Supporting Information

Shoval et al. 10.1073/pnas.1002352107

SI Text

Generalized Fold-Change Detection: Symmetry Invariance. We study general systems with inputs and outputs of the following form (Eq. S1 *A* and *B*):

$$\dot{x} = f(x, u) \quad \textbf{[S1}A\textbf{]}$$

$$y = h(x, u) \quad \textbf{[S1}B\textbf{]}$$

where $u = u(t)$ is a stimulus, excitation, or input function and $y = y(t)$ is a response or output function. We are using here the standard control-theory formalism: typically, y represents a selection of one of the state variables x_i, which quantifies the response of the system. This variable, which is one of the coordinates of x, satisfies a differential equation, and the output map h is of the form $y = h(x) = x_i$. Eqs. **3** and **4** are for this special case, which is also discussed in detail in *Coordinate projection*.

As usual, Eq. **S1** is meant as shorthand for

$$\begin{aligned} \frac{dx}{dt}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t), u(t)) \end{aligned}.$$

The spaces of states, input values, and output values, $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{Y}$, respectively, are subsets of Euclidean spaces $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^q$, respectively, and $u : [0, \infty) \to \mathbb{U}, x : [0, \infty) \to \mathbb{X}, y : [0, \infty) \to \mathbb{Y}$.

Our central question is as follows. Suppose that we are interested in understanding how a certain set of transformations $\mathcal{P}$ or symmetries (Eq. **S2**)*

$$\pi : \mathbb{U} \to \mathbb{U}, u \mapsto \pi u \quad \textbf{[S2]}$$

acting on the space $\mathbb{U}$ of input values affects the response of the system. The set $\pi \in \mathcal{P}$ might constitute, for example, a group of rotations, translations, and/or dilations in an image-recognition system.

Specifically, we are interested in what one might call response invariance to symmetries in $\mathcal{P}$: the system response that is observed after a jump from some constant value of u to a new input $v(t)$ will be the same as if we started instead with the constant value $\pi(u)$ and then jumped to $\pi(v(t))$. For example, suppose that we are watching a distant static image, and suddenly, a target appears in the visual field. Our response should be identical (if response invariance to translations is valid) when we observe the image from a displaced location.

One particular example of interest is $\mathcal{P} = \mathbb{U} = \mathbb{R}_{>0}$ (positive real numbers) and πu = multiplication. The requirement that the response should be the same when jumping from u to v as when jumping from pu to pv, for any $p > 0$ means that the response only depends on the fold change or ratio v/u. We use the terminology fold-change detection (FCD) because of this motivation.

Technical assumptions. We take the functions f, h to be differentiable and make the assumption that, for each input u:$[0, \infty) \to \mathbb{U}$ and each initial state $\xi \in \mathbb{X}$, there is a (unique) solution of the initial value problem (Eq. S1) with initial condition $x(0) = \xi$.† We denote this solution as:

$$\phi(t, \xi, u)$$

and the corresponding output as:

$$\psi(t, \xi, u) = h(\phi(t, \xi, u), u(t))$$

We also make the assumption that, for each constant input u, there is a unique steady state, which we denote as $\sigma(u)$. That is to say, there is a unique solution of $f(x, u) = 0$ given by $x = \sigma(u)$ (Eq. **S3**):

$$f(\sigma(u), u = 0) \quad \textbf{[S3]}$$

We will say that the system is stable‡ if, in addition, it holds that every trajectory approaches $\sigma(u)$ when the constant input $u(t) = u$ is used, which is to say:

$$\lim_{t \to \infty} \phi(t, \xi, u) = \sigma(u) \text{ for all } \xi \in \mathbb{X}, u \in \mathbb{U}.$$

Here and later, we make the abuse of notation of viewing an element $u \in U$ both as an input value $u(t) \in \mathbb{U}$ and as a constant input function u:$[0, \infty) \to \mathbb{U}$; the meaning should be clear from the context.

