ACUSTICA
S. HIRZEL VERLAG · STUTTGART

Vol. 17 Heft 5 1966

Repetition Pitch: Monaural Interaction of a Sound with the Repetition of the Same, but Phase Shifted, Sound

by F. A. BILSEN

Department of Technical Physics, Technological University of Delft, Netherlands

Summary

When a sound and the repetition of the same sound after a delay τ are presented together, monaurally, one perceives a "coloration", accompanied by a pitch sensation. This pitch, *RP* (Repetition Pitch), appears to correspond to the reciprocal $1/\tau$ of the delay time τ. However, when all the frequency components of the delayed sound are shifted in phase by 90°, 180° and 270° resp., the *RP* changes to $1.07/\tau$, $0.88/\tau$ (or $1.14/\tau$) and $0.94/\tau$ resp. (In our case the sound was white noise, delayed by a tape recorder.) The same pitches are perceived when listening to a random sequence of pulse pairs (pulse distance τ), where the first or (and) the second pulse is shifted in phase. These results cannot be explained simply by means of autocorrelation or frequency analysis.

Wiederholungs-Tonhöhe: Monaurales Zusammenwirken eines Schallsignals mit demselben, aber phasenverschobenen Signal

Zusammenfassung

Wenn ein Schall und die Wiederholung dieses Schalles nach der Zeit τ zusammen monaural angeboten werden, beobachtet man eine „Färbung", begleitet von einer Tonhöheempfindung. Diese Tonhöhe, *RP* (Repetition Pitch), hat offenbar den reziproken Wert $1/\tau$ der Verzögerungszeit τ. Wenn man aber allen Frequenzkomponenten des verzögerten Schalles eine Phasenverschiebung von 90°, 180° oder 270° gibt, so ändert sich die *RP* in $1{,}07/\tau$, $0{,}88/\tau$ (oder $1{,}14/\tau$), beziehungsweise $0{,}94/\tau$. (Der Schall war in unserem Fall weißes Rauschen, verzögert durch ein Tonbandgerät.) Die gleiche Tonhöhe wird wahrgenommen, wenn das Testsignal eine unperiodische Aufeinanderfolge von Impulspaaren (Impulsabstand τ) ist, wovon der erste oder (und) der zweite Impuls phasenverschoben ist. Diese Resultate können nicht einfach mittels Autokorrelation oder Frequenzanalyse erklärt werden.

Ton de répétition: Interaction monaurale d'un son et de la répétition déphasée de ce son

Sommaire

Quand un son et la répétition du même son après le temps τ sont présentés ensemble, monauralement, on perçoit une «coloration» accompagnée par une sensation de ton. Il se trouve que ce ton, *RP* (Repetition Pitch), a la valeur réciproque $1/\tau$ du temps de retardation τ. Cependant, quand toutes les fréquences du son retardé sont soumises à un changement de phase respectivement de 90°, 180° et 270°, le *RP* devient respectivement $1{,}07/\tau$, $0{,}88/\tau$ (ou $1{,}14/\tau$) et $0{,}94/\tau$. (Le signal était du bruit blanc, retardé par un magnétophone.) On perçoit des tons identiques, quand on écoute une succession non-périodique des couples des impulsions (distance τ entre les deux impulsions), dont la première ou (et) la deuxième impulsion est soumise au changement de phase. Ces résultats ne peuvent pas être expliqués simplement par autocorrelation ou par analyse de fréquence.

1. Introduction

Listening to music or speech one perceives a remarkable change of timbre, when an echo with a delay time of about 30 ms or less is added to the original signal. Several authors have investigated this so called coloration of the original sound. A first impression is that of a reinforcement of a small frequency band of the spectrum of the signal. A trained listener, however, can hear a tone with the timbre of a periodic pulse. Indeed, the coloration appears to be accompanied by a pitch sensation which we propose to call: "*RP*" "Repetition Pitch", as it is perceived only when a signal and the repetition (echo) of the same signal are presented simultaneously to a listener.

ATAL, SCHROEDER and KUTTRUFF have studied the problem of coloration, (not in particular *RP*) in search of an explanation for man's inability to detect the great irregularities in the frequency response of a room [1].

