RESPONSE-FIELD DYNAMICS IN THE AUDITORY PATHWAY

D.A. Depireux, Powen Ru, S.A. Shamma and J.Z. Simon

Center for Auditory and Acoustic Research
Institute for Systems Research
University of Maryland
College Park MD 20742 U.S.A.

I. INTRODUCTION

Natural Sounds are characterized by loudness, pitch and timbre (i.e. the dynamic envelope of the spectrum).

Our Question: how is timbre encoded in primary auditory cortex (AI)?

Our Approach: beyond the sensory epithelium, principles used by neural systems are universal. So we view the basilar membrane as a 1-D retina, and use the method of gratings to study single units in AI.

Important Concepts

• *Response Field* (RF): range of frequencies that influence a neuron (as a function of time).

• *Ripple*: broadband sound of sinusoidally modulated spectral envelope ("auditory grating").

• Data analysis based on **linear systems theory** to characterize response field. By varying ripple frequency and velocity, we measure the transfer function. The inverse Fourier transform gives the *spectro-temporal* RF (STRF).

We Find:

• Cells can be characterized by an STRF, separable or not.

• Cells behave like a linear system: when presented with a sound made of up the sum of several profiles, the response of the cell is the sum of the responses to the individual profiles.

• Response fields in AI tend to have characteristic shapes both spectrally and temporally.

• Cortical cells with all center frequencies, all spectral symmetries, bandwidths, latencies and temporal impulse response symmetries.

We Show predictions of single-unit responses in AI to complex spectra, verifying:

• **Linearity** of AI responses to all types of dynamic ripples: responses to up and down moving ripples can be superimposed linearly to predict responses to arbitrary combinations of these ripples.

• **Separability** of spectral and temporal measurements of the responses: spectral properties can be measured independently of the temporal properties.

We Conclude: Because of linearity of cortical responses with respect to spectral envelope, we can use the ripple method to characterize auditory cortical cell responses to dynamic, broadband sounds. AI decomposes the input spectrum into different spectrally and temporally tuned channels. Another view is that a population of such cells effectively represents the

input spectrum at multiple scales. AI performs a multi-dimensional, multi-scale wavelet transform of the auditory spectrum. The combined spectro-temporal decomposition in AI can be described by an affine wavelet transformation of the input, in concert with a similar temporal decomposition.

II. THEORY

A. Spectro-Temporal Fourier Transform

Since the cochlea performs (to first order) a Fourier transform along the log frequency axis, we measure spectral distance in log(frequency). Since the Fourier transform is time-windowed, we also require a time axis. For this reason we will focus attention on two-dimensional functions of log(frequency) and time. For linear systems, the spectro-temporal domain and its Fourier domain are equivalent. Analysis is often conceptually simpler in the Fourier domain. Real functions in the spectro-temporal domain give rise to complex conjugate symmetric functions in Fourier space.

The next figure illustrates the envelope of a speech fragment (*"Water all year"*), in both its spectro-temporal and Fourier representations. In the Fourier representation, the function is highly concentrated near zero.
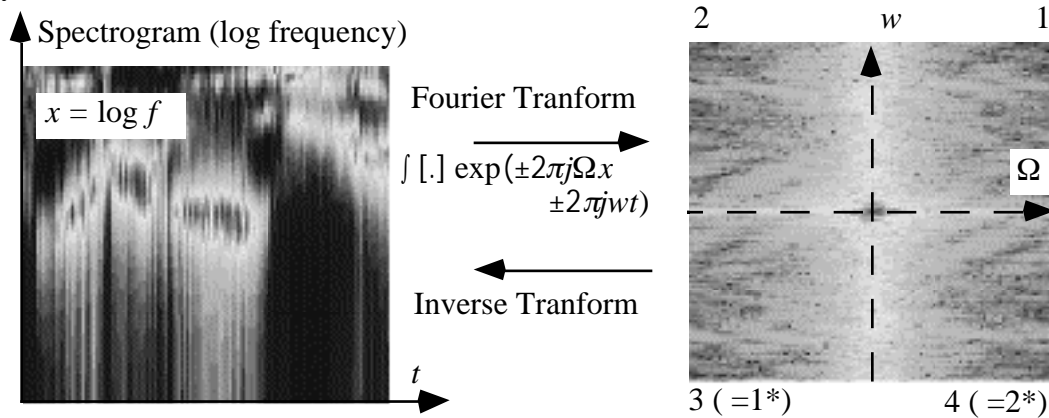

