Communicated by Steve Lisberger

Generalization and Analysis of the Lisberger–Sejnowski VOR Model

Ning Qian
Center for Neurobiology and Behavior, Columbia University, 722 W. 168th Street, New York, NY 10032 USA

Lisberger and Sejnowski (1992) recently proposed a computational model for motor learning in the vestibular–ocular reflex (VOR) system. They showed that the steady-state gain of the system can be modified by changing the ratio of the two time constants along the feedforward and the feedback projections to the Purkinje cell unit in their model VOR network. Here we generalize their model by including two additional time constant variables and two synaptic weight variables, which were set to fixed values in their original model. We derive the stability conditions of the generalized system and thoroughly analyze its steady-state and transient behavior. It is found that the generalized system can display a continuum of behavior with the Lisberger–Sejnowski model and a static model proposed by Miles *et al.* (1980b) as special cases. Moreover, although mathematically the Lisberger–Sejnowski model requires two precise relationships among its parameters, the model is robust against small perturbations from the physiological point of view. Additional considerations on the gain of smooth pursuit eye movement, which is believed to share the positive feedback loop with the VOR network, suggest that the VOR network should operate in the parameter range favoring the behavior studied by Lisberger and Sejnowski. Under this condition, the steady-state gain of the VOR is found to depend on all four time constants in the network. The time constant of the Purkinje cell unit should be relatively small in order to achieve effective VOR learning through the modifications of the other time constants. Our analysis provides a thorough characterization of the system and could thus be useful for guiding further physiological tests of the model.

1 Introduction

The VOR provides an important mechanism for stabilizing visual images on our retinas when we rotate our heads. During each head turn, the system generates a nearly equal and opposite eye movement. The gain of the system, defined as the ratio of the eye velocity to the head velocity, is close to one under normal conditions. It has been demonstrated

that the VOR can be recalibrated when it is inaccurate and images move across the retina during head turns. The recalibration is usually induced in the laboratory by fitting subjects with miniaturizing or magnifying glasses (Miles and Fuller 1974; Gonshor and Jones 1976). The gain of the VOR can become significantly above or below unity. This form of motor learning has been the subject of many recent studies and the main neural circuitry involved has been identified (see Lisberger 1988 for a review). To reconcile apparently contradictory data on the site of VOR learning (Ito 1972; Miles and Lisberger 1981), Lisberger and Sejnowski (1992) recently proposed a computational model based on the simplified VOR network shown in Figure 1. The units in the figure are assumed to be linear and the input–output relationship of each unit is described by a single time constant (see equation 2.1). Here V represents the input head velocity signal from the vestibular system, unit B represents the brain stem neurons that generate the output eye velocity signals, and unit P represents a group of Purkinje cells in the cerebellum that project to the brain stem. There is a feedforward and a feedback pathway to the Purkinje cells. Units T and F represent relay stations along these two pathways. With the simplifying assumption that time constants for P and B (τ_P and τ_B) equal zero and that the connection weights from F to P and from P to B (W_1 and W_2 in Fig. 1) equal one, Lisberger and Sejnowski (1992) showed that the steady-state output of the network is proportional to the ratio of the time constant of unit T to the time constant of unit F. They therefore proposed that modifications of these time constants may provide a major contribution to VOR learning. In this paper, we remove their simplifying assumptions on the values of τ_P, τ_B, W_1, and W_2, and investigate the stability, the steady-state behavior and the transient behavior of the generalized system. We will describe some interesting new features found in the generalized system and discuss their biological implications. We will show that the generalized system can display a continuum of behavior with the original Lisberger–Sejnowski model and a static model proposed by Miles *et al.* (1980b) as special cases. Our analysis of the network transient behavior further establishes the robustness of the Lisberger–Sejnowski model under physiological conditions. A few related results on smooth pursuit will also be presented.

2 Formulation and Results

Following Lisberger and Sejnowski (1992), the output of each unit in Figure 1 is determined from its total input according to the linear relationship:

$$o(t) = i(t) * e^{-t/\tau} \tag{2.1}$$

where $*$ denotes convolution, $i(t)$ and $o(t)$ represent the total input to and the output from the unit, respectively, and τ is the time constant of

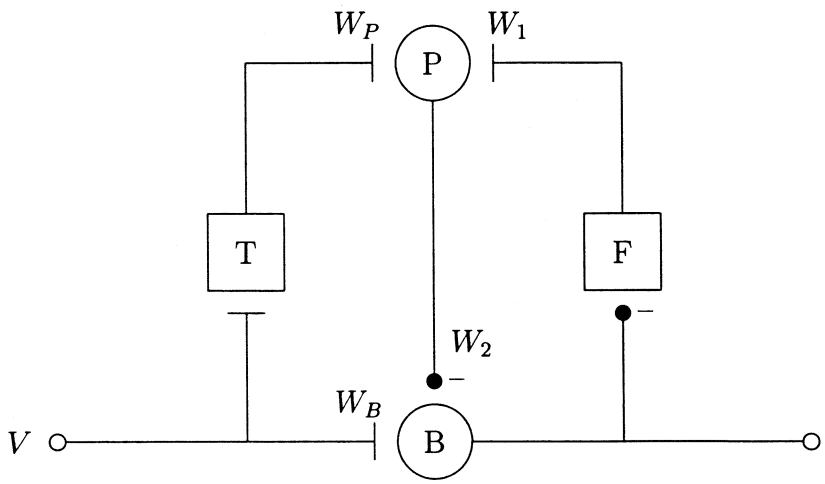

Figure 1: The model VOR network used by Lisberger and Sejnowski (1992). Unit V represents the head velocity signal from the vestibular system, unit B the brain stem neurons which generate eye velocity signals as the output of the system, and unit P a group of Purkinje cells in the cerebellum. Units T and F represent the relay stations along the feedforward and the feedback pathways to the Purkinje cells. W's stand for the magnitude of synaptic weights between the units. The inhibitory connections are shown as filled dots and are labeled with a negative sign.