KUHL reported on his experience with small, non-reverberant, recording studios, where the direct sound is often followed by a strong first reflection, which results in a coloration [2].

Bassett and Eastmond obtained a delayed signal in an anechoic room with the aid of a plane reflector. They found that *RP* depends on the interference pattern of the sound field; *RP* appeared to have the value $1/\tau$, where τ is the time interval between the original and the delayed signal at the ear of the listener [3].

Fourcin was the first who made a more detailed study of *RP* with white noise as signal delayed by means of a shift register. He, too, came to the conclusion that *RP* equals $1/\tau$. But, by changing the polarity of the delayed noise *RP* was lowered a little more than a full tone; the new *RP* appeared to be $7/8\,\tau$. He did not, however, propose an explanation for this behaviour of *RP* [4].

In the hope of gaining some insight into the problem of *RP* we undertook a series of listening tests in which the spectral components of the delayed or (and) undelayed signal were shifted in phase by 0°, 90°, 180° and 270°. Two test signals were used. In the first experiment white noise is delayed by τ and shifted in phase, mixed with the undelayed (thus correlated) noise and presented monaurally. In the second experiment a test signal was presented to the subject consisting of a random sequence of pulse pairs (pulse distance τ), where the second or (and) the first pulse eventually was shifted in phase.

Signals, like our latter signal, consisting of pairs of identical pulses (without phase shifting) have been investigated already by Small jr. and McClellan. They found *TSP* (Time Separation Pitch) to correspond to $1/\tau$ [5], [6].

Both our signals yielded identical results, notwithstanding their different natures. This may be explained to a certain extent on considering the mathematical analysis (next Section) of the signals. Their considerable conformity may be understood directly by considering that white noise is a random sequence of (Dirac) pulses; thus white noise with a repetition is a random sequence of pulse pairs.

2. Mathematical description

2.1. *Pulse and frequency response*

The pulse response of a delay system which provides the original signal (at $t=0$) and the repetition of the (same) signal after the delay time τ (at $t=\tau$) (see Fig. 1) is given by

$$h_0(t) = \delta_0(t) + \delta_0(t-\tau) \tag{1}$$

where $\delta_0(t)$ is the Dirac pulse.

The frequency response of this system is

$$H_0(\omega) = 1 + \mathrm{e}^{-\mathrm{j}\omega\tau}. \tag{2}$$

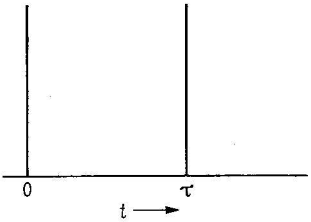

Fig. 1. The pulse response of a delay system.

If the system, moreover, shifts the phase of all the frequency components of $\delta_0(t)$ by an angle φ_1 and those of $\delta_0(t-\tau)$ by an angle φ_2, then the frequency response will be

$$H_\varphi(\omega) = \mathrm{e}^{\mathrm{j}\varphi_1} + \mathrm{e}^{\mathrm{j}\varphi_2}\,\mathrm{e}^{-\mathrm{j}\omega\tau}$$

or:

$$H_\varphi(\omega) = 2\cos\left(\frac{\varphi_1}{2} - \frac{\varphi_2}{2} + \frac{\omega\tau}{2}\right) \times \\ \times \exp\mathrm{j}\left(\frac{\varphi_1}{2} + \frac{\varphi_2}{2} - \frac{\omega\tau}{2}\right) \tag{3}$$

and:

$$\left|H_\varphi(\omega)\right| = 2\left|\cos\left(\frac{\varphi_1}{2} - \frac{\varphi_2}{2} + \frac{\omega\tau}{2}\right)\right|$$

and the pulse response

$$h_\varphi(t) = \cos\varphi_1\,\delta_0(t) + \sin\varphi_1\delta_{90}(t) + \\ + \cos\varphi_2\,\delta_0(t-\tau) + \sin\varphi_2\,\delta_{90}(t-\tau) \tag{4}$$

where $\delta_0(t)$ and $\delta_{90}(t)$ are given by

$$\delta_0(t) = \frac{1}{\pi}\int_0^\infty \cos\omega t\,\mathrm{d}\omega; \tag{5}$$

$$\delta_{90}(t) = -\frac{1}{\pi}\int_0^\infty \sin\omega t\,\mathrm{d}\omega = -\frac{1}{\pi t}. \tag{6}$$

Note that $\delta_{90}(t)$ arises from $\delta_0(t)$ by shifting the phase 90°; it is a hyperbole which starts at $t=-\infty$, becomes infinite at $t=0$ and makes a jump from $+\infty$ to $-\infty$; it dies away at $t=+\infty$.