Figure 1: w = ripple velocity, Ω = ripple frequency

B. Spectro-Temporal Response and the Fourier Transform (Transfer Function)

Properties of AI cells are typically derived using pure tones or clicks akin to using dots of light or flashes to study cells in the visual pathway. We use the auditory version of drifting gratings[1] to characterize response properties of cells to dynamic broadband sounds, so as to gain insight to how timbre is encoded. The method presented here allows us to simultaneously determine temporal and spectral properties, using the same set of stimuli for a variety of cells. We use the Response Field (RF), a function measured using broadband sounds. It is given in the form of a function, with positive values describing excitation) and negative values inhibition.
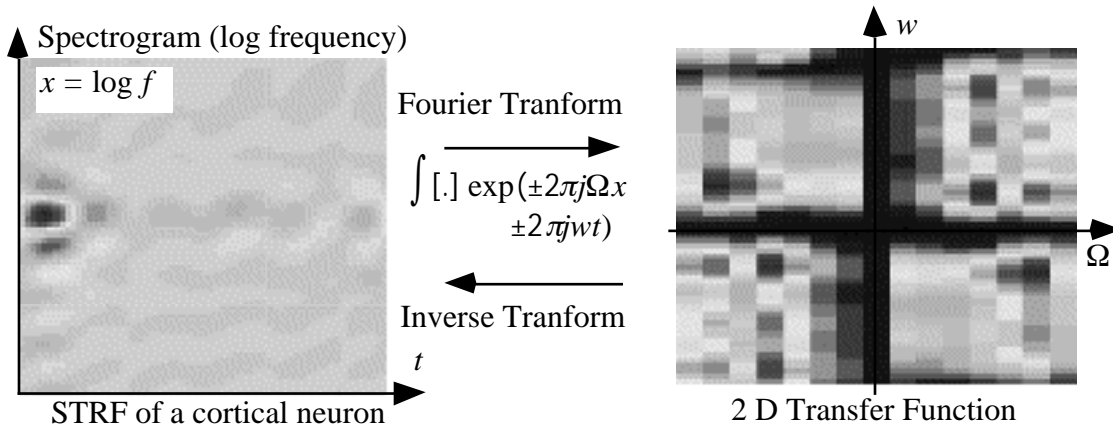

Figure 2: Spectro-temporal RF of a neuron, and its Fourier dual, the transfer function.

C. Spectro-Temporal Stimulus and the Fourier Transform

Natural sounds, such as environmental sounds and speech, are classified along several perceptual axes: loudness, pitch and timbre. Pitch is what changes when we pronounce the same vowel with different tonal heights. Timbre is what changes when, keeping the same tonal height, we pronounce different vowels. In this work we address timber only. Figure 3 illustrates the spectral envelope of a sound, i.e. its timbre. It can be viewed as a low-order polynomial fit of the (time-windowed) spectrum of the sound. A common method for the extraction of the envelope is the Linear Predictive Method (LPC).[2]
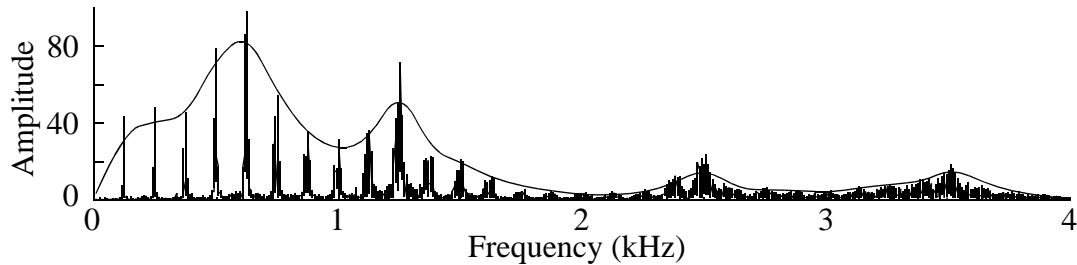

