Mathematical framework for place coding in the auditory system

Alex D. Reyes[1]

[1]Center for Neural Science, New York University, New York, NY, USA

Abstract

In the auditory system, tonotopy is postulated to be the substrate for a place code, where sound frequency is encoded by the location of the neurons that fire during the stimulus. Though conceptually simple, the computations that allow for the representation of intensity and complex sounds are poorly understood. Here, a mathematical framework is developed in order to define clearly the conditions that support a place code. To accommodate both frequency and intensity information, the neural network is described as a topological space with elements that represent individual neurons and clusters of neurons. A bijective mapping is then constructed from acoustic space to neural space so that frequency and intensity are encoded, respectively, by the location and size of the clusters. Algebraic operations -addition and multiplication- are derived to elucidate the rules for representing, assembling, and modulating multi-frequency sound in networks. The predicted outcomes of these operations are consistent with network simulations as well as with electrophysiological and psychophysical data. The analyses show how both frequency and intensity can be encoded with a purely place code, without the need for rate or temporal coding schemes. The algebraic operations are used to describe loudness summation and suggest a mechanism for the critical band. The mathematical approach complements experimental and computational approaches and provides a foundation for interpreting data and constructing models.

Author Summary

One way of encoding sensory information in the brain is with a so-called place code. In the auditory system, tones of increasing frequencies activate sets of neurons at progressively different locations along an axis. The goal of this study is to elucidate the mathematical principles for representing tone frequency and intensity in neural networks. The rigorous, formal process ensures that the conditions for a place code and the associated computations are defined precisely. This mathematical approach offers new insights into experimental data and a framework for constructing network models.

Introduction

Many sensory systems are organized topographically so that adjacent neurons have small differences in the receptive fields. The result is that minute changes in the sensory features causes an incremental shift in the spatial distribution of active neurons. This is has led to the notion of a place code where the location of the active neurons provides information about sensory attributes. In the auditory system, the substrate for a place code is tonotopy, where the preferred frequency of each neuron varies systematically along one axis [1]. Tonotopy originates in the cochlea [2, 3] and is inherited by progressively higher order structures along the auditory pathway [4]. The importance of a place code [5] is underscored by the fact that cochlear implants, arguably the most successful brain-machine interface, enable deaf patients to discriminate tone pitch simply by delivering brief electrical pulses at points of the cochlea corresponding to specific frequencies [6].

Although frequency and intensity may be encoded in several ways [7], there are regimes where place-coding seems advantageous. Humans are able to discriminate small differences in frequencies and intensities even for stimuli as brief as 5-10 ms [8, 9, 10, 11, 12]. Therefore, the major computations have already taken place within a few milliseconds. This is of some significance because in this short time interval, neurons can fire only bursts of 1-2 action potentials [13, 14], indicating that neurons essentially function as binary units. Therefore, it seems likely that neither frequency nor intensity can be encoded via the firing rate of individual cells since the dynamic range would be severely limited. Similarly, coding schemes based on temporal or 'volley' schemes is difficult to implement at the level of cortex because neurons can phase-lock only to low frequency sounds [15, 16, 17].

There are, however, several challenges with implementing a place coding scheme. First, the optimal architecture for representing frequency is not well-defined. Possible functional units include individual neurons, cortical columns [18, 19], or overlapping neuron clusters [20]. The dimension of each unit ultimately determines the range and resolution at which frequencies and intensities that can be represented and discriminated. Second, how both frequency and intensity can be encoded with a place code is unclear, particularly for brief stimuli when cells function mostly as binary units. Other variables such as firing rate or spike timing are commonly incorporated into models [21]. Third, the rules for combining multiple stimuli is lacking. Physiological sounds are

composed of pure tones with differing frequencies and intensities, resulting in potentially complex spatial activity patterns in networks. Finally, the role of inhibition in a place coding scheme has not been established.

Here, a mathematical model is developed in order to gain insights into: 1) the functional organization of the auditory system that supports a place coding scheme for frequency and intensity: and 2) the computations that can be performed in networks. The approach is to use principles from set theory and topology to construct the acoustic and neural spaces, find a mapping between the spaces, and then develop the algebraic operations. The predictions of the math model are then tested with simulations.

Results

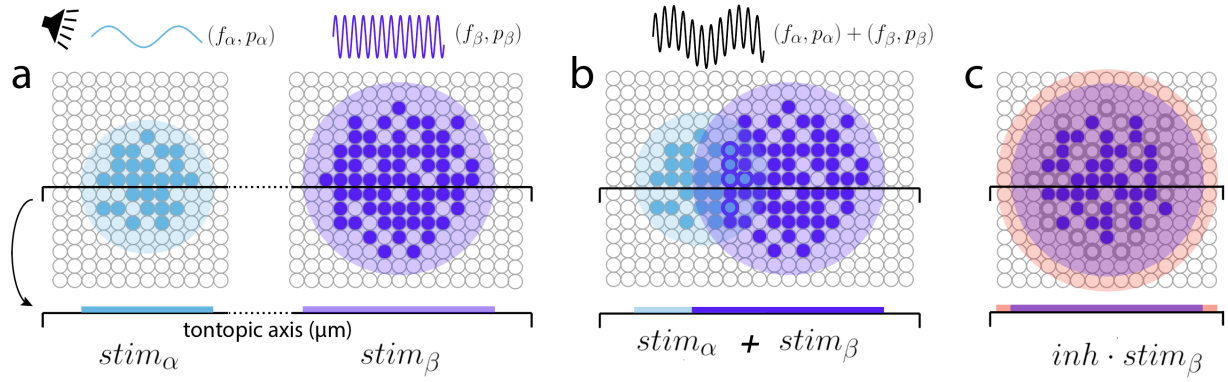


Figure 1: **Computations with a place code**. ***a, left***, hypothetical neural representation of a low frequency (f_α), small amplitude (p_α) pure tone stimulus in a two-dimensional neural network. A stimulus-evoked synaptic field covers a circular area (cyan disk) and causes a subset of cells to fire (filled circles). Projection of the synaptic field onto the tonotopic axis (cyan bar) gives the location and size of the activated area. ***right***, synaptic field generated by a tone with higher frequency (f_β) and sound pressure (p_β). ***b***, synaptic field generated by a sound composed of the two tones. ***c***, modulation by inhibition (red).

The mathematical model is subject to the following biological constraints. First, the neural network inherits the tonotopic organization of the cochlea [2, 3] so that the preferred frequency of neurons changes systematically with location along one axis. Second, a pure tone activates a population of neurons within a confined area [22], with the location of the area varying with the tone frequency. Third, the area grows with sound intensity [22], paralleling the increase in the

response area of the basilar membrane [23, 24]. The model is broadly related to the "spread-of-excitation" class of models [7, 25]

The basic computations that can be performed using a place code are shown schematically in a 2-dimensional network of neurons (Fig. 1). In response to a pure tone stimulus, a synaptic field from a presynaptic population of neurons is generated within an enclosed area of the network (a, left panel, cyan disk), causing a population of cells to fire (filled circles). A tone with a higher frequency and intensity activates a larger area at a different location (right panel). A sound composed of the two pure tones activates both regions simultaneously; the regions may overlap if the difference in frequencies is small (Fig. 1b). Finally, excitatory synaptic fields and the activated neuron clusters are modulated by inhibition (Fig. 1c). The mathematical basis for these computations is developed below. For clarity, only the main results are shown (details in Supporting Text S1).

Neural Space

Although the brain has three spatial dimensions, only one dimension -that corresponding to the tonotopic axis- is relevant for a place code. Thus, the circular synaptic field in Fig. 1 is projected onto the tonotopic axis (bars). For simplicity, the neural space is defined as a single row of neurons (Fig. 2a). In the presence of sound, afferents from an external source generates a synaptic field that covers a contiguous subset of neurons. In the following, a mathematical description of the neural space will be developed that accommodates the neural elements and synaptic fields.

The neural space is defined as an interval, bounded by minimum and maximum values x_{min} and x_{max} (Fig. 2b, magenta). To incorporate neurons, the neural space is partitioned into N_{cell} sections. Both the neural space and the cells are 'half-open' intervals, which are closed ("[") on one end and open on the other (")"). This is convenient mathematically because neurons can be packed next to each other with no space in between, allowing for a formal definition of a partition (see Supporting Text S1). Each neuron can therefore be expressed as non-overlapping subintervals of the form $[x, x + \Delta x)$, where Δx is the width of each neuron. Note that each neuron can be uniquely identified by the point at the closed end, which also gives its location along the tonotopic axis. The set containing these points is:

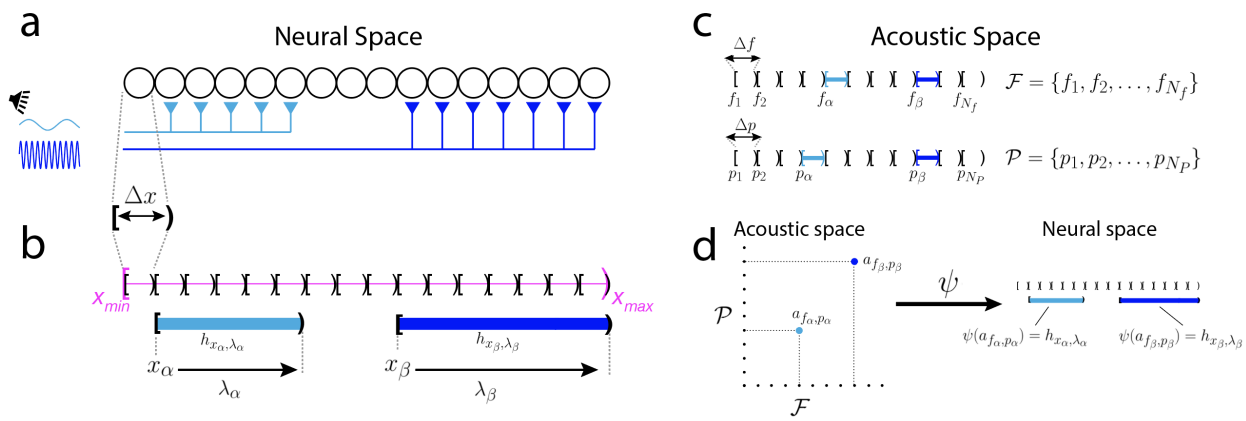


Figure 2: **Mathematical representation of neural and acoustic Spaces.** ***a***, neurons on the tonotopic axis are positioned next to each other with no space in between. Pure tone stimuli activate afferents onto a subset of neurons (cyan, blue). ***b***, mathematical representation of neural space. The tonotopic axis is a half-open interval (magenta) partitioned into smaller intervals that represent the space (Δx) taken up by neurons. The synaptic fields ($h_{x_\alpha,\lambda_\alpha}, h_{x_\beta,\lambda_\beta}$) are also half-open intervals. ***c***, The acoustic space has two dimensions, with frequency as one axis and pressure as the other. The frequency and pressure spaces are partitioned into half-open intervals of length $\Delta f, \Delta p$, respectively. ***d***, Mapping two tones in the acoustic space to intervals in neural space via a function ψ.

$$\mathcal{X} = \{x_1, x_2, \ldots, x_{N_{cell}}\} \tag{1}$$

A synaptic field spans an integral number (n_λ) of neurons and is also defined as a half-open interval with length $\lambda = n_\lambda \Delta x$. The length λ ranges from Δx (1 cell) to a maximum $\lambda_{max} = n_{\lambda,max}\Delta x$. Each synaptic interval, designated as $h_{x,\lambda} = [x, x+\lambda)$, is uniquely identified by the location of the cell at the closed end (x_α, x_β in Fig. 2b) and by its length ($\lambda_\alpha, \lambda_\beta$). The set of starting points $\mathcal{X}_\lambda$ and the set of achievable lengths Λ are given by:

$$\begin{aligned} &\mathcal{X}_\lambda = \{x_1, x_2, \ldots, x_{N_\lambda}\}, N_\lambda = N_{cell} - n_{\lambda,max} \\ &\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_{max}\} \end{aligned} \tag{2}$$

As will be shown below, $\mathcal{X}_\lambda$ will contain information about the frequency while Λ will contain information about the sound pressure.

The set of all synaptic intervals in neural space is a basis for topology on the neural space, which will be used to define algebraic operations and construct mapping from acoustic space (see below, Supporting Text S1).

Representing sound in neural space

The elements of acoustic space $\mathcal{A}$ are pure tones, each of which is characterized by a sine wave of a given frequency (f) and amplitude corresponding to sound pressure (p). Theoretically, frequencies and pressures are unbounded and can take an uncountable number of values but under physiological conditions, the audible range is likely bounded by minimum and maximum values and is finite.

The acoustic space has two dimensions with frequency as one axis and sound pressure as the other. As was done for neural space, the frequency and pressure axes are defined as a half-open intervals and divided into non-overlapping subintervals expressed as $[f, f + \Delta f)$ and $[p, p + \Delta p)$, respectively (Fig. 2c). Physiologically, Δf and Δp is loosely related the minimum frequency and pressure difference that can be discriminated (frequency and intensity difference limens, see Discussion) [8, 9, 10, 11, 12]). For example, two tones with frequencies f_α, f_β that are in the interval $[f_1, f_1 + \Delta f)$ will both be 'assigned' to f_1 (see Supporting Text S1). Therefore, the sets of audible frequencies and pressures are given by the first points of each interval:

$$\mathcal{F} = \{f_1, f_2, \ldots, f_{N_f}\} \tag{3}$$
$$\mathcal{P} = \{p_1, p_2, \ldots, p_{N_p}\}$$

The number of audible frequencies and pressures are limited by the number of synaptic intervals that fit into the tonotopic axis (N_λ in eq. 3 and the maximum number of cells that fit into a single synaptic interval ($n_{\lambda,\max}$), respectively.

Features of the acoustic space are represented in neural space via a mapping ψ (Fig. 2d; see Supporting Text S1 for formal treatment). A pure tone $a_{f,p} \in \mathcal{A}$ is mapped to an interval $h_{x,\lambda}$ by first mapping the components f to x and p to λ via $\psi_f(f) = x$ and $\psi_p(p) = \lambda$.

$$\psi(a_{f,p}) = h_{\psi_f(f),\psi_p(p)} = h_{x,\lambda} \tag{4}$$

By adjusting the number of elements in frequency and pressure spaces to match the number of synaptic intervals (N_λ) and maximum number of cells within a synaptic interval ($n_{\lambda,\max}$), respectively, the mapping becomes bijective (one-to-one, onto).

A mapping from acoustic space to intervals that are inhibitory can be similarly defined. The mapping is complicated by the fact that the number and receptive field properties of excitatory ($\boldsymbol{E}$) and inhibitory ($\boldsymbol{I}$) cells differ. The mapping is under ongoing investigation but for purposes of the present analyses, the mapping is taken to be identical to that for $\boldsymbol{E}$.

Algebraic operations with synaptic intervals

Having formally described the mathematical structure of neural space, it is now possible to define the algebraic operations -addition and multiplication- for combining and modulating synaptic intervals (Fig.1b,c). An important condition that must be satisfied is that operations be 'closed'; that is, the result obtained with addition or multiplication is an interval (or intervals) that remains in neural space. To simplify notation, the intervals will henceforth be identified with a single subscript or superscript.