2.2. *Autocorrelation function*

The autocorrelation function $\Phi(t)$ of the output signal for a Dirac pulse-input of a system with the pulse response $h(t)$ is given by

$$\Phi(t) = \int_{-\infty}^{+\infty} h(t+\vartheta)\,h(\vartheta)\,\mathrm{d}\vartheta. \tag{7}$$

Substituting eq. (4) in eq. (7) we find

$$\Phi_\varphi(t) = 2\cos\varphi_1\,\delta_0(t) + 2\sin\varphi_1\,\delta_{90}(t) + \\ + \cos\varphi_2\,\delta_0(t-\tau) + \sin\varphi_2\,\delta_{90}(t-\tau) + \\ + \cos\varphi_2\,\delta_0(t+\tau) + \sin\varphi_2\,\delta_{90}(t+\tau). \tag{8}$$

On taking into account that: a) the autocorrelation function is the Fourier transform of the power spectrum and b) the power spectrum at the output follows from the power spectrum at the input by multiplication by the (frequency-dependent) transmission factor, we conclude that all input signals having identical power spectra will yield identical power spectra and autocorrelation functions for the output signals.

Now the ensemble averages of the power spectra for white noise and for a series of randomly distributed pulses are identical with the spectrum of the Dirac pulse and thus eq. (8) may be taken to represent the autocorrelation function for the former two test signals.

3. First experiment

3.1. Experimental apparatus

White noise from a noise generator (NG) (General Radio) is shifted in phase by 90° eventually by a so called "Hilbert Transformer" (HT) and delayed by means of a specially constructed tape recorder (delay τ). The undelayed and delayed noise are added and fed through an amplifier to an electrostatic loudspeaker (Quad) in the anechoic room. A block diagram is given in Fig. 2.

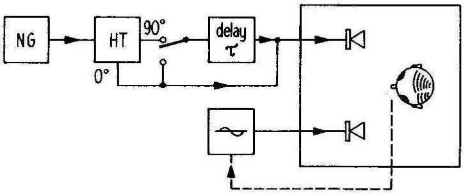

Fig. 2. Block diagram of the first experiment.

3.1.1. The delay mechanism

The tape recorder used in the experiment has three separate channels, each using a separate track on the tape and having a separate head used for both recording and play-back. The heads may be moved along a track. Time intervals between the three channels during playback may thus be introduced by shifting two of the heads from their recording positions (see Fig. 3). The smallest delay is no longer limited by the size of the heads; any delay, even negative and zero, can be obtained. In this experiment only two channels have been used.

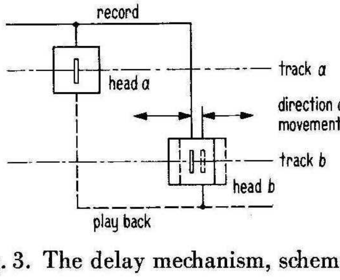

Fig. 3. The delay mechanism, schematic.

3.1.2. The phase shifter

The phase shift of 90° is achieved with help of a "Hilbert Transformer". This consists essentially of a delay line, fed from a travelling wave source, terminated by a negligibly small resistor, the voltage over this resistor being one of the output signals, namely the 0°-signal. The 90°-signal is obtained by integrating the voltage along the line with a weighting function inversely proportional to the distance from the termination.

In the practical circuit (see Fig. 4) the delay line consists of a finite number of segments of constant delay τ_0, the integration being replaced by a summation. The resistors R_n may be adjusted to obtain the flattest possible frequency response.

