ACUSTICA
S. HIRZEL VERLAG · STUTTGART

Vol. 22 1969/70 Heft 2

Repetition Pitch and Its Implication for Hearing Theory

by F. A. Bilsen

Applied Physics Department, Delft University of Technology, the Netherlands

and R. J. Ritsma

Instituut voor Perceptie Onderzoek, Eindhoven, the Netherlands

Summary

When a sound and the (delayed) repetition of that sound are listened to, monaurally, a subjective tone is perceived with a pitch corresponding to the reciprocal value of the delay time. This pitch, Repetition Pitch, undergoes a relative change, independent of the delay time, when all the frequency components of the delayed sound are shifted in phase by a constant amount. Evidence is presented underlining the concept that Repetition Pitch is the result of a combined frequency- and time-analysis in the hearing organ. Thereby the perceived pitch appears to correspond to the reciprocal value of the time interval between two prominent positive peaks in the temporal fine structure of the displacement waveform, evoked by the signal, at a dominant frequency region on the basilar membrane. For each pitch value, this dominant region appears to be situated around a frequency of about four times the frequency value of the pitch. Finally, it is concluded that the perception of Repetition Pitch and identical pitch effects, known in literature as "Reflection Tone", "Sweep Pitch", "Time Difference Tone", "Time Separation Pitch", etc., is based on the same principles as the perception of the pitch of complex periodic sounds, known in literature as "Residue Pitch" and "Periodicity Pitch".

Hauteur du Son de Répétition et Son Implication Dans la Théorie de l'Audition

Sommaire

Quand on écoute un son et la répétition (retardée) de ce son, on perçoit un ton subjectif de hauteur correspondant à la valeur réciproque du temps de retard. Cette hauteur, hauteur du son de répétition, supporte un changement relatif, indépendant du temps de retard, quand toutes les composantes de la fréquence du son retardé sont modifiées en phase par une montée constante. Il est évident que la hauteur du son de répétition est le résultat d'une analyse de fréquence et de temps combinés dans l'organe d'audition. Par ce moyen la hauteur du son perçu semble correspondre à la valeur réciproque de l'intervalle de temps entre deux pointes positives saillantes dans la structure délicate de la forme d'onde de déplacement, évoquée par le signal, à une région de fréquence dominante sur la membrane basilaire. Pour chaque valeur de hauteur du son cette région dominante semble se situer autour d'une fréquence d'environ quatre fois la valeur de fréquence de la hauteur du son. Finalement on conclut que la perception de la hauteur du son de répétition et des effets de hauteur de son identiques connus sous le nom de «Ton de réflexion», «Hauteur du son de balayage», «Ton de différence de temps», «Hauteur de la séparation de temps» etc.... est basée sur les mêmes principes que la perception de la hauteur de sons périodiques complexes connus sous le nom de «Hauteur de son résiduelle» et «Hauteur de son de périodicité».

Die Wiederholtonhöhe und ihre Bedeutung für die Hörtheorie

Zusammenfassung

Beim einohrigen Hören eines Schallsignals und seiner (verzögerten) Wiederholung wird ein Ton wahrgenommen, dessen Höhe dem Kehrwert der Verzögerungszeit entspricht. Diese Wiederholtonhöhe erfährt von der Verzögerungszeit unabhängige relative Veränderungen, sobald alle Frequenzkomponenten des verzögerten Signals um einen konstanten Betrag phasenverschoben werden. Es werden Ergebnisse gezeigt, die für die Auffassung sprechen, daß die Wiederholtonhöhe das Ergebnis einer kombinierten Frequenz- und Zeitanalyse im Gehör ist. Dabei scheint die empfundene Tonhöhe dem Kehrwert des Zeitintervalls zwischen zwei hervortretenden positiven Spitzen in der zeitlichen Feinstruktur der Auslenkung zu entsprechen, die das Signal in einem vorherrschenden Frequenzbereich auf der Basilarmembran hervorruft. Für jede Tonhöhe scheint dieser vorherrschende Frequenzbereich bei einer Frequenz von ungefähr dem vierfachen des Frequenzwertes des Tones zu liegen. Schließlich ergibt sich der Schluß, daß die Empfindung der Wiederholtonhöhe und damit identischer Tonhöheneffekte, bekannt als „Reflection Tone", „Sweep Pitch", „Time Difference Tone", „Time Separation Pitch" und so weiter auf denselben Grundlagen beruht wie die Wahrnehmung der Höhe von Tongemischen, in der Literatur als „Residue Pitch" und „Periodicity Pitch" bezeichnet.

1. Introduction

In a previous paper (Bilsen [1]) one of the authors described the subjective pitch effect due to the interaction of a sound with its repetition, called Repetition Pitch (RP). When a sound (speech, music, noise or merely a pulse) and the repetition of that sound after a delay time τ are added and presented to a listener, he may perceive a coloration and even a pitch. The delay and addition may be effectuated naturally (for example: in roomacoustics) or quite artificially (for example: in psychoacoustics).

To the notion "repetition" we must add the following restrictions. Roughly speaking, only those repetitions which fuse with the original sound to a single sound impression, in general, will be able to evoke a marked coloration and a pitch sensation. Here, it will be very enlightening to think in terms of roomacoustics. There, one distinguishes between reflections and echoes of a sound, the echo being defined as a late reflection which can be heard as a separate entity, as a distinct reiteration of the original sound. The early reflections, however, are not perceived separately but are integrated with the direct (original) sound and perceived as a whole. For speech these are the reflections within about 50 ms after the direct sound. The interval from 0 to 50 ms is also the range in which a marked coloration and even a pitch sensation may be perceived.

Further, we must realize that the notion "repetition" is meant to indicate that the original and repeated sound must be sufficiently correlated. Highest possible correlation is obtained when both original and repeated sound are transmitted through the same acoustic channel. This manner of reproduction, of course, includes monaural presentation to the listeners. Moreover, only monaural presentation is of importance. Binaural presentation, for example: the original sound to the left ear and the repetition to the right ear, sometimes may also give rise to a pitch sensation which is, however, much less intense than in the monaural case. In all our experiments we used the monaural presentation, whilst the original and repeated sound always had the same power spectrum.

