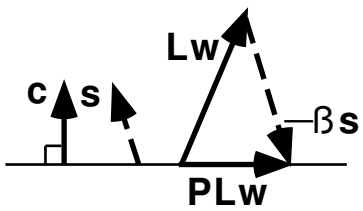

Figure 2: **Projection Onto the Constraint Surface**
The projection operator is $\mathbf{P} = \mathbf{1} - \frac{\mathbf{s}\mathbf{c}^{\mathrm{T}}}{\mathbf{s}\cdot\mathbf{c}}$. This acts on the unconstrained derivative $\mathbf{C}\mathbf{w}$ by removing its $\mathbf{c}$ component, projecting the dynamics onto the constraint surface $\mathbf{c}\cdot\mathbf{P}\mathbf{C}\mathbf{w} = \mathbf{c}\cdot\frac{d}{dt}\mathbf{w} = 0$. This constraint surface is shown as the line perpendicular to $\mathbf{c}$. The constraint is enforced through subtraction of a multiple of $\mathbf{s}$: $\mathbf{P}\mathbf{C}\mathbf{w} = \mathbf{C}\mathbf{w} - \beta\mathbf{s}$ where $\beta = \mathbf{c}\cdot\mathbf{C}\mathbf{w}/\mathbf{c}\cdot\mathbf{s}$. For multiplicative constraints, $\mathbf{s} = \mathbf{w}$; for subtractive constraints, $\mathbf{s} = \mathbf{n}$.

method of constraint enforcement: $\mathbf{s} = \mathbf{w}$ for multiplicative constraints, while $\mathbf{s} = \mathbf{n}$ for subtractive constraints. The constraint vector $\mathbf{c}$ determines the constraint that is enforced: the dynamics remain on the constraint surface $\mathbf{c}\cdot\mathbf{w} = \text{constant}$. When the total synaptic strength is conserved (type 1 constraint), the constraint vector is $\mathbf{c} = \mathbf{n}$; while if the sum–squared strength is conserved (type 2 constraint), the constraint vector is $\mathbf{c} = \mathbf{w}$. Note in Fig. 2 that S1 ($\mathbf{s} = \mathbf{c} = \mathbf{n}$) represents perpendicular projection onto the constraint surface;[5] M1 ($\mathbf{s} = \mathbf{w}$, $\mathbf{c} = \mathbf{n}$) is projection along $\mathbf{w}$; and M2 ($\mathbf{s} = \mathbf{c} = \mathbf{w}$) is both perpendicular projection and projection along $\mathbf{w}$.

2.3 Dynamical effects of multiplicative and subtractive constraints

We begin by illustrating in Fig. 3 the typical dynamics under M1, S1, and M2 in the plane formed by the principal eigenvector of $\mathbf{C}$, $\mathbf{e}^0$, and one other eigenvector with positive eigenvalue, $\mathbf{e}^1$. We have illustrated two typical cases for M1 and S1 constraints:

- In Fig. 3A we have drawn the principal eigenvector $\mathbf{e}^0$ close in direction to the M1 and S1 constraint vector $\mathbf{n}$. This is typical for Hebbian learning when there are only positive correlations and the total synaptic sum is conserved. Positive correlations lead to a principal eigenvector in which all weights have a single sign; this is close in direction to the DC constraint vector $(1, 1, \ldots, 1)^{\mathrm{T}}$, which conserves total synaptic strength. Multiplicative and subtractive constraints lead to very different outcomes in this case: multiplicative constraints lead to convergence to $\mathbf{e}^0$, whereas subtractive constraints lead to unstable flow. The outcome in this case was illustrated in Figs. 1A,B.

[5] One could also consider a non-perpendicular form of S1 constraints, S1′, in which $\mathbf{s}$ and $\mathbf{c} = \mathbf{n}$ are constants but $\mathbf{s} \neq \mathbf{c}$ (and $\mathbf{s}\cdot\mathbf{c} \neq 0$). The results for S1 constraints can be generalized for S1′ constraints without major changes. S1 constraints may turn into S1′ when viewed in alternative synaptic representations; see Appendix B.

$$\frac{d}{dt}\mathbf{w}(t) \;=\; \mathbf{C}\mathbf{w}(t) - \epsilon(\mathbf{w})\mathbf{n} \qquad \text{(Subtractive Constraints)} \tag{4}$$

The vector $\mathbf{n}$ is a constant. For example, if all synapses have equal subtractive decay rate, then $\mathbf{n}=(1,1,\ldots,1)^{\mathrm{T}}$ in the synaptic basis; we refer to this vector, $(1,1,\ldots,1)^{\mathrm{T}}$, as the DC vector.

