Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion

Martin Ebeling[a)]
Peter-Cornelius-Conservatory of Music, Mönchengladbach D-55122 Mainz, Germany

(Received 7 December 2007; revised 5 July 2008; accepted 13 July 2008)

A mathematical model is presented here to explain the sensation of consonance and dissonance on the basis of neuronal coding and the properties of a neuronal periodicity detection mechanism. This mathematical model makes use of physiological data from a neuronal model of periodicity analysis in the midbrain, whose operation can be described mathematically by autocorrelation functions with regard to time windows. Musical intervals produce regular firing patterns in the auditory nerve that depend on the vibration ratio of the two tones. The mathematical model makes it possible to define a measure for the degree of these regularities for each vibration ratio. It turns out that this measure value is in line with the degree of tonal fusion as described by Stumpf [Tonpsychologie (Psychology of Tones) (Knuf, Hilversum), reprinted 1965]. This finding makes it probable that tonal fusion is a consequence of certain properties of the neuronal periodicity detection mechanism. Together with strong roughness resulting from interval tones with fundamentals close together or close to the octave, this neuronal mechanism may be regarded as the basis of consonance and dissonance.
 [DOI: 10.1121/1.2968688]

PACS number(s): 43.75.Cd, 43.60.Ek, 43.64.Bt, 43.66.Ba [DD] Pages: 2320–2329

I. INTRODUCTION

A. The consonance theories of von Helmholtz and Stumpf

In the second half of the 19th century, two important theories of consonance were established. Von Helmholtz (1877) proposed a theory based on roughness, an unpleasant sensation produced by rapid sound fluctuations. He argued that roughness between partials of two tones forming a musical interval unconsciously leads to the sensation of dissonance. However, if two partials coincide, they cannot cause roughness. Thus, if many partials coincide, only a slight roughness is evoked, and, conversely, if many partials do not coincide, the interval sounds rough. Von Helmholtz (1877) believed that roughness explains Phythagoras' rule: In the case of simple vibration ratios, many partials coincide and the interval is consonant, but in the case of complex vibration ratios, only a few partials coincide, and the interval sounds dissonant.

In contrast to von Helmholtz's psychoacoustical approach, the German psychologist Stumpf (1930) tried to give a psychologically motivated definition of consonance. He had observed that consonant intervals show a tendency to cohere into a single sound image, which he called an entity ("Einheit," Stumpf, 1890). He termed this phenomenon tonal fusion ("Tonverschmelzung," Stumpf, 1890). Consonant intervals show a stronger tendency to fuse than less consonant or dissonant intervals. Stumpf (1930), however, could not give a psychoacoustical or physiological explanation for this phenomenon, so roughness-based concepts of dissonance and consonance are still currently widely accepted in many models (e.g., Terhardt, 1976, 1977, 1984; Plomp and Levelt, 1965; and Zwicker and Fastl, 1999; see Burns and Ward, 1982). Unfortunately, these models fail to explain consonance and dissonance in musical intervals without any roughness, such as intervals of pure tones or of primaries with only few partials. This gap is sometimes bridged by a learning hypothesis ("Lernhypothese," Terhardt, 1976, 1977; see Burns and Ward, 1982) which states that in the case of pure tones, the perception of consonance and dissonance and of the specific interval quality has been learned beforehand from musical intervals with complex tones. Von Helmholtz (1877) argued that in the case of pure component intervals, roughness is evoked by ear overtones. However, Plomp (1965) demonstrated that "for usual listening levels, the ear's distortion is sufficiently low to avoid audible combination tones" (Plomp, 1965, p. 1123). This contradicts Helmholtz's hypothesis and within all theories based on this hypothesis.

B. Neuronal code and pitch

At the end of the 19th century, scientists had little knowledge about auditory processing in the brain. However, since then, many insights into the neuronal code and processing of auditory signals have been gained. For example, neurophysiologists can demonstrate that pitch is coded in the interspike intervals (ISIs) of neuronal firing patterns in the auditory nerve and brain stem (Rhode, 1995). The inner ear provides a frequency analysis mechanism which transforms incoming sound into a neuronal code (Zhang *et al.*, 2001). Due to the mechanics of the basilar membrane and the frequency selectivity of the hair cells, sound induced pressure waves traveling through the *cochlea* are converted into neural impulses, representing the different resolvable frequency components of the sound (Goldstein and Scrulovic, 1977).

[a)]Electronic mail: mar.ebeling@arcor.de

In the case of a single pure tone, the traveling pressure wave maximally activates hair cells at a certain place on the basilar membrane and causes them to react with a periodic firing pattern. According to the volley principle of Wever (1949), this results in a running spike train in the auditory nerve with a period corresponding to the reciprocal of the frequency of the tone. The period is equal to the distance of two adjacent spikes. As the pitch becomes higher, the distance becomes smaller. One may say that neuronally, pitch corresponds to the period of the tone. The described mechanism shows that the fibers respond to the instantaneous phase of the motion of the basilar membrane. At higher frequencies, above about 4 Hz, this synchrony disappears. (Zwicker and Fastl, 1999). However, those high frequencies are beyond the range of musically relevant pitches.

Langner (2005) assumed "… since all frequency components (harmonics) of a harmonic sound are multiples of its frequency, the period of the fundamental is also encoded in the cochlea in amplitude modulations resulting from superposition of frequency components above the third harmonic. As a consequence, the period of the fundamental is coded temporally in spike intervals in the auditory nerve and analysed by neurons in the auditory brain stem (cochlear nucleus: CN) and midbrain (inferior colliculus: IC)."

The time between neural spikes, the so-called ISI, can be measured either between successive discharges ("first-order ISI") or between both successive and nonsuccessive spikes ("all-order ISI"). Counting all ISIs in a discharge pattern leads to histograms that show the ISI distributions recorded from the auditory nerve. Cariani and Delgutte (1996) have shown that ISI-histograms ("autocorrelograms") computed from all-order ISIs show high peaks for periods corresponding to the pitch. This demonstrated that "the most frequent all-order ISI corresponded to the pitch heard" (Cariani and Delgutte, 1996, p. 1698). This rule of pitch estimation holds for lower frequencies. In cats, Cedolin and Delgutte (2005) found an upper boundary of about 1300 Hz for pitch estimation from pooled ISI distributions and proposed a complementing rate-place profile for frequencies higher than 400 Hz.

As pitch perception is closely related to the perception of musical intervals, it is obvious that the neuronal firing patterns evoked by consonant and dissonant intervals should be studied to understand the phenomenon of consonance. Thus, Tramo *et al.* (2001) analyzed the neural responses to harmonic intervals. They used stimuli of isolated harmonic intervals formed by complex tones. Each of the two complex tones in the musical intervals (minor second, perfect fourth, triton, and perfect fifth) contained the first six harmonics with equal amplitude and equal phase.

They observed that for consonant intervals, the fine timing of auditory nerve fiber responses contains strong representations of harmonically related pitches implied by the interval (e.g., Rameau's fundamental bass) in addition to the pitches of notes actually present in the interval. Moreover, all or most of the partials can be resolved by finely tuned neurons throughout the auditory system. This finding can be compared with the results of psychoacoustic experiments testing the audibility of partials in complex tones. Moore and Glasberg (1986) measured the frequency of difference limens (DLFs) for the individual components in a complex tone. "For a complex tone containing the first 12 harmonics at equal amplitude, the DLFs were small (between 0.2 and 1.0%) for harmonic number up to four, but the DLFs increased rather abruptly around the fifth to seventh harmonic" (Moore and Glasberg, 1986, p. 283).

