Does the brain process sound in the time domain?

Our typical “Machine Hearing” system makes good use of fine time structure

Respect the Auditory Nerve

“Artificial Auditory Recognition in Telephony”

Auditory nerve carries fine time structure

Cochlear model: CAR-FAC

(cascade of asymmetric resonators with fast-acting compression)

Tonotopic projection to cortex, but…

What is the other dimension in auditory cortex?

What corresponds to image retinotopy in vision?

Auditory images…
Correlograms are movies; this one is based on Roy Patterson’s “Stabilized Auditory Image” (SAI)

Timing explains pitch of aperiodic sounds
• “Strike note of a chime” demo from ASA Auditory Demonstrations CD.
• Correlogram (SAI) shows near alignment at interval corresponding to the musical pitch:

Temporal coincidence detection theories:
Jeffress 1948 binaural, Licklider 1951 pitch

A trigger method for Patterson’s triggered temporal integration approach to making stabilized auditory images:
One trigger per segment, at max; channel 60:
Wavelength with trigger points marked

Without trigger alignment, integrating (averaging) segments doesn’t make sense:

Trigger points in 20 ms cochleagram segment

Align to triggers, and temporally integrate
SAI and Summary SAI

Licklider’s autocorrelator neural circuit

\[ g(t, \tau) = (f(t)f(t - \tau)) \ast w(t) \]

Fig. 1. - Basic scheme of neuronal autocorrelator. \( A \) is the input nerve, \( B_1, B_2, B_3, \ldots \) is a delay chain. The original signal and the delayed signal are multiplied when \( A \) and \( B_1 \) feed \( C_1 \), and a running integral of the product is obtained at the output. When \( C_1 \) and \( B_2 \) feed \( C_2 \), the two delayed signals are compared at \( C_2 \), which could be the difference of \( C_1 \). The comparison is repeated in a similar way. Since the output filter is proportional to the convolution of running autocorrelation, the excitatory stimuli of \( B_1, B_2, B_3, \ldots \) provide a display of the running autocorrelation functions of the input time function, the temporal course of the stimulus of \( A \).

**A Auditory Nerve**

AVCN bushy cells extract trigger points?

“Primary-like” bushy cells sync well to stimulus period

**B Bushy Cell**

Cross-corr versions

\[ g_1(t, \tau) = (f(t)f(t - \tau)) \ast w(t) \]

\[ g_2(t, \tau) = (f(t)f(t - \tau)) \ast w(t) \]

Lord Rayleigh’s 1907 “duplex theory” of binaural lateralization: interaural intensity and phase differences

Phase is ambiguous above about 650 Hz.

Thus, although there might be right and left sensations from sources obliquely situated, these sensations would fall when most needed, that is when the sound is really in the line of the ears. In this case a perception of phase-differences would seem to do more harm than good. At a pitch a little higher, ambiguity of a misleading and dangerous kind would necessarily enter. For example, the same sounds might arise from a sound a little on the left and from another fully on the right.

On the whole it appears that the sensation of lateralness due to phase-difference disappears in the region of pitch where there would be danger of its becoming a misleading guide. ... It is fortunate that when difference of phase fails, difference of intensity comes to our aid.

Demo break

real-time
AIM-C movies of music and stuff
then binaural...
But sounds other than sinusoids are easier...

Jens Blauert Spatial Hearing 1997

Mallock 1908 observation of ITD cue

A sound which is caused by the detached waves, such as those which accompany a bullet, can scarcely be said to have a pitch, but the wave-length is certainly small compared with the distance between the ears, and is indeed comparable with the dimensions of the bullet itself. It would seem, therefore, that the ears can determine the direction of a sound, not only by difference of phase, but by the actual difference in the times at which a single pulse reaches them.

1920 ITD compensator

World War I era directional listener using a linear compensator to compute arrivals (Drysdale, 1920). The dashed line on the left is a waveform.

Rotating ITD eliminator measures angles directly

World War I era acoustic goniometer for locating "invisible aeroplanes" (Ferry, 1921).

HRTFs and HRIRs “encode” direction; horizontal plane:

Head-related impulse responses of a dummy head, for sounds from various directions.

Daniel Tollin at U. Colo. Denver
Barn Owl’s ITD system in N.L.

In order to explain the existence of a movable image of the sound within this zone, we may suppose that the transmission of the sound impulse through some specialized part of the auditory apparatus or brain takes a definite time from each ear, and that the point where the impulses meet is the focus that gives rise to the sensation of a sound-image.

This was regarded as an “unwelcome hypothesis”.

Lloyd Jeffress’s model...
Bowlker 1908

HRIRs for “frontal plane” (over top of head)

Head-related impulse responses of a dummy head, for sounds from various directions.

Binaural circuits
fig. from Daniel J. Tollin 2003

Specialized fast cells and synapses for triggering and correlation: endbulb of Held, calyx of Held

Henrique von Gersdorff & J. Gerard G. Borst 2002
Tom C. Yin’s lab: confirming Jeffress-style delay lines in MSO

Highest peaks are not best arrival-time triggers

Binaural Auditory Image

Hartung & Trahiotus 2001: an appropriate hair cell model gives it some "precedence effect"

Monaural cross-corr versions can easily be changed to binaural

\[ g_1(t, \tau) = (\hat{f}(t)f(t-\tau)) * w(t) \]

\[ g_2(t, \tau) = (f(t)\hat{f}(t-\tau)) * w(t) \]
Binaural test: speech and ping; Lyon 1983

Figure 2. Cork diagrams of test signals. Top: left and right channels of speech sound. Middle: left and right channels of introducing ping sound, with reverberation. Bottom: left and right composite sounds, the inputs to the beamformer separation task.

Binaural localization/separation outputs

Figure 3. Separation results. Top: left and right separated sound streams. Bottom: left and right echo, or reverberation.