Multiplicative or subtractive constraints represent two methods of *enforcing* a constraint, that is, of maintaining the weight vector on some constraint surface. We now consider the type of constraint to be enforced. We will focus on two types. First, a constraint may conserve the total synaptic strength $\sum_i w_i$. We refer to this as a type 1 constraint, and to a multiplicative or subtractive constraint of this type as M1 or S1 respectively. These are frequently used in modeling studies, (*e.g.* M1: [5, 40, 22, 23, 33, 34, 41, 43, 44]; S1: [26, 27]). We define a type 1 constraint more generally as one that conserves the dot product $\mathbf{n}\cdot\mathbf{w}$ for some constant constraint vector $\mathbf{n}$; when $\mathbf{n}$ is the DC vector, the conserved quantity is the total synaptic strength. A type 1 constraint corresponds to a hyperplane constraint surface. We choose the subtracted vector $\mathbf{n}$ in Eq. 4 to be the same as the constraint vector $\mathbf{n}$. Then type 1 constraints can be achieved by choosing

$$\text{M1:} \qquad \gamma(\mathbf{w}) \;=\; \mathbf{n}\cdot\mathbf{C}\mathbf{w}/\mathbf{n}\cdot\mathbf{w} \qquad (\text{with } \mathbf{n}\cdot\mathbf{w}(t=0)\neq 0) \tag{5}$$

$$\text{S1:} \qquad \epsilon(\mathbf{w}) \;=\; \mathbf{n}\cdot\mathbf{C}\mathbf{w}/\mathbf{n}\cdot\mathbf{n} \tag{6}$$

These choices yield $\mathbf{n}\cdot\frac{d}{dt}\mathbf{w} = \frac{d}{dt}(\mathbf{n}\cdot\mathbf{w}) = 0$ under Eqs. 3 or 4 respectively.

Second, we consider a constraint that conserves the sum–squared synaptic strength, $\sum_i w_i^2 = \mathbf{w}\cdot\mathbf{w}$; this corresponds to a hypersphere constraint surface. We refer to this as a type 2 constraint (the numbers '1' and '2' refer to the exponent p in the constrained quantity $\sum w_i^p$). This constraint, while not biologically motivated, is often used in theoretical studies, *e.g.* [17, 31]. We will consider only multiplicative enforcement of this constraint,[4] called M2. M2 can be achieved by choosing

$$\text{M2:} \qquad \gamma(\mathbf{w}) \;=\; \mathbf{w}\cdot\mathbf{C}\mathbf{w}/\mathbf{w}\cdot\mathbf{w} \tag{7}$$

This yields $2\mathbf{w}\cdot\frac{d}{dt}\mathbf{w} = \frac{d}{dt}(\mathbf{w}\cdot\mathbf{w}) = 0$ under Eq. 3.

Each form of constrained dynamics can be written $\frac{d}{dt}\mathbf{w} = \mathbf{P}\mathbf{C}\mathbf{w}$, where $\mathbf{P}$ is a projection operator that projects the unconstrained dynamics onto the constraint surface. For S1, the projection operator is $\mathbf{P} = \mathbf{1} - \frac{\mathbf{n}\mathbf{n}^{\mathrm{T}}}{\mathbf{n}\cdot\mathbf{n}}$; for M1, it is $\mathbf{P} = \mathbf{1} - \frac{\mathbf{w}\mathbf{n}^{\mathrm{T}}}{\mathbf{w}\cdot\mathbf{n}}$; and for M2, it is $\mathbf{P} = \mathbf{1} - \frac{\mathbf{w}\mathbf{w}^{\mathrm{T}}}{\mathbf{w}\cdot\mathbf{w}}$. We can write these operators as $\mathbf{P} = \mathbf{1} - \frac{\mathbf{s}\mathbf{c}^{\mathrm{T}}}{\mathbf{s}\cdot\mathbf{c}}$, where $\mathbf{s}$ is the *subtracted* vector, $\mathbf{c}$ the *constraint* vector, and $\mathbf{1}$ the identity matrix (Fig. 2). The projection operator removes the $\mathbf{c}$ component of the unconstrained derivative $\mathbf{C}\mathbf{w}$, through subtraction of a multiple of $\mathbf{s}$. The subtracted vector $\mathbf{s}$ represents the