By contrast, for dissonant intervals, Tramo *et al.* (2001) observed auditory nerve fiber activity that does not contain strong representations of constituent notes or related bass notes. Many partials of the two complex tones are too close together to be resolved. "Consequently, they interfere with one another, cause coarse fluctuations in the firing of peripheral and central auditory neurons, and give rise to the perception of roughness and dissonance" (Tramo *et al.*, 2001, p. 92).

II. BACKGROUND

A. Autocorrelation functions and hearing theories

Tramo *et al.* (2001) determined the ISI distributions embedded in the responses of axons throughout the auditory nerve during stimulation with musical intervals. Comparing these ISI histograms with the computed autocorrelation functions (the primaries consisting of six equally strong harmonics), they found the same periodicity patterns in both the autocorrelation and the ISI distributions.

From a logical point of view, measuring and counting the all-order ISI are an analysis in the time domain equivalent to the computation of an autocorrelation function. The autocorrelation function shows peaks for all periods of a signal. As periods are distances in time, the autocorrelation function has to be regarded as an analysis in the time domain. From the investigations of Cariani and Delgutte (1996), it seemed probable that a neuronal autocorrelation mechanism for the detection of the periods of running spike trains in the auditory system provides the sensation of pitch. It must be pointed out that the autocorrelation function is as powerful a means for sound analysis as the Fourier transform. The (famous) theorem of Wiener–Khnitchine (Wiener, 1930; Hartmann, 2000) states that the autocorrelation function is the Fourier transform of the power spectrum (energy spectral density) (see also Papoulis, 1962, p. 246). As a consequence, the autocorrelation analysis is equivalent to a Fourier analysis of a signal. The Fourier analysis is used for spectral analysis in the frequency domain; the autocorrelation analysis is a periodicity analysis in the time domain. The power spectrum shows all frequencies inherent in the signal but no phase; the autocorrelation function shows all periods inherent in the signal including all subharmonic periods but also no phase information. This is an important feature of the autocorrelation process: Runtime differences between different spike trains in the auditory system are nullified, thus facilitating the highest possible coincidence rate between two correlated spike trains.

Neuronal spike patterns in the auditory system can be represented mathematically by pulse sequences. All information about their periodicities and ISIs can be provided by forming their autocorrelation functions (Papoulis, 1962, p.

249). Thus, the "existence of a central processor capable of analyzing these interval patterns could provide a unified explanation for many different aspects of pitch perception" (Cariani and Delgutte, 1996, p. 1698). Since Licklider (1951), a lot of auditory theories operating in the time domain have presumed an autocorrelation mechanism or a related model to detect the periodicity of the stimuli (for an overview, see Hartmann, 2000; de Cheveigné, 2005). These models have been tested psychoacoustically or with computer simulations (e.g., Meddis and Hewitt, 1991; Patterson *et al.*, 1995), using different stimuli. Psychoacoustical tests are interpreted as evidences for (e.g., Yost *et al.*, 1996) and sometimes against (e.g., Kaernbach and Demany, 1998) an autocorrelation mechanism in the auditory system. Few models are based on physiological data using properties of neuronal circuits in the auditory pathway (Langner, 1983; Hewitt and Meddis, 1994; Voutsas *et al.*, 2005). The present paper refers to properties of Langner's model of periodicity detection in the *inferior colliculus* (IC) (Langner, 2007) The mathematical model in this paper emulates the underlying logical structure of Langner's neuronal periodicity detection model for pitch perception and applies it to the perception of musical intervals. So, Langner's model will be briefly presented in Sec. II B.

B. Langner's neuronal correlator model

Langner (1983) measured the responses of neurons in the *cochlear nucleus* (CN) and IC to amplitude modulated signals and proposed a model that performs a correlation between signal fine-structure and modulation envelope.

As Langner (1983) explained, its processing elements are a trigger, an oscillator, a reducer, and a coincidence neuron, which are supposed to have their counterparts in well-described on-type, chopper neurons, and pauser neurons in the CN and disk cells in the IC. The trigger unit synchronizes the responses of oscillator and reducer cycle to the modulation. While the oscillator responds with short bursts of regular intrinsic oscillations to each modulation period, the reducer generates intervals precisely related to the signal fine structure. By integrating synchronized activity of many nerve fibers, the reducer is able to code frequencies to the upper limit of phase coupling. The coincidence unit is activated by simultaneous inputs from oscillator and reducer and responds best when signal fine structure and envelope are correlated and the envelope period matches the reducer delay. Thereby it responds best to a periodically modulated sound (BMF) and is simultaneously representing a certain frequency and a certain pitch. (Langner, 2005, p. 51)

Three different periods are crucial for coincidence detection: (a) τ_m the period of the envelope; (b) τ_c the period of a carrier frequency: in other words, the fine-structure of the sound; (c) τ_o the period of intrinsic oscillation.

Langner (2007) assumed that the detection of the envelope period yields to the sensation of pitch, whereas the timbre of the sound corresponds to the fine structure of the sound, represented by τ_c.

The intrinsic oscillation provides a time slot of coincidence. The distribution of intervals shows the highest peaks for periods of τ_o=0.8 ms, τ_o=1.2 ms, τ_o=1.6 ms, τ_o=2.0 ms, and τ_o=2.4 ms or generally τ_o=0.8 ms $+k\cdot 0.4$ ms (Langner and Schreiner, 1988, p. 1813). Mathematically, these three periods correlate if there are small integers n and m, so that the "periodicity equation" is valid (see Langner and Schreiner, 1988, p. 1818):

$$m\cdot\tau_m + n\cdot\tau_c + \tau_o = 0. \tag{1}$$

The intrinsic oscillation with period τ_o may be interpreted as a time window for the coincidence detection as intrinsic oscillation raises the coincidence neuron onto an excitation level closely under the threshold. Therefore, one may hypothesize that the shortest possible oscillation interval of 0.8 ms is also the size of the shortest possible coincidence window. At the beginning of stimulation, each coincidence neuron shows a response characteristic of a comb filter which makes the whole bank of coincidence circuits act as an autocorrelator for the modulation frequencies. However, after about 30 ms, inhibition triggered by the onset neuron converts the coincidence neuron to a bandpass filter (Voutsas *et al.*, 2005) so that an unambiguous pitch discrimination is granted.

III. MATHEMATICAL MODEL OF GENERALIZED COINCIDENCE

A. Correlation functions

Applying autocorrelation functions makes it necessary to classify functions (signals) according to their average power that is defined by

$$\overline{f^2}(t) = \lim_{T\to 2T}\frac{1}{2T}\int_{-T}^{T}|f(t)|^2 dt. \tag{2}$$

The proposed model of generalized coincidence makes use only of functions with finite energy, which means that $\overline{f^2}(t)=0$. Nevertheless, it can easily be extended to finite power functions, which have the property that $0<\overline{f^2}(t)<\infty$ (Papoulis, 1962, p. 240).

In the case of functions with finite energy, the correlation functions of two functions $f_1(t)$ and $f_2(t)$ are defined by the following.