Main definitions. Suppose a system (Eq. **S1**) and a set of symmetries $\pi \in \mathcal{P}$ as in Eq. **S2**.§

Definition: property FCD is satisfied if the equality

$$\psi(t, \sigma(u), v) = \psi(t, \sigma(\pi u), \pi v) \quad \text{(FCD)}$$

holds for all constants $u \in \mathbb{U}$, all input functions v: $[0, \infty) \to \mathbb{U}$, all $\pi \in \mathcal{P}$, and all $t \geq 0$.¶

A consequence of FCD is as follows. Suppose that we use $v(t) = v$ (constant function) in the definition of FCD . Then, evaluating at $t = 0$ and using that, by definition, $\psi(0, \sigma(u), v) = h(\varphi(0, \sigma(u), v), v(0)) = h(\sigma(u), v)$ (Eq. **S4**):

$$h(\sigma(u), v) = h(\sigma(\pi u), \pi v) \text{ for all } u \in \mathbb{U}, v \in \mathbb{U}, \pi \in . \quad \textbf{[S4]}$$

Definition: the system perfectly adapts to constant inputs if there exists some value $y_0 \in \mathbb{Y}$ so that

$$h(\sigma(u), u) = y_0 \text{ for all } u \in \mathbb{U}$$

Remark: suppose that a system perfectly adapts to constant inputs and also, that it is stable in the sense previously defined. This means that, given any initial state $\xi \in \mathbb{X}$ and any constant input u, $\phi(t, \xi, u) \to \sigma(u)$ as $t \to \infty$. It then follows that

$$\psi(t, \xi, u) = h(\phi(t, \xi, u), u) \to h(\sigma(u), u) = y_0.$$

This stronger property of output convergence to the same value y_0, independent of initial state, is often taken as the definition of perfect adaptation.

*The transformations π are allowed to be nonlinear. We write πu for notational simplicity but use $\pi(u)$ when there may be a possible confusion.

†For the purposes of this note, we may think of inputs $u(t)$ as piecewise-continuous functions and solutions $x(t)$ as continuous and piecewise differentiable. More generally, one could consider Lebesgue-measurable locally essentially bounded inputs u, and the definition of solution is that $x(t)$ is an absolutely continuous function for which the differential equation holds almost everywhere. See ref. 1 for details.

‡A more proper mathematical term is attracting, because this weak definition of stability does not rule homoclinic phenomena.

§To be precise, we should require that $\pi(v(t))$ be a piecewise-continuous function (or more generally, Lebesgue-measurable) whenever $v(t)$ is a piecewise-continuous function. Asking that every $\pi \in \mathcal{P}$ be continuous is enough to guarantee this requirement.

¶The expression πv on the right side of FCD means the input $w(t) = \pi(v(t))$. We could just require the property to hold only for $t > 0$, but the property would be equivalent, taking limits as $t \to 0^+$.

FCD Implies Perfect Adaptation and Weber's Law. ***2.1 Perfect adaptation.*** We say that the action is transitive on inputs if the following property holds: for each pair of distinct $u, v \in \mathbb{U}$, there is some $\pi = \pi_{u,v}$ such that $v = \pi u$.

The most interesting example of transitive action in our context is as follows: $\mathbb{U} = \mathcal{P} = \mathbb{R}^{\Downarrow}_{>\prime}$ (m vectors consisting of positive entries) and $\pi(u) = (p_1 u_1, \ldots, p_m u_m)^T$, which we write as pu, if $\pi = (p_1, \ldots, p_m)$. Clearly, $\pi_{u,v} = (v_1/u_1, \ldots, v_m/u_m)$ achieves $\pi u = v$.

Lemma 1: suppose that the action of is transitive on inputs. Then, FCD implies perfect adaptation.

Proof: pick an arbitrary element $u_0 \in \mathbb{U}$ and define $y_0 := h(\sigma(u_0), u_0)$. Now, pick an arbitrary $w \in \mathbb{U}$. By transitivity, there exists some $\pi \in \mathcal{P}$ such that $\pi u_0 = w$. We now apply Eq. **S4** with $u = u_0$ and also $v = u_0$: $h(\sigma(w), w) = h(\sigma(\pi u_0), \pi u_0) = h(\sigma(u_0), u_0) = y_0$, as required for adaptation.