[4] Subtractive enforcement, S2, does not work in the typical case (Theorem 4) in which the fixed points are unstable. The constraint fails where $\mathbf{n}$ is tangent to the constraint hypersphere, *i.e.* at points where $\mathbf{w}\cdot\mathbf{n}=0$. Such points form a circumference around the hypersphere. The S2 dynamics flow away from the unstable fixed points, at opposite poles of the hypersphere, and flow into this circumference unless prevented by the bounds on synaptic weights. S2 constraints work if weights are non-negative ($w_{\min}=0$) and $\mathbf{n}=(1,1,\ldots,1)^{\mathrm{T}}$; but typically type 2 constraints are used in situations in which weights may take either sign.

2.1 The effect of constant decay terms

Inclusion of synaptic decay results in the equation

$$\frac{d}{dt}\mathbf{w}(t) = \mathbf{C}\mathbf{w} - \epsilon\mathbf{n} - \gamma\mathbf{w} \tag{1}$$

or, in index notation, $\frac{d}{dt}w_i(t) = \sum_j C_{ij}w_j - \epsilon n_i - \gamma w_i$. The term $-\epsilon\mathbf{n}$ is a subtractive decay term; $\mathbf{n}$ is a constant vector, for example $\mathbf{n} = (1, 1, \ldots, 1)^T$ if all synapses decay at the same rate. The term $-\gamma\mathbf{w}$ is a multiplicative decay term.

To examine the effect of these decay terms, we work in the eigenvector basis. Write the initial condition of Eq. 1 as $\mathbf{w}(t=0) = \sum_a w_a^0\mathbf{e}^a$. Write the fixed point of the equation, where $\frac{d}{dt}\mathbf{w} = 0$, as $\mathbf{w}^{\mathrm{FP}} = \sum_a w_a^{\mathrm{FP}}\mathbf{e}^a$. Then the solution to Eq. 1 can be written

$$\mathbf{w}(t) - \mathbf{w}^{\mathrm{FP}} = \sum_a \mathbf{e}^a(w_a^0 - w_a^{\mathrm{FP}})e^{(\lambda^a-\gamma)t} \tag{2}$$

We can compute[2] $\mathbf{w}^{\mathrm{FP}} = \epsilon(\mathbf{C} - \gamma\mathbf{1})^{-1}\mathbf{n} = \sum_a \left(\frac{\epsilon\mathbf{e}^a\cdot\mathbf{n}}{\lambda^a-\gamma}\right)\mathbf{e}^a$, so $w_a^{\mathrm{FP}} = \frac{\epsilon\mathbf{e}^a\cdot\mathbf{n}}{\lambda^a-\gamma}$.

This solution illustrates the basic difference explained in the introduction between multiplicative and subtractive decay terms. The multiplicative decay term lowers the growth rate of all eigenvectors by the same amount γ, thus reducing smaller eigenvalues by a greater proportion. This favors the principal eigenvector. The subtractive term moves the fixed point of the equation, $\mathbf{w}^{\mathrm{FP}}$. This alters the growth rate only of eigenvectors $\mathbf{e}^a$ for which $w_a^{\mathrm{FP}} \neq 0$; these are the eigenvectors with non-zero component in the direction $\mathbf{n}$, that is with $\mathbf{e}^a \cdot \mathbf{n} \neq 0$. That is, constant subtractive decay only alters growth in directions coupled to $\mathbf{n}$, and leaves growth unaltered in all directions perpendicular to $\mathbf{n}$.

Note that neither type of constant decay achieves selectivity: all patterns with $\lambda^a > \gamma$ grow exponentially until the limits on synaptic weights are reached, and all synapses typically reach $w_{\max}$ if the leading eigenvector has no changes of sign. To achieve selectivity, a constraint is needed.