the unit. It is convenient to analyze the system using Laplace transform, which converts convolutions into multiplications. Throughout the paper, we will use $B(t)$, $P(t)$ etc. to represent the time-dependent responses of the units, and $B(s)$, $P(s)$ etc. to represent the corresponding Laplace transforms. The set of equations governing the dynamics of the network in Figure 1 is then

$$T(s) = V(s)\frac{1}{s\tau_T + 1} \tag{2.2}$$

$$P(s) = [W_P T(s) + W_1 F(s)]\frac{1}{s\tau_P + 1} \tag{2.3}$$

$$B(s) = [W_B V(s) - W_2 P(s)]\frac{1}{s\tau_B + 1} \tag{2.4}$$

$$F(s) = -B(s)\frac{1}{s\tau_F + 1} \tag{2.5}$$

where s is the variable of the Laplace transform, τs are the time constants for the four types of units, and Ws are the synaptic weights as labeled in Figure 1. Note that the connections from input V to unit T or from unit B to unit F are assumed to be fixed at a magnitude of 1. It is not necessary to introduce additional variables to represent these two weights because the linearity of the units would allow them to be absorbed into W_P and W_1, respectively. They would only affect the responses of the T and F units but would not add anything new to the behavior of the network output represented by unit B. Also note that all the τs and Ws are assumed to be positive. The negative signs for inhibitory connections are explicitly expressed in the above set of equations, instead of being absorbed in the Ws. Solving for $B(s)$, we obtain

$$B(s) = H_{\text{vor}}(s)V(s) \tag{2.6}$$

where the VOR transfer function $H_{\text{vor}}(s)$ is given by

$$H_{\text{vor}}(s) = \frac{(s\tau_F + 1)\{W_B[s^2\tau_T\tau_P + s(\tau_T + \tau_P)] + W_B - W_2W_P\}}{(s\tau_T+1)[s^3\tau_B\tau_F\tau_P+s^2(\tau_B\tau_F+\tau_F\tau_P+\tau_P\tau_B)+s(\tau_B+\tau_F+\tau_P)+1-W_1W_2]} \tag{2.7}$$

The temporal responses of the units in the system can be obtained through inverse Laplace transforms.

We now state and derive our results for the VOR network in Figure 1.

2.1 Stability Conditions. An essential requirement for any useful system is stability. We therefore first determine the stability conditions of the network. The stability is defined in this study as the boundedness of the response of the system. That is, the system is considered stable if the response does not diverge with time. We consider the case of W_1W_2 equal to 1 and W_1W_2 not equal to 1 separately.

2.1.1 VOR.

Result 1. *When $W_1W_2 = 1$, the VOR network is stable under sustained and finite vestibular input if and only if $W_B = W_2W_P$.*

The VOR input signal is sustained during the head turn. When $W_1W_2 = 1$, the positive feedback loop BFPB in the network acts as a perfect integrator and the temporal integration of a sustained input would cause divergence with time. Result 1 states that this problem can be resolved by properly adjusting the feedforward pathway to the Purkinje cell unit in the VOR network.

We first prove the necessary condition by noting that to avoid integration in the VOR network, the VOR transfer function should not contain a $1/s$ term in its expansion. This is possible only when there is an s factor in the numerator to cancel that in the denominator in equation 2.7. This

in turn requires $W_B = W_2 W_P$. To prove the sufficient condition, we note that the roots of the denominator in equation 2.7 are

$$s_1 = -\frac{1}{\tau_T} \tag{2.8}$$

$$s_{2,3} = \frac{-(\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B) \pm \sqrt{\Delta}}{2\tau_B\tau_F\tau_P} \tag{2.9}$$

where Δ is defined as

$$\Delta \equiv (\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B)^2 - 4\tau_B\tau_F\tau_P(\tau_B + \tau_F + \tau_P) \tag{2.10}$$

The real parts of all roots can be easily shown to be less than zero. This ensures the stability of the system.

Result 2. *With finite vestibular input $V(t)$, the VOR network is stable when $W_1W_2 < 1$. It is in general unstable when $W_1W_2 > 1$, except when an additional condition is satisfied.*

It is intuitively clear that $W_1W_2 > 1$ should normally cause divergence since W_1W_2 is the gain of the positive feedback loop BFPB in Figure 1. The fact that $W_1W_2 < 1$ guarantees stability is somewhat less obvious. To prove these results mathematically, we consider the roots of the denominator in equation 2.7. We show in the Appendix that when $W_1W_2 < 1$ the real parts of all roots are negative. This ensures the stability of the system. When $W_1W_2 > 1$ positive real part(s) occurs and this will in general cause divergence. The only exception is when there happens to be an identical root in the numerator of the transfer function that cancels the offensive root in the denominator. The condition for the occurrence of this type of cancellation will be derived in Section 2.3.

Note that the stability condition in Result 1 is a generalization of that used by Lisberger and Sejnowski (1992), who let $W_1 = W_2 = 1$ and $W_B = W_P$ (see also Lisberger *et al.* 1994b). There are two equalities in this condition. Since it is unlikely for a biological network to maintain a precise relationship between its parameters we next examine what happens to the stability of the system when the equalities are slightly violated (see also Section 2.3). If W_1W_2 drops below one the system will still be stable because of Result 2. If W_1W_2 becomes larger than one, however, the system will diverge with time exponentially. With perturbation method, the time constant of the divergence can be found to be approximately:

$$\tau \approx \frac{\tau_B + \tau_F + \tau_P}{W_1W_2 - 1} \tag{2.11}$$

Clearly, the divergence is slow when W_1W_2 is just slightly above one. For example, with $W_1W_2 = 1.01$, and with $\tau_F = 70$ msec and $\tau_B = \tau_P = 0$ as used by Lisberger and Sejnowski (1992), it will take about 7 sec for the divergent term to become significant, much longer than the time scale of the VOR, which is typically less than a second.

