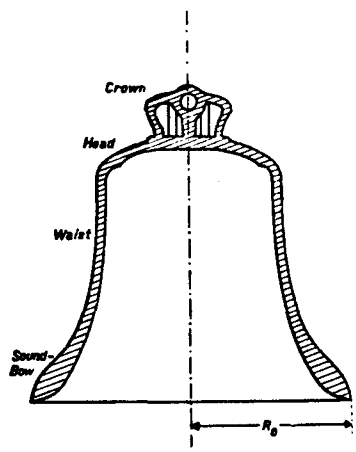


Figure 3.1.: Cross-section of a bell, cut along a plane through the axis of rotation. The bell is suspended on the crown.The waist is nearly cylindrical shaped. The lower part of the bell, called the sound-bow has a greater wall thickness, and is hit by the clapper or the hammer (van Heuven,1949).

The physical sound spectrum of a bell contains a great number of characteristic eigenfrequencies, the so-called partials (see figure 3.2). Each partial corresponds to a certain mode of vibration of the bell. The partials of a bell are, as distinct from most other musical instruments, not harmonically related. Despite this inharmonicity one perceives a clear pitch when a tuned bell is struck, usually referred to as the strike note pitch. The first eight vibrational modes of a correctly tuned bell, and their corresponding names and frequencies are shown in figure 3.3. Usually the eigenfrequencies are expressed relative to the strike note pitch. (f_s) If the frequencies of the partials are given relative to the "hum", the following typical series of frequency ratios can be seen : 1,2,2.4,3,4,5, 6,8 . The minor third is the only non-integer ratio of the series. Being quite prominent in amplitude, the minor third gives the bell it´s distinct sound quality.

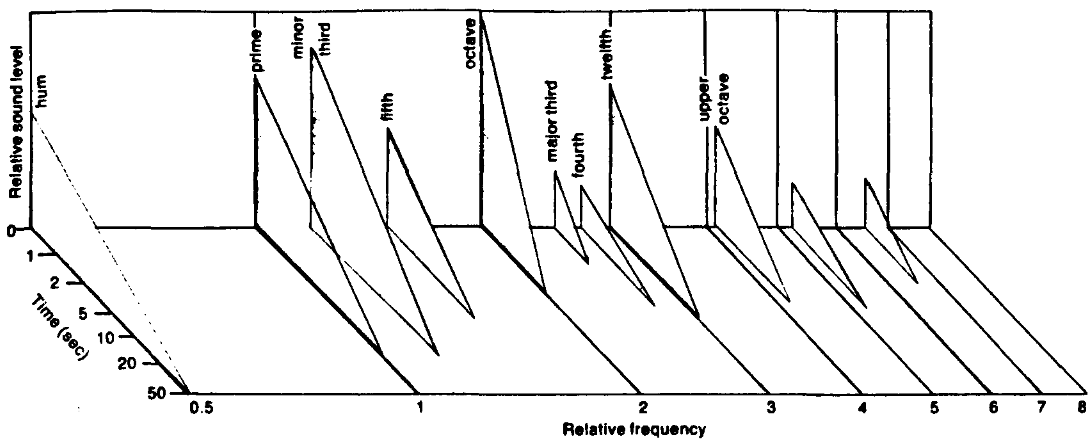


figure 3.2.: The sound spectrum of a bell contains a great number of partials, each with its own frequency, initial amplitude and decay time. The tonal character of the sound changes with time elapsed after the bell has been struck (Rossing,1984).

The mathematical description of the vibrations of a bell is rather complex. However, the vibration pattern of a bell is assumed to be composed of a set of different normal modes of vibration. These normal modes of vibration are described by their latitudinal (circular) nodes and longitudinal (meridian) nodes. There is no motion normal to the surface on the nodal lines. The first eight vibrational modes of a bell are depicted in figure 3.3 .

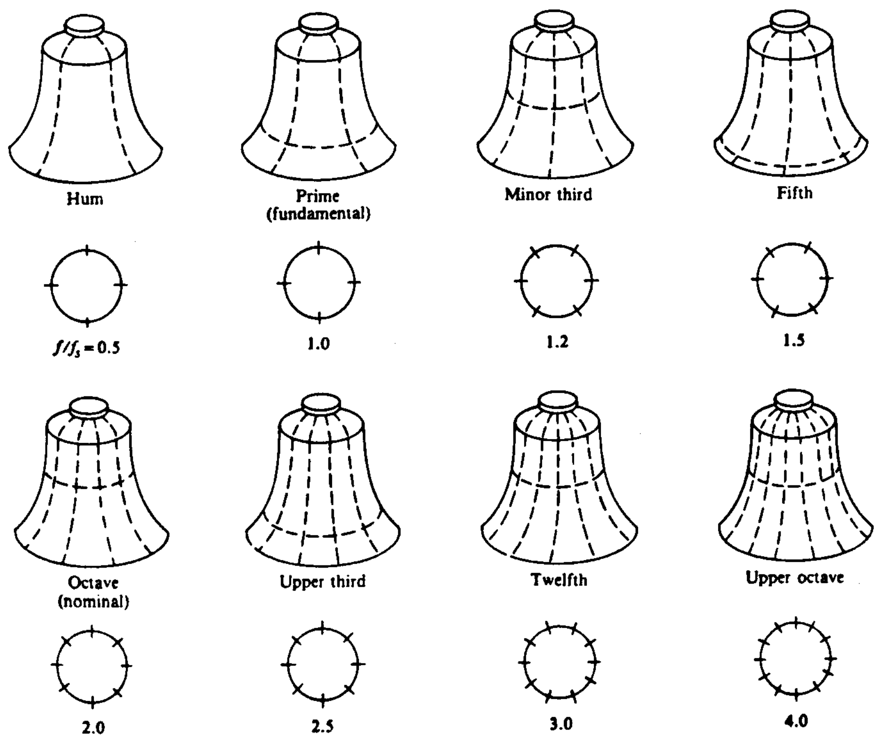


figure 3.3.: The first eight vibrational modes of a tuned bell. The approximate locations of nodes are indicated by dotted lines. The partial frequencies are given relative to the strike note pitch f_s (Rossing,1982).

According to Lehr (1965) the hum tone has two complete "over-the-crown" nodes, which devide the bell surface into four regions. These four regions move alternately inward and outward. Usually, the normal modes are catagorized into different groups, according to the number and location of their circular and meridian nodal lines (Tyzzer,1930; Grutzmacher et al., 1965; Lehr,1965; Perrin et al.,1983). Table 3.1 gives the classification of the bell modes of figure 3.3 (according to Lehr,1965).

Name of mode	Ratio to strike note	Classification
Hum	0.5	I-2 *)
Prime	1.0	II-2 **)
Minor third	1.2	I-3
Fifth	1.5	II-3
Octave	2.0	I-4
Upper third	2.5	II-4
Twelfth	3.0	I-5
Upper octave	4.0	I-6

Table 3.1.: Classification of the first eight vibrational modes of a tuned bell (After Lehr,1965).
*) Experiments on bell profiles showed that in case of the hum note the nodal circle, characteristic for group I, is shifted to the crown of the bell.
**)The same experiments showed that the nodal circle for group II in a conical bell coincides with the nodal circle for the prime in the normal bell.

Group I : The only nodal circle is located in the waist of the bell. An antinode occurs at the soundbow.

Group II : The nodal circle lies at the soundbow and an antinode is located in the waist

3.1 Psychophysical aspects

Schouten and ´t Hart (1965) describe three different characteristics of the total sound of a bell. In the first place, one can hear the very short atonal strike sound of metal on metal. The strike note is the second characteristic of the total bell sound. It has a sharp timbre and dies out quickly. The strike note has a clear pitch which is almost the same as the pitch of the prime, the third part of the bell sound that can be heard. The prime has a soft timbre and a longer decay time.

Lord Rayleigh (1890) carried out some of the earliest investigations on bells. He found that in most of the bells he examined the strike note pitch didn´t coincide with one of the partials. As a result of his research, he put forward the "octave rule", for the pitch determination of a bell. According to this rule, the strike note pitch of a bell is perceived one octave below the fifth partial (The octave, I-4). The octave rule was an empirical rule which didn´t gave any explanation of the origin of the strike note. Jones (1930) proved that the strike note was a subjective tone, which could not be amplified by a Helmholtz resonator. In the following, a historical survey of the possible explanations of the strike note will be given.

The explanation of the "misjudged octave" comes from Jones (1930). This theory of the strike note suggests that the pitch of the strike note is determined by the strong fifth partial. Due to the presence of the second (prime) and third (minor third) partials, the ear is mislead and the fifth partial is perceived an octave too low.

Meyer and Klaes (1933) showed that a sound consisting of the lower five partials could not evoke the strike note pitch sensation. Only when the seventh partial (twelfth) was also added, the strike note was perceived. They showed that the strike note was related to the higher partials, so the "misjudged octave" hypothesis of Jones was rejected. Meyer and Klaes formulated the "difference tone" rule. According to this rule, the subjective strike note was formed by nonlinear distortion in the inner ear as a difference tone between the strong fifth and seventh partials. Arts

(1938, 1939) described many practical situations where the strike note corresponded much better to the octave rule than the difference rule of Meyer and Klaes.

A completely different explanation was suggested by van Heuven (1949). He assumed that the strike note pitch was a result of the interaction between the clapper and the bell. Jones (1930) had also studied the contact time between the clapper and the bell. The subjective nature of the strike note seemed to contradict this hypothesis.

Schouten (1940) performed experiments on complex tones which lacked a fundamental frequency. Although the fundamental frequency was missing, one could hear a clear pitch corresponding to the pitch of the fundamental. In order to explain his findings, Schouten formulated the "residue" theory. (see section 2.1.1) According to the residue theory the strong partials, named the octave (I-4), the twelfth (I-5) and the upper octave (I-6) with the frequency ratios 2:3:4, form a residue. This tonal residue will be perceived as a whole, with a pitch corresponding to the fundamental frequency of its harmonic constituent components. This fundamental frequency is perceived as being the strike note. The theory also predicts the existence of a secondary strike note, created by the I-6,I-7,I-8 and I-9 partials with the frequency ratios 3:4:5:6. The residue theory was strongly supported by Pfundner (1962).

Recent investigations of bells were carried out by Greenhough and Terhardt. Greenhough (1976) systematically determined the realative importance of the first eight partials of a bell-like sound, with respect to the strike note pitch. Greenhough used bell-like sounds, generated on a computer by means of additive synthesis. The amplitudes and decay times were equal for every one of the eight partials. However, the frequencies of the partials resembled those of a real bell. By measuring subjective pitch shift of the strike note as a function of systematic frequency shift in selected partials, Greenhough found that, besides the octave (I-4), the double octave (I-6) was an even more important partial (see figure 3.4). It would be interesting to investigate whether the validity of the results of Greenhough could be extended to real bells.

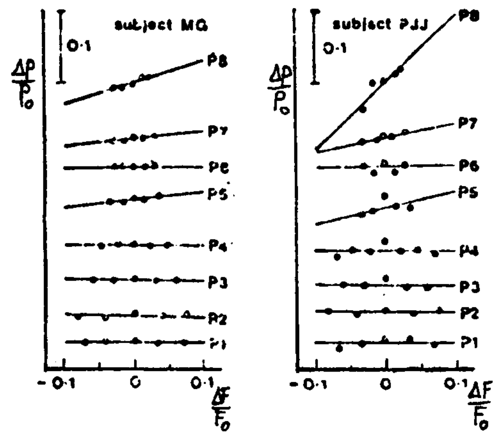


figure 3.4.: Relation between the shift P of the original strike note pitch P_0 and the shift F of the original partial frequency F_i .(i=1,2,...,8) The vertical separation of the curves is for convenience only (Greenhough, 1976).

Terhardt (1984) compared experimental determined bell pitches to theoretical predictions, made by the virtual pitch theory.(sections 2.2.2 and 2.3.2) The subjects were instructed to sing a tone which had a pitch equal to the pitch of the bell heard. The first sung pitch was recorded and matched by the experimenter with a sine tone. It should be noted that it´s not at all an easy task to reproduce at once the pitch of a bell by means of singing. In 79% of 137 historical church bells, the musical category of strike notes was correctly identified by the virtual pitch algorithm of Terhardt (1979, 1982).

4 ANALYSIS AND SYNTHESIS OF BELL STIMULI

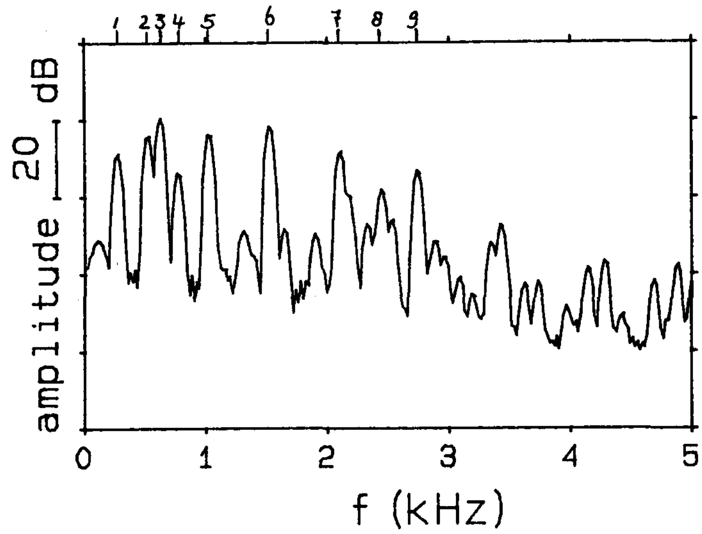
A number of bell recordings were placed at our disposal by the Royal Eysbouts bell foundry. A particular bell of this set was used as a basis for creating the bell stimuli involved in the experiments. This "C2"-bell was cast in 1675 by P.Hemony and belonged to the carillon of Gouda (Main Church). The recording was made on a Sony PCM-F1 digital audio recorder using a Senheisser microphone (MKH 816).

4.1 Analysis

The bell recording was low-pass filtered at 4.3 kHz and stored in a VAX 11/780 computer using a 12-bit analog-to-digital converter and a sampling rate of 10 kHz. A FORTRAN IV computer program was written which calculated the sound spectrum of the bell using a Fast Fourier Transform (FFT) subroutine of the ILS-library (Interactive Laboratory System for signal processing).

A FFT is an algorithm which calculates the Discrete Fourier Transform (DFT) in a fast and efficient way. The DFT can be calculated on a computer, because it is discrete in both time and frequency domain. However, the FFT can work only on a small part of the time signal. The original input signal therefore has to be segmented in time before the analysis. The result of the analysis may be influenced by this time limitation. The so-called leakage effect and the picket-fence effect are caused by this truncation of the input signal in the time domain (Brigham,1974;Thrane,1979). In order to reduce these time limitation effects, a window function was applied to the input signal. A Hanning or a Hamming window function could be chosen.

A 2048-point FFT representing 204.8 ms of the bell sound was calculated. Next, the computer program determined the local maxima of the discrete bell spectrum. More precise estimates of the peak amplitudes and frequencies were obtained by means of a parabolic interpolation. The results of the analysis are shown in figure 4.1 and table 4.1.



<u>figure 4.1.</u>: Mean amplitude spectrum of the Gouda "C2"-bell,
FFT : 256 points, 10 kHz samplefrequency, Hamming window,
Note: the mean of 60 spectra is plotted. This picture was used for visual inspection of the peaks only. More accurate estimates of the peak amplitudes and frequencies were calculated using a 2048-point FFT.

No.	Name of partial	Frequency (Hz)	Note	Ratio to strike note	
				ideal	real
1	Hum	250.6 ± 0.5	C4 -72 cents	.5	.51
2	Prime	500.8 ± 0.5	C5 -75 cents	1.0	1.01
3	Minor Third	603.0 ± 0.5	Es5-54 cents	1.2	1.21
4	Fifth	749.9 ± 0.5	G5 -77 cents	1.5	1.51
5	Octave	1004.8 ± 0.5	C6 -70 cents	2.0	2.02
6	Twelfth	1506.0 ± 0.5	G6 -70 cents	3.0	3.03
7	Upper Octave	2083.4 ± 0.5	C7 - 8 cents	4.0	4.19
8	-	2426.8 ± 0.5	ES7-44 cents	-	4.88
9	Double Undeciem	2721.0 ± 0.5	F7 -46 cents	5.33	5.47

Strike note of this bell is 497 Hz.

<u>Table 4.1.</u>: Frequency analysis Gouda "c2"-bell (cast by P.Hemony, 1675),
FFT : 2048 points, 10 kHz samplefrequency, Hamming window,
Note: 100 cents = 1 semitone .

4.2 The synthesis of bell stimuli

The aim of this project was to measure the pitch of the strike note of a bell as a function of a systematic frequency shift of selected partials. The seperate partials were shifted in frequency using a digital filtering technique. The very good preservation of the sound quality of the original bell was the main reason for choosing this digital filtering technique. The following steps can be distinguished in the synthesis of bell stimuli :

I. A specific partial is removed from the original bell sound using a digital band reject filter.

A digital filter is a linear time-invariant discrete system (LTI-system), which can be described in the time domain by its discrete impulse response h(n) (n is a integer). A digital filter can also be characterized by its discrete transfer function H(z). This transfer function is the z-transform of the impulse response of the filter ; $H(z) = \sum_{n=-\infty}^{n=\infty} h(n).z^{-n}$, where z is a complex variable (Rabiner and Gold,1975).

We used a digital filter with an infinite impuls response (IIR-filter). The transfer function H(z) of the filter is given by :

$$H(z) = \sum_{i=0}^{N} a_i.z^{-i} \;/\; [1+\sum_{i=1}^{N} b_i.z^{-i}] \qquad (4.1)$$

where N is the order of the filter.
It can be shown that a multiplication with z^{-1} in the z-domain corresponds to a one-sample delay in the time domain. The filter described by formula 4.1 therefore can be realized using seperate delay elements for both the input and output sequences of samples of the filter. The structure of the filter is shown in figure 4.2.

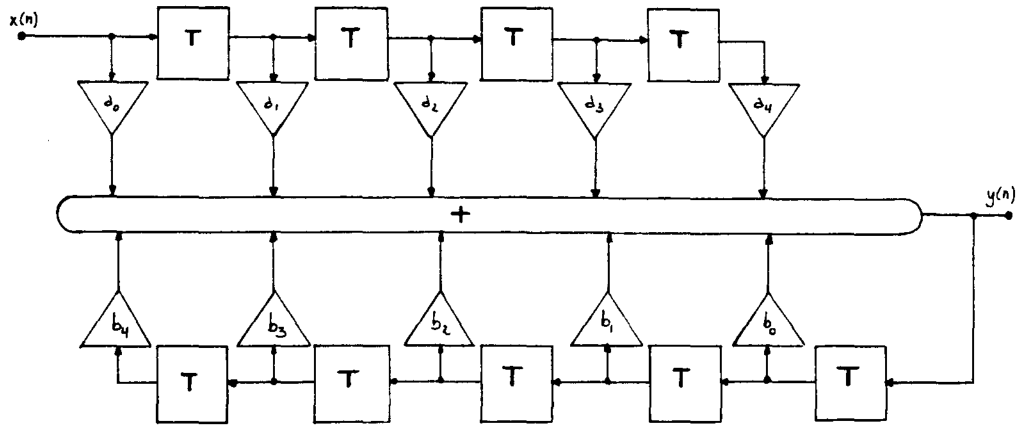


figure 4.2.: Structure of the digital filter used. This structure is called the "Direct form 1" (Van den Enden, Verhoeckx, 1985).

The digital filter is designed using the EFI-subroutine of the ILS-library. The EFI subroutine designs an elliptical filter as close to the input specifications as possible. EFI calculates the numerator (a_i) and the denominator coefficients (b_i) of formula 4.1. Besides the order of the filter and the sample frequency, one has to specify the frequency response of the IIR-filter. The characteristics of the magnitude response of the IIR-filter are described by the passband and stopband frequencies, the passband ripple and the stopband attenuation (see figure 4.3).

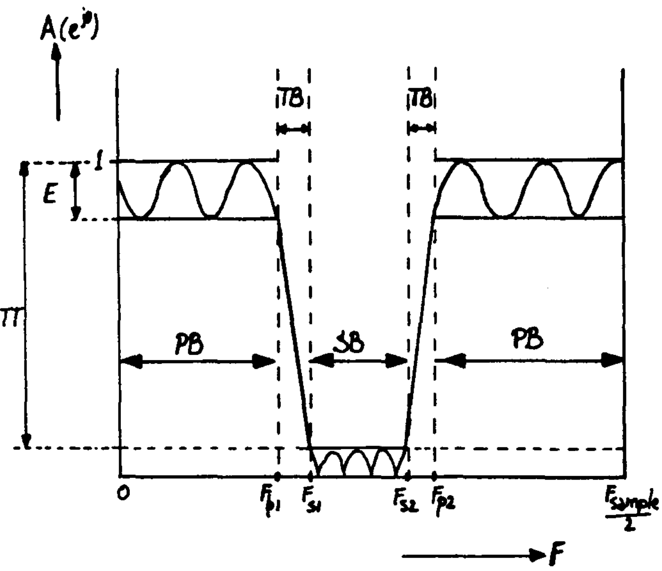


figure 4.3.: Specification of the magnitude response of a digital filter. E : Passband ripple, ATT: Stopband attenuation, PB : Passband, TB : Transition band, SB : Stopband, F_{p1}, F_{s1}, F_{p2}, F_{s2}, are frequencies of the edges of the various regions.

The EFI routine uses the following input parameters :

- Order of the filter N
- Passband ripple (in milli-dB)
- Sampling frequency (in Hz)
- Lower passband limit (F_{p1}), (in Hz),
- Upper passband limit (F_{p2}), (in Hz),
- Stopband attenuation, "dB-down" (in dB).

The subroutine gives a warning when input specifications lead to an unstable filter design. The following criteria were used to specify the input parameters :

- The filter should affect the selected partial only, i.e. the stopband, the transition band and the passband ripple have to be as small as possible.
- The filtering should be as effective as possible, i.e. a maximum stopband attenuation.
- The filter has to be stable.

EXAMPLE 1 :

This example describes the digital filter used to filter the minor third partial from the original bell sound.

Input specification : Order= 2, E= 1000 milli-dB, Sample frequency= 10 kHz, ATT= 100 dB down, F_{p1}= 570 Hz, F_{p2}= 640 Hz.

Result EFI routine : Order= 4, E= 1000 milli-dB, ATT= 100.273 dB down, F_{p1}= 570.00 Hz, F_{p2}= 640.00 Hz, F_{s1}= 603.88 Hz, F_{s2}= 604.19 Hz.

Note: In case of a bandreject filter, the order of the filter as calculated by the EFI-routine is always twice the specified order.

A fourier transform of the impulse response was calculated in order to obtain the frequency response of this filter. The impulse and frequency response are shown in figure 4.4.

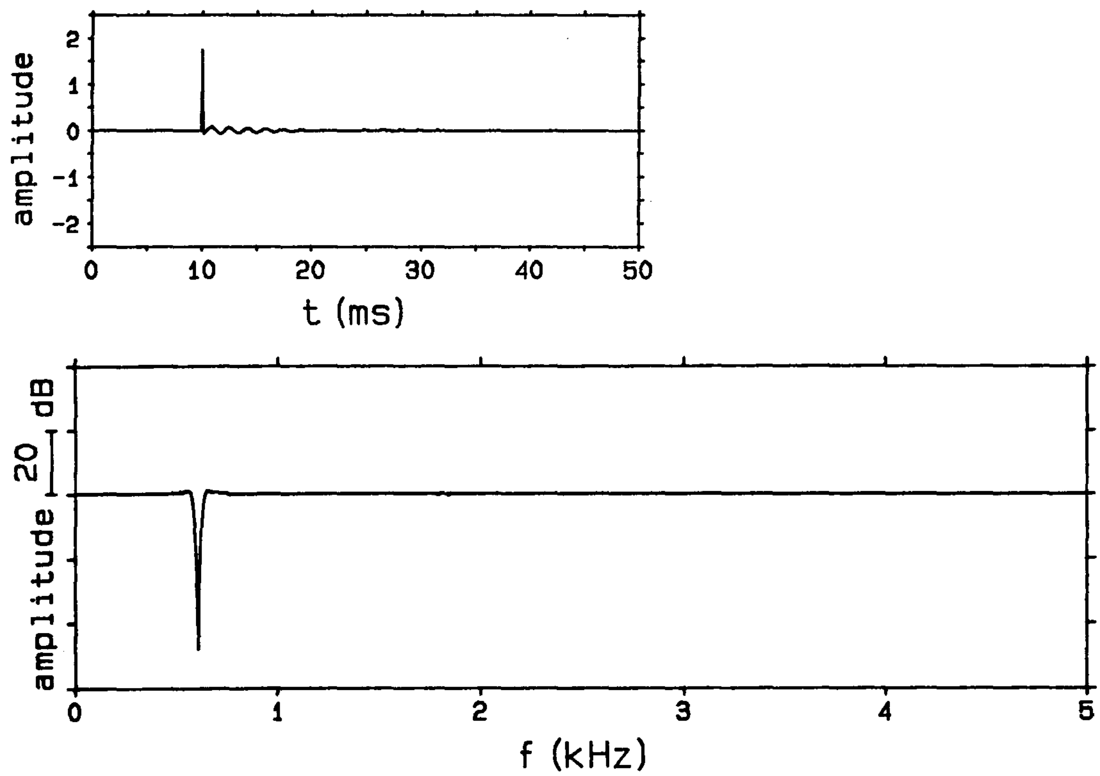


figure 4.4.: Digital filter used to remove the minor third partial from the original bell sound. a.Impulse response, b.Frequency response.