2.2 Formulation of multiplicative and subtractive constraints

We now replace the decay terms above by either a multiplicative or a subtractive constraint term. By multiplicative or subtractive constraints, respectively, we refer to time-varying decay terms $\gamma(t)\mathbf{w}$ or $\epsilon(t)\mathbf{n}$ that move $\mathbf{w}$, after application of $\mathbf{C}$, toward a constraint surface. We assume γ and ϵ are determined by the current weight vector $\mathbf{w}(t)$ and do not otherwise depend on t, so we write them as $\gamma(\mathbf{w})$ or $\epsilon(\mathbf{w})$. Thus, the constrained equations are[3]

$$\frac{d}{dt}\mathbf{w}(t) \quad = \quad \mathbf{C}\mathbf{w}(t) - \gamma(\mathbf{w})\mathbf{w}(t) \qquad \text{(Multiplicative Constraints)} \tag{3}$$

[2] We have assumed here that $\mathbf{n}$ has no components in the nullspace of $\mathbf{C}-\gamma\mathbf{1}$, so that $\left(\mathbf{C}-\gamma\mathbf{1}\right)^{-1}\mathbf{n}$ is well-defined. This is easily generalized.

[3] To understand why the term $-\gamma(\mathbf{w})\mathbf{w}(t)$ represents multiplicative constraints, consider a multiplicatively constrained equation $\mathbf{w}(t+\Delta t) = \beta(\mathbf{w})\left[\mathbf{w}(t) + \mathbf{C}\mathbf{w}(t)\Delta t\right]$, where $\beta(\mathbf{w})$ achieves the constraint. This is identical to $\frac{\mathbf{w}(t+\Delta t)-\mathbf{w}(t)}{\Delta t} = \mathbf{C}\mathbf{w}(t) - \gamma(\mathbf{w})\mathbf{w}(t+\Delta t)$ where $\gamma(\mathbf{w}) = \frac{1-\beta(\mathbf{w})}{\beta(\mathbf{w})\Delta t}$. For $\Delta t \to 0$ this becomes Eq. 3.

visual, teacher signal is dominant, the Hebbian rule becomes $\frac{d}{dt}\mathbf{w} = \langle s\mathbf{x} \rangle$. Here the right-hand side is a constant vector, $\mathbf{t} \equiv \langle s\mathbf{x} \rangle$, which measures the correlation between the auditory inputs and the visual teacher input. We consider the case in which there is initially a rough auditory map, so that the correlation of auditory inputs with the visual teacher decreases as a Gaussian from the center of the arbor. In the absence of constraints, all synapses again reach $w_{\max}$, so no selectivity develops. Multiplicative constraints lead the weights to match the correlations, giving a moderate advantage to the best-correlated inputs. Subtractive constraints lead to a sharpening of the map so that only the best correlated inputs survive.

In summary, unconstrained Hebbian equations often lead all synapses to saturate at the maximal allowed value, destroying selectivity. Multiplicative constraints then lead the inputs to develop graded strengths. Subtractive constraints lead synapses to saturate at either the maximal or minimal allowed value, and can result in a sharpening to a few best-correlated inputs. They also can allow ocular dominance segregation to develop in circumstances where multiplicative constraints do not. These differences between subtractive and multiplicative constraints are easily understood, as we now show.

2 Multiplicative and Subtractive Constraints for a Single Output Cell

We begin with a general linear synaptic plasticity equation without decays, $\frac{d}{dt}\mathbf{w}(t) = \mathbf{C}\mathbf{w}(t)$. $\mathbf{C}$ is a matrix that drives the dynamics. We assume that $\mathbf{C}$ has a complete set of orthonormal eigenvectors $\mathbf{e}^a$ with corresponding eigenvalues λ^a (that is, $\mathbf{C}\mathbf{e}^a = \lambda^a \mathbf{e}^a$), as occurs if $\mathbf{C}$ is symmetric. In Hebbian learning, $\mathbf{C}$ is a correlation matrix: C_{ij} represents the correlation in activity between inputs i and j, so $\mathbf{C}$ is symmetric.[1] Typically most or all of the eigenvalues of $\mathbf{C}$ are positive; for example, if $\mathbf{C}$ is the covariance matrix of the input activities then all its eigenvalues are positive.

