Biol. Cybern. 79, 377–391 (1998)

Biological Cybernetics

Nonuniformity in the linear network model of the oculomotor integrator produces approximately fractional-order dynamics and more realistic neuron behavior

Thomas J. Anastasio

Beckman Institute and Department of Molecular and Integrative Physiology,
University of Illinois at Urbana/Champaign, Urbana, IL 61801, USA

Received: 8 August 1997 / Accepted in revised form: 18 June 1998

Abstract. The oculomotor integrator is a network that is composed of neurons in the medial vestibular nuclei and nuclei prepositus hypoglossi in the brainstem. Those neurons act approximately as fractional integrators of various orders, converting eye velocity commands into signals that are intermediate between velocity and position. The oculomotor integrator has been modeled as a network of linear neural elements, the time constants of which are lengthened by positive feedback through reciprocal inhibition. In this model, in which each neuron reciprocally inhibits its neighbors with the same Gaussian profile, all model neurons behave as identical, first-order, low-pass filters with dynamics that do *not* match the variable, approximately fractional-order dynamics of the neurons that compose the actual oculomotor integrator. Fractional-order integrators can be approximated by weighted sums of first-order, low-pass filters with diverse, broadly distributed time constants. Dynamic systems analysis reveals that the model integrator indeed has many broadly distributed time constants. However, only one time constant is expressed in the model due to the uniformity of its network connections. If the model network is made nonuniform by removing the reciprocal connections to and from a small number of neurons, then many more time constants are expressed. The dynamics of the neurons in the nonuniform network model are variable, approximately fractional-order, and resemble those of the neurons that compose the actual oculomotor integrator. Completely removing the connections to and from a neuron is equivalent to eliminating it, an operation done previously to demonstrate the robustness of the integrator network model. Ironically, the resulting nonuniform network model, previously supposed to represent a pathological integrator, may in fact represent a healthy integrator containing neurons with realistically variable, approximately fractional-order dynamics.

Correspondence to: T.J. Anastasio, University of Illinois, Beckman Institute, 405 North Mathews Ave., Urbana, IL 61801, USA (Fax: 217-244-5180, e-mail: tja@uiuc.edu)

1 Introduction

The function of the oculomotor system is to move the eye. In order to do so, the eye muscles must overcome both the viscous and the elastic forces that tend to keep the eye from moving. The viscous and elastic forces are proportional to eye rotational velocity and position, respectively. Therefore, the motoneurons must provide the eye muscles with a command combining both the velocity and the position components of the intended eye movement. The velocity component is available because oculomotor commands originate as velocity signals. For example, input to the vestibulo-ocular reflex, which stabilizes gaze during head rotation, is provided by semicircular canal afferents that encode head rotational velocity. Similarly, saccades, which rapidly redirect gaze, are driven by reticular burst neurons that encode eye rotational velocity. The position component is computed from velocity by a neural network known as the oculomotor neural integrator. (For an overview of the oculomotor system, see Robinson 1989a.

The purpose of this paper is to analyze neural network models of the oculomotor integrator in order to gain insight into the dynamics of the neurons that compose the actual integrator. Experimental evidence indicates that there are separate integrators for horizontal and vertical eye movements (Fukushima and Kaneko 1995). This study will focus on the neurons that compose the horizontal integrator, which are located in the medial vestibular nuclei and the nuclei prepositus hypoglossi (MVN/NPH) in the brainstem. An overview will explain the functioning of model oculomotor neural integrators and will present the problem to be addressed here.

1.1 Overview of oculomotor integrator models

Previous models have suggested a mechanism by which neurons could integrate oculomotor velocity signals. Essentially, they use neural elements to construct a low-pass filter. A first-order, low-pass filter has dynamics described by the following first-order differential equation:

$$\tau\dot{y} + y = u \tag{1}$$

where y and u are functions of time t. The transfer function of a first-order, low-pass filter can be written after Laplace transformation of (1), assuming that the initial condition is 0 (Milsum 1966):

$$\frac{Y(s)}{U(s)} = \frac{1}{\tau s + 1} \tag{2}$$

where $U(s)$ and $Y(s)$ are the Laplace transforms of the input to and output from the low-pass filter, and the Laplace variable s is complex frequency ($s = \omega\sqrt{-1}$ where ω is frequency in rad/s). For any frequency ω, the value of the transfer function is a complex number. That number can be represented as a vector in the complex plane. The magnitude and angle of that vector represent, respectively, the gain (output amplitude/input amplitude) and the phase difference (output phase – input phase) of the frequency response described by the transfer function.

The single time constant τ of the first-order, low-pass filter defines its break frequency, which is $1/\tau$ rad/s or $1/(2\pi\tau)$ cycles/s (Hz). The low-pass filter will pass inputs at frequencies lower than its break frequency. For lower frequencies, the gain and the phase difference of the low-pass filter are 1 and 0, respectively. For frequencies higher than its break frequency, the gain is proportional to $1/\omega$, and the phase difference is $-\pi/2$ rad (a phase lag of 90 deg). Because the integral of $\sin(\omega t)$ is $1/\omega \sin(\omega t - \pi/2)$, the first-order, low-pass filter acts as an integrator for frequencies higher than its break frequency.

To a first approximation, a single neuron will act as a first-order, low-pass filter with τ equal to its membrane time constant. With $\tau = 5$ ms (a typical value), the single-neuron, low-pass filter would integrate inputs at frequencies above its break frequency of about 32 Hz. This is too high for the single neuron to be useful as an oculomotor integrator, since most eye movements occur at much lower frequencies. The time constant of the oculomotor integrator (τ') has been measured to be 20 s or more in human and cat (Becker and Klein 1973; Robinson 1974), corresponding to a break frequency of 0.01 Hz or lower. Thus, the integrator time constant τ' is at least 4000 times longer than the single-neuron time constant τ.

Kamath and Keller (1976) suggested that the single-neuron τ could be lengthened to the neural-integrator τ' through positive feedback. A simplified, lumped version of their model is schematized in Fig. 1A, in which a neuron receives input with weight v and sends an excitatory collateral to itself with weight w. To allow the use of linear systems analysis here, all neural inputs and outputs are scalars that represent neural firing rate, and neural nonlinearities and synaptic delays are neglected. Let us assume that $v = 1$, and that the neuron by itself,

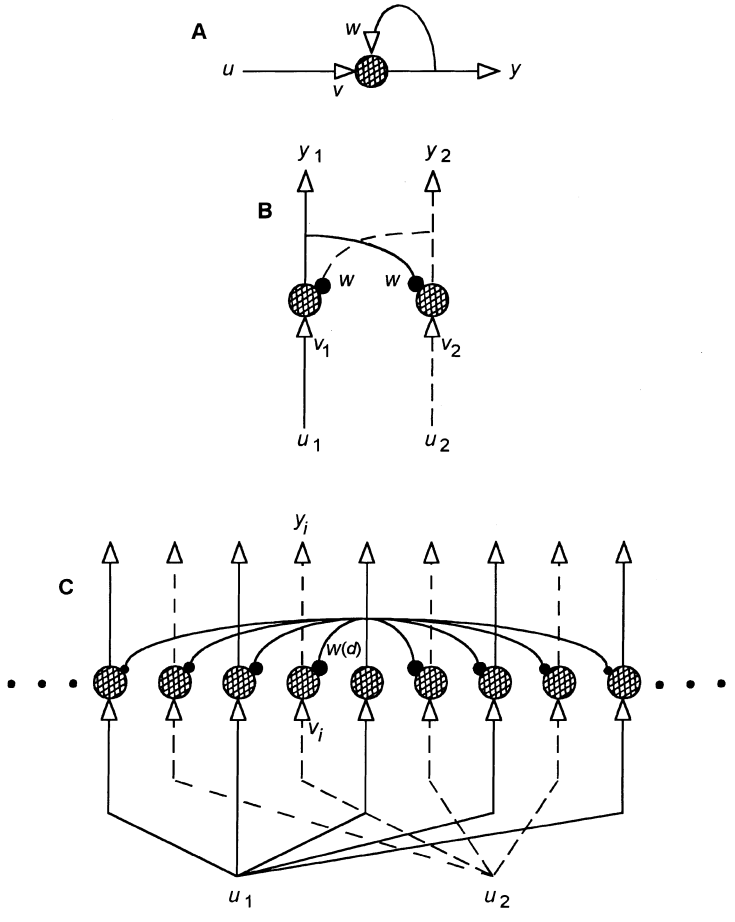

Fig. 1A–C. Schematic diagrams of three neural network models of the oculomotor neural integrator. In all three networks, each *hatched circle* represents a single neuron. *Open arrows* and *filled circles* represent excitatory and inhibitory synaptic connections, respectively. Input and recurrent connection weights are labeled with v's and w's, respectively. Inputs to and outputs from the networks are labeled with u's and y's, respectively. **A** A1-neuron integrator in which the neuron sends an excitatory connection to itself. **B** A 2-neuron integrator in which the neurons reciprocally inhibit each other. **C** A 32-neuron integrator in which each neuron reciprocally inhibits its neighbors, and connection weights $w(d)$ vary with distance according to a Gaussian profile. The 32 neurons are arranged in a circle so that neuron 1 and neuron 32 are nearest neighbors. **A** is adapted from Kamath and Keller (1976) and **B, C** from Cannon, Robinson, and Shamma (1983)

in the absence of positive feedback, has first-order, low-pass filter dynamics described by (1, 2) with $\tau = 5$ ms. Then the feedback equation (Milsum 1966) can be used to show that, with positive feedback of weight w, the transfer function is again that of a first-order, low-pass filter (2) but with a new time constant $\tau' = \tau/(1 - w)$. Thus, τ' can be made arbitrarily long by choosing w close to 1; for example, $\tau' = 20$ s with $w = 0.99975$. (Positive feedback will also affect the gain, but because the focus here is on dynamics, the frequency-independent gain will be ignored – see Results).