We next consider how fast the system will diverge with time if the requirement $W_B = W_2W_P$ is not precisely satisfied. By calculating the coefficient of the $1/s$ term in the expansion of $B(s)$, it can be shown that the diverging term is given by

$$\frac{W_B - W_2W_P}{\tau_F + \tau_P + \tau_B} \int_0^t V(t)\,dt \tag{2.12}$$

which is equal to

$$(W_B - W_2W_P)V_0 \frac{t}{\tau_F + \tau_P + \tau_B} \tag{2.13}$$

for constant vestibular input $V(t) = V_0$ (for $t > 0$). For the VOR with short durations relative to $\tau_F + \tau_P + \tau_B$ and for relatively small $W_B - W_2W_P$ this term will not pose a serious problem. For example, with $W_B - W_2W_P = 0.01$, and with $\tau_F = 70$ msec and $\tau_P = \tau_B = 0$, the term is equal to 10% of V_0 after 700 msec. We conclude that small violations of the stability condition in Result 1 will not completely break down the VOR system over the time scale of the normal VOR.

2.1.2 Smooth Pursuit. Evidence suggests that the positive feedback loop BFPB in Figure 1 is also involved in maintaining smooth pursuit eye movements (Lisberger and Fuchs 1978a,b; Lisberger *et al.* 1987). Smooth pursuit is activated by retinal image motion generated by an outside moving target. Under the normal closed-loop condition, the retinal image velocity is equal to the target velocity $U(t)$ minus the eye tracking velocity $B(t)$. If we assume that this error signal enters the network through unit B with a weighting factor W'_B, the set of equations governing the dynamics of pursuit can then be obtained by replacing $V(s)$ with $U(s) - B(s)$ and W_B with W'_B, and by setting $W_P = 0$ in equations 2.2–2.5. (If, instead, we assume that the error signal $B(S) - U(S)$ enters the network through unit P with a weighting factor W'_P (Lisberger *et al.* 1987), results identical to those shown in Results 3 and 6 below can be obtained except W'_B should be replaced by $W_2W'_P$. All conclusions remain the same.) The new set of equations so obtained can then be solved to obtain

$$B(s) = H_{\text{sp}}(s)U(s) \tag{2.14}$$

where the pursuit transfer function is given by

$$H_{\text{sp}}(s) = \frac{W'_B(s\tau_F + 1)(s\tau_P + 1)}{(s\tau_B + 1)(s\tau_F + 1)(s\tau_P + 1) + W'_B(s\tau_F + 1)(s\tau_P + 1) - W_1W_2} \tag{2.15}$$

Using the same method for proving Result 2, we obtain the following stability condition for smooth pursuit:

Result 3. *The smooth pursuit system is stable when* $W_1W_2 < 1 + W'_B$ *and is unstable when* $W_1W_2 \geq 1 + W'_B$.

Note that smooth pursuit is stable over a wider parameter range than the VOR. The difference is caused by the fact that while the vestibular input to the VOR is sustained during the head turn, the retinal error signals to the pursuit system decrease as soon as the eye starts tracking. This effective negative feedback of pursuit increases its parameter range of stability. However, to ensure the stability of both the VOR and smooth pursuit it is necessary to require that $W_1W_2 \leq 1$.

Unlike the VOR case (see Result 2), the pursuit system is always unstable when $W_1W_2 \geq 1 + W_B'$ without exception. The cancellation of the offensive root in the denominator of the pursuit transfer function will not happen since the offensive root is nonnegetive while the two roots in the numerator are both negative.

2.2 Steady-State Behavior. We now investigate the steady-state responses of the system under each of the stability conditions stated above.

2.2.1 VOR.

Result 4. *Under the condition that* $W_1W_2 = 1$ *and* $W_B = W_2W_P$, *the steady-state gain of the VOR is*

$$G_{\text{vor}} = \frac{\tau_T + \tau_P}{\tau_F + \tau_B + \tau_P} W_B \tag{2.16}$$

for constant vestibular input.

Consider the constant vestibular input $V(t) = V_0$ (for $t > 0$). The Laplace transform of this input is simply $V(s) = V_0/s$. Equations 2.6 and 2.7 then become

$$B(s) = \frac{(s\tau_F + 1)(s\tau_T\tau_P + \tau_T + \tau_P)W_BV_0}{s(s\tau_T+1)[s^2\tau_B\tau_F\tau_P + s(\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B) + \tau_B + \tau_F + \tau_P]} \tag{2.17}$$

To derive the above gain expression, let $f(s)$ denote the numerator and $g(s)$ the denominator of equation 2.17. The steady-state output of the VOR network is then given by

$$B(t \to \infty) = \frac{f(0)}{g'(0)} = \frac{\tau_T + \tau_P}{\tau_F + \tau_P + \tau_B} W_BV_0 \tag{2.18}$$

Equation 2.16 is obtained using the fact that the steady-state VOR gain is defined as $B(t \to \infty)/V_0$.