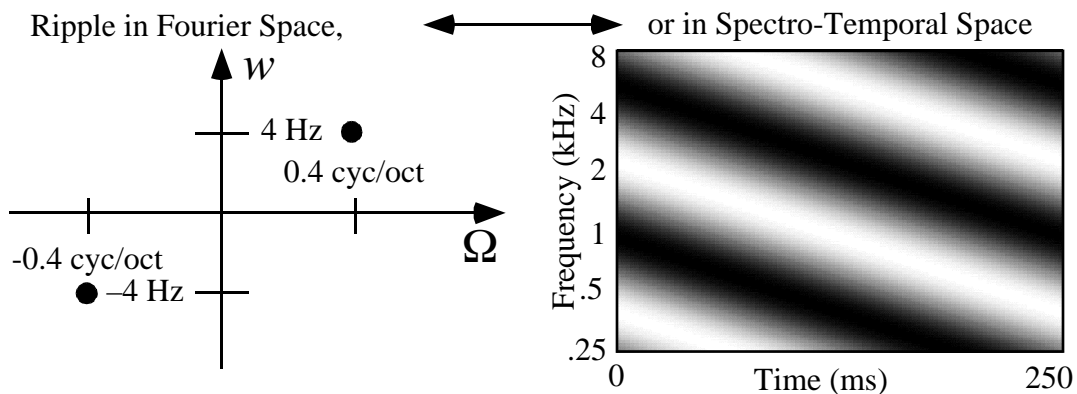**Figure 3**: The spectrum of /aa/ spoken by one author, with the spectral envelope superimposed.

Figure 4: Points in the Fourier space correspond to broadband sounds with a sinusoidally modulated spectral and temporal envelope. The Fourier transform of a ripple has support only on a single point (and its conjugate).

D. Quadrant Separability

An STRF can fall into one of three categories:

• **Non-separable**: The transfer function is an arbitrary function of ripple frequency and ripple velocity.

• **Quadrant separable**: The transfer function within each quadrant is a product of a function of ripple frequency and a function of ripple velocity. The *envelope* of the STRF is the product of a function of spectrum and a function of time.

• **Fully separable**: The transfer function is the product of a function of ripple frequency and ripple velocity everywhere. The resulting STRF is a product of a function of spectrum and a function of time.

E. Linearity

The guiding principle behind our research program is that cells behave like a linear system with respect to the spectral envelope. The proof of linearity is that when cells are presented with a sound made of up the sum of several spectral envelopes, the response, as measured assuming a rate code, is the sum of the responses to the individual envelopes. A response linear in frequency and time is characterized by a two-dimensional impulse response (or time-dependent response field) or equivalently, its two-dimensional Fourier transform.