To satisfy closure, 'addition' of synaptic intervals is defined as their union. Let h_α, h_β be two synaptic intervals. Then,

$$h_\alpha + h_\beta \stackrel{\text{def}}{=} h_\alpha \cup h_\beta \tag{5}$$

The addition operation yields two possible results (see Supporting Text S1 for details). If the two synaptic intervals do not overlap (h_β, h_γ in Fig. 3ai), the union yields a set with the same two intervals (Fig. 3aii, bottom). If the two intervals overlap (h_α and h_β, h_α and h_γ), the result is a single interval whose length depends on the amount of overlap (Fig. 3aii, top two traces). Note that summation is sublinear: e.g. $|h_\alpha + h_\gamma| \leq |h_\alpha| + |h_\gamma|$. Moreover, if the starting points are the same (h_α and h_β), the length is equal to that of longer summand ($|h_\alpha + h_\beta| = |h_\alpha|, |h_\alpha| \geq |h_\beta|$).

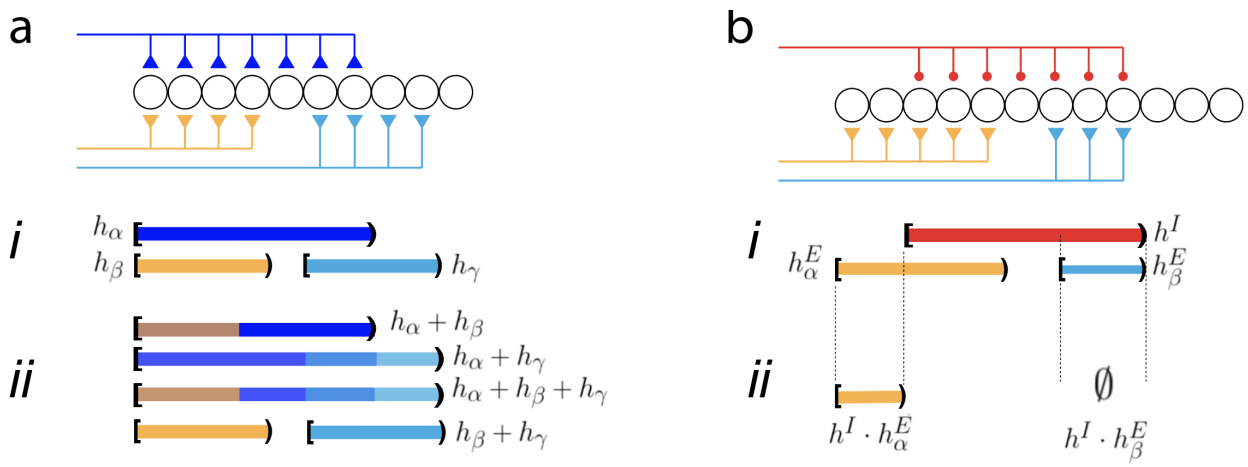


Figure 3: **Addition and Multiplication.** ***a***, schematic of network receiving three excitatory afferent inputs. ***i***, half-open intervals representing the synaptic fields. ***ii***, addition (union) of different combinations of intervals. **b**, schematic of network receiving two excitatory inputs (orange, cyan) and an inhibitory input (red). ***i***, half-open intervals associated with the activated afferents. ***ii***, multiplication (set minus) of each excitatory intervals by the inhibitory interval.

It is noteworthy that there is some ambiguity from a decoding perspective because the addition operation will fuse overlapping intervals into a larger, single interval (e.g. $h_\alpha + h_\gamma$ in Fig. 3a). It would not be possible to determine whether a synaptic interval is a result of a single high intensity pure tone, multiple low intensity pure tones with small differences in frequencies, or band limited noise (see Discussion).

Closure is satisfied because the union of an arbitrary number of overlapping intervals results in an interval that is always within the neural space. The union of any number of disjoint intervals is also within the neural space (more accurately, is a subset of the topology defined in neural space, see Supporting Text S1). A notable property that will be important later is that there is no additive inverse that can decrease the interval length.

One example of addition that occurs under biological conditions is when a pure tone arrives simultaneously to the two ears. The signal propagates separately through the auditory pathway but eventually converges at some brain region. Because each input is due to the same tone, the resultant synaptic intervals will be at same location (i.e. have the same starting points) on the tonotopic axis, though their lengths may differ because of interaural intensity differences (orange and blue intervals in Fig. 3a). Fig. 4a (top panel) shows the predicted total length when two intervals with different lengths are added (the length of the cyan interval is fixed while that of the

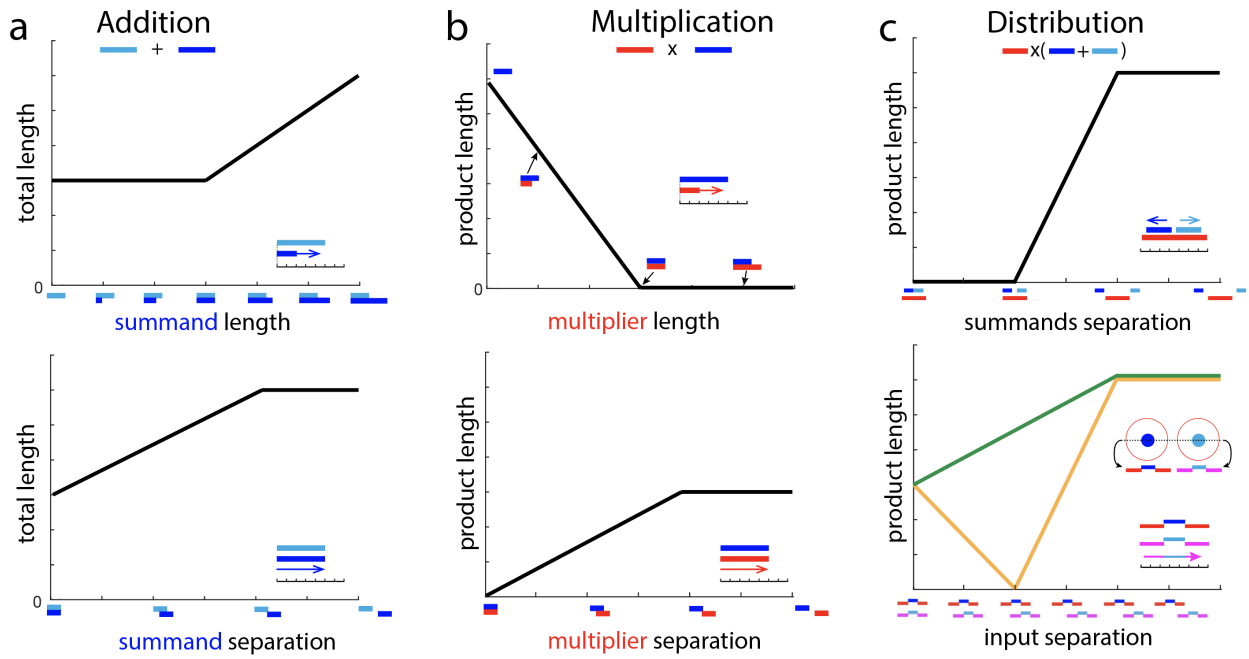


Figure 4: **Predicted effects of addition and multiplication**. ***a***, ***top***, addition of two intervals (cyan, blue) when the length of one interval (blue) is increased (top). ***bottom***, one interval (blue) is shifted to the right of the other (cyan). ***b*, top**, multiplication of an excitatory interval (blue) by an inhibitory interval of increasing length (red). ***bottom***, multiplication when the inhibitory interval is shifted to the right. ***c*, top**, effects of inhibition (red) on two excitatory intervals (blue, cyan). With the inhibitory interval at a fixed location, the distance between the excitatory intervals is increased systematically. ***bottom***, Effects of two inputs configured as center-surround where the excitatory intervals (blue, cyan) are each flanked by two inhibitory intervals (red, magenta). One of the inputs is shifted systematically to the right of the other. Predicted product length when the inputs occur simultaneously (orange) and when calculated with inputs delivered sequentially.

blue is increased). The total length is equal to the length of the longer interval: hence, it is initially constant and equal to that of the cyan interval but then increases linearly when the length of the blue interval becomes longer.

Addition also takes place when the sound is composed of two pure tones with different frequencies. This would generate two synaptic intervals with different starting points and possibly different lengths (e.g. orange and cyan intervals in Fig. 3a). The total length depends on the degree of overlap between the intervals. In Fig. 4a (bottom panel), the location of one interval (cyan) is fixed while the other (blue) is shifted rightward. When the two intervals completely overlap, the total length is equal to the length of one interval. As the blue interval is shifted, the total length increases linearly and plateaus when the two intervals become disjoint.

Inhibition decreases the excitability of the network and would be expected to reduce the size

of the synaptic interval. This is not possible with the addition operation because there is no inverse (i.e. 'subtraction' is not defined). Therefore, to incorporate the effects of inhibition, a 'multiplication' operation is introduced (Fig. 3b). Multiplication ('$\cdot$') of an excitatory synaptic interval h^E by an inhibitory interval h^I is defined as:

$$h^I \cdot h^E \stackrel{\text{def}}{=} h^E \setminus h^I \tag{6}$$

The set minus operation "\" eliminates the points from the multiplicand h^E that it has in common with the multiplier h^I (Fig. 3b*i,ii*) thereby decreasing the multiplicand's length. Note that this differs from the more common set theoretic definition of multiplication, where the product of sets is given by their intersection. Multiplication yields several results, depending on the relative locations and size of the multiplicand and the multiplier (see Supporting Text S1). If the excitatory ($\boldsymbol{E}$) and inhibitory ($\boldsymbol{I}$) intervals do not overlap, then the $\boldsymbol{E}$ interval is unaffected. If the intervals overlap (h^E_α and h^I), then the $\boldsymbol{E}$ interval is shortened ($h^I \cdot h^E_\alpha$ in *ii*). If the $\boldsymbol{E}$ interval (h^E_β) is completely within the $\boldsymbol{I}$ interval, the product is the empty set, indicating complete cancellation ($h^I \cdot h^E_\beta$ in *ii*). Multiplication can also change the starting points of the intervals and split an interval into two separate intervals (not shown, see Supporting Text S1). Multiplication satisfies the closure property as the resulting interval is a half-open interval that is always in neural space. Other algebraic properties are discussed in Supporting Text S1.

If the excitatory and inhibitory inputs are co-activated (as in feedforward circuits), then the $\boldsymbol{E}$ and $\boldsymbol{I}$ intervals will be at the same location (same starting points) on the tonotopic axis but may have different lengths. Fig. 4b (top panel) plots the predicted length of the product when the $\boldsymbol{E}$ interval is multiplied by $\boldsymbol{I}$ intervals of increasing lengths. The product length decreases and becomes zero when length of the $\boldsymbol{I}$ interval exceeds that of the $\boldsymbol{E}$ interval.

If the $\boldsymbol{E}$ and $\boldsymbol{I}$ inputs are independent of each other, the synaptic intervals could be at different locations on the tonotopic axis. Figure 4b (bottom panel) plots the length of the product when the starting point of the $\boldsymbol{I}$ interval (red) is shifted systematically to the right of the $\boldsymbol{E}$ interval (blue). The product length is zero when the $\boldsymbol{E}$ and $\boldsymbol{I}$ intervals overlap completely (separation = 0) and increases linearly as the overlap decreases, eventually plateauing to a constant value when

the intervals become disjoint.

With addition and multiplication defined, the rules for combining the two operations can now be determined. A simple case is when a network receives two excitatory inputs that results in synaptic intervals (h^E_α, h^E_β) and a single inhibitory input that result in an inhibitory interval (h^I). This scenario would occur if binaural excitatory inputs that converge in a network are then acted on by local inhibitory neurons. When all three inputs are activated simultaneously, the intervals combine in neural space as $h^I \cdot (h^E_\alpha + h^E_\beta)$. It can be shown that multiplication is *left* distributive so that (see Supporting Text S1):

$$h^I \cdot (h^E_\alpha + h^E_\beta) = (h^I \cdot h^E_\alpha) + (h^I \cdot h^E_\beta) \tag{7}$$

Intuitively, this means that the effect of a single inhibitory input on two separate excitatory inputs can be calculated by computing the inhibitory effects on each separately and then adding the results. Multiplication, however, is not *right* distributive. Thus, given two inhibitory intervals (h^I_α and h^I_β) acting on a single excitatory interval (h^E):

$$(h^I_\alpha + h^I_\beta) \cdot h^E \neq (h^I_\alpha \cdot h^E) + (h^I_\beta \cdot h^E) \tag{8}$$

Fig. 4 c (top panel) plots the predicted length when two $\boldsymbol{E}$ intervals are multiplied by an $\boldsymbol{I}$ interval as per eq. 8 ($|h^I \cdot (h^E_\alpha + h^E_\beta)|$). The two $\boldsymbol{E}$ intervals (blue, cyan) are shifted, respectively, left- and rightward relative to the $\boldsymbol{I}$ interval (red). The product length is zero as long as the two $\boldsymbol{E}$ intervals are within the $\boldsymbol{I}$ interval. When the two $\boldsymbol{E}$ intervals reach and exceed the borders of the $\boldsymbol{I}$ interval, the product length increases and reaches a plateau when the $\boldsymbol{E}$ and $\boldsymbol{I}$ intervals become disjoint.

A common physiological scenario is when sound is composed of two pure tones and each tone results in an excitatory synaptic field surrounded by an inhibitory field (center surround inhibition, Fig. 4c, bottom panel). The corresponding composite interval contains an excitatory interval that is flanked by two inhibitory intervals (inset). Letting the $\boldsymbol{I}$ -$\boldsymbol{E}$ -$\boldsymbol{I}$ interval triplet generated by each

tone be $(h^I_\alpha + h^I_\beta) \cdot h^E_\epsilon$ and $(h^I_\gamma + h^I_\delta) \cdot h^E_\zeta$, the expression when both occur simultaneously is:

$$(h^I_\alpha + h^I_\beta + h^I_\gamma + h^I_\delta) \cdot (h^E_\epsilon + h^E_\zeta) \tag{9}$$

In Fig. 4c (bottom panel, orange curve), the location of one composite interval is shifted rightward. When the composite intervals coincide (separation =0), the product length is equal to that of a single excitatory interval. With increasing separation, the product length decreases towards zero but then increases, reaching a plateau when the excitatory and inhibitory components of each composite interval no longer overlap.

It is worth noting that the effect of introducing two tones simultaneously cannot be predicted by introducing each separately and then combining the results (see Supporting Text S1 for proof). That is,

$$(h^I_\alpha + h^I_\beta + h^I_\gamma + h^I_\delta) \cdot (h^E_\epsilon + h^E_\zeta) \neq (h^I_\alpha + h^I_\beta) \cdot h^E_\epsilon + (h^I_\gamma + h^I_\delta) \cdot h^E_\zeta \tag{10}$$

The green curve in Fig. 4 c (bottom panel) is the predicted product length when the ***I*** -***E*** -***I*** triplet pairs are delivered separately and their product lengths subsequently summed. Intuitively, the curves differ because the effects of inhibition on the adjacent excitatory interval is absent; indeed, the result resembles that of adding two excitatory intervals (Fig. 4a, bottom panel). A practical implication is that the intervals due to complex sound cannot be predicted by presenting individual tones separately (see Discussion).