In Section 2, the historical review, we shall give a summing up of the experiences and experiments on this subject. In Section 3, the general features of the pitch effect in study will be dealt with, while a part of the experiments summarized in Section 3 will be repeated in Section 5 for narrow-band signals. In Sections 4 and 6 an explanation of the pitch behaviour will be given and in Section 7 the main conclusions will be drawn.

2. Historical review

2.1. Experiments in applied acoustics

In 1693 Huygens, in a letter to de la Hire [2] describes the following observation which he made at the castle at Chantilly de la Cour in France. Standing at the foot of a big stony staircase leading to the garden he noticed that the noisy sound coming from a fountain near the staircase was producing a certain pitch. He concluded that the pitch was caused by the periodic reflections of the sound of the fountain against the steps of the staircase (see Fig. 1). We cite here the particular part of his letter:

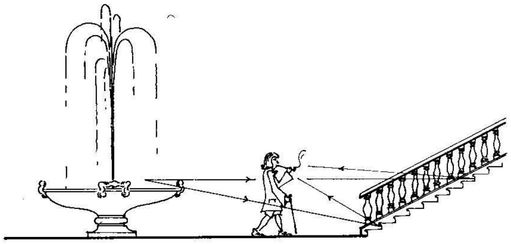

Fig. 1. The pitch due to the periodic reflections of the noise of the fountain against the steps of the staircase (Huygens).

«Je veux ajouter ici au sujet de la réflexion du son une observation assez singulière, que j'ai fait autrefois étant à la belle maison de Chantilly de la Cour où est la statue Equestre on descend avec un degré large de ...... marches dans le parterre où il y a une fontaine de celles qu'on appelle gerbe d'eau, qui fait un bruit continuel. Quand on est descendu en bas et qu'on se tient entre le degré et la fontaine, on entend du coté du degré une résonance qui a un certain ton de musique qui dure continuellement, tant que la gerbe jette de l'eau. On ne savait pas d'où venait ce son ou en découvrait des causes peu vraisemblables ce qui me donna envie d'en chercher une meilleure. Je trouvai bientôt qu'il procédait de la réflexion du bruit de la fontaine contre les pierres du degré. Car, comme tout son, ou plutôt bruit, réitéré à des intervalles égaux et très petits, fait un ton de musique et que la longueur d'un tuyau d'orgue détermine le ton qu'il a par sa longueur, parce que les battements de l'air arrivent également dans les petits intervalles de temps que ses ondoiements emploient à faire deux fois la longueur du tuyau à savoir quand il est fermé par le bout, ainsi je concevois que chaque bruit tant soit peu distingué que venait de la fontaine, étant réfléchi contre les marches du degré, devait arriver à l'oreille de chacune d'autant plus tard qu'elle soit plus éloignée, et celà par des différences de temps justement égales à celle que les ondoiements de l'air emploient à aller et venir autant qu'était la largeur d'une marche. Ayant mesuré cette largeur qui est de 17 pouces, je fis un rouleau de papier qui avait cette longueur, et je trouvai qu'il avait le même ton qu'on entendait au bas du degré. Je trouvai, comme j'ai dit, que la gerbe n'allant point l'on cessait d'entendre ce ton. Et ayant eu l'occasion d'aller à Chantilly pendant l'hiver, qu'il était tombé beaucoup de neige qui ôtait la forme aux marches, je remarquai qu'on entendait rien quoique la gerbe allât et fît du bruit à l'ordinaire.»

From the physical point of view HUYGENS gives an adequate explanation of this pitch effect. Rightly, he points to the analogy between the system of the organ pipe and the system "fountain-staircase". Both produce a periodic signal. Assuming that each periodic signal produces a pitch corresponding to the inverse of the period, he solved the physical side of the problem.

However, the problem becomes interesting when the number of reflections is diminished and at last only one repetition follows a sound event. In that case no periodic signal is formed, but a pitch is still present. In literature, the observation of a pitch due to one repetition has quite often been reported. MINNAERT [3] describes how a steam locomotive, halting at the platform of a station and blowing off steam, produces a hissing sound in which a certain tone predominates (see Fig. 2). This tone is discerned most easily when one approaches the locomotive; then one can perceive how the pitch gets lower and lower. MINNAERT concludes that the effect may be compared with the "mirror-experiment of FRESNEL" in which the pitches correspond to the coloured lines of interference.

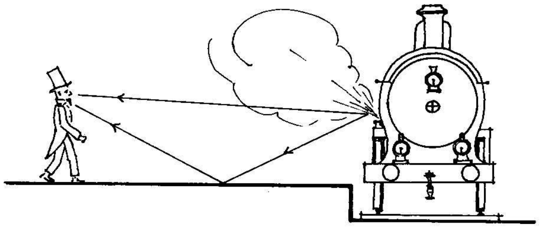

Fig. 2. The pitch due to the interference of the hissing sound from the locomotive with the reflection from the platform (MINNAERT).

Keen observers like BAUMGARTEN [4], PFAUNDLER [5] and HERMANN [6] heard identical effects in the murmur of waterfalls, flood-control dams and all the cases (reflecting) walls were in the neighbourhood of the "noise"-source. HERMANN [6], investigating both the "Reflection Tone" and the Interruption Tone being the low pitch due to the periodic interruption of a high pure tone, looked for an explanation of both pitch effects based on time separation. In Fig. 3 one of his experimental set-ups has been outlined.

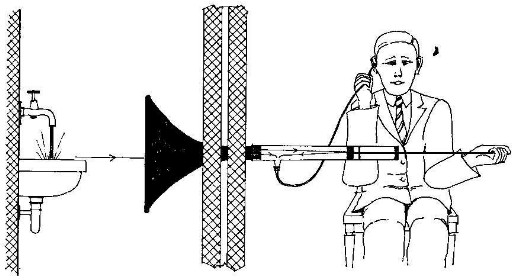

Fig. 3. The pitch due to the interaction of the noise from the running water with its reflection against the cork at the end of the tube (HERMANN).

A series of experiments performed at Cornell University by SUPA, COTZIN and DALLENBACH [7] showed that blind men use reflected sound to locate objects. It appeared that an observer approaching a flat, sound-reflecting surface in the presence of certain sounds, perceives a "broad tone" with an associated pitch which seems to be superimposed on the sound. The pitch varies inversely with distance from the surface.