The model of Kamath and Keller (1976) demonstrates how an integrator could be constructed from neurons using positive feedback. The model is extremely sensitive to the value of w and becomes unstable with $w > 1$. Weight sensitivity and the potential for instability are drawbacks common to all integrator models that rely on some form of positive feedback to achieve neural time-constant lengthening (see Discussion). The particular drawback of the Kamath and Keller model is its failure to discriminate between the oculomotor velocity signals, which occur as modulations of a background firing rate or carrier, and the carrier itself. The carrier component can be thought of as a step up to some constant level. The 1-neuron integrator would integrate the carrier as well as the modulated component, and its output would ramp up to a very high, unusable level (the ramp is the integral of the step).

This drawback was overcome by Cannon, Robinson, and Shamma (1983) using the mechanism of reciprocal inhibition. They took advantage of the fact that the oculomotor system is arranged bilaterally and operates in push-pull. The simplest bilateral, reciprocal network is composed of two neurons, as schematized in Fig. 1B. Both neurons are first-order, low-pass filters, as in (2) with $\tau = 5$ ms. The neurons reciprocally inhibit each other with reciprocal connections of weight w, and receive inputs of weight v_1 and v_2. The inputs u_1 and u_2 have the same carrier level and are modulated in push-pull by the same amount. Thus, velocity commands would modulate the carriers of both inputs by the same amount but in opposite directions. An illustration of an opposite-direction impulse input modulation is given in Fig. 2A. The modulations could occur about a carrier of any level, but a carrier of 0 is chosen for simplicity.

Cannon and co-workers (1983) showed that the neurons in the 2-neuron integrator act as identical, first-order, low-pass filters, but with a time constant that differs for opposite- and same-direction input modulations (Fig. 2A,B). The single-neuron time constant is lengthened for opposite-direction input modulations $[\tau' = \tau/(1 - w)]$, as for the 1-neuron integrator] but shortened for same-direction input modulations $[\tau' = \tau/(1 + w)]$. For example, if w is set to 0.99975 as for the 1-neuron integrator, then τ' would equal 20 s and 2.5 ms for opposite- and same-direction input modulations, respectively. Thus, both neurons respond to an opposite-direction impulse input modulation (Fig. 2A) with opposite-direction exponential decays having a long time constant of 20 s (Fig. 2C). This approximates a step change in the output (the step is the integral of the impulse). In contrast, both neurons respond to a same-direction step input modulation (Fig. 2B) with a same-direction exponential change to a new steady state having a short time constant of 2.5 ms (Fig. 2D). Quickly coming to a new steady state rather than producing a ramp change in the output indicates that the neurons don to integrate same-direction input modulations and would not non-0 carriers.

The 2-neuron integrator can discriminate between opposite- and same-direction input modulations, but as it is composed of only two neurons, it would fail completely if it lost even one neuron. To make the integrator more robust, Cannon and co-workers (1983) expanded the reciprocal integrator network to 32 neurons. The expanded network is schematized in Fig. 1C. As before,

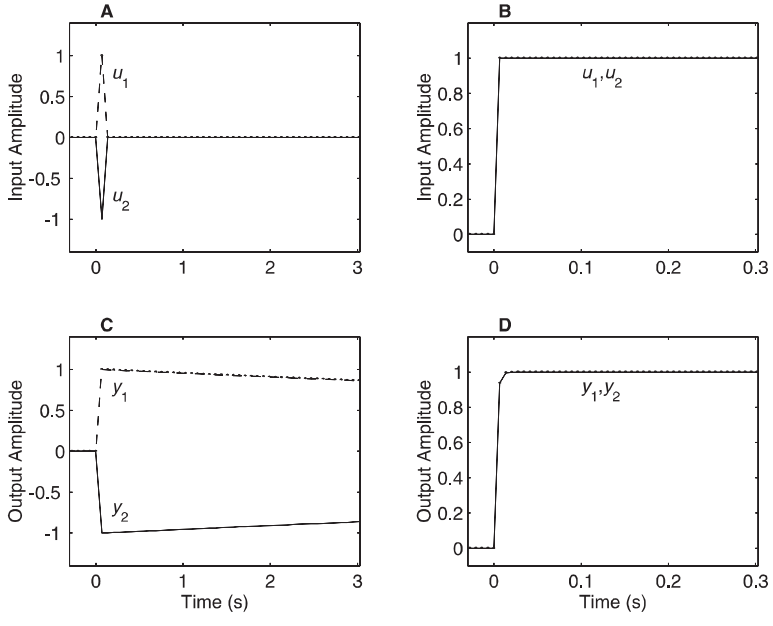

Fig. 2A–D. Impulse and step responses of the 2-neuron integrator. **A, B** Inputs u_1 (*dashed lines/dot symbols*) and u_2 (*solid lines*) have carrier rates of 0 that are modulated either as opposite-direction impulses (**A**) or as same-direction steps (**B**). **C, D** Outputs y_1 (*dashed lines/dot symbols*) and y_2 (*solid lines*) respond to opposite-direction impulses (**A**) with opposite-direction exponential decays having time constants of 20 s (**C**) or respond to same-direction steps (**B**) with same-direction exponential changes in steady state having time constants of 2.5 ms (**D**). Time scales in **A, C** differ from those in **B, D**

the neurons by themselves, in the absence of network connections, are first-order, low-pass filters as in (2) with $\tau = 5$ ms. In the 32-neuron network, the neurons reciprocally inhibit their neighbors with connections of weight $w(d)$. These weights have a Gaussian profile:

$$w(d) = \exp\left[-\frac{1}{2}\left(\frac{d}{\sigma^2}\right)^2\right] \quad (3)$$

where d is the distance in the network, measured in numbers of neurons, and σ^2 is the variance of the Gaussian. The inhibitory projections of any neuron vary with distance according to the Gaussian profile (3), but the network is uniform in that each of the 32 neurons inhibits its neighbors with exactly the same profile. To prevent edge effects, the neurons are arranged in a circle so that neuron 1 and neuron 32 are immediate neighbors. Each of the 32 neurons receives an input of weight v_i, either from u_1 or u_2 which alternately project to odd- and even-numbered neurons, respectively. As before, u_1 and u_2 have the same carrier level and are modulated in push-pull by the same amount. Like the 2-neuron integrator, the 32-neuron integrator discriminates between opposite- and same-direction input modulations. Unlike the 2-neuron network, the 32-neuron network is robust in that the integrating properties of many of the neurons are relatively unaffected by eliminating one neuron (or by removing its reciprocal connections – see Results).

Despite its complicated connectivity, each neuron in the uniform 32-neuron integrator acts as a first-order, low-pass filter as in (2). All the neurons have identical time constants, one long and one short, for opposite- and same-direction input modulations, respectively (Cannon et al. 1983). The frequency responses for all 32 neurons in the uniform network with opposite-direction input modulations are shown in Fig. 3A,B. The frequency responses are represented as Bode plots (Milsum 1966) in which log gain (Fig. 3A) and phase difference (Fig. 3B) are plotted separately against log frequency in Hz. The frequency range (0.01–10 Hz) is experimentally relevant (see below). The Bode plots show that all 32 neurons have gain slopes of −1 and phase lags of 90 deg over most of this range, implying that they are acting as first-order integrators. These simple, identical, first-order dynamics do *not* correspond to the variable, approximately fractional-order dynamics of the MVN/NPH neurons that make up the actual oculomotor neural integrator.

1.2 MVN/NPH neurons act approximately as fractional integrators

The frequency responses of MVN/NPH neurons have been studied over the range extending from 0.01 to 10 Hz. The maximum phase lag they attain is 90 deg or less. The most salient features of MVN/NPH frequency responses are near-parallel shifts of up to 45 deg or more between the phase characteristics of individual neurons that are maintained over several decades of frequency (Shinoda and Yoshida 1974; Blanks et al. 1977; Keller and Precht 1979; Lopez-Barneo et al. 1979; McFarland and Fuchs 1992; Stahl and Simpson 1995). The uniform 32-neuron integrator model does not reproduce these phase shifts.

Parallel shifts in phase lag between 0 and 90 deg, similar to those observed for MVN/NPH neurons, are characteristic of fractional integrators with orders between 0 and 1 (Anastasio 1994). While a phase lag of 90 deg implies 1 order of integration, a phase lag between 0 and 90 deg implies an order of integration between 0 and 1 (a fractional order). In the Laplace domain, a first-order integrator can be represented as s^{-1}, and a fractional integrator can be represented as s^{-k} where $0 < k < 1$ (Oldham and Spanier 1974). The

