Diversity within a Birdsong

Rodrigo Laje and Gabriel B. Mindlin*
Departamento de Física, FCEN, UBA, Ciudad Universitaria, Pabellon I (1428), Buenos Aires, Argentina
(Received 28 May 2002; published 27 December 2002)

We present a model for the activities of neural circuits in a nucleus found in the brains of songbirds: the robust nucleus of the archistriatum (RA). This is a fore brain song control nucleus responsible for the phasic and precise neural signals driving vocal and respiratory motor neurons during singing. Driving a physical model of the avian vocal organ with the signals generated by the neural model, we produce synthetic songs. This allows us to show that certain connectivity architectures in the RA give rise to a wide range of different vocalizations under simple excitatory instructions.

DOI: 10.1103/PhysRevLett.89.288102 PACS numbers: 87.19.–j, 43.70.+i

The songs of birds constitute some of the richest sounds in nature. Beyond their beauty, the songs of songbirds (some 4000 species of birds out of the approximately 9000 known to exist) are a test bench for studying learned behavior. In fact, songbirds develop their songs in a way which has many parallels with the acquisition of speech by humans. For this reason, extensive research has been carried out in order to identify the location and functions of different areas of the brain in songbirds (for a subset of a rich literature, [1–4]). These studies revealed the structure of the basic neural pathway for the control of song. The ultimate role of this complex pathway is to generate the appropriate activities of the muscles controlling the avian vocal organ (syrinx) and the muscles involved in respiration in order to produce a given vocalization.

The syrinx generates sound through flow-induced oscillation of the lateral labia, a bilateral structure of tissue folds which open and close the air passage from the bronchi to the trachea. Recently, it was shown that a large variety of song elements can be generated by the avian vocal organ by means of slight changes in a generic gesture of two driving parameters: bronchial pressure and fold tension [5,6]. In order to produce the characteristic repetitive elements in a song, both the driving bronchial pressure and the tension of the folds have to be swept cyclically (at a much slower rate than the self-sustained oscillation of the folds). The basic quantity determining the sound features of the resulting song element is the phase difference between the cycles of these two driving parameters [5]. Since a song is made out of a wide variety of elements, the question naturally arises of how does the precise sequence of phase differences emerge from the neural architecture of the motor pathway.

In this work, we build a model for the activities of a neuron subpopulation in the robust nucleus of the archistriatum (RA), one of the major nuclei of the vocal motor pathway. It is in the RA that the key phase difference between instructions driving vocal and respiratory motor neurons is established [7]. Our model, based on a detailed neurological study performed by Spiro and co-workers [7], allows us to show that simple instructions from excitatory afferents to the RA nucleus are able to produce, for a given architecture of connections, a rich variety of elements. Moreover, we show that some instructions are able to generate sets of elements (syllables).

The brain structure which controls the production of birdsong in songbirds is made out of discrete sets of neurons called nuclei, and axons that project into other nuclei forming a pathway [8]. The pathway runs from the high vocal center (HVc) to the RA nucleus. In the RA there are neurons projecting to the motor neurons innervating the muscles of the syrinx, and neurons that project to areas that control respiration [3,7,9,10]. In fact, the bursting activity of the neurons at RA allows the prediction of element identity [4,11]. Recently, experiments in singing birds revealed that individual HVc neurons projecting onto RA neurons present sparse bursts. Interneurons in HVc, on the other hand, present activities of tonic nature, representative of the population activity of RA-projecting HVc neurons [12].

Recently, the structure of RA was described in detail [7]. It was reported that besides the excitatory projection neurons described above, there are long-range inhibitory interneurons that project widely throughout RA without exiting the nucleus. These inhibitory interneurons are believed to provide the inhibition to projection neurons needed to create the patterns of bursts and pauses characteristic of RA during singing [7]. It is known that neurons in HVc are active during particular syllables [13], probably exciting some specific subpopulations in RA. But which is the range of possible elements that a given subpopulation of RA neurons can generate?