COTZIN and DALLENBACH [8] repeated the experiments in a more detailed manner: subjects listened with earphones to the sound from a moving loudspeaker, reflected from a wall, and picked up by a microphone moving on the same carriage with the loudspeaker. When the sound was thermal noise, the pitch was quite clear and a sort of siren effect was noticed as the carriage drew nearer to the wall and back again.

BASSETT and EASTMOND [9] carried out the following experiment. They used an experimental set-up very similar to that of COTZIN and DALLENBACH. In an anechoic room a loudspeaker, producing white noise, was placed opposite to a flat reflector. A microphone could be moved to and fro between the loudspeaker and the reflector. The signal from the microphone was analysed by a sound spectrograph. At a certain microphone position a pattern was present which was due to the interference between incident and reflected waves, causing a cancellation of certain frequencies and an augmentation of others. The frequency interval between the maxima of this pattern appeared to correspond to the reciprocal value of the delay time between the incident and the reflected wave. The perceived pitch also corresponded to this value. From that correspondence BASSETT and EASTMOND concluded that the subjective tone is generated in the ear as a difference tone (see Fletcher [10]) corresponding to the frequency spacing between the frequencies thus augmented by the interference.

Bassett and Eastmond experienced the pitch effect under more or less ideal test circumstances. However, also under less ideal circumstances, for example in a concert hall or studio with many other reflections, a coloration and even a pitch may be perceivable due to a strong early reflection (Kuhl [11], Somerville et al. [12], Klimenko [13], and Müller [14]).

Experiments on the critical relative level of the repetition with respect to the original sound for a coloration to be perceivable, were carried out by Somerville et al. [12], Klimenko [13], Atal, Schroeder and Kuttruff [15], and Bilsen [16].

2.2. *Experiments in psychoacoustics*

The observations made by Baumgarten [4] and Pfaundler [5] inspired Pfaundler [5] and Hermann [6] to carry out further investigations with help of a special kind of siren, consisting of a rotating disc with a number of narrow holes situated in a circle and separated from each other by irregular distances. The disc was blown by two small pipes conducting compressed air. Changing the position of one pipe with respect to the other, a gliding pitch could be perceived. Hermann points out that this pitch cannot be the result of frequency analysis as promoted by Helmholtz's resonance theory:

„Zwei Impulse von gegebenem Zeitintervall reichen also, wenn nur der Vorgang sich recht häufig in beliebigen, womöglich unregelmäßig wechselnden Intervallen wiederholt, ganz sicher hin, die dem Intervall des Impulspaares entsprechende Tonempfindung hervorzubringen. Ich sehe absolut keine Möglichkeit, diese Tonempfindung darauf zurückzuführen, daß ein Resonator von der Periode des Paarintervalls angesprochen wird." (See Fig. 4.)

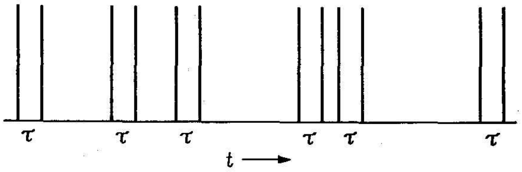

Fig. 4. Pitch due to a stochastic series of pulse pairs with inter-pulse time τ (Hermann).

Later, Thurlow and Small [17] reported an identical experiment done with electronic means. Shifting a sequence of band-pass filtered pulses (a, b, ...) continuously in time with respect to a second identical sequence (a′, b′, ...) a "Sweep Pitch" was perceived that, at any given instant, appeared to be related to the time interval a-a′ or a′-b, depending on which was shorter.

In addition, Thurlow [18] found that when altering the polarity of one of the sequences of pulses a pitch jump upwards or downwards was perceived. The antiphasic condition, thus, has two pitches (ambiguity), one a little higher, the other a little lower than for the cophasic condition.

Nordmark [19] showed that the time compensation required to offset this pitch shift could be compared with the time compensation required to offset the lateralization shift of the central sound image caused by antiphasic pulses at the left and right ear (Flanagan [20]). The similarity between the pitch and lateralization phenomena seemed to support the hypothesis that the sweep pitch (or "Time Difference Tone", as Nordmark called it) is related to the time pattern of neural firings in the auditory nerve.

An extensive series of experiments dealing with matchings of "Time Separation Pitch" has been reported by Small and McClellan [21], [22], [23]. They used different kinds of signals: a periodic sequence of pulse pairs, a random sequence of pulse pairs, periodic sequences of pairs of different (uncorrelated) or identical (correlated) noise bursts, and single pulse pairs. Because they gave the original sound and the repetition always the same polarity, they could not decide whether spectral cues or timing cues are responsible for the perception of Time Separation Pitch. One thing, however, appeared to be evident, namely that time separation between successive sound events alone is insufficient for the evocation of pitch. Obviously successive sound events must be highly correlated. Further, they concluded that for the test signals listed above the amount of information available per unit time is not particularly important, because the distributions of pitch matchings appeared to be about the same for the different test signals.

The pitch character of the phenomenon was confirmed once more by Kylstra [24] who reported that a periodic sequence of pulse pairs with changing interval between the two pulses of a pulse pair can produce a melody. The statement of Jenkins [25], that what Thurlow and Small [17] reported as a sweep pitch was, in reality, a timbre or quality change, is debatable.

Some years ago, the pitch effect due to continuous noise with its repetition was reported again by Fourcin [26] and Wilson [27]. The delay was obtained by means of a shift register. The pitch appeared to correspond to $1/\tau$, the reciprocal value of the delay time τ. However, when the delayed noise was added with negative sign (Fourcin [26]), the pitch appeared to have the value $7/(8\,\tau)$. Experiments by Bilsen [1] were carried out in such a manner as to provide a deeper insight into these findings.

Explanations of Repetition Pitch (which name, in fact, is a synonym for Reflection Tone, Sweep Pitch, Time Difference Tone, Time Separation Pitch) were presented by Bilsen and Ritsma [28], [29] and by Müller [14]. We will refer to them in further detail in the following Sections.

3. "Place" Pitch or "Time" Pitch?

In the following, qualitative and quantitative experiments are summarized in order to get an idea about the general characteristics of and the important factors in the perception of Repetition Pitch.