(a) Autocorrelation function [see Papoulis, 1962, p. 241 (12–7)]:

$$\rho_i(\tau) = \int_{-\infty}^{\infty} f_i(t) f_i(t+\tau) dt. \tag{3}$$

(b) Cross-correlation functions [see Papoulis, 1962, p. 244 (12–20)]:

$$\rho_{12}(\tau) = \int_{-\infty}^{\infty} f_1(t) f_2(t+\tau) dt, \tag{4}$$

$$\rho_{21}(\tau) = \int_{-\infty}^{\infty} f_2(t) f_1(t+\tau) dt. \tag{5}$$

Substituting $t'=t+\tau \Leftrightarrow t=t'-\tau$ in Eqs. (3) and (4) shows that

$$\rho_i(\tau) = \rho_i(-\tau), \tag{6}$$

$$\rho_{12}(\tau) = \rho_{21}(-\tau). \tag{7}$$

From Eq. (7), it follows that the sum of both cross-correlation functions is even or symmetric:

$$(\rho_{12} + \rho_{21})(\tau) = \rho_{12}(\tau) + \rho_{21}(\tau) = \rho_{21}(-\tau) + \rho_{12}(-\tau)$$
$$= (\rho_{21} + \rho_{12})(-\tau). \tag{8}$$

Let $F_i(\omega)$ be the Fourier transforms of $f_i(t)$. With $\bar{F}_i(\omega)$ as the conjugate complex of $F_i(\omega)$, definitions (3)–(5) are equivalent to

$$\rho_i(\tau) \leftrightarrow F_i(\omega)\bar{F}_i(\omega) =: A^2(\omega), \tag{9}$$

$$\rho_{12}(\tau) \leftrightarrow \bar{F}_1(\omega)F_2(\omega) =: E_{12}(\omega), \tag{10}$$

$$\rho_{21}(\tau) \leftrightarrow \bar{F}_2(\omega)F_1(\omega) =: E_{21}(\omega), \tag{11}$$

with the arrows indicating Fourier transform [see Papoulis, 1962, p. 242 (12–9) and p. 244 (12–17)]. A finite energy function is real if and only if (see Papoulis, 1962, p. 11)

$$\bar{F}_i(\omega) = F_i(-\omega). \tag{12}$$

Obviously, $A^2(\omega)$ is real. As Eq. (12) holds for $E_{12}(\omega)$ and $E_{21}(\omega)$, both cross-correlation functions are also real functions.

Sometimes it is more convenient to calculate the correlation functions from these formulas.

Let $S(t)=f_1(t)+f_2(t)$ be the sum of two functions $f_1(t)$ and $f_2(t)$. Using definition (3) immediately leads to the sum formula of autocorrelation functions:

$$\rho_S(\tau) = \int_{-\infty}^{\infty} S(t)S(t+\tau)dt$$
$$= \int_{-\infty}^{\infty} [f_1(t) + f_2(t)][f_1(t+\tau) + f_2(t+\tau)]dt$$
$$= \int_{-\infty}^{\infty} f_1(t)f_1(t+\tau)dt + \int_{-\infty}^{\infty} f_2(t)f_2(t+\tau)dt$$
$$+ \int_{-\infty}^{\infty} f_1(t)f_2(t+\tau)dt + \int_{-\infty}^{\infty} f_2(t)f_1(t+\tau)dt$$
$$= \rho_1(\tau) + \rho_2(\tau) + \rho_{12}(\tau) + \rho_{21}(\tau), \tag{13}$$

B. Sequence representation of a tone

In the auditory system, pitch is represented by periodic pulse trains which can be mathematically written as a sequence of equally spaced pulses (M positive integer or ∞):

$$x_\mu(t) = \sum_{m=-M}^{M} I_\mu(t - mT). \tag{14}$$

The constant T is the period of the pulse train, and it is the reciprocal of the frequency corresponding to the perceived pitch. The function $I_\mu(t)$ describes the pulse form. A neuronal pulse is built from several neuronal discharges randomly distributed around time mT. Therefore, we demand $I_\mu(t)$ to be a density function, which means

$$\text{(i)} \quad I_\mu(t) \geq 0 \quad \text{for every } t, \tag{15}$$

$$\int_{-\infty}^{\infty} I_\mu(t)dt = 1. \tag{16}$$

Furthermore, the spread of the single discharges is determined by the dimensionless parameter μ describing the "width" of the pulse $I_\mu(t)$. The pulse width is crucial for the model. Later, it is chosen to meet the probability of coincidence of two different pulse sequences, processed in an autocorrelator such as the one described by Langner (1983) (Sec. II B). Taking μ as a real number, $I_\mu(t)$ becomes a family of functions with the generalized limit $\delta(t)$ (Papoulis, 1962, p. 277). Thus, a third property of $I_\mu(t)$ follows:

$$\text{(iii)} \quad \lim_{\mu\to 0} I_\mu(t) = \delta(t). \tag{17}$$

The limit

$$\lim_{\mu\to 0} x_\mu(t) = \sum_{m=-M}^{M} \delta(t - mT) \tag{18}$$

is the idealized case of all neuronal discharges occurring exactly at time mT. Examples for $I_\mu(t)$ may be the rectangular pulse with the width μ, the Gaussian pulse with μ to determine the variance or the half-wave rectified cosine pulse with 2μ to determine the period of the cosine (see Papoulis, 1962, p. 279 and Hartmann, 2000, p.156; compare Fig. 1).

If two pulses fulfill properties (i)–(iii), their cross-correlation functions also fulfill properties (i)–(iii). Considering definition (4), properties (i) and (ii) become obvious for the cross-correlation function. Property (iii) can be proved by applying the definition of the generalized limit: If

$$\chi_{\mu\nu}(\tau) = \int_{-\infty}^{\infty} I_\mu(t)J_\nu(t+\tau)dt \tag{19}$$

is the cross-correlation function of two pulses $I_\mu(t)$ and $J_\nu(t)$, it can be shown that

$$\lim_{\nu\to 0} \lim_{\mu\to 0} \int_{-\infty}^{\infty} \chi_{\mu\nu}(\tau)\phi(\tau)d\tau = \phi(0), \tag{20}$$

for every continuous test function $\phi(t)$. By definition of the generalized limit, this is equivalent to property (iii) (Papoulis, 1962, p. 277).

As by definition the autocorrelation function is a special case of a cross-correlation function, properties (i)–(iii) also hold for the autocorrelation functions of $I_\mu(t)$ and $J_\nu(t)$.

C. Sequence representation of a musical interval

In the mathematical model presented here, the sum of two simultaneously running pulse trains is the mathematical description of the spike train representation of a musical interval in the auditory system. Let $I_\mu(t)$ and $J_\nu(t)$ be the two families of pulse functions with properties (i)–(iii) as above. The two tones of the interval shall be represented by the two sequences, as follows:

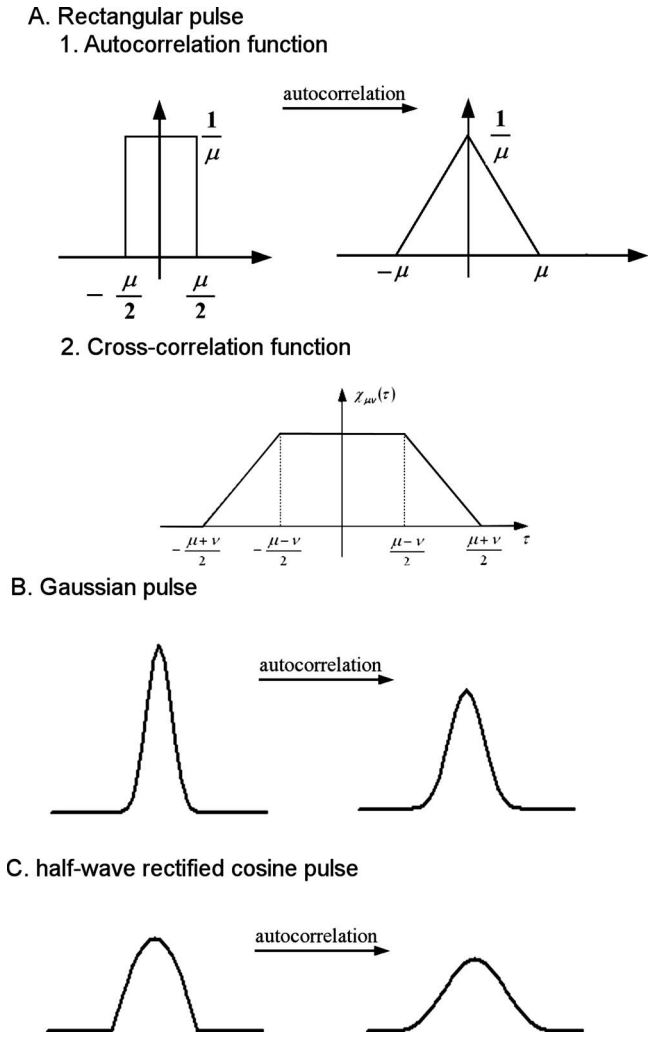

FIG. 1. The correlation functions of tree pulse forms. (A1) The autocorrelation function of a rectangular pulse is a triangular pulse. (A2) shows the cross-correlation function of two rectangular pulses with the different widths μ and ν. (B) The autocorrelation function of a Gaussian pulse is a Gaussian pulse with double the width. (C) shows the autocorrelation function of a half-wave rectified cosine pulse.

$$x_\mu(t) = \sum_{m=-M}^{M} I_\mu(t - mT_1), \tag{21}$$

$$x_\nu(t) = \sum_{n=-N}^{N} J_\nu(t - nT_2). \tag{22}$$

Their sum

$$S(t) = x_\mu(t) + x_\nu(t) \tag{23}$$

is the mathematical representative of the musical interval of this tones. This simple mathematical definition of a musical interval presumes that two neuronal pulse sequences add up without any disturbances and implies complete linearity of the auditory system. Actually, two simultaneously presented tones interfere in the auditory system, which generally does not behave like a linear time invariant (LTI) system. Carefully considering some restrictions, the processing of the auditory system may be approximated by LTI systems (for discussion, see Sec. IV).

Furthermore, let $\alpha_\mu(\tau)$ be the autocorrelation function of the pulse $I_\mu(t)$, $\alpha_\nu(\tau)$ be the autocorrelation function of the pulse $J_\nu(t)$, and $\chi_{\mu\nu}(\tau)$ and $\chi_{\nu\mu}(\tau)$ be the cross-correlation functions of the pulses $I_\mu(t)$ and $J_\nu(t)$. From Eq. (7), it follows

$$\chi_{\mu\nu}(-\tau) = \chi_{\nu\mu}(\tau). \tag{24}$$

If the pulses of both pulse sequences are of the same form and width, then it is valid that $I_\mu(t)=J_\nu(t)$. With Eqs. (3)–(5), it therefore follows that the cross-correlation functions are equal to the autocorrelation function:

$$I_\mu(t) = J_\nu(t) \Rightarrow \chi_{\mu\nu}(\tau) = \chi_{\nu\mu}(\tau) = \alpha_\mu(\tau) = \alpha_\nu(\tau). \tag{25}$$

From the definitions (3)–(5) and the sum formula (13) (the linearity of integration, respectively), it follows that the autocorrelation function of $x_\mu(t)$ is the sequence

$$\rho_\mu(\tau) = \sum_{m=-2M}^{2M} (2M + 1 - |m|)\alpha_\mu(\tau - mT_1). \tag{26}$$

The autocorrelation function of $x_\nu(t)$ is the sequence

$$\rho_\nu(\tau) = \sum_{n=-2N}^{2N} (2N + 1 - |n|)\alpha_\nu(\tau - nT_2). \tag{27}$$

The cross-correlation functions of $x_\mu(t)$ and $x_\nu(t)$ are

$$\rho_{\mu\nu}(\tau) = \sum_{m=-M}^{M} \sum_{n=-N}^{N} \chi_{\mu\nu}(\tau - nT_2 + mT_1), \tag{28}$$

$$\rho_{\nu\mu}(\tau) = \sum_{n=-N}^{N} \sum_{m=-M}^{M} \chi_{\nu\mu}(\tau - mT_1 + nT_2). \tag{29}$$

As the parameters m and n are as well positive and negative and as $\chi_{\mu\nu}(-\tau)=\chi_{\nu\mu}(\tau)$ from Eq. (24), both cross-correlation functions are equal:

$$\rho_{\mu\nu}(\tau) = \rho_{\nu\mu}(\tau). \tag{30}$$

The pulse sequences $x_\mu(t)$ and $x_\nu(t)$ are real functions. Their correlation functions are also real functions [see Eqs. (9)–(12)]. The imaginary part of a signal describes its phase composition. Thus, as $\rho_\mu(\tau)$, $\rho_\nu(\tau)$, $\rho_{\mu\nu}(\tau)$, and $\rho_{\nu\mu}(\tau)$ are real, it follows that they have no phase shifts, or as one may say, they are all in phase. We must keep in mind that the correlation functions of pulse sequences are pulse sequences themselves. As the correlation functions are in phase, it is granted that at least one pulse of each sequence coincides with at least one pulse of all other sequences, namely, for $\tau=0$. The occurrence of further coincidences depends exclusively on the vibration ratio s. Thus, to calculate the degree of coincidence, the variable s must be introduced.

D. Autocorrelation function of an interval

As above, let $x_\mu(t)$ and $x_\nu(t)$ be the mathematical descriptions of the two pulse sequences that are assumed to be the neuronal representations of a musical interval with the vibration ratio s. If ν_1 and ν_2 are the frequencies of the two tones constituting the interval, and T_1 and T_2 are the corresponding periods, the vibration ratio of the interval is

$$s = \frac{\nu_2}{\nu_1} = \frac{T_1}{T_2}. \tag{31}$$

As the autocorrelation function also depends on the vibration ratio, s is introduced as a second variable. Applying Eq. (13) to $S(t)=x_\mu(t)+x_\nu(t)$ with regard to Eq. (31), it follows from Eqs. (26)–(30) that

$$\begin{aligned}\rho_S(\tau,s) &= \rho_\mu(\tau) + \rho_\nu(\tau) + \rho_{\mu\nu}(\tau) + \rho_{\nu\mu}(\tau)\\ &= \rho_\mu(\tau) + \rho_\nu(\tau) + 2\rho_{\mu\nu}(\tau)\\ &= \sum_{m=-2M}^{2M} (2M+1-|m|)\alpha_\mu(\tau - mT_1)\\ &\quad + \sum_{n=-2N}^{2N} (2N+1-|n|)\alpha_\nu(\tau - ns^{-1}T_1)\\ &\quad + 2\sum_{m=-M}^{M}\sum_{n=-N}^{N} \chi_{\mu\nu}(\tau - (ns^{-1}-m)T_1)\end{aligned} \tag{32}$$

is the autocorrelation function of the musical interval with the vibration ratio s. If the pulses of both pulse sequences are of the same form and width, then together with Eq. (25) the autocorrelation function of Eq. (32) becomes

$$\begin{aligned}\rho_S(\tau,s) &= \sum_{m=-2M}^{2M} (2M+1-|m|)\alpha_\mu(\tau - mT_1)\\ &\quad + \sum_{n=-2N}^{2N} (2N+1-|n|)\alpha_\mu(\tau - ns^{-1}T_1)\\ &\quad + 2\sum_{m=-M}^{M}\sum_{n=-N}^{N} \alpha_\mu(\tau - (ns^{-1}-m)T_1).\end{aligned} \tag{33}$$

E. Definition of the generalized coincidence function

As only positive periods up to a certain length $D>0$ are of interest, the generalized coincidence function (GCF) is defined as the integral

$$K(s) := \int_0^D \rho_S^2(\tau,s)d\tau. \tag{34}$$

For each vibration ratio s, $K(s)$ is a measure value of overall coincidence between the two tones of the musical interval with regard to pulse forms and pulse widths. It should be mentioned that only a finite number of pulses are considered, because $D<\infty$.