As indicated for a 4 Hz ripple in Figure 5, the response of a cell as a function of time is modulated at the same (temporal) frequency as that of the stimulus. Therefore, we just have to extract the phase and the amplitude of the response.
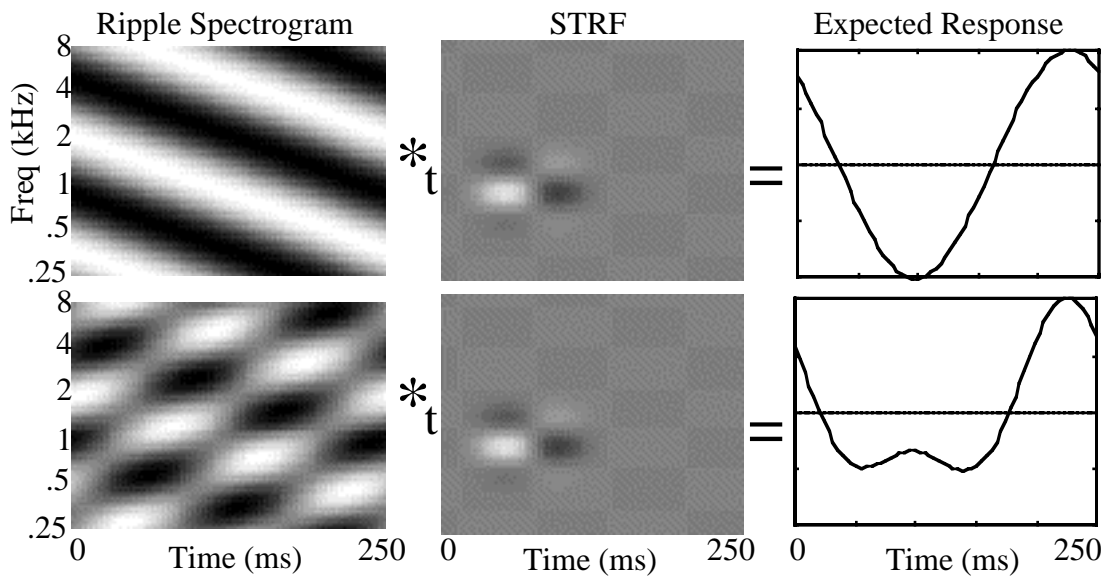
Figure 5: Assuming linearity, the STRF predicts the response to any broadband dynamic stimulus, including single ripples moving in either direction (first row) and combinations of upward and downward moving ripples.
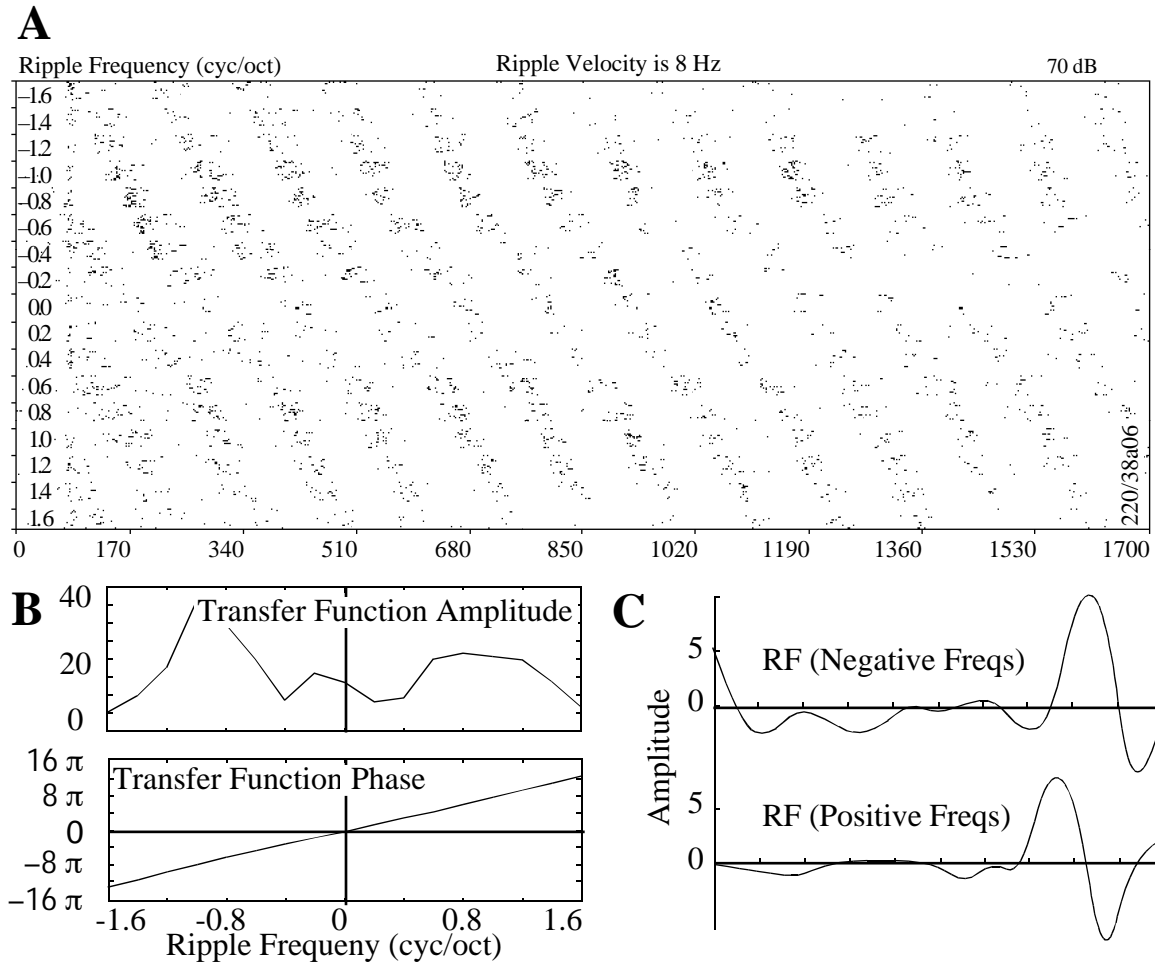
III. EXPERIMENT AND RESULTS

Figure 6: Data analysis using ripples of fixed velocity and varying frequencies. **A**: Raster plot of responses. Each point represents an action potential, and each paradigm is presented 15 times. **B**: Magnitude and phase of the period histogram fits. **C**: Separate inverse Fourier transforms for positive and negative ripple frequencies of **B**, obtaining a slice of the RF.