We use indices i, j to refer to the synaptic basis, and a, b to refer to the eigenvector basis. The strength of the i^{th} synapse is denoted by w_i. The weight vector $\mathbf{w}$ can also be written as a combination of the eigenvectors, $\mathbf{w} = \sum_a w_a \mathbf{e}^a$, where the components of $\mathbf{w}$ in the eigenvector basis are $w_a = \mathbf{w} \cdot \mathbf{e}^a$. We assume as before that the dynamics are linear up to hard limits on the synaptic weights, $w_{\min} \leq w_i(t) \leq w_{\max}$; we will not explicitly note these limits in subsequent equations.

As discussed in the introduction, a constraint can be enforced through time-varying synaptic decay, where the rate of the decay varies to maintain the constraint. Before studying constraints, it is instructive to review the effects of constant decay terms.

[1] We work in a representation in which each synapse is represented explicitly, and the density of synapses is implicit. Equivalently, the dynamics may be written in a representation in which the synaptic density or "arbor" function is explicitly represented. The current analysis also applies in these representations, as described in Appendix B.

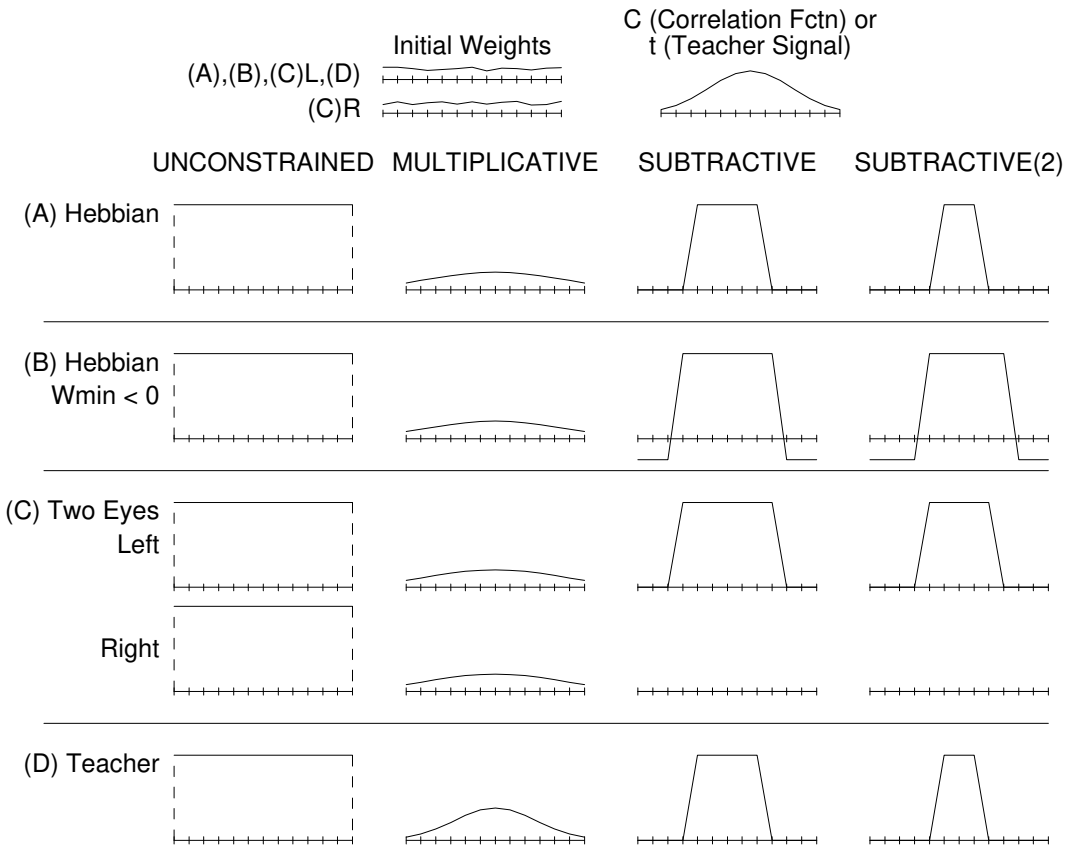

Figure 1: **Outcomes of Development Without Constraints and Under Multiplicative and Subtractive Constraints**