IV. APPLICATIONS OF THE MODEL TO DIFFERENT PULSE FORMS

A. The general coincidence functions calculated from rectangular pulses

1. Autocorrelation function of the rectangular pulse

To apply the model, the degree of coincidence shall be calculated for all vibration ratios within an octave, $1<s<2$. A pulse function to fulfill properties (i)–(iii) is the rectangular pulse. The parameter μ determines the width of the pulse.

$$I_\mu(t) := \begin{cases} \frac{1}{\mu} & \text{if } |t| < \frac{\mu}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{35}$$

Its autocorrelation function is the triangular pulse (see Fig. 1; see also Papoulis, 1962, p. 243).

$$\alpha_\mu(\tau) = \Delta_\mu(\tau) := \begin{cases} \frac{1}{\mu}\left(1 - \frac{|\tau|}{\mu}\right) & \text{if } |\tau| < \mu \\ 0 & \text{otherwise} \end{cases}. \tag{36}$$

2. Cross-correlation function of the rectangular pulse

The cross-correlation functions of $I_\mu(t)$ and $I_\nu(t)$ can be found from the Fourier transforms of the rectangular pulses $I_\mu(t)$ and $I_\nu(t)$ that are given by (see Papoulis, 1962, p. 21)

$$I_\mu(t) \leftrightarrow \frac{1}{\mu}\frac{2\sin\left(\omega\frac{\mu}{2}\right)}{\omega} =: F_\mu(\omega), \tag{37a}$$

$$I_\nu(t) \leftrightarrow \frac{1}{\nu}\frac{2\sin\left(\omega\frac{\nu}{2}\right)}{\omega} =: F_\nu(\omega). \tag{37b}$$

As $\sin(ax)\sin(bx)=\frac{1}{2}(\cos((a-b)x)-\cos((a+b)x))$, it follows that

$$F_\mu(-\omega)F_\nu(\omega) = \frac{2}{\mu\nu\omega^2}\left(\cos\left(\frac{\mu-\nu}{2}\omega\right) - \cos\left(\frac{\mu+\nu}{2}\omega\right)\right). \tag{38}$$

Note that this product is even. Its Fourier transform is the cross-correlation function [see Eq. (10)]:

$$\begin{aligned}\chi_{\mu\nu}(\tau) &= \frac{1}{2\pi\mu\nu}\int_{-\infty}^{\infty}\frac{1}{\omega^2}\cos\left(\left(\frac{\mu-\nu}{2}-\tau\right)\omega\right)d\omega\\ &\quad + \frac{1}{2\pi\mu\nu}\int_{-\infty}^{\infty}\frac{1}{\omega^2}\cos\left(\left(\frac{\mu-\nu}{2}+\tau\right)\omega\right)d\omega\\ &\quad - \frac{1}{2\pi\mu\nu}\int_{-\infty}^{\infty}\frac{1}{\omega^2}\cos\left(\left(\frac{\mu-\nu}{2}-\tau\right)\omega\right)d\omega\\ &\quad - \frac{1}{2\pi\mu\nu}\int_{-\infty}^{\infty}\frac{1}{\omega^2}\cos\left(\left(\frac{\mu-\nu}{2}+\tau\right)\omega\right)d\omega.\end{aligned} \tag{39}$$

The integrals can be solved using the equation

$$\int_{-\infty}^{\infty}\frac{\cos(a\omega)}{\omega^2}d\omega = -\pi|a|, \tag{40}$$

which can be proved by the calculus of residue. As a result, the cross-correlation function of the two pulses $I_\mu(t)$ and $I_\nu(t)$ is the function [see Fig. 1 A2]

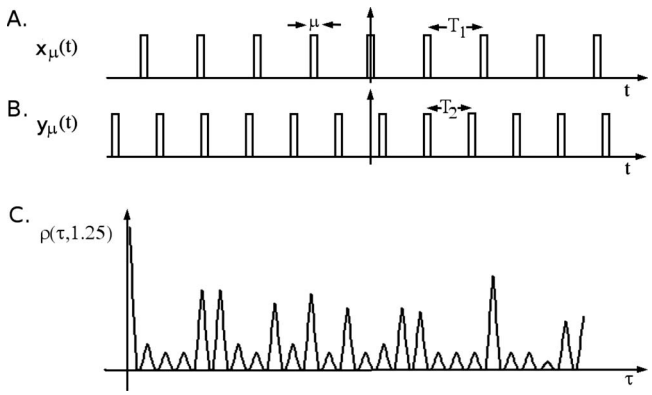

FIG. 2. (A) $x_\mu(t)$ is a sequence of rectangular pulses with a pulse width of μ. It has the period T_1. (B) $y_\mu(t)$ is a sequence of rectangular pulses also with a pulse width of μ. It has the period T_2 and is delayed to $x_\mu(t)$. (C) The ratio $s=(T_1/T_2)=(5/4)=1.25$ is the vibration ration of the major third. The graph shows the autocorrelation function $\rho(\tau,1.25)$ of the sum $S(t)=x_\mu(t)+y_\mu(t)$ of both pulse sequences neuronally representing a major third.

$$\chi_{\mu\nu}(\tau) := \begin{cases} \frac{1}{\mu\nu}\left(\frac{\mu+\nu}{2}-|\tau|\right) & \text{if } \frac{\mu-\nu}{2}<|\tau|\le\frac{\mu+\nu}{2} \\ \frac{1}{\mu} & \text{if } |\tau|\le\frac{\mu-\nu}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{41}$$

3. Autocorrelation function of a musical interval represented by rectangular sequences

Equations (21) and (22) can be applied to rectangular pulses as defined in Eq. (35). The parameter μ now describes a time window. Thus, the neuronal representation of the two tones are now mathematically described as sequences of rectangular pulses [see Figs. 2(a) and 2(b)] The autocorrelation function of the musical interval, represented by $S(t)=x_\mu(t)+x_\nu(t)$, is a sequence of partly coinciding triangular pulses [see Fig. 2(c)]:

$$\begin{aligned}\rho(\tau,s) = &\sum_{m=-2M}^{2M}(2M+1-|m|)\Delta_\mu(\tau-mT_1) \\ &+\sum_{n=-2N}^{2N}(2N+1-|n|)\Delta_\nu(\tau-ns^{-1}T_1) \\ &+2\sum_{m=-M}^{M}\sum_{n=-N}^{N}\chi_{\mu\nu}(\tau-(ns^{-1}-m)T_1).\end{aligned} \tag{42}$$

4. Calculation of the generalized coincidence function for musical intervals represented by rectangular sequences

To compute the GCF, the constants are set as follows.

- Audible frequencies range from about 20 to about 20 000 Hz, which corresponds to periods with lengths from 0 to 50 ms; thus we set $D:=50$.