In order to address this issue, we build a model for the activities of a subpopulation of RA neurons. This subpopulation consists of three sets of neurons, where each set contains several neurons behaving similarly. Two of these sets contain excitatory neurons, while the third set consists of long-range inhibitory neurons. By activity we denote the average number of action potentials per unit time of the neurons behaving similarly within a set. The variables of this model are x_k, x_p, and y, which stand for

the activities of the two sets of excitatory neurons controlling (through the paths described above) the syringeal muscles and respiratory muscles (x_k and x_p, respectively), and the activity of the set of inhibitory interneurons (y). Clearly, a description of the dynamics within this nucleus in terms of averaged activities will miss the details of acoustic features determined by the actual bursting [11] of individual RA neurons. Yet, many of the features of the note are determined by the relative phases of the cyclic parameters driving pressure and tension of labia in the syrinx [5], and this is the scale of description at which we perform our study.

In order to relate the variables in our description, we write the following additive model [14]:

$$\begin{aligned} \dot{x}_p &= 30[-x_p + S(\rho_1 + Ax_p - By)], \\ \dot{y} &= 30[-y + S(\rho_2 + Cx_p - Dy + \alpha x_k)], \\ \dot{x}_k &= 120[-x_k + S(\rho_3 + Ex_k - \beta y)], \end{aligned} \tag{1}$$

where S denotes a saturating function $S(x) = 1/(1 + e^{-x})$, and ρ_i $(i = 1, 2, 3)$ are the excitatory inputs from HVc into the sets of neurons with activities x_p, y, and x_k, respectively. In this mean-field model, the excitatory inputs are the population activities of HVc neurons projecting onto RA (reflected in the activities of individual HVc interneurons [12]). The scaling is chosen in such a way that an oscillation in the equations fits temporally a typical song syllable. In Fig. 1 we draw schematically the sets of neurons in the subpopulation of RA neurons under study. The synaptic connections are represented by lines, and the letters in the figure associate the connections with the coefficients in the model. According to Spiro and co-workers [7], inhibitory neurons are of long range. In our model, this is reflected by the fact that both excitatory sets have nonzero synaptic connections with the neurons in the inhibitory set. On the other hand, the excitatory sets cannot connect directly with each other. In other words, it is the inhibitory set of interneurons which ultimately locks otherwise unconnected projection excitatory neurons [7].

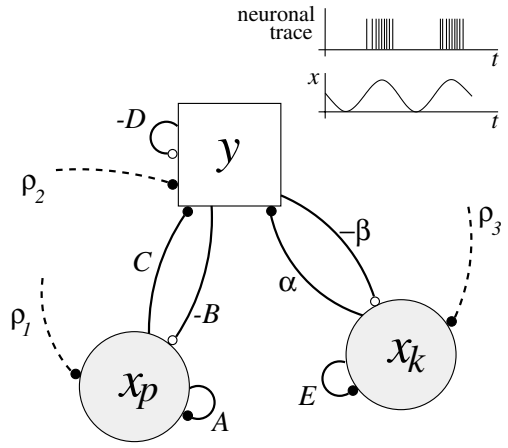

FIG. 1. The neural sets within RA: two excitatory sets (shaded circles) connected to a long-range inhibitory one (open square). All sets receive excitatory inputs from HVc (the high vocal center), denoted here by dashed lines. Inset: averaged activity of one of the populations (bottom), and spiking activity consistent with the averaged activity (top).

A systematic study of these equations is rather complex. For α and β equal to zero, the first two equations conform the well-studied Wilson-Cowan oscillator [14], while the third equation is a one-neuron additive model. In order to illustrate possible solutions of the complete model, we fix the parameters describing the synaptic weights $A, B, C, D, E, \alpha, \beta$ (see Fig. 2 caption for values) and explore the dynamics as one of the parameters denoting external input (from neurons in HVc) is varied.