(A), (B): Outcome of a simple Hebbian development equation: unconstrained equation is $\frac{d}{dt}\mathbf{w} = \mathbf{C}\mathbf{w}$. Initial synaptic weights are shown at the top left. The correlation matrix $\mathbf{C}$ is a Gaussian function of the separation between two synapses (shown at top right). (A): $w_{\min} = 0$; (B): $w_{\min} = -2$. (C): Outcome of a similar equation but with two identical sets of inputs, representing left- and right-eye inputs. Within each eye, correlations are the same as in (A); between the eyes there is zero correlation. Unconstrained equations are $\frac{d}{dt}\mathbf{w}^L = \mathbf{C}\mathbf{w}^L$; $\frac{d}{dt}\mathbf{w}^R = \mathbf{C}\mathbf{w}^R$. (D). Outcome of a simplified model of the alignment of one map with another: a Hebbian rule in the limit in which the output cell's activity is determined by a "teacher signal". Unconstrained equation is $\frac{d}{dt}\mathbf{w} = \mathbf{t}$ (see text). The constant vector $\mathbf{t}$ is a Gaussian centered at the center of the arbor (shown at top right). Note that the resulting weights under multiplicative constraints are proportional to $\mathbf{t}$.

All results are from simulations of a two-dimensional receptive field consisting of a diameter-13 circle of inputs drawn from a 13 by 13 square grid. The resulting receptive fields were approximately circularly symmetric; the figures show a slice horizontally through the center of the field. All simulations used $w_{\max} = 8$; all except (B) used $w_{\min} = 0$. The left three columns show results for $w_{\text{init}} = 1$. The right column (subtractive(2)) uses $w_{\text{init}} = 0.5$, which halves the conserved total synaptic strength w_{tot}.

Consider first a simple equation for Hebbian synaptic plasticity, $\frac{d}{dt}\mathbf{w} = \mathbf{C}\mathbf{w}$, where $\mathbf{C}$ is a matrix describing correlations among input activities [18, 20, 27, 31]. Suppose this correlation is a Gaussian function of the separation of two inputs. We assume first that $w_{\min} = 0$. The final outcomes of development under this equation are shown in Fig. 1A. With no constraint (column 1), all synapses saturate at $w_{\max}$, so all selectivity in the cell's response is lost. Under a multiplicative constraint (column 2), synaptic strengths decrease gradually from center to periphery. The final synaptic pattern in this case is proportional to the principal eigenvector of $\mathbf{C}$. Under a subtractive constraint (columns 3 and 4), a central core of synapses saturate at or near strength $w_{\max}$, while the remaining synapses saturate at $w_{\min}$. If $w_{\max}$ is increased, or the total conserved synaptic strength w_{tot} is decreased by decreasing w_{init}, the receptive field is sharpened under a subtractive constraint (column 4). The final number of nonzero synapses in this case (with $w_{\min} = 0$) is approximately $w_{\text{tot}}/w_{\max}$. In contrast, under the multiplicative constraint or without constraints, the shape of the final receptive field is unaltered by such changes in $w_{\max}$ and w_{tot}.

If $w_{\min}$ is decreased below zero, center-surround receptive fields can result under subtractive constraints (Fig. 1B). Results under multiplicative constraints or without constraints are unaltered by this change. This mechanism of developing center-surround receptive fields underlies the results of Linsker [18], as explained in [20, 21] and in section 2.4. Again, an increase in $w_{\max}$ or decrease in w_{tot} leads to sharpening of the positive part of the receptive field under subtractive constraints (column 4).

Next, consider ocular dominance segregation [27] (Fig. 1C). We suppose the output cell receives two equivalent sets of inputs: left-eye inputs and right-eye inputs. A Gaussian correlation function $\mathbf{C}$ describes correlations within each eye as before, while between the two eyes there is zero correlation; and $w_{\min} = 0$. Now results are much as in (A), but a new distinction emerges. Under subtractive constraints, ocular dominance segregation occurs: the output cell becomes monocular, receiving input from only a single eye. Under multiplicative constraints, there is no ocular dominance segregation: the two eyes develop equal innervations to the output cell. Segregation can only occur under multiplicative constraints if there are anticorrelations between the two eyes, as will be explained in section 2.5.