1. By filtering out, or masking the lower frequency range of the signal the place of the basilar membrane corresponding to the frequency of the perceived pitch can be made inactive. However, the pitch continues to exist even at the lowest sensation levels (Thurlow and Small [17]). Thus, Repetition Pitch is not due to a frequency component or a distortion product equal in frequency to the pitch perceived; it behaves itself like a "residue" pitch (Schouten [30]).

2. Presenting the original sound to the left (or right) ear and the repetition to the right (or left) ear together with uncorrelated white noise, only an extremely faint pitch effect is evoked, much less intense than in the monaural case (see also Fourcin [26]).

3. No pitch is perceived when the original sound is fed through a high-(or low-)pass filter and the repetition through a low-(or high-)pass filter in the absence of a common passband (Bilsen [16]). Obviously, there must be detection of the original and repeated sound in a common region of the basilar membrane and the peripheral nervous system (see also 2).

4. It is important to note that, in the case of continuous noise with its repetition, Repetition Pitch cannot possibly result from a process of detection of a temporal envelope because this is, essentially, missing (several authors assume the temporal envelope to be the most important pitch clue for "Periodicity" Pitch).

5. By reversal of the polarity of the repetition (addition with negative sign) an upwards as well as a downwards pitch jump (ambiguity) can be observed (Thurlow [18]; Fourcin [26]).

Experiments in which the pitch was determined, accurately by matching to a pure tone, as a function of the phase shift of all the spectral components of the delayed sound were reported by Bilsen [1], [31]. Two types of signals were used, viz. continuous wide-band noise with its repetition after a delay time τ (see Fig. 5), and a random sequence of pulse pairs with a time interval τ between the pulses of each pulse pair (see Fig. 4). Repetition Pitch was determined for a large range of τ-values (1 ms to 10 ms) and for four values of the overall phase shift, viz. 0°, 90°, 180°, and 270°. The experiments delivered the following results:

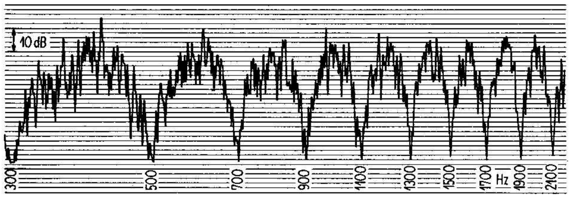

Fig. 5. A section of the spectrum of a noise signal added to its repetition after a delay time of 5 ms.

$$\mathrm{RP}^0_0 = \frac{1.00}{\tau}, \quad \mathrm{RP}^0_{90} = \frac{1.07}{\tau}, \quad \mathrm{RP}^0_{180} = \frac{0.88}{\tau}$$

$$\text{and} \quad = \frac{1.14}{\tau}, \quad \mathrm{RP}^0_{270} = \frac{0.94}{\tau}. \qquad (1)$$

It was discussed that there is no point in looking for a simple explanation in the time domain, because notwithstanding the different values of the overall phase shift the time separation is always equal to τ; the pitches, however, are not equal to $1/\tau$ except for the 0°-case.

Furthermore, Repetition Pitch cannot possibly be explained as being a place pitch corresponding to the centre frequency of the lowest "cosine mountain" in the continuous frequency spectrum proper to this kind of signals, or as being a difference tone which could only have the value $1/\tau$.

4. Three hypotheses

As was pointed out in the foregoing Section, Repetition Pitch cannot be explained, simply, in terms of time or frequency analysis alone. In this section we show that an explanation is possible if one assumes three hypotheses [28], [31]:

(a) A signal with a wide-band frequency spectrum undergoes a spectral analysis on the basilar membrane. Due to the limited resolving power of the membrane, however, the signal is not completely analysed into its frequency components. Therefore, on each place of the basilar membrane a temporal displacement waveform is present which is due to the interaction of several neighbouring frequency components.

(b) The temporal displacement waveform present at each place on the basilar membrane is able to evoke a pitch corresponding to its temporal fine structure by triggering of the neurons ending on a restricted region around that place.

(c) The integral activity originating from one of these regions of the basilar membrane is dominant with respect to the perception of pitch.

ad (a) Spectral analysis

The hypothesis of spectral analysis is confirmed for the greater part by the work of VON BÉKÉSY [32]. He showed that each frequency has a place of maximum stimulation along the basilar membrane. The width and form of the stimulation patterns measured suggest limited frequency resolving power. In agreement with this fact, it appeared from experiments with models of the basilar membrane that for wide-band signals at each place of the basilar membrane a waveform is present which may be considered as being the response of a band-pass filter to the wide-band signal (see for example FLANAGAN [33]). The working of the basilar membrane as a set of band-pass filters had already been suggested by SCHOUTEN [30].

ad (b) Fine-structure detection

The hypothesis of fine-structure detection has been tested extensively for amplitude-modulated sinusoids and frequency-modulated sinusoids (DE BOER [34]; SCHOUTEN, RITSMA, CARDOZO [35]; RITSMA, ENGEL [36]). For these types of signals it has been shown that as a first approximation the time interval between two positive peaks in the fine structure of the acoustical waveform near two successive crests in the time envelope is an adequate parameter of the pitch perceived.

These types of signals with narrow frequency spectrum produce a so called "ready made" envelope at the corresponding place on the basilar membrane and, thus, we may be sure that the displacement waveform at that place is about the same as the acoustical waveform. Further, the neurons ending on that place are able to be triggered by the displacement waveform and to produce a spike pattern exhibiting a remarkable similarity with the unipolar-peak pattern in the temporal fine structure of the displacement waveform (KIANG [37]).

Therefore, the hypothesis of fine-structure detection can be extended to a relation between the frequency of the low-pitch perceived and the spike pattern in the neurons ending on a specific place along the basilar membrane.

ad (c) Spectral dominance

Due to the analysing function of the basilar membrane, signals with a wide-band spectrum will give displacement waveforms at different places along the membrane which differ markedly from each other. It is hypothesized that one of these places is dominant with respect to pitch and, thus, the fine structure at that place (region) determines the pitch of the wide-band signal.