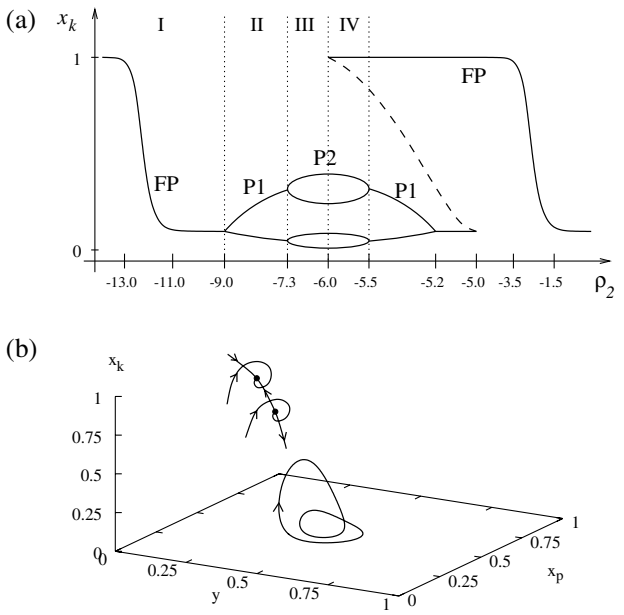

FIG. 2. (a) A bifurcation diagram obtained as ρ_2 is varied, for the neural model presented in the text [Eq. (1)]. FP = fixed point; P1 = period-one limit cycle; P2 = period-two limit cycle. The P1 oscillation is represented by two points indicating maximum and minimum of the oscillation, and the P2 oscillation is represented by four points indicating two local maxima and two local minima. Notice that the coexistence of solutions implies that at bifurcating points the system presents solutions that are qualitatively different. The fixed point solutions generate neural signals that drive the avian vocal organ to produce tonal sounds (i.e., constant frequency vocalizations). The complex period-two solution drives it in such a way that a complex solution is produced: an alternation of two distinct notes. Solid line: stable solutions; dashed line: unstable solutions. Throughout this work we use $A = 10$, $B = 10$, $C = 10$, $D = -2$, $E = 4$, $\alpha = 2$, $\beta = 20$, $\rho_1 = 0$, and $\rho_3 = 6$. (b) Qualitative aspect of the phase space within region IV. Two stable solutions coexist: an attracting fixed point (top) and a period-two limit cycle (bottom), separated by a saddle fixed point (middle).

In Fig. 2(a) we show a bifurcation diagram as the external input ρ_2 is varied. Qualitative changes in the solutions for the neural model take place as the control parameter is changed slightly. The excitatory sets represent neurons projecting to motor neurons innervating syrinx and to areas in lateral medulla controlling respiration. Therefore, the qualitative differences between the solutions will translate into diversity in the song. With a physical model of the syrinx, we will show that this is the case.

Based on the detailed experiments by Goller, Suthers, and Larsen [15,16], a physical model for the syrinx was recently built [5,6]. Using the activities of our neural model to drive the physical model of the syrinx, we can generate synthetic songs. As in [6], we write a minimal equation for the departure of the midpoint of a lateral labium at one side of the syrinx from the stationary position as

$$\ddot{x} - [p(t) - b]\dot{x} + k(t)x + cx^2\dot{x} = 0, \qquad (2)$$

where $p(t) - b$ stands for the difference between bronchial pressure and linear dissipation, $k(t)$ for the fold stiffness, and c for a nonlinear dissipation constant, all per unit mass of the folds [6]. To generate a song with an excitatory instruction given by HVc, we integrate Eqs. (1) to generate the neural signals that allow the driving of the syrinx model [17]. Using for p and k linear functions of the activities x_p and x_k, respectively, we integrate Eq. (2) which allows us to compute the pressure sound wave associated to a song [5,6]. The linear functions we used to relate neural activities with syrinx driving parameters are $p \equiv p(t) = p_1 x_p(t) + p_0$ and $k \equiv k(t) = k_1 x_k(t) + k_0$ (see Fig. 3 caption for linear parameter values). The chosen solutions of the neural model, transformed in pressure and tension instructions ready to drive the physical model of the syrinx, are displayed in Fig. 3. In [6] the model describing the dynamics of a lateral labium includes a force (controlled by the siringealis dorsalis muscle), responsible for active opening and closing of the lumen. Since this dynamics is locked to the respiratory activity, we omit it here for simplicity.

In Fig. 4(a), we display a recorded song [18] of a white-crowned sparrow (*Zonotrichia leucophrys*). The left inset shows the sound wave as a function of time, while the right inset displays the sonogram (i.e., the instantaneous spectral content of the sound wave as a function of time). The displayed segment consists of four distinct syllables. The first two are tonic sounds, followed by a very fast trill, and finally there is repetition of a two-element syllable. Figure 4(b) shows a synthetic song, obtained by driving the physical model of the syrinx with the solutions of the neural model shown in Fig. 3. Notice that the connectivity coefficients A, B, C, D, E, α, and β, the external inputs ρ_1 and ρ_3, and the linear transformation parameters are fixed. Different syllables were