II. The amplitude envelope of the filtered partial is approximated by an amplitude function.

Two methods were used to deterimine the amplitude envelope of the filtered partial :

1. The total bell signal (1536 ms) was devided into 51.2-ms segments. The amplitude of the partial within each segment was determined by means of a FFT.

2. A digital passband filter was designed using the ILS-subrotine EFI. This filter was used to extract the selected partial from the bell signal. All partials of the Gouda "C2"-bell shown in table 4.1 could be approximated by an amplitude function which consisted of two exponentials :

$$A(t)=c_1 + c_2.\exp(c_4.t) + c_3.\exp(c_5.t) \qquad (4.2)$$

where $A(t)$ is the amplitude function, t is time and $c_1,c_2,..,c_5$ are constants.

EXAMPLE 2 :

Figure 4.5 shows the original amplitude envelope of the minor third (method 2) and the approximated envelope function.

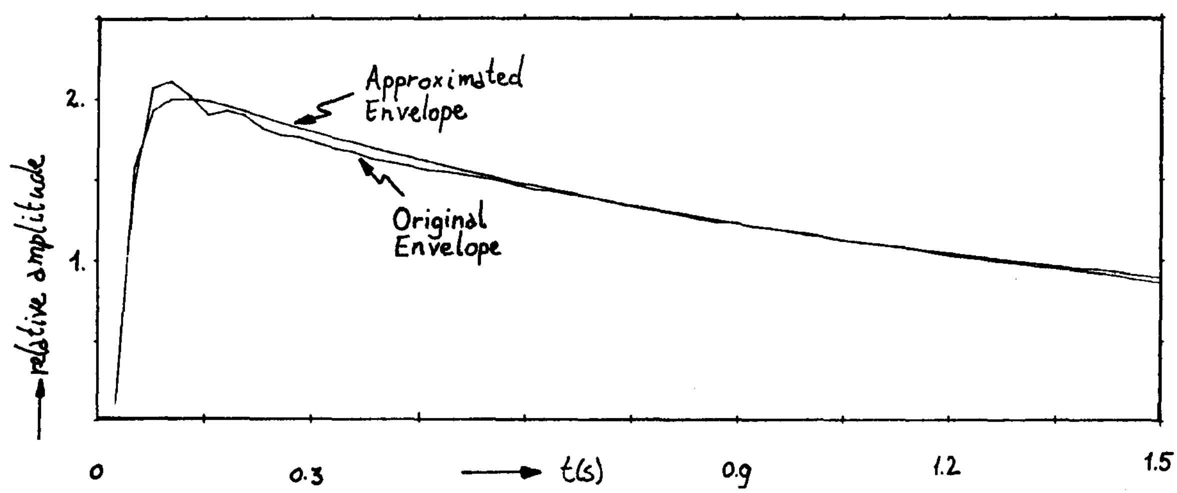


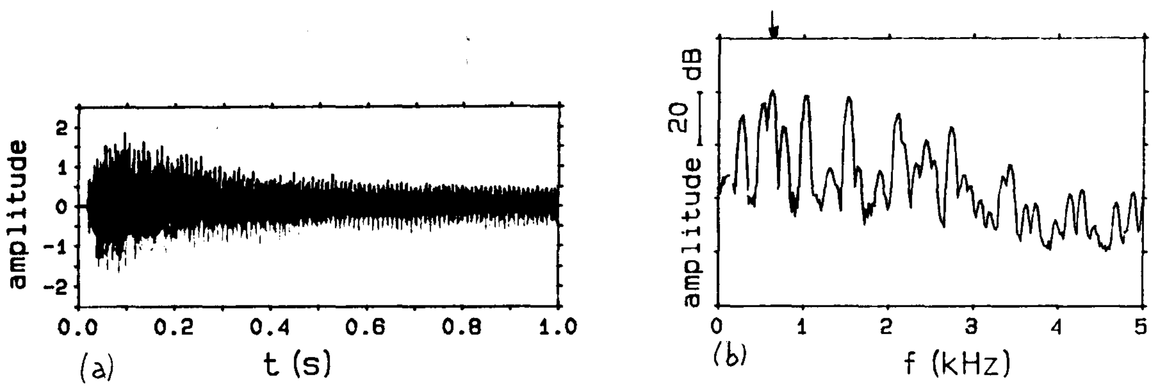
figure 4.5.: Approximation of the amplitude envelope of the minor third partial.

III. Synthesis of a bell stimulus.

A new partial is created by multiplying a sinusoid with the approximated amplitude function. This new partial is added to the filtered bell sound. By varying the frequency of the sinusoid, the pitch of the strike note can be measured systematically as a function of the frequency shift of the selected partial.

EXAMPLE 3 :

Application of the digital filtering technique in case of the minor third partial is shown in figure 4.6. The relative frequency shift of the minor third is 0.1 (frequency minor third : 603 Hz ---> 663,3 Hz).



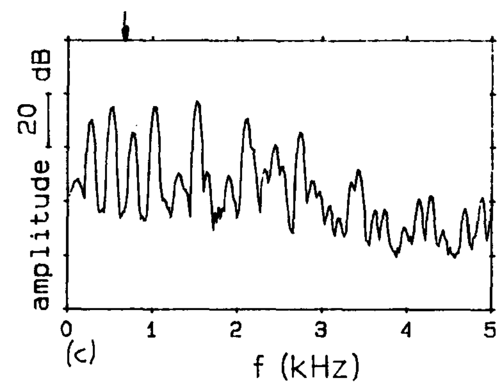
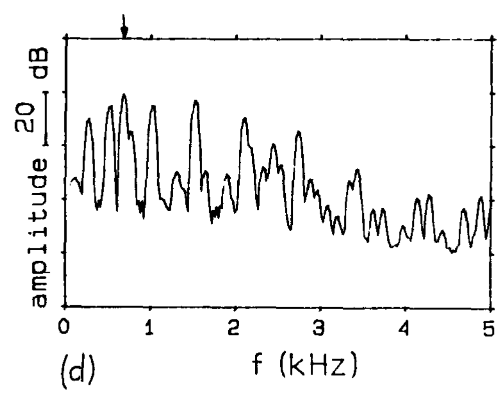


figure 4.6.: Digital filtering technique : a. Time plot of the bell signal, b. spectrum of the original bell, c. Minor third partial (603 Hz) is removed from the sound spectrum, d. Shifted partial is added to the spectrum of c.



Each of the nine partials shown in table 4.1 was replaced by eleven new partials. The relative shift ($\Delta F/F$) of these partials took one of the following values : $\Delta F/F$=-0.10,-0.08,-0.06,....,0.00,0.02,....,0.10 . In this way a set of 99 bell stimuli was generated. The stimuli in this set were called the "long" bells, because of their duration of 1536 ms. A second set of 99 stimuli was formed by taking the first 100 ms of every stimulus of the first set. This new set contained the so-called "short" bells.

In the second part of the experiment, two or three partials were shifted in frequency at the same time. The partials of the following groups were shifted simultaneously : (5,6),(5,7),(6,7),(5,6,7). Again two sets of stimuli were produced, a "long" bell and a "short" bell set, each containing 44 bell stimuli.

5 EXPERIMENT

An experiment was designed to measure the shift of the strike note pitch of a bell stimulus (test tone) as a function of the frequency shift of a partial or a group of partials of the bell. Four subjects had to match the pitch of a sinusoidal comparison tone to the pitch of a test tone. The apparatus is discussed in section 5.1. Section 5.2 will describe the experimental procedure and the results will be shown in section 5.3.

5.1 Apparatus

An experimental set-up was designed around a Philips P800 minicomputer system (see figure 5.1.). After preparation on a VAX 11/780 computer, the bell stimuli were down-loaded to the P800 and stored on a 40 MB hard disc. The acoustic stimuli were generated using a 12-bit digital-to-analog converter (DAC). The output of the DAC was low-pass filtered at 4.3 kHz using two sections of a Krohn-Hite 3343 filter (overall slope 96 dB/oct). The comparison tone was a sinusoid generated by a Philips PM5190 programmable oscillator. Both test and comparison tone were fed into an amplifier and transmitted to the experimentation room. The subject was sitting in a sound-proofed chamber and could hear the sound stimuli through Senheisser HD424 earphones. Measurements with a small probe microphone at the ear canal entrance with the earphones in place showed that the frequency response of the earphones was flat within $\pm$ 3.5 dB from 60-5000 Hz.

The time pattern of presentation of the reference and the comparison tone depended on the duration of the test tone (the bell stimulus). The duration of a bell stimulus could be 1536 ms (a so-called "long" bell) or 100 ms (a "short" bell). The time patterns of presentation are shown in figure 5.2. for both cases. Each tone had 5 ms onsets and offsets in order to avoid clicks produced by switching the signals on and off.

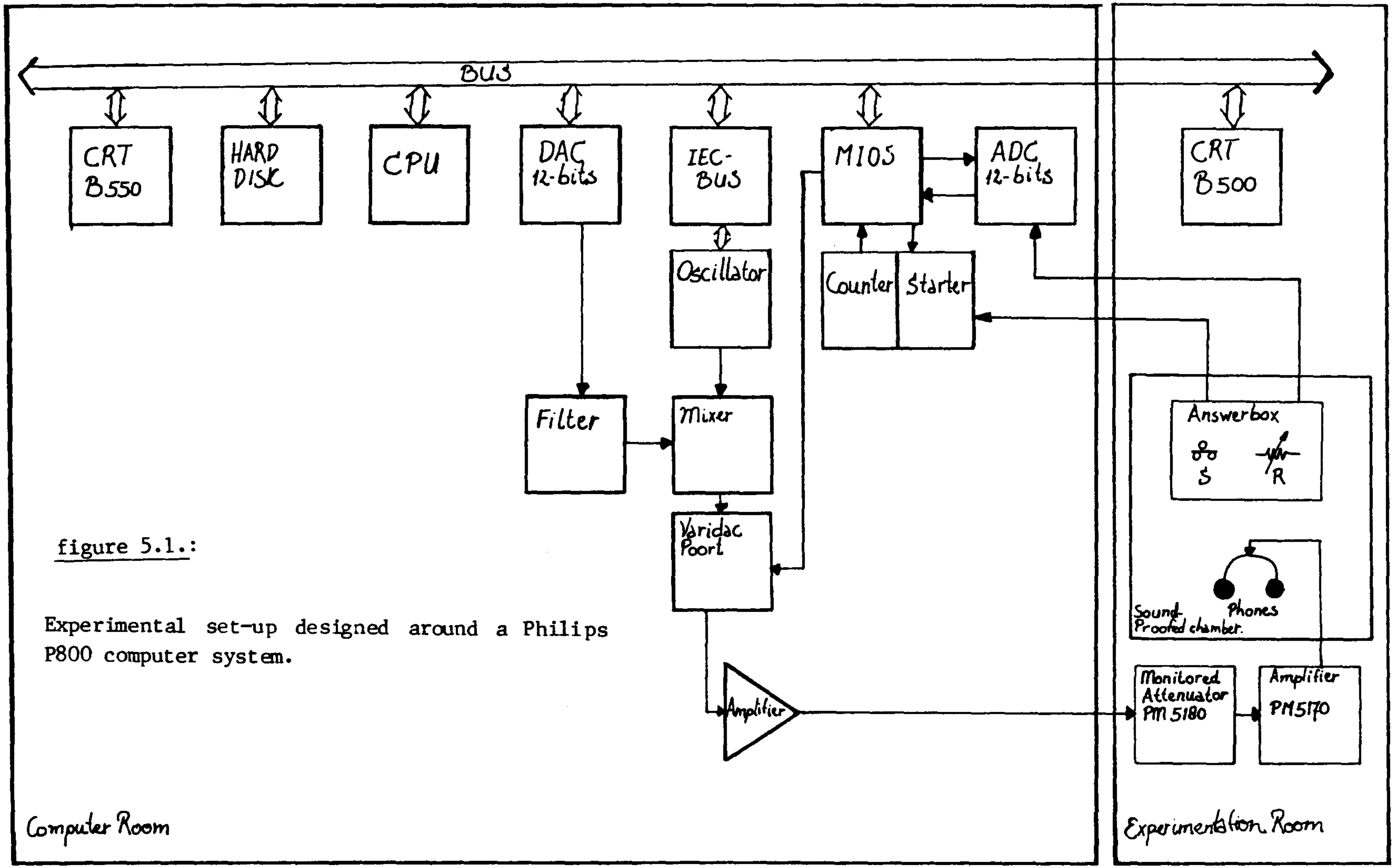


figure 5.1.:

Experimental set-up designed around a Philips P800 computer system.

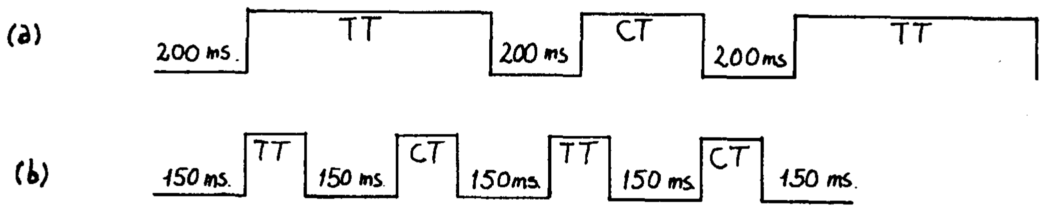


figure 5.2.: The time pattern of presentation of the test tone (TT) and the comparison tone (CT) :
a. the "long" bell run ; TT=1536 ms., CT=500 ms.
b. the "short" bell run ; TT=100 ms., CT=100 ms.

An answer box provided the communication from the subject to the computer. This answer box contained a ten-turn potentiometer and a push button. The subject could control the frequency of the comparison tone by adjusting the potentiometer. The position of the potentiometer corresponded to a certain DC voltage. This voltage was converted into a binary number using a 12-bit analog-to-digital converter (ADC). The subject could start or stop the experiment by pushing the button. A change of the status of the pushbutton was memorized by a special counter-starter configuration. The ADC, the counter, the starter and a varidac port were interfaced to the P800 data bus via MIOS (Multiple Input Output System). The total management of the experiment was controlled by a FORTRAN IV computer program. The communication between the experimenter and the computer was provided by two Beehive (B550/B500) CRT-terminals. One CRT-display was located in the computer room, the other in the experimentation room. The experimenter could attenuate or amplify the sound stimuli by means of a monitored attenuater and a amplifier in the experimentation room. The stimuli were checked with a Data Precision spectrum analyser (DATA 6000).

5.2 Procedure

Four subjects participated in the experiments. Three of them, PB, HK, and BE, had some musical experience. The fourth subject AH was highly experienced in psychoacoustical tasks. The task of the subjects was to adjust the pitch of a comparison tone to match that of the test tone. There was no limit to the time needed by the subjects to produce a pitch match. When the subject was satisfied about the match, he pressed a push button. The matching frequency was stored by the computer program and a new test tone was presented to the subject. Before the next trial began a small random increment or decrement was applied to the offset of the potentiometer. In this way, the subject could not use the position of the potentiometer as a cue. "Long" bells ans "short" bells were presented in different runs. Each run contained 35 trials.

The stimuli were presented diotically (both ears received the same acoustical information) at 20 dB SL. A low sensation level was chosen in order to minimize analytic perception (see section 3.2.2). At the begin of the experiments, the threshold of audibility of the sound stimuli was determined using a DAVEN attenuator. The subject adjusted the attenuator to make the sound stimulus just audible. The stimuli were presented 20 dB above this threshold.

5.3 Results

In a first series of experiments, the pitch shift of the strike note was measured as a function of the frequency shift of one partial of the bell. A second experiment measured the pitch shift of the strike note in case two or three partials were shifted in frequency simultaneously.

The pitch matches of the four subjects were pooled, because the inter-subject consistency proved to be good. The results for the individual subjects are shown in appendix A2. A least-squares technique was used to fit straight lines to the data (Subroutine E04FDF of the NAG-library). The results of the first experiment are shown in figure 5.3. The relative pitch shift of the strike note ($\Delta P/P$) is shown as a function of the relative frequency shift of a partial ($\Delta F/F$). The results for the "long" bells are shown on the left side of the figure. The results for the "short" bells can be seen on the right part of the figure. In the "long" bell figures, each data point is based on 8 pitch measurements. Each "short" bell data point is based on 10 pitch matches. It should be noted that in case of the sixth partial (the twelfth), there are only pitch matches of the strike note in the central region of the figure (e.g. a small frequency shift of the twelfth). When the absolute frequency shift of the twelfth $\Delta F/F$ is greater than 0.04, most comparison tones are matched to a subharmonic of the twelfth. The analytic matches of the twelfth are presented in figure 5.4. The measured pitch shift is expressed relative to the frequency of the subharmonic. (Note: pitch matches of the strike note are absent in the central part of figure 5.4.) The results of the first experiments are summerized in table 5.1.

Two or three partials were shifted simultaneously in frequency in the second experiment. The results for the "long" bell stimuli and the "short" bell stimuli are presented in figure 5.5. The pitch matches of the different subjects were pooled again (the results for the individual subjects are shown in appendix A2). Each data point is based on 16 pitch matches. In case the sixth (the twelfth) and the seventh (the upper octave) partial are shifted together, analytic matches occurred again (see figure 5.6). The results of experiment two are summerized in table 5.2.

Partial	Strike note pitch	Number of points	Regression Coefficient	Regression Constant
L1	497.4	88	-.01 ± .02	-.000 ± .001
L2	496.4	88	.05 ± .03	-.001 ± .002
L3	498.5	83	-.02 ± .02	-.001 ± .001
L4	497.0	83	-.01 ± .02	.001 ± .001
L5	497.4	83	.56 ± .06	-.005 ± .004
L6	496.5	29	.5 ± .1	-.000 ± .004
LA6	376.5	29	1.06 ± .03	-.012 ± .003
L7	497.1	88	.16 ± .03	.001 ± .002
L8	497.3	87	-.02 ± .02	.003 ± .001
L9	497.3	87	.00 ± .02	.001 ± .001
S1	496.8	110	-.02 ± .01	-.003 ± .001
S2	495.8	109	-.02 ± .02	-.001 ± .001
S3	495.0	108	-.01 ± .02	.001 ± .001
S4	494.1	108	.00 ± .02	.003 ± .001
S5	495.5	93	.75 ± .06	.001 ± .003
S6	495.4	31	.6 ± .1	.001 ± .003
SA6	376.5	44	1.02 ± .02	-.018 ± .002
S7	495.9	105	.08 ± .03	-.003 ± .002
S8	495.1	107	-.01 ± .02	.002 ± .001
S9	498.0	108	-.03 ± .02	-.007 ± .001

Table 5.1 : Results experiment 1.

Partial	Strike note pitch	Number of points	Regression Coefficient	Regression Constant
L56	496.5	159	1.02 ± .02	-.004 ± .001
L57	496.5	170	.98 ± .03	.009 ± .002
L67	495.8	70	.58 ± .06	.002 ± .001
LA67	376.5	55	1.12 ± .02	-.012 ± .002
L567	497.3	167	1.09 ± .02	-.002 ± .001
S56	493.2	158	1.00 ± .03	.001 ± .002
S57	490.4	159	.88 ± .04	.017 ± .002
S67	492.2	66	.5 ± .1	.007 ± .004
SA67	376.5	48	1.10 ± .03	-.018 ± .003
S567	491.6	171	1.00 ± .03	.012 ± .002

Table 5.2 : Results experiment 2.

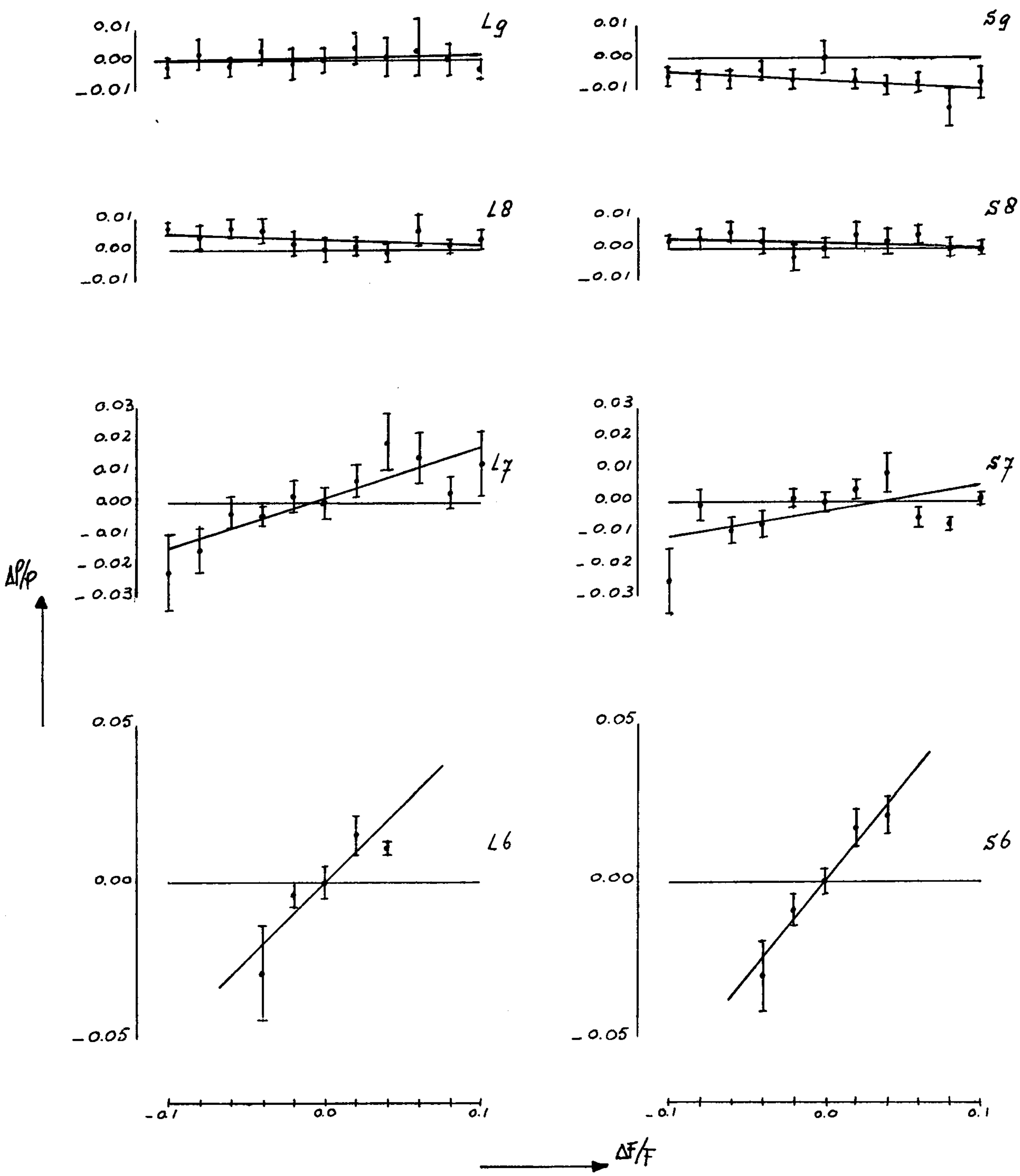


figure 5.3.: Relative shift of the strike note pitch (ΔP/P) as a function of the relative frequency shift of a selected partial (ΔF/F). The shift of the strike note pitch of the "long" bells (L) is shown on the left of the figure; the results for the "short" bells (S) are plotted on the right. The numeric symbol attached to the L or S indicates which partial is shifted in frequency. Averages plus/minus one standard deviation of the mean of data from 4 subjects. Figure 5.3 is to be continued on the next page.