Simulations with spiking neurons

Key features of the mathematical model were examined with simulations performed on a 2 dimensional network model of spiking excitatory and inhibitory neurons [26]. Both ***E*** and ***I*** neuron population receive a Gaussian distributed excitatory drive from an external source (Fig. 5a); the ***E*** cells in addition receive feedforward inhibitory inputs from the ***I*** cells. Stimulation evokes Gaussian distributed excitatory inward currents in both populations and also inhibitory currents in the ***E***

cells (profiles of currents shown in insets). With brief stimuli, the recurrent connections between neurons [27] do not contribute significantly to activity in auditory cortical circuits [26] and were omitted. The region encompassing neurons that fire is henceforth referred to as the activated area (Fig. 5b, top panel). The underlying synaptic field (bottom panel) is described by the area of the network where the *net* synaptic inputs to cells exceeded (were more negative than) rheobase, the minimum current needed to evoke an action potential (I_{Rh}, inset). As defined, the synaptic field is a composite of all the inputs, both excitatory and inhibitory, that are evoked during a stimulus. Both the activated area and the synaptic field are quantified either by the diameters of circles fitted to the boundary points (magenta) or by the length of their projections to one axis (orange bar). Note that the spatial dimensions have units of *cell number* (see Methods to convert to microns).

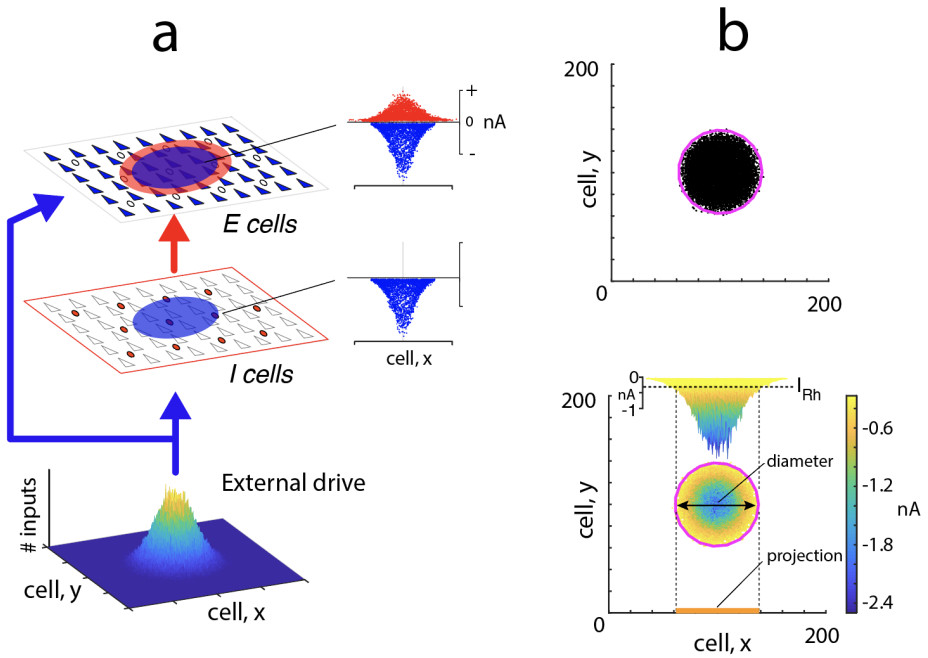


Figure 5: **Simulations with spiking neurons.** ***a***, Network consists of excitatory and inhibitory neuron populations. An external drive evokes excitatory inputs in both populations (blue disks) and inhibitory inputs to the ***E*** cells (red disk). ***insets***, profiles of excitatory (blue) and inhibitory (red) currents evoked in the ***E*** and ***I*** populations. ***b, bottom***, synaptic field evoked in the network during a stimulus. The spatial extent of the synaptic field is quantified either by the diameter of a circle fitted to its outermost points (magenta) or by the length of its projection to the tonotopic axis (orange bar). ***inset***, profile of net synaptic current generated in the ***E*** cell population. The perimeter of the synaptic field encompasses cells whose net synaptic current input exceeded rheobase (I_{Rh}). ***top***, activated area contains cells that fired action potentials (dots).

To test the addition operation, two external excitatory drives were delivered to the center of

the network simultaneously (without inhibition). Increasing the width (by increasing the standard deviation σ_α of the external drive) of one stimulus, while keeping that (σ_β) of the other fixed, increased the diameters of the synaptic field and activated areas (Fig. 6a, top panel, *i-iii*). As predicted, the diameters of the synaptic field (bottom panel, orange) and activated area (black) initially did not change but then increased as σ_α continued to widen. However, because the synaptic currents were Gaussian distributed (Fig. 5b, bottom panel), the curve started to increase before σ_α became equal to σ_β ($\frac{\sigma_\alpha}{\sigma_\beta} = 1$). When delivered simultaneously, the magnitude of the composite current increased, causing the region that exceeded rheobase to widen (Fig 6 a, top panel, compare synaptic field evoked with a single stimulus (i) to that evoked with 2 stimuli (ii)). The diameter can be calculated from the standard deviations of the two inputs (diameter $\propto \sqrt{\sigma_\alpha^2 + \sigma_\beta^2}$, not shown).

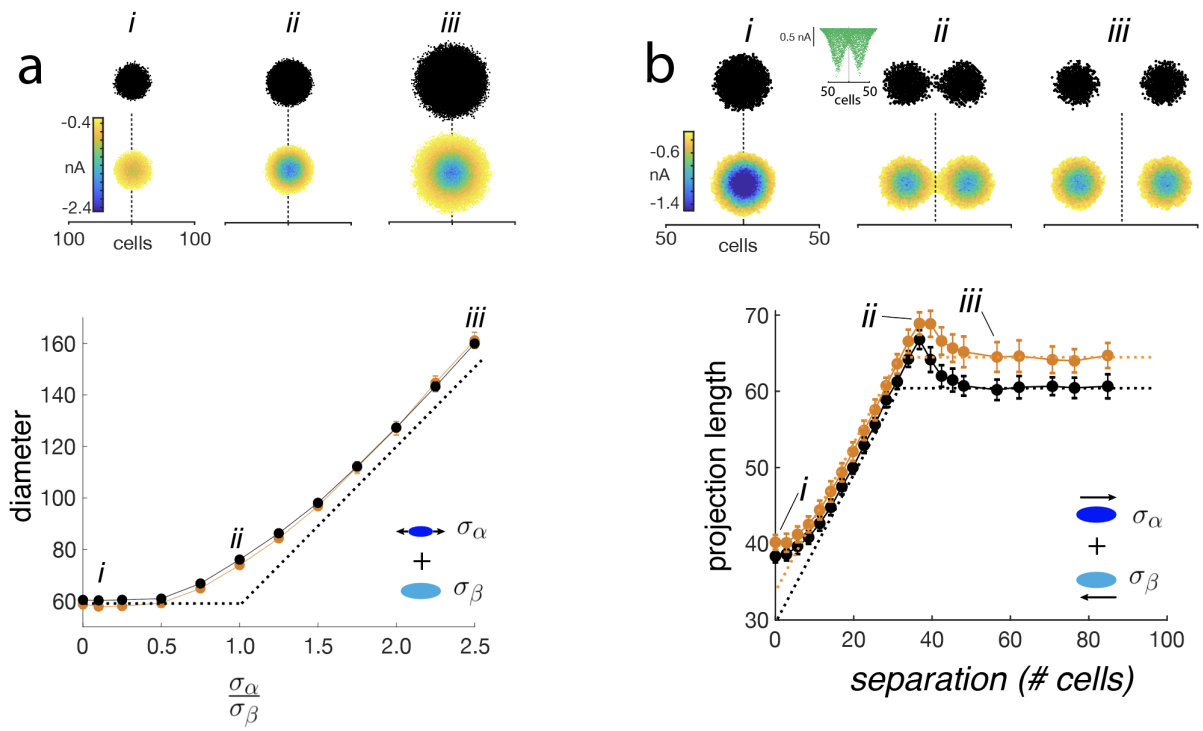


Figure 6: **Test of addition.** ***a, top***, Activated area (top; 1 sweep) and underlying synaptic field (bottom; average of 25 sweeps). ***i***, one stimulus. ***ii-iii***, two stimuli delivered to the center of the network. The width of one input was systematically increased (σ_α: 2- 45 cells) while that of the other ($\sigma_\beta = 20$ cells) was kept constant. ***bottom***, plot of synaptic field (orange) and activated area (black) diameters vs $\frac{\sigma_\alpha}{\sigma_\beta}$. Dashed curve is predicted relation. ***b***, addition of two spatially separated excitatory inputs ($\sigma_\alpha = \sigma_\beta = 10$). ***top, i - iii***, activated areas and synaptic fields with increasing stimulus separation. Inset in ***ii*** shows example of excitatory synaptic current profiles. ***bottom***, projection length vs. separation distance for synaptic field (orange) and activated area (black). Dashed curves are predicted changes.

To examine the addition of spatially disparate synaptic fields, two excitatory inputs were delivered at different distances from each other (Fig. 6b, top panel). Consistent with the prediction, the projection lengths of the synaptic field (bottom panel, orange) and activated areas (black) increased with stimulus separation and reached a plateau when the two inputs became disjoint (*iii*). The projection lengths were greater than predicted (dashed lines) when the separation was small (< 10 cells) and when the intervals were just becoming disjoint (at separation ~ 40 cells, *ii*) due to the summation properties of the Gaussian distributed inputs discussed above.

To test the multiplication operation, the $\boldsymbol{E}$ and $\boldsymbol{I}$ neurons were stimulated simultaneously, resulting in excitatory and inhibitory synaptic currents in the $\boldsymbol{E}$ cells (inset in top panel of Fig. 7a *ii*). The width of the excitatory input (σ_{exc}) was kept constant while that of the inhibitory input (σ_{inh}) was increased systematically. As predicted, the diameter of the synaptic field (bottom panel, orange) and activated area (black) decreased with increasing σ_{inh}. However, the diameter asymptoted towards a non-zero value. Because the network was feedforward, the inhibitory input was delayed relative to excitation by about 10-15 ms; as a result, there was always a time window where excitation dominated. The excitatory synaptic input was not canceled even when the inhibition was twice as wide (top panel, *iii*).

To examine multiplication of spatially disparate $\boldsymbol{E}$ and $\boldsymbol{I}$ inputs, the excitatory input was shifted systematically to the right of the inhibitory input (Fig. 7b). As predicted, the diameters of the synaptic field (bottom panel, orange) and activated area (black) increased with the $\boldsymbol{E}$ -$\boldsymbol{I}$ separation and plateaued when the $\boldsymbol{E}$ and $\boldsymbol{I}$ inputs became disjoint (*iii*).

To examine how multiplication distributes over addition, two excitatory inputs and one inhibitory input were delivered simultaneously to the network (Fig. 8a). This is the analog of the left hand side of (eq. 8). All three inputs were initially at the center and then with the inhibition stationary, the two excitatory inputs were shifted left and right (Fig. 8a, top panel, *i*-*iii*). As predicted, the projection length of the synaptic field increased towards an asymptotic value (bottom panel, orange). To reproduce the right hand side of eq. 8, simulations were performed with inhibition, first with one of the excitatory inputs and then with the other; the resultant projection lengths of each were then summed (green). As was observed with simultaneous stimulation, the projection length increased with separation. The match was poor at small separations <10 cells (*i*) and at separation of ~ 30 cells (*ii*) because the interaction between the Gaussian excitatory

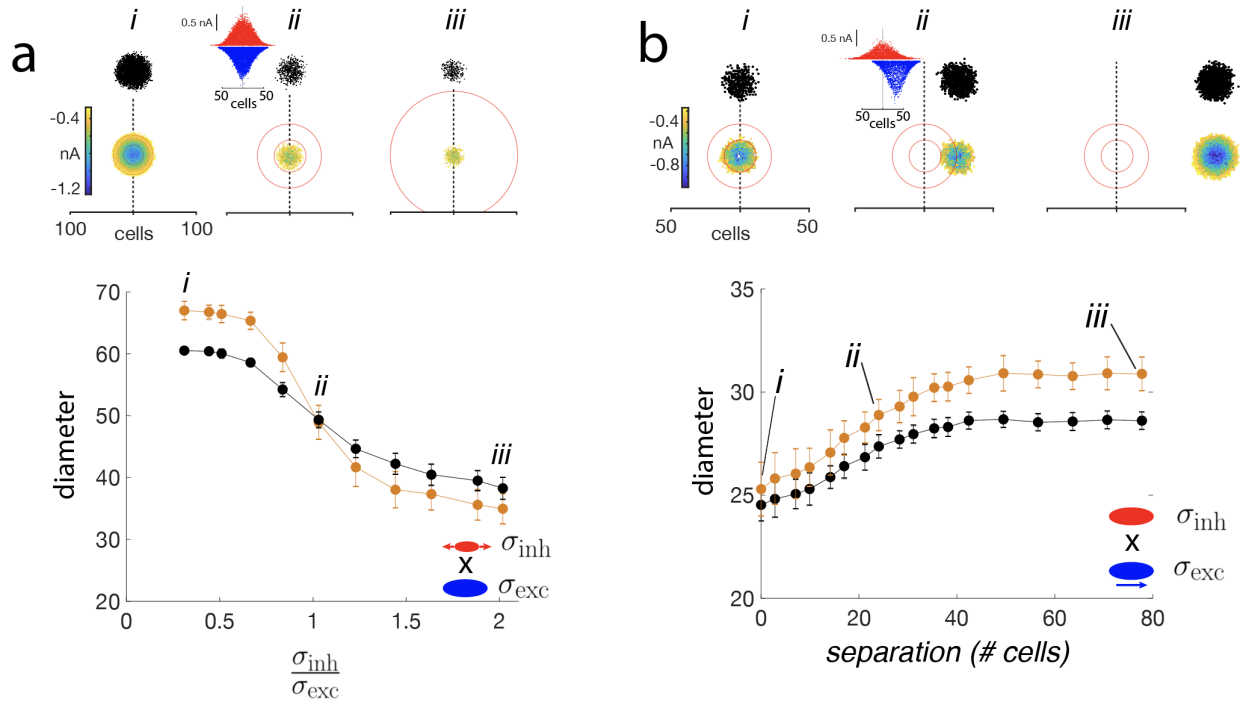


Figure 7: **Test of multiplication.** ***a, top, i-iii*** activated areas and synaptic fields evoked with excitatory ($\sigma_{exc} = 20$) and inhibitory ($\sigma_{inh} = 2 - 45$) inputs. The spatial extent of the inhibition is demarcated by the red circles (inner circle: 1 σ_{inh}; outer: 2 σ_{inh}). Inset in ***ii*** shows an example of excitatory (blue) and inhibitory (red) synaptic current profiles. ***bottom***, plot of activated area (black) and synaptic field (orange) vs $\frac{\sigma_{inh}}{\sigma_{exc}}$. ***b***, Same as in ***a*** except that the excitatory input ($\sigma_{exc} = 10$) was shifted systematically to the right of inhibition ($\sigma_{inh} = 10$). ***bottom***, Diameters of activated area (black) and synaptic field (orange) plotted against separation between excitatory and inhibitory synaptic fields

.

currents (see above) did not factor in when each input was delivered separately. The two curves were nearly identical at separations of 40-80 cells. In this range the ***E*** inputs were disjoint (as indicated by the plateauing of the excitation-only curve (dashed orange)) but still overlapped with the inhibitory synaptic field (the orange and green curves were below the excitation-only curve).