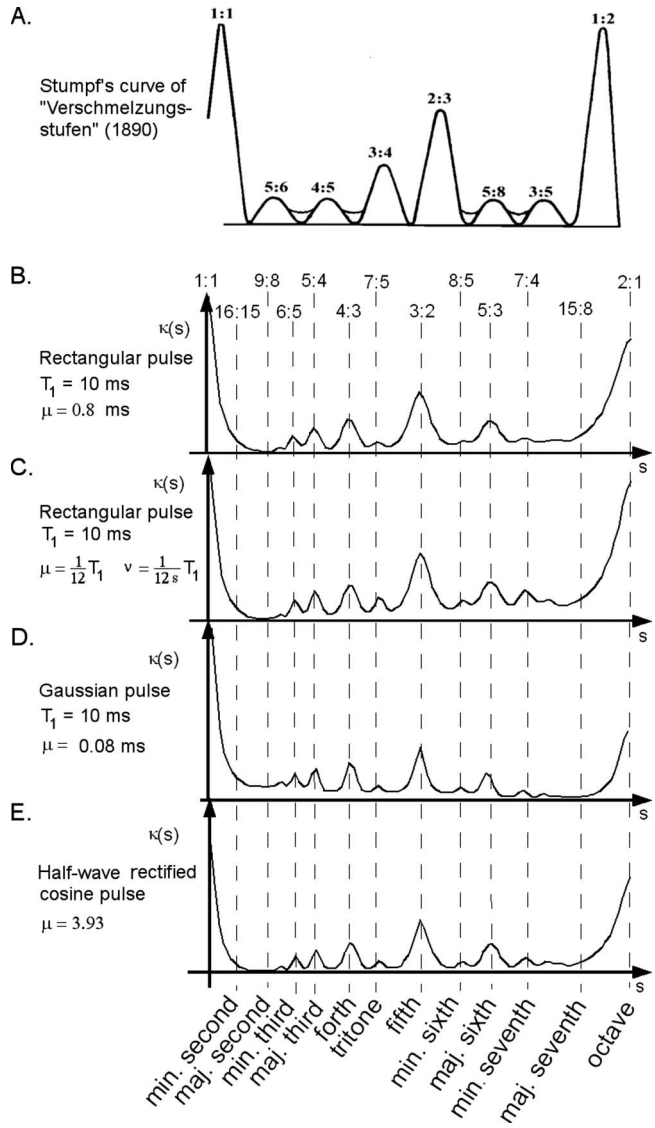

FIG. 3. (A) Stumpf's curve showing the system of the levels of tonal fusion ("System der Verschmelzungsstufen in einer Curve," 1890, 1965). It shows the degree of tonal fusion for all intervals within the range of an octave from the prime (1:1) to the octave (1:2) and the reciprocals of the vibration ratios of the musically important intervals. [(B)–(E)] The computed GCFs $\kappa(s)$ for different pulse forms: at the bottom, the 13 intervals within an octave are indicated and on the top the respective vibration ratios s are plotted.

- As at the most $M=D/T_1$ and $N=D/s^{-1}T_1$ pulses of $x_1(t)$ and $x_2(t)$, respectively, fit into the time window of D, we set $M=\text{floor}(D/T_1)$ and $N=\text{floor}(sD/T_1)$.
- In the first instance, let $\mu=\nu=0.8$ ms and $T_1=10$ ms so that the lowest tone has a frequency of 100 Hz. The relative pulse width can be described by the ratio

$$r := \frac{\mu}{T_1}. \tag{43}$$

In this case, the relative pulse width is $r=0.08$. The computer calculation leads to the GCF of Fig. 3(b) calculated from Eq. (33) with $\alpha_\mu(\tau)=\Delta_\mu(\tau)$ [see Eq. (25)].

- Again, let be $T_1=10$ ms and set $\mu=\frac{1}{12}T_1=0.898$ ms and $\nu=\frac{1}{12}T_2$. This case of a frequency dependent upper pulse width leads to the GCF of Fig. 3(c).

The graphs show maxima for musically important intervals within an octave. As can clearly be seen, the simpler the vibration ratio of a musical interval, the higher the corresponding peak in the GCF.

B. Two examples of the general coincidence functions calculated from other pulse forms

1. Generalized coincidence function calculated on the basis of the Gaussian pulse

The Gaussian pulse

$$I_\mu(t) = \frac{1}{\sqrt{\mu\pi}} e^{-1/\mu t^2} \tag{44}$$

is a distribution fulfilling properties (i)–(iii). The parameter μ is dimensionless and determines the width of the pulse. The autocorrelation function of the Gaussian pulse can be calculated from Eq. (9). Its Fourier transform is [see Papoulis, 1962, p. 25 (2–68)]

$$F(\omega) = e^{-\mu/4\omega^2}. \tag{45}$$

Thus, according to Eqs. (9) and (12) the Fourier transform of the product $F(\omega)F(-\omega)$ is the autocorrelation function of the Gaussian pulse [see Papoulis, 1962, p. 25 (2–68)]:

$$\alpha_\mu(\tau) = \frac{1}{\sqrt{2\mu\nu}} e^{-1/2\mu\tau^2} \leftrightarrow F(\omega)F(-\omega) = e^{-\mu/2\omega^2}. \tag{46}$$

It can be shown that $\alpha_\mu(\tau)$ is a distribution fulfilling properties (i)–(iii). Considering only the case that the pulses are of the same variance (width) for both pulse sequences, the autocorrelation function $\rho_S(\tau,s)$ of the sum of both sequences can be obtained from Eq. (33) with $\alpha_\mu(\tau)$ as in Eq. (46). Setting $\mu=0.08$ gives a Gaussian pulse that fits well into a rectangular pulse with a width of 0.8 ms, the Gaussian pulse with $\mu=0.102$ has the same center of gravity as the rectangular pulse with a width of 0.8. Again, set $T=10$ ms.The calculated GCF for $\rho_S(\tau,s)$ in the case of Gaussian pulses with $\mu=0.08$ is shown in Fig. 3(d).

2. Generalized coincidence function calculated on the basis of a half-wave rectified cosine pulse

The half-wave rectified cosine pulse, given by

$$I_\mu(t) := \begin{cases} \frac{1}{2\mu}\cos\left(\frac{1}{\mu}t\right) & \text{if } |t| < \frac{\pi}{2}\mu \\ 0 & \text{else} \end{cases}, \tag{47}$$

is a distribution fulfilling properties (i)–(iii). In this case, the parameter μ^{-1} can be interpreted as a frequency and determines the width of the cosine pulse. From definition (3), its autocorrelation function calculates to

$$\alpha_\mu(\tau) = \begin{cases} \frac{1}{8\mu^2}\left[(\pi\mu - |\tau|)\cos\left(\frac{|\tau|}{\mu}\right) + \mu\sin\left(\frac{|\tau|}{\mu}\right)\right] & \text{if } |\tau| < \pi\mu \\ 0 \quad \text{if } |\tau| \geq \pi\mu \end{cases}. \tag{48}$$

Again, it can be shown that $\alpha_\mu(\tau)$ is a distribution fulfilling properties (i)–(iii). Considering only the case that the pulses are of the same frequency (and thus of the same width) for both pulse sequences, the autocorrelation function $\rho_S(\tau,s)$ of the sum of both sequences can be obtained from Eq. (33) with $\alpha_\mu(\tau)$ as in Eq. (48). Setting $\mu=0.3142$ results in a half-wave rectified pulse with the same center of gravity as the rectangular pulse with a width of 0.8; choosing $\mu=0.364$ gives a half-wave rectified cosine pulse with an amplitude half down for $t=0.4$, and the half-wave rectified cosine pulse with $\mu=0.8\pi^{-1}$ fits well into a rectangular pulse with a width of 0.8 ms. Again set $T=10$ ms. The calculated GCF for $\rho_S(\tau,s)$ in the case of a half-wave rectified cosine pulse with $\mu=0.8\pi^{-1}$ is shown in Fig. 3(e).