This hypothesis has been tested for periodic white pulse trains (RITSMA [38]) and for white noise with its repetition (RITSMA, BILSEN [39], [31]). It has been found that the temporal information taken at a specific place along the basilar membrane is dominant in the perception of pitch over the temporal information present at other places along the membrane. The most effective spectral area influencing pitch varies more or less with the subject. For pitches in the range of 100 Hz to 400 Hz, the most effective spectral areas cover the region which is about four times the pitch value, e.g. the third, fourth and fifth harmonic in the case of a periodic signal.

5. Experiments

In order to check the hypothesis of fine-structure detection we undertook a series of listening tests with third-octave filtered pulse pairs of which the second pulse was given a phase shift of 90°, 180° or 270°.

5.1. Signal description

With good approximation, the response of the third-octave filter (Brüel and Kjaer 2112) used in the experiments can be described in terms of the response of an ideal band-pass filter with centre frequency f_0 and bandwidth Δf. The response to a DIRAC pulse $\delta_0(t)$ follows from the FOURIER integral by replacing the limits of integration, thus

$$\delta'_0(t) = \frac{1}{\pi}\int\limits_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} \cos \omega t \, d\omega = \frac{2}{\pi t} \cos \omega_0 t \sin \frac{\Delta\omega}{2} t . \tag{2}$$

In the same way we find for the response to a 90°-phase shifted DIRAC pulse

$$\delta'_{90}(t) = -\frac{1}{\pi}\int\limits_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} \sin \omega t \, d\omega = -\frac{2}{\pi t} \sin \omega_0 t \sin \frac{\Delta\omega}{2} t . \tag{3}$$

For $t=0$ the envelope $2\sin(\Delta\omega\, t/2)/\pi t$ is at maximum; thus, in this vicinity the fine structure has its greatest peaks. In particular, $\delta'_0(t)$ has its major positive peak for $t=0$, $\delta'_{90}(t)$ for $t=-1/(4f_0)$, $\delta'_{270}(t)$ for $t=1/(4f_0)$ and $\delta'_{180}(t)$ has two major peaks, one for $t=1/(2f_0)$ and the other for $t=-1/(2f_0)$. The actual response of the filter used in our experiments has been photographed as it appeared on an oscilloscope and is reproduced in Fig. 6. For each value of the phase φ_2 of the second pulse the lower trace represents the input signal to the filter and the upper trace the response of the filter. These responses are in accordance with the eqs. (2) and (3), in particular the relative position of the major positive peaks. The differences between the theoretical and the actual filter lie in the fact that the theoretical filter does not fulfil the causality conditions of physical realizable systems, e.g. the theoretical system responds before an input signal is present.

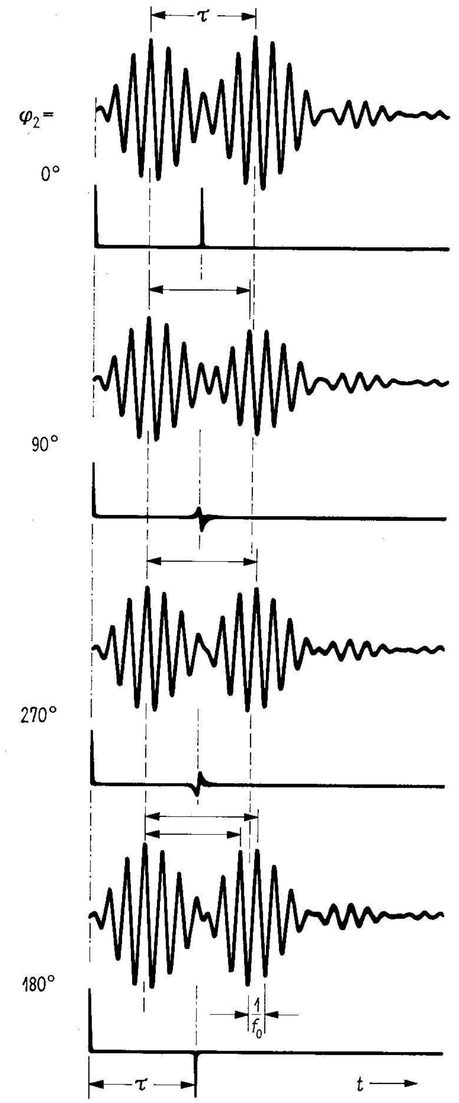

Fig. 6. The response of the third-octave band-pass filter with centre frequency f_0 (each upper trace) to pulse pairs consisting of a 0°-pulse and a 0°-, 90°-, 180°-, or 270°-pulse with a time interval τ between the pulses (each lower trace).

Due to the fact that the vibrations of the basilar membrane are highly damped (see the pulse responses found by FLANAGAN [33]), the very undamped responses (high Q-factor) of the B&K-filter will produce displacement waveforms at the corresponding place on the basilar membrane which are very similar to the B&K-filter output waveform.

5.2. *Theoretical pitch values*

If, due to the interaction of two successive filtered DIRAC pulses (pulse distance τ), Repetition Pitch is indeed correlated with the reciprocal value of the time distance between the major, positive, peaks of the fine structure, the pitch-behaviour for that place on the basilar membrane which corresponds to the frequency f_0, must fulfil the following relations

$$\begin{aligned}
[\mathrm{RP}^0_0]_{f0} &= 1/\tau\,, \\
[\mathrm{RP}^0_{90}]_{f0} &= 1/[\tau-1/(4f_0)]\,, \\
[\mathrm{RP}^0_{180}]_{f0} &= 1/[\tau\pm1/(2f_0)]\,, \\
[\mathrm{RP}^0_{270}]_{f0} &= 1/[\tau+1/(4f_0)]\,.
\end{aligned} \tag{4}$$

These relations are represented in Figs. 8 and 9 by the solid lines.

In fact, every combination of positive peaks in the temporal fine structure is a physical parameter for pitch. Thus also pitch values are to be expected which fulfil equations like eq. (4) with a factor $\pm n/f_0$ added in the denominator.