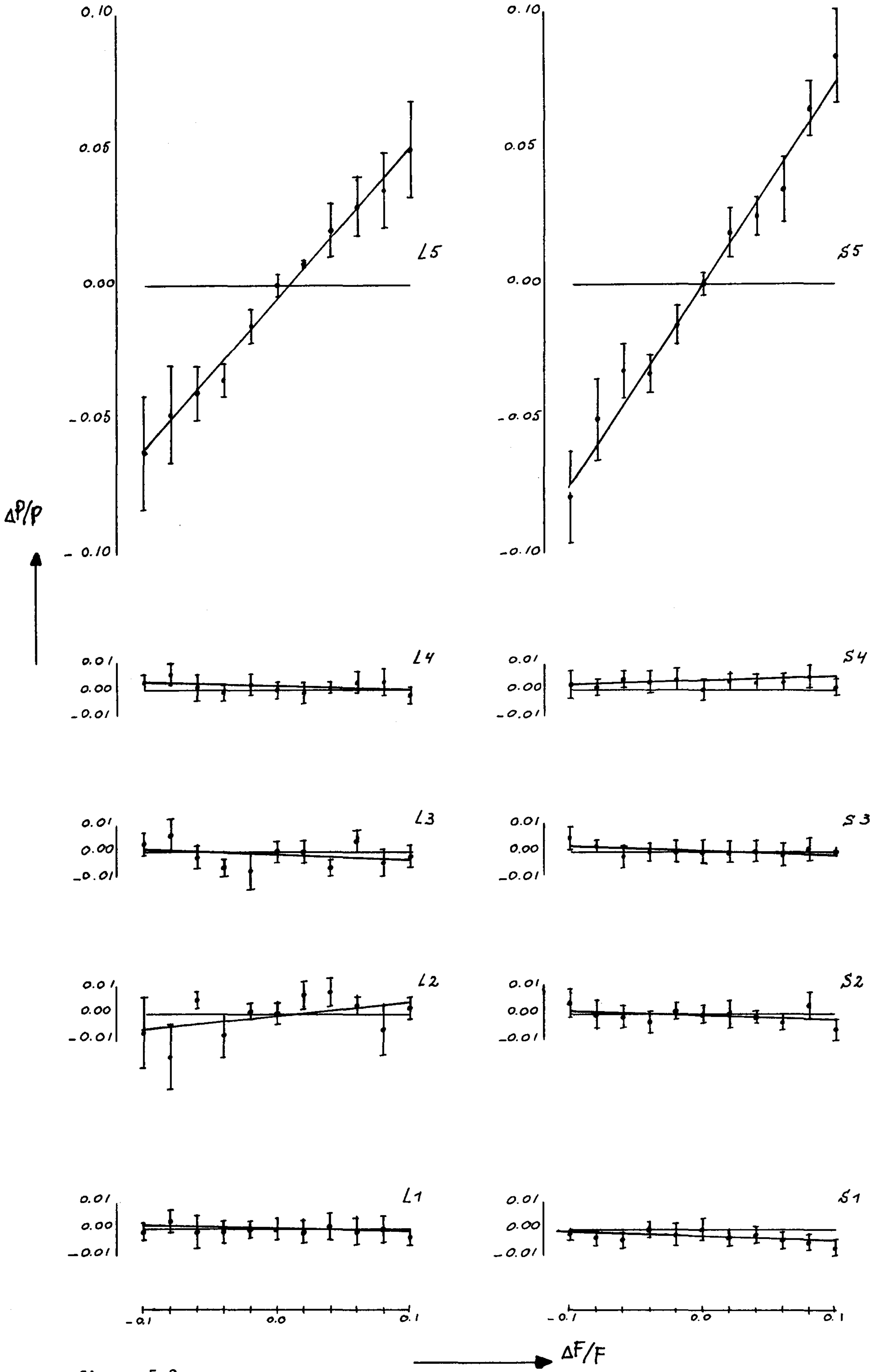

0.10
0.05
0.00
-0.05
-0.10
L5
S5
ΔP/P
L4
S4
L3
S3
L2
S2
L1
S1
0.01
0.00
-0.01
-0.1
0.0
0.1
ΔF/F

figure 5.3.

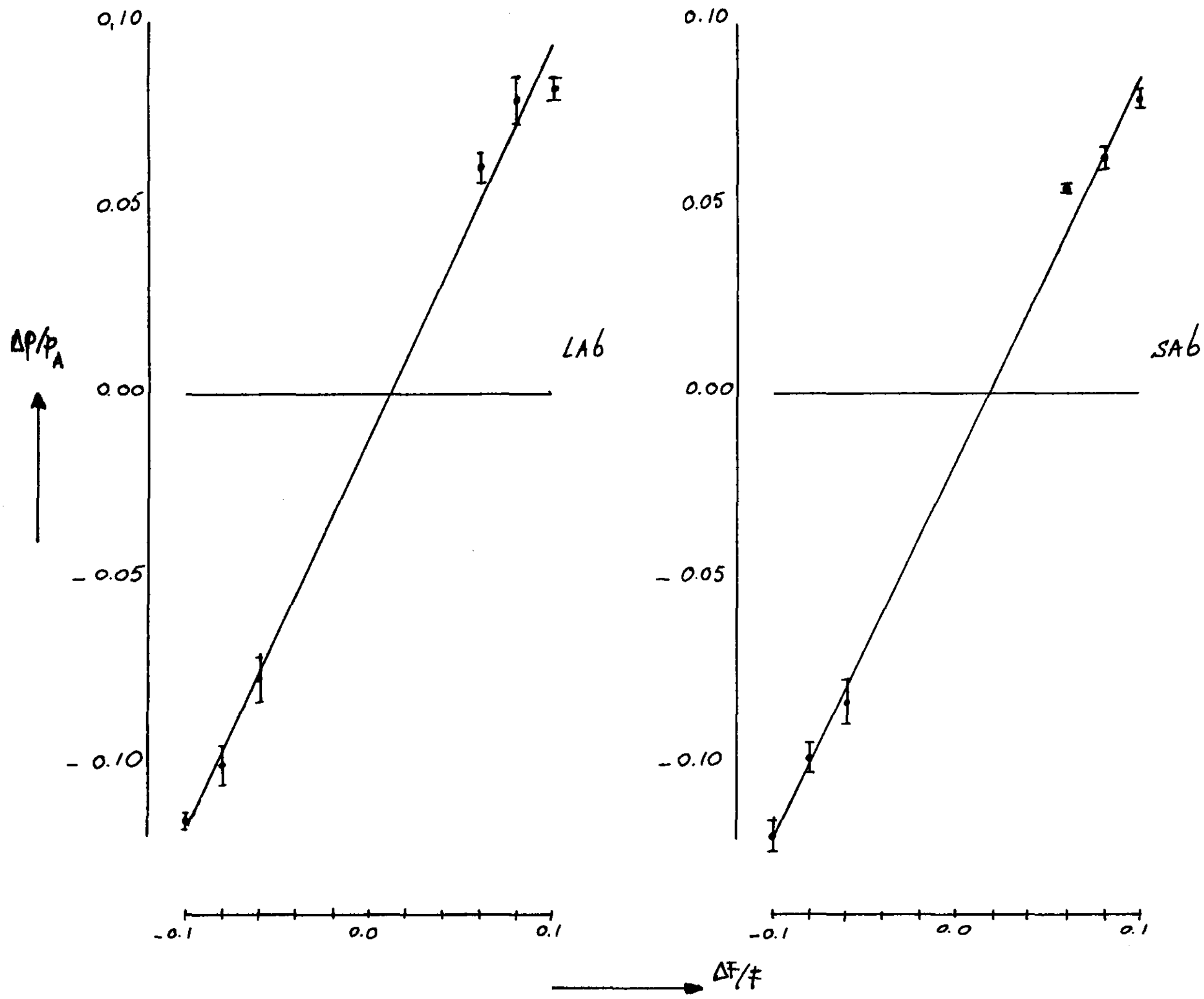


figure 5.4.: When the frequency of the 6th partial (twelfth) is shifted too much, the subjects match the comparison tone to a subharmonic of the twelfth and not to the strike note of the bell (analytic perception). Averages plus/minus one standard deviation of the mean of data from 3 subjects.
$P_A = 376.5$ Hz.

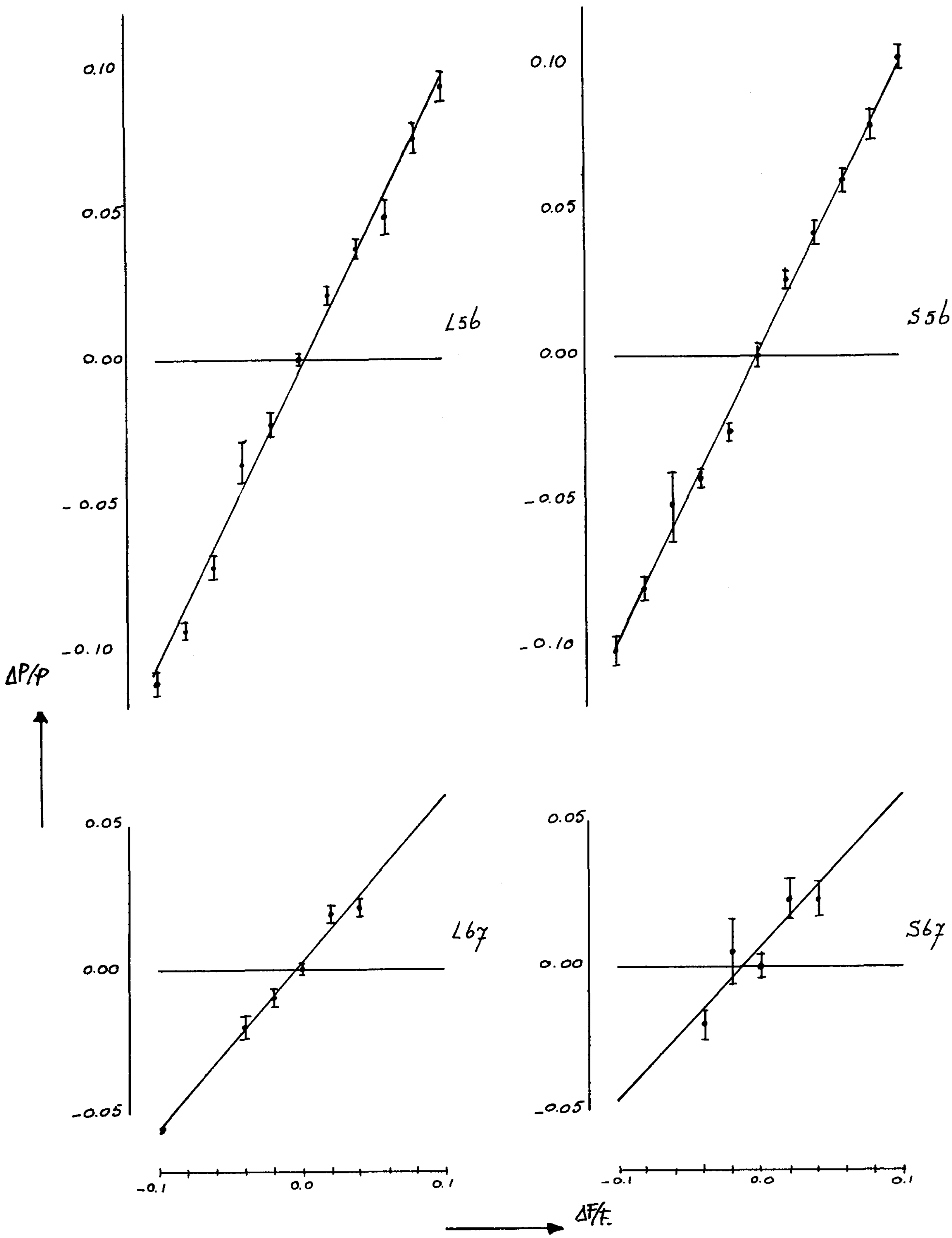


figure 5.5.: Relative shift of the strike note pitch (ΔP/P) as a function of the relative frequency shift of a group of partials (ΔF/F). The shift of the strike note pitch of the "long" bells (L) is shown on the left of the figure; the results for the "short" bells (S) are plotted on the right. The numeric symbol attached to the L or S indicates which partials are shifted in frequency. Averages plus/minus one standard deviation of the mean of data from 4 subjects. Figure 5.5 is to be continued on the next page.

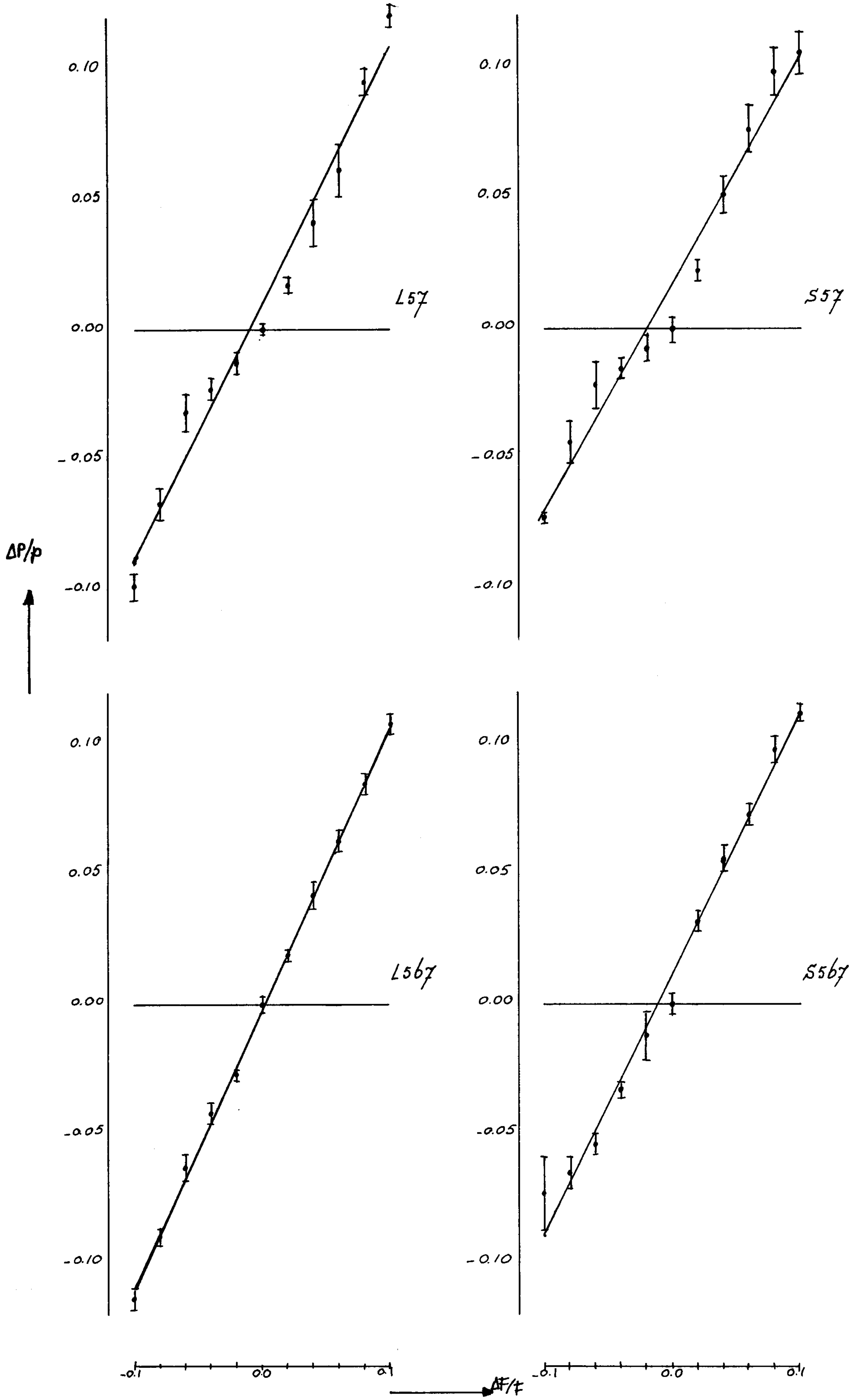
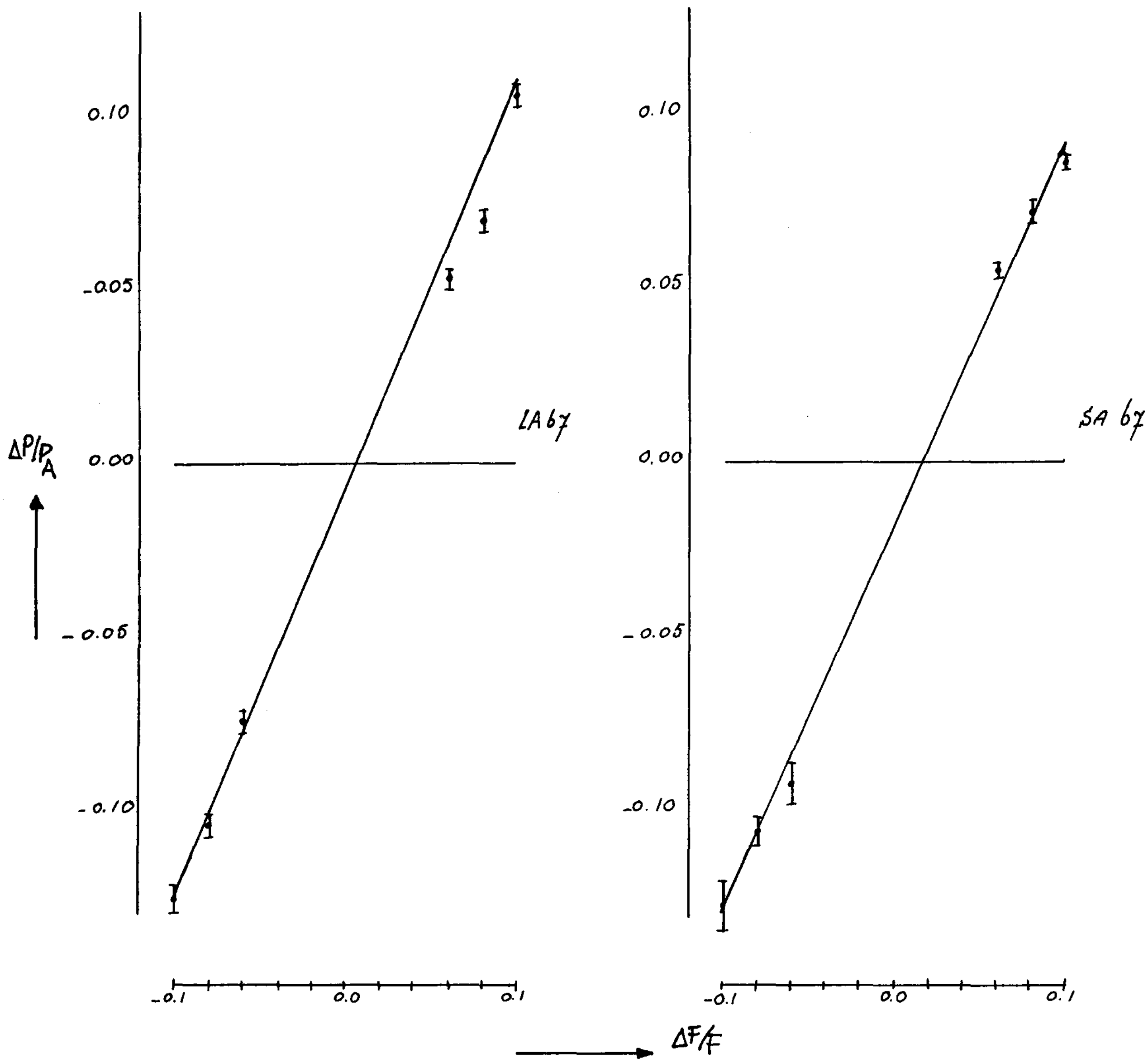


figure 5.5.



<u>figure 5.6.</u>: When the frequency of the 6th and 7th partials are shifted too much, the subjects match the comparison tone to a subharmonic of the 6th partial (twelfth) and not to the strike note of the bell (analytic perception). Averages plus/minus one standard deviation of the mean of data from 3 subjects. $P_A = 376.5$ Hz.

6 DISCUSSION AND CONCLUSIONS

The results presented in chapter 5 show a good agreement between the pitch matches of the "long" bell and "short" bell stimuli. The acoustical information concerning the strike note pitch seems to be contained within the first 100 ms of the bell sound.

The octave (L5,S5), the twelfth (L6,S6) and the upper octave (L7,S7) have a significant effect on the strike note pitch. The other partials exert no significant effect. It should be noted that the octave, the twelfth and the upper octave are the only partials with a frequency which is an integral multiple of the strike note pitch. These results are consistent with the experimental findings of Greenhough (1976). However, Greenhough found the upper octave to be the most important partial, whereas our results show a much weaker influence of the upper octave on the stike note pitch. The frequencies of the partials of the synthetic bell stimuli used by Greenhough were almost identical to the partial frequencies of the Gouda "C2"-bell used in our experiments. It should be noted that the upper octave was the highest partial (in frequency) in the synthetic bell stimuli used by Greenhough. An edge effect could have played an important role in the case of the upper octave. In contrast with real bell sounds, Greenhough´s stimuli consisted of eight partials with equal amplitude. In the case of the Gouda "C2"-bell the amplitude of the upper octave was lower relative to the amplitudes of the octave and twelfth partials (see figure 4.1). The level of a partial relative to the other partials seems to have an effect on its relative dominance within the complex sound.

This result is confirmed by Moore et al. (1985). They investigated the relative dominance of individual partials in determining the pitch of complex tones. They found that the level of a harmonic relative to adjacent harmonics could have a significant effect on its dominance. The optimum processor theory of Goldstein (1973) doesn´t account for this finding. This model therefore needs modification. The virtual pitch theory of Terhardt (1979) is more satisfactory in this respect.

The data show that the relative shift of the pitch of the strike note is approximately a lineair function of the relative shift in frequency of the selected partials. According to Moore et al. (1985) the slope of this function may be considered as an estimate of the weight of the harmonic (partial) in contributing to the residue pitch (strike note pitch). When one assumes that the value of the strike note pitch is derived as a weighted average of pitch cues derived from the individual partials, the sum of the slopes should be close to one. The sum of the slopes of the octave, the twelfth and the upper octave equals 1.2 (± 0.2) for the "long" bell and 1.4 (± 0.2) for the "short" bell stimuli. In the case of the "long" bells the sum of the slopes is almost one within the limits of experimental error. However, the sum of the slopes is somewhat large for the "short" bells.

When the slopes of the measured functions of the first experiment are considered as estimates of the weights in which the partials contribute to the strike note pitch, the partials should have the following weights :

No.	Name of partial	Weight of partial in determining the strike note pitch	
		"long" bells	"short" bells
1	Hum	0	0
2	Prime	0	0
3	Minor Third	0	0
4	Fifth	0	0
5	Octave	.56 ± .06	.75 ± .06
6	Twelfth	.5 ± .1	.6 ± .1
7	Upper Octave	.16 ± .03	.08 ± .03
8	-	0	0
9	Double Undeciem	0	0

Table 6.1.: Estimates of the weights of the partials of the Gouda "C2"-bell, in contributing to the strike note pitch.

In the second experiment, two or three of the most important partials (octave, twelfth and upper octave) were shifted in frequency simultaneously. The measured slopes were compared with the slopes as predicted by the weights of table 6.1. The results are shown in table 6.2.

Shifted Partials	Predicted slope	Measured slope
L56	1.1 ± .2	1.02 ± .02
L57	.72 ± .09	.98 ± .03
L67	.7 ± .1	.58 ± .06
L567	1.2 ± .2	1.09 ± .02
S56	1.4 ± .2	1.00 ± .03
S57	.83 ± .09	.88 ± .04
S67	.7 ± .1	.5 ± .1
S567	1.4 ± .2	1.00 ± .03

Table 6.2.: Comparison between the measured slopes of the bell stimuli of the second experiment and the predicted slopes calculated from table 6.1.

Table 6.2. shows a reasonable good agreement between the predicted and measured slopes. One can conclude that the relative dominance of the twelfth is almost equal to the relative dominance of the octave in determining the strike note pitch.

According to Ritsma (1967) the frequency band consisting of the third, fourth and fifth harmonics of a fundamental frequency range 100-400 Hz, tends to dominate the pitch sensation. In the case of the Gouda "C2"-bell this would imply a dominant role for the twelfth (3 times the strike note pitch) and the upper octave (4 times the strike note pitch) partials. Although the twelfth is indeed very important in determining the strike note pitch, the upper octave is less important. The octave (2 times the strike note pitch) plays a dominant role, contrary to the classical concept of the dominant region (Ritsma, 1967). The dominant role of partials lower than the third to fifth harmonics is consistent with the findings of Moore et al. (1985).

The results shown in chapter 5 clearly demonstrate the subjective character of the strike note pitch. The data disprove the older theories of the strike note pitch. According to the data, the "misjudged octave" hypothesis could equally well have been called the "misjudged twelfth" hypothesis; a shift in frequency of the twelfth only causes a shift of the strike note pitch. When the octave is shifted upwards in frequency, the difference frequency between the octave and the twelfth becomes smaller. According to the "difference tone" hypothesis the pitch of the strike note should exhibit a downward shift. However, our data show an upward shift of the strike note pitch. The "residue" theory of Schouten (1940) requires that the auditory system fails to resolve the octave, the twelfth and the upper octave partials from the acoustical signal. Modern psychophysical as well as physiological experiments (Plomp,1964; Kiang,1965) have shown conclusively that such frequency components are resolved in the cochlea. In this view, the octave, the twelfth and the upper octave do not show the temporal characteristics of the strike note.

The strike note phenomenon was confronted with modern pitch theories by measuring the pitch of the bell stimuli using some pitch meters based on these theories (see chapter 2). The DWS-pitch meter is based on Goldstein´s optimum processor theory. This meter didn´t give any pitch estimates in the neighbourhood of the strike note pitch at all. This can be explained by the fact that this pitch meter is originally constructed for determining the pitch of speech signals. The partials of a speech signal are harmonically related, whereas for a bell sound the partials form a inharmonic series. It should be noted that the DWS-pitch meter extracts only the lowest six components of the sound signal. In the case of our stimuli, the twelfth is rejected by the DWS algorithm. The TVP-meter gives a great number of virtual pitches. The strongest virtuals pitches are not found in the immediate vicinity of the strike note pitch. They are found one or two octaves lower. However, there are virtual pitches in the range 400-600 Hz. According to Gerson and Goldstein (1978) the "natural" operation of the pitch processor includes a restriction in the range of expected pitches. In our experimental paradigm the pitch of a bell stimulus seemed to be expected in a close range around the previously determined strike note pitch. This concept of restricted pitch range can be applied to the TVP-meter. When only the virtual pitches in the range 400-600 Hz are considered, the TVP-meter gives a good prediction of the shift of the strike note pitch when the octave,twelfth, and the upper octave (L567, S567) or merely the octave and the twelfth (L56, S56) are shifted simultaneously in frequency (see figure 6.1). However, no shift of the strike note pitch is predicted when (5,7) and (6,7) are shifted.