Finally, the interaction of inputs with center-surround inhibition (eq. 9) was examined by delivering two excitatory inputs, each with associated inhibitory components (inset in Fig, 8b, top panel, *iii*), to the network. The distance between the inputs were then increased systematically and the projection lengths measured (bottom panel, orange curve). At separations > 20 cells, the projection length of the synaptic field decreased to a minimum (*ii*) and then increased towards a plateau (*iii*), consistent with the prediction (Fig. 4c, bottom panel). However, at separations < 20 cells, the length increased instead of decreasing; this is likely due to the interaction of the Gaussian

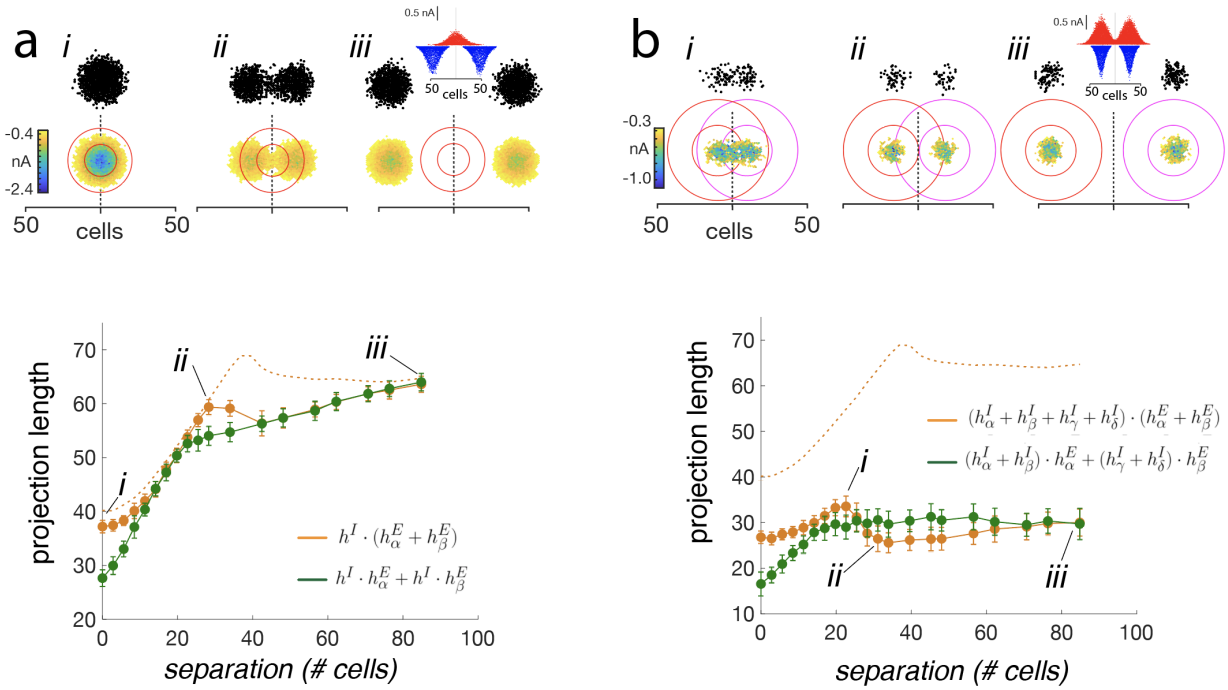


Figure 8: **Test of distributive properties.** ***a, top, i-iii*** Representative activated areas and synaptic fields generated by two excitatory inputs and a single inhibitory input ($\sigma_{exc} = \sigma_{inh} = 10$). With the location of inhibition fixed, the two excitatory inputs were separated systematically. The spatial extent of the inhibition is demarcated by the red circles (inner circle: 1 σ_{inh}; outer: 2 σ_{inh}). Inset in ***iii*** shows the excitatory (blue) and inhibitory (red) synaptic current profiles. ***bottom***, plot of synaptic field projection lengths (orange) vs separation of the excitatory inputs. Green symbols are projection lengths obtained with the sequential stimulation protocol (see text). Dotted curve is with no inhibition. ***b***, Simulations with two excitatory-inhibitory pairs, each with center-surround configuration (see inset in ***iii***). ***bottom***, legend as in ***a*** except that the green traces plot the projected lengths obtained when each input ($\sigma_{exc} = 10, \sigma_{inh} = 17$) was delivered sequentially (see text).

distributed excitatory fields discussed above. To confirm that the same result cannot be obtained by presenting each stimulus separately (right hand side of eq. 10), each ***E*** -***I*** pair was delivered sequentially and the individual projection lengths summed. Unlike with simultaneous stimulation and consistent with the prediction (green curve in Fig. 4c,bottom panel), the projection length increased monotonically to a plateau without a dip (green curve).

Discussion

The aim of this study was to develop a mathematical framework for a place-code and derive the underlying principles for how tones of varying frequencies and intensities are represented, assembled, and modulated in networks of excitatory and inhibitory neurons. The analyses are not intended to replicate the detailed aspects of biological networks but rather to clarify the minimal conditions that must be met for a viable place coding scheme, to aid in the interpretation of experimental data, and to provide a blueprint for developing computational models. The formal approach ensures that the terms of the model are defined precisely and the proposed computations are constrained strictly by well-established mathematical rules.

Place code framework in auditory processing: evidence and implications

The model has several implications with regards to auditory processing. In this section, the advantages of the place coding framework are discussed and experimental data are interpreted within the context of the mathematical framework. The formal approach is then applied a well-known psychophysical phenomena to elucidate the underlying computations and network mechanisms.

Representation of frequency and sound pressure

A key feature of the model is that the 'functional unit' of neural space is a set of contiguous neurons that have flexible borders. The associated mathematical architecture, or topology, is a collection of half-open intervals of varying lengths. The model provides a framework for encoding both frequency and intensity (or sound pressure) with a purely place-coding scheme. This is advantageous for brief stimuli where firing rate and spike timing [7, 25, 11] may not be available (see Introduction). Frequency and intensity discrimination does improve with stimulus duration, suggesting that the other variables play complementary roles in improving coding and perception [8, 9, 10, 11, 28].

A network with flexible functional units is also advantageous for maintaining both high resolution frequency and pressure representations. This can be appreciated by comparing the resolutions attainable with the classical columnar organization [18, 20] (the stimulus is assumed brief so that firing rate information is unavailable; see Introduction). In this scheme, the neural space is divided into non-overlapping columns with fixed dimensions and distinct borders. The frequency of

a stimulus is encoded by the location of the active column and sound pressure by the number of active neurons within the column (i.e. population rate code). The disadvantage is that there is a trade-off between the achievable frequency and sound pressure resolutions. Intuitively, to maximize the number of frequency levels, the columns should be as small as possible so that more can fit along the tonotopic axis; however, this reduces the number of pressure levels that can be encoded because there are fewer neurons within a column. As shown in Supporting Text S1, the number of frequency levels that can be represented is inversely proportional to the number of pressure levels ($|\mathcal{F}| = \frac{N_{cell}}{|\mathcal{P}|}$), where N_{cell} is the number of cells in the network. In contrast, for a network with flexible borders, the relation is linear: $|\mathcal{F}| = N_{cell} - |\mathcal{P}|$. Thus, compared with columns, fewer neurons are needed to have an equal number of frequency and pressure levels.

The advantage of a columnar organization is that the components of a multi-frequency stimulus remain separated in neural space. With flexible units, two intervals generated by two tones with small frequency differences and/or high intensities can fuse into a single interval. In such a case, the inverse mapping (ψ^{-1}) yields a single tone. As discussed below, ambiguities in perception of complex stimulus are more consistent with a flexible unit organization.

Relation to difference limens

The Δf, which is used to discretize acoustic space into equivalence classes, is loosely related to minimum discriminable frequency difference measured experimentally (frequency difference limen or fDL) [8, 9, 10, 11]. As defined, Δf is a constant but the measured fDL will be $0 < fDL \leq \Delta f$, depending on where the frequencies (f_α, f_β) of the test tones fall within an equivalence class. For example, given two adjacent equivalence classes $[f_0, f_0 + \Delta f)$ and $[f_1, f_1 + \Delta f)$, fDL will be small if f_α and f_β are close $f_0 + \Delta f$, near the border: initially, both f_α and f_β are classified (or 'heard') as f_0 but just a small increase in f_β will put it in the next equivalence class and be classified as f_1. If the two frequencies are closer to f_0, a larger increase in f_β is needed for the jump to occur. In principle, this prediction can be tested by presenting two indiscriminable test tones with small differences in frequencies ($\epsilon << \Delta f$); increasing both frequencies with ϵ constant should cause the tones to be discriminable at intervals of $\sim \Delta f$.

Addition operation

The addition operation applied to synaptic intervals is defined as their union: $h_\alpha + h_\beta \stackrel{\text{def}}{=} h_\alpha \cup h_\beta$. An important consequence is that if the intervals overlap, they will fuse into a single, longer interval. Under physiological conditions, this would occur if tones of a multifrequency stimulus have small differences in frequencies. This is in line with psychophysical experiments, which show that subjects perceive tones with small differences in frequencies as a single tone [29, 30] and have difficulties distinguishing the individual components of a multi-frequency stimulus [31, 32].

Another consequence is that addition of two overlapping intervals is sublinear: $|h_\alpha| + |h_\beta| > |h_\alpha + h_\beta|$. If one interval is also a subset of the other ($h_\alpha \subset h_\beta$), then the sum is equal to the larger of the two intervals: $|h_\alpha| + |h_\beta| = |h_\beta|$. This scenario would occur when binaural inputs converge onto a common site. Consistent with the prediction, electrophysiological recordings from neurons in inferior colliculus show that the FRAs (assumed to be representative of activity spread, see above) evoked binaurally is equal to larger of two responses evoked monaurally [33] (see Supporting Fig. S1). Similarly, assuming that loudness perception is linked to the length of the interval, a possible psychophysical analog is that a tone presented binaurally to a subject is perceived to be less than twice as loud as monaural stimulation [34]. The apparent sublinear effects may be due to the properties of addition operation, though inhibitory processes are also likely to contribute.

Multiplication operation and distributive properties

Multiplication of two synaptic intervals is defined as the set minus operation: $h_\alpha \cdot h_\beta = h_\beta \setminus h_\alpha$. The operation removes from the multiplicand (h_β) elements that it has in common with the multiplier (h_α), thereby shortening it. The effect of inhibition can be inferred from the FRAs of neurons. Applying GABA blockers causes the FRAs to widen [35, 36]. If, as mentioned above, the FRA can be used as a proxy for the spatial extent of activated neurons, then the result is consistent with inhibition shortening the synaptic intervals.

The manner in which multiplication distributes over addition has important implications for combining information from multiple sources. In auditory cortex, excitatory pyramidal neurons receive convergent afferent inputs from the thalamus and other pyramidal cells [37, 38]. The two afferents also appear to innervate a common set of local inhibitory neurons [27, 38]. The fact that

multiplication is left distributive (eq. 7) means that the effect can be estimated by measuring the effects of inhibition (h^I) on each excitatory inputs (h^E_{ctx}, h^E_{thal}) separately and then summing the results: $|h^I \cdot (h^E_{ctx} + h^E_{thal})| = |h^I \cdot h^E_{ctx}| + |h^I \cdot h^E_{thal}|$. However, because multiplication is not right distributive (eq. 8) a similar approach cannot be used to examine two sources of inhibition acting on a single excitatory interval. In cortical circuits, for example, excitatory neurons can be innervated by several classes of local inhibitory interneurons [39].

More generally, the representation of complex sound with a place coding scheme cannot be predicted by combining the representations of individual components if the inhibition generated by each component interact. As shown in Eq. 10 and Fig. 4c (bottom), the response of two tones presented simultaneously is not a simple combination of the responses to each tone separately. This is in line with experimental data showing that the response of neurons to pure tones and white noise stimuli differ [40]. It should be emphasized that this conclusion was derived mathematically from the distributive properties; it not trivially related to non-linearities contributed by e.g. inhibitory conductances or voltage gated channels since the model has no biophysical variables.

Transmission of acoustic information

At a more abstract level, the mathematical structures of the place coding scheme set the rules for the transfer of signals across brain regions along the auditory pathway. For example, faithful transmission of frequency and pressure information might require that the mapping be homeomorphic, which would necessitate that the topology of the source and target brain regions be the same. If the algebraic operations also need to be preserved, then the mapping should also be a homomorphism. The mathematical conditions set biological constraints on the circuitry and connection patterns of the brain regions and can be used a guide for experiments.

Algebra of loudness summation

In the auditory system, the perception of simultaneously presented tones or band limited noise depends on whether the frequencies are within the so-called critical band (CB) of frequencies [41, 42]. Figure 9a shows an experiment where subjects were asked to judge the loudness of a band limited white noise [43, 44]. As the bandwidth is increased, the perceived loudness is initially constant but then increases when the bandwidth exceeds CB (vertical line). Increasing the intensity

of the stimulus shifts the curves up but does not change CB. The origin of the CB is unclear and there is debate as to whether it is peripheral involving mainly excitatory processes [45, 46], or central, which may also recruit inhibition [47, 48, 49]. The tonotopic axis is often divided into 24 CBs, each uniquely identified by the center frequency [42]. In the following, algebraic operations are used to describe features of loudness summation and to suggest network mechanisms.

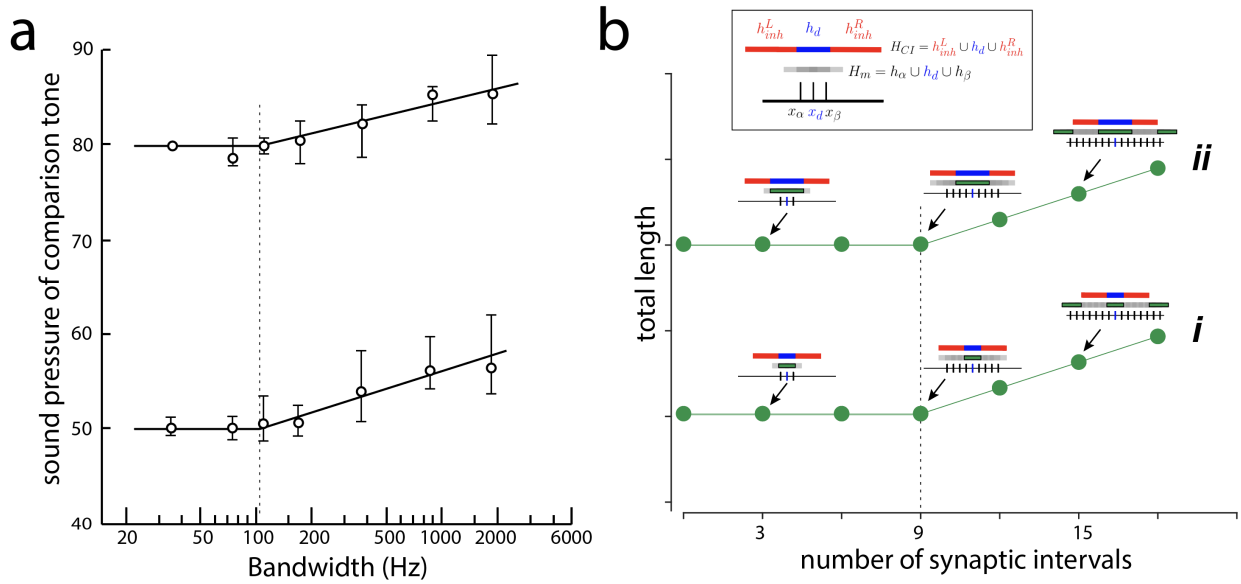


Figure 9: **Algebra of loudness summation** ***a***, Perceived loudness when subjects were presented with white noise of increasing band width (abscissa) at 2 sound pressure levels (lower and upper traces). Subjects were asked to increase the intensity of a pure tone until it was as loud as the band-limited white noise (ordinate). Dotted vertical line corresponds to critical band. Figure adapted from [44]. ***b***, Predicted interval lengths resulting from the interaction of multi-tone stimulus delivered simultaneously. ***Boxed inset***, overlapping synaptic intervals (H_m, gray) generated by stimulus with 3 frequency components. Tic marks show location of interval centers (x_α, x_d, x_β) along tonotopic axis. The dominant tone (blue) also generates two inhibitory side bands (red). Plot shows resultant length (h_l) after the operations (see text) as the number of intervals in H_m is increased (abscissa). Green bars in insets show portion of H_m that was not cancelled by inhibition. Dotted vertical line marks deviation from constant value.