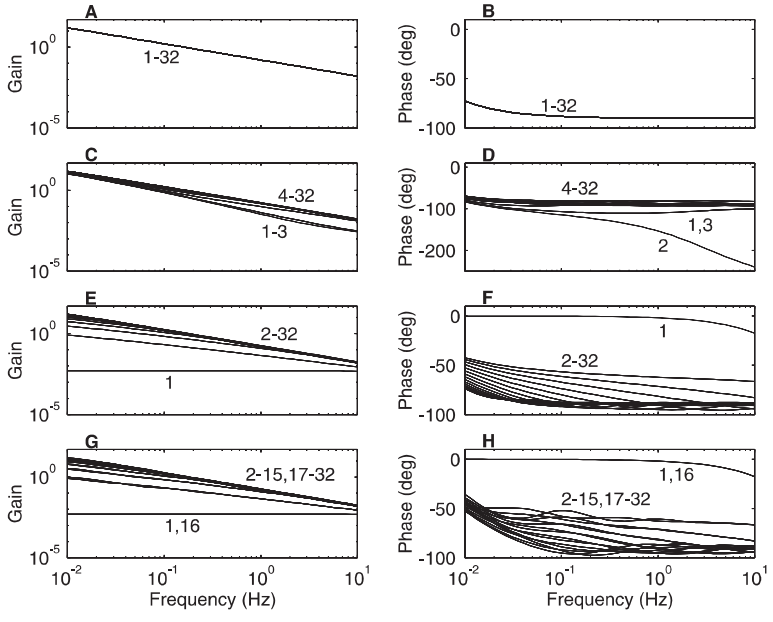

Fig. 3A–H. Bode plots of the frequency responses of the 32-neuron integrator with various connections removed. All plots cover the same, experimentally relevant frequency range (10^{-2} to 10^{1} Hz). For clarity, this range is marked only on the *bottom plots* (**G**, **H**). In each case, the gain (**A, C, E, G**) and phase (**B, D, F, H**) of all 32 neurons in the integrator are plotted. *Numerals* near the curves denote individual neurons. All responses are computed with opposite-direction input modulations. The phase difference of 180 deg between oppositely modulated outputs has been removed to bring all phases into the same range. **A, B** Bode plot for the uniform 32-neuron integrator. **C, D** Bode plot for the 32-neuron integrator with the input connections (weights v_i) to neurons 1, 2, and 3 removed. **E, F** Bode plot for the 32-neuron integrator with the reciprocal connections [weights $w(d)$] to and from neuron 1 removed. **G, H** Bode plot for the 32-neuron integrator with the reciprocal connections to and from neurons 1 and 16 removed

Bode plot of a fractional integrator has a gain slope of $-k$ and a phase lag of $k \times 90$ deg. Fractional integrators were used recently to model the dynamics of MVN/NPH neurons (Anastasio 1994), where it was also shown that a fractional integrator could be approximated as a weighted sum of first-order, low-pass filters:

$$s^{-k} \approx \frac{\Delta(\log \tau)}{\Gamma(k)\Gamma(1-k)} \sum \tau^k \frac{1}{\tau s + 1} \tag{4}$$

where Γ is the gamma function. The time constants τ have logarithmic spacing $\Delta(\log \tau)$ and would need to span at least as many orders of magnitude as the frequency range of the < 90 deg phase shifts to be produced. Each low-pass filter in the sum is weighted by its time constant τ raised to the power k, which is the order of the fractional integrator. In practice, the precise spacing and weighting of the low-pass filters is not critical for the production of near-constant phase shifts < 90 deg in the sum. The critical features are that the set of low-pass filters must have time constants that are roughly evenly distributed over many orders on a logarithmic scale, and they must act in parallel so that their outputs sum. A sum of many first-order, low-pass filters, as indicated on the right-hand side of (4), would produce a high-order transfer function with order equal to the number of low-pass filters in the sum. Thus, a fractional-order integrator can be approximated by an integer-order transfer function of very high order. The approximation in (4) can also work in reverse. For certain cases, the weighted sum of many first-order, low-pass filters with diverse time constants may be approximated as a fractional-order integrator.

1.3 Reconciling the fractional integrator and neural network approaches

Fractional integrators capture well the essential features of the dynamics of the MVN/NPH neurons that compose the actual oculomotor integrator (Anastasio 1994). The neural network models proposed to date do not reproduce those features (see Discussion for more details). The 32-neuron integrator (Cannon et al. 1983) successfully integrates opposite- but not same-direction input modulations and is robust. However, despite having a complicated architecture, the dynamics of the neurons in the 32-neuron integrator are those of identical, first-order, low-pass filters. As such, the uniform 32-neuron integrator model does not reproduce the variable, approximately fractional-order dynamics that are observed for the MVN/NPH neurons that compose the actual oculomotor integrator.

The purpose of this paper is to reconcile the fractional integrator and neural network approaches. It will be shown, using dynamic systems analysis, that the linear 32-neuron integrator model can be made to express the kind of high-order dynamics that are needed to approximate fractional-order integration. This improves the ability of the model to simulate the behavior of the MVN/NPH neurons that compose the actual oculomotor integrator. The key to the solution lies in disrupting the uniformity of the connections in the 32-neuron integrator. This is done most effectively by disconnecting or eliminating a small number of neurons. These are operations which, ironically, were previously supposed to produce a pathological integrator.

2 Methods

Previous models of the oculomotor neural integrator were developed using linear control systems theory (e.g., Milsum 1966). That development will be extended here using the more modern state-space approach to dynamic systems (Luenberger 1979; Kailath 1980). All matrix manipulations were carried out numerically in MATLAB (The MathWorks, Inc.).

Any linear, dynamic system can be represented as a system of coupled, first-order differential equations:

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 u \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 u \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{n_n}x_n + b_n u \end{aligned} \tag{5}$$

where coefficients a_{ij} couple states x_i to each other, and coefficients b_i couple input u to the states. In this case, there are n states and n coupled first-order differential equations, but only one input. The system can be put into state-space form using matrix notation:

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}u \\ y &= \boldsymbol{c}^T \boldsymbol{x} \end{aligned} \tag{6}$$

where matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ contain the a_{ij} and b_i coefficients, respectively. Note that $\boldsymbol{A}$ is an $n \times n$ square matrix and $\boldsymbol{b}$ is an n-element column vector. The n-element vector $\boldsymbol{c}$ (a column transposed to a row, where superscript T denotes the transpose operation) is used to specify the output of the system. Here c is used to select one state at a time as the specified output y.

The system can be converted from state-space form into transfer function form:

$$\frac{Y(s)}{U(s)} = \boldsymbol{c}^T [s\boldsymbol{I} - \boldsymbol{A}]^{-1} \boldsymbol{b} \tag{7}$$

where $\boldsymbol{I}$ is the identity matrix. A transfer function can be expressed as a ratio of polynomials in s:

$$\frac{Y(s)}{U(s)} = \frac{g_m s^m + \cdots + g_1 s + g_0}{h_n s^n + \cdots + h_1 s + h_0} \tag{8}$$

where the time constants are functions of the numerator and denominator coefficients g_i and h_i. In addition to describing frequency-domain responses (see Introduction), the transfer function can be used to find the responses of the system in the time domain. This is done by first parameterizing the input [for example, $U(s) = 1$ for an impulse and $U(s) = 1/s$ for a step] and then performing a reverse Laplace transformation.

The advantage of the state-space approach lies in its ability to make the dynamic structure of a system transparent. This can be done through a similarity transformation of the system to Jordan form. The system of (6), with state variables $\boldsymbol{x}$, is transformed to a new but similar system with state variables $\boldsymbol{z}$, where $\boldsymbol{x} = \boldsymbol{X}\boldsymbol{z}$. Matrix $\boldsymbol{X}$ is the matrix of eigenvectors $\boldsymbol{e}_i$ which satisfy the eigenvalue equation $\boldsymbol{A}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$ with eigenvalues λ_i. If system matrix $\boldsymbol{A}$ is symmetric, as it is in all cases here, then the eigenvalues are all real numbers. The eigenvalues of $\boldsymbol{A}$, also called the modes of $\boldsymbol{A}$, determine the dynamics of the system. Real eigenvalues are the opposite reciprocals of the time constants of the system ($\tau_i = -1/\lambda_i$).

The system can be transformed using the eigenvector matrix $\boldsymbol{X}$:

$$\begin{aligned} \dot{\boldsymbol{z}} &= \boldsymbol{X}^{-1}\boldsymbol{A}\,\boldsymbol{X}\,\boldsymbol{z} + \boldsymbol{X}^{-1}\boldsymbol{b}\,u \\ y &= \boldsymbol{c}^T\,\boldsymbol{X}\,\boldsymbol{z} \end{aligned} \tag{9}$$

where superscript -1 denotes the matrix inverse. For a symmetric matrix $\boldsymbol{A}$, matrix $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$ is a diagonal matrix with the eigenvalues λ_i (or system dynamic modes) along the diagonal and 0 elsewhere. This is shown in an example system in Jordan form where $\boldsymbol{b}' = \boldsymbol{X}^{-1}\boldsymbol{b}$ and $\boldsymbol{c}'^T = \boldsymbol{c}^T\boldsymbol{X}$:

$$\dot{\boldsymbol{z}} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & & \\ 0 & 0 & \lambda_3 & 0 & & \\ 0 & 0 & 0 & \lambda_4 & & \\ \vdots & & & & \ddots & \\ 0 & & & & & \lambda_n \end{bmatrix} \boldsymbol{z} + \begin{bmatrix} b'_1 \\ b'_2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u \tag{10}$$

$$y = \begin{bmatrix} c'_1 & 0 & c'_3 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{z}$$

The coefficients of the states of the transformed system are the eigenvalues (modes) of the original system. Also, the transformed states z_i are uncoupled from each other, so the contribution of each mode to system dynamics can be studied in isolation.

All modes can influence the dynamics of the system as it relaxes from some non-0 initial conditions. All modes need not influence the system transfer functions, which describe the effect of input u on any selected output y when the initial conditions are 0. Of interest here are the transfer function dynamics of the model neurons, because they can be compared with the input/output dynamics that are observed for the MVN/NPH neurons that compose the actual oculomotor integrator. In the integrator network model, the modes that contribute a corresponding time constant to the neuron transfer functions are those that are both controllable by the input u and observable at the output y. Controllable and observable modes are identified in the transformed system as those associated with non-0 elements in $\boldsymbol{b}' = \boldsymbol{X}^{-1}\boldsymbol{b}$ and $\boldsymbol{c}'^T = \boldsymbol{c}^T\boldsymbol{X}$, respectively. In example system (10), only eigenvalue λ_1 is both controllable and observable. The others are either unobservable (λ_2), uncontrollable (λ_3), or both (λ_4 through λ_n). The transfer functions of example system (10) could potentially be nth order with n time constants, because it has n modes. Instead, they are first-order with only one time constant $\tau_1 = -1/\lambda_1$, because only mode λ_1 is both controllable and observable.