In our research the bell stimuli were presented at a low sensation level (20 dB SL) in order to avoid the analytic mode of perception (see paragraph 2.2.2) and the presence of aural combination tones. According to Houtsma (1979) it should be apparently difficult for a subject to isolate individual tone components when the complex sounds are large and spectrally dense, and are presented at low sensation levels. Indeed, none of the subjects could hear consciously the individual partials "out" of the total bell sound. Despite the low sensation level of the stimuli, analytic pitch matches occurred in our experiments, especially in the case of the twelfth. It could be possible that the character of the comparison tone, being a sinusoid, favoured the analytic perception mode. It should be noted that most of the analytic matches of the twelfth occured when the relative shift was greater than 4% . In these cases the pitch of the bell stimulus was matched to a subharmonic of the twelfth. According to Gerson and Goldstein (1978) such "subharmonic" matches may occur, only if the "normal" match is outside the expected range of pitches. The virtual pitch theory of Terhardt (1974) describes the two modes of pitch perception. The synthetic mode of pitch perception is described by the virtual pitch pattern, wheras the spectral pitch pattern represents the analytic listen mode. In most of the cases, the twelfth has the largest weight of the spectral pitch pattern. The octave and the upper octave show no analytic matches because they have an octave relation with the strike note pitch. It seems that the competition between both perceptual modes depends on the frequency of the specific partial as well as on the inharmonicity of the total percept of the bell stimulus. The data of the individual subjects (see Appendix A2) show that not all subjects are equally sensitive to analytic matches. Subject HK made no analytic matches, whereas BE produced analytic matches even in the case two or three partials were shifted in frequency simultaneously.

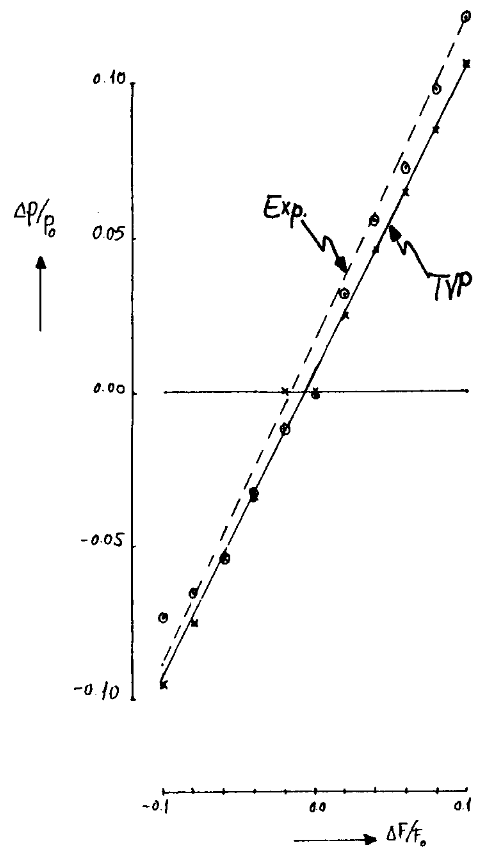


figure 6.1.: Theoretical × (TVP-meter) and experimental ⊙ determined shift of the strike note pitch as a function of a shift in frequency of the octave, the twelfth and upper octave partials (S567).

It would be very interesting to repeat the experiments carried out in this project for bells which have a much lower or higher strike note pitch, in order to see if the results of this project have general validity. It would also be interesting to search if there is any influence of the character of the comparison tone on the pitch matches made.

At the end of this report the following conclusions are drawn :

1. The first 100 ms of a bell sound contain all relevant acoustical information needed by the human auditory system in order to perceive a strike note pitch.

2. The octave and the twelfth are equally important in determining the strike note pitch.

3. The upper octave is less important with respect to the strike note pitch. This fact is caused by its amplitude level relative to the levels of the octave and twelfth.

4. The minor third, which is the strongest spectral component, does not contribute to the strike note pitch.

5. The DWS- and the TVP-pitch meters based on modern pitch theories give poor predictions of the measured shift of the strike note pitch as a function of systematic frequency shift of selected partials. When the concept of an expected pitch range is applied to the TVP-meter, only partially correct predictions are shown.

6. The strike note is clearly a subjective tone.

8. The data contradict the "misjudged octave" hypothesis, as well as the "difference tone" hypothesis.

9. Analytic perception can play an important role in psycho-acoustical experiments, even when the stimuli are spectrally dense and presented at low sensation levels. Analytic perception occurs when partials assume an inharmonic frequency relationship to the fundamental.

7 ACKNOWLEDGMENTS

At the end of this report I wish to express my thanks to all who have cooperated in making this study possible. Especially to Dr. A.J.M. Houtsma for his support and helpfull comments during this project. To Professor Dr. H. Bouma of the Institute for Perception Research for permission to make use of the laboratory. To Professor Dr. J.A. Poulis of the Department of Physics at Eindhoven University of Technology for his encouraging support. I wish to express my appreciation to Paul Braat, Hans Kwaspen and Dr. A.J.M. Houtsma who participated as subjects in the experiments. I want to thank the research associates and the research staff of the Hearing and Speech group at the I.P.O., especially Theo de Jong who assisted me with the technical setup of the experiments and John Beerends for helpfull discussions.

8 REFERENCES

Allik,J.,Mihkla,M.,Ross,J.(1984)."Comment on "measurements of the pitch in speech: an implementation of Goldstein's theory of pitch perception", "
J.Acoust.Soc.Am.75,1855-1857.

Arts,J.(1938)."The sound of bells. Jottings from my experiences in the domain of the sound of bells,"
J.Acoust.Soc.Am.9,344-347.

Arts,J.(1939)."The sound of bells. Jottings from my experiences with the sound of bells. The secondary strike note,"
J.Acoust.Soc.Am.10,327-329.

Beerends,J.G.,Houtsma,A.J.M.(1986)."Pitch identification of simultaneous dichotic two-tone complexes,"
To be published in the J.Acoust.Soc.Am.

Bekesy,G.von (1960).Experiments in hearing,(Mc GrawHill, New York).

Brigham,E.O.(1974).The Fast Fourier Transform,(Prentice-Hall, New Jersey).

Duifhuis,H.,Willems,L.F.,Sluyter,R.J.(1982)."Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception,"
J.Acoust.Soc.Am.71,1568-1580.

van den Enden,A.W.M.,Verhoeckx,N.A.M.(1985)."Digitale signaal bewerking: theoretische achtergronden,"
Philips Techn.T.42,120-156.

Gerson,A.,Goldstein,J.L.(1978)."Evidence for a general template in central optimal processing for pitch of complex tones,"
J.Acoust.Soc.Am.63,498-510.

Goldstein,J.L.(1973)."An optimum processor theory for the central formation of the pitch of complex tones,"
J.Acoust.Soc.Am.54,1496-1516.

Goldstein,J.L.,Gerson,A.,Srulovicz,P.,Furst,M.(1978). "Verification of the Optimal Probabilistic Basis of Aural Processing in Pitch of Complex Tones,"
J.Acoust.Soc.Am.63,486-497.

Greenhough,M.(1976)."The analysis and synthesis of bell sounds,"
Proceedings of the Institute of Acoustics, sept. 1976.

Grützmacher,M.,Kallenbach,W.,Nellessen,E.(1965)."Acoustical Investigations on Church Bells,"
Acustica 16,34-45.

von Helmholtz,H.L.F.(1863).On the Physiological Basis for the Theory of Music,(English Translation, Dover, New York).

van Heuven,E.W.(1949).Acoustical Measurements on Church-bells and Carillons,(Gebroeders van Cleef, ´S-Gravenhage).

Houtsma,A.J.M.,Goldstein,J.L.(1972)."The central origin of the pitch of complex tones: evidence from musical interval recognition,"
J.Acoust.Soc.Am.51,520-529.

Houtsma,A.J.M.(1979)."Musical pitch of two-tone complexes and predictions by modern theories,"
J.Acoust.Soc.Am.66,87-99.

Jones,A.T.J.(1930)."The strike note of bells,"
J.Acoust.Soc.Am.1,373-381. Repr. in Rossing 1984.

Kiang,N.Y.S.(1965).Discharge Patterns of Single Fibers in the Cat´s Auditory Nerve,(M.I.T.Press, Cambridge, Mass.).

Lehr,A.J.(1965)."Hedendaagse nederlandse klokkengietkunst,"
Neth.Acoust.Soc.7,20-49.

Lehr,A.J.(1976).Leerboek der campanologie,(Nationaal Beiaardmuseum, Asten, the Netherlands).

Meyer,E.,Klaes,J.(1933)."On the strike note of bells," Naturwissenschaften 39,697-701. Repr. in Rossing 1984.

Moore,B.C.J.(1982).An Introduction to the Psychology of Hearing, (Academic Press, London).

Moore,B.C.J.,Glasberg,B.R.(1985)."Relative dominance of individual partials in determining the pitch of complex tones," J.Acoust.Soc.Am.77,1853-1860.

Ohm,G.S.(1843)."Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen," Ann.Phys.Chem.59,513-565.

Perrin,R.,Charnley,T.,Depont,J.(1983)."Normal modes of the modern english church bells," J.Sound Vib. 90,29-49.

Pfundner,J.(1962)."On the strike note of bells," Acustica 12,153-157.

Plomp,R.(1964)."The Ear as a Frequency Analyzer," J.Acoust.Soc.Am.36,1628-1636.

Rabiner,L.R.,Gold,B.(1975).Theory and Application of Digital Signal Processing, (Prentice-Hall, New Jersey).

Rabiner,L.R.,Cheng,M.J.,Rosenberg,A.E. and Mc Gonegal,C.A.(1976). "A comperative performance study of several pitch detection algorithms," IEEE Trans. on Acoust.,Speech, Signal Processing 24,399-418.

Rayleigh (J.W.Strutt,Lord)(1890)."On bells," Phil.Mag.29,1-17. Repr. in Rossing 1984.

Ritsma,R.J.(1967)."Frequencies dominant in the perception of the pitch of complex sounds,"
J.Acoust.Soc.Am.42,191-198.

Rossing,T.D.(1982).Science of sound, (Addison-Wesley, Phillipines).

Rossing,T.D.(1984)."The acoustics of bells,"
American Scientist 72,440-447.

Rossing,T.D.(1984).Acoustics of bells, (Van Nostrand Reinhold, New York).

Schouten,J.F.(1940)."De toonhoogtegewaarwording,"
Philips Techn. Tijdschrift 5,298-306.

Schouten,J.F.,Ritsma,R.J.,Cardozo,B.L.(1962)."Pitch of the Residue,"
J.Acoust.Soc.Am.34,1418-1424.

Schouten,J.F., Hart,J.t´ (1965)."De slagtoon van klokken,"
Neth.Acoust.Soc.7,8-19.

Seebeck,A.(1841)."Beobachtungen uber einige Bedingungen der Entstehung von Tonen,"
Ann.Phys.Chem.53,417

Slaymaker,F.H.,Meeker,W.F.(1954)."Measurements of the tonal characteristics of carillon bells,"
J.Acoust.Soc.Am.26,515-522.

Terhardt,E.(1974)."Pitch, consonance and harmony,"
J.Acoust.Soc.Am.55,1061-1096.

Terhardt,E.(1979)."Calculating virtual pitch,"
Hear.Res.1,155-182.

Terhardt,E.,Stoll,G.,Seewann,M.(1982)."Pitch of complex signals according to virtual-pitch theory : tests,examples, and predictions," J.Acoust.Soc.Am.71,671-678.

Terhardt,E.,Stoll,G.,Seewann,M.(1982)."Algorithm for extraction of pitch and pitch salience from complex tonal signals," J.Acoust.Soc.Am.71,679-688.

Terhardt,E.von,Seewann,M.(1984)."Auditive und objektive Bestimmung der Schlagtonhöhe von historischen Kirchenglocken," Acustica 54,129-144.

Thrane,N.(1979)."The Discrete Fourier Transform and FFT analysers," Bruel & Kjaer Tech.Rev.1,1-25.

van Trees,H.L.(1968).Detection estimation and modulation theory,(Wiley, New York).

Tyzzer,F.G.(1930)."Characteristics of Bell Vibrations," J.Franklin Inst. 210,55-66. Repr. in Rossing 1984.

Wever,E.G.(1949).Theory of Hearing, (Wiley, New York).

APPENDIX A1. Formulas Virtual Pitch Theory.

The sound pressure level exess LX_i of the ith tonal component is given by :

$$LX_i = L_i - 10\log_{10}\left[\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} 10^{L_{Ej}(f_i)/20}\right)^2 + I_{Ni} + 10^{L_{TH}(f_i)/10} \right]$$

where L_i is the SPL of the component, $L_{Ej}(f_i)$ is the excitation level which is produced at the frequency f_i by the jth tonal component (note: j is incremented from one to N, skipping i). I_{Ni} is the noise intensisty in the critical band around the considered tonal component. $L_{TH}(f_i)$ is the hearing threshold at the frequency f_i.

$L_{Ej}(f_i)$ is specified by :

$$L_{Ej}(f_i) = L_j - s(z_j - z_i)$$

where L_j is the SPL of the jth tonal component.z represents the critical-band rate; thus z_i and z_j are the critical band rates of the ith and jth components, respectively. z is related to frequency :

$$z = \{ 13.\arctan(0.76\ f) + 3.5\arctan(f/7.5)^2 \} \text{ Bark}$$

where f is the frequency of the component in Hz. Equation A1.2 represents the triangular shape of the excitation level critical-band rate pattern, where s depicts the steepness of the slope.

$$s = 27 \text{ dB/Bark, if } f_i \leqslant f_j;$$

$$s = [-24 - (0.23/f_j) + (0.2L_j)], \text{ if } f_i > f_j$$

where L_j is level in dB and f_j frequency in kHz.

The noise intensity I_{Ni} is obtained by adding the sound intensities of those spectrum samples which correspond to the particular critical-band rate interval extending from (z_i - 0.5 Bark) to (z_i + 0.5 Bark), skipping the five central samples of every tonal component which which eventually has been detected in that critical band.

The threshold of hearing is specified by:

$$L_{TH}(f_i) = \{3.64\ f_i^{-.08} - 6.5\exp[-0.6(f_i-3.3)^2] + 10^{-3}\ f^4\ \}$$

The individual pitch of th ith component is specified by:

$$H_i = f_i(1 + v_i)$$

where:

$$v_i = 2\cdot 10^{-4}(L_i-60)(f_i-2) + 1.5\ 10^{-2}\exp(-LX_i'/20) \times (3-\ln(f_i))$$
$$3\cdot 10^{-2}\exp(-LX_i''/20)\ (0.36 + \ln(f_i))$$

where L_i and f_i. are the SPL(in dB) and the frequency(in kHz) of the considered component. LX_i' and LX_i'' are given by :

$$LX_i' = L_i - 20\log_{10}\sum_{j=1}^{i-1} 10^{L_{Ej}(f_i)/20}$$
$$LX_i'' = L_i - 20\log_{10}\sum_{j=i+1}^{N} 10^{L_{Ej}(f_i)/20}$$

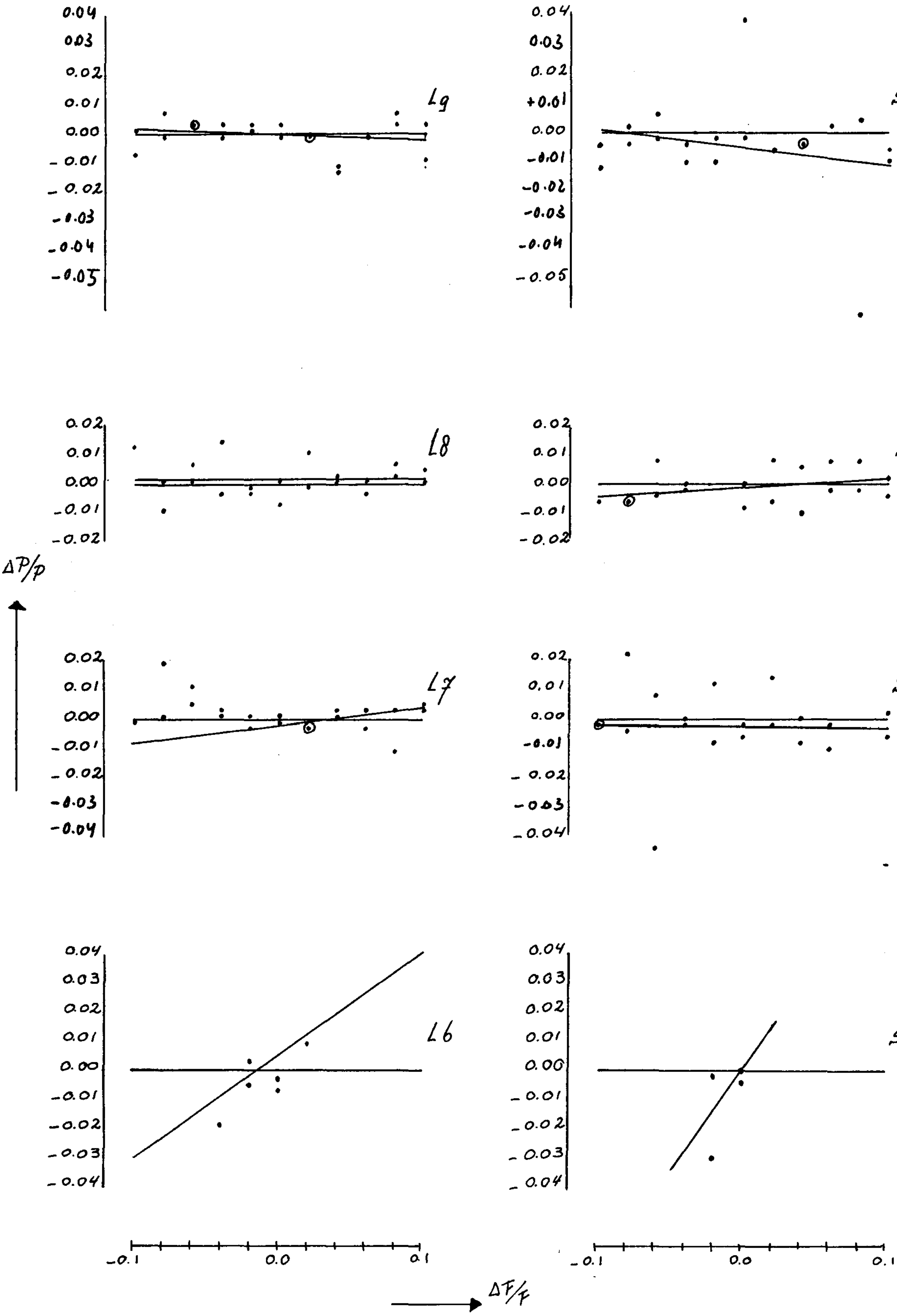
A formula for the true virtual pitch is given by:

$$\underline{H}_{im} = m^{-1}\ f_i(1 + v_i - \text{sign}(m-1)\ 10^{-3}\ \{\ 18 + 2.5m - (50-7m)\ f_im^{-1} +$$
$$0.1[m^{-1}f_i]^{-2}\})$$

APPENDIX A2 Results of the individual subjects

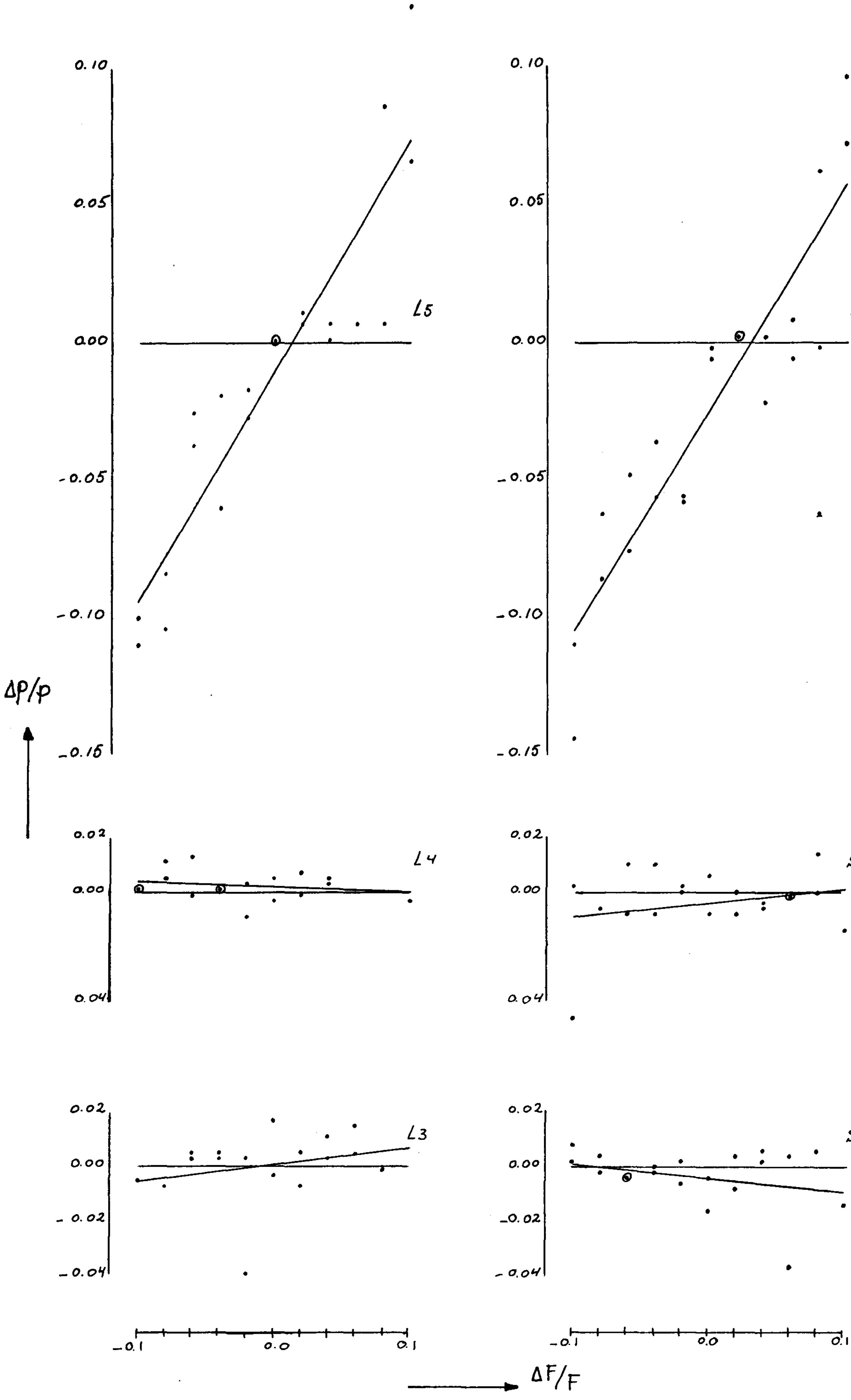
Relative shift of the strike note pitch ($\Delta P/P$) as a function of the relative frequency shift of a selected partial ($\Delta F/F$). The shift of the strike note pitch of the "long" bells (L) is shown on the left of the figures: the results for the "short" bells (S) are plotted on the right. The numeric symbol attached to the L or S indicates which partial is shifted in frequency. When the frequency of the 6th partial (twelfth) is shifted too much, the subjects match the comparison tone to a subharmonic of the twelfth and not to the strike note of the bell (analytic perception).