A band-limited noise stimulus, or more generally a complex stimulus with multiple tones, may be expressed, after discretization, as a set of sequential frequencies: $F_m = \{f_i \mid i = (1, 2, \ldots, n); f_{i+1} - f_i = \Delta f)$. The 'bandwidth' is defined as the difference between the highest and lowest frequency components ($BW = f_n - f_1$). In neural space, the stimulus results in a set of excitatory intervals that is the union of individual excitatory intervals generated by each tonal component: $H_m = \bigcup_{i=1}^{n} h_i^E = [x_i, x_i + \lambda)$, where λ is the length of each interval and is the same for all intervals. Note

that because the frequencies for a band limited noise stimulus is a contiguous set, H_m is a single interval: $H_m = [x_i, x_n + \lambda)$.

The model assumes that in the multi-tone stimulus, one of the tones with frequency $f_d \in F_m$ is dominant and generates inhibitory intervals (h_L^I and h_R^I) that abut the excitatory interval $h_d^E \in H_m$ with no overlap ($h_d^E \cap h_L^I \cap h_R^I = \emptyset$), as in a so-called lateral inhibitory configuration (see Suppl. text for formal definitions). The other tones may also generate inhibitory intervals but they should be much smaller. Physiologically, the dominant tone may be the one with the lowest frequency, which has been shown to mask higher frequency components [30], or defined as the tone at the center of a CB [42]. The boxed inset in Fig. 9b shows the relationship between H_m (gray, composed of 3 intervals), h_d^E (blue), and the two inhibitory intervals (red). The length of the interval h_l that results from the interaction of these 4 intervals is taken to be a proxy for loudness perception and is described by (see Suppl. text for derivation and proofs):

$$\begin{aligned} h_l &= (h_L^I + h_R^I) \cdot (H_m) \\ &= h_d^E + (h_d^E + h_L^I + h_R^I) \cdot H_m \\ &= h_d^E + H_{CI} \cdot H_m \end{aligned} \tag{11}$$

where $H_{CI} = (h_d^E + h_L^I + h_R^I)$ is henceforth termed the critical interval. Intuitively, the second term on the right hand side means that portions of H_m that overlap with critical interval are cancelled by multiplication. The first term on the right is the synaptic interval generated by the dominant tone, which is part of H_m and does not overlap with the inhibition. Fig. 9b shows the result graphically when H_m is lengthened by adding more tones to the stimulus. The length is constant ($|h_l| = |h_d^E|$) until the number of components is such that H_m exceeds the boundaries of the critical interval (dotted vertical line). In this example, the deviation occurs when the number of intervals, and hence the number of frequency components, exceed 9. The CB is then $f_9 - f_1$.

Increasing the intensity of the stimulus amounts to increasing the lengths of the excitatory synaptic intervals. As shown in the supplementary text, the CB will not change provided the inhibitory interval lengths do not change. Under these conditions (see Supporting text), lengthening of intervals only affects the first term of eq. 11, resulting in an increase in baseline (upward shift

of curves) without a change in CB (compare curves ***i,ii*** in Fig. 9b).

An all-excitatory version without inhibition will not reproduce the data: the critical interval would then be h_d^E and since $h_d^E \in H_m$, h_l will be greater than h_d^E if H_m has more than one component and will grow with increasing number of tonal components. Unlike the data, the curves would have no flat region. An intermediate version is to weaken the inhibition so that the intervals in H_m are merely shortened rather than completely cancelled. The critical interval is still h_d^E but because the other intervals are reduced in length (less than $|h_d^E|$), H_m can have more than one interval and still be within the critical interval. However, the CB will be narrower than with strong inhibition.

The operations have also been used to describe a related experiment where instead of white noise, 4 tones with equally spaced frequencies were used as stimuli. Increasing the frequency difference (spacing) between the tone produced curves similar to those of Fig. 9a [44] (see supplementary Fig. S2).

Equation 11 is not intended to suggest specific circuit mechanisms but rather to elucidate the general requirements for loudness summation. While there is some evidence for a dominant tone [30] and inhibitory processes [49], the extent of e.g. inhibitory intervals is less clear and is likely to reflect the combined effect of the individual excitatory and inhibitory intervals generated by other tones in the stimulus. The precise mechanisms needs to be systematically explored with more detailed analyses, simulations, and experiments. The two assumptions -the existence of dominant tone and an associated inhibitory interval that doesn't vary with intensity- need to be verified experimentally.

Assumptions and limitations

The mathematical model is based on two salient features of the auditory system. One is that the neural space is organized tonotopically. Tonotopy has been described in most neural structures in the auditory pathway, from brainstem areas [4, 50, 51] to at least layer 4 of primary auditory cortex. Whether tonotopy is maintained throughout cortical layers is controversial, with some studies (all in mice) showing clear tonotopy [52, 53, 54, 55] and others showing a more 'salt-and-pepper' organization [56, 57, 55]. A salt-and-pepper organization suggests that the incoming afferents are distributed widely in the neural space rather than confined to a small area. The model

needs a relatively prominent tonotopy to satisfy the requirement that synaptic intervals encompass a contiguous set of cells.

A second requirement is that the size of the synaptic interval and activated area increase with the intensity of the sound. Intensity-related expansion of response areas occurs in the cochlea [23, 24, 58] and can also be inferred from the excitatory frequency-response areas (FRAs) of individual neurons. The excitatory FRAs, which document the firing of cells to tones of varying frequencies and intensities, are typically "V-shaped" (see Supporting Figure S1 for examples). At low intensities, neurons fire only when the tone frequencies are near its preferred frequency (tip of the V). At higher intensities, the range of frequencies that evoke firing increases substantially [59, 53]. If adjacent neurons have comparably-shaped FRAs but have slightly different preferred frequencies, an increase in intensity would translate to an increase in the spatial extent of activated neurons.

For mathematical convenience, the location of the synaptic intervals was identified by the left-most point (closed end) of the interval, with increases in intensity signaled by a lengthening of the interval in the rightward (high frequency) direction. Similar behavior has been observed in the cochlea albeit in the opposite direction: an increase in the intensity causes response area to increase towards low frequency region of the basilar membrane while the high frequency cutoff remains fixed [58, 24, 3]. The choice of the leftmost point to tag the interval is arbitrary and any point in the interval will suffice provided an analogous point can be identified uniquely in each interval in the set. Experimentally, using the center of mass of active neurons as the identifier might be more practical.

For simplicity, both Δf and Δp are kept constant along the tonotopic axis, which is inaccurate because the range of frequencies and sound pressure changes with frequency and sound pressure level. To represent the full ranges, the frequency and pressure can be transformed into an octave and decibel scales prior to mapping to neural space.

The algebraic operations were derived from set theoretic operations and the magnitude of the underlying synaptic inputs were irrelevant. Under biological conditions, the input magnitude determines the degree to which biophysical, synaptic, and network processes become engaged, which will affect the length of the synaptic intervals and activated areas. Not surprisingly, the results of the network simulations deviated quantitatively from the mathematical predictions in some regimes (compare Fig. 4 to Figs 6, 7). Most of the discrepancies in the simulations were because

the magnitudes of the synaptic inputs were Gaussian distributed along the tonotopic axis. In biological networks, the discrepancies may be exacerbated by the presence of threshold processes such as regenerative events [60, 61]. The underlying algebraic operations may be obscured in regimes such as these.

The model incorrectly assumes that the strength of inhibition is sufficiently strong to fully cancel excitation. This facilitated analysis because the effect of multiplication depends solely on the overlap between the multiplicand and multiplier. As the simulations with the feedforward network showed, the excitation cannot be fully canceled by inhibition owing to synaptic delay. Moreover, the balance may be spatially non-homogeneous: in center-surround suppression, excitation dominates at the preferred frequency with inhibition becoming more prominent at non-preferred frequencies [59, 35, 36]. To apply multiplication to biological systems, it may be necessary to define empirically an "effective" inhibitory field that takes into account for $\boldsymbol{E}$ -$\boldsymbol{I}$ imbalances.

Though the simulations that were used to test the analyses predictions used a network model based on cortical circuits [26, 27], the results should generalize to other network types provided the stimuli are brief (50 ms) so that cells fire only a single action potential. The mathematical model treats neurons as binary units and so only the first action potential is important. Hence, if the stimulus is brief and suprathreshold, the results obtained with a network consisting of e.g. repetitively firing cortical neurons [14, 27] or transiently firing bushy cells [62] will be qualitatively similar. The results are likely to differ with longer duration stimuli, which would allow various time- and voltage-dependent channels to become active and engage recurrent connections.

Methods and Materials

Simulations were performed with a modified version of a network model used previously [26]. Briefly, the model is a 200 x 200 cell network composed of 75% excitatory ($\boldsymbol{E}$) and 25% inhibitory ($\boldsymbol{I}$) neurons. The connection architecture, synaptic amplitudes/dynamics, and intrinsic properties of neurons were based on experimental data obtained from paired whole-cell recordings of excitatory pyramidal neurons and inhibitory fast-spiking and low threshold spiking interneurons [27]. For this study, the low-threshold spiking interneurons and the recurrent connections between the different cell types were removed, leaving only the inhibitory connections from fast spiking interneurons to

pyramidal neurons. The connection probability between the inhibitory fast-spiking cells and the excitatory pyramidal cells was Gaussian distributed with a standard deviation of 75 μm and peak of 0.4 [27].

Both $\boldsymbol{E}$ and $\boldsymbol{I}$ cells received excitatory synaptic barrages from an external source. The synaptic barrages to each cell (50 ms duration) represented the activity of a specified number of presynaptic neurons. The average number ($n_{in}(x, y)$) of inputs that each neuron at location x, y received followed a Gaussian curve so that cells at the center of the network received more inputs (Fig.5 a, bottom). For each run, the number was randomized by drawing a number from a Gaussian distribution with mean $n_{in}(x, y)$ and a standard deviation $0.25 * n_{in}(x, y)$ so that the synaptic fields and activated areas varied from trial to trial. Excitatory synaptic currents were evoked in the $\boldsymbol{E}$ and $\boldsymbol{I}$ cell populations and inhibitory synaptic currents in the $\boldsymbol{E}$ cell population after the $\boldsymbol{I}$ cells fired (insets in Fig. 5a). The spatial extents of the synaptic inputs were varied by changing the standard deviations of the external drive. In some simulations, the $\boldsymbol{E}$ and $\boldsymbol{I}$ cell populations were uncoupled and received separate inputs that could be varied independently of each other.

The neurons are adaptive exponential integrate-and-fire units with parameters adjusted to replicate pyramidal and fast spiking inhibitory neuron firing (see [26] for the parameter values). The cells fired at most one action potential when stimulated with the synaptic barrage.

The synaptic field was defined as the area of the network where the net synaptic currents to the cells exceeded rheobase, the minimum current needed to evoke an action potential in the $\boldsymbol{E}$ cells (I_{Rh}, inset in Fig. 5b, bottom panel). I_{Rh} was estimated by calculating the net synaptic current near firing threshold (V_θ): $I_{\text{net}} = g_{exc} * (V_\theta - E_{exc}) + g_{inh} * (V_\theta - E_{inh})$ where g_{exc}, g_{inh} are the excitatory and inhibitory conductances, respectively, and $E_{exc} = 0$ mV, $E_{inh} = -80$ mV are the reversal potentials. For the $\boldsymbol{E}$ cells, rheobase is approximately -0.27 nA.

The spatial extent of the synaptic field or activated area was quantified as the diameter of a circle fitted to the outermost points (maroon circles in Fig. 5b). In simulations with multiple components, the spatial extents were quantified as the total length of the projection onto the tonotopic axis (orange bar in Fig. 5b, bottom panel). The diameters and lengths have units of cell number but can be converted to microns by multiplying by 7.5 μm, the distance between $\boldsymbol{E}$ cells in the network. For all plots, the data points are plotted as mean +/- standard deviation compiled from 20-100 sweeps.

References

[1] William M Siebert. "Frequency discrimination in the auditory system: Place or periodicity mechanisms?" In: *Proceedings of the IEEE* 58.5 (1970), pp. 723–730.

[2] Georg Von Bekesy and Ernest Glen Wever. *Experiments in hearing*. Vol. 8. McGraw-Hill New York, 1960.

[3] Jozef J Zwislocki and My Nguyen. "Place code for pitch: A necessary revision". In: *Acta oto-laryngologica* 119.2 (1999), pp. 140–145.

[4] Troy A Hackett et al. "Linking topography to tonotopy in the mouse auditory thalamocortical circuit". In: *Journal of Neuroscience* 31.8 (2011), pp. 2983–2995.

[5] Brian CJ Moore and Stephan MA Ernst. "Frequency difference limens at high frequencies: evidence for a transition from a temporal to a place code". In: *The Journal of the Acoustical Society of America* 132.3 (2012), pp. 1542–1547.

[6] Robert P Carlyon et al. "Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear". In: *Journal of the Association for Research in Otolaryngology* 11.4 (2010), pp. 625–640.

[7] Bertrand Delgutte. "Physiological models for basic auditory percepts". In: *Auditory computation*. Springer, 1996, pp. 157–220.

[8] JM Doughty and WR Garner. "Pitch characteristics of short tones. II. Pitch as a function of tonal duration." In: *Journal of Experimental Psychology* 38.4 (1948), p. 478.

[9] Brian CJ Moore. "Frequency difference limens for short-duration tones". In: *The Journal of the Acoustical Society of America* 54.3 (1973), pp. 610–619.

[10] Christophe Micheyl, Li Xiao, and Andrew J Oxenham. "Characterizing the dependence of pure-tone frequency difference limens on frequency, duration, and level". In: *Hearing Research* 292.1-2 (2012), pp. 1–13.

[11] Christophe Micheyl, Paul R Schrater, and Andrew J Oxenham. "Auditory frequency and intensity discrimination explained using a cortical population rate code". In: *PLoS computational biology* 9.11 (2013), e1003336.

[12] Mary Florentine. "Level discrimination of tones as a function of duration". In: *The Journal of the Acoustical Society of America* 79.3 (1986), pp. 792–798.