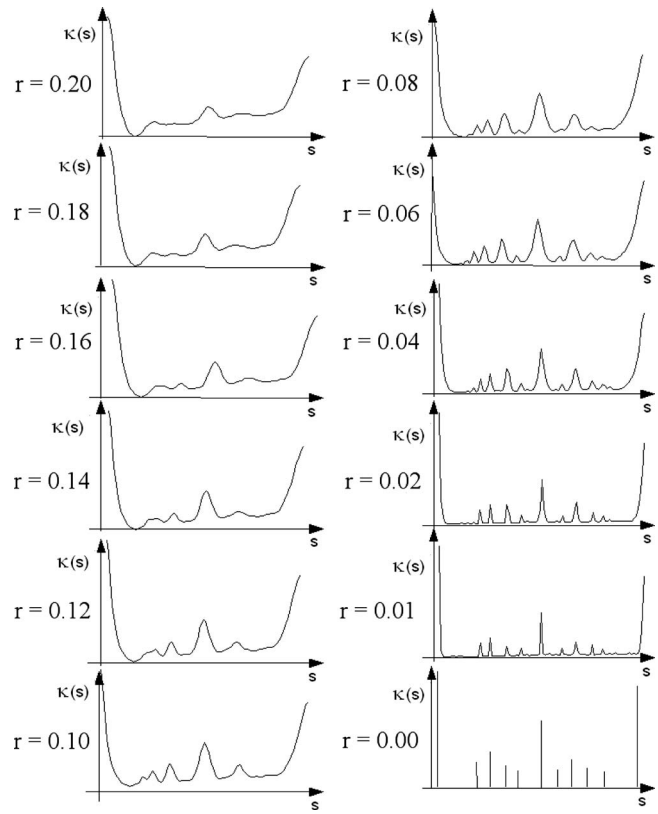

FIG. 4. (Color online) 12 GCFs are computed from sequences of rectangular pulse with the relative pulse width r converging from $r=0.2$ to zero, thus illustrating the convergence of the GCF.

C. The convergence of the generalized coincidence function

All pulses of the pulse sequences above are distributions with property (iii) so that every pulse has the generalized limit $\delta(t)$. Considering this property of the pulses, it becomes clear that all GCFs computed from different pulse forms approach the same limit function as parameter μ approaches zero. In Fig. 4, the convergence is illustrated for GCFs calculated from rectangular pulses of different widths with relative pulse widths ranging from $r=0.2$ to $r=0$. Taking into account the convergence against the same limit function, it becomes clear that the GCFs calculated from different pulse forms must become quite similar if the parameter μ is sufficiently small.

V. DISCUSSION

A. Properties of the generalized coincidence functions

The GCF shows a ranking of the musical intervals that is obviously equal to the traditional degrees of consonance as described by music theorists. Moreover, every consonant interval is surrounded by a region of still high coincidence. This is in accordance with the observation that slightly mistuned consonances preserve the characteristic of the interval and still sound consonant.

Though the model is formed in accordance with auditory pulse trains, the pulses of the model as applied in this paper do not reflect neuronal spiking but represent the probability of coincidence in an auditory autocorrelation process. Thus, the pulse shape is determined by properties of an auditory autocorrelator. The pulse width depends on the width of the coincidence window provided by the neuronal autocorrelation mechanism.

Provided that the pulses have a sufficiently small width, the shapes of the GCFs are only slightly affected by the pulse form. In spite of a presumable randomness of pulse forms in the auditory autocorrelation mechanism, the GCF obviously grants a great stability in coincidence detection.

Only symmetric pulses have been discussed so far. From Eq. (6), the symmetry of the autocorrelation function of any pulse is granted and from Eq. (8) it becomes clear that two related cross-correlation functions sum up to a symmetric pulse. So, on the level of the autocorrelation function, only symmetric pulses occur. Thus, it is totally sufficient to consider symmetric pulse forms exclusively.

As demonstrated in Sec. IV C, the shape of the GCF is sensitive to the relative pulse width, which is the ratio of pulse width to period. On the other hand, the number of considered spikes has a slight influence only on the shape of the GCF: The coincidences in the autocorrelation function of the interval are heightened. The squaring in the GCF leads to a more pronounced curve, but no general changes occur.

The assumption—that the sum of two pulse sequences [see Eq. (23)] is an adequate representation of firing patterns, evoked by musical intervals—is a linear assumption. Furthermore, the GCF is based on correlation functions in the time domain. The underlying mathematical techniques imply that the auditory system acts as a LTI system. Quite often LTIs are applied to describe neuronal processes. "In general, the success of most of these time series analysis methods in physiology is surprising considering that physiological processes are known to include significant nonlinearities. The explanation for this relative success is perhaps due to the cases in which physiological systems can be studied in a state where the linear behavior is most prominent …" (van Drongelen, 2007, p. 279). Each stage on the auditory pathway is a potential site of nonlinearity (see Hartmann, 2000, p. 511). However, first and foremost, nonlinearity affects signal strength and not the time patterns. If the effects of nonlinearity are comparatively small, especially in the dynamic range of music, the consideration of a LTI system is justified.

Wever *et al.* (1941), who discovered the cochlear microphonic, found that "the peripheral mechanism of the ear transmits vibrations with high fidelity" and considered "the most probable source of distortion to be the processes of the inner ear through which mechanical vibrations are transformed into electrical effects" (Wever *et al.*, 1941). The nonlinearities of the inner ear become evident from combination tones and suppression (de Boer, 1984). Plomp (1965) investigated the audibility of combination tones and determined that "all mean detectability thresholds found exceeded 40 dB, corresponding to a nonlinear distortion of below 1%." In experiments with musical interval stimuli, Tramo *et al.* (2001) found a strong correspondence between the pooled autocorrelogramms of neural spike rates and the computed autocorrelation functions of the stimuli, so that no significant suppression effect could be observed. Moreover, neuronal nonlinear effects, such as otoacoustic emissions (Zwicker and Fastl, 1999) or inhibition (Buzsáki, 2006), are believed to even enhance neural signal quality. Last, but not least, the evident correspondence between the GCF and Stumpf's "Stufen der Tonverschmelzung" may be regarded as a cue for a linearity in the perception of harmony.

B. The generalized coincidence functions and Stumpf's tonal fusion

Figure 3(a) shows the curve which Stumpf (1897) proposed to illustrate the levels of tonal fusion (Stufen der Tonverschmelzung) as he had investigated. The graph shows the degree of fusion for all intervals over a range of an octave. Not only do the consonant intervals, which are the prime (1:1), the minor third (5:6), the major third (4:5), the pure fourth (3:4), the pure fifth (2:3), the minor sixth (5:8), the major sixth (3:5), and the pure octave (1:2)—indicated by their vibration ratio—have a higher degree of fusion, but the slightly mistuned intervals nearby the consonant intervals also do.

The similarities of Stumpf's curve and the graph of the GCF are obvious, especially for relative pulse width in the range of $r=0.10-0.04$. Stumpf's curve is a schematic sketch showing an equal degree of coincidence for both sixths and thirds. He felt unsure about their order but had some evidence that the major third fuses more intensively than the minor third. The order of the minor and major sixth was not clear to him (Stumpf, 1897). In this respect, all GCFs in the range of $r=0.10-0.04$ clearly show a higher tonal fusion of the major sixth.