PB



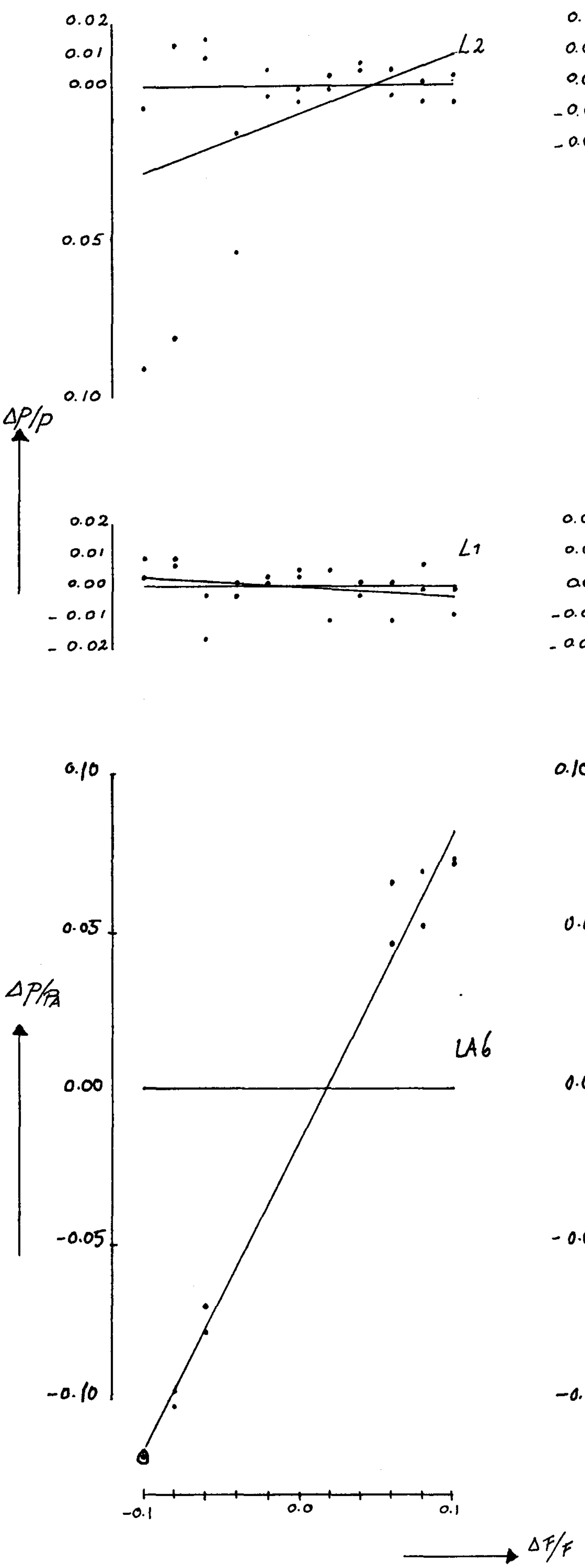
A-4

PB



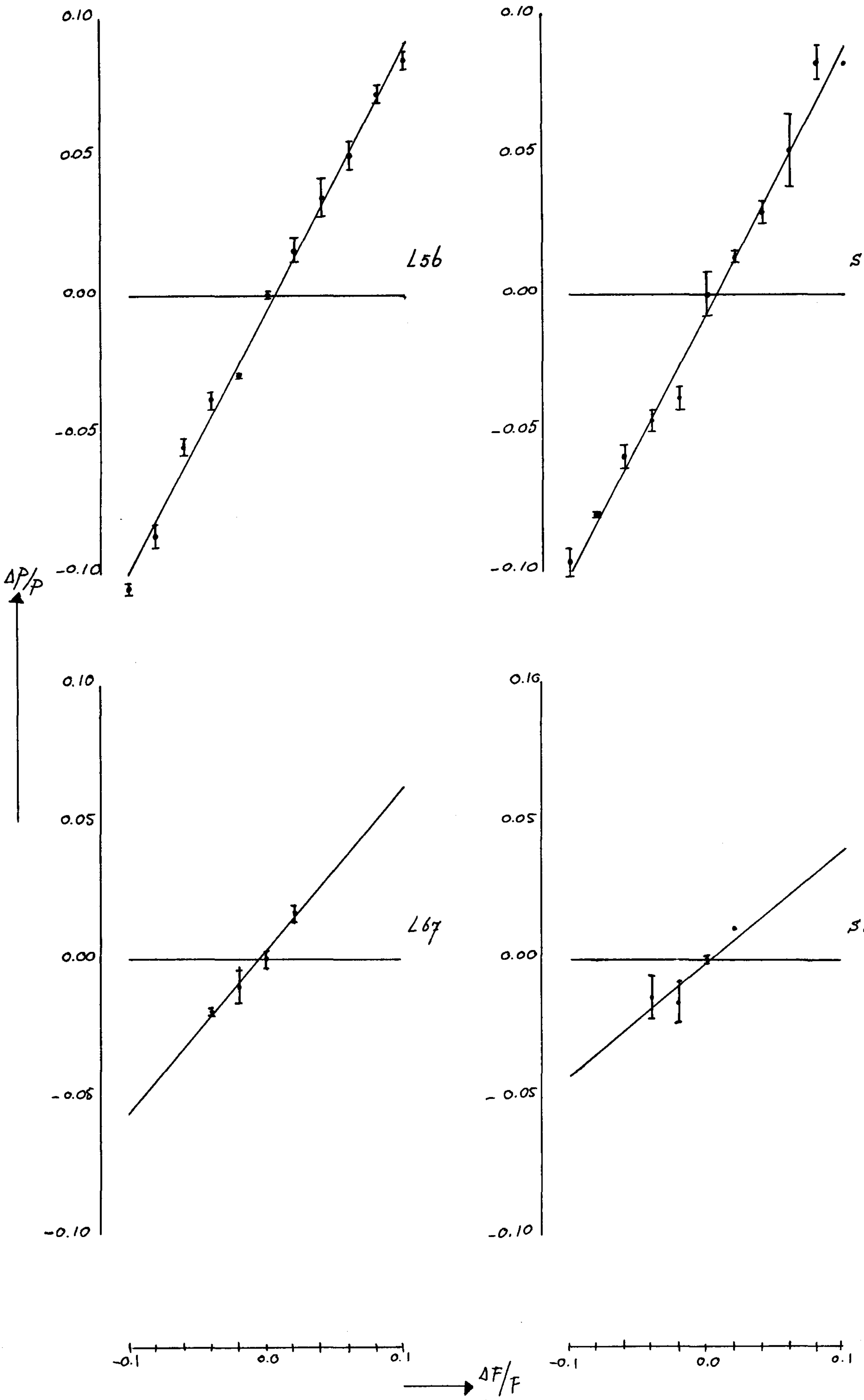
A-5

PB



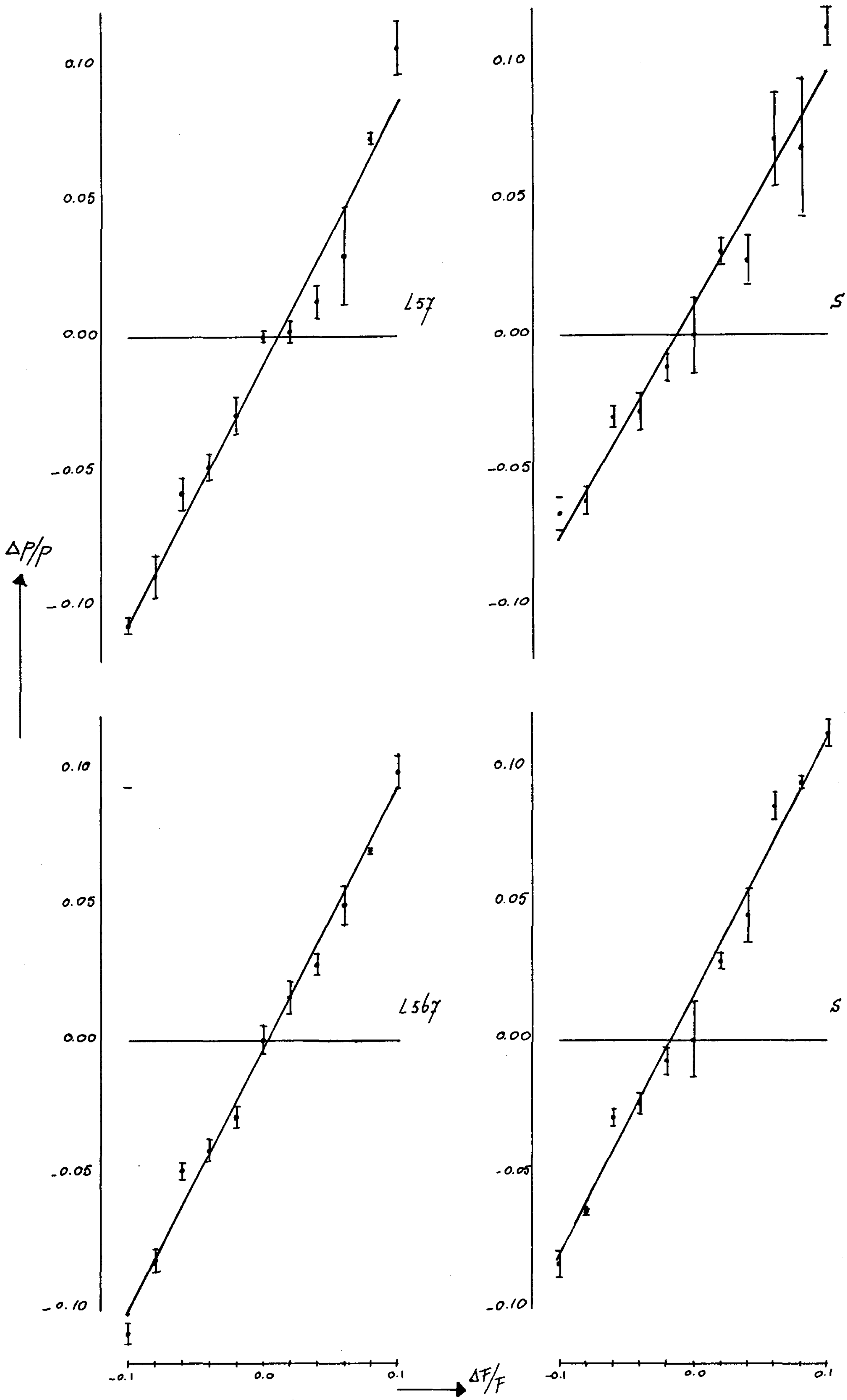
L2
S2
L1
S1
LA6
SA6
ΔP/P
ΔP/PA
ΔF/F
0.02
0.01
0.00
-0.01
-0.02
0.05
0.10
-0.05
-0.10
-0.1
0.0
0.1

A-6



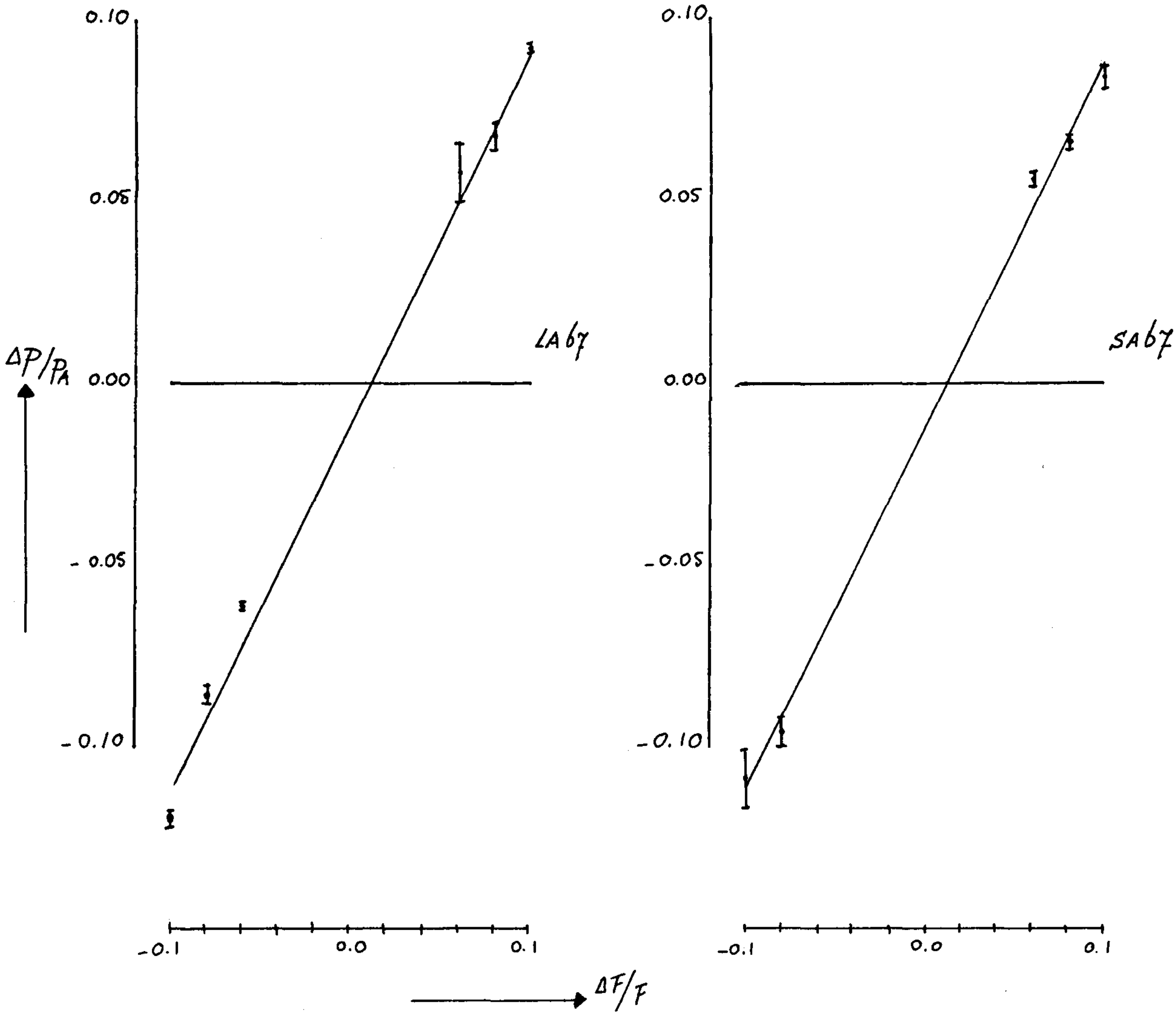
A-7

PB

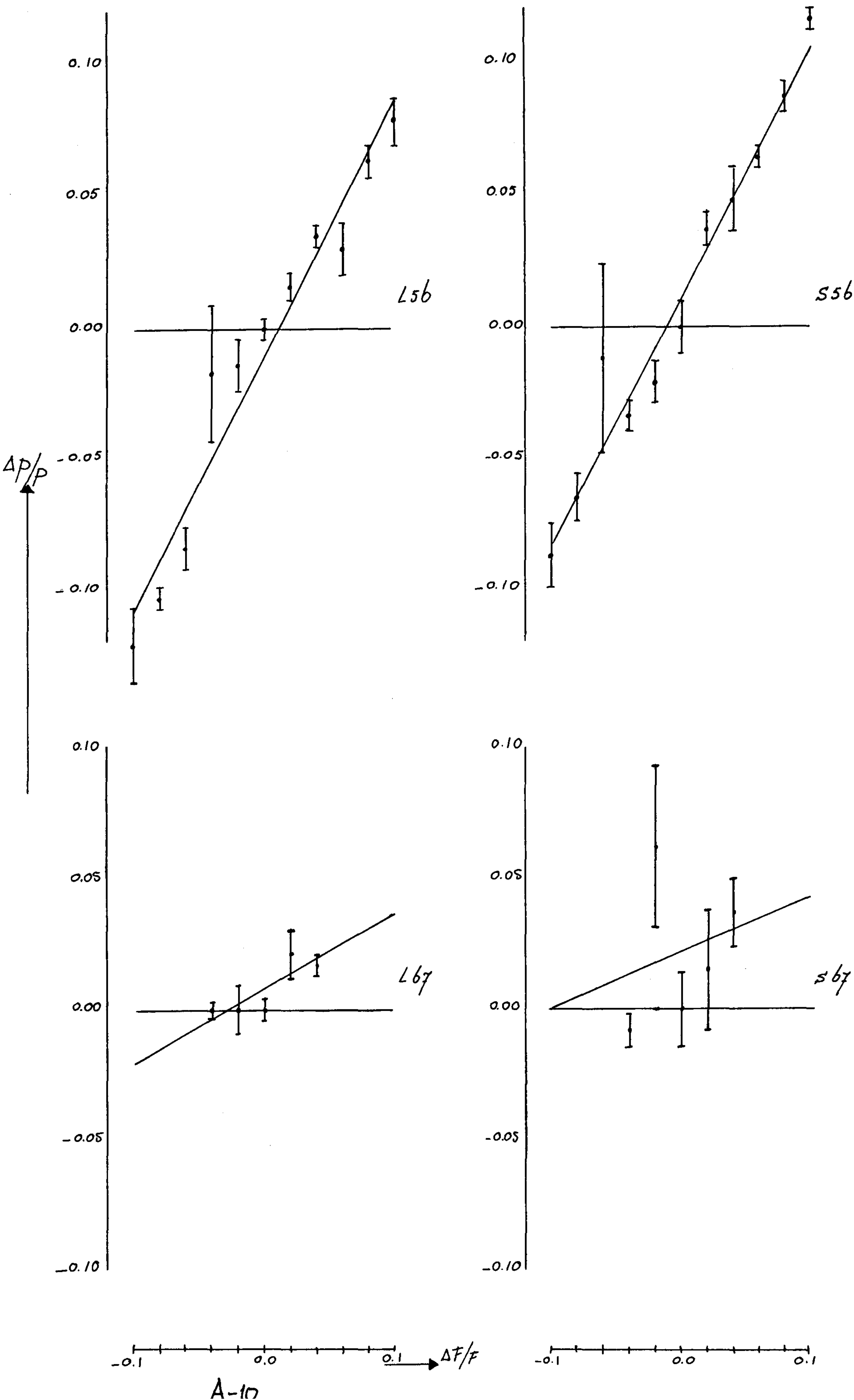


A-8

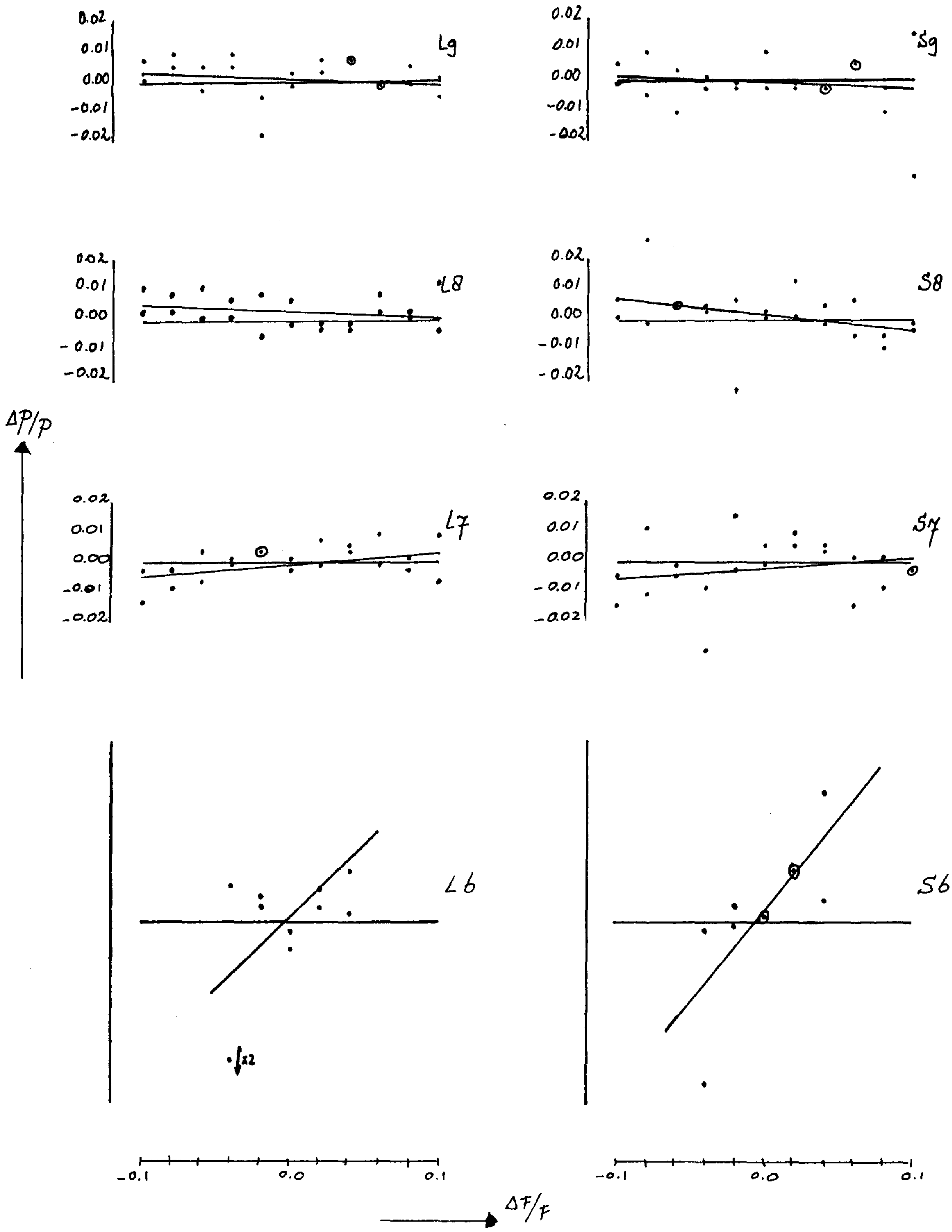
PB



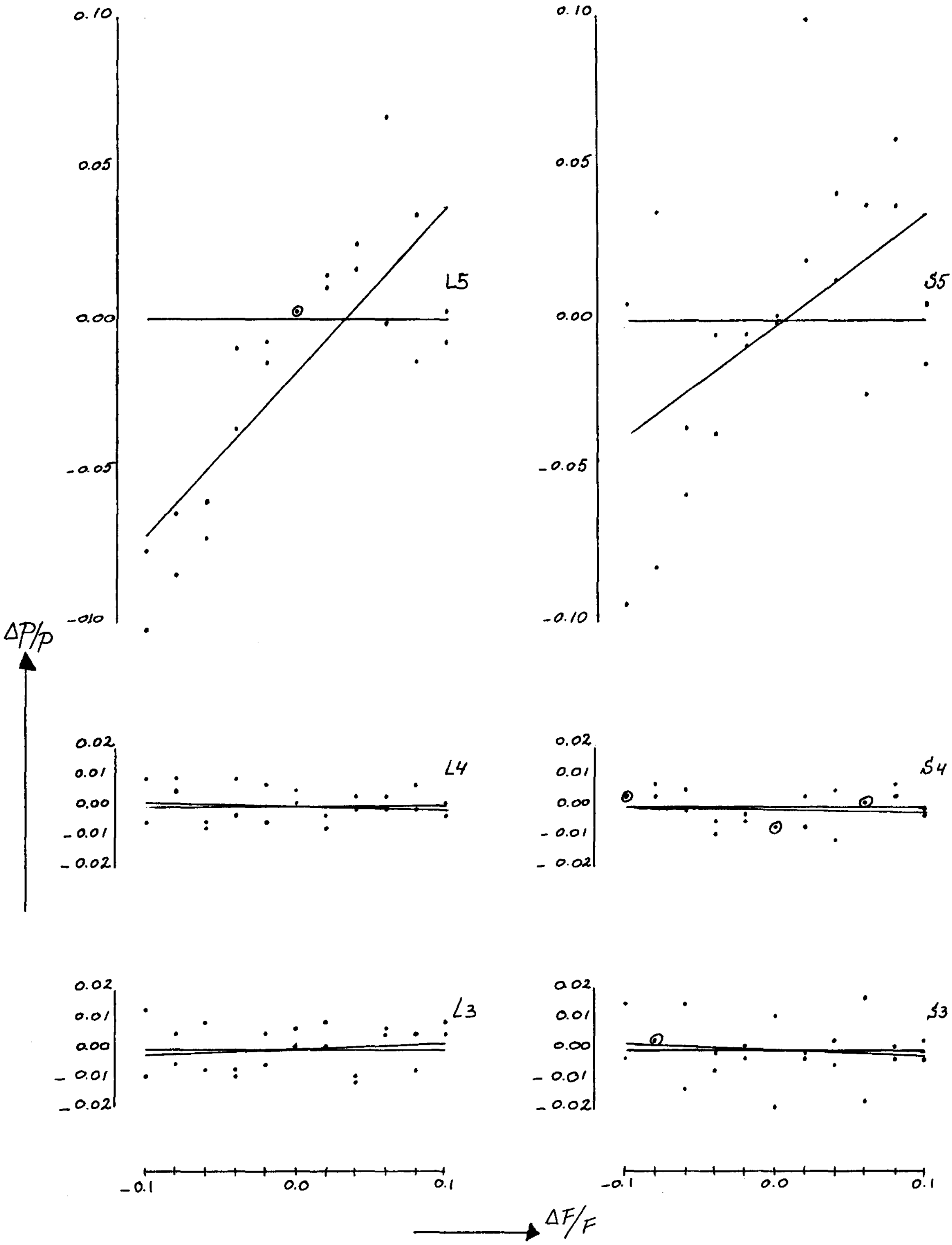
ΔP/PA
0.10
0.05
0.00
- 0.05
- 0.10
LA 67
-0.1
0.0
0.1
ΔF/F
0.10
0.05
0.00
-0.05
-0.10
SA 67
-0.1
0.0
0.1