As we mentioned before, the condition that $W_1W_2 = 1$ and $W_B = W_2W_P$ is a generalization of that used by Lisberger and Sejnowski (1992) who assumed $W_1 = W_2 = 1$ and $W_B = W_P$. If we let $\tau_P = \tau_B = 0$, the generalized gain expression, equation 2.16, reduces to the special case $G = W_B\tau_T/\tau_F$, found by Lisberger and Sejnowski. Our result indicates

that the steady-state VOR gain depends on all four time constants in the network. This is perhaps not surprising because the gain is determined by the relative balance between the inhibitory feedforward and the positive feedback loops in Figure 1, and this balance is influenced by all the four time constants. The time constant of the Purkinje cell unit τ_P is special in that it appears on both the numerator and the denominator of the gain expression in equation 2.16. This is a reflection of the fact the projection from unit P to unit B is part of both the feedforward and the feedback loops. Consequently, modifications of τ_P would be the least effective in changing the VOR gain. Furthermore, the value of τ_P determines how effectively the VOR gain can be changed through modifications of the other three time constants. Large τ_P will render any changes in those other time constants insignificant. Our analysis therefore generates the testable prediction that τ_P should be significantly smaller than τ_T, or $\tau_F + \tau_B$, or both, if the modifications of the time constants indeed contribute significantly to the VOR plasticity, as proposed by Lisberger and Sejnowski (1992). A quantitative test of the Lisberger–Sejnowski model requires measurement of all four time constants in the network.

Result 5. *When $W_1W_2 < 1$, the steady-state gain of the VOR is given by*

$$G = \frac{W_B - W_2W_P}{1 - W_1W_2} \tag{2.19}$$

for constant vestibular input.

This result can be derived in the same way as the previous one. It emphasizes the fact that the dependence of the steady-state gain on the time constants shown in equation 2.16 occurs only when $W_1W_2 = 1$ and $W_B = W_2W_P$. If $W_1W_2 < 1$, i.e., if the positive feedback loop is a leaky integrator, we will not require $W_B = W_2W_P$ for the stability of the system and the steady-state gain will be given by equation 2.19.

It is interesting to note that equation 2.19 is essentially identical to the gain expression derived by Miles and co-workers (Miles *et al.* 1980a,b), who used a static model with connectivity similar to Figure 1. Our analysis therefore provides a connection between the dynamic model of Lisberger and Sejnowski and the static model of Miles *et al.* Both of them can be viewed as special instances of the generalized system. We conclude that the difference between the two models is mainly due to Lisberger and Sejnowski's assumption that $W_1W_2 = 1$ and the consequent requirement that $W_B = W_2W_P$. A natural question to ask is how the system switches from one steady-state behavior to the other as W_1W_2 varies from 1 to just slightly below 1. More importantly, is the Lisberger–Sejnowski model robust against small perturbations of the weight variables? These questions will be examined in detail in Section 2.3.

2.2.2 Smooth Pursuit.

Result 6. *When $W_1W_2 < 1 + W'_B$, the steady-state gain of smooth pursuit is given by*

$$G_{sp} = \frac{W'_B}{1 + W'_B - W_1W_2} \tag{2.20}$$

for constant target velocity.

The derivation of this result is the same as that for Result 4.

We showed earlier that the stability of the VOR network requires $W_1W_2 \leq 1$ (see Results 1 and 2). Under this constraint, the smooth pursuit gain $G_{sp} \leq 1$. It is equal to 1 only when $W_1W_2 = 1$. In reality, pursuit gain is close to 1, which implies that W_1W_2 should be close to 1.

2.3 Transient Behavior. We showed in Section 2.1 that the VOR network is stable under two different conditions: (1) $W_1W_2 = 1$ and $W_B = W_2W_P$, or (2) $W_1W_2 < 1$. We further demonstrated that the steady-state gains of the system under these two conditions are quite different and they are given by equations 2.16 and 2.19, respectively. Two related issues need to be addressed to interpret these results correctly. First, the steady-state gains tell us only the asymptotic behavior of the system. They do not tell us how long it takes for the system to settle down to the final states. If a state can be reached only after a time period much longer than the time scale of the VOR, then it is not physiologically relevant. Second, the first stability condition involves two equalities. While it is conceivable that biological learning algorithms could maintain the two equalities approximately, it is unlikely for the equalities to be satisfied exactly. We have shown in Section 2.1 that small violations of the first stability condition will not break down the system over the time scale of the VOR. It is also important to find out how the system behaves when the equalities are only approximately satisfied. In particular, when W_1W_2 changes its value from 1 to slightly below 1, how will the system change its gain from that given by equation 2.16 to that given by equation 2.19? Is the Lisberger–Sejnowski model, which is based on the two equalities of the first stability condition, robust against small changes in the network parameters? We address these issues in this section by studying the transient behavior of the network.

2.3.1 A Special Case with $\tau_P = \tau_B = 0$. To simplify the analysis and to illustrate the key points, we first consider the special case of $\tau_P = \tau_B = 0$. For constant vestibular input $V(t) = V_0$ (for $t > 0$), it can be shown that the exact solution of the network output $B(t)$ is given by

$$\frac{B(t)}{V_0} = W_B\frac{\tau_T}{\tau_F} + W_B\left(1 - \frac{\tau_T}{\tau_F}\right)e^{-t/\tau_T} + (W_B - W_2W_P)\left(\frac{t}{\tau_F} + 1\right)$$
$$\text{for } W_1W_2 = 1 \tag{2.21}$$

and by

$$\frac{B(t)}{V_0} = \frac{W_B - W_2W_P}{1 - W_1W_2} + \frac{(\tau_F - \tau_T)W_2W_P}{\tau_F - \tau_T(1 - W_1W_2)}e^{-t/\tau_T} - \frac{W_1W_2[\tau_T W_B(1 - W_1W_2) - \tau_F(W_B - W_2W_P)]}{(1 - W_1W_2)[\tau_F - \tau_T(1 - W_1W_2)]}e^{-t(1-W_1W_2)/\tau_F} \quad \text{for } W_1W_2 \neq 1 \tag{2.22}$$

First note that these expressions confirm our general conclusions on the VOR stability conditions and steady-state gains stated in Results 1, 2, 4, and 5. When $W_1W_2 = 1$ the system is stable only if $W_B = W_2W_P$ and it will reach the steady-state gain $W_B\tau_T/\tau_F$. The system is also stable when $W_1W_2 < 1$, and the steady-state gain will become $(W_B - W_2W_P)/(1 - W_1W_2)$.