In addition to matrix $\boldsymbol{A}$, which determines matrix $\boldsymbol{X}$, controllability and observability in single-input/single-output systems as studied here are dependent upon vectors $\boldsymbol{b}$ and $\boldsymbol{c}$, respectively [see (9)]. For the 2- and 32-neuron integrators, vector $\boldsymbol{c}$ can be chosen arbitrarily, so all modes can be made observable. Therefore, which modes will contribute time constants to the transfer functions of the 2- and 32-neuron integrators depends, upon which modes are controllable. That in turn depends upon matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$, which define the structure of the neural integrator. The analytical techniques outlined in this section will be used to show how changes in matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ can affect the dynamics of the neurons in the integrator network.

3. Results

The purpose of this paper is to use dynamic systems analysis to explore the 32-neuron integrator and to show how it can be modified so that the dynamics of its constituent neurons approximate fractional-order integration [see (4)]. As such, the modified 32-neuron integrator model can simulate better the dynamics that are observed for the MVN/NPH neurons that compose the actual oculomotor integrator. The concept of controllability is pivotal to this analysis. The use of this concept is best illustrated by analyzing the 2-neuron integrator.

3.1 Analyzing the 2-neuron integrator model

The dynamics of both of the neurons that compose the 2-neuron integrator (Fig. 1B) are governed either by one long or by one short time constant, which pertain, respectively, for opposite- or same-direction input modulations (see Introduction). Linear dynamic systems analysis can be used to show that this difference in dynamics is due to differences in controllability for the two types of inputs. To simplify the analysis and to make the system single-input/single-output, there will be only one input u, and it will be opposite or same when v_1 and v_2 have opposite signs or the same sign, respectively. The dynamics of each neuron in the 2-neuron integrator can then be described by the following set of two coupled differential equations:

$$\begin{aligned} \tau\dot{x}_1 + x_1 &= -wx_2 + v_1 u \\ \tau\dot{x}_2 + x_1 &= -wx_1 + v_2 u \end{aligned} \tag{11}$$

where x_1 and x_2 denote the states (firing rates) of neurons 1 and 2, respectively. The equations have the same form as (1) except that each neuron, in addition to an input from u, has an input from the other neuron. Equations (11) can be put into matrix form:

$$\tau\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}+\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0 & -w\\ -w & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}v_1\\ v_2\end{bmatrix}u \tag{12}$$

and (12) can then be rearranged into state-space form:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=-\frac{1}{\tau}\begin{bmatrix}1 & w\\ w & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\frac{1}{\tau}\begin{bmatrix}v_1\\ v_2\end{bmatrix}u \tag{13}$$

which can be parameterized. As for the 1-neuron integrator (see Introduction), the single-neuron time constant τ is 5 ms and the feedback weight w is 0.99975. Because the focus here is on dynamics, the convention of Cannon and co-workers (1983) will be adopted in which the input weights (after scaling by $1/\tau$) are given the absolute value of 1 for simplicity. The resulting frequency-independent gain will be ignored, although it could be controlled precisely through the input weights (ibid). To have the input modulated in opposite directions, $v_1/\tau = 1$ and $v_2/\tau = -1$.

The parameterized system can be transformed numerically into Jordan form:

$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\end{bmatrix}=\begin{bmatrix}-0.05 & 0\\ 0 & -399.95\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+\begin{bmatrix}-1.41\\ 0\end{bmatrix}u \tag{14}$$

in which the eigenvalues ($\lambda_1 = -0.05$, $\lambda_2 = -399.95$) fall along the diagonal of the transformed system matrix $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$ [see (9) and (10)]. The eigenvalues are negative real numbers, implying that system dynamics are stable and nonoscillatory. The eigenvalues correspond to time constants $\tau_1 = -1/\lambda_1 = 20$ s and $\tau_2 = -1/\lambda_2 = 2.5$ ms. Of the two elements in the transformed input vector $\boldsymbol{X}^{-1}\boldsymbol{b}$, only the first is non-0. Therefore, only the corresponding eigenvalue λ_1 is controllable, and τ_1 is the only time constant that will appear in the transfer functions of the system when the input is modulated in opposite directions. Using the $\boldsymbol{c}$ vector to choose either x_1 or x_2 as the output y (6) and converting to transfer function form (7) show that the outputs of the neurons have opposite sign, and both have first-order, low-pass dynamics as in (2) with $\tau' = 20$ s. In contrast, for same-direction input modulations, where $v_1/\tau = v_2/\tau = 1$, only λ_2 is controllable. In that case, the outputs of the neurons have the same sign, and both have first-order, low-pass dynamics as in (2) with $\tau' = 2.5$ ms. Thus, the ability of the 2-neuron integrator to discriminate opposite- and same-direction input modulations can be interpreted as differences in the abilities of the two types of input modulations to control the dynamic modes of the system. A similar approach can be applied to provide insight into the dynamic structure of the 32-neuron integrator network.

3.2 Analyzing the 32-neuron integrator model

The 32-neuron integrator (Fig. 1C) can be represented as a system of 32 coupled differential equations (5), where each equation describes the dynamics of the state (firing rate) of one neuron. This system can be represented in state-space form (6) and parameterized. The absolute values of the inhibitory, reciprocal connection weights $w(d)$ have a Gaussian profile (3) with variance $\sigma^2 = 1.51$. (This variance is slightly larger than that used by Cannon et al. 1983 – see Discussion.) In the uniform network, every neuron inhibits its neighbors with exactly the same connectivity profile, and matrix A is symmetric. The network has a single input u. The scaled input weights v_i/τ have absolute value 1. Odd-numbered input weights are always positive, while even-numbered input weights are negative for opposite- and positive for same-direction input modulations. Vector $\boldsymbol{c}$ selects one neural state x_i at a time as the single output y.

For opposite-direction input modulations, each neuron in the uniform 32-neuron integrator, with all reciprocal connections present, has identical, first-order, low-pass filter dynamics (2) with time constant $\tau' = 50.8$ s. Bode plots of the frequency responses of all 32 neurons are shown in Fig. 3A,B, in which the phase difference of 180 deg between oppositely modulated outputs has been removed to simplify the presentation of the data. The Bode plots show gain slopes of -1 and phase lags of 90 deg over most of the frequency range, implying that the neurons are acting as first-order integrators over that range. It is surprising that the neurons connected together in a complicated network like the 32-neuron integrator should all have simple, identical, first-order dynamics. An analysis of the modal structure of the 32-neuron integrator provides insight into the dynamics of this network.

The state-space representation of the 32-neuron integrator can be transformed into Jordan form, in which its dynamic modes and their controllability are revealed. The 32-neuron integrator will have as many dynamic modes (eigenvalues) as it has states (neurons). Because the system matrix for the 32-neuron integrator is symmetric, all eigenvalues will be real numbers. The 32-neuron integrator has 32 real eigenvalues, which are plotted in Fig. 4A ($\times$ symbols). All eigenvalues are negative, indicating that the system is stable. The most impressive feature of the plot is that the eigenvalues of the uniform 32-neuron integrator are distributed roughly evenly on a logarithmic scale over many orders of magnitude.

These real eigenvalues correspond to the opposite reciprocals of the time constants. If the broadly distributed time constants of the 32-neuron integrator contributed to transfer function dynamics as low-pass filters acting in parallel, then the dynamics of the neurons in the network could approximate those of fractional integrators [see (4)]. The range in eigenvalues corresponds to a range in time constants that could account for the constant phase lags < 90 deg that are observed across frequencies for the MVN/NPH neurons that compose the actual oculomotor integrator. The reason why this diverse set of modes does not contribute to the transfer function dynamics of the 32-neuron integrator is because only one mode is controllable (Fig. 4A, + symbol). That mode corresponds to the eigenvalue with the smallest absolute value, and so also to the longest time constant. Thus, for opposite-direction inputs, every neuron in the uniform 32-neuron

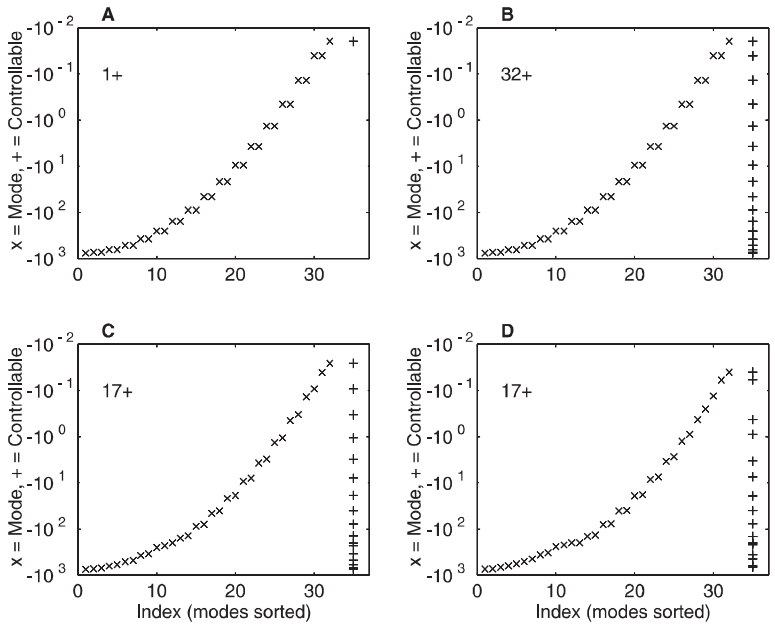

Fig. 4A–D. Dynamic modes and controllability of the 32-neuron integrator with various connections removed. All modes (eigenvalues) are computed with opposite-direction inputs. In each case, the 32 modes are first sorted and then plotted with an × symbol. Controllable modes are designated with a + symbol in a *column on the right-hand side* of each plot, at the same level as the controllable mode. The number of controllable modes is also indicated. **A** Modes and controllability of the uniform 32-neuron integrator. Only one mode, the smallest in absolute value, is controllable. **B** Modes and controllability of the 32-neuron integrator with the input connections (weights v_i) to neurons 1, 2, and 3 removed. All 32 modes are controllable. **C** Modes and controllability of the 32-neuron integrator with the reciprocal connections [weights $w(d)$] to and from neuron 1 removed. Seventeen modes (roughly half) are controllable. **D** Modes and controllability of the 32-neuron integrator with the reciprocal connections to and from neurons 1 and 16 removed. Again, 17 modes are controllable

integrator has simple, first-order, low-pass filter dynamics as in (2) with $\tau' = 50.8$ s (Fig. 3 A,B).