2.2 Weber's law. We now discuss connections between the FCD property, relative to the symmetries $u \mapsto pu$ ($\mathbb{U} = \mathcal{P} = \mathbb{R}_{>\prime}$)$\mapsto$ and the Weber or Weber-Fechner law of perception.

There are several versions of Weber's law. The textbook (1) provides two relevant definitions (a third one, based on steady-state sensitivity, is irrelevant to systems that perfectly adapt). The main definition used in ref. 1 can be phrased, using our notations, as follows.

Consider the maximum deviation of the output in response to a step from an input value u to an input value v:

$$\Psi(v,u) = \max_{t \geq 0} |\psi(t, \sigma(u), v) - y_0|,$$

where $y_0 = h(\sigma(u), u)$ is the adapted value of the output. Suppose that Ψ is differentiable and introduce the sensitivity of the response

$$S(u) := \left.\frac{\partial \Psi(v,u)}{\partial v}\right|_{v=u}.$$

With these concepts, ref. 1 defines the Weber law as asserting that $S(u)$ is (approximately) inversely proportional to u, which we formalize as there exists a constant k such that

$$S(u) = \frac{k}{u}.$$

FCD implies Weber's law. FCD implies that $\Psi(v, u) = f(v/u)$, for some function f, which we assume is differentiable, and therefore,

$$S(u) = \left.\frac{\partial f(v/u)}{\partial v}\right|_{v=u} = \frac{f'(1)}{u}$$

and Weber's law is indeed satisfied with $k = f'(1)$.

An intuitive way to restate this property is as follows. We expand Ψ to first order around $v = u$, and therefore,[‖]

$$\Psi(v,u) = \Psi(u,u) + S(u)(v-u) + o(v-u).$$

If the system perfectly adapts, then $\Psi(u, u) = 0$, and therefore, Weber's law amounts to the property $\Psi(v,u) \approx \frac{k(u-v)}{u}$. If we write $\Delta y = \Psi(v, u)$ to represent a maximal response change in output and $\Delta u = v - u$, we can write

$$\Delta y \approx k \frac{\Delta u}{u}.$$

More generally, one can prove that the entire response has the same proportionality property. Take any two constant input values u and v. Picking $p = 1/u$ in the FCD condition $\psi(t, \sigma(u), v) = \psi(t, \sigma(\pi u), \pi v)$, we conclude that $\psi(t, \sigma(u), v) = \psi(t, \sigma(1), w) = Q(t, w)$ where $w = v/u$. We expand $Q(t, w) = Q(t, 1) + M(t)(w - 1) + o(w - 1)$ to first order, where $M(t) = \frac{\partial Q}{\partial w}(t, 1)$, and observe that $Q(t, 1) = \psi(t, \sigma(1), 1) = y_0$ for all t, where $y_0 := h(\sigma(1), 1)$ is the adapted value of the output. Note that $y(t) = \psi(t, \sigma(u), v)$ is the output that results after the input jumps from u to v. Writing $\Delta u = v - u$ and $w - 1 = \Delta u/u$, we conclude:

$$\Delta y(t) = y(t) - y_0 = M(t)\frac{\Delta u}{u} + o(\frac{\Delta u}{u}),$$

which is one way to formalize $\Delta y \approx k\frac{\Delta u}{u}$ for all t. The function $M(t)$ can be computed explicitly, as follows:

$$M(t) = c(e^{tA} - I)A^{-1}b + d$$

where

$$A = \frac{\partial f}{\partial x}(\xi, 1), B = \frac{\partial f}{\partial u}(\xi, 1), c = \frac{\partial h}{\partial x}(\xi, 1), d = \frac{\partial h}{\partial u}(\xi, 1)$$

is a matrix and vectors of sizes $n \times n$, $n \times 1$, $q \times n$, and $q \times 1$, respectively, and $\xi = \sigma(1)$. This follows from the fact that that the derivative is computed by solving the variational differential equation $\dot{z} = Az + bu$ with output $cz + du$ (see the proof of theorem 1 in ref. 1). Observe that, when $M(t) = 0$, one can expand to higher order, in which case $\Delta y(t)$ becomes proportional to a power $(\Delta u/u)^k$.