We next examine in detail the transient behavior of the system. It is easy to show that the initial gain of the system, $B(t = 0)/V_0$, is equal to W_B, according to either equation 2.21 or 2.22. This is expected since the direct VOR pathway with a gain of W_B takes immediate action when τ_B is 0. The subsequent behavior of the system depends on the other parameters. Under the first stability condition the system will settle from the initial gain W_B to the final steady-state gain $W_B\tau_T/\tau_F$ through a single exponential decay with a time constant τ_T. Under the second stability condition, on the other hand, the gain will reach the final steady-state value $(W_B - W_2W_P)/(1 - W_1W_2)$ after exponential decay of two terms with time constants $\tau_1 = \tau_T$ and $\tau_2 = \tau_F/(1 - W_1W_2)$, respectively. When W_1W_2 is just slightly below 1, the second time constant is much greater than the first. It can be shown with equation 2.22 that when time t is much smaller than τ_2 the gain of the system is approximately

$$\frac{B(t)}{V_0} \approx W_B\frac{\tau_T}{\tau_F} + W_B\left(1 - \frac{\tau_T}{\tau_F}\right)e^{-t/\tau_T} \tag{2.23}$$

accurate to the first order of $1 - W_1W_2$ and $W_B - W_2W_P$. That is, the solution is approximately equal to that under the first stability condition. Therefore, when W_1W_2 is slightly below 1, the system first settles into a quasi-steady-state similar to that described by Lisberger and Sejnowski (1992) after the quick decay of the fast transient with time constant τ_T. It will then slowly reach the final steady-state identical to that of the static model of Miles *et al.* (1980b) with the decay of the slow transient with time constant $\tau_F/(1-W_1W_2)$. As W_1W_2 is getting closer to 1, it will take longer for the system to reach the steady-state of Miles *et al.*, and consequently the system will spend more time in the quasi-steady-state of Lisberger and Sejnowski. When W_1W_2 becomes 1, the quasi-steady-state becomes a real steady-state and the system will behave like that described by Lisberger and Sejnowski exactly. We conclude that although the steady-state responses of the network under the two different stability conditions are very different, the system is not discontinuous when the product

W_1W_2 changes from 1 to slightly below 1. For the Lisberger–Sejnowski model to work, the system does not have to maintain $W_1W_2 = 1$ and $W_B = W_2W_P$ precisely. So long as the two equalities are approximately satisfied, the system will follow the behavior of the Lisberger–Sejnowski model over the time period $t \ll \tau_F/(1 - W_1W_2)$. The final real steady-state may not be relevant over the time scale of the VOR. For example, assuming that $W_1W_2 = 0.98$ and $\tau_F = \tau_T = 70$ msec as in the original Lisberger–Sejnowski model, the time constant for the slow decay to the final steady-state is 3.5 sec.

Smooth pursuit can also function when W_1W_2 is less than but close to one. The eye position will lag significantly behind the visual target only after extended periods of time and this could be corrected by sacades.

Root Cancellation. The second exponential term in equation 2.22, which is responsible for the final slow decay when W_1W_2 is close to 1, comes from the root $s = (W_1W_2 - 1)/\tau_F$ in the denominator of the VOR transfer function. When $W_1W_2 = 1$, this root becomes zero and the stability of the VOR network requires that $W_B = W_2W_P$ to have a zero root in the numerator to cancel this root in the denominator (see Result 1). Similar cancellation between the numerator and the denominator can also occur for the case of $W_1W_2 \neq 1$ (although it is not required for the sake of stability when $W_1W_2 < 1$) and when this happens the second exponential term in equation 2.22 will disappear. By setting either one of the two roots in the numerator equal to $(W_1W_2 - 1)/\tau_F$, it is easy to show that the exact conditions for the cancellation are

$$\frac{1 - W_1W_2}{\tau_F} = \frac{W_B - W_2W_P}{\tau_T W_B} \tag{2.24}$$

or

$$W_1W_2 = 0 \tag{2.25}$$

(The same conditions can also be obtained by setting the coefficient of the second exponential term in equation 2.22 to zero.) The second condition is not interesting since it implies complete lesion of the positive feedback loop. The first condition makes intuitive sense since it implies the identity of the gain expression for the Lisberger–Sejnowski model ($W_B\tau_F/\tau_T$) and that for the static model of Miles *et al.* $[(W_B - W_2W_P)/(1 - W_1W_2)]$. Under this condition, the network response settles to the final steady-state after a single exponential decay with the time constant τ_T, regardless of whether W_1W_2 is near 1 or not.

It is also worth pointing out that when equation 2.24 is satisfied, the network is stable even if W_1W_2 is larger than 1 because the canceled root is the one that would otherwise cause divergence. It is the additional condition mentioned in Result 2.

2.3.2 The General Case. Similar analysis on the time courses of the responses can be carried out for the general case. The exact solutions could also be derived but the expressions would be too complex to be useful. Since our results on smooth pursuit suggest that the product W_1W_2 should be very close to 1 for the pursuit gain to be close to 1, we choose to examine the transient behavior of the VOR system when W_1W_2 is around 1. Using perturbation method, we found that the transient behavior of the general system is characterized by the following four time constants:

$$\tau_1 = \tau_T \tag{2.26}$$

$$\tau_2 \approx \frac{1}{s_2}\left[1 - \frac{1 - W_1W_2}{s_2^2(\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B) + 2s_2(\tau_B + \tau_F + \tau_P)}\right] \tag{2.27}$$

$$\tau_3 \approx \frac{1}{s_3}\left[1 - \frac{1 - W_1W_2}{s_3^2(\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B) + 2s_3(\tau_B + \tau_F + \tau_P)}\right] \tag{2.28}$$

$$\tau_4 \approx \frac{\tau_B + \tau_F + \tau_P}{1 - W_1W_2} \tag{2.29}$$

where s_2 and s_3 are given by equation 2.9. The approximations are accurate to the first order of $1 - W_1W_2$. The behavior of the system is similar to the special case discussed above. When W_1W_2 is very close to 1 and W_B is approximately equal to W_2W_P, the system will first reach the quasi-steady-state given by equation 2.16 after the decay of the three fast transients characterized by the time constants τ_1 to τ_3. It will then slowly reach the final steady-state given by equation 2.19 with the time constant τ_4. When W_1W_2 is close to 1, τ_4 will be so large that the final steady-state will have no physiological relevance.