4

Data were collected from the auditory cortex of domestic ferrets anesthetized with ket-amine and xylazine; with sounds presented in the contralateral ear. AI cells were typically isolated in cortical layers III and IV[3]. For details see Shamma et al.[3]

A. Obtaining the Transfer Functions

We measure cells' transfer functions by presenting, at a fixed ripple frequency, ripples of varying velocities; then, for a fixed velocity, we present ripples of varying ripple frequencies.

A typical example of the analysis is shown in Figure 6. Ripples were presented at 8 Hz, for ripples frequencies from –1.6 cyc/oct to 1.6 cyc/oct in steps of 0.2 cyc/oct, with the ripple starting to move at t = 0 ms, and being acoustically turned on starting at 50 ms. Each ripple is presented 15 times. Once the onset activity has died away, the cell goes into a steady-state response. For each ripple frequency, we compute a period histogram excluding the onset response. To assess the strength and phase of the phase-locked response, we compute the phase and the strength of the response of the cell by Fourier transforming a 16 bin period his-togram of the response, extracting the phase and amplitude of $T(\Omega, w = 8 \text{ Hz})$ from the first component of the transform

The magnitude and phase of the transfer function is shown in panel **B**. In **C**, we have inverse Fourier transformed separately the transfer function in quadrant 1 and 2, or equiva-lently for down- and up-moving ripples, after removing the constant (temporal) phase factor $2\pi w\tau_d + \theta$, where $w = 8$ Hz.

The extraction of the temporal cross-section of the transfer function as in Figure 6 would proceed the same way. Ripples are presented at 0.4 cyc/oct, for ripple velocities from –24 Hz to 24 Hz in steps of 4 Hz. For each ripple frequency, we compute a period histogram to assess the strength and phase of the phase-locked response. The amplitude and phase of the response is then evaluated by performing a Fourier transform of the data, and extracting the phase and the amplitude of $T(\Omega = 0.4 \text{ cyc/oct}, w)$ from the first component of the Fourier transform. We inverse Fourier transformed separately the transfer function for down- and up-moving ripples, after removing the constant (spectral) phase factor $2\pi\Omega x_m + \phi$, where $\Omega = 0.4$ cyc/oct.
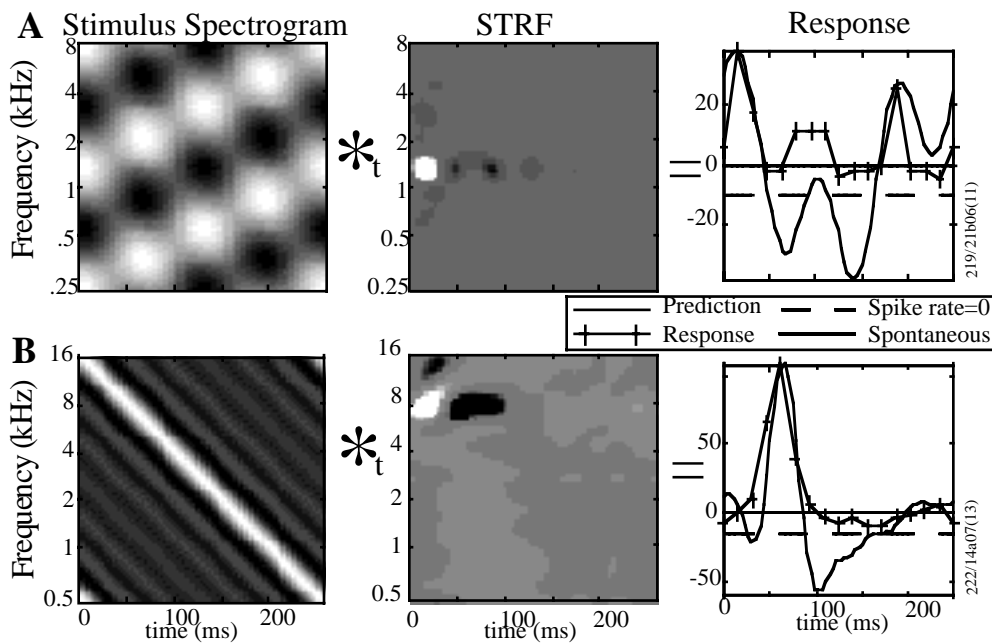
B. Separability and Linearity