Psychophysical sensitivity. There is a second possible definition of Weber's law, also discussed in ref. 1, based on psychophysical sensitivity and defined as follows. We let r be the smallest possible observable response (in a subjective sense of an individual responding to a stimulus or of a given physical measurement) and let $R(u)$ be the smallest value of the constant input v for which $\Psi(v, u) = r$. Thus, v represents the smallest input that elicits an observable response. Now, the sensitivity $S(u)$ is defined as $1/R(u)$, and Weber's law is once again the property that $S(u) = \frac{k}{u}$ for some k. We prove that FCD implies this version of Weber's law as well.

Indeed, let f be as defined, and therefore, $R(u) = \inf_v \{f(v/u) = r\}$. We assume that f is monotonic before reaching its global maximum or minimum (which is satisfied when there is a unimodal response) and introduce the function g as the inverse of f in its initially monotonic interval. Thus,

$$R(u) = \inf_v \left\{v/u = g(r)\right\} = ug(r) = \frac{u}{k}$$

with $k := 1/g(r)$. Therefore, Weber's law in this psychophysical sensitivity sense holds true, because $S(u) = 1/R(u) = \frac{k}{u}$.

Sufficient Conditions for FCD. We discuss here a technique for verifying the FCD property.

We will call a mapping $\rho : \mathbb{X} \to \mathbb{X}$ an equivariance associated to a given symmetry $\pi \in \mathcal{P}$ if it is differentiable and satisfies the following properties (Eq. **S5**):

$$f(\rho(x), \pi u) = \rho_*(x) f(x, u) \qquad \textbf{[S5]}$$

and (Eq. **S6**)

$$h(\rho(x), \pi u) = h(x, u) \qquad \textbf{[S6]}$$

for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$, where ρ_* denotes the Jacobian matrix of ρ.

Note that we are using a slightly more compact notation than in the paper: we write $\rho(x)$ instead of $\phi(p, x)$ if ρ is the equivariance associated to a symmetry parametrized by p. Thus, $\rho_*(x)$ is the same as $\partial_x \phi(p, x)$.

Lemma 2: the steady-state mapping σ interlaces π and its associated ρ as follows (Eq. **S7**):

‖ The notation $o(x)$ means that $o(x)/x \to 0$ as $x \to 0$; in other words, $o(x) \ll x$ for small x.

$$\rho(\sigma(u)) = \sigma(\pi(u)) \text{ for all } u \in U. \quad \textbf{[S7]}$$

Proof: indeed, we use Eq. **S5** with any u and $x = \sigma(u)$:

$$f(\rho(\sigma(u)), \pi u) = \rho_*(x) f(\sigma(u), u) = 0,$$

because $f(\sigma(u), u) = 0$, by definition of $\sigma(u)$; this means that $\rho(\sigma(u))$ is the steady state $\sigma(\pi u)$ corresponding to the constant input πu, which is what Eq. **S7** asserts.

Lemma 3: suppose that for each $\pi \in \mathcal{P}$, there is an associated equivariance ρ. Then, FCD holds.

Proof: pick any $\pi \in \mathbb{X}$, any constant $u \in \mathbb{U}$, and any input function v: $[0, \infty) \to \mathbb{U}$. Consider the two solutions $x(t) = \varphi(t, \sigma(u), v)$ and $z(t) = \varphi(t, \sigma(\pi u), \pi v)$. We need to show that (Eq. **S8**)

$$\begin{aligned} h(x(t),\ v(t)) &= \psi(t,\ \sigma(u),\ v) = \psi(t,\ \sigma(\pi u), \pi v) \\ &= h(z(t),\ \pi v(t)) \end{aligned} \quad \textbf{[S8]}$$

for every $t \geq 0$.

Take an equivariance ρ associated to π and define $\hat{x}(t) := \rho(x(t))$. Because

$$\hat{x}(0) = \rho(x(0)) = \rho(\sigma(u)) = \sigma(\pi(u))$$

(using Eq. **S7**) and

$$(d/dt)\hat{x}(t) = \rho_*(x(t)) f(x(t), v(t)) = f(\hat{x}(t), \pi v(t))$$

(using the chain rule and then Eq. **S5**), it follows, by definition of φ, that

$$z(t) = \hat{x}(t) = \rho(x(t)).$$

Therefore, Eq. **S8** becomes:

$$h(x(t),\ v(t)) = h(\rho(x(t)),\ \pi v(t)).$$

This property is the second equivariance condition (Eq. **S6**). For controllable and observable systems, the condition in Lemma 3 is necessary as well as sufficient, as follows from uniqueness results in minimal realization theory in control theory [3].