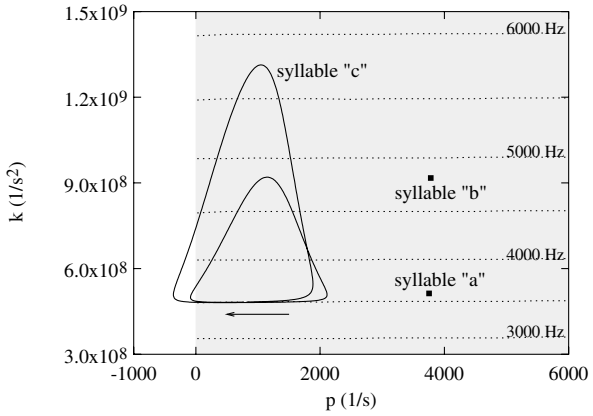

FIG. 3. (p, k) parameter space for the physical model of the syrinx [Eq. (2)], showing the solutions of the neural model (after adjusting scales) that drive the syrinx to generate three syllables of the white-crowned sparrow. The shaded region is the syrinx oscillation region. Dashed-line curves are isofrequency contours. Syllable "a": $\rho_2 = -11.0$; syllable "b": $\rho_2 = -11.8$; syllable "c": $\rho_2 = -7.1$ (syllable labels refer to corresponding labels in Fig. 4). Scaling parameters are set to $p_0 = -2200\ \mathrm{s}^{-1}$, $p_1 = 7000\ \mathrm{s}^{-1}$, $k_0 = 4.8 \times 10^8\ \mathrm{s}^{-2}$, and $k_1 = 1.4 \times 10^9\ \mathrm{s}^{-2}$ throughout this work.

obtained by changing only the value of excitatory input ρ_2 of the neural model (corresponding to input from the neurons in other nuclei, as HVc). Three out of the four syllables present in the recorded song could be reproduced by this subpopulation of neural units. Remarkably, a constant excitatory instruction driving the neural model is able to produce the two-element syllable as a whole. Notice that while the instruction from HVc (the activity of the population of RA-projecting HVc neurons) is constant, it is not until the local circuitry of RA is involved that the two-element syllable emerges. This rich syllable is then the result of a bifurcation of the solutions of the neural model.

For the chosen values of connectivity parameters in the neural model, it was possible to reproduce a large variety of solutions under changes of the excitatory input, showing that many qualitatively different syllables in birdsong can be the result of the interaction of a simple instruction with the local circuits of RA. This allows us to predict that several syllables involve a unique subpopulation of RA neurons. In particular, a syllable such as "c" which is composed by the alternation of two elements is, according to our model, generated by a period doubled solution of the activity of a unique subpopulation of neurons. Therefore the local subpopulation of neurons in the rostral part of the RA (projecting to neurons activating syringeal muscles) that presents activity during the vocalization of the first element, will be active during the vocalization of the second element. A simultaneous measurement in HVc interneurons (representative of the population activity of RA-projecting HVc neurons) will show a constant tonic activity during the two elements. Moreover, if the

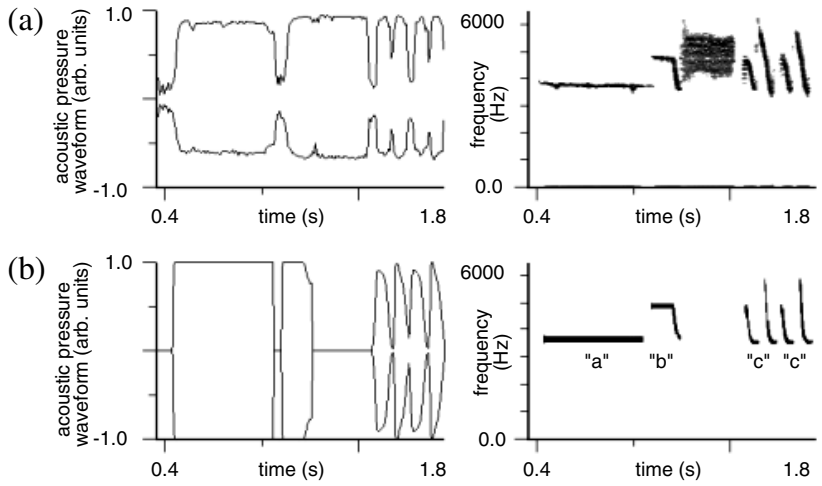

FIG. 4. (a) A recorded song of a white-crowned sparrow (*Zonotrichia leucophrys*), and (b) the synthetic song. One of the syllables cannot be generated by this subpopulation. Parameter values in the model of the syrinx [Eq. (2)] are $b = 1000\ \mathrm{s}^{-1}$ and $c = 10^8\ \mathrm{s}^{-1}\ \mathrm{cm}^{-2}$ throughout this work.