5.3. *Experimental setup*

A block diagram of the experimental setup is given in Fig. 7. Two double-pulse generators (DPG) are triggered by the pulses of a single-pulse generator (PG) which in turn is triggered externally by a random-noise generator (NG) or internally by a periodic signal. The inter-pulse times (τ_1 and τ_2) can be adjusted separately. Moreover, the second pulse of pulse pair 2 may be shifted at will in

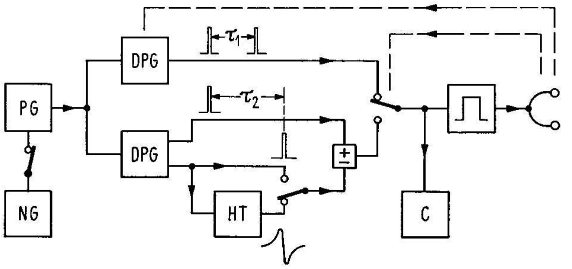

Fig. 7. Block diagram of the experiment

phases by a HILBERT transformer (HT) producing a wide-band-phase shift of 90° (see BILSEN [1]). Thus, at the switch manipulated by the subject (lower dashed line), two signals are available. Both are of the same character:

a) The test signal built up of a periodic or random sequence of pulse pairs consisting of a 0°-pulse and a 90°-, 180°-, or 270°-pulse with inter-pulse time τ_2;

b) the matching signal built up of a periodic or random sequence of 0°-pulse pairs with inter-pulse time τ_1. The period duration amounted to 50 ms (except during the control measurements). Before being fed to the headphones (Sennheiser HD 110) the pulse pairs are filtered by a third-octave band-pass filter (Brüel and Kjaer 2112).

The subject heard the signals at a sensation level of about 40 dB. He was asked to match the pitch of the test signal to the pitch of the matching signal by adjusting the inter-pulse time τ_1 (upper dashed line); he could switch over between test and matching signal as long as he thought this was necessary for a good pitch matching. After the setting was made the experimenter read off the time τ_1 on a digital counter (C). Two subjects participated in the listening tests.

5.4. Experimental results

The results of the listening tests are represented in Figs. 8, 9 and 10. Each measured point is the average of ten pitch matchings, five by each subject. Fig. 8 gives the frequency value of the perceived pitch – RP ($=1/\tau_1$) in Hz – due to a third-octave band-pass filtered 0°- and φ_2-pulse, as a function of the third-octave centre frequency f_0; τ_2 is held constant at 5 ms (thus $\mathrm{RP}_0^0 = 200$ Hz). Fig. 9 represents the results for $\tau_2 = 2$ ms (thus $\mathrm{RP}_0^0 = 500$ Hz).

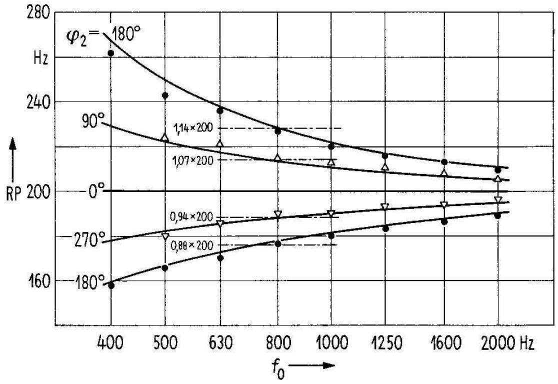

Fig. 8. Repetition Pitch (RP) due to a third-octave filtered 0°- and φ_2-pulse as a function of the third-octave band-pass centre frequency f_0, for a time interval of 5 ms between the pulses.

Moreover, a special series of pitch matchings was carried out in view of the fact that it was always possible to assign more than one pitch to the test signal; sometimes even four different pitches could be perceived and matched with great precision. Fig. 10 confirms the existence of this multivalence; here $f_0 = 2000$ Hz and $\tau_2 = 2$ ms.

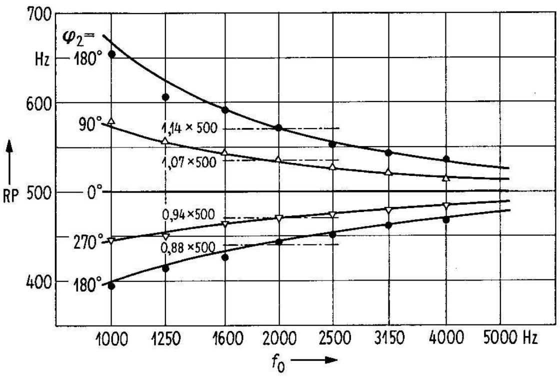

Fig. 9. Repetition Pitch (RP) due to a third-octave filtered 0°- and φ_2-pulse as a function of the third-octave band-pass centre frequency f_0, for a time interval of 2 ms between the pulses.

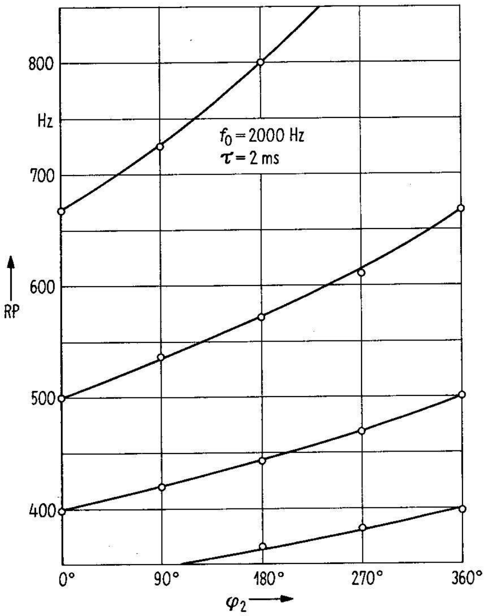

Fig. 10. Repetition Pitch (RP) due to a third-octave filtered 0°- and φ_2-pulse, for a third-octave band-pass centre frequency of 2000 Hz and a time interval of 2 ms between the pulses.

Some control measurements were carried out for the purpose of confirming the fact that the way of repeating the pulse pair, periodically or at random, is not of significant influence on the perceived pitch (see Table I).

Table I.
Values of RP^0_{180} in Hz for $\tau = 5$ ms.

f_0 Hz	Average time interval between pulse pairs		
	1 s	100 ms	50 ms random
630	171	170	171
	236	236	239
1250	183	183	183
	217	216	217

It can be concluded that the measured points fit the theoretical curves very closely. Evidently, the hypothesis of fine-structure detection is confirmed by the experiment.