L56
S56
L67
S67
0.10
0.05
0.00
-0.05
-0.10
ΔP/P
-0.1
0.0
0.1
ΔF/F

H K

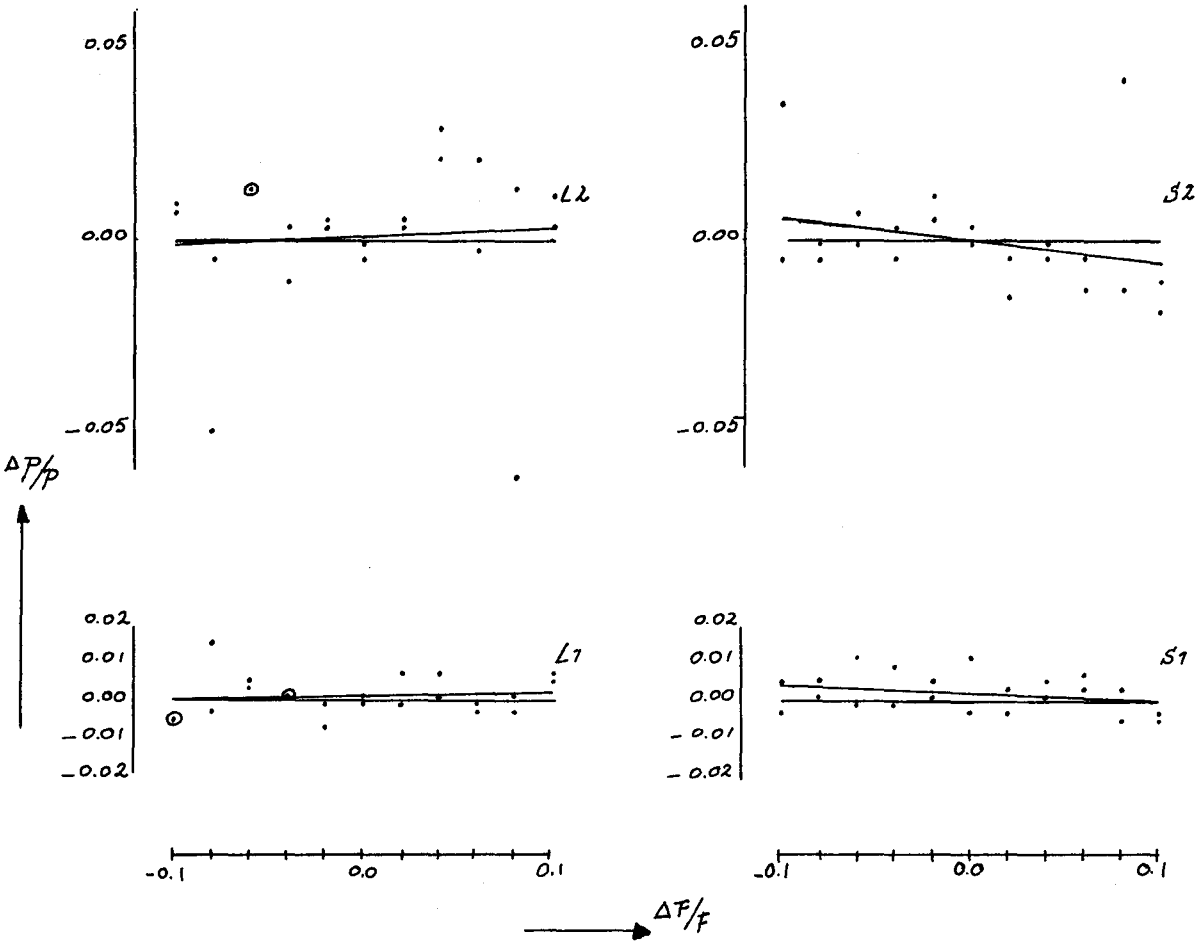


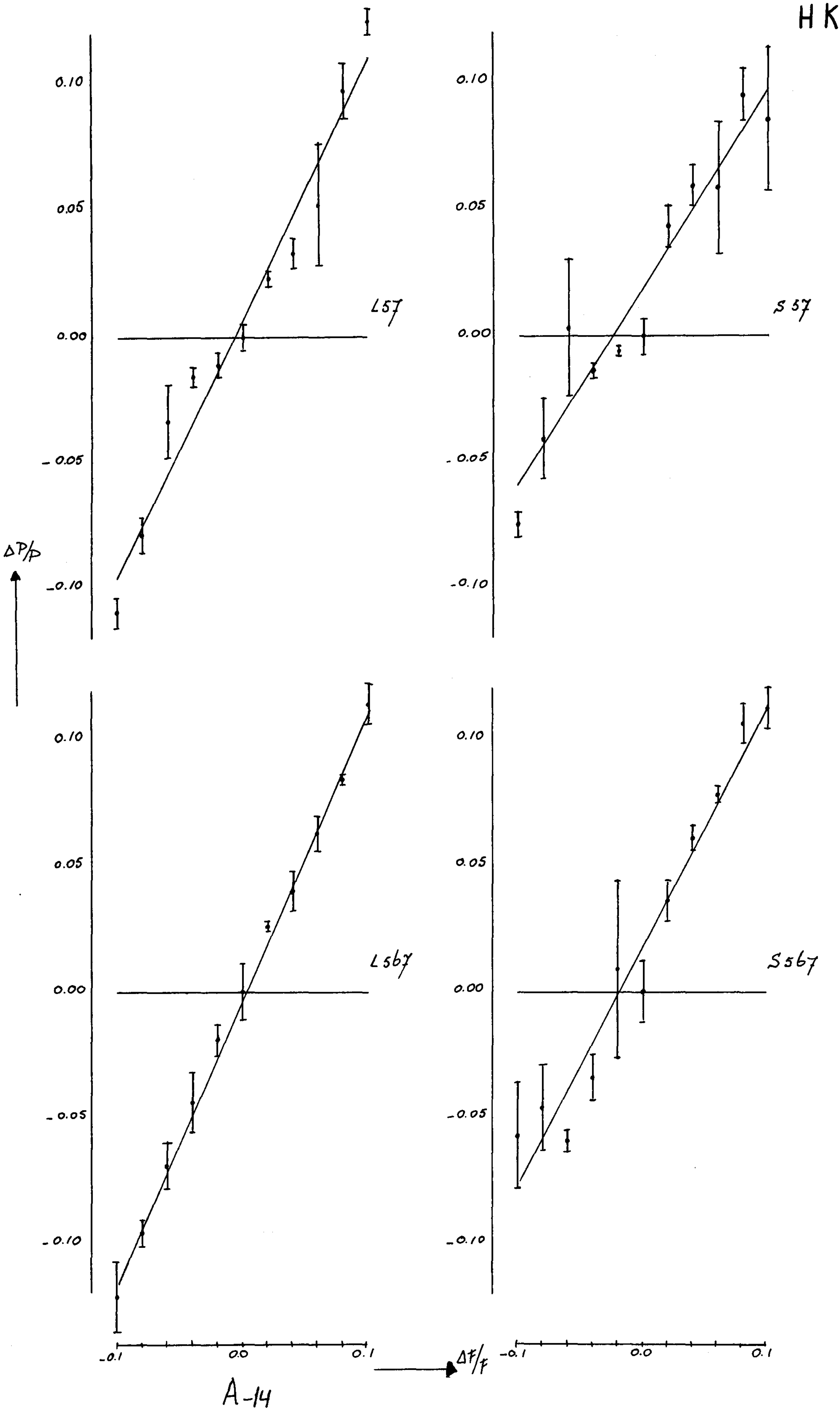
H K

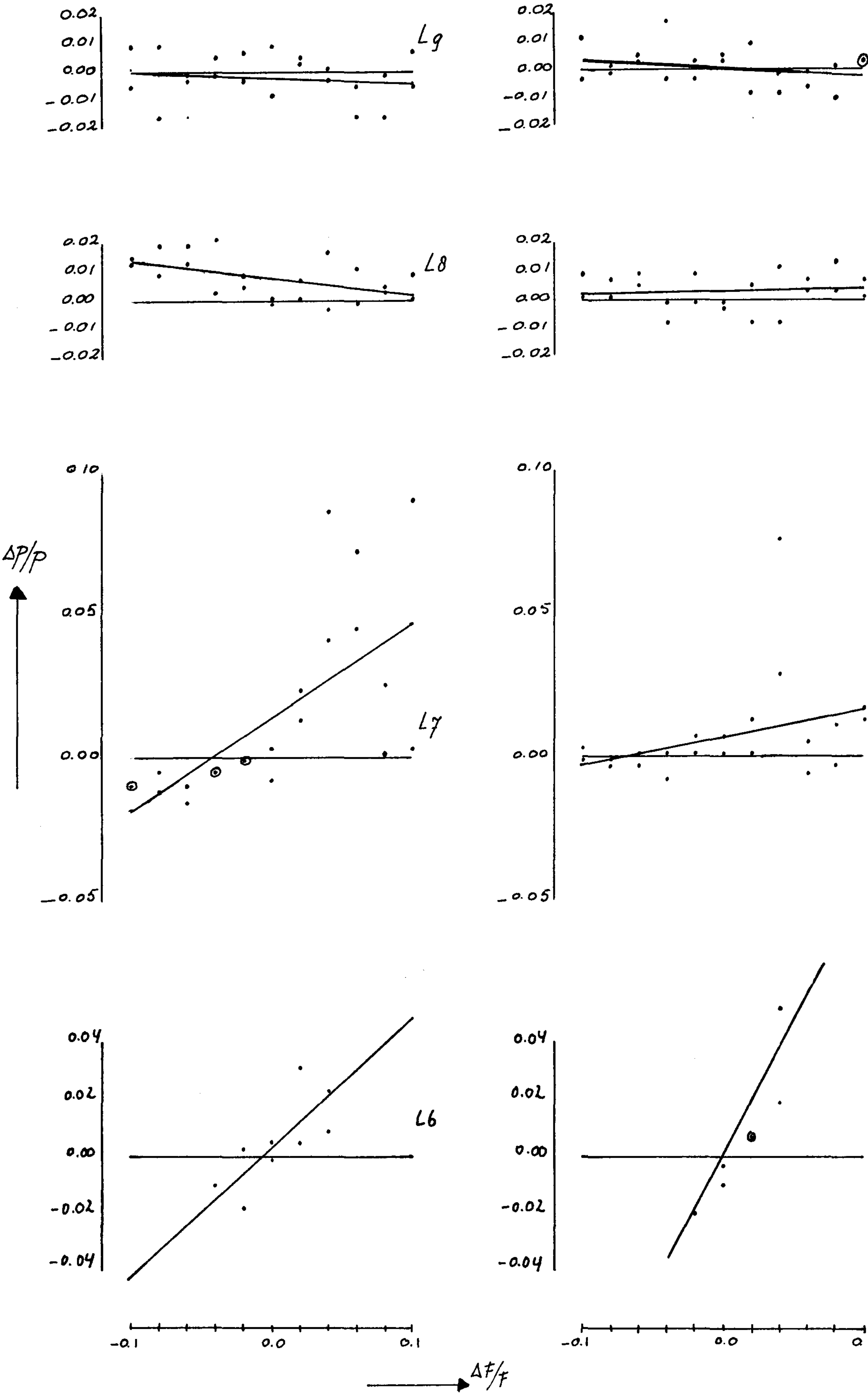


0.10
0.05
0.00
-0.05
-0.10
L5
S5
ΔP/P
0.02
0.01
0.00
-0.01
-0.02
L4
S4
L3
S3
-0.1
0.0
0.1
ΔF/F

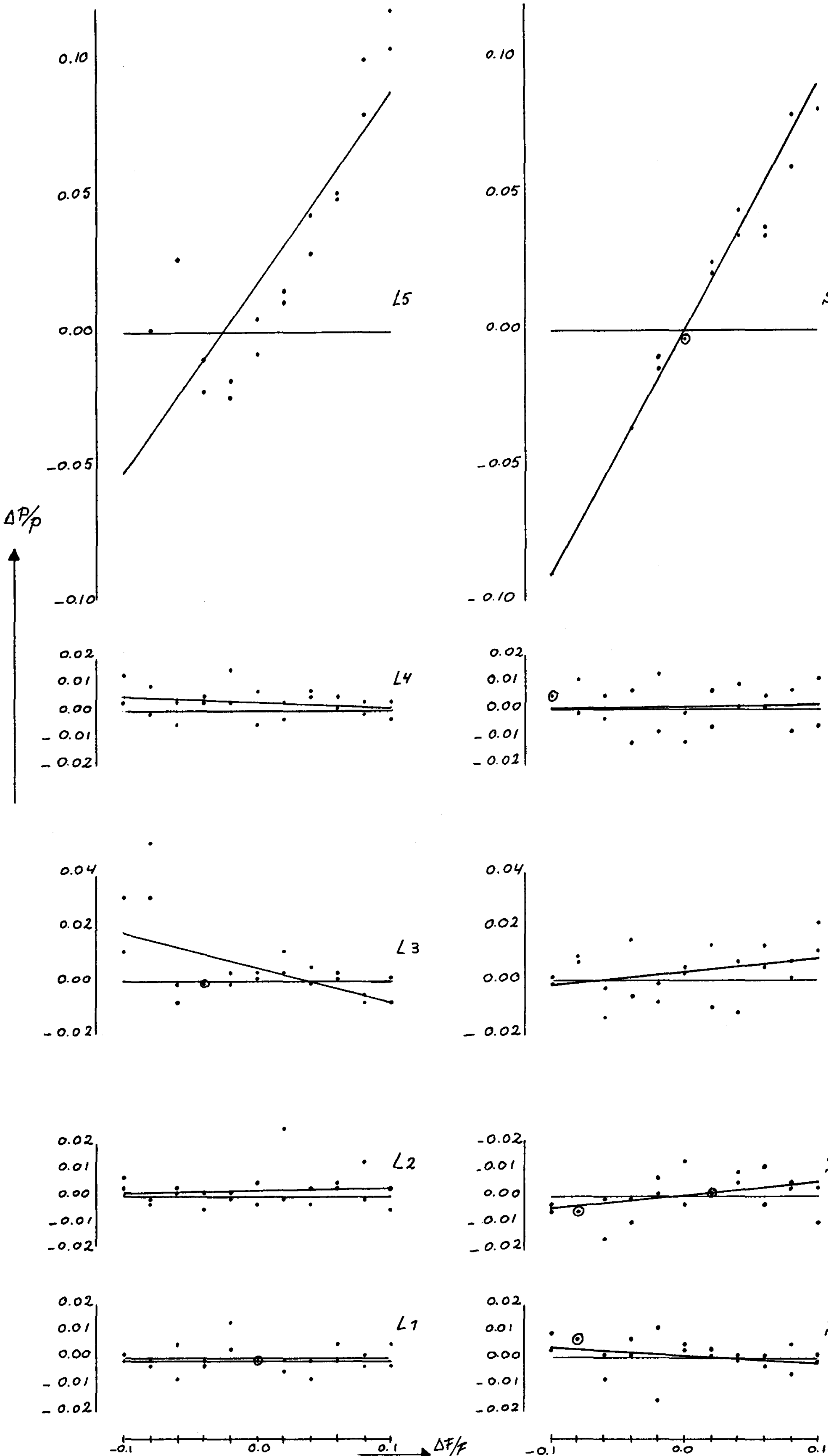
A-12


0.05
0.00
-0.05
L2
S2
ΔP/P
0.02
0.01
0.00
-0.01
-0.02
L1
S1
-0.1
0.0
0.1
ΔF/F


HK
L57
S57
L567
S567
ΔP/P
ΔF/F
0.10
0.05
0.00
-0.05
-0.10
-0.1
0.0
0.1
A-14

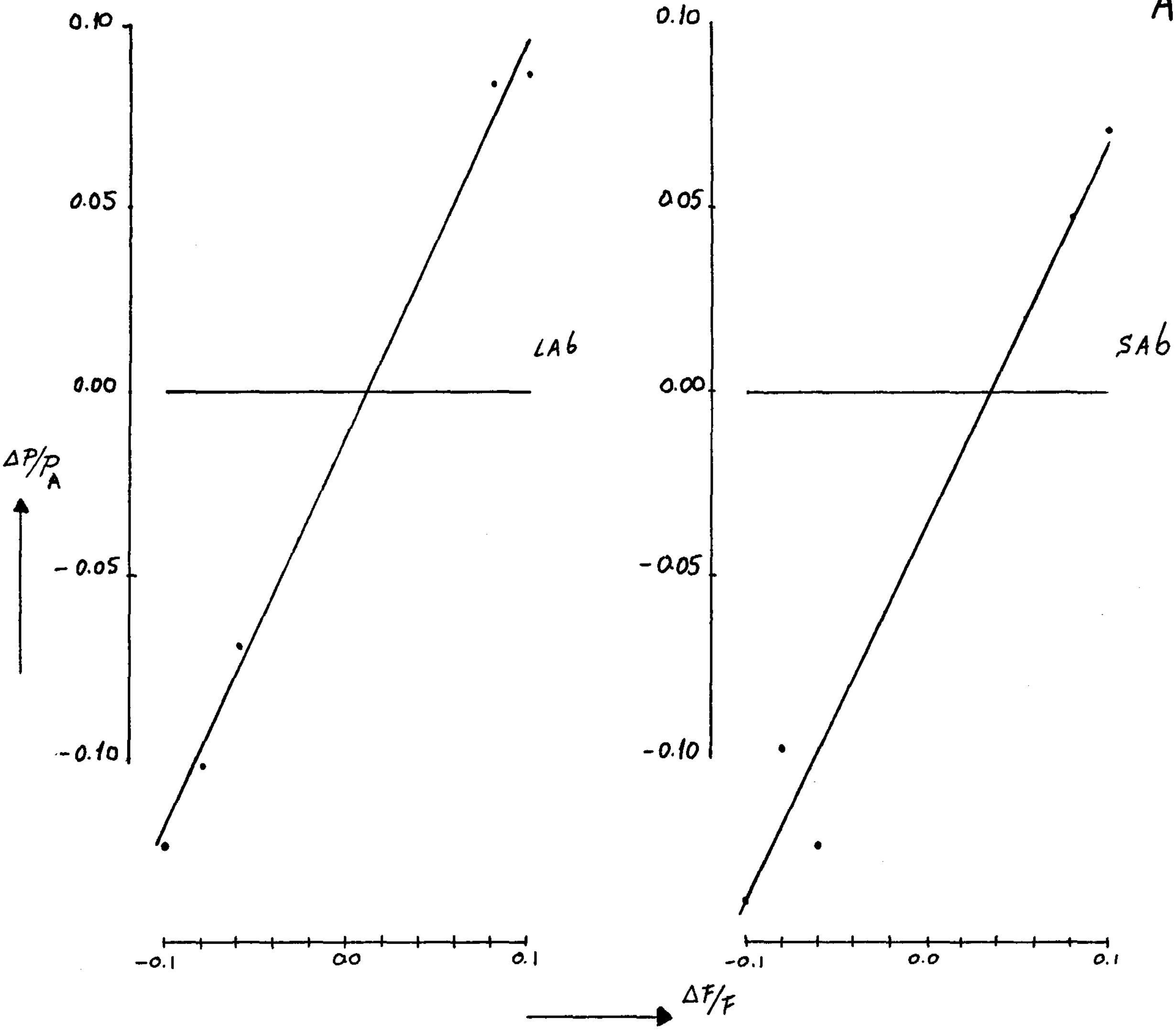

L9
S9
L8
S8
L7
S7
L6
S6
ΔP/P
ΔF/F
0.02
0.01
0.00
-0.01
-0.02
0.10
0.05
-0.05
0.04
-0.04
-0.1
0.0
0.1

AH

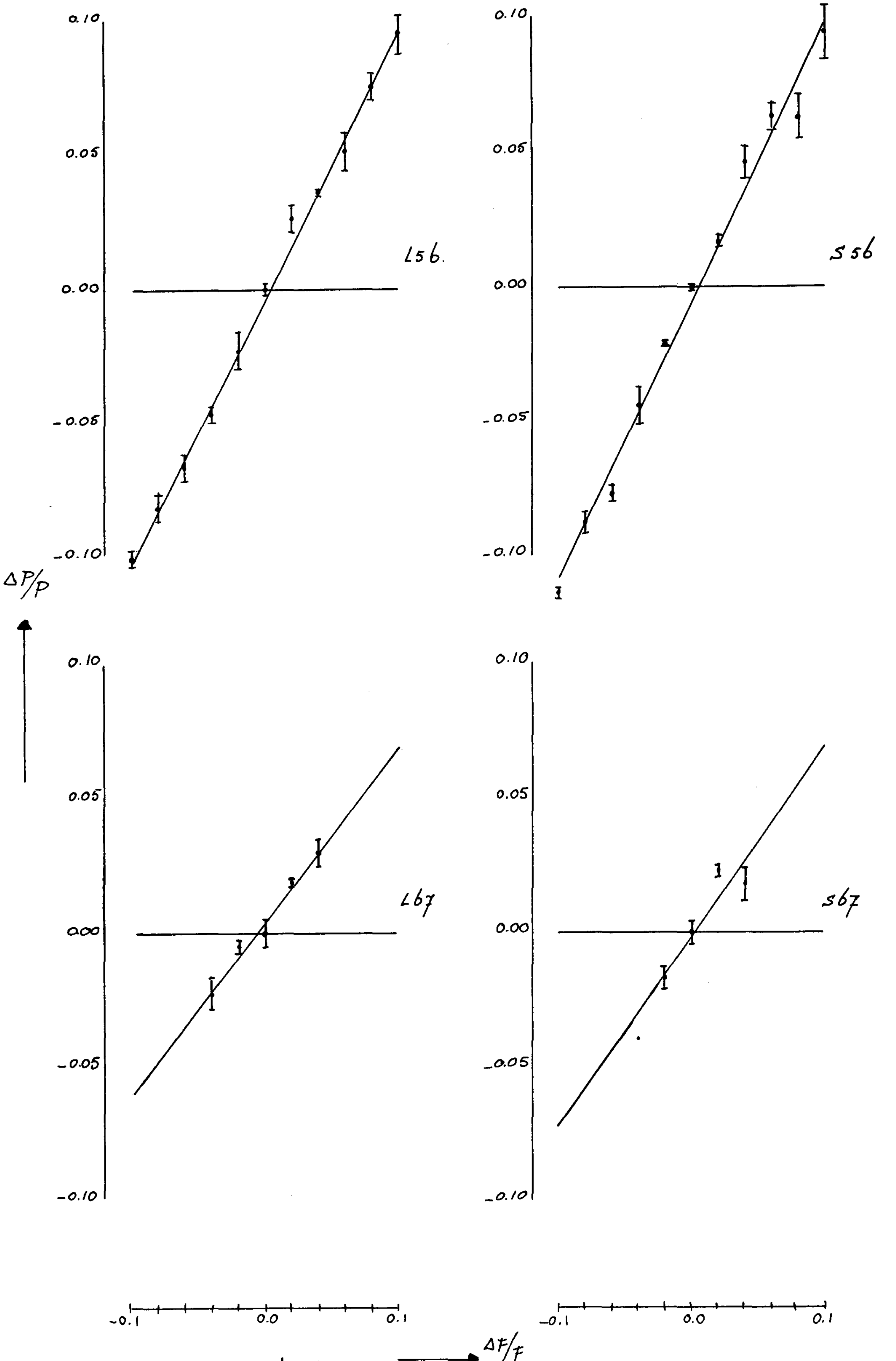


L5
S5
L4
S4
L3
S3
L2
S2
L1
S1
ΔP/P
ΔF/F
0.10
0.05
0.00
-0.05
-0.10
0.02
0.01
-0.01
-0.02
0.04
-0.1
0.0
0.1

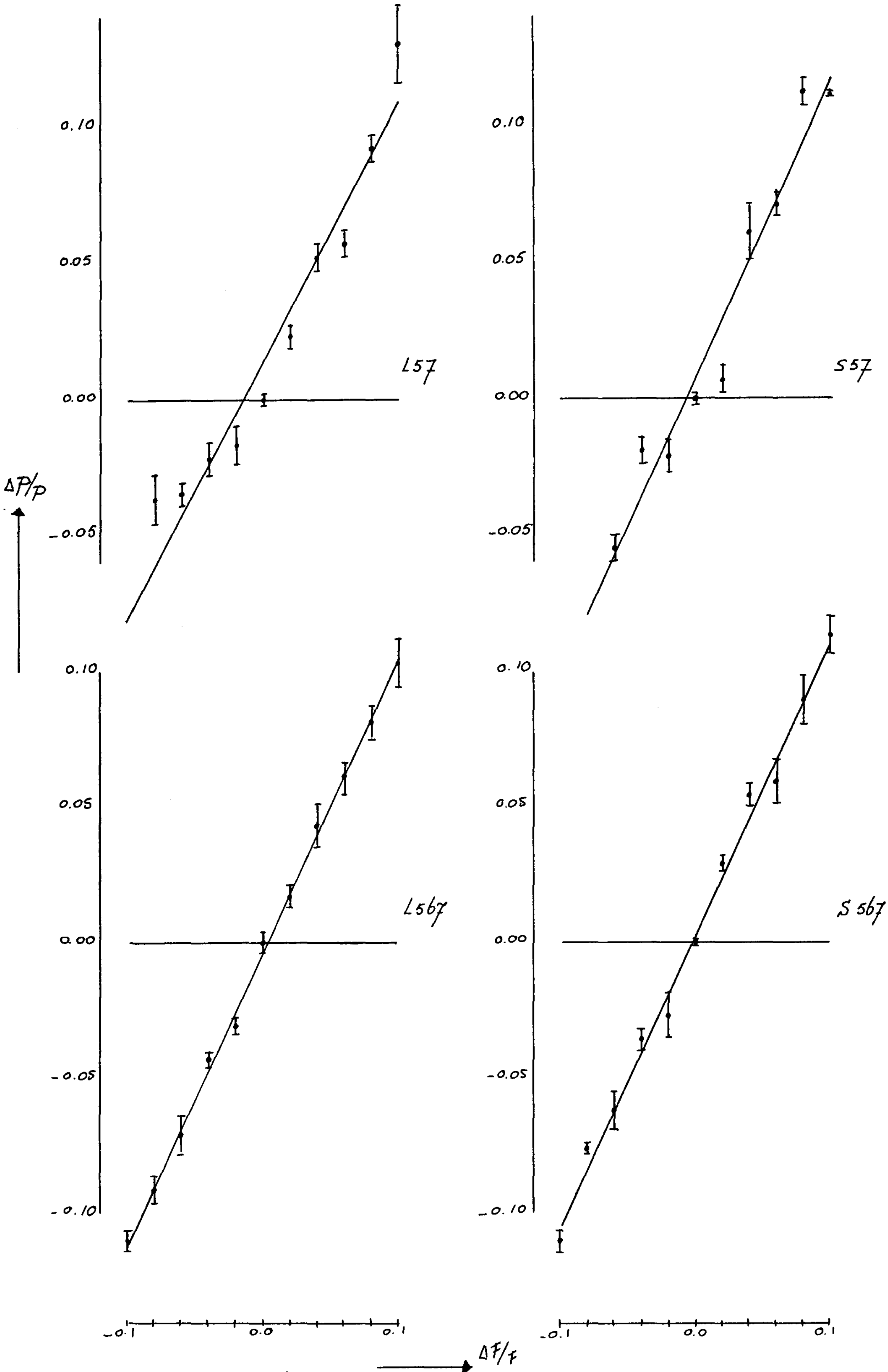
A H



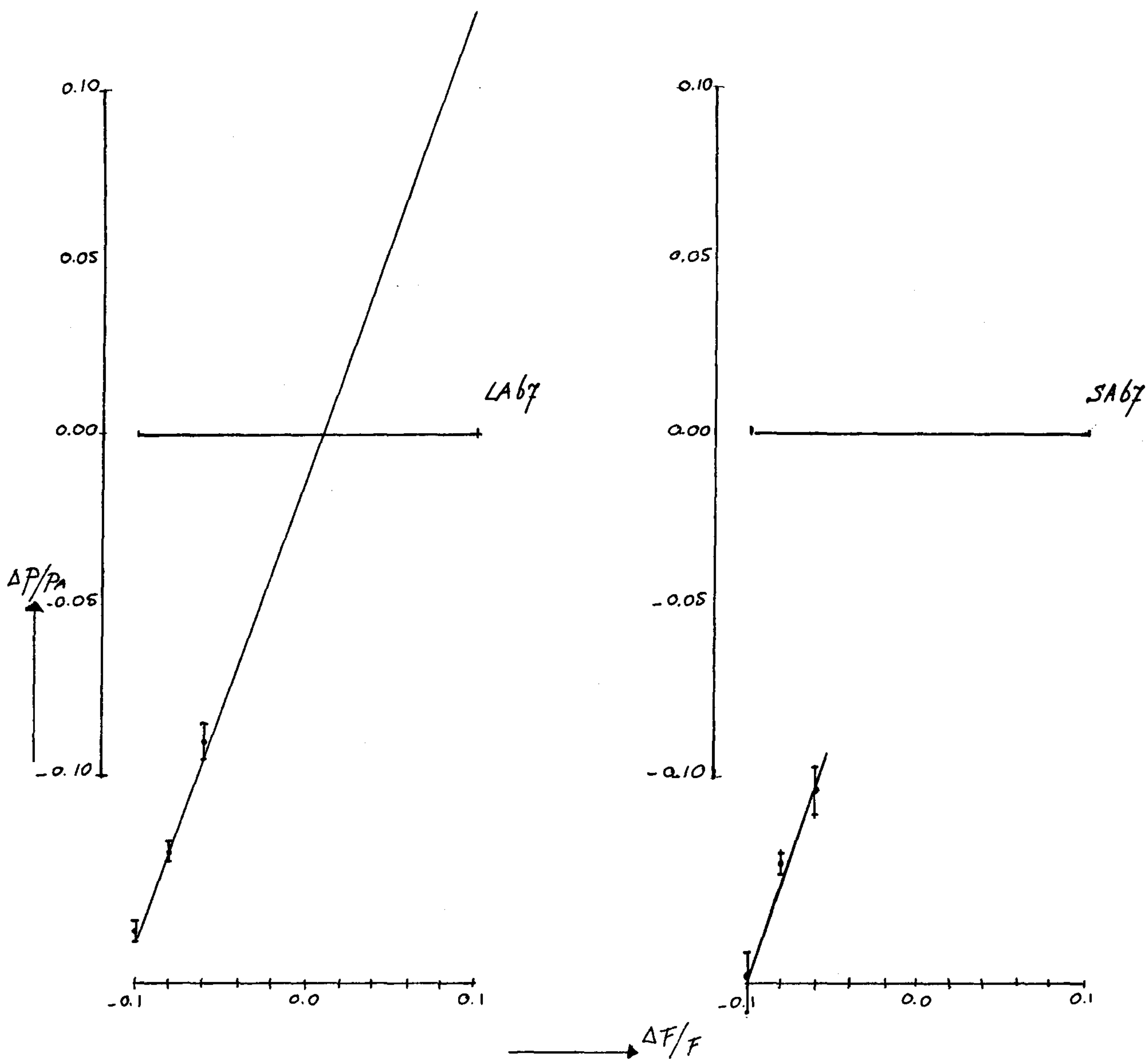
A H



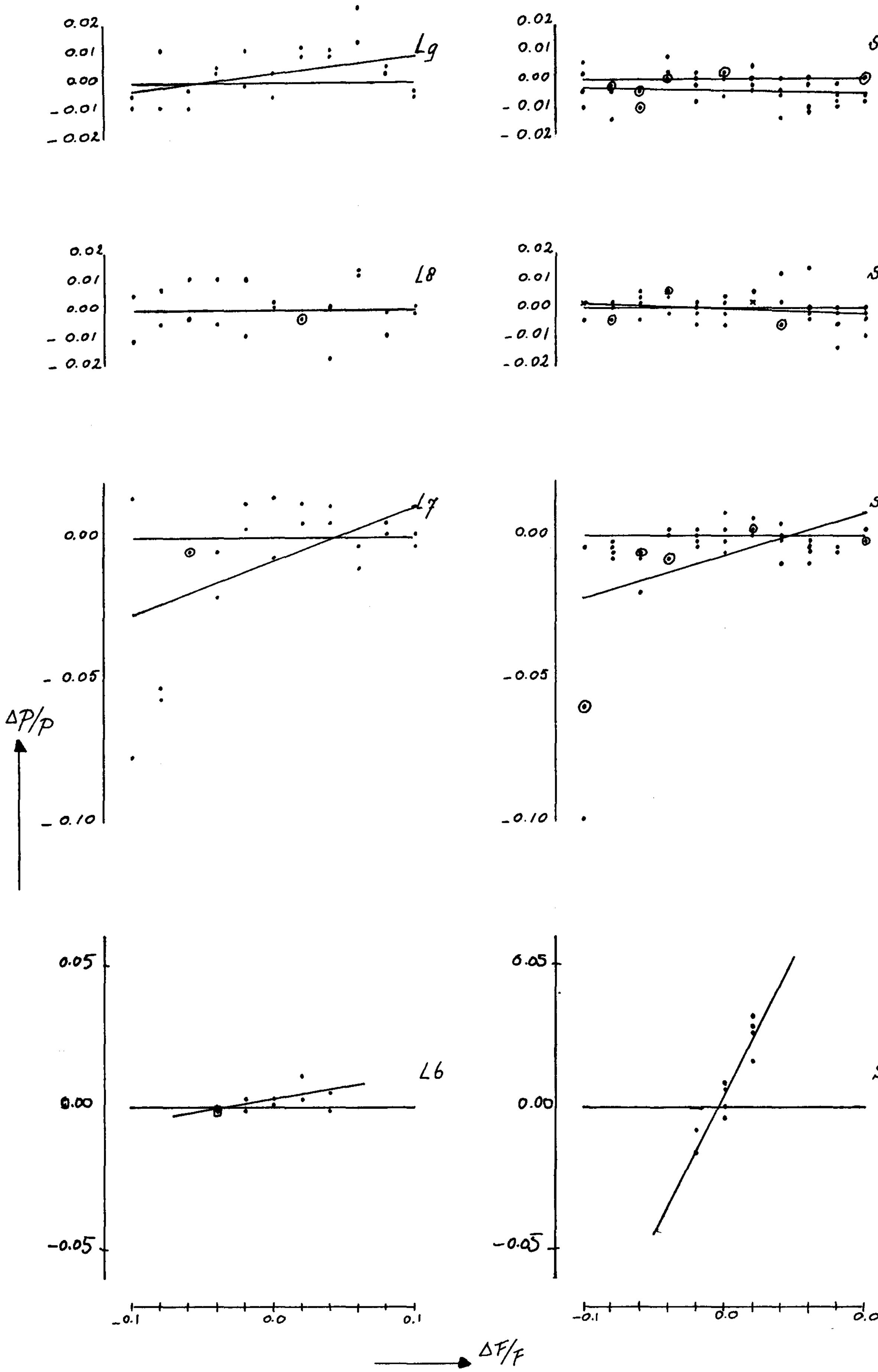
A H

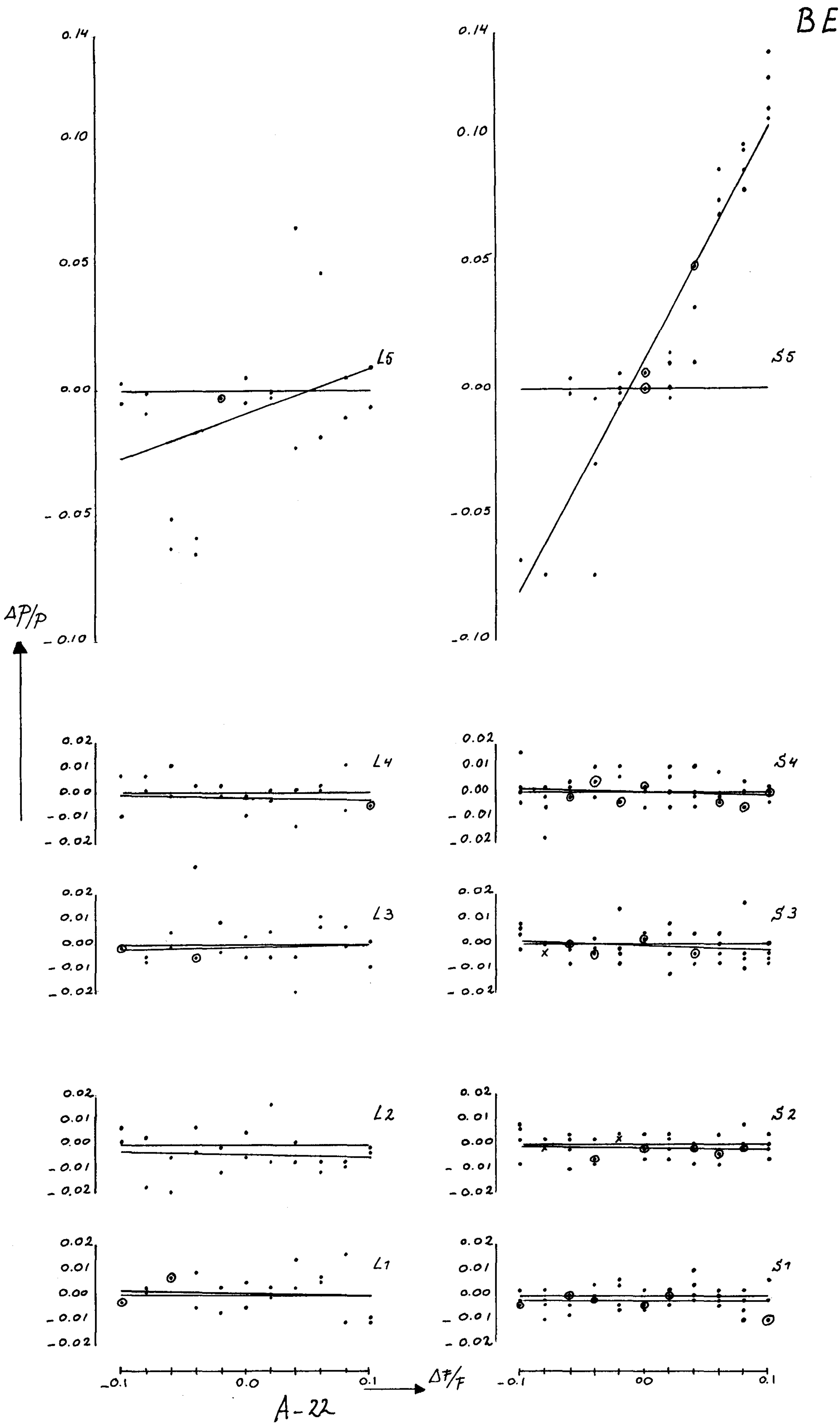


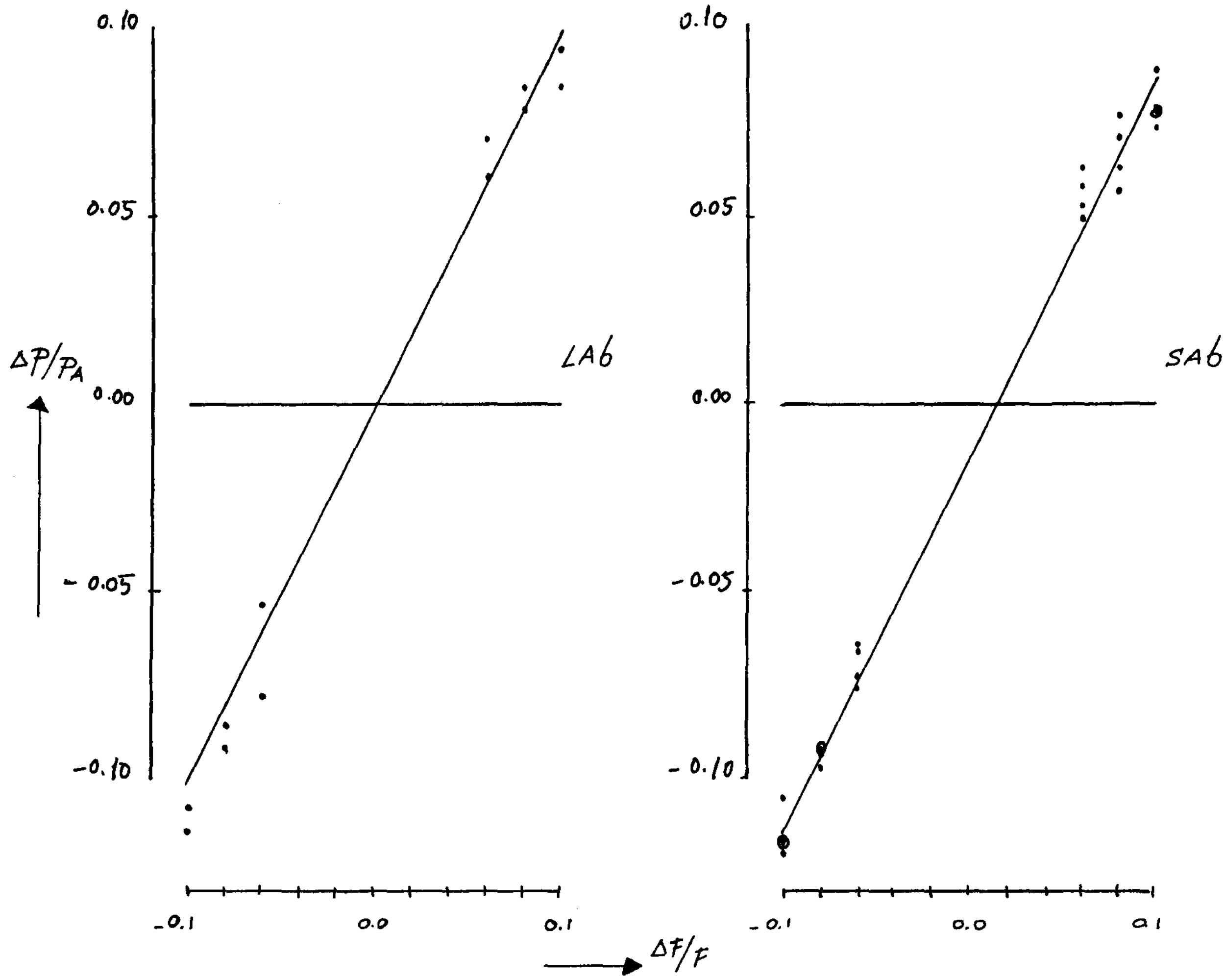
A-19

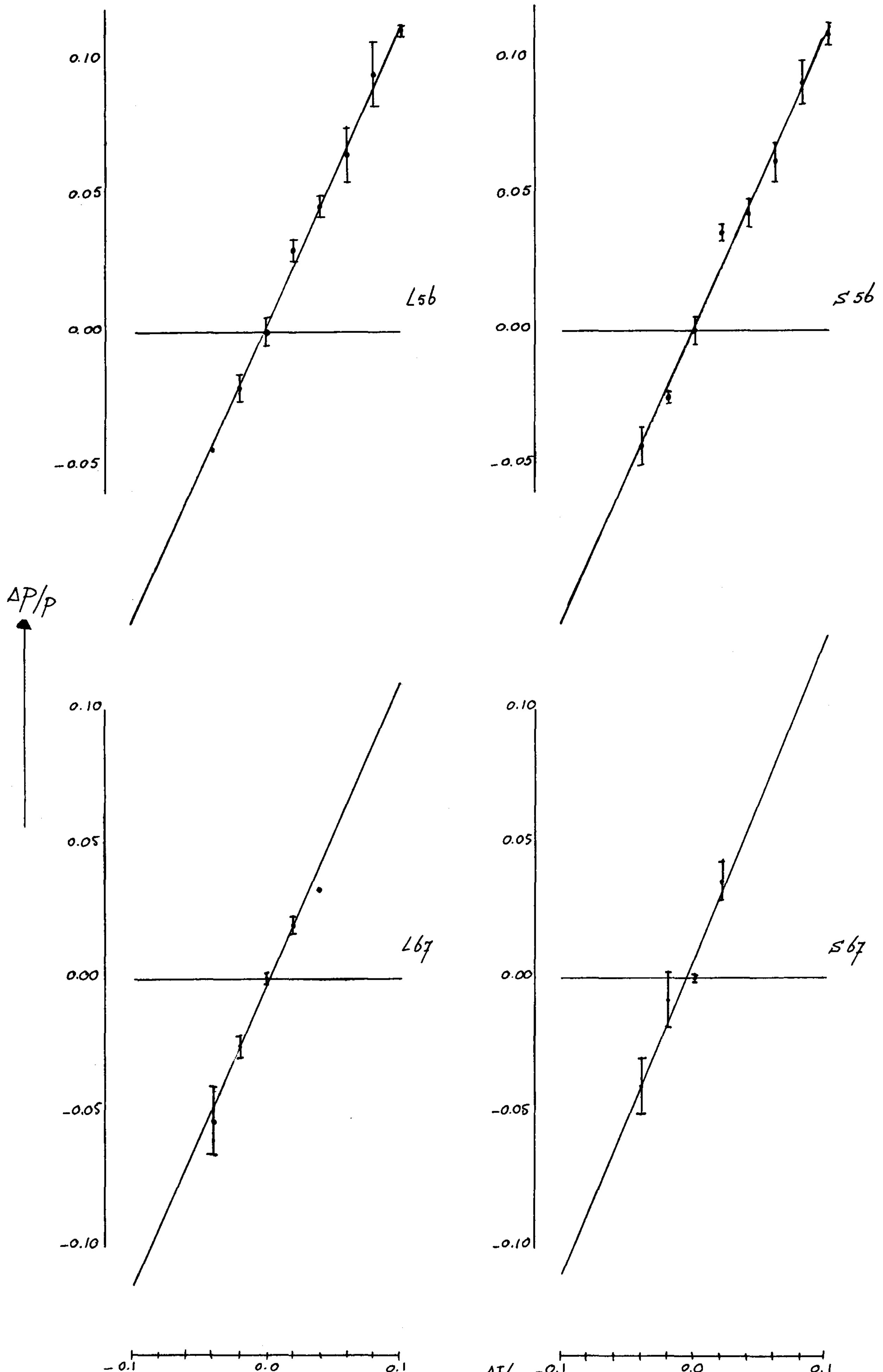


LA67
SA67
0.10
0.05
0.00
-0.05
-0.10
$\Delta P/P_A$
-0.1
0.0
0.1
$\Delta F/F$


L9
S9
L8
S8
L7
S7
L6
S6
0.02
0.01
0.00
- 0.01
- 0.02
- 0.05
- 0.10
0.05
-0.05
ΔP/P
-0.1
0.0
0.1
ΔF/F

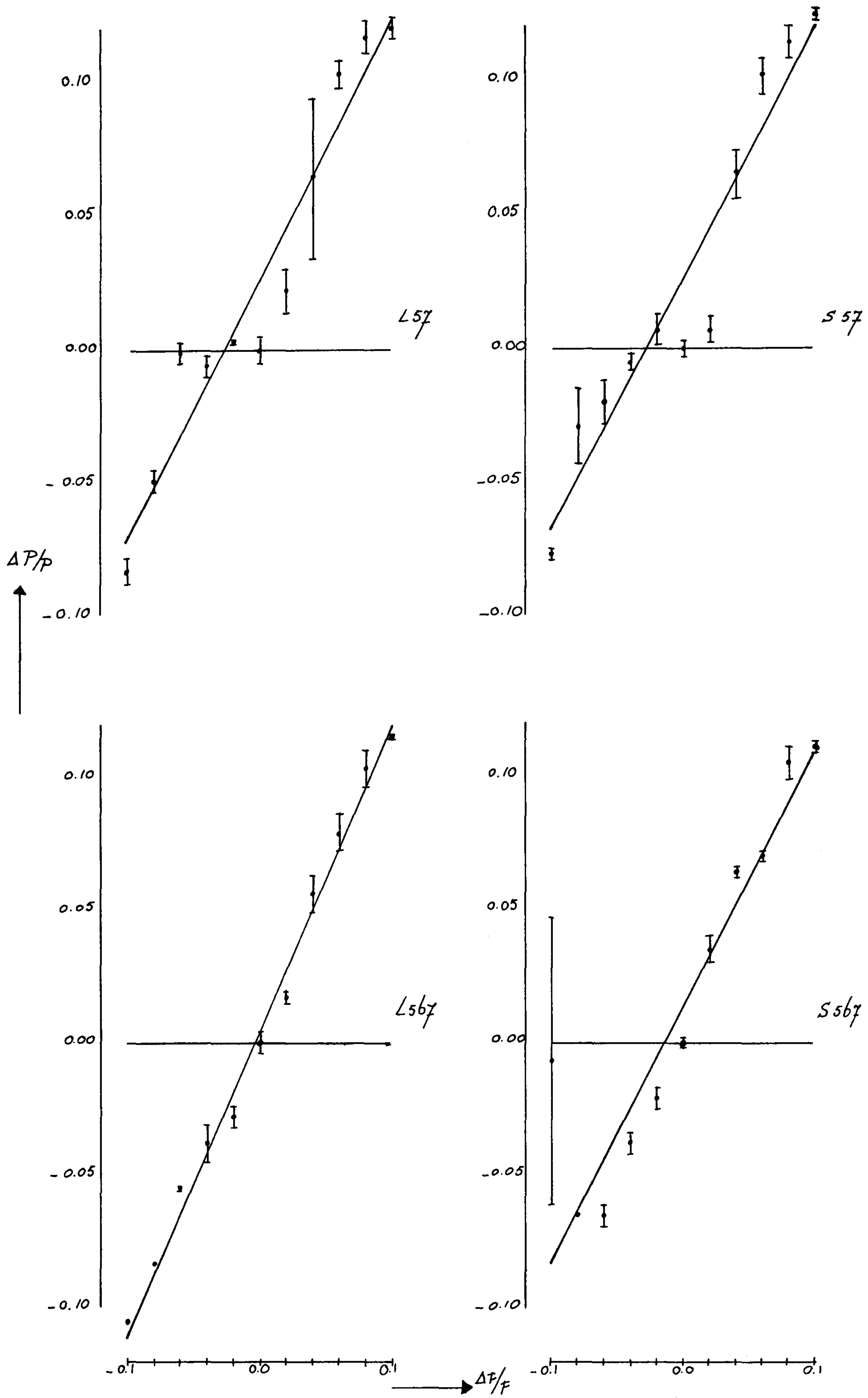

BE
L5
S5
L4
S4
L3
S3
L2
S2
L1
S1
ΔP/P
ΔF/F
A-22

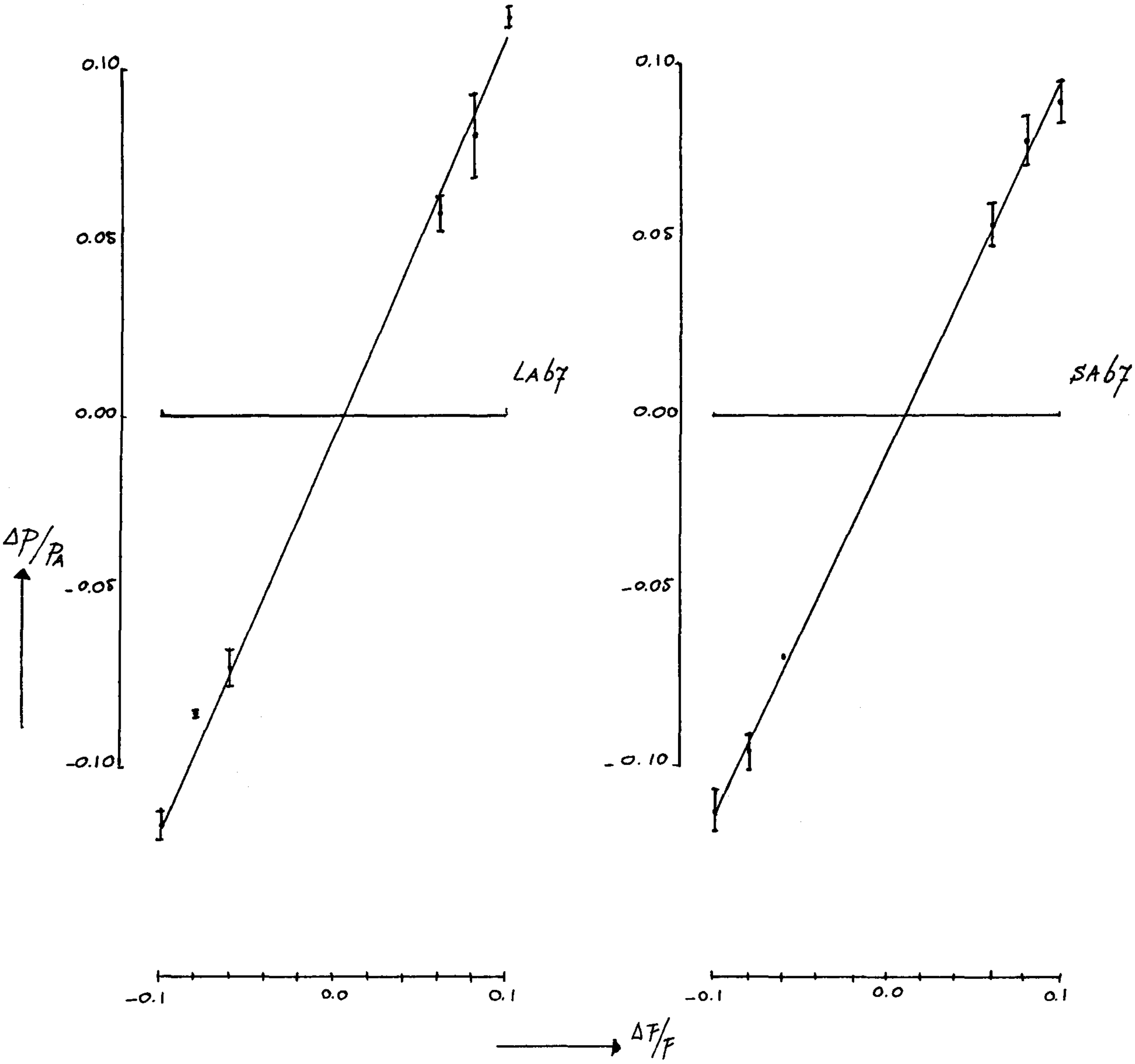

0.10
0.05
0.00
-0.05
-0.10
ΔP/PA
LAb
SAb
-0.1
0.0
0.1
ΔF/F



ΔP/P
0.10
0.05
0.00
-0.05
-0.10
L56
S56
L67
S67
-0.1
0.0
0.1
ΔF/F

BE




0.10
0.05
0.00
-0.05
-0.10
LA67
SA67
ΔP/P_A
-0.1
0.0
0.1
ΔF/F

```
c       ANALYSE.FOR
c
*************************************************************************
* This algorithm calculates the spectral pitch-pattern according to *
* Terhardt. (1982)                                                      *
*                                                                       *
*    J.H.EGGEN  1985                                                    *
*************************************************************************
c
c
       PROGRAM Analyse
       IMPLICIT integer*2 (i-n)
       DIMENSION naln(15),nas(15),nav(15),idr(256)
       DIMENSION arrmax(200),relamp(200),dbmax(200)
       DIMENSION rin(2048),x(2048),y(2048),result(2048),dbres(2048)
       DIMENSION zbr(1024),dmask(200),dlex(200),pshift(200),ws(200)
       INTEGER*4 idr4(128)
       BYIE kar,prfl
       CHARACTER*30 nasa,nava
       EQUIVALENCE (idr(1),idr4(1))
       EQUIVALENCE (nas(1),nasa(1:1))
       EQUIVALENCE (nav(1),nava(1:1))
       DATA nasa(:8)/'DUM1.DAT'/
       DATA nava(:8)/'DUM2.DAT'/
       DATA nx/2h  /
c
c
       io=6
       print *
       n=INAL('Printer output  (yes=y) ?',1,prfl)
       if (prfl.eq.'y') then
         io=11
    5    print *
         nas(1)='D:'
         n=INAL('Name DATA-file :                    D:',28,nas(2))
         nasa(n+3:n+7)='.DAT'
         call FLXST(nas,30,ierr,lln)
         if (ierr.eq.0) then
           print '(/,x,''***** FILE ALREADY EXISTS *****'')'
           goto 5
         endif
         OPEN(unit=io,file=nas,access='sequential',status='new',err=1C
       endif
c
c
   10 naln(1)='N:'
       print *
       n=INAL('Specify N-file :                    N:',28,naln(2))
       if (n.eq.0) then
         print '(/,x,''***** N-FILE SPECIFICATION PLEASE !!! *****'')'
         goto 10
       endif
       call FLXST(naln,30,ierr,lln)
       if (ierr.ne.0) then
         print '(/,x,''***** FILE NOT FOUND !!! *****'')'
         goto 10
       endif
c
c
       iv=10
```

```
 15 print *
    nav(1)='D:'
    n=INAL('Name SP-pattern DATA-file :   D:',28,nav(2))
    nava(n+3:n+7)='.DAT'
    call FLXST(nav,30,ierr,lln2)
    if (ierr.eq.0) then
      print '(/,x,''***** FILE ALREADY EXISTS *****'')'
      goto 15
    endif
    OPEN(unit=iv,file=nava,access='sequential',status='new',err=100)
c
c
    call FLRED(naln,30,idr4,0,1)          ! Read ID-record
    lln=lln-idr4(65)                      ! Subtract extra ID-records
    sf=float(idr4(62))                    ! Sample-frequency
    print *
    write(io,'(10x,''*************************'')')
    write(io,'(10x,''*'',7x,8a2,''*'')')(naln(i),i=1,8)
    write(io,'(10x,''*************************'')')
    write(io,'(/,x,''N-file contains '',i4,'' records.'',/)')lln-1
    write(io,'(x,''1 record  ='',f9.1,'' ms.'')')(256/sf)*1000.
    do i=2,8
      write(io,'(i2,'' records ='',f9.1,'' ms.'')')i,i*(256/sf)*1000.
    enddo
    write(io,'(/)')
    n=ININ('Analysis-window, in records            (max. 8) ? ',1,iar)
    n=ININ('Total number of records to be analyzed          ? ',1,itr)
    n=ININ('Start record                              ( >0 ) ? ',1,ibr)
    n=INRE('dB level strongest component (real, <cr>=0.dB) ? ',1,dni)
    if (n.eq.0) dni=0.
    n=INRE('Max. amp. difference in dB  (real, <cr>=43.dB) ? ',1,grdb)
    if (n.eq.0) grdb=43.
    iap=iar*256
    ant=float(iap)
    k=inint(alog10(ant)/alog10(2.))
    deltaf=sf/ant
    print *
    write(io,'(x,''Samplefrequency    ='',f9.2,'' Hz.'')')sf
    write(io,'(x,''Delta-f            ='',f9.2,'' Hz.'')')deltaf
    write(io,'(x,''Minimal amplitude  ='',f9.2,'' dB.'')')
   @          dni-grdb
c
c
    print *
    n=inal('Apply window                     (yes=y or <cr>) ? ',1,kar)
    if (kar.eq.'y'.or.n.eq.0) then
      n=inal('Hamming (=m) or Hanning (=n)                   ? ',1,kar)
      if (kar.eq.'m') then
        write(io,'(/)')
        write(io,'(x,''***** HAMMING window applied !!! *****'')')
      elseif (kar.eq.'n') then
        write(io,'(/)')
        write(io,'(x,''***** HANNING window applied !!! *****'')')
      else
        write(io,'(/)')
        write(io,'(x,''***** NO WINDOW APPLIED !!! *****'')')
        write(io,'(/)')
      endif
    endif
c
c
```

```
      lmaal=itr/iar
      do 40 m=0,lmaal-1
        ilr=ibr+m*iar                                  ! Current start-rec
        call LNFL (naln,rin,ilr,iar)                   ! Read data
        if (kar.eq.'m') then
          call HAMMING(iap,rin)                        ! Windowing
        elseif (kar.eq.'n') then
          call HANNING(iap,rin)
        endif
        call FFTR(rin,x,y,iap,k,ffsq,1,nprv,lprv)      ! FFT on rin
        np=2**(k-1)
        do i=1,np
          x(i)=x(i)*2/2**k                             ! Multiply with
          y(i)=y(i)*2/2**k                             ! Scale factor
          result(i)=sqrt(x(i)**2 + y(i)**2)            ! Amplitude spectru
        enddo
        call LOKMAX(np,result,arrmax)                  ! Local maxima
        call PARINT(deltaf,result,arrmax)              ! Parabolic interp.
        totmax=0.
        iamax=inint(arrmax(1))                         ! Number local max.
        do i=1,iamax
          j=2*i
          if (arrmax(j+1).gt.totmax) totmax=arrmax(j+1)
        enddo
        do i=1,np
          dbres(i)=20.*alog10(result(i)/totmax)        ! Result of FFT in
          dbres(i)=dbres(i)+dni                        ! dB + ref. level
          frq=(i-1)*deltaf
          zbr(i)=13.*atan(.76*frq/1000.) +             ! Table of z-values
     @           3.5*atan((frq/(7.5*1000.))**2)
        enddo
        call RELMAX(totmax,arrmax,relamp,dbmax,dni)    ! Relative maxima
        call MASK(np,zbr,dbres,dbmax,dmask,dlex)       ! Masking
        call SHIFT(dmask,dbmax,dlex,pshift)            ! Pitch shift
        call WEIGHT(dlex,ws)                           ! Component weights
c
c       **** OUTPUT ****
c
        iamax=inint(arrmax(1))                         ! Number of local ma
        trec=(256/sf)*1000                             ! Time 1 record
        venster=iar*trec
        write(io,'(//)')