subpopulation of neurons in RA being recorded presents activity during this type of cyclic behavior, it will also be active during the constant frequency vocalizations.

Not every choice of the parameters characterizing the connections leads to this variety of vocalizations. In order to find parameter values appropriate to display a large variety of vocalizations, we analyzed the bifurcation diagrams of the solutions of our model, and fixed connectivity parameters in a value where the bifurcation structure was rich. In particular, the coexistence of solutions that eventually dissapear in local bifurcations allows important changes in the behavior of the system as the control parameters are slightly changed. A detailed analysis of our minimal neural model will be reported elsewhere [19]. Recently, sensory-motor learning in songbirds was studied in terms of the activities of interacting nuclei [20]. Building in this direction, it should be possible to integrate learning and production pathway models in order to obtain a complete description of song development.

We thank Doug Nelson for providing us with white-crowned sparrow recordings. Enlightening discussions with Franz Goller, Tim Gardner, and Pablo Jercog are also acknowledged. This work was partially funded by UBA and CONICET.

*Present address: Institute for Nonlinear Science, UCSD, 9500 Gilman Drive, La Jolla, CA.

[1] F. Nottebohm, T. M. Stokes, and C. M. Leonard, J. Comp. Neurol. **165**, 457–486 (1976).
[2] J. M. Wild, M. N. Williams, and R. A. Suthers, J. Comp. Neurol. **423**, 413–426 (2000).
[3] M. S. Brainard and A. J. Doupe, Nature Rev. Neurosci. **1**, 31–40 (2000).
[4] A. C. Yu and D. Margoliash, Science **273**, 1871–1875 (1996).
[5] T. Gardner, G. Cechi, M. Magnasco, R. Laje, and G. B. Mindlin, Phys. Rev. Lett. **87**, 208101 (2001).
[6] R. Laje, T. J. Gardner, and G. B. Mindlin, Phys. Rev. E **65**, 051921 (2002).
[7] J. E. Spiro, M. B. Dalva, and R. Mooney, J. Neurophysiol. **81**, 3007–3020 (1999).
[8] C K. Catchpole and P. J. B. Slater, *Bird Song, Biological Themes and Variations* (Cambridge University Press, Cambridge, 1995).
[9] D. S. Vicario, NeuroReport **4**, 983–986 (1993).
[10] J. M. Wild, J. Comp. Neurol. **338**, 225–241 (1993).
[11] Z. Chi and D. Margoliash, Neuron **32**, 899–910 (2001).
[12] R. H. R. Hahnloser, A. A. Kozhevnikov, and M. S. Fee, Nature (London) **419**, 65–70 (2002).
[13] J. S. McCasland, J. Neurosci. **7**, 23–39 (1987).
[14] F. Hoppensteadt and E. M. Izhikevich, *Weakly Connected Neural Networks* (Springer, New York, 1997).
[15] R. Suthers and F. Goller, *Motor Correlates of Vocal Diversity in Song Birds*, in Current Ornitology Vol. 14, edited by V. Nolan, Jr., *et al.* (Plenum Press, New York, 1997).
[16] F. Goller and O. N. Larsen, Proc. Natl. Acad. Sci. U.S.A. **94**, 14 787–14 791 (1997).
[17] In [6] the model describing the dynamics of a lateral labium includes a force (controlled by the siringealis dorsalis muscle), responsible for active opening and closing of the lumen. Since this dynamics is locked to the respiratory activity, we omit it here for simplicity.
[18] The recording was performed and kindly provided to us by D. Nelson.
[19] R. Laje, P. Jercog, and G. B. Mindlin (to be published).
[20] T.W. Troyer and A. J. Doupe, J. Neurophysiol. **84**, 1204–1223 (2000).