[13] Brenda J Hefti and Philip H Smith. "Anatomy, Physiology, and Synaptic Responses of Rat Layer V Auditory Cortical Cells and Effects of Intracellular GABAABlockade". In: *Journal of Neurophysiology* 83.5 (2000), pp. 2626–2638.

[14] Anne-Marie M Oswald and Alex D Reyes. "Maturation of intrinsic and synaptic properties of layer 2/3 pyramidal neurons in mouse auditory cortex". In: *Journal of neurophysiology* 99.6 (2008), pp. 2998–3008.

[15] AD Reyes, EW Rubel, and WJ Spain. "Membrane properties underlying the firing of neurons in the avian cochlear nucleus". In: *Journal of Neuroscience* 14.9 (1994), pp. 5352–5364.

[16] Xiaoqin Wang et al. "Neural coding of temporal information in auditory thalamus and cortex". In: *Neuroscience* 154.1 (2008), pp. 294–303.

[17] Mounya Elhilali et al. "Dynamics of precise spike timing in primary auditory cortex". In: *Journal of Neuroscience* 24.5 (2004), pp. 1159–1172.

[18] Vernon B Mountcastle. "Modality and topographic properties of single neurons of cat's somatic sensory cortex". In: *Journal of neurophysiology* 20.4 (1957), pp. 408–434.

[19] David H Hubel and Torsten N Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". In: *The Journal of physiology* 160.1 (1962), pp. 106–154.

[20] Arnold L Towe. "Notes on the Hypothesis of Columnar Organization in Somatosensory Cerebral Cortex; pp. 32–47". In: *Brain, Behavior and Evolution* 11.1 (1975), pp. 32–47.

[21] Andrew J Oxenham. "Revisiting place and temporal theories of pitch". In: *Acoustical science and technology* 34.6 (2013), pp. 388–396.

[22] Christoph E Schreiner. "Spatial distribution of responses to simple and complex sounds in the primary auditory cortex". In: *Audiology and Neurotology* 3.2-3 (1998), pp. 104–122.

[23] Tianying Ren, Wenxuan He, and David Kemp. "Reticular lamina and basilar membrane vibrations in living mouse cochleae". In: *Proceedings of the National Academy of Sciences* 113.35 (2016), pp. 9910–9915.

[24] Monita Chatterjee and Jozef J Zwislocki. "Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch". In: *Hearing research* 111.1-2 (1997), pp. 65–75.

[25] William M Siebert et al. "Stimulus transformations in the peripheral auditory system". In: *Recognizing patterns* 104.133 (1968), pp. 602–615.

[26] Robert B Levy and Alex D Reyes. "Coexistence of lateral and co-tuned inhibitory configurations in cortical networks". In: *PLoS computational biology* 7.10 (2011), e1002161.

[27] Robert B Levy and Alex D Reyes. "Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex". In: *Journal of Neuroscience* 32.16 (2012), pp. 5609–5619.

[28] Andrew J Oxenham. "How we hear: The perception and neural coding of sound". In: *Annual review of psychology* 69 (2018).

[29] WR Thurlow and S Bernstein. "Simultaneous Two-Tone Pitch Discrimination". In: *The Journal of the Acoustical Society of America* 29.4 (1957), pp. 515–519.

[30] RLF Wegel and CE Lane. "The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear". In: *Physical review* 23.2 (1924), p. 266.

[31] Willard R Thurlow and Irving L Rawlings. "Discrimination of number of simultaneously sounding tones". In: *The Journal of the Acoustical Society of America* 31.10 (1959), pp. 1332–1336.

[32] Christophe NJ Stoelinga and Robert A Lutfi. "Discrimination of the spectral density of multitone complexes". In: *The Journal of the Acoustical Society of America* 130.5 (2011), pp. 2882–2890.

[33] Xiaorui R Xiong et al. "Interaural level difference-dependent gain control and synaptic scaling underlying binaural computation". In: *Neuron* 79.4 (2013), pp. 738–753.

[34] Michael Epstein and Mary Florentine. "Binaural loudness summation for speech presented via earphones and loudspeaker with and without visual cues". In: *The Journal of the Acoustical Socient of America* 131 (2012), pp. 3981–3988.

[35] Jian Wang, Donald Caspary, and Richard J Salvi. "GABA-A antagonist causes dramatic expansion of tuning in primary auditory cortex". In: *Neuroreport* 11.5 (2000), pp. 1137–1140.

[36] Fiona EN LeBeau, Manuel S Malmierca, and Adrian Rees. "Iontophoresis in vivo demonstrates a key role for GABAA and glycinergic inhibition in shaping frequency response areas in the inferior colliculus of guinea pig". In: *Journal of Neuroscience* 21.18 (2001), pp. 7303–7312.

[37] Heather J Rose and Raju Metherate. "Auditory thalamocortical transmission is reliable and temporally precise". In: *Journal of neurophysiology* 94.3 (2005), pp. 2019–2030.

[38] Max L Schiff and Alex D Reyes. "Characterization of thalamocortical responses of regular-spiking and fast-spiking neurons of the mouse auditory cortex in vitro and in silico". In: *Journal of neurophysiology* 107.5 (2012), pp. 1476–1488.

[39] Alex Reyes et al. "Target-cell-specific facilitation and depression in neocortical circuits". In: *Nature neuroscience* 1.4 (1998), pp. 279–285.

[40] Xiaoqin Wang et al. "Sustained firing in auditory cortex evoked by preferred stimuli". In: *Nature* 435.7040 (2005), pp. 341–346.

[41] Brian C J Moore. "Relation between the critical bandwidth and the frequency-difference limen". In: *The Journal of the Acoustical Society of America* 55.2 (1974), pp. 359–359.

[42] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics: facts and models*. Vol. 22. Springer Science & Business Media, 2006.

[43] Bertram Scharf. "Complex sounds and critical bands." In: *Psychological Bulletin* 58.3 (1961), p. 205.

[44] Eberhard Zwicker, Georg Flottorp, and Stanley Smith Stevens. "Critical band width in loudness summation". In: *The Journal of the Acoustical Society of America* 29.5 (1957), pp. 548–557.

[45] Eberhard Zwicker and Bertram Scharf. "A model of loudness summation." In: *Psychological review* 72.1 (1965), p. 3.

[46] EF Evans et al. "Comparisons of physiological and behavioural properties: auditory frequency selectivity". In: *Auditory physiology and perception*. Elsevier, 1992, pp. 159–169.

[47] JO Pickles. "Auditory-nerve correlates of loudness summation with stimulus bandwidth, in normal and pathological cochleae". In: *Hearing research* 12.2 (1983), pp. 239–250.

[48] Michael G Heinz, John B Issa, and Eric D Young. "Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment". In: *Journal of the Association for Research in Otolaryngology* 6.2 (2005), pp. 91–105.

[49] Bertrand Delgutte. "Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers". In: *The Journal of the Acoustical Society of America* 87.2 (1990), pp. 791–809.

[50] Imme Stiebler and Gunter Ehret. "Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution". In: *Journal of Comparative Neurology* 238.1 (1985), pp. 65–76.

[51] Feng Luo et al. "Three-dimensional tonotopic organization of the C57 mouse cochlear nucleus". In: *Hearing research* 257.1-2 (2009), pp. 75–82.

[52] I Stiebler et al. "The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation". In: *Journal of Comparative Physiology A* 181.6 (1997), pp. 559–571.

[53] Wei Guo et al. "Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types". In: *Journal of Neuroscience* 32.27 (2012), pp. 9159–9172.

[54] John B Issa et al. "Multiscale optical Ca2+ imaging of tonal organization in mouse auditory cortex". In: *Neuron* 83.4 (2014), pp. 944–959.

[55] Daniel E Winkowski and Patrick O Kanold. "Laminar transformation of frequency organization in auditory cortex". In: *Journal of Neuroscience* 33.4 (2013), pp. 1498–1508.

[56] Gideon Rothschild, Israel Nelken, and Adi Mizrahi. "Functional organization and population dynamics in the mouse primary auditory cortex". In: *Nature neuroscience* 13.3 (2010), p. 353.

[57] Sharba Bandyopadhyay, Shihab A Shamma, and Patrick O Kanold. "Dichotomy of functional organization in the mouse auditory cortex". In: *Nature neuroscience* 13.3 (2010), p. 361.

[58] JJ Zwislocki. "What is the cochlear place code for pitch?" In: *Acta oto-laryngologica* 111.2 (1991), pp. 256–262.

[59] ML Sutter et al. "Organization of inhibitory frequency receptive fields in cat primary auditory cortex". In: *Journal of Neurophysiology* 82.5 (1999), pp. 2358–2371.

[60] Rishidev Chaudhuri and Ila Fiete. "Computational principles of memory". In: *Nature neuroscience* 19.3 (2016), p. 394.

[61] Jeremie Barral, Xiao-Jing Wang, and Alex D Reyes. "Propagation of temporal and rate signals in cultured multilayer networks". In: *Nature communications* 10.1 (2019), pp. 1–14.

[62] Xiao-Jie Cao, Shalini Shatadal, and Donata Oertel. "Voltage-sensitive conductances of bushy cells of the mammalian ventral cochlear nucleus". In: *Journal of neurophysiology* 97.6 (2007), pp. 3961–3975.

[63] Jonathan C Horton and Daniel L Adams. "The cortical column: a structure without a function". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1456 (2005), pp. 837–862.

S1 Appendix

Acoustic Space

The general approach for defining acoustic space is as follows. The elements of acoustic space are pure tones of varying frequencies and sound pressures and may be expressed as a product of frequency space $S^{\mathcal{F}}$ and pressure space $S^{\mathcal{P}}$: $S^{\mathcal{A}} = S^{\mathcal{F}} \times S^{\mathcal{P}}$. Biologically, the range of audible and discriminable frequencies and sound pressures are limited. To reflect this, the frequency and pressure spaces will each be defined as a topological space where the underlying set is an interval but the topology is coarser than the usual topology.

The frequency space ($S^{\mathcal{F}}$) is defined as a half-open interval on the positive real line: $S^F = [f_{min}, f_{max}) \subset \mathbb{R}^+$, where f_{min}, f_{max} are the minimum and maximum audible frequencies, respectively. To account for the fact that discriminability is limited (i.e. tones of similar frequencies are perceived as the same), $S^{\mathcal{F}}$ is partitioned into N_f half-open intervals $h_f = [f_\alpha, f_\beta) = [f_\alpha, f_\alpha + \Delta f)$ where $\Delta f = \frac{1}{N_f}$ is the length of the interval and loosely represents biologically the minimum frequency difference that can be discriminated. The use of half-open intervals ensures that there is no 'gap' between adjacent frequency intervals. The set of all half-open intervals is:

$$\mathcal{H}^{\mathcal{F}} = \{h_f = [f_{min} + i\Delta f, f_{min} + (i+1)\Delta f) \mid f \in S^{\mathcal{F}}, i \in \mathbb{N}_0, i < N_f\} \tag{1}$$

The topology $\tau^{\mathcal{F}}$ is generated by the elements of $\mathcal{H}^{\mathcal{F}}$:

$$\tau^{\mathcal{F}} = \{\bigcup H_f \mid H_f \subseteq \mathcal{H}^{\mathcal{F}}\} \tag{2}$$

The set of intervals is a partition where each interval is an equivalence class $[f_j]$: for $f_\alpha, f_\beta \in S^F, f_\alpha \sim f_\beta$ if $f_\alpha, f_\beta \in [f_i, f_i + \Delta f)$. This may be extended so that $f_\alpha \sim f_\beta$ if $f_\alpha, f_\beta > f_{max}$ and $f_\alpha \sim f_\beta$ if $f_\alpha, f_\beta \leq f_{min}$. Frequencies f_α, f_β are 'assigned' to f_i if they are between f_i and $f_i + \Delta f$. Thus, f_α, f_β are indistinguishable from each other since both are classified as f_i. Each frequency interval is uniquely identified by the minimum point at the closed end. The set of these points is:

$$\begin{aligned}\mathcal{F} &= \{f_i \mid f_i = f_{min} + i\Delta f, i \in \mathbb{N}_0, i < N_f\} \subset S^{\mathcal{F}} \\ &= \{f_0, f_1, \ldots, f_{N_f-1}\}\end{aligned} \tag{3}$$

$\mathcal{F}$ is a discretization of $S^{\mathcal{F}}$ and may be thought of as a set of audible frequencies.

The sound pressure space is constructed similarly to yield:

$$\begin{aligned}&a. \quad \mathcal{H}^{\mathcal{P}} = \{h_p = [p_{min} + i\Delta p, p_{min} + (i+1)\Delta p) \mid p \in S^{\mathcal{P}}, i \in \mathbb{N}_0, i < N_p\} \\ &b. \quad \tau^{\mathcal{P}} = \{\bigcup H_p \mid H_p \subseteq \mathcal{H}^{\mathcal{P}}\} \text{ (topology of } S^{\mathcal{P}}) \\ &c. \quad \mathcal{P} = \{p_0, p_1, \ldots, p_{N_P-1}\} \subset S^{\mathcal{P}} \text{ (set of physiological pressures)}\end{aligned} \tag{4}$$

The basis for a topology in $\mathcal{A} = S^{\mathcal{F}} \times S^{\mathcal{P}}$ is:

$$\mathcal{B}^{\mathcal{A}} = \{U_f \times U_p \mid U_f \subset \mathcal{H}^{\mathcal{F}}, U_p \subset \mathcal{H}^{\mathcal{P}}\} \tag{5}$$

The topology for $\mathcal{A}$ is:

$$\tau^{\mathcal{A}} = \{\bigcup B_A \mid B_A \subseteq \mathcal{B}^{\mathcal{A}}\} \tag{6}$$

If only the sets of audible frequencies and pressures are considered, then the discretized acoustic space is:

$$\mathcal{A} = \mathcal{F} \times \mathcal{P} \tag{7}$$

with the discrete topology.

Neural Space

The neural space $S^{\mathcal{X}}$ is constructed in a similar fashion as $\mathcal{F}$ and $\mathcal{P}$, though the topology is slightly different. The idea is to partition $S^{\mathcal{X}}$ into subintervals that represent neurons and then to combine the subintervals contiguously to form synaptic intervals.