        write(io,'(x,''Analyse      :      '',15a2)')naln
        write(io,'(x,''Block        :      '',i3,'' -'',i3)')ilr,ilr+iar-1
        write(io,'(x,''Time length :'',f9.1,'' ms.'')')venster
        write(io,'(x,''Time         :'',f9.1,'' -'',f9.1,'' ms.'')')
     @           (ilr-1)*trec,(ilr-1)*trec+venster
        write(io,'(/)')
        write(io,'(10x,''***** LOCAL MAXIMA  *****'')')
        write(io,'(/)')
        write(io,'(x,''Freq.  (Hz)'',5x,''Amp.'',6x,
     @           ''Amp.  (dB)'',3x,''Lex  (dB)'')')
        do i=1,iamax
          j=2*i
          if (dbmax(j+1).gt.dni-grdb) then
            write(io,'(x,f7.1,5x,f7.1,5x,f7.1,5x,f7.1)')
     @           arrmax(j),arrmax(j+1),dbmax(j+1),dlex(j+1)
          endif
        enddo
        write(io,'(///)')
        write(io,'(22x,''***** SPECTRAL PITCH PATTERN *****'')')
```

```
         write(io,'(/)')
         write(io,'(x,''Nominal Spectral Pitch (Hz)'',
     @               5x,''Pitch shift (Hz)'',
     @               5x,''Spectral Pitch Weight'')')
         do i=1,iamax
           j=2*i
           if (dlex(j+1).gt.0.) then
             write(io,'(8x,f7.1,20x,f7.1,17x,f7.1)')
     @       arrmax(j),pshift(j+1),ws(j+1)
             write(iv,'(3f7.2)')arrmax(j),pshift(j+1),ws(j+1)
           endif
         enddo
         dum=-1.
         write(iv,'(3f7.1)')dum,dum,dum
   40 continue
c
       write(io,'(/)')
       write(io,'(x,''ANALYSIS FINISHED !!!'')')
c
       if (prfl.eq.'y') then
         CLOSE(io,err=110)
       endif
       CLOSE(iv,err=110)
       stop 'OK.'
c
  100 stop '***** ERROR OPENING FILE *****'
  110 stop '***** ERROR CLOSING FILE *****'
c
       END
c
c
**********************************************************************
* Subroutine LNFL reads data from N-file.                            *
**********************************************************************
c
c
       SUBROUTINE LNFL (naln,rin,ibr,iar)
       IMPLICIT INTEGER*2 (I-N)
       DIMENSION naln(1),isampl(2048),rin(1)
       LOGICAL*1 ifout
       DATA nx/2h  /
c
       call FLRED(naln,30,isampl,ibr,iar)
       iap=iar*256
       do i=1,iap
         rin(i)=float(isampl(i))
       enddo
       RETURN
       END
c
c
**********************************************************************
* Subroutine HANNING multiplies array arr with the Hanning-function *
**********************************************************************
c
c
       SUBROUTINE HANNING (iap,arr)
       IMPLICIT INTEGER*2 (I-N)
       DIMENSION arr(1)
c
       pi=acos(-1.)
```

```
      ant=float(iap)
      do i=1,iap
        arr(i)=arr(i)*(1-cos(2*pi*float(i-1)/ant))
      enddo
      RETURN
      END
c
c
*************************************************************************
* Subroutine HAMMING multiplies array arr with the Hamming-function *
*************************************************************************
c
c
      SUBROUTINE HAMMING (iap,arr)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION arr(1)
c
      pi=acos(-1.)
      ant=float(iap)
      do i=1,iap
        arr(i)=arr(i)*(.54-.46*cos(2*pi*float(i-1)/ant))
        arr(i)=arr(i)/.54
      enddo
      RETURN
      END
c
c
*************************************************************************
* Subroutine LOKMAX determines local maxima of array arr.              *
*                                                                      *
* If arr(i) is a local maximum :                                       *
* 1. arr(i)-arr(i-1) >= 0  and arr(i)-arr(i+1) >= 0                    *
* 2. 20.*log10(arr(i)/arr(i+m)) >= 7.dB     m=-3,-2,2,3 .              *
*                                                                      *
* arrmax(1) contains the number of local maxima.                       *
* arrmax(j) contains the index of the (j/2 -1)-th local maximum.       *
* arrmax(j+1) contains the amplitude of the (j/2 -1)-th local          *
* maximum. j is even.                                                  *
*************************************************************************
c
c
      SUBROUTINE LOKMAX(np,arr,arrmax)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION arr(1),arrmax(1)
c
      j=2
      do i=4,np-3
        a=arr(i)-arr(i-1)
        b=arr(i)-arr(i+1)
        c1=20.*alog10(arr(i)/arr(i+2))
        c2=20.*alog10(arr(i)/arr(i-2))
        c3=20.*alog10(arr(i)/arr(i+3))
        c4=20.*alog10(arr(i)/arr(i-3))
        if ((a.ge.0.and.b.ge.0).and.
     @      (c1.ge.7..and.c2.ge.7..and.c3.ge.7..and.c4.ge.7.)) then
          arrmax(j)=i                             ! Index local maximum.
          arrmax(j+1)=arr(i)                      ! Amplitude local maximum.
          j=j+2
        endif
        arrmax(1)=(j/2)-1                         ! Number of local maxima.
      enddo
```

```
      RETURN
      END
c
c
*******************************************************************
* Subroutine PARINT provides more accurate values of the frequen-  *
* cies and amplitudes of the local maxima, by means of a parabolic *
* interpolation.                                                   *
*                                                                  *
* arrmax(1) contains the number of local maxima.                   *
* arrmax(j) contains the frequency of the (j/2 -1)-th local maximum.*
* arrmax(j+1) contains the amplitude of the (j/2 -1)-th local      *
* maximum. j is even.                                              *
*******************************************************************
c
c
      SUBROUTINE PARINT(deltaf,arr,arrmax)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION arr(1),arrmax(1)
c
      iamax=arrmax(1)
      do i=1,iamax
        j=2*i
        index=inint(arrmax(j))
        a=2*arr(index)-arr(index-1)-arr(index+1)
        b=arr(index+1)-arr(index-1)
        arrmax(j)=( (arrmax(j)-1) + b/(2*a) )*deltaf
        arrmax(j+1)=arrmax(j+1) + (b**2)/(8*a)
      enddo
      RETURN
      END
c
c
*******************************************************************
* Subroutine RELMAX determines relative amplitudes of the local    *
* maxima of arrmax.                                                *
* relamp contains the relative local maxima .                      *
* dbmax contains the local maxima in dB.                           *
*******************************************************************
c
c
      SUBROUTINE RELMAX(totmax,arrmax,relamp,dbmax,dni)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION arrmax(1),relamp(1),dbmax(1)
c
      iamax=arrmax(1)
      relamp(1)=arrmax(1)
      dbmax(1)=arrmax(1)
      do i=1,iamax
        j=2*i
        relamp(j)=arrmax(j)
        relamp(j+1)=arrmax(j+1)/totmax
        dbmax(j)=arrmax(j)
        dbmax(j+1)=20*(alog10(arrmax(j+1))-alog10(totmax))
        dbmax(j+1)=dbmax(j+1)+dni
      enddo
      RETURN
      END
c
c
*******************************************************************
```

```
* De subroutine MASK calculates masking-effects according to          *
* TERHARDT.                                                           *
* 1. Masking.                                                         *
* 2. Hearing threshold.                                               *
* 3. Noise                                                            *
***********************************************************************
c
c
      SUBROUTINE MASK(np,zbr,dbres,dbmax,dmask,dlex)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION zbr(1),dbres(1),dbmax(1),dmask(1),dlex(1)
c
      iamax=dbmax(1)
      dmask(1)=dbmax(1)
c
      do i=1,iamax
        j=2*i
        dmask(j)=13.*atan(.76*dbmax(j)/1000.) +
     @            3.5*atan((dbmax(j)/(7.5*1000.))**2)      ! Calculate z
