

THE RESIDUE PHENOMENON AND ITS IMPACT ON THE THEORY OF HEARING

J. F. Schouten

The Residue^{1) 2) 3) 4)} is the joint perception of a number of neighbouring Fourier-components of a sound as a single percept of sharp timbre and more or less pronounced low pitch. It occurs when the Fourier-components are too narrowly spaced for them to be resolved and perceived separately. The pitch, if present, is roughly equal to the spacing of the components.

The residue cannot be explained in terms of a difference tone generated within the inner ear by non-linear distortion. First, it also occurs at moderate loudness levels where non-linear distortion can be proved to be negligible. Second, the phenomena observed in mutual beating and masking are all in accordance with the hypothesis that, at these loudness levels, the inner ear behaves as a linear Fourier-analyser, be it with a limited resolving power. Thus a residue consisting of the frequencies 1800, 2000 and 2200 cps, and perceived with a pitch of 200 cps, must be interpreted as being generated at the location of its constituent frequencies on the basilar membrane.

This necessitates the supposition that each area of the basilar membrane may give rise to sensations of widely different pitch, depending upon the pattern of stimulation^{3) 5)}.

This leads to the following model. First a Fourier-analysis along the basilar membrane with a restricted resolving power. Second, the possibility of a sharpening mechanism improving the resolving power⁶⁾. Third, a set of pitch extractors determining the pitch of the stimulated area. Both the sharpening and the pitch extracting mechanisms are probably of neural origin.

A crucial experiment⁷⁾ consists in shifting all components an equal amount along the frequency axis. If the frequencies of the components thus reach values of e.g. 1850, 2050 and 2250 cps, leading of course to a non-harmonic sound, a residue is still clearly heard with a well-defined pitch of about 205 cps. This pitch shift again rules out the explanation in terms of a difference tone. Interpreting the pitch extractor as a time measuring instrument it also rules out the envelope as the relevant parameter in determining pitch, since the frequency of the envelope remains 200 cps irrespective of the frequency shift. The occurrence of the pitch shift is completely in line, however, with the supposition that the actual time measurement is based on the distance in time between the high frequency peaks at or near the maxima of the low frequency envelope⁸⁾. The predicted ambiguity of pitch³⁾ was recently confirmed experimentally⁴⁾.

The residue phenomena do not occur when the constituent frequencies

are higher than 4000 cps. Moreover, for a given average frequency f of the components, the residue loses its tonal character when the spacing between the components is less than about one twentieth of that value⁹⁾. Thus, the pitch extractor involved has a limited range of measurable time-differences ranging between $1/f$ and about $20/f$ secs. The latter value might be related to the length of the delay-line inherently involved in measuring time-differences.

The experimentally established ambiguity of pitch leads to the conclusion that the tonal character of the residue has to be determined within a few periods of the envelope. This again may give a clue for the mechanism of the pitch extractor.

A third clue may be obtained from measurements on the acuity of pitch of the residue, also as a function of its duration⁹⁾.

No theory of pitch is complete unless it accounts for the great similarity in sensation between two tones or residues an octave apart. It should indicate that our subjective scale of pitch cannot be represented by a straight line as is possible with objective frequencies.

In conclusion it may be stated that the property of pitch of the tonal residue, within its frequency bounds, cannot be accounted for except by a pitch extractor, operating in the time domain, which has to be assumed as coming after an analysing mechanism operating in the frequency domain. The limited resolving power in the frequency domain (critical bandwidth, Frequenzgruppen etc), although an indispensable prerequisite for the existence of the residue as a separate percept, has no immediate bearing on the acuity of pitch perception.

In experiments with residues, however, average frequency and periodicity are linked closely together and thus provide meagre clues regarding their relative influence.

In experiments with residues, however, average frequency and periodicity are widely different and may be varied fairly independently, thereby providing us with major clues regarding the operation of the various mechanisms in the frequency and the time domain respectively.

LE PHENOMENE DU RESIDU ET SES CONSEQUENCES SUR LA THEORIE DE L'OUIE

Le résidu^{1) 2) 3) 4)} est la perception commune d'un certain nombre de composantes de Fourier voisines en tant qu'un son unique de timbre aigu et de ton basse plus ou moins marquée. Il se produit lorsque les composantes de Fourier sont trop proches l'une de l'autre pour être résolues et perçues séparément. La hauteur, si elle est présente, est sensiblement égale à l'espacement des composantes.

On ne peut pas expliquer le résidu comme étant un son de différence engendré dans l'oreille interne par une distorsion non-linéaire. Tout d'abord il se produit à des intensités moyennes pour lesquelles la distorsion non-linéaire peut être considérée comme négligeable. Ensuite, les phénomènes du battement et du cache concordent tous avec l'hypothèse selon laquelle, à ces intensités, l'oreille interne se comporte comme un analyseur linéaire

de Fourier, avec une puissance de résolution limitée. Ainsi, un résidu comprenant les fréquences 1800, 2000 et 2200 pps, et perçu à une hauteur de 200 pps doit être interprété comme étant engendré à l'endroit des fréquences de ses composantes sur la membrane basilaire.