The neural space $S^{\mathcal{X}} \subset \mathbb{R}^+$ is a half-open interval representing the one dimensional tonotopic axis: $S^{\mathcal{X}} = [x_{min}, x_{max})$. $S^{\mathcal{X}}$ is partitioned into N_{cell} half-open intervals, each of width Δx and expressed as $[x_\alpha, x_\beta), x_\beta > x_\alpha$ or $[x_\alpha, x_\alpha + \Delta x)$. Each interval represents the length (Δx) of a single cell on the tonotopic axis. The set of intervals is expressed as:

$$\mathcal{H}^{\mathcal{X},1} = \{h_{x,1} = [x_{min} + i\Delta x, x_{min} + (i+1)\Delta x) \mid i \in \mathbb{N}_0, i < N_{cell}\} \tag{8}$$

The set $\mathcal{H}^{\mathcal{X},1}$ is a partition, which effectively discretizes $S^{\mathcal{X}}$ (see section above for a similar treatment). Each interval is uniquely identified by the point at the closed end, which also gives it's location along the tonotopic axis. The set $\mathcal{X}$ containing these points is:

$$\begin{aligned} \mathcal{X} &= \{x_i \mid x_i = x_{min} + i\Delta x \mid x_i \in S^{\mathcal{X}}, 0 < i < N_{cell}\} \\ &= \{x_0, x_1, \ldots, x_{N_{cell}-1}\} \end{aligned} \tag{9}$$

A synaptic interval is also a half-open interval with length λ expressed as:

$$h_{x,\lambda} = [x, x+\lambda),\ x \in \mathcal{X}, \lambda \in S^{\Lambda}$$

where $S^{\Lambda} = [\lambda_{\min}, \lambda_{\max}] \subset \mathbb{R}^+$ is the range of values that the length can take. The length of the synaptic interval is defined to be a contiguous set of cells so that $h_{x,\lambda}$ is an integral multiple of cell length Δx. Letting $n_\lambda \in \mathbb{N}^+$ be the number of cells in the interval, the intervals may be re-expressed as:

$$h_{x,\lambda} = [x, x+\lambda),\ x \in \mathcal{X}, \lambda = n_\lambda \Delta x, n_\lambda \leq n_{\lambda,max} \leq N_{cells} \tag{10}$$

where $n_{\lambda,max}$ is the maximum number of cells in the interval. The set containing all possible interval lengths is :

$$\begin{aligned} \Lambda &= \{\lambda_{n_\lambda} \mid \lambda_{n_\lambda} = n_\lambda \Delta x, n_\lambda \leq n_{\lambda,max} \leq N_{cell}\} \subset S^\Lambda \\ &= \{\lambda_1, \lambda_2, \ldots, \lambda_{max}\} \end{aligned} \tag{11}$$

As shown in the next section, sound pressure will be encoded as the length of the intervals. In order to ensure a full representation of sound pressure, a 'buffer' of $N_{cell} - n_{\lambda,\max}$ cells is needed. Without the buffer, the maximum value of last interval is 1 cell. The set of identifying points for synaptic intervals is then:

$$\mathcal{X} = \{x_1, x_2, \ldots, x_{N_\lambda}\} \subset \mathcal{X}; N_\lambda = N_{cell} - n_{\lambda,max} \tag{12}$$

The set of synaptic intervals is expressed as:

$$\mathcal{H}^{\mathcal{X},\Lambda} = \{h_{x,\lambda} = [x, x+\lambda) \mid x \in \mathcal{X}, \lambda \in \Lambda\} \tag{13}$$

The topology of $S^\mathcal{X}$ generated by the elements of $\mathcal{H}^{\mathcal{X},\Lambda}$ is:

$$\tau^\mathcal{X} = \{\bigcup H_{x,\lambda} \mid H_{x,\lambda} \subseteq \mathcal{H}^{\mathcal{X},\Lambda}\} \tag{14}$$

Mapping from acoustic space to neural space

The aim is to construct a mapping between the acoustic and neural spaces in such a way that allows for both encoding and decoding sounds in neural space.

Note that a pure tone $a_{f,p}$ corresponds to the ordered pair $(f,p) \in S^{\mathcal{F}} \times S^{\mathcal{P}}$ and an interval $h_{x,\lambda} = [x,x+\lambda)$ corresponds to the ordered pair $(x,\lambda) \in S^{\mathcal{X}} \times S^{\Lambda}$. The mapping (ψ) from acoustic to neural space is obtained by first defining mappings from frequency space $(S^{\mathcal{F}})$ to neural space $(S^{\mathcal{X}})$ and from pressure space $(S^{\mathcal{P}})$ to interval length space (S^{Λ}):

$$\begin{aligned} &\psi_f : S^{\mathcal{F}} \to S^{\mathcal{X}}, \psi_f(f) = x; \\ &\psi_p : S^{\mathcal{P}} \to S^{\Lambda}, \psi_p(p) = \lambda; \\ &\psi : S^{\mathcal{F}} \times S^{\mathcal{P}} \to S^{\mathcal{X}} \times S^{\Lambda}; \psi(f,p) = (\psi_f(f), \psi_p(p)) = (x,\lambda) \end{aligned} \tag{15}$$

After partitioning, the spaces are effectively discretized into $\mathcal{F} \subset S^{\mathcal{F}}, \mathcal{P} \subset S^{\mathcal{P}}, \mathcal{X}_\lambda \subset S^{\mathcal{X}}, \Lambda \subset S^{\Lambda}$. These are the functionally and biological relevant sets as they contain the audible frequencies and pressures that can be mapped into the compartmentalized neural space. Therefore, the mapping can be restricted to these sets:

$$\begin{aligned} &\psi_f\restriction_{\mathcal{F}} : \mathcal{F} \to \mathcal{X}, \psi_f\restriction_{\mathcal{F}}(f) = x; \\ &\psi_p\restriction_{\mathcal{P}} : \mathcal{P} \to \lambda, \psi_p\restriction_{\mathcal{P}}(p) = \lambda; \\ &\psi\restriction_{\mathcal{F},\mathcal{P}} : \mathcal{F} \times \mathcal{P} \to \mathcal{X} \times \Lambda; \psi\restriction_{\mathcal{F},\mathcal{P}}(f,p) \\ &= (\psi_f\restriction_{\mathcal{F}}(f), \psi_p\restriction_{\mathcal{P}}(p)) = (x,\lambda) \end{aligned} \tag{16}$$

The maps are bijective if $|\mathcal{F}| = |\mathcal{X}_\lambda|$ and $|\mathcal{P}| = |\Lambda|$ and continuous because the restricted spaces have the discrete topology.

Algebraic Operations in Neural Space

In neural space, addition ("+") is also defined as the union. For $h_{x_\alpha\lambda_\alpha}, h_{x_\beta\lambda_\beta} \in (S^{\mathcal{X}}, \tau^{\mathcal{X}})$

$$\begin{aligned} &+ : (S^{\mathcal{X}}, \tau^{\mathcal{X}}) \times (S^{\mathcal{X}}, \tau^{\mathcal{X}}) \to (S^{\mathcal{X}}, \tau^{\mathcal{X}}); \\ &h_{x_\alpha\lambda_\alpha} + h_{x_\beta\lambda_\beta} \stackrel{\text{def}}{=} h_{x_\alpha\lambda_\alpha} \cup h_{x_\beta\lambda_\beta} \end{aligned} \tag{17}$$

Generally, for any two synaptic intervals $(h_{x\lambda_\alpha}, h_{x\lambda_\beta})$, the addition operation yields two possi-

bilities:

$$h_{x_\alpha\lambda_\alpha} + h_{x_\beta\lambda_\beta} = \begin{cases} h_{x\lambda}, x = \min(x_\alpha, x_\beta), \lambda = |h_{x\lambda_\alpha} \cup h_{x\lambda_\beta}|, & h_{x\lambda_\alpha} \cap h_{x\lambda_\beta} \neq \emptyset \\ \{h_{x_\alpha\lambda_\alpha}, h_{x_\beta\lambda_\beta}\}, & h_{x\lambda_\alpha} \cap h_{x\lambda_\beta} = \emptyset \end{cases} \tag{18}$$

For more than 2 synaptic intervals, the total length λ is given by the Inclusion-Exclusion principle.

The algebraic structure on $(S^{\mathcal{X}}, \tau^{\mathcal{X}})$ with the addition operation is a monoid since it inherits the algebraic properties of the union operation. In particular, addition is closed because the result is either a longer, half-open interval or two disjoint half-open intervals, either of which is an element of $\tau^{\mathcal{X}}$. Addition is commutative, associative, has the empty set as the identity element, and does not have an inverse.

Multiplication "$\cdot$" of synaptic intervals is defined as the set minus operation ($\setminus$). Given two intervals $h_{x_\alpha,\lambda_\alpha}, h_{x_\beta,\lambda_\beta} \in (S^{\mathcal{X}}, \tau^{\mathcal{X}})$ where $h_{x_\alpha,\lambda_\alpha}$ is the multiplier and $h_{x_\beta,\lambda_\beta}$ is the multiplicand,

$$\begin{aligned} &\bullet : (S^{\mathcal{X}}, \tau^{\mathcal{X}}) \times (S^{\mathcal{X}}, \tau^{\mathcal{X}}) \to (S^{\mathcal{X}}, \tau^{\mathcal{X}}); \\ &h_{x_\alpha,\lambda_\alpha} \cdot h_{x_\beta,\lambda_\beta} \stackrel{\text{def}}{=} h_{x_\beta,\lambda_\beta} \setminus h_{x_\alpha,\lambda_\alpha} \end{aligned} \tag{19}$$

Letting $\hat{\lambda} = |h_{x_\beta,\lambda_\beta} \setminus h_{x_\alpha,\lambda_\alpha}|$, the possible results of multiplication are:

$$
h_{x_\alpha,\lambda_\alpha} \cdot h_{x_\beta,\lambda_\beta} =
\begin{cases}
1.\, h_{x_\beta,\lambda_\beta}, & h_{x_\alpha,\lambda_\alpha} \cap h_{x_\beta,\lambda_\beta} = \emptyset \\
2.\, [x_\beta, x_\alpha) = [x_\beta, x_\beta + \hat{\lambda}), & h_{x_\alpha,\lambda_\alpha} \cap h_{x_\beta,\lambda_\beta} \neq \emptyset, x_\beta < x_\alpha, h_{x_\alpha,\lambda_\alpha} \not\subset h_{x_\beta,\lambda_\beta} \\
3.\, [x_\delta, x_\delta + \hat{\lambda}), & h_{x_\alpha,\lambda_\alpha} \cap h_{x_\beta,\lambda_\beta} \neq \emptyset, x_\beta \geq x_\alpha, h_{x_\beta,\lambda_\beta} \not\subset h_{x_\alpha,\lambda_\alpha} \\
\quad x_\delta = x_\beta + (\lambda_\beta - \hat{\lambda}) & \\
4.\, \emptyset, & h_{x_\beta,\lambda_\beta} \subset h_{x_\alpha,\lambda_\alpha} \\
5.\, \{[x_\beta, x_\alpha), [x_\alpha + \lambda_\alpha, x_\beta + \lambda_\beta)\}, & h_{x_\alpha,\lambda_\alpha} \subset h_{x_\beta,\lambda_\beta}
\end{cases}
\tag{20}
$$

The algebraic structure on $(S^{\mathcal{X}}, \tau^{\mathcal{X}})$ with the multiplication operation is a magma, with properties inherited from the set minus operation. Multiplication is closed because the set minus operation yields an interval that is either the empty set or a half-open interval(s) with a minimum length(s) of Δx and maximum length of λ_{max}. In either case, the resulting interval is in $(S^{\mathcal{X}}, \tau^{\mathcal{X}})$. While the $\emptyset$ is a left identity element, there is no unique inverse element since the result is the $\emptyset$ whenever the length of multiplier is greater than or equal to that of the multiplicand. It is also neither commutative nor associative.

Proposition 0.0.1. The *multiplication* operation is left but not right distributive over addition.

Proof. **Left distributive**: Let $h_{x_\alpha,\lambda_\alpha}, h_{x_\beta,\lambda_\beta}, h_{x_\gamma,\lambda_\gamma} \in (S^{\mathcal{X}}, \tau^{\mathcal{X}})$

$$
\begin{aligned}
&= h_{x_\gamma,\lambda_\gamma} \cdot (h_{x_\alpha,\lambda_\alpha} + h_{x_\beta,\lambda_\beta}) \\
&= (h_{x_\alpha,\lambda_\alpha} \cup h_{x_\beta,\lambda_\beta}) \setminus h_{x_\gamma,\lambda_\gamma} \\
&= h^c_{x_\gamma,\lambda_\gamma} \cap (h_{x_\alpha,\lambda_\alpha} \cup h_{x_\beta,\lambda_\beta}), \text{ 'c' is complement} \\
&= (h^c_{x_\gamma,\lambda_\gamma} \cap h_{x_\alpha,\lambda_\alpha}) \cup (h^c_{x_\gamma,\lambda_\gamma} \cap h_{x_\beta,\lambda_\beta}) \\
&= h_{x_\alpha,\lambda_\alpha} \setminus h_{x_\gamma,\lambda_\gamma} \cup h_{x_\beta,\lambda_\beta} \setminus h_{x_\gamma,\lambda_\gamma} \\
&= h_{x_\gamma,\lambda_\gamma} \cdot h_{x_\alpha,\lambda_\alpha} + h_{x_\gamma,\lambda_\gamma} \cdot h_{x_\beta,\lambda_\beta}
\end{aligned}
$$

Not Right distributive:

$$
\begin{aligned}
&(h_{x_\alpha,\lambda_\alpha} + h_{x_\beta,\lambda_\beta}) \cdot h_{x_\gamma,\lambda_\gamma} \\
&= h_{x_\gamma,\lambda_\gamma} \setminus (h_{x_\alpha,\lambda_\alpha} \cup h_{x_\beta,\lambda_\beta}) \\
&= (h_{x_\gamma,\lambda_\gamma} \setminus h_{x_\alpha,\lambda_\alpha}) \cap (h_{x_\gamma,\lambda_\gamma} \setminus h_{x_\beta,\lambda_\beta}) \\
&\neq h_{x_\alpha,\lambda_\alpha} \cdot h_{x_\gamma,\lambda_\gamma} + h_{x_\beta,\lambda_\beta} \cdot h_{x_\gamma,\lambda_\gamma}
\end{aligned}
$$

□

Proposition 0.0.2. Let the representation of a pure tone in neural space be an excitatory interval (h^E_α) flanked by two inhibitory intervals (h^I_α, h^I_β) and that of another pure tone be also an excitatory (h^E_β) flanked by two inhibitory intervals (h^I_γ, h^I_δ), then the combined effect when both tones are introduced simultaneously is not equal to presenting each tone separately and then combining the result. That is, $(h^I_\alpha + h^I_\beta + h^I_\gamma + h^I_\delta) \cdot (h^E_\alpha + h^E_\beta) \neq (h^I_\alpha + h^I_\beta) \cdot h^E_\alpha + (h^I_\gamma + h^I_\delta) \cdot h^E_\beta$.