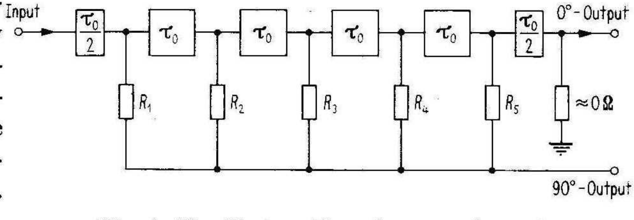

Fig. 4. The Hilbert Transformer, schematic.

The bandwidth of the system used in our experiments is 500 Hz to 7000 Hz within 3 dB, this being the bandwidth of the test signals. A photograph of the pulse response is given in Fig. 5; note that the resemblance with a hyperbole is very clear. For a more detailed description of the "Hilbert Transformer" we refer to [7], for example.

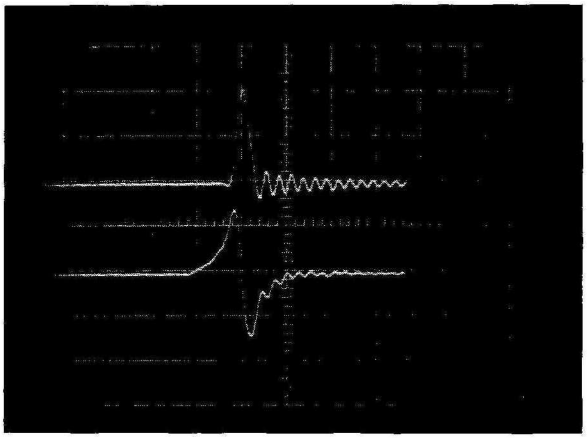

Fig. 5. Pulse response of the Hilbert Transformer. Upper trace: 0°-signal; lower trace: 90°-signal.

3.2. Task of the subject

The subject, who was seated in the anechoic room, heard the test signal at a *SPL* of 30 dB (re $2 \cdot 10^{-5}\,\mathrm{Nm}^{-2}$). He was asked to match the pitch of a pure tone to the perceived *RP* by adjusting the frequency of the pure tone himself. He could vary the level of the pure tone as desired. (It would have been more accurate to match *RP* holding the frequency of the pure tone constant, because one's attention is directed more easily to *RP*. However, this was impossible with the apparatus available).

There were two subjects, both trained in pitch matching.

3.3. *Results*

First, we introduce the symbol $RP^{\varphi_1}_{\varphi_2}$, where φ_1 denotes the phase shift of the undelayed sound (in the second experiment: the first pulse) and φ_2 the phase shift of the delayed sound (in the second experiment: the second pulse).

In Fig. 6 the measured points have been gathered belonging to RP^{0}_{180}; every measured point represents the average of four pitch matchings (two from each subject). Here the very remarkable fact emerges that there are two *RP*'s belonging to the same signal.

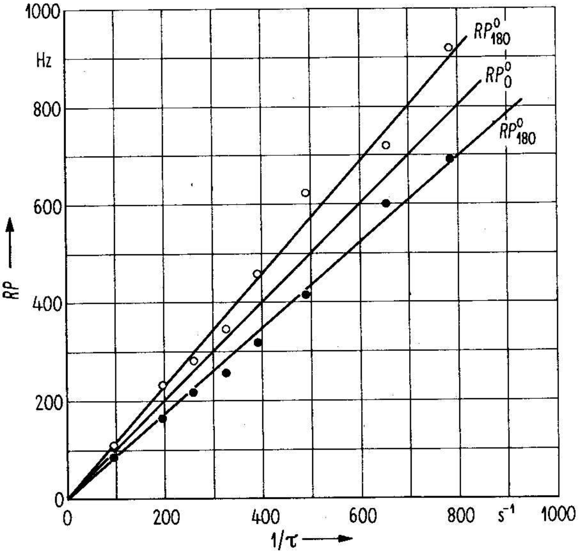

Fig. 6. RP^{0}_{180} as a function of $1/\tau$.

For clearness' sake the measured points of RP^{0}_{0}, RP^{0}_{90} and RP^{0}_{270} have been omitted in this Figure. They show the same trend as the corresponding results of the second experiment.

From about a hundred and fifty pitch matchings we have calculated the ratios $RP^{0}_{180}/RP^{0}_{0} = 0{,}87$ (and 1,14), $RP^{0}_{90}/RP^{0}_{0} = 1{,}08$.