Figure 7: Predictions of response to complex dynamic spectra using the STRF. **A** A prediction is computed by convolution (along t) of the STRF with the spectrogram The stimulus shown consists of 2 ripples (0.4 cyc/oct at 12 Hz and –4 Hz). The prediction is shown juxtaposed with the actual response (crosses) over one stimulus period. **B** Another example: the stimulus consists of a combination of ripples with ripple frequencies 0.2 cyc/oct at 4 Hz, 0.4 cyc/oct at 8 Hz, ... 1.2 cycles/octave at 24 Hz, in cosine phase, resulting in an FM-like stimulus.

In vision, some cortical simple cells are fully separable,[4] but all are at least quadrant separable.[5] We have found both types in AI as well; Figure 8 shows examples of each. A fully separable cell has an STRF that is a simple product of an RF and an IR, as in the left two examples. A quadrant separable cell, as in the right two examples, does not, since it has different responses for upward and downward moving ripples: the STRF is not symmetric about x_m. The separability of a cell does not affect the linearity of responses to ripple combinations.
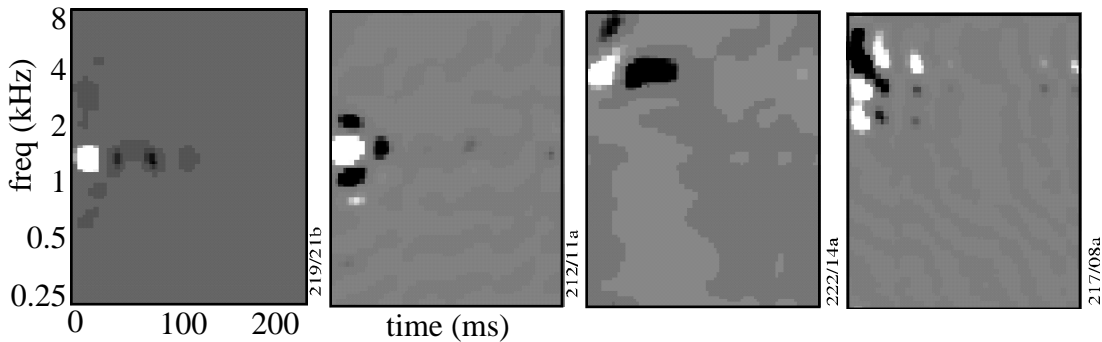

Figure 8: Examples of Spectro-Temporal Response Fields.

IV. ACKNOWLEDGEMENTS

V. REFERENCES

1. R.L. De Valois and K.K. De Valois, Spatial Vision, Oxford University Press, New-York (1988).

2. L.R. Rabiner and R.W. Schafer, Digital processing of speech signals, Prentice-Hall, New-Jersey (1978).

3. S.A. Shamma, J.W. Fleshman, P.R. Wiser and H. Versnel, Organization of response areas in ferret primary auditory cortex, J. Neurophys. **69**, 367-383 (1993).

4. J. McLean and L.A. Palmer, Organization of simple cell responses in the three-dimensional frequency domain. Vis. Neurosc. **11**, 295-306 (1994). G.C. DeAngelis, I. Ohzawa and R.D. Freeman, Receptive-field dynamics in the central visual pathways. Trends Neurosc. **18**, 451–458 (1995).

5. B.W. Andrews and D.A. Pollen, Relationship between spatial frequency selectivity and receptive field profile of simple cells, J. Physiol. (London) **287**, 163–176 (1979). S.M. Friend and C.L. Baker, Spatio-temporal frequency separability in area 18 neurons of the cat, Vision Res. **33**, 1765–1771 (1993).