Changing the inputs to the uniform 32-neuron integrator from opposite to same direction changes the one controllable mode from the eigenvalue with the smallest to that with the largest absolute value (not shown). Thus, for same-direction inputs, every neuron in the uniform 32-neuron integrator has simple, first-order, low-pass filter dynamics as in (2) with $\tau' = 1.3$ ms (not shown). Interestingly, the 30 eigenvalues other than those with the largest and smallest absolute values are grouped in pairs having precisely the same value. This is associated with the uniform structure of the network (see below).

3.3 Analyzing the 32-neuron integrator with nonuniform input connections

In order for it to simulate the high-order, approximately fractional-order dynamics exhibited by MVN/NPH neurons, it is necessary to increase the number of modes that contribute to the transfer function dynamics of the 32-neuron integrator by increasing the number of modes that are controllable. It is counterintuitive that, for a system in which every element receives an input, only one dynamic mode is controllable by the input. This lack of controllability is due to the uniformity of the input. The 32-neuron integrator actually becomes more controllable as inputs to individual neurons are removed. An analogy might help to make this idea more concrete.

Imagine a large, cubic box that slides on the floor when pushed. Potentially, the box could be translated in two dimensions and rotated in one dimension (about the vertical axis) on the planar floor. Imagine that the input comes from two individuals, centered on opposite sides of the box, who are constrained to push at the same time and with the same force. Obviously, with this input configuration, the box will not move. However, if an individual is removed from one side, then the box will translate in the direction of that side. Because of the particular uniformity in the configuration of the inputs, the box actually becomes more controllable when an input is removed.

Removing the input to even one neuron in the 32-neuron integrator makes all 32 modes controllable, as determined through numerical computation of the Jordan form. The dynamic effects of increasing controllability through input disconnection are most clearly illustrated in a 32-neuron integrator, with opposite-direction inputs, from which the inputs to neurons 1, 2, and 3 are removed. Eigenvalues and controllability in this case are shown in Fig. 4B. The eigenvalues themselves are the same as in the uniform network (Fig. 4A), because the system matrix has not been changed, but now all 32 modes are controllable. These dynamic modes, distributed over many orders, are now free to contribute to individual neuron transfer functions. It is possible that the resulting high-order dynamics of the neurons in the 32-neuron integrator would approximate those of fractional integrators, and thus better match the dynamics of their counterparts in the MVN/NPH. Such is not the case.

Bode plots of the frequency responses of all 32 neurons in the 32-neuron integrator, with opposite-direction inputs but with the input connections to neurons 1, 2, and 3 removed, is shown in Fig. 3C, D. For neurons 4 through 32, gain (Fig. 3C) and phase (Fig. 3D) vary slightly but are otherwise similar to the case with all input connections present (Fig. 3A, B). For neurons 1, 2, and 3, the gain slopes are more negative than -1, implying orders of integration > 1. The phase lag for neurons 1 and 3 increases to 180 deg at frequencies above the range, and phase lag for neuron 2 increases to more than 200 deg within the range, again implying orders of integration > 1. Phase lag within the range

continues to increase as the input connections to additional neurons are removed (not shown). This type of phase behavior is not observed for the MVN/NPH neurons that compose the actual oculomotor integrator.

Following input removal from the 32-neuron network, it appears that neurons are acting as low-pass filters in series, with the disconnected neurons downstream from the connected neurons. The lack of order > 1 dynamics in the actual integrator suggests that the neurons composing it are well innervated by inputs that have not yet been integrated. Neurons that compose the actual oculomotor integrator have dynamics that approximate those of fractional integrators with orders between 0 and 1 (Anastasio 1994). Therefore, rather than acting in series, it appears that the modes that govern the dynamics of MVN/NPH neurons are acting as low-pass filters in parallel [see (4)].

3.4 Analyzing the 32-neuron integrator with nonuniform reciprocal connections

In a system like the uniform 32-neuron integrator, making the structure of the system itself less regular can also increase controllability. Returning to the box analogy, again imagine that there are two individuals, centered on opposite sides of the box, both pushing at the same time and with the same force. But now the structure of the cubic box is made nonuniform by changing its two horizontal sides from squares into parallelograms. The two inputs will cause the reconfigured (nonuniform) box to rotate on the planar floor. The box becomes more controllable when its structure is made less regular.

The uniformity of the 32-neuron integrator can be disrupted simply by removing the reciprocal connections to and from one neuron in the network. This is equivalent to eliminating one neuron. This operation was performed by Cannon and co-workers (1983) to demonstrate the robustness of the 32-neuron integrator. Their findings on the opposite-direction impulse response of the 32-neuron integrator, with all inputs present but with the reciprocal connections to and from neuron 1 removed, are reproduced in Fig. 5A. The impulse responses of the neurons farthest away from the disconnected neuron decay as slowly as they do in the uniform 32- or 2-neuron network (Fig. 2C), showing that the network is robust. The decay rate increases with proximity to the disconnected neuron (that of disconnected neuron 1 itself returns to baseline with the single-neuron time constant). Inspection of the impulse responses of the near neighbors suggests that they are composed of sums of exponential decays with broadly distributed time constants. This implies that removing the reciprocal connections of neuron 1 has increased the number of controllable modes, which then contribute to

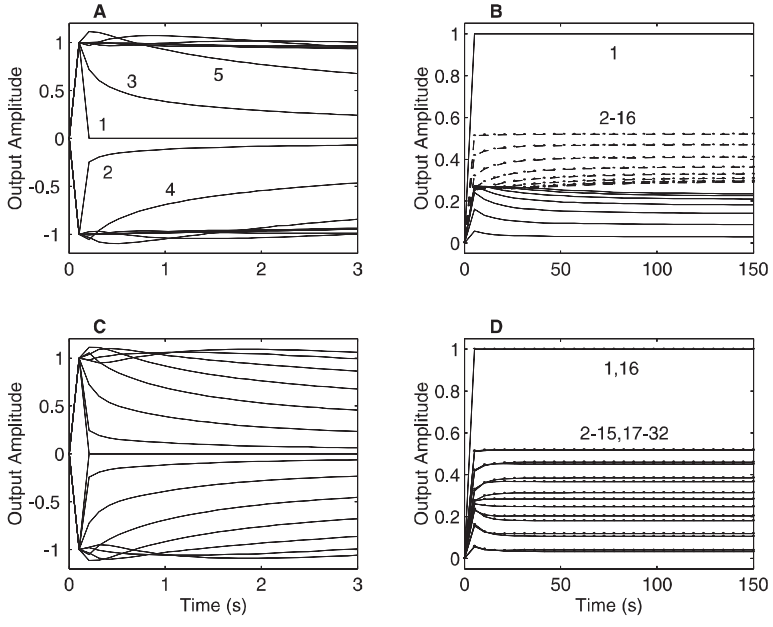

Fig. 5A–D. Impulse and step responses of the 32-neuron integrator with various connections removed. Impulse (**A, C**) and step (**B, D**) responses are computed with opposite- and same-direction inputs, respectively. *Numerals near the curves* denote individual neurons and are included where space permits. **A** Opposite-direction impulse response of the 32-neuron integrator with the reciprocal connections [weights $w(d)$] to and from neuron 1 removed. The responses of neurons 1 through 16 are plotted (those of neurons 17 through 32 are identical). **B** Same-direction step response of the 32-neuron integrator with the reciprocal connections to and from neuron 1 removed. The responses of neurons 1 through 16 are plotted (those of neurons 17 through 32 are identical). Step responses of the odd-numbered neurons are plotted with *solid lines*, while those of the even-numbered neurons are plotted with *dashed lines* and *dot symbols*. **C** Opposite-direction impulse response of the 32-neuron integrator with the reciprocal connections to and from neurons 1 and 16 removed. The responses of neurons 1 through 16 are plotted (those of neurons 17 through 32 are nearly identical). **D** Same-direction step response of the 32-neuron integrator with the reciprocal connections to and from neurons 1 and 16 removed. The responses of all 32 neurons are plotted. Step responses of the odd-numbered neurons are plotted with *solid lines*, while those of the even-numbered neurons are plotted with *dot symbols*

the transfer function dynamics of the neurons in the 32-neuron integrator. This implication is borne out by analysis of the Jordan form.

Eigenvalues and controllability for the 32-neuron integrator, with opposite-direction input connections all present but with the reciprocal connections to and from neuron 1 removed, are shown in Fig. 4C. Like the uniform 32-neuron integrator, the network with the reciprocal connections of neuron 1 removed has all negative eigenvalues, implying that it is stable. Also, all eigenvalues are real, because the system matrix is still symmetric after the network is made nonuniform by removing the reciprocal connections to and from neuron 1. Unlike the uniform network, the eigenvalues in the network with the reciprocal connections of neuron 1 removed are no longer grouped precisely in pairs, and 17 of the 32 modes spanning the range are now controllable. This indicates that the pairing and uncontrollability of modes was due to regularity in the structure of the uniform network. The transfer function dynamics of the neurons in the nonuniform network will be governed by many modes, associated with time constants distributed over many orders of magnitude in value. As such, the network has the potential to produce the kind of high-order dynamics that approximates fractional integration, thereby better simulating the dynamics that have been observed for MVN/NPH neurons.