        dmask(j+1)=-24.-(230./dbmax(j))+(.2*dbmax(j+1))    ! Calculate s2
      enddo
c
      do 40 i=1,iamax
c
c     ***** MASKING *****
c
        som=0
        mu=i*2
        if (i.eq.1) goto 25
        do j=1,i-1
          nu=j*2
          dum=dbmax(nu+1)-dmask(nu+1)*(dmask(nu)-dmask(mu))
          amp=10**(dum/20.)
          som=som+amp
        enddo
        if (i.eq.iamax) goto 35
   25   do j=(i+1),iamax
          nu=j*2
          dum=dbmax(nu+1)-27.*(dmask(nu)-dmask(mu))
          amp=10**(dum/20.)
          som=som+amp
        enddo
   35   som=som**2
c
c     ***** HEARING THRESHOLD *****
c
        dremp=3.64*((dbmax(mu)/1000.)**-.8)
     @        -6.5*exp(-.6*((dbmax(mu)/1000.)-3.3)**2)
     @        +.001*((dbmax(mu)/1000.)**4)
        som=som+ 10**(dremp/10.)
c
c     ***** NOISE *****
c
        zm=dmask(mu)
        zl=zm-.5
        zh=zm+.5
        call ZTOF(np,zbr,zm,im)
        call ZTOF(np,zbr,zl,il)
        call ZTOF(np,zbr,zh,ih)
        sn=0.
        do j=il,im-3
```

```
          snl=10.**(dbres(j)/10.)
          sn=sn+snl
        enddo
        do j=im+3,ih
          snl=10.**(dbres(j)/10.)
          sn=sn+snl
        enddo
        som=som+ sn
c
      dlex(1)=iamax
      dlex(mu)=dbmax(mu)
      dlex(mu+1)=dbmax(mu+1)-10.*alog10(som)
c
   40 continue
      RETURN
      END
c
c
      SUBROUTINE ZTOF(np,zbr,z,index)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION zbr(1)
c
      dmin=.05
      do 10 i=1,np
        afst=abs(zbr(i)-z)
        if (afst.lt.dmin) then
          index=i
          dmin=afst
        endif
   10 continue
      END
c
c
*********************************************************************
* De subroutine SHIFT  determines pitch-shifts.                     *
*********************************************************************
c
c
      SUBROUTINE SHIFT(dmask,dbmax,dlex,pshift)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION dmask(1),dbmax(1),dlex(1),pshift(1)
c
      iamax=dlex(1)
      pshift(1)=iamax
      do 40 i=1,iamax
        mu=i*2
        pshift(mu)=dbmax(mu)
        if (dlex(mu+1).gt.0.) then
          if (i.eq.1) goto 25
          dlx1=0.
          do j=1,i-1
            nu=j*2
            dum=dbmax(nu+1)-dmask(nu+1)*(dmask(nu)-dmask(mu))
            amp=10**(dum/20.)
            dlx1=dlx1+amp
          enddo
          dlx1=dbmax(mu+1)-20.*alog10(dlx1)
          if (i.eq.iamax) goto 35
   25     dlx2=0.
          do j=(i+1),iamax
            nu=j*2
```

```
          dum=dbmax(nu+1)-27.*(dmask(nu)-dmask(mu))
          amp=10**(dum/20.)
          dlx2=dlx2+amp
        enddo
        dlx2=dbmax(mu+1)-20.*alog10(dlx2)
35      dum1=.0002*(dbmax(mu+1)-60.)*((dbmax(mu)/1000.)-2.)
        dum2=.015*exp(-1.*dlx1/20.)*(3.-alog(dbmax(mu)/1000.))
        dum3=.03*exp(-1.*dlx2/20.)*(.36+alog(dbmax(mu)/1000.))
        pshift(mu+1)=dbmax(mu)*(dum1+dum2+dum3)
      else
        pshift(mu+1)=0.
      endif
40  continue
    RETURN
    END
c
c
**********************************************************************
* De subroutine WEIGHT accounts for the principal of spectral        *
* dominance.                                                         *
**********************************************************************
c
c
      SUBROUTINE WEIGHT(dlex,ws)
      IMPLICIT INTEGER*2 (I-N)
      DIMENSION dlex(1),ws(1)
c
      iamax=dlex(1)
      ws(1)=iamax
      do i=1,iamax
        mu=2*i
        ws(mu)=dlex(mu)
        if (dlex(mu+1).ge.0.) then
          dum1=1.-exp(-1.*dlex(mu+1)/15.)
          dum2=(dlex(mu)/700.-700./dlex(mu))**2
          dum3=(1.+.07*dum2)**-.5
          ws(mu+1)=dum1*dum3
        else
          ws(mu+1)=0.
        endif
      enddo
      RETURN
      END
```

```
c       TVP.FOR
c
**********************************************************************
* This algorithm calculates the  [illegible]  virtual pitch-pattern   *
* according to Terhardt (1982).                                       *
**********************************************************************
c
c
      PROGRAM TVP
      IMPLICIT integer*2 (i-n)
      DIMENSION spf(16),sps(16),spw(16)
      DIMENSION vpf(16,12),vpw(16,12)
      DIMENSION freq(400),wght(400)
      DIMENSION nav(15)
      CHARACTER*30 nava
      EQUIVALENCE (nav(1),nava(1:1))
      DATA nava(:8)/'DUM2.DAT'/
      DATA nx/2h  /
c
c
      iv=10
   15 print *
      nav(1)='D:'
      n=INAL('Name SP-pattern DATA-file :   D:',28,nav(2))
      nava(n+3:n+7)='.DAT'
      OPEN(unit=iv,file=nava,access='sequential',status='old',err=100)
c
c     **** INPUT SYSTEM PARAMETERS ****
c
      write(6,'(//,x,''Virtual pitch range :'')')
      n=INRE('Lower bound                    (<cr>=400 Hz.) :  ',1,freql)
      if (n.eq.0) freql=400.
      n=INRE('Upper bound                    (<cr>=600 Hz.) :  ',1,freqh)
      if (n.eq.0) freqh=600.
      write(6,'(/)')
      n=ININ('Maximum suharmonic number     (<cr>=12) :  ',1,msh)
      if (n.eq.0) msh=12
      n=INRE('Width coincidence interval (<cr>=0.08) :  ',1,delta)
      if (n.eq.0) delta=0.08
      write(6,'(//)')
c
c     **** INPUT SP-DATA ****
c
    6 i=0
    7 i=i+1
      read(iv,'(3f7.2)')spf(i),sps(i),spw(i)
      if (spf(i).ne.-1.) goto 7
      nsp=i-1
c
c     **** SORT SP-DATA IN TERMS OF DESCENDING WEIGHT ****
c
      do isp=1,nsp-1
        dumw=spw(isp)
        dumf=spf(isp)
        dums=sps(isp)
        iwm=isp
        wmax=0.0
        do i=isp,nsp
          if (spw(i).gt.wmax) then
            wmax=spw(i)
            iwm=i
```

```
          endif
        enddo
        spw(isp)=spw(iwm)
        spf(isp)=spf(iwm)
        sps(isp)=sps(iwm)
        spw(iwm)=dumw
        spf(iwm)=dumf
        sps(iwm)=dums
      enddo
c
c     **** PRINT SORTED SP-DATA ****
c
      write(6,'(x,''      Freq.     Shift    Weight'')')
      do i=1,nsp
        print '(2x,3f9.1)',spf(i),sps(i),spw(i)
      enddo
c
c     **** CALCULATING THE VIRTUAL PITCH-PATTERN ****
c
      do 10 i=1,nsp
        do 20 m=1,msh
          w=0.0
          do 30 j=1,nsp
            n=iint(m*spf(j)/spf(i) + 0.5)
            if (i.eq.j.or.n.eq.0) goto 37
            dum0=(n*spf(i))/(m*spf(j))
            gamma=abs(dum0-1.)
            if (n.gt.20.or.gamma.gt.delta) then
              cij=0.0
            else
              dum1=sqrt(spw(i)*spw(j)/m*n)
              cij=dum1*(1.-gamma/delta)
            endif
            w=w+cij
 37         dummy=dum
 30       continue
          him=spf(i)/m
          beta=1./(1.+ (him/800)**4)
          wim=w*beta
          vpf(i,m)=him
          vpw(i,m)=wim
 20     continue
 10   continue
c
c     **** OUTPUT VIRTUAL PITCH-PATTERN ****
c
      do i=1,nsp
         do m=1,12
           j=(i-1)*12 + m
           freq(j)=vpf(i,m)
           wght(j)=vpw(i,m)
        enddo
      enddo
      mvp=12*nsp
      do ivp=1,mvp-1
        dumf=freq(ivp)
        dumw=wght(ivp)
        wmax=0.0
        do i=ivp,mvp
          if (wght(i).gt.wmax) then
            wmax=wght(i)
```

```
            imv=i
          endif
        enddo
        wght(ivp)=wght(imv)
        wght(imv)=dumw
        freq(ivp)=freq(imv)
        freq(imv)=dumf
      enddo
      print *
      write(6,'(x,''VP-range                          : '',f6.0,'' -'',
     @          f6.0,'' Hz.'')')freql,freqh
      write(6,'(x,''Max. Subharmonic number     : '',i4)')msh
      write(6,'(x,''Width Coincidence interval : '',f8.3)')delta
      print *
      do i=1,mvp
        if (freq(i).ge.freql.and.freq(i).le.freqh) then
          print '(x,''VP-nr., Freqency, Weight : '',i,2f9.1)',
     @            i,freq(i),wght(i)
        endif
      enddo
c
      print *
      stop ' OK.'
  100 stop ' **** FILE OPENING ERROR ****'
      END
```