Ceci laisse à supposer que chaque zone de la membrane basilaire puisse donner lieu à des sensations de hauteurs très différentes, selon le schéma de stimulation ^{3) 5)}.

Ceci conduit au modèle suivant. D'abord une analyse de Fourier lelong la membrane basilaire, ayant une capacité de résolution restreinte. Ensuite, le cas échéant, un dispositif d'aigus améliorant ladite capacité ⁶⁾. En troisième lieu, un jeu d'extracteurs de hauteurs déterminant la hauteur de la zone stimulée. Ces derniers mécanismes sont probablement d'origine nerveale.

Une expérience cruciale ⁷⁾ consiste à déplacer toutes les composantes une distance égale lelong de l'axe de fréquence. Ainsi, si les nouvelles composantes deviennent p. ex. 1850, 2050 et 2250 pps, introduisant évidemment un son anharmonique, on entend encore clairement un résidu ayant une hauteur nettement marquée d'environ 205 pps. Cette déviation de hauteur élimine l'explication basée sur un son de différence. En interprétant l'extracteur comme un instrument de mesure de temps, on élimine également l'enveloppe en tant que paramètre correspondant pour la détermination de la hauteur, étant donné que la fréquence de l'enveloppe rest de 200 pps., indépendamment de la variation de fréquence. Cependant, on est sur la même voie d'explication qu'en supposant que la mesure de temps réel est réalisée. Ceci est cependant tout à fait conforme à la supposition selon laquelle la mesure de temps réel est effectuée entre la pointe de haute fréquence au maximum, auprès du maximum de l'enveloppe de basse fréquence ⁸⁾. L'ambiguïté de hauteur prévue ³⁾ a été récemment confirmée par expérience ⁴⁾.

On n'observe aucun résidu lorsque les fréquences, constitutives sont supérieures à 4000 pps. En outre, pour une fréquence moyenne donnée f des composantes, le résidu perd son caractère tonal parce que l'espacement devient inférieur à environ un vingtième de f ⁹⁾. Ainsi l'extracteur de hauteur en question a une gamme limitée de différences de temps mesurables, s'étalant entre $1/f$ et environ $20/f$ secondes. Cette dernière valeur pourrait être associée à la longueur de la ligne de retard impliquée dans la mesure des différences de temps.

Les hauteurs ambiguës observées permettent d'arriver à la conclusion selon laquelle le caractère tonal du résidu est établi dans quelques périodes de l'enveloppe. Ceci peut de même mener à un indice pour le mécanisme de l'extracteur de hauteur.

Un troisième indice peut être obtenu à partir des mesures réalisées sur l'acuité de hauteur du résidu, ainsi qu'en fonction de sa durée ¹⁰⁾.

Aucune théorie de hauteur n'est complète à moins que l'on ne tienne compte de la grande similitude de sensation entre deux tons ou résidus distants d'un octave, en indiquant que notre échelle de hauteur subjective ne peut pas être représentée par une droite comme les fréquences objectives.

En conclusion, on peut indiquer que la propriété de hauteur du résidu tonal dans sa région d'existence ne peut pas être prise en considération, sauf par un extracteur de hauteur opérant dans le domaine de temps, à la suite

d'un mécanisme d'analyse fonctionnant dans le domaine de fréquence. La capacité de résolution limitée dans le domaine de fréquence (critical bandwidth, Frequenzgruppen, etc.) bien que déterminant l'existence-même du résidu en tant que perception séparée n'a pas d'influence immédiate sur l'acuité de la perception de hauteur.

Dans des expériences sur les tons purs ou avec des bandes limitées de bruits, la fréquence et la périodicité sont étroitement associées et fournissent ainsi de maigres indices concernant leur influence relative.

Dans des expériences portant sur les résidus cependant, la fréquence moyenne et la périodicité sont très différentes et peuvent varier d'une façon relativement indépendante. Nous fournissons par là même des indices importants quant au fonctionnement des différents mécanismes dans les domaines respectifs de fréquence et de temps.

Prof. Dr J. F. Schouten,

Instituut voor Perceptie Onderzoek,
Eindhoven - the Netherlands.

-
- 1) J. F. Schouten, Proc. Koninklijke Nederlandse Akad. van Wetenschappen, **XLI**, 1086-1093 (1938)
 - 2) J. F. Schouten, Proc. Koninklijke Nederlandse Akad. van Wetenschappen, **XLIII**, 356-365 (1940)
 - 3) J. F. Schouten, Proc. Koninklijke Nederlandse Akad. van Wetenschappen, **XLIII**, 991-999 (1940)
 - 4) J. F. Schouten, R. J. Ritsma, B. Lopes Cardozo, On the Pitch of the Residue, J.A.S.A. in press
 - 5) J. C. R. Licklider, Auditory Frequency Analysis. Proc. 3rd London Symposium on Information Theory, London, 1956
 - 6) N.V. Franssen, Some considerations on the mechanism of directional hearing. Ac. thesis, Delft, 1960
 - 7) J. F. Schouten, Philips Techn. Review **3**, 286-294 (1940)
 - 8) E. de Boer, On the „Residue“ in Hearing, Thesis, Amsterdam 1956
 - 9) To be published by R. J. Ritsma