Finally, consider the problem of map alignment (Fig. 1D). In the barn owl optic tectum, an auditory map comes to be aligned with a visual map under the guidance of visual experience [16]. Consider a toy model of this, as follows. Two input maps innervate the tectum. One map (the auditory) is adaptive, while the other map (the visual) is fixed and serves as a teacher to the auditory map. We consider a single output cell in the tectum, with activity y. Let $\mathbf{w}$ be the auditory map's synaptic weight vector, and $\mathbf{x}_n$ its vector of input activities corresponding to the n^{th} stimulus. Let s_n be the synaptic input from the visual map in response to the same stimulus: the visual map is topographic, so the largest responses s_n will come from stimuli in a particular topographic position in visual space. Suppose a simple activation rule, $y = \mathbf{w} \cdot \mathbf{x} + s$, is combined with a Hebbian rule, $\frac{d}{dt}w_i = \langle y x_i \rangle$. The brackets indicate an average over input patterns. In the limit in which the

strength conservation rule in the context of perceptron learning. Subsequently, many authors have used one form or another of similar constraints, *e.g.* [5, 17, 18, 27, 31, 41].

A constraint that conserves total synaptic strength over a cell can be enforced through nonspecific decay of all synaptic strengths, provided the rate of this decay is set for the cell as a whole to cancel the total increase due to specific, Hebbian plasticity. Two simple types of decay can be considered. First, each synapse might decay at a rate proportional to its current strength; this is called multiplicative decay. Alternatively, all synapses might decay at an equal rate; this is called subtractive decay. The message of this paper is that the dynamical effects of a constraint depend significantly on whether it is enforced via multiplicative or subtractive decay. We have noted this briefly in previous work [20, 21, 24, 27, 28].

Intuitively, the difference between these two methods is that multiplicative enforcement of a constraint (a "multiplicative constraint") suppresses the growth of all patterns of synaptic weights, while subtractive enforcement (a "subtractive constraint") only suppresses patterns whose growth would violate the constraint. In each short time Δt, a multiplicative constraint multiplies all weights, or equivalently all weight patterns, by a constant. If the unconstrained dynamics are linear in the synaptic strengths, the dynamical equations are invariant under such a rescaling. Then the dynamics under the multiplicative constraint are equivalent to those of unbounded linear growth. So the synaptic pattern that would grow the fastest in the absence of the constraint (the principal eigenvector of the operator determining unconstrained development) will eventually emerge, while all other components will eventually be suppressed. In contrast, in each short time, a subtractive constraint subtracts a small constant from all weights. This selectively alters the amplitude of weight patterns whose growth would violate the constraint — those that have a nonzero total synaptic strength — while leaving the amplitude of all other weight patterns unaffected.

In this paper, we analyze the effects of subtractive and multiplicative constraints on the dynamics of linear learning rules. We then briefly discuss some theoretical and biological methods of implementing constraints.

1 Simple Examples of the Effects of Constraints

A few simple examples will illustrate that strikingly different outcomes can result from the subtractive or multiplicative enforcement of a constraint. The remainder of the paper will present a systematic analysis of these differences.

Consider synaptic plasticity on a single postsynaptic cell. Let $\mathbf{w}$ be the vector of synaptic weights onto this cell; the i^{th} component, w_i, is the synaptic weight from the i^{th} input. We assume synaptic weights are initially randomly distributed with mean w_{init}, and are limited to remain between a maximum value w_{max} and minimum value w_{min}. We consider the effect of a constraint that conserves the total synaptic strength, $\sum_i w_i$, implemented either multiplicatively or subtractively.

Development in many neural systems appears to be guided by "Hebbian" or similar activity-dependent, correlation-based rules of synaptic modification (reviewed in [24]). Several lines of reasoning suggest that constraints limiting available synaptic resources may play an important role in this development.