We have performed some computer simulations with the general system and the results are shown in Figures 2 and 3. The parameters used in Figure 2 are $W_1 = W_B = W_P = 1$, $\tau_B = 14$ msec, $\tau_F = 70$ msec, $\tau_P = 2$ msec, and $\tau_T = 41$ msec. The seven curves were obtained with W_2 equal to 1, 0.99, 0.98, 0.96, 0.9, 0.8, and 0.5, respectively. Under this set of parameters, the quasi-steady-state gain is approximately 0.5 and the final steady-state gain is 1.0. We see from the figure that when W_1W_2 is close to 1, the system first reaches a gain of about 0.5, which is then slowly moving toward the value 1. As W_1W_2 gets smaller, the decay to the final state becomes faster. A similar set of simulations is shown in Figure 3 where $W_2 = W_B = W_P = 1$, $\tau_B = 14$ msec, $\tau_F = 70$ msec, $\tau_P = 2$ msec, and $\tau_T = 41$ msec. The seven curves were obtained with W_1 equal to 1, 0.99, 0.98, 0.96, 0.9, 0.8, and 0.5, respectively. Under this set of parameters, the quasi-steady-state gain is 0.5 and the final steady-state gain is 0. Again, when W_1W_2 is close to 1, the system first reaches a gain of around 0.5 and then very slowly decays toward a gain of 0.

Root Cancellation. Similar to the special case described in Section 2.3.1, in the general case, the root in the denominator of the transfer func-

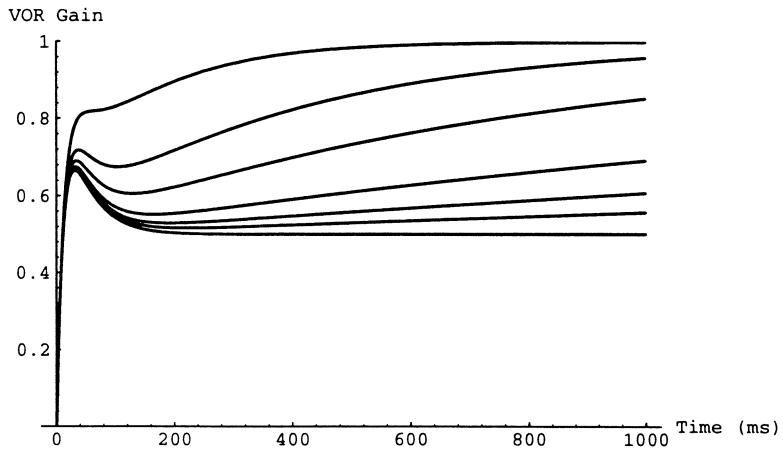

Figure 2: The VOR gains under a step vestibular input. The network parameters are so chosen (see text) such that the expressions in equations 2.16 and 2.19 are equal to 0.5 and 1, respectively. The seven curves counting from the bottom are generated with W_1W_2 equal to 1, 0.99, 0.98, 0.96, 0.9, 0.8, and 0.5, respectively. It can be seen that when $W_1W_2 < 1$ the gain will eventually approach the value of equation 2.19. However, for W_1W_2 very close to 1, the system practically behaves like equation 2.16 over a time period of less than a second.

tion that generates the slow decay term in the inverse transform may also be canceled by one of the three roots in the numerator in equation 2.7. The three corresponding conditions for the cancellation can be easily derived under the approximation that $1 - W_1W_2$ and $W_B - W_2W_P$ are small. Two of the three cancellation conditions are not physiologically plausible, however, because they require the positive feedback loop to become a negative feedback loop. The remaining condition is a generalization of equation 2.24 and is given by

$$\frac{1 - W_1W_2}{\tau_F + \tau_B + \tau_P} \approx \frac{W_B - W_2W_P}{(\tau_P + \tau_T)W_B} \tag{2.30}$$

Similar to equation 2.24 for the special case above, this condition also implies that the gain expression in equation 2.16 for the first VOR stability condition is equal to that in equation 2.19 for the second VOR stability condition. For this reason, we suspect that the condition is probably exact even though it was derived with an approximation method. Under this condition, the system quickly settles to the final real steady-state independent of the value of W_1W_2.

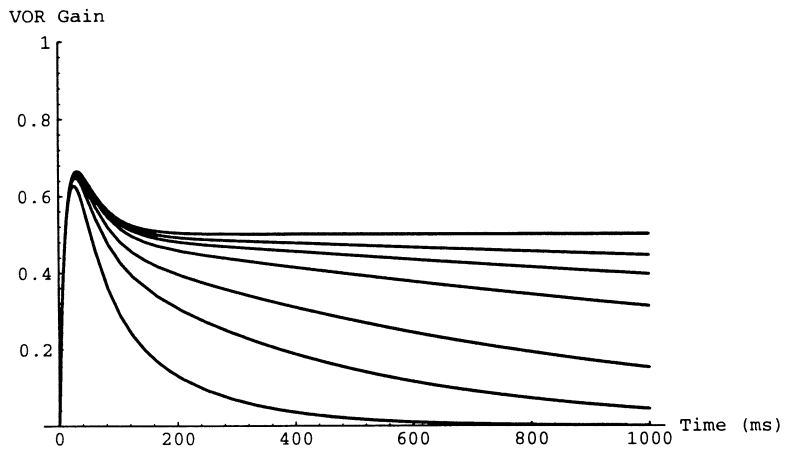

Figure 3: This figure is the same as Figure 2 except that the network parameters are so chosen (see text) such that the expressions in equations 2.16 and 2.19 are equal to 0.5 and 0, respectively. The seven curves counting from the top are generated with W_1W_2 equal to 1, 0.99, 0.98, 0.96, 0.9, 0.8, and 0.5, respectively. It can be seen that when $W_1W_2 < 1$ the gain will eventually approach the value of equation 2.19. However, for W_1W_2 very close to 1, the system practically behaves like equation 2.16 over a time period of less than a second.