6. Final explanation

The hypotheses having been proved, we may try to explain the wide-band behaviour of Repetition Pitch, expressed by the relations (1). If, indeed, the wide-band pitch is equal to the pitch belonging to a certain dominant frequency band, we can find this band by the substitution of the relations (1) in eqs. (4). We get, for example, for RP^0_{90}

$$RP^0_{90} = \frac{1.07}{\tau} = [RP^0_{90}]_{f_0\,\mathrm{dom.}} = \frac{1}{[\tau - 1/(4\,f_{0\,\mathrm{dom.}})]}$$

Solving for $f_{0\,\mathrm{dom.}}$ we find

$$f_{0\,\mathrm{dom.}} = 3.8/\tau = 3.8\,RP^0_0 . \qquad (5)$$

Solving in this manner for all values of Repetition Pitch, and thereafter averaging we get the average dominant centre-frequency

$$f_{0\,\mathrm{av.\,dom.}} = (3.9 \pm 0.2)\,RP^0_0 . \qquad (6)$$

From this result we may conclude that there is no question of only one specific spectral region on the basilar membrane dominant for various values of pitch, but rather of one dominant region, the position or centre frequency of which depends on the value of the pitch via a factor of about 4. Thus the dominant spectral region of $RP^0_0 = 200$ Hz lies at about 800 Hz; this is denoted by the dash-dotted lines in Fig. 8. Likewise, the dominant region of $RP^0_0 = 500$ Hz is around 2000 Hz (see Fig. 9).

The value of eq. (6) is in full accordance with the result of experiments on spectral dominance (Ritsma and Bilsen [39]), namely that an octave band around 1000 Hz is most dominant for values of Repetition Pitch of about 250 Hz. The results also completely agree with the results of tests with periodic pulses (Ritsma [38]) where the third, fourth and fifth harmonic proved to be dominant.

Fig. 11 is meant to summarize the whole story; it indicates that Repetition Pitch is the result of a process of fine-structure detection in the temporal displacement waveform present at a (relative positioned) dominant spectral region on the basilar membrane.

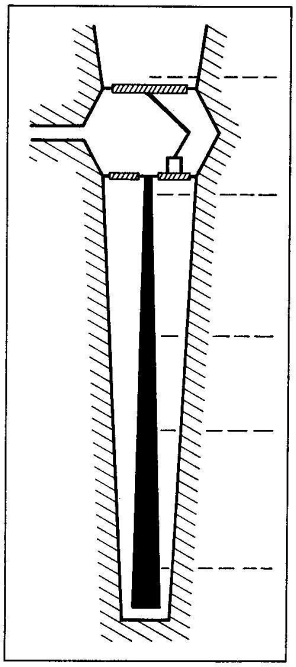

Fig. 11 a

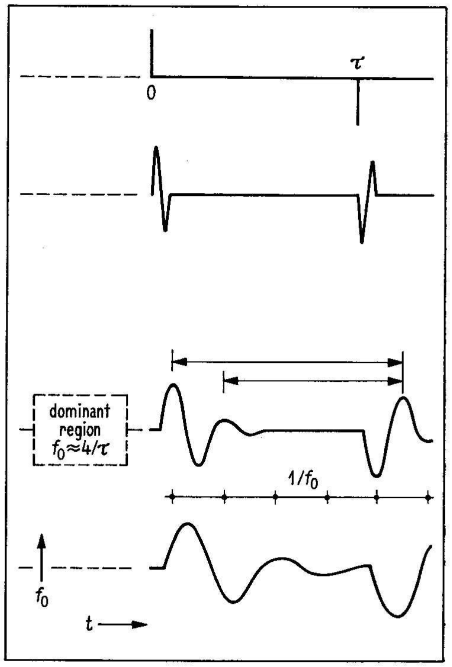

Fig. 11 b

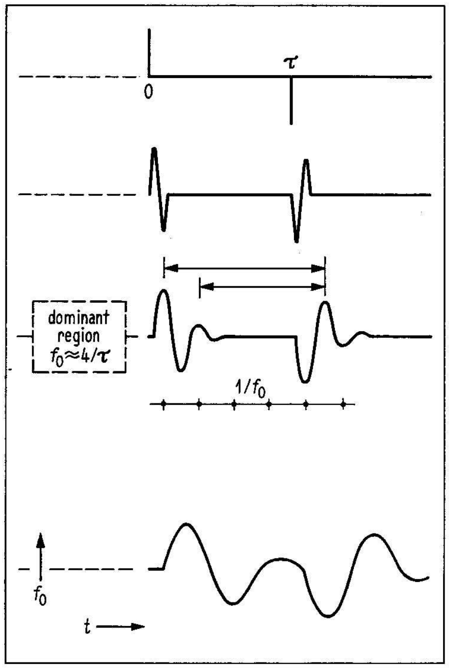

Fig. 11 c

Fig. 11. Repetition Pitch corresponds to the time interval between two prominent positive peaks in the temporal fine structure of the displacement waveform at a dominant frequency region on the basilar membrane.

$$\mathrm{RP}^0_{180} = \frac{1}{\longleftrightarrow} = \frac{1}{(\tau \pm 1/2\, f_0)_{\mathrm{dom.}}} = \frac{0.88}{\tau} \text{ or } \frac{1.14}{\tau} .$$

The explanation given by MÜLLER [14] and based on the build-up of a residue pitch from four hypothetical equidistant spectral components of optimal energy, positioned nearest to the tops of the second to the fifth or the third to the sixth cosine mountain, seems to be debatable. This follows from the fact, mentioned in the foregoing Section, that third-octave filterd pulses, having a spectrum containing only about one cosine mountain in the dominant region, are able to evoke pitch.

7. Conclusions

We dealt with the following phenomenon observed in the free air, in psychoacoustics and applied acoustics: when a sound and its (delayed) repetition are added together and listened to, a subjective tone is evoked with a pitch corresponding to the reciprocal value of the delay time. From the literature we learned that this pitch effect had manifested itself, in the past, in different ways: "Pitch heard in a fountain's noise reflected against the steps of a stair" (HUYGENS [2]), "Reflection Tone" (HERMANN [6]), "Sweep Pitch" (THURLOW and SMALL [17]), "Time Difference Tone" (NORDMARK [19]), "Time Separation Pitch" (SMALL and MC CLELLAN [21], [22], [23]), "Pitch of noise with periodic spectral peaks" (FOURCIN [26]).