4. Second experiment

4.1. *Experimental apparatus*

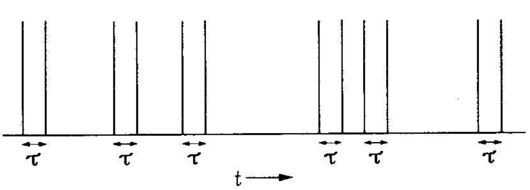

Fig. 7. A random sequence of pulse pairs.

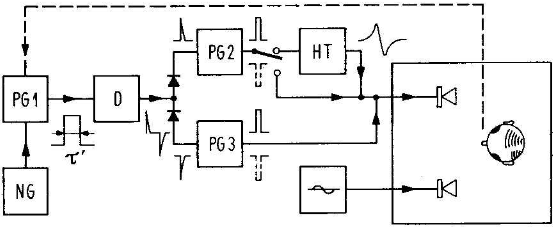

Fig. 8. Block diagram of the second experiment.

For obtaining a random sequence of pulse pairs, as represented in Fig. 7, we made use of three pulse generators (Philips), thus having the facility of shifting the phase of one or both pulses of the pulse pair independently. A block diagram is given in Fig. 8. A (General Radio) noise generator (NG) provides the triggering signal for a pulse generator (PG 1). The pulses from PG 1, thus having a random distribution, are differentiated (D) and fed to the other pulse generators (PG 2 and PG 3). PG 2 reacts only to positive triggering spikes, PG 3 to negative spikes. The pulse from PG 2, phase shifted or not by HT, is added to the pulse of PG 3. Both are fed through an amplifier to the electrostatic loudspeaker in the anechoic room, thus forming the desired pulse pair.

The width τ' of the pulse PG 1 regulates the distance τ between the two pulses of the pulse pair. The width of each of the two pulses is held constant at 20 μs, thus having a frequency spectrum as flat as possible in the audio range.

4.2. *Task of the subjects*

The subject, seated (again) in the anechoic room, heard the test signal at a *SPL* of 30 dB (re $2 \cdot 10^{-5}\,\mathrm{Nm}^{-2}$). He was asked to match *RP* to the pitch of a pure tone by adjusting the width τ' of the pulse PG 1. He could vary the level of the pure tone as desired. The subjects were those of the first experiment.

4.3. *Results*

The results of this experiment are given in Fig. 9. It must be said that, this time, neither of the two subjects made two different pitch matchings for RP^{0}_{180}. *RP* as a function of τ is taken from this Figure; we found: $RP^{0}_{0} = 1/\tau$ (s^{-1}), $RP^{0}_{90} = 1{,}07/\tau$ (s^{-1}), $RP^{0}_{180} = 0{,}88/\tau$ (s^{-1}) and $RP^{0}_{270} = 0{,}94/\tau$ (s^{-1}).

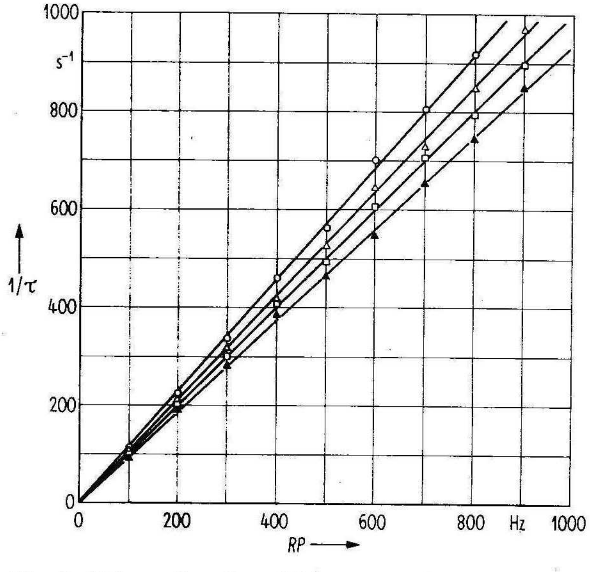

Fig. 9. *RP* as a function of $1/\tau$.