Oscillation. Another type of transient behavior the system may have is oscillation. This happens when the/some roots of the transfer function are complex. For example, when Δ in equation 2.10 is negative, the roots s_2 and s_3 in equation 2.9 form a complex conjugate pair. Sinusoidal functions with frequency $\beta = \sqrt{-\Delta}/2\tau_B\tau_F\tau_P$ will then appear in the temporal response $B(t)$. However, since there is always a negative real part associated with each of the complex roots, the oscillation will be quickly dumped. In fact, we performed simulations with a wide range of parameters and the effect of oscillation is never very significant.

VOR Overshoot. Finally, we consider the experimental observation that when the gain of the VOR is low, the initial eye velocity during a head turn overshoots its steady-state level (Lisberger 1988). This overshoot disappears under the high gain condition. Lisberger and Sejnowski (1992) noticed that their network model displayed similar behavior. Here we would like to provide a simple mathematical explanation for this observation. For simplicity we assume $\tau_P = \tau_B = 0$ as in the original model of Lisberger and Sejnowski. The solution to the network output is given

in equation 2.21. Under the stability requirement stated in Result 1, the third term, which diverges with time, becomes zero. The first two terms are the steady-state and the transient VOR responses, respectively. Using the steady-state gain expression

$$G = W_B \frac{\tau_T}{\tau_F} \tag{2.31}$$

equation 2.21 can be written as

$$B(t) = G + (W_B - G)e^{-t/\tau_T} \tag{2.32}$$

When the gain G is low, $(W_B - G)$ is positive. There is therefore a transient overshoot that decays away exponentially. With increased VOR gain, the amplitude of the overshoot, $(W_B - G)$, decreases. The overshoot disappears when the gain reaches $G = W_B$, and further increases in gain will generate an undershoot.

3 Discussion

In this paper, we generalized the Lisberger–Sejnowski model for VOR learning by removing their simplifying assumptions $\tau_P = \tau_B = 0$ and $W_1 = W_2 = 1$, and investigated the properties of the generalized network analytically. We found that the generalized system can display a continuum of behavior including that of the original Lisberger–Sejnowski model and the static model proposed by Miles *et al.* (1980b). Specifically, we showed that the VOR network is stable under either one of the two conditions: (1) $W_1W_2 = 1$ and $W_B = W_2W_P$, or (2) $W_1W_2 < 1$. The first condition is a generalization of that used by Lisberger and Sejnowski. Under this condition, the steady-state gain is given by equation 2.16, which depends on the time constants of all the units in the network. Under the second stability condition, on the other hand, the steady-state VOR gain is given by equation 2.19, which is equivalent to the gain of the static model proposed by Miles *et al.* The difference between the Lisberger–Sejnowski model and that of Miles *et al.* is therefore mainly due to Lisberger and Sejnowski's assumption that $W_1W_2 = 1$ and the consequent requirement that $W_B = W_2W_P$.

Although the steady-state VOR responses under the two stability conditions are quite different, the system is not discontinuous when W_1W_2 varies from 1 to slightly below 1. Our analyses and simulations of the network transient behavior demonstrate that when W_1W_2 is very close to 1 the model works in almost the same way as when W_1W_2 is exactly 1, over the time scale of the normal VOR. Although the system will eventually decay to a different steady-state, the process is too slow to have any physiological significance. Thus, for practical purposes, the network behavior described by Lisberger and Sejnowski is robust against small perturbations in the model parameters. Of course, when W_1W_2 becomes

significantly smaller than 1 the system's behavior will be very different. Under this condition, the network can quickly settle to the steady-state identical to that of the static model described by Miles *et al.* (1980b).

Since the positive feedback loop in the VOR network is believed to be involved in smooth pursuit eye movement as well, we also calculated the pursuit gain under the closed-loop condition. The result shown in equation 2.20 suggests that W_1W_2 should be close to 1 to keep the pursuit gain close to 1. Thus the VOR network is likely to operate in the mode similar to that described by Lisberger and Sejnowski (1992). Under this condition, we found that the VOR gain depends on all four time constants in the system as shown in equation 2.16. The gain could thus be changed through modifications of many possible combinations of the time constants. However, the fact that the VOR learning does not affect smooth pursuit indicates that the three time constants involved in the temporal response of pursuit, τ_B, τ_F, and τ_P, are not modified (Lisberger 1994) during the learning process. Since τ_P appears on both numerator and denominator in equation 2.16, we require its value to be small so that the modification of the fourth time constant τ_T could effectively change the gain. In this connection, it is interesting to note that stimulation of the flocculus and the ventral paraflocculus (corresponding to the P unit in the model network) evokes an inhibitory response in the flocculus target neurons in the brain stem (corresponding to the B unit) with a time delay of only about 2 msec (Lisberger *et al.* 1994a). The physiological value of τ_P could therefore be as small as a couple of milliseconds.

Our analysis provides a fairly complete picture of how the model VOR network behaves over the entire parameter range. It could thus be useful for guiding further physiological tests of the model and for interpreting new physiological data on the VOR learning.