An inspection of these pitch effects revealed that all are, in fact, one and the same pitch effect. The only necessary condition for the pitch to occur appeared to be that original and repeated sound are sufficiently correlated; in other words: there must be a repetition in the true sense of the word. Therefore, we called this pitch effect: "Repetition Pitch".

Many efforts, made in the past, to explain this pitch effect in the frequency domain or in the time domain failed. However, we have shown that this pitch effect can be explained qualitatively as well as quantitatively by a combination of the frequency and time principle. The explanation goes as follows: A sound with a wide- band frequency spectrum, interacting with its repetition, undergoes a spectral analysis on the basilar membrane. Due to the limited resolving power of the membrane, on each place of the membrane a temporal displacement waveform is generated which is the result of the interaction of a limited "region" of frequencies. One of these regions of which the wide-band signal is built up, is dominant with respect to the perception of pitch, namely the region lying around a frequency of about four times the frequency value of the pitch perceived. This pitch corresponds to the time interval between two pronounced positive peaks in the temporal fine structure of the displacement waveform belonging to the dominant spectral region.

A comparison with studies on the pitch of periodic signals made clear that the Repetition Pitch-effect is based on the same principles as the residue effect (SCHOUTEN [30]), because there exists a complete similarity as regards fine-structure detection (DE BOER [34]; SCHOUTEN, RITSMA and CARDOZO [35]) as well as spectral dominance (RITSMA [38]).

Acknowledgements

We wish to thank Prof. Dr. C. W. KOSTEN (Delft University of Technology) and Prof. Dr. J. F. SCHOUTEN (Instituut voor Perceptie Onderzoek) for their useful comments during the preparation of the manuscript.

We are indebted to Mr. J. A. W. ROEP (Delft University of Technology) for his help during the experiments.

(Received April 15th, 1969.)

References

[1] Bilsen, F. A., Acustica **17** [1966], 295.
[1] Huygens, C., Oeuvres Complètes. Vol. 10, 570.
[3] Minnaert, M., De natuurkunde van 't vrije veld. Thieme en Cie, Zutphen 1941.
[4] Baumgarten, A., Ber. d. Naturw. Med. Vereins in Innsbruck **7** [1876], 116.
[5] Pfaundler, S., Sitzungsber. d. Wiener Akad. Math. Naturw. **76** [1877], 561.
[6] Hermann, L., Arch. ges. Physiol. **146** [1912], 249.
[7] Supa, M., Cotzin, M., and Dallenbach, K. M., Amer. J. Psychol. **57** [1944], 133.
[8] Cotzin, M., and Dallenbach, K. M., Amer. J. Psychol. **63** [1950], 485.
[9] Bassett, J. G., and Eastmond, E. J., J. Acoust. Soc. Amer. **36** [1964], 911.
[10] Fletcher, H., Phys. Rev. **23** [1924], 427.
[11] Kuhl, W., Rundfunktech. Mitt. **9** [1965], 170.
[12] Somerville, T., Gilford, C. L. S., Spring, N. F., and Negus, R. D. M., J. Sound Vib. **3** [1966], 127.
[13] Klimenko, G. K., Soviet Phys.-Acoustics **12** [1966], 98.
[14] Müller, L., Proc. 6th Intern. Congr. Acoust., Tokyo 1968, E 61 – 64.
[15] Atal, B. S., Schroeder, M. R., and Kuttruff, K. H., 4th Intern. Congr. Acoust., Copenhagen 1962, H 31.
[16] Bilsen, F. A., Acustica **19** [1967/68], 27.
[17] Thurlow, W. R., and Small jr., A. M., J. Acoust. Soc. Amer. **27** [1955], 132.
[18] Thurlow, W. R., Amer. J. Psychol. **71** [1958], 448.
[19] Nordmark, J., J. Acoust. Soc. Amer. **35** [1963], 1544.
[20] Flanagan, J. L., Bell System Tech. J. **41** [1962], 959.
[21] Small jr., A. A., and McClellan, M. E., J. Acoust. Soc. Amer. **35** [1963], 1246.
[22] McClellan, M. E., and Small jr., A. M., J. Acoust. Soc. Amer. **40** [1966], 570.
[23] McClellan, M. E., and Small jr., A. M., J. Acoust. Soc. Amer. **41** [1967], 690.
[24] Kylstra, P. H., Proc. Kon. Ned. Akad. W., series C, **67** [1964], 351.
[25] Jenkins, R. A., J. Acoust. Soc. Amer. **33** [1961], 1550.
[26] Fourcin, A. J., 5e Congr. Intern. d'Acoust., Liège 1965, B 42.
[27] Wilson, J. P., Proc. Congr. Les Systèmes Sonars Animaux. Frascati 1966, 89.
[28] Bilsen, F. A., and Ritsma, R. J., Acustica **19** [1967/68], 114.
[29] Ritsma, R. J., Proc. 6th Intern. Congr. Acoust., Tokyo 1968, A 101 – 104.
[30] Schouten, J. F., Five articles on the perception of sound. IPO, Eindhoven 1960.
[31] Bilsen, F. A., On the interaction of a sound with its repetitions. Acad. Thesis, Delft 1968.
[32] von Békésy, G., Phys. Z. **29** [1928], 793.
[33] Flanagan, J. L., J. Acoust. Soc. Amer. **34** [1962], 1370.
[34] deBoer, E., On the residue in hearing. Acad. Thesis, Amsterdam 1956.
[35] Schouten, J. F., Ritsma, R. J., and Cardozo, B. L., J. Acoust. Soc. Amer. **34** [1962], 1418.
[36] Ritsma, R. J., and Engel, F. L., J. Acoust. Soc. Amer. **36** [1964], 1637.
[37] Kiang, N. Y.-S., Discharge patterns of single fibers in the cat's auditory nerve. M. I. T. Press, Cambridge 1965.
[38] Ritsma, R. J., J. Acoust. Soc. Amer. **42** [1967], 191.
[39] Ritsma, R. J., and Bilsen, F. A., To be published in Acustica.