○ RP^0_{180},
□ RP^0_{270},
▲ RP^0_0,
△ RP^0_{90},

When a third pulse is added to the pulse pair one perceives three *RP*'s in accordance with the three pulse distances τ_1, τ_2, τ_3 (see Fig. 10).

5. Discussion

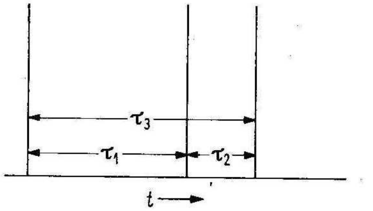

Fig. 10. Configuration of three pulses.

The results of the first and the second experiment agree very well indeed, although the spread in the results is much smaller in the second experiment (at the same number of pitch matchings for each measured point); this, however, is explicable because of the difference in the test procedure.

The results are in very good agreement with the findings of FOURCIN [4] and SMALL JR. and MCCLELLAN [5], [6], too. Compare the relations $1/\tau = f_{\text{subj.}}$ (FOURCIN) and $TSP = 1/\tau$ (SMALL JR. and MCCLELLAN) to our $RP^0_0 = 1/\tau$; our $RP^0_{180} = 0{,}88/\tau$ is very close to $1/\tau = 8\,f_{\text{subj.}}/7$ (FOURCIN).

For clearness' sake we have gathered different configurations of the 0°-, 90°-, 180°- and 270°-pulses together with their frequency (amplitude) spectrum, determined of eq. (3), and their *RP* in one Figure (Fig. 11). From this Figure we conclude that neither the time pattern, nor the frequency spectrum can give a direct explanation for *RP*. The only thing we can derive from this Figure is that as a condition for two configurations to have the same *RP* their amplitude spectra must be the same, though their phase spectra may be different.

All these results have been collected in one simple Figure (Fig. 12). Here *RP* is given as a function of

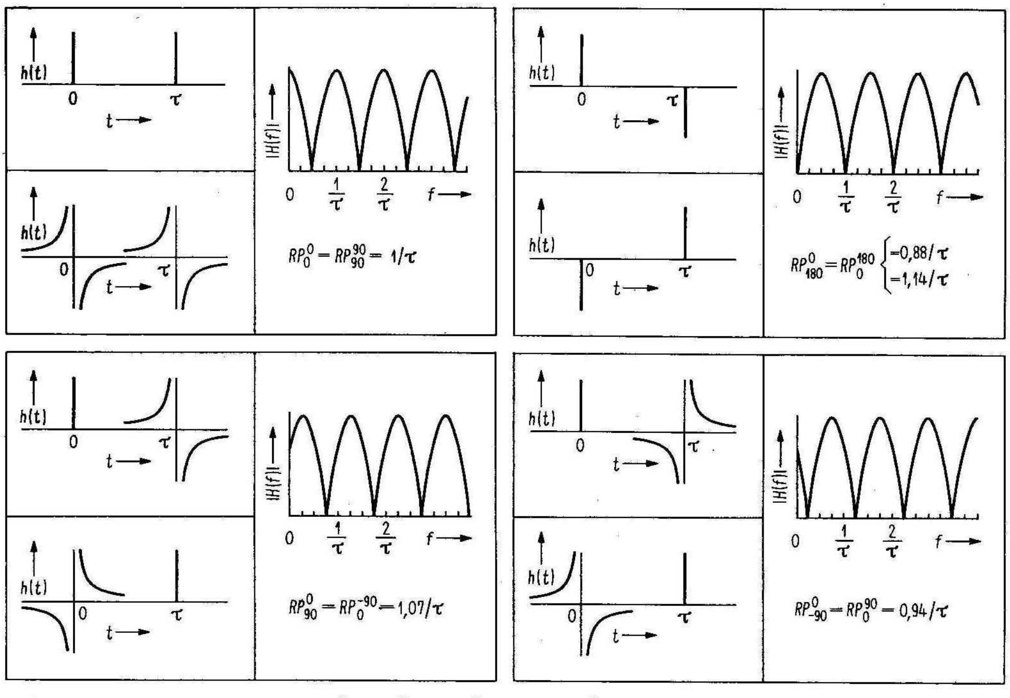

Fig. 11. Different configurations of the 0°-, 90°-, 180°- and 270°-pulse with their spectrum and *RP*.