Bode plots of the frequency responses of all 32 neurons in the 32-neuron integrator, with opposite-direction inputs all present but with the reciprocal connections to and from neuron 1 removed, are shown in Fig. 3E,F. For the disconnected neuron (neuron 1), gain (Fig. 3E) is flat and phase (Fig. 3F) is 0 over most of the frequency range, implying that the disconnected neuron is passing its inputs without integrating them. The frequency responses of many of the other, connected neurons (neurons 2–32) approximate those of fractional integrators over a broad frequency range. This implies that the controllable modes of the 32-neuron network with one neuron disconnected are acting as low-pass filters in parallel. The gain slopes for many of the other neurons are less negative than -1, implying fractional orders of integration between 0 and 1. Correspondingly, the phases show across-frequency lags < 90 deg, with near-parallel shifts in phase between individual neurons spanning 45 deg or more, that resemble the phase characteristics of the MVN/NPH neurons that compose the actual oculomotor neural integrator.

The nonuniform 32-neuron network, with one neuron disconnected, could be used as a model for the actual integrator, except that it is unbalanced. The neurons that compose the actual integrator are situated on opposite sides of the brainstem, and the two sides receive inputs of opposite polarity. This is equivalent to an arrangement of the model in which the odd- and even-numbered neurons are segregated onto opposite sides (Cannon et al. 1983). By disconnecting neuron 1, the even-numbered neurons are inhibited less than the odd-numbered neurons, and this unbalances the network. This is shown for the same-direction step responses of the 32-neuron network, with one neuron disconnected, in Fig. 5B. The disconnected neuron (neuron 1) comes to a new steady state (of 1) with the single-neuron time constant $\tau = 5$ ms. As for opposite-direction inputs, the 32-neuron network with one neuron disconnected also has 17 controllable modes for same-direction inputs (not shown). Therefore, instead of coming to a new steady state quickly with one short time constant, as in the uniform 32- and 2-neuron integrators (Fig. 2D), the 32-neuron network with neuron 1 disconnected takes over 1 min to reach steady state (see Discussion). This steady state is higher for the even-numbered (dashed lines/dot symbols) than for the odd-numbered (solid lines) neurons, and this imbalance would send an inappropriate eye movement command to the eye muscle motoneurons. Fortunately, this imbalance can be corrected simply by also disconnecting the neuron on the opposite side of the network.

The same-direction step responses of the 32-neuron network, with neurons 1 and 16 disconnected, are shown in Fig. 5D. The disconnected neurons (neurons 1 and 16) come to a new steady state (of 1) rapidly. The other neurons reach steady state more rapidly with two neurons than with one neuron disconnected, because the further loss of positive feedback from the network reduces the length of the longest time constants (see below and Discussion). Importantly, for the network with neurons 1 and 16 disconnected, the steady states reached by the even-numbered (dot symbols) and odd-numbered (solid lines) neurons are identical. Thus, imbalance can be avoided by removing reciprocal connections from neurons on opposite sides, and fortunately, the desirable properties of the nonuniform network are retained.

The opposite-direction impulse response of the 32-neuron integrator, with all inputs connected but with the reciprocal connections to and from neurons 1 and 16 removed, are reproduced in Fig. 5C. As for the network with only one neuron disconnected, the decays of the neurons closer to the disconnected neurons are more rapid and are composed of sums of exponential decays with broad distributions of time constants. The eigenvalues (the opposite reciprocals of the time constants) of the nonuniform network with two neurons disconnected are shown in Fig. 4D. The absolute value of the smallest eigenvalue is not quite as small, and consequently the longest time constant is not quite as long, in the network with two as in the network with one neuron disconnected (Fig. 4C). This results because the removal of an additional set of reciprocal connections reduces the amount of positive feedback in the network. As in the network with one neuron disconnected, the eigenvalues in the network with two neurons disconnected are unpaired, and 17 of the 32 modes spanning the range are controllable. This increased number of controllable modes is due to the nonuniformity of the network with two neurons disconnected, and these controllable modes contribute to the transfer function dynamics of the neurons.

Bode plots of the frequency responses of all 32 neurons in the nonuniform 32-neuron integrator, with opposite-direction input connections all present but with

the reciprocal connections to and from neurons 1 and 16 removed, are shown in Fig. 3G,H. They are similar to those in the network with only one neuron disconnected in that they also approximate the frequency responses of fractional integrators over a broad frequency range. In the network with two neurons disconnected, the phase characteristics of the other, connected neurons (Fig. 3H) also show across-frequency lags < 90 deg and near-parallel shifts between individual neurons of up to 45 deg or more. They resemble the phase characteristics of the MVN/NPH neurons that compose the actual oculomotor neural integrator. This similarity in approximately fractional-order dynamics suggests that the structure of the actual integrator, although balanced, may also be nonuniform, in the sense that a regular pattern of reciprocal inhibition is interrupted in an even-sided way by unconnected (or absent) neurons.

4 Discussion

The modeling and analysis results show that making the linear, 32-neuron network nonuniform causes its dynamics to become variable and high-order. For certain cases of the nonuniform network, the dynamics of the individual neurons approximate those of fractional-order integrators. The MVN/NPH neurons that compose the actual oculomotor integrator also have dynamics that approximate those of fractional integrators (Anastasio 1994). Thus, nonuniformity in the linear, 32-neuron network model of the oculomotor integrator makes the dynamics of its constituent neurons more realistic.

4.1 Dynamics of MVN/NPH neurons

The frequency responses of MVN/NPH neurons have been studied during sinusoidal head rotation, eliciting the vestibulo-ocular reflex, or during sinusoidal pursuit. Their responses to impulse-like inputs have been studied during saccades. The most salient feature of the frequency responses of those neurons are near-constant shifts in phase lag < 90 deg that are maintained across several decades of frequency (Shinoda and Yoshida 1974; Blanks et al. 1977; Keller and Precht 1979; Lopez-Barneo et al. 1979; McFarland and Fuchs 1992; Stahl and Simpson 1995). Their impulse responses are described as a summation of many exponential decays with broadly distributed time constants (Escudero et al. 1992; McFarland and Fuchs 1992; Stahl and Simpson 1995). This behavior is characteristic of fractional integrators. It is reproduced by nonuniform 32-neuron networks in which the reciprocal inhibitory connections [weights $w(d)$ in Fig. 1C] to and from either one or two neurons have been removed (Figs. 3F,H and 5A,C).

The dynamics of the neurons in the actual integrator are influenced not only by the integrator itself, but also by the vestibular, pursuit, saccadic, and other oculomotor subsystems that provide inputs to the integrator. Oculomotor subsystem dynamics have been excluded from the network model so that the analysis of integrator dynamics would not be obscured, but this makes it difficult to compare the modeling results directly with experimental data. The analysis shows that neurons in the nonuniform network model are similar to the MVN/NPH neurons that compose the actual integrator in that they both have dynamics that approximate fractional-order integration. Formal models of the oculomotor system, which include oculomotor subsystem dynamics and incorporate fractional-order integrator terms (s^{-k}), have been shown to provide superior fits to experimental data (Anastasio 1994).

The critical features shared by model and actual integrator neurons are phase lags < 90 deg that are nearly constant across frequencies and impulse responses composed of many exponential decays with time constants that span the range from very short to very long. These behaviors have been described as approximate fractional integrations of eye velocity, producing signals that are intermediate between velocity and position components (Anastasio 1994). For example, phase lags < 90 deg are intermediate between those of velocity (phase lag $= 0$) and position (phase lag $= 90$ deg). Also, compound impulse responses, with time constants spanning the range from very short to very long, are intermediate between the velocity component (pulse, approximated by a decay with a very short time constant) and the position component (step, approximated by a decay with a very long time constant).

In the traditional view, the neurons that compose the actual oculomotor integrator, such as MVN/NPH neurons, were seen as combining separate eye velocity and eye position components into a so-called pulse-step response (Robinson 1989a). The original integrator network model of Cannon and co-workers (1983) consisted of a single layer of inhibitory neurons. As shown above, the neurons in this (uniform) network produce a first-order integration of velocity signals and carry only an eye position component. In a subsequent model, Cannon and Robinson (1985) introduced a second, excitatory layer and connected it to the first, inhibitory layer. Neurons in this two-layer network carried separate eye position and eye velocity components. It was considered to be more realistic than the one-layer network in which the neurons carried only the eye position component. More recently, the responses of MVN/NPH neurons have been described as those resulting from fractional integrations of velocity signals, producing signals intermediate between velocity and position. Fractional integrator models match the data better than models in which velocity and position are combined as separate components (Anastasio 1994). Therefore, the one-layer, nonuniform, reciprocal inhibitory integrator network, with neurons having dynamics that approximate fractional integration, may obviate the two-layer excitatory/inhibitory network that Cannon and Robinson (1985) proposed in order to produce a network with neurons that carried separate velocity and position components.

4.2 Model integrator dynamics

Analysis of the 32-neuron integrator provides insight into how a linear network can produce the kind of high-order dynamics that approximates fractional integration [see (4)] and thus better simulates the behavior observed for the neurons composing the actual oculomotor integrator. These high-order dynamics involve many modes that are distributed roughly evenly over many orders on a logarithmic scale (Fig. 4). The 32-neuron integrator can produce this modal diversity because its reciprocal inhibitory connections are both variable and local. With variance $\sigma^2 = 1.51$ as used here [see (3)], connection weights vary with a Gaussian profile over a distance of about 12 neurons and fall nearly to 0 for more distant neurons. For comparison, the modal structure can be examined of networks that also have local profiles (spanning 10, 12, or 14 neurons) but have weights that do not vary (boxcar profile), or of networks that also have Gaussian profiles but with variance so large that each profile encompasses the entire network (global connectivity). Both of these types of networks have modes that are not evenly distributed, and dynamics that do not approximate fractional integration.