Appendix

Proof of Result 2. As discussed in the text, in order to prove Result 2, we need to examine the roots of the denominator of equation 2.7. If the real parts of all roots are negative, the system is stable. On the other hand, if any root has a positive real part, the system is unstable unless the root is canceled by an identical term in the numerator. We need the following theorem (Korn and Korn 1961) to prove our result: All the roots of the nth degree algebraic equation with real coefficients:

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (a_0 \neq 0) \tag{A.1}$$

have negative real parts, if and only if the same is true for the $(n-1)$st degree equation:

$$a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n - \frac{a_0}{a_1}a_3x^{n-2} - \frac{a_0}{a_1}a_5x^{n-4} - \cdots = 0 \tag{A.2}$$

The first term in the denominator of equation 2.7 always gives a negative root. We focus on the roots of the second term:

$$\tau_B\tau_F\tau_P s^3 + (\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B)s^2 + (\tau_B + \tau_F + \tau_P)s + 1 - W_1W_2 = 0 \quad \text{(A.3)}$$

Applying the theorem, we examine the second-order equation:

$$as^2 + bs + 1 - W_1W_2 = 0 \quad \text{(A.4)}$$

with

$$a = \tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B \quad \text{(A.5)}$$

$$b = \frac{\tau_B^2(\tau_F + \tau_P) + \tau_F^2(\tau_P + \tau_B) + \tau_P^2(\tau_B + \tau_F) + 2 + W_1W_2}{\tau_B\tau_F + \tau_F\tau_P + \tau_P\tau_B} \quad \text{(A.6)}$$

The two roots of equation A.4 are given by

$$s_{1,2} = \frac{-b \pm \sqrt{b^2 - 4a(1 - W_1W_2)}}{2a} \quad \text{(A.7)}$$

Since $a > 0$ and $b > 0$, it is clear that if $(1 - W_1W_2) > 0$, both roots have negative real parts, and that if $(1 - W_1W_2) < 0$, then one root is positive, the other negative. This completes the proof.

Acknowledgments

I am grateful to Dr. Richard Andersen for his support and encouragement.

I would also like to thank Drs. Larry Snyder, Steve Lisberger, Terry Sejnowski, and three anonymous reviewers for their helpful comments to early versions of the manuscript. The author is supported by a research grant from the McDonnell-Pew Program in Cognitive Neuroscience. The early part of the work was supported by Office of Naval Research Contract N00014-89-J1236 and NIH Grant EY07492, both to Richard Andersen.

References

Gonshor, A., and Jones, M. G. 1976. Short-term adaptive changes in the human vestibulo-ocular reflex arc. *J. Phsiol.* **256**, 361–376.

Ito, M. 1972. Neural design of the cerebellar motor control system. *Brain Res.* **40**, 80–84.

Korn, G. A., and Korn, T. M. 1961. *Mathematical Handbook for Scientists and Engineers*. McGraw-Hill, New York.

Lisberger, S. G. 1988. The neural basis for learning of simple motor skills. *Science* **242**, 728–735.

Lisberger, S. G. 1994. Neural basis for motor learning in the vestibulo-ocular reflex of primates. III. Computational and behavioral analysis of the sites of learning. *J. Neurophysiol.* **72**, 974–998.

Lisberger, S. G., and Fuchs, A. F. 1978a. Role of primate flocculus during rapid behavioral modification of vestibulo-ocular reflex. I. Purkinji cell activity during visually guided horizontal smooth-pursuit eye movements and passive head rotation. *Neurophysiol.* **41**, 733–763.

Lisberger, S. G., and Fuchs, A. F. 1978b. Role of primate flocculus during rapid behavioral modification of vestibulo-ocular reflex. II. Mossy fiber firing patterns during horizontal head rotation and eye movement. *J. Neurophysiol.* **41**, 764–777.

Lisberger, S. G., and Sejnowski, T. J. 1992. Motor learning in a recurrent network model based on the vestibulo-ocular reflex. *Nature (London)* **360**, 159–161.

Lisberger, S. G., Morris, E. J., and Tychsen, L. 1987. Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annu. Rev. Neurosci.* **10**, 97–129.

Lisberger, S. G., Pavelko, T. A., and Broussard, D. M. 1994a. Responses during eye movements of brain stem neurons that receive monosynaptic inhibition from the flocculus and ventral paraflocculus in monkeys. *J. Neurophysiol.* **72**, 909–927.

Lisberger, S. G., Pavelko, T. A., Bronte-Stewart, H. M., and Stone, L. S. 1994b. Neural basis for motor learning in the vestibuloocular reflex of primates. II. Changes in the responses of horizontal gaze velocity Purkinje cells in the cerebellar flocculus and ventral paraflocculus. *J. Neurophysiol.* **72**, 954–973.

Miles, F. A., and Fuller, J. H. 1974. Adaptive plasticity in the vestibulo-ocular responses of the rhesus monkey. *Brain Res.* **80**, 512–516.

Miles, F. A., and Lisberger, S. G. 1981. Plasticity in the vestibulo-ocular reflex: A new hypothesis. *Annu. Rev. Neurosci.* **4**, 273–299.

Miles, F. A., Braitman, D. J., and Dow, B. M. 1980a. Long-term adaptive changes in primate vestibulo-ocular reflex. IV. Electrophysiological observations in flocculus of adapted monkeys. *J. Neurophysiol.* **43**, 1477–1493.

Miles, F. A., Fuller, J. H., Braitman, D. J., and Dow, B. M. 1980b. Long-term adaptive changes in primate vestibulo-ocular reflex. III. Electrophysiological observations in flocculus of normal monkeys. *J. Neurophysiol.* **43**, 1437–1476.

Received September 20, 1994; accepted November 8, 1994.