Experimentally, such development often appears to be *competitive*. That is, the fate of one set of inputs depends not only on its own patterns of activity, but on the activity patterns of other, competing inputs. A classic example is given by the experiments of Wiesel and Hubel on the effects of monocular versus binocular visual deprivation in young animals ([42]; see also [6]). If neural activity is reduced in one eye, inputs responding to that eye lose most of their connections to the visual cortex, while the inputs responding to the normally active, opposite eye gain more than their normal share of connections. If activity is reduced simultaneously in both eyes for a similar period of time, normal development results: each eye's inputs retain their normal cortical innervation. Such competition appears to yield a roughly constant final total strength of innervation regardless of the patterns of input activity, although the distribution of this innervation among the inputs depends upon neural activities.

There is no direct biological evidence for the existence of competitive limits on the total synaptic strength supported by a cell. However, some support for the idea that a cell's synaptic resources are limited comes from biological evidence for intrinsic limits to the total *number* of synapses supported by a cell. In the retinotectal system in several species, if retinal cells are forced to innervate only a partial tectum, rather than their usual target of the entire tectum, the number of synapses per tectal cell remains normal [8, 30, 32]. This suggests that there are a limited number of *postsynaptic* sites per target cell, for which incoming neurons must compete. If inputs are given a *larger* than normal target structure, they innervate target cells with less than normal density, indicating intrinsic limits in the total *presynaptic* innervation by an input cell [9] (this also occurs in other systems [3, 4, 36]). In several systems, synapse number appears to be conserved regardless of the patterns of neural activity, even while synaptic rearrangements occur that vary with neural activity patterns. These systems include the monkey visual cortex during ocular dominance column development [2] and the goldfish optic tectum during post-regeneration map refinement [11, 10].

The existence of constraints limiting synaptic resources is also suggested on theoretical grounds. Development under simple correlation-based rules of synaptic modification typically leads to instability. Either all synapses grow to the maximum allowed value, or all synapses decay to zero strength. To achieve the results found biologically, a Hebbian rule must instead lead to the development of selectivity, so that some synaptic patterns grow in strength while others correspondingly shrink. One theoretical method of achieving selectivity is to use constraints to enforce a competition. Von der Malsburg [40] first proposed that constraints conserving the total synaptic strength supported by each input or output cell could be used to stabilize a Hebbian rule, a proposal also made by Perez et al. [33]. Much earlier, Rochester et al. [34] used a model that conserved total synaptic strength over all synapses in a Hebbian network. Rosenblatt [35] proposed a synaptic

Abstract

Models of unsupervised correlation-based (Hebbian) synaptic plasticity are typically unstable: either all synapses grow until each reaches the maximum allowed strength, or all synapses decay to zero strength. A common method of avoiding these outcomes is to use a constraint that conserves or limits the total synaptic strength over a cell. We study the dynamical effects of such constraints. Two methods of enforcing a constraint are distinguished, multiplicative and subtractive. For otherwise linear learning rules, multiplicative enforcement of a constraint results in dynamics that converge to the principal eigenvector of the operator determining unconstrained synaptic development. Subtractive enforcement, in contrast, leads to a final state in which almost all synaptic strengths reach either the maximum or minimum allowed value. This final state is often dominated by weight configurations other than the principal eigenvector of the unconstrained operator. Multiplicative enforcement yields a "graded" receptive field in which most mutually correlated inputs are represented, whereas subtractive enforcement yields a receptive field that is "sharpened" to a few maximally-correlated inputs. If two equivalent input populations (*e.g.* two eyes) innervate a common target, multiplicative enforcement prevents their segregation (ocular dominance segregation) when the two populations are weakly correlated; whereas subtractive enforcement allows segregation under these circumstances. An approach to understanding constraints over input and over output cells is suggested, and some biological implementations are discussed.

CALIFORNIA INSTITUTE OF TECHNOLOGY
COMPUTATION AND NEURAL SYSTEMS PROGRAM

OCTOBER 9, 1992 CNS MEMO 19

THE ROLE OF CONSTRAINTS IN HEBBIAN LEARNING

Kenneth D. Miller† and David J. C. MacKay*

† Division of Biology 216-76 and
* Computation and Neural Systems Program
California Institute of Technology
Pasadena, CA 91125

† ken@cns.caltech.edu
* mackay@ras.phy.cam.ac.uk
*Current address: Radio Astronomy, Cavendish Laboratory,
Madingley Road, Cambridge, CB3 0HE. United Kingdom

Submitted, slightly abridged, to *Neural Computation*

PASADENA, CALIFORNIA 91125