Although the uniform 32-neuron integrator with local, Gaussian reciprocal inhibition has a varied modal structure, the diverse modes do not contribute to network dynamics because only one mode (that associated either with the short or the long time constant) is controllable. This lack of controllability results because each neuron inhibits the others in a rigidly regular way. Just as in the box analogy, in which structural uniformity causes inputs to work against each other (see Results), structural uniformity in the 32-neuron network causes neurons to work against each other to prevent expression of the diverse modes. Making the network nonuniform by removing the reciprocal connections of a small number of neurons allows many dynamic modes to become controllable. Breaks in the regular and local innervation pattern of the network unleash the inherent variability of the Gaussian connectivity profiles, and this causes variation between neurons in the weightings of the modes that govern their transfer function dynamics. The modes act as low-pass filters in parallel, and the resulting high-order dynamics approximate those of fractional integrators over a broad range of frequency.

Contrast that with the modal structure of integrator networks in which all reciprocal connection weights have the same absolute value. There are various ways to set up such same-weight networks. It is easiest to think of them as segregated so that neurons receiving inputs of opposite polarity are located on opposite sides of the network. Then, integrator networks can be set up in which the reciprocal connections between neurons on opposite sides are inhibitory, while those between neurons on the same side are excitatory, inhibitory, or nonexistent. These networks also have as many modes as there are neurons. However, in same-weight networks, the modes are associated with one of only three time constants: one long, one short, and one at or near the single-neuron time constant. The absolute value of the reciprocal connection weights, which is the same for all weights, can be adjusted to achieve any desired value for the long time constant.

Making these same-weight networks nonuniform by removing the reciprocal connections of one or more neurons can change the values of some of the modes. For example, neuron disconnection shortens the long time constant by reducing the amount of positive feedback in the network. Neuron disconnection also increases the number of controllable modes. However, it does not alter the simple three-mode structure of these same-weight networks. Thus, whether uniform or nonuniform, these same-weight networks lack the modal diversity needed to approximate fractional integrators and simulate the responses of MVN/NPH neurons.

Integrator networks in which all reciprocal connection weights have the same absolute value are conceptually the simplest. It is easy to conceive of developmental strategies whereby these same-weight networks could self-organize (Anastasio 1997). Integrator networks have been trained using adaptive algorithms to control an eye with first-order dynamics (Arnold and Robinson 1991, 1997). They contain neurons that, in the absence of network connections, have dynamics described by (2) with $\tau = 5$ ms and are linear except that negative states are disallowed. Those networks develop a structure in which opposite- and same-side reciprocal connections are inhibitory and excitatory, respectively, and the weights all have nearly the same absolute value. Because of the sameness of their connection weights, the neurons in those networks would lack the dynamic diversity needed to approximate fractional integrators and simulate the responses of MVN/NPH neurons. It has been suggested that the approximately fractional-order dynamics of MVN/NPH neurons are needed to control the approximately fractional-order dynamics of the eye (Anastasio 1994). It would be interesting to examine the structure of a neural network that had been trained to control an eye with more realistic, fractional-order dynamics.

4.3 Mechanisms of nonuniformity

The linear network model of the oculomotor neural integrator has Gaussian reciprocal connectivity profiles (Cannon et al. 1983) and a diverse modal structure (see Results). The modes contribute to neuron dynamics when they are made controllable, and this can be done by making the network nonuniform. When the network is made nonuniform by removing the reciprocal connections of one or two neurons, the other, connected neurons exhibit approximately fractional integrator dynamics. Similar dynamics are obtained if the reciprocal connections of a small number of neurons are reduced in strength instead of being removed completely, but the resulting approximate fractional integrator dynamics are not maintained over as broad a range of frequency.

There are, of course, other mechanisms for making the uniform network nonuniform. These other mechanisms also increase the number of controllable modes, but the modes contribute to network dynamics in different ways. Exactly how a certain nonuniformity will affect the dynamics depends specifically on the nature of the change made to network structure. For example, randomization of the variances of the Gaussian weight profiles of each neuron causes all of the modes to become controllable while maintaining modal diversity. Perturbing the variances of all 32 Gaussians in the network by adding a uniformly distributed number between $\pm$ 0.5 can produce some spread in neuron phase characteristics, but the frequency range is narrow, and the curves often cross each other and are otherwise far from parallel. Larger perturbations of variance produce complex (as distinct from real) modes and more erratic and unrealistic behavior.

An obvious way of making the 32-neuron integrator network nonuniform is by randomizing its reciprocal connection weights. This makes the network completely controllable, but produces unstable and complex modes that contribute erratic and oscillatory dynamics. Unrealistically erratic and unstable dynamics are observed even with random weight perturbations of only 10%. In networks like the 32-neuron integrator, which rely on some form of positive feedback to achieve time constant lengthening, some connections will have to be held to tight tolerances in order to produce long time constants while avoiding oscillations and instabilities.

Among the many structural nonuniformities that were examined, removing the reciprocal connections of one or two of the neurons in the 32-neuron integrator produced dynamics that most closely approximated fractional integration. Disconnection of three or more neurons also produces a spread in phase lag among the other, connected neurons, which approximates fractional integrators over a broad frequency range. While the phase shifts are near parallel for many of the other neurons, some can develop more complicated phase behavior. Depending upon the specific combination, removing the reciprocal connections of three or more neurons can also cause some neurons to exhibit phase lags in excess of 90 deg, as with input removal, but this is unrealistic (see Results). The reciprocal connections of proportionally more neurons can be removed in larger networks and still produce approximate fractional integration in the other neurons without causing any to exhibit phase lags in excess of 90 deg.

The modeling and analysis suggest that modal diversity results from variability in network connections, as provided by the local, Gaussian connectivity profiles in the linear 32-neuron integrator. The diverse modes act as low-pass filters in parallel, producing approximate fractional integrator dynamics, when the network is made nonuniform by disconnecting one or two neurons. These dynamics resemble those of MVN/NPH neurons, suggesting that the structure of the actual oculomotor neural integrator is also nonuniform in some analogous way. The structure of the actual oculomotor neural integrator remains unknown.

4.4 Consequences of neuron disconnection

Increased modal controllability is not the only consequence of neuron disconnection in the integrator network model. Due to loss of positive feedback, the maximum time constant is shortened when the reciprocal connections of individual neurons are removed, and this also decreases the spread in time constant values. Following disconnection or elimination of individual neurons, the maximum time constant and time constant spread of the nonuniform integrator network can be brought up again by increasing the variance of the Gaussians or the absolute values of the remaining reciprocal connection weights. The variance of the Gaussians ($\sigma^2 = 1.51$) is slightly higher than that ($\sigma^2 = 1.50$) used by Cannon and co-workers (1983) in the original, uniform integrator network model. This slightly larger variance ensures that the maximum time constant will still be long even after one or two neurons are disconnected. Disconnecting neuron 1, or neurons 1 and 16, decreases the maximum time constant from 50.8 to 38.6 or 25.0 s, respectively. These time constant values are typical of those reported for the actual integrator (Becker and Klein 1973; Robinson 1974). Increases in variance are similar to increases in the absolute values of the remaining reciprocal connection weights in that both increase the amount of positive feedback in the network. Increases that are too large will cause the network to become unstable.

4.5 Modeling the lesioned integrator

The focus of this analysis is on the neural integrator that controls horizontal eye movements (see Introduction), but the results may apply to some extent to the vertical integrator as well. For example, Crawford and Vilis (1993) present data showing the effects of incompletely lesioning the interstitial nucleus of Cajal, the brain region responsible for vertical oculomotor integration. These lesions cause a postsaccadic drift in which the eye drifts back, in the direction opposite the saccade, with a multiplicity of time constants. Some of these time constants are quite long, so the eye appears to drift back to a new baseline with each subsequent saccade. Crawford and Vilis (1993) explain these data using a set of separate, parallel subintegrators, each having first-order, low-pass dynamics but differing in their arbitrarily assigned time constants. These data can also be explained using fractional integrators and nonuniform integrator networks.

The role of the neural integrator is to allow the oculomotor system, whose commands are encoded as velocity, to control eye movements (Robinson 1989a,). For example, if premotor circuits, including the neural integrator, and the eye work together as a perfect integrator, then a saccadic command encoded as a velocity pulse will be converted into the desired step of eye position. The neural integrator and the eye can be modeled as fractional-order integrators in series (Anastasio 1994). A perfect integrator can be represented as s (the Laplace

variable) raised to the power −1. A fractional integrator can be represented as s raised to the power $-k$ where $0 < k < 1$. Laplace operators in series multiply, so the order of the combined system can be found simply by adding the exponents of s denoting the fractional-orders of the neural integrator and the eye. If the order of the neural integrator ($-ki$) and the order of the eye ($-ke$) sum to −1, then the overall system will act as a perfect integrator, and a saccadic pulse command will be converted to the desired step of eye position. If a lesion lowers the order of the fractional neural integrator (i.e., makes $-ki$ less negative, and so makes the neural integrator more leaky) then $-ki$ and $-ke$ will sum to a value between −1 and 0, and the overall system will act as a fractional integrator. A fractional integrator would convert a saccadic pulse command into an eye position decay that would appear to be governed by a multiplicity of time constants, some of which are quite long (Anastasio 1994). Thus, the fractional integrator model provides an elegant explanation for the lesion data of Crawford and Vilis (1993).