On the whole, one can conclude that tonal fusion and the degree of coinciding periods in the firing patterns for musical intervals are equivalent. Stumpf's concept of tonal fusion was an attempt to define consonance and dissonance psychologically. Considering the theory of periodicity detection in the midbrain as described by Langner (1983; Langner and Schreiner, 1988) and calculating the coincidence of musical intervals with a coincidence window of 0.8 ms as suggested by this theory result in a GCF closely resembling the curve Stumpf deduced from his extensive hearing experiments. Thus, Langner's model of periodicity analysis in the midbrain may provide a physiological correlate of tonal fusion.

Nevertheless, roughness is also an important sensation that is used to distinguish between consonance and dissonance. Pure tones produce roughness if they are close together or if the upper tone is close to the octave (Zwicker and Fastl, 1999). The intervals of the minor second, the minor seventh, and the minor ninth sound very harsh but show a high degree of coincidence, as can be seen in Figs. 3 and 4. They are dissonant only because of their roughness and in spite of their high degree of tonal fusion. On the other hand, we can also distinguish the degrees of consonance in the absence of roughness: for example, if an interval is formed by pure tones. In the absence of roughness, the sensation of consonance must be a consequence of tonal fusion. This leads to the conclusion that both roughness and tonal fusion determine the sensation of consonance and dissonance.

Burns, E. M., and Ward, W. D. (**1982**). "Intervals, scales, and tuning," in *The Psychology of Music*, edited by D. Deutsch (Academic, London), pp. 241–269.

Buzsáki, G. (**2006**). *Rhythms of the Brain* (Oxford, New York).

Cariani, P. A., and Delgutte, B. (**1996**). "Neural correlates of the pitch of complex tones. I. Pitch and pitch Salience, II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch," J. Neurophysiol. **76**, 1698–1734.

Cedolin, L., and Delgutte, B. (**2005**). "Representation of the pitch of complex tones in the auditory nerve," in *Auditory Signal Processing*, edited by D. Pressnitzer, A. de Cheveigné, S. McAdams, and L. Collet (Springer, Danvers).

de Boer, E. (**1984**). "Auditory physics. physical principles in hearing theory. II," Phys. Rep. **105**, 141–226.

de Cheveigné, A. (**2005**). "Pitch perception models" in *Pitch: Neural Coding and Perception*, edited by C. Plack, R. R. Fay, A. J. Oxenham, and A. N. Popper (Springer, New York), pp. 169–233.

Goldstein, J. L., and Scrulovic, P. (**1977**). "Auditory-nerve spike intervals as an adequate basis for aural spectrum analysis," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York), pp. 337–345.

Hartmann, W. M. (**2000**). *Signal, Sound, and Sensation*, 4th ed. (Springer, New York).

Hewitt, M. J., and Meddis, R. (**1994**). "A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus," J. Acoust. Soc. Am. **95**, 2145–2159.

Kaernbach Chr., and Demany L. (**1998**). "Psychophysical evidence against the autocorrelation theory of auditory temporal processing," J. Acoust. Soc. Am. **104**, 2298–2306.

Langner, G. (**1983**). "Evidence for neuronal periodicity detection in the auditory system of the guinea fowl implications for pitch analysis in the time domain," Exp. Brain Res. **52**, 333–355.

Langner, G. (**2005**). "Neuronal mechanisms underlying the perception of pitch and harmony," Ann. N. Y. Acad. Sci. **1060**, 50–52.

Langner, G. (**2007**). "Temporal processing of periodic signals in the auditory system: neuronal representation of pitch, timbre, and harmonicity," Z. Audiologie **46**, 8–21.

Langner, G., and Schreiner, C. E. (**1988**). "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms, II. Topographical organization," J. Neurophysiol.. **60**, 1799–1822 and 1823–1840.

Licklider, J. C. R. (**1951**). "A duplex theory of pitch perception," Cell. Mol. Life Sci. **7/4**, 128–134.

Meddis, R., and Hewitt, M. J. (**1991**). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. II: Phase sensitivity," J. Acoust. Soc. Am. **89**, 2866–2894.

Moore, B. C. J., and Glasberg, B. R. (**1986**). "The role of frequency selectivity in the perception of loudness, pitch and time," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).

Papoulis, A. (**1962**). *The Fourier Integral and Its Applications* (McGraw-Hill, New York).

Patterson, R. D., Allerhand, M. H., and Giguère Chr. (**1995**). "Time-domain modelling of peripheral auditory processing: A modular architecture and a software platform," J. Acoust. Soc. Am. **98**, 1890–1894.

Plomp, R. (**1965**). "Detectability threshold for combination tones," J. Acoust. Soc. Am. **37**, 1110–1123.

Plomp, R., and Levelt, W. J. M. (**1965**). "Tonal consonance and critical bandwidth," J. Acoust. Soc. Am. **38**, 548–560.

Rhode, W. S. (**1995**). "Interspike intervals as correlate of periodicity pitch in cat cochlear nucleus," J. Acoust. Soc. Am. **97**, 2414–2429.

Stumpf, C. (**1890**). *Tonpsychologie ("Psychology of Tones")* (Knuf, Hilversum), reprinted 1965.

Stumpf, C. (**1897**). "Neueres über Tonverschmelzung," Zeitschift für Psychologie und Physiologie der Sinnesorgane **15**, 280–303.

Stumpf, C. (**1930**). "Autobiography of Carl Stumpf," in *History of Psychology in Autobiography*, edited by C. Murchison (Clark University Press, Worchester, MA), Vol. **1**, pp. 389–441.

Terhardt, E. (**1976,1977**). "Ein psychoakustisch begründetes Konzept der Musikalischen Konsonanz," ("A psychoacoustically substantiated concept of musical consonance,"),," Acustica **36**, 121–137.

Terhardt, E. (**1984**). "The concept of musical consonance: A link between music and psychoacoustics," Music Percept. **1**, 276–226.

Tramo, M. J. P., Cariani, A., Delgutte, B., and Braida, L. D. (**2001**). "Neurobiological foundations for the theory of harmony in Western tonal music," in *The Biological Foundations of Music*, edited by J. Zatorre et I. Peretz (Annals of the New York Academy of Sciences, New York), Vol. **930**, pp. 92–116.

van Drongelen, W. (**2007**). *Signal Processing for Neuroscientists* (Elsevier, Amsterdam).

von Helmholtz H. (**1877**). in *On the Sensations of Tone*, edited by J. A. Ellis (Dover, New York), reprinted 1885 and 1954.

Voutsas, K., Langner, G., Adamy, J., and Ochse, M. (**2005**). "A brain-like neural network for periodicity analysis," IEEE Trans. Syst., Man, Cybern., Part B: Cybern. **35**, 12–22.

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

Wever, E. G., Bray, C. W., and Lawrence, M. (**1941**). "Combination tones: their nature and origin in the auditory mechanism," J. Acoust. Soc. Am. **12**, 468.

Wiener, N. (**1930**). "Generalized harmonic analysis," Acta Math. **55**, 117–258.

Yost, W. A., Patterson, R., and Sheft, St. (**1996**). "A time domain description for the pitch strength of iterated rippled noise," J. Acoust. Soc. Am. **99**, 1066–1077.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). "A phenomenological model for the responses of auditory–nerve fibers," J. Acoust. Soc. Am. **109**, 648–670.

Zwicker, E., and Fastl, H. (**1999**). *Psychoacoustics*, 2nd ed. (Springer, Berlin).