Proof. $(h^I_\alpha + h^I_\beta + h^I_\gamma + h^I_\delta) \cdot (h^E_\alpha + h^E_\beta)$ can be re-expressed as:

$$
\begin{aligned}
&(h^E_\alpha \cup h^E_\beta) \setminus (h^I_\alpha \cup h^I_\beta \cup h^I_\gamma \cup h^I_\delta) \\
&\quad = (h^I_\alpha \cup h^I_\beta \cup h^I_\gamma \cup h^I_\delta)^c \cap (h^E_\alpha \cup h^E_\beta)) \text{ where “c” is the complement}
\end{aligned}
$$

Using a similar argument, $(h^I_\alpha + h^I_\beta) \cdot h^E_\alpha + (h^I_\gamma + h^I_\delta) \cdot h^E_\beta$ may be re-expressed as:

$$
\begin{aligned}
&(h^E_\alpha \setminus (h^I_\alpha \cup h^I_\beta)) \cup (h^E_\beta \setminus (h^I_\gamma \cup h^I_\delta)) \\
&= ((h^I_\alpha \cup h^I_\beta)^c \cap h^E_\alpha) \cup ((h^I_\gamma \cup h^I_\delta)^c \cap h^E_\beta)
\end{aligned}
$$

For an element $y \in ((h^I_\alpha \cup h^I_\beta)^c \cap h^E_\alpha)$, the following statements are true:

$$1. y \in (h^I_\alpha \cup h^I_\beta)^c$$

$$2. y \in (h^I_\alpha \cup h^I_\beta)^c \cup ((h^I_\gamma \cup h^I_\delta)^c \cap h^E_\beta)$$

$$3. y \in (h^I_\alpha \cup h^I_\beta)^c \cup (h^I_\gamma \cup h^I_\delta)^c$$

$$4. y \in h^E_\alpha$$

$$5. y \in h^E_\alpha \cup ((h^I_\gamma \cup h^I_\delta)^c \cap h^E_\beta)$$

$$6. y \in h^E_\alpha \cup h^E_\beta$$

Since $y \in ((h^I_\alpha \cup h^I_\beta)^c \cap h^E_\alpha)$, then from lines 3 and 6, $y \in ((h^I_\alpha \cup h^I_\beta)^c) \cup ((h^I_\gamma \cup h^I_\delta)^c) \cap (h^E_\alpha \cup h^E_\beta)$. Therefore,

$$((h^I_\alpha \cup h^I_\beta)^c \cap h^E_\alpha) \cup ((h^I_\gamma \cup h^I_\delta)^c \cap h^E_\beta) = ((h^I_\alpha \cup h^I_\beta)^c) \cup ((h^I_\gamma \cup h^I_\delta)^c) \cap (h^E_\alpha \cup h^E_\beta)$$

$$= ((h^I_\alpha \cup h^I_\beta) \cap (h^I_\gamma \cup h^I_\delta))^c \cap (h^E_\alpha \cup h^E_\beta)$$

$$\neq (h^I_\alpha \cup h^I_\beta \cup h^I_\gamma \cup h^I_\delta)^c \cap (h^E_\alpha \cup h^E_\beta))$$

Therefore, $(h^I_\alpha + h^I_\beta + h^I_\gamma + h^I_\delta) \cdot (h^E_\alpha + h^E_\beta) \neq (h^I_\alpha + h^I_\beta) \cdot h^E_\alpha + (h^I_\gamma + h^I_\delta) \cdot h^E_\beta$ $\square$

Comparison with columnar organization

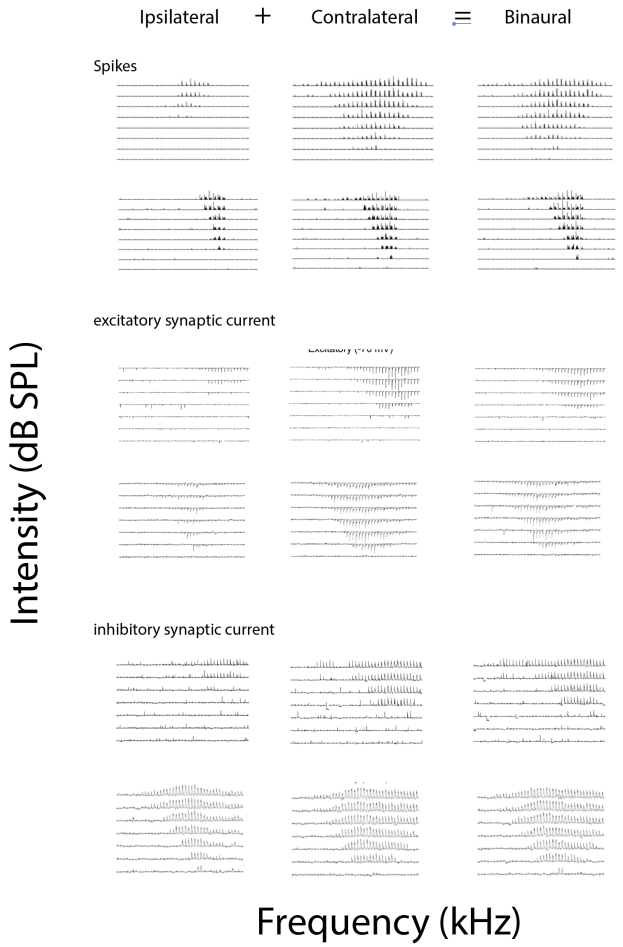
In the mathematical model, the 'functional' unit is an interval that can vary in size, depending on the sound intensity. The indistinct borders are consistent with topographical organization of afferents [20]. The expanding border facilitates high resolution representation of both frequency and pressure. Assuming a bijective mapping between acoustic and neural spaces, the maximum number of frequency levels that can be represented is proportional to the number of cells along the tonotopic axis (N_{cell} and the maximum number of pressure levels is proportional to the maximum number of cells in each interval ($n_{\lambda,max}$). The relation between the number of frequency and pressure levels is then given by $|\mathcal{F}| = |\mathcal{A}| - |\mathcal{P}| = N_{cell} - n_{\lambda,max}$ (see eq. 12 in appendix). Hence, the number of frequency levels decreases linearly with the number of pressure levels.

In contrast, consider an alternative neural space that consists of classical columns [18]. The definition of a column has evolved [63] but for illustrative purposes, columns are defined here as having distinct borders; the enclosed neurons have identical receptive fields and hence function as a single unit. In this scheme, the location of the active column encodes frequency and sound pressure is encoded as the number of neurons that become active during the stimulus (assuming a brief stimulus so that neurons are effectively binary).

Construction of the neural space is as above except that the flexible synaptic intervals are replaced with half-open intervals, termed columns, spanning a *fixed* number of cells n_{col}. The number of columns that fit into the neural space with N_{cell} neurons is therefore $N_{col} = \frac{N_{cell}}{n_{col}}$. Assuming a bijective mapping from frequency space to neural space, the number of frequency levels $N_{\mathcal{F}} = N_{col} = \frac{N_{cell}}{n_{col}}$. Because the number of cells in a column (n_{col}) is fixed, pressure will be encoded as the number of cells that become active within a column during a stimulus. Assuming a bijective map, the number of pressure levels is $N_{\mathcal{P}} = n_{col}$.

Hence, the number of frequency levels is inversely proportional to the number of pressure levels: $N_{\mathcal{F}} = \frac{N_{cell}}{n_{col}} = \frac{N_{cell}}{N_{\mathcal{P}}}$. Compared with the model with flexible synaptic intervals, the number of frequency levels drops off much faster with the number of pressure levels.

Experimental evidence



Supporting Fig. S1: Experimental evidence for Addition operation Data adapted from reference [33]. Extracellular (top) and whole-cell voltage clamp (middle, bottom) recording from the central nucleus of the Inferior colliculus, which receives binaural inputs. The frequency response areas (FRA) were constructed by presenting pure tones of varying frequencies (abscissas) and intensities (ordinates) and documenting the responses. Tones were delivered ipsilaterally (left column), contralaterally (middle), and binaurally (right). The FRAs can be taken as an indicator of the spatial extent of synaptic field and activated area in the following manner. Note that the FRAs are "V-shaped". At low intensities, neurons fire only when the tone frequencies are near its preferred frequency (tip of the V). At higher intensities, the range of frequencies that evoke firing increases substantially . If adjacent neurons have comparably-shaped FRAs but have slightly different preferred frequencies, an increase in intensity would translate to an increase in the spatial extent of activated neurons. The model predicts that because the contralateral response is greater than the ipsilateral response, the response evoked with binaural stimulation should equal that of contralateral stimulation. In all cases, the binaural FRAs was nearly identical to the contralateral FRAs. *Permission pending.*

Algebra of loudness summation

A multi-frequency stimulus is described as a set of frequencies $F_m = \{f_i, f_{i+1}, \ldots, f_n\} \subset \mathcal{F}$. In neural space, the stimulus generates a set of excitatory intervals H_m that is the union of the individual intervals generated by each tonal component: $H_m = \bigcup_{i=1}^{n}(h_i^E = [x_i, x_i + \lambda)) \subset \mathcal{H}^{\mathcal{X}}$, where λ is the length of each interval.

Definition 0.0.1. A tone with frequency f_d in multi-frequency stimulus is 'dominant' if, in addition to an excitatory interval h_d^E, it generates an inhibitory interval h^I that is larger than those generated by others in neural space.

Physiologically, the dominant frequency f_d could be the center frequency that identifies a CB [42] or the lowest frequency tone of a complex stimulus [30]. It is assumed that every multifrequency stimulus has a dominant frequency, which generates the synaptic interval $h_d^E \subset H_m$.

Definition 0.0.2. The inhibitory interval is in the lateral inhibitory configuration if $h^I \cap h_d^E \neq \emptyset$ and there exists a half-open interval $\hat{h} \supset (h_d^E \cup h^I)$ so that if $x \in \hat{h}$, then $x \in h_d^E \cup h^I$. This ensures that the excitatory and inhibitory intervals abut each other with no space in between. Note that this definition applies whether h^I is to one side of h_d^E or is split into two that occupies either side (as in the main text) $h^I = h_L^I \cup h_R^I$.

The interaction of h_d^E and H_m described by the algebraic operation: $h_l = h^I \cdot H_m$. The length $|h_l|$ is taken to be the proxy for loudness perception.

Proposition 0.0.3. In a lateral inhibition configuration, increasing the size of H_m does not increase h_l as long as H_m remains within (is a subset of) a 'critical interval' h_{CI}.

Proof. H_m can be divided in 3 disjoint sets according to their relation to h_d^E and h^I.

$$
\begin{aligned}
H_m &= h_d^E \cup h^I \cup H_{out} \text{ where} \\
&\quad h_d^E = H_m \cap h_d^E = h_d^E \qquad \text{since } h_d^E \in H_m \\
&\quad h^I = H_m \cap h^I \\
&\quad H_{out} = H_m \setminus (h_d^E \cup h^I) \\
\Rightarrow h_l &= h^I \cdot H_m = h^I \cdot h_d^E + h^I \cdot h^I + h^I \cdot H_{out} \\
&= h_d^E + \emptyset + H_{out} = h_d^E + H_{out} \\
&= h_d^E + (h_d^E + h^I) \cdot H_m
\end{aligned}
\tag{21}
$$

Thus, h_l is equal to h_d^E as long as $H_{out} = \emptyset$ or when $H_m \subseteq (h_d^E \cup h^I)$. The critical interval is therefore $h_{CI} = h_d^E \cup h^I$. □

Definition 0.0.3. The critical band (CB) is difference between the highest frequency and lowest frequency in F_m that satisfies $H_m \subseteq h_{CI}$.

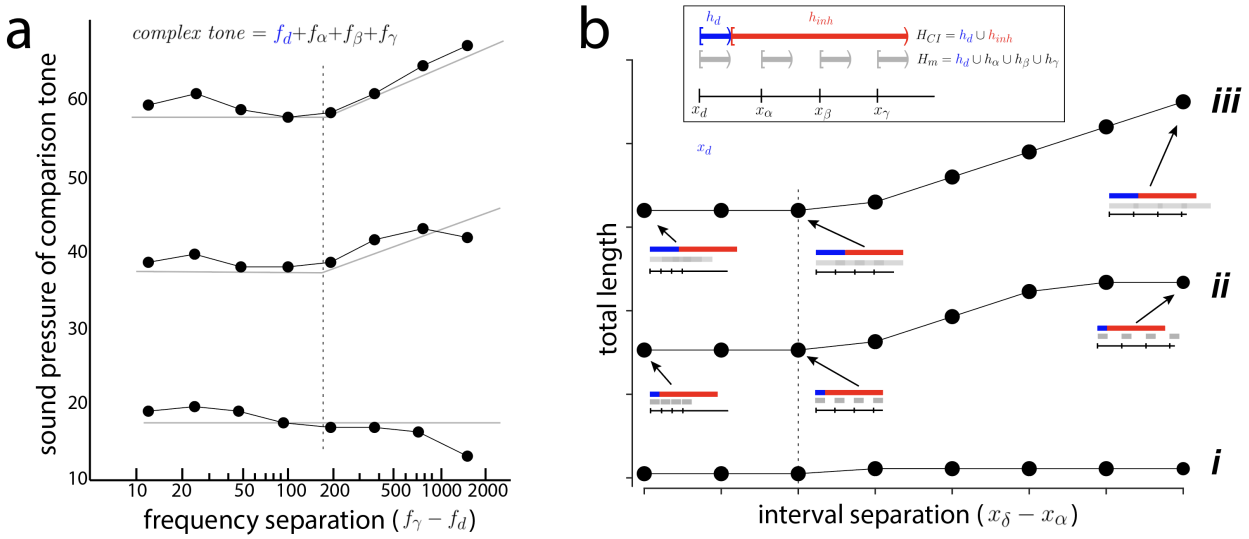
Proposition 0.0.4. The critical band does not change with stimulus intensity. In the following, an increase in stimulus intensity is represented by an increase in the length of each excitatory interval by an amount Δh.

Proof. Let $h_{CI} = (h_d^E \cup h^I)$, which is a single interval $[x_\alpha, x_\alpha + \lambda_{d,inh}), \lambda_{d,inh} = |h_d^E \cup h^I|$ if in the lateral inhibition configuration. As per above, $H_m = \bigcup_{i=1}^{n}(h_i = [x_i, x_i + \lambda))$. For $H_m \subseteq (h_d^E \cup h^I)$ to be true, the starting point of the first interval $x_1 \geq x_\alpha$ and the starting point x_n of the last point must be such that $x_n + \lambda \leq x_\alpha + \lambda_{d,inh}$. In frequency space, x_1 is associated with a frequency f_1 and x_n with f_n so the CB is $f_n - f_1$.

Increasing the lengths all of the excitatory intervals by Δh effectively increases $h_d^E \cup h^I$ also by Δh. For example, the union of two successive intervals $h_1 = [x_1, x_1 + \lambda), h_2 = [x_2, x_2 + \lambda)$ is $[x_1, x_1 + \lambda) \cup [x_2, x_2 + \lambda) = [x_1, x_2 + \lambda)$. Increasing each interval by Δh results in $[x_1, x_1 + \lambda + \Delta h) \cup [x_2, x_2 + \lambda + \Delta h) = [x_1, x_2 + \lambda + \Delta h)$. Therefore, the condition for starting point of the first interval is still $x_1 \geq x_\alpha$ and that for the starting point of the last interval is $x_n + \lambda + \Delta h \leq x_\alpha + \lambda_{d,inh} + \Delta h$. Hence, n does not change. If H_m is due to a set of frequencies $F_m = \{f_1, f_2, \ldots, f_n\}$, then then CB

is still $f_n - f_1$. □

The total length (h_l) resulting from the operations are shown graphically for band-limited white noise with increasing bandwidth (main text) and for a complex stimulus with 4 tones with increasing spacing between frequencies (Supporting Fig. S2)



Supporting Figure S2: Algebra of loudness summation ***a***, Perceived loudness when subjects were presented with a complex tone with 4 frequency components spaced evenly apart at 3 sound pressure levels. Subjects were asked to increase the intensity of a single tone (ordinate) until it was as loud as the complex tone. The frequency separation was systematically increased (abscissa). Dotted vertical line corresponds to critical band. Figure adapted from [44]. ***b***, Predicted interval lengths resulting from the interaction of 4 tones delivered simultaneously (inset). ***Boxed inset***, The tones generate 4 excitatory intervals (H_m, gray) one of which is dominant ($h_d^E, blue$). The dominant tone also generates an inhibitory synaptic interval (red). Plot shows resultant length (h_l) after the operations as the frequency spacing between intervals in H_m is increased (abscissa). Dotted vertical line marks deviation from constant value.