The purpose of this analysis is to show that neurons in the nonuniform network model of the integrator have dynamics that approximate those of fractional integrators. Individual neurons in the nonuniform network approximate fractional integrators of different orders. The same appears to be true for MVN/NPH neurons (Anastasio 1994), suggesting that the actual oculomotor integrator network is also nonuniform. The network model is made nonuniform by disconnecting one or two neurons, but this is equivalent to eliminating them. Eliminating additional neurons from the network could simulate lesions to the nonuniform integrator. Elimination of additional neurons could cause some of the remaining neurons to exhibit unrealistic dynamics (such as phase lags in excess of 90 deg – see above). However, most of them would continue to have approximately fractional integrator dynamics, but with an average order lower than it was before the additional neurons were eliminated. This can be appreciated by comparing the impulse responses of neurons in the networks in which one or two neurons have been disconnected (eliminated). There are more leaky neurons in the network in which two neurons (Fig. 5C), rather than one neuron (Fig. 5A), have been disconnected. Thus, the average fractional order of integration in the network is lower. Average fractional-order continues to decrease with the disconnection (elimination) of additional neurons from the network (not shown).

The nonuniform integrator model differs in a crucial way from other models that represent the lesioned integrator as a summation over a set of separate, parallel subintegrators, each having first-order, low-pass dynamics but differing in their arbitrarily assigned time constants (Abel et al. 1978; Crawford and Vilis 1993). Presumably, individual neurons in those subintegrators would also have first-order dynamics. In the nonuniform integrator, time constant diversity is an emergent property of the network, and it characterizes the dynamics not only of the overall network, but also of the individual neurons that compose it.

4.6 A prediction of the nonuniform model

In the nonuniform network, numerous modes govern the dynamics of the responses to both opposite- and same-direction input modulations. In uniform 32- and 2-neuron integrator networks, the neurons respond to same-direction modulations by quickly coming to a new steady state (Fig. 2B,D). In contrast, in the nonuniform 32-neuron network, same-direction modulations bring the neurons to a new steady state more slowly, with the slowest neurons requiring about 1 min (Fig. 5B,D). Same-direction modulations represent bilateral, same-direction changes in the carrier levels of the inputs to the actual integrator that might be brought about by changes in arousal or metabolism. Because these changes are themselves likely to be slow, the slowness of nonuniform integrators in responding to them is not problematic. A prediction of the nonuniform integrator model is that individual neurons composing the actual integrator, in the MVN/NPH, should reach a new steady state following same-direction input modulations on time scales ranging up to 1 min (e.g., Fig. 5B,D).

4.7 Evidence for positive feedback

Several lines of evidence suggest that the neural integrator resides in the MVN/NPH, and that it is mediated by positive feedback brought about through reciprocal inhibition (Robinson 1989b). The neurons in these nuclei carry signals appropriate to the task of oculomotor neural integration (Anastasio 1994). In addition, electrolytic lesions or pharmacological inactivations of the MVN/NPH produce deficits in oculomotor neural integration in monkeys and cats (Cheron et al. 1986; Cannon and Robinson 1987; Cheron and Godaux 1987; Straube et al. 1991; Mettens et al. 1994). In monkeys, brainstem commissural fibers, which could mediate reciprocal inhibition in the integrator, cross just caudal to the level of the abducens nuclei (McCrea et al. 1987). Focal electrolytic lesions at this level in monkeys produce pronounced integrator deficits (Anastasio and Robinson 1991; Arnold and Robinson 1997). Lesions of parts of the cerebellum also produce integrator deficits (Robinson 1974; Zee et al. 1981), suggesting that positive feedback loops may pass through and may be calibrated by the cerebellum.

4.8 Conclusion

Experimental findings support the network model of Cannon, Robinson, and Shamma (1983), in which oculomotor integration is brought about by positive feedback through reciprocal inhibition. In that network, every neuron inhibits its neighbors with the same Gaussian profile, but every neuron also has the same first-order dynamics. The specific purpose of this analysis is to show that, when that network is made nonuniform by breaking its regular innervation pattern, the model neurons have diverse, high-order dynamics

that can approximate fractional integrators. The MVN/NPH neurons that compose the actual oculomotor integrator also have dynamics that approximate fractional integrators (Anastasio 1994). Thus, the dynamics of the neurons in the linear network model of the oculomotor integrator can be made more realistic by making the network nonuniform.

Acknowledgements. I thank Drs. Jure Medanic and Petros Voulgaris for consultation and many helpful discussions, and Drs. Medanic, Joseph Malpeli, and Ernst Dow for comments on the manuscript. This work was supported by National Institutes of Health grant MH50577.

References

Abel LA, Dell'Osso LF, Daroff RB (1978) Analog model for gaze-evoked nystagmus. IEEE Trans Biomed Eng 25:71–75

Anastasio TJ (1994) The fractional-order dynamics of brainstem vestibulo-oculomotor neurons. Biol Cybern 72:69–79

Anastasio TJ (1997) Symmetry and self-organization of the oculomotor neural integrator. In: Mira J, Moreno-Díaz R, Cabestany J (eds) Biological and artificial computation: from neuroscience to technology. Springer, Berlin Heidelberg New York, pp 116–123

Anastasio TJ, Robinson DA (1991) Failure of the oculomotor neural integrator from a discrete midline lesion between the abducens nuclei in the monkey. Neurosci Lett 127:82–86

Arnold DB, Robinson DA (1991) A learning network model of the neural integrator of the oculomotor system. Biol Cybern 64:447–454

Arnold DB, Robinson DA (1997) The oculomotor integrator: testing of a neural network model. Exp Brain Res 113:57–74

Becker W, Klein HM (1973) Accuracy of saccadic eye movements and maintenance of eccentric eye positions in the dark. Vision Res 13:1021–1034

Blanks RHI, Volking R, Precht W, Baker R (1977) Response of cat prepositus hypoglossi neurons to horizontal angular acceleration. Neuroscience 2:391–403

Cannon SC, Robinson DA (1985) An improved neural-network model for the neural integrator of the oculomotor system: more realistic neuron behavior. Biol Cybern 53:93–108

Cannon SC, Robinson DA (1987) Loss of the neural integrator of the oculomotor system from brain stem lesions in monkey. J Neurophysiol 57:1383–1409

Cannon SC, Robinson DA, Shamma S (1983) A proposed neural network for the integrator of the oculomotor system. Biol Cybern 49:127–136

Cheron G, Godaux E (1987) Disabling of the oculomotor neural integrator by kainic acid injections in the prepositus-vestibular complex of the cat. J Physiol 394:267–290

Cheron G, Godaux E, Laune JM, Vanderkelen B (1986) Lesions in the cat prepositus complex: effects on the vestibulo-ocular reflex and saccades. J Physiol 372:75–94

Crawford DJ, Vilis T (1993) Modularity and parallel processing in the oculomotor integrator. Exp Brain Res 96:443–456

Escudero M, de la Cruz RR, Delgado-Garcia JM (1992) A physiological study of vestibular and prepositus hypoglossi neurons projecting to the abducens nucleus in the alert cat. J Physiol 458:539–560

Fukushima K, Kaneko CRS (1995) Vestibular integrators in the oculomotor system. Neurosci Res 22:249–258

Kailath T (1980) Linear systems. Prentice-Hall, New Jersey

Kamath BY, Keller EL (1976) A neurological integrator for the oculomotor control system. Math Biosci 30:341–352

Keller EL, Precht W (1979) Adaptive modification of central vestibular neurons in response to visual stimulation through reversing prisms. J Neurophysiol 42:896–911

Lopez-Barneo J, Darlot C, Berthoz A (1979) Functional role of the prepositus hypoglossi nucleus in the control of gaze. Prog Brain Res 50:667–679

Luenberger DG (1979) Introduction to dynamic systems. John Wiley, New York

McCrea RA, Strassman A, May E, Highstein SM (1987) Anatomical and physiological characteristics of vestibular neurons mediating the horizontal vestibulo-ocular reflex of the squirrel monkey. J Comp Neurol 264:547–570

McFarland JL, Fuchs AF (1992) Discharge patterns in nucleus prepositus hypoglossi and adjacent medial vestibular nucleus during horizontal eye movement in behaving macaques. J Neurophysiol 68:319–332

Mettens P, Godaux E, Cheron G, Galiana HL (1994) Effects of muscimol microinjections into the prepositus hypoglossi and the medial vestibular nuclei on cat eye movements. J Neurophysiol 72:785–802

Milsum JH (1966) Biological control systems analysis. McGraw-Hill, New York

Oldham KB, Spanier J (1974) The fractional calculus. Theory and applications of differentiation and integration to arbitrary order. Academic Press, New York

Robinson DA (1974) The effect of cerebellectomy on the cat's vestibulo-ocular integrator. Brain Res 71:195–207

Robinson DA (1989a) Control of eye movements. In: Brooks VB (ed) Handbook of physiology, Sect 1: the nervous system, vol II, part 2. American Physiology Society, Bethesda, pp 1275–1320

Robinson DA (1989b) Integrating with neurons. Annu Rev Neurosci 12:33–45

Shinoda Y, Yoshida K (1974) Dynamic characteristics of responses to horizontal head angular acceleration in vestibuloocular pathway in the cat. J Neurophysiol 37:653–673

Stahl JS, Simpson JI (1995) Dynamics of rabbit vestibular nucleus neurons and the influence of the flocculus. J Neurophysiol 73:1396–1413

Straube A, Kurzan R, Büttner U (1991) Differential effects of bicuculline and muscimol microinjections into the vestibular nuclei on simian eye movements. Exp Brain Res 86:347–358

Zee DS, Yamazaki A, Butler PH, Gucer G (1981) Effects of ablation of flocculus on eye movements in primate. J Neurophys 46:878–899