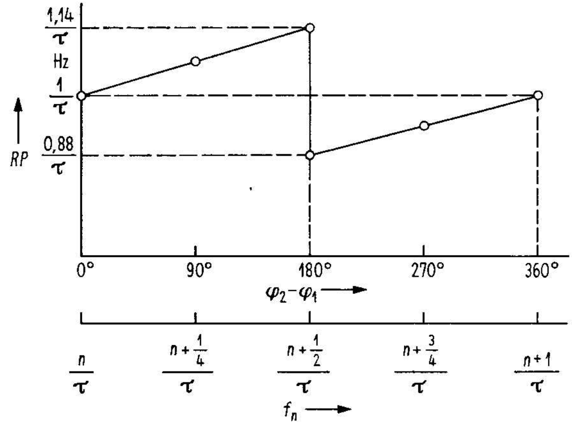

Fig. 12. *RP* as a function of $(\varphi_2-\varphi_1)$, resp. the spectral maxima f_n $(n=1, 2, 3, \ldots)$.

the phase shift difference between the repetition and the original signal: $\varphi_2-\varphi_1$, $1/\tau$ being the parameter; and, moreover, as a function of the spectral maxima f_n, where f_n symbolizes the frequencies f_0, f_1, f_2, ... being the frequencies belonging to the maxima of the amplitude spectrum.

As we proceed in this Figure from 0° to 180°, the amplitude spectrum moves to the right in the frequency domain and *RP* increases continuously. However, as we proceed from 360° to 180°, the amplitude spectrum moves to the left and *RP* decreases continuously; we could not expect otherwise on taking into account that at 360° and 0° the signals (and thus their amplitude spectra) are, naturally, identical. Thus the *RP*-jump at 180°, although not yet explained, does not appear to us to be so strange after all (we have perceived it, although a preference proved to exist in our experiments for the lower *RP*).

Finally, as a point of discussion we refer to the findings of DE BOER [8] and SCHOUTEN et al. [9]; they determined the residue pitch of an amplitude modulated signal. Their results show a remarkable analogy between the qualitative behaviour of the pitch of their test signal and *RP*. We are inclined not to be surprised at this result, as their and our test signals show a certain spectral conformity, notwithstanding their different natures.

I wish to thank Prof. Dr. Ir. C. W. KOSTEN and Ir. D. W. VAN WULFFTEN PALTHE for reading the manuscript and stimulating interest.

I am indebted to Mr. J. A. W. ROEP for his help during the experiments.

(Received March 28th, 1966.)

References

[1] ATAL, B. S., SCHROEDER, M. R., and KUTTRUFF, K. H., Perception of coloration in filtered gaussian noise; short-time spectral analysis by the ear. Fourth International Congress on Acoustics, Copenhagen 1962, H. 31.

[2] KUHL, W., Das Zusammenwirken von direktem Schall, ersten Reflexionen und Nachhall bei der Hörsamkeit von Räumen und bei Schallaufnahmen. Rundfunktech. Mitt. **9** [1965], 170.

[3] BASSETT, J. G. and EASTMOND, E. J., Echolocation: Measurement of pitch versus distance for sounds reflected from a flat surface. J. Acoust. Soc. Amer. **36** [1964], 911.

[4] FOURCIN, A. J., The pitch of noise with periodic spectral peaks. Cinquième Congrès International d'Acoustique, Lüttich 1965, B 42.

[5] SMALL JR., A. M. and MCCLELLAN, M. E., Pitch associated with time delay between two pulse trains. J. Acoust. Soc. Amer. **35** [1963], 1246.

[6] MCCLELLAN, M. E. and SMALL JR., A. M., Pitch perception of randomly triggered pulse pairs. J. Acoust. Soc. Amer. **35** [1963], 1881.

[7] VON URFF, C. A. and ZONIS, F. I., The square-law single-sideband system. I.R.E. Trans. Comm. Syst. **10** [1962], 257.

[8] DE BOER, E., On the residue in hearing. Diss. Amsterdam, 1956.

[9] SCHOUTEN, J. F., RITSMA, R. J., and LOPES CARDOZO, B., Pitch of the residue. J. Acoust. Soc. Amer. **34** [1962], 1418.