A Subset of Conditions That Is Sufficient for Weber's law. We consider now the very special case of systems with two variables in which the second variable is the output:

$$\dot{x}_1 = f_1(x_1, x_2, u) \quad \textbf{[S9}\textbf{\textit{A}}\textbf{]}$$

$$\dot{x}_2 = f_2(x_1, x_2, u) \quad \textbf{[S9}\textbf{\textit{B}}\textbf{]}$$

$$y = x_2. \quad \textbf{[S9}\textbf{\textit{C}}\textbf{]}$$

We assume that the system adapts ($h(\sigma(u), u) = y_0$ for all u), which translates in this special case to the following property: $\sigma_2(u) = y_0$ for the second component of the steady-state map σ.

We impose the following property for the second component f_2 of f, but no assumptions are made for f_1:

$$f_2(px_1,\ y_0,\ pu) = f_2(x_1,\ y_0,\ u) \quad \textbf{[S10]}$$

for all $u \in \mathbb{R}_{>0}$ (as with the other Weber's Law results, we are restricting attention to the special symmetries $u \mapsto pu$ with U = P = $\mathbb{R}_{>0}$).

We claim that, *for small times t and small* $\Delta u = v - u$, *there holds the approximate Weber's Law*:

$$\Delta y(t) \approx c\frac{\Delta u}{u}$$

where $\Delta y(t) = \psi(t, \sigma(u), v) - y_0$, for an appropriate constant c (which is linearly dependent on t: $c = kt$). Note that $y(t) = \psi(t, \sigma(u), v)$ can be expanded to first order as $y(t) = y_0 + \dot{y}(0)t + o(t)$, and that $\dot{y}(0) = f_2(\sigma_1(u), y_0, v)$. Thus, we now give the precise statement:

Proposition 1: Suppose that Eq. **S10** holds, that f_2 is a differentiable function, and that σ is a continuous function. Then, there is a constant k such that

$$f_2\left(\sigma_1\left(u\right), y_0, v\right) = k\frac{v-u}{u} + o\left(\frac{v-u}{u}\right)$$

for all u, v.

Proof: Eq. **S10** applied with $p = 1/u$, means that $f_2(x_1, y_0, u) = F(x_1/u)$: $= f_2(x_1/u,\ y_0,\ 1)$ for all x_1, u. Thus, our objective is to show that, for some constant k:

$$F\left(\frac{\sigma_1(u)}{v}\right) = k\frac{v-u}{u} + o\left(\frac{v-u}{u}\right) \quad \textbf{[S11]}$$

for all v, u. Since σ is by definition the steady state map, we have that $f_2(\sigma_1(u), y_0, u) = 0$ for all $u \in \mathbb{R}_{>0}$, that is,

$$F\left(\frac{\sigma_1(u)}{u}\right) = 0 \quad \textbf{[S12]}$$

for all y.

So Eq. **S11** can be restated as:

$$\left.\frac{\partial}{\partial v}\right|_{v=u} F\left(\frac{\sigma_1(u)}{v}\right) = \frac{k}{u}$$

for all u. Because of the chain rule, we need to show that:

$$-F'\left(\frac{\sigma_1(u)}{u}\right)\frac{\sigma_1(u)}{u^2} = \frac{k}{u}$$

or, equivalently, that:

$$F'\left(\frac{\sigma_1(u)}{u}\right)\frac{\sigma_1(u)}{u} \text{ is constant.}$$

Let us write $\alpha(u)$: $= \sigma_1(u)/u$ (this is a continuous function defined on the positive reals). We need to show that $F'(\alpha(u))\alpha(u)$ is constant, knowing (from Eq. **S12**) that $F(\alpha(u))$ is constant.

It is a general fact that $F(\alpha(u))$ constant implies $F'(\alpha(u))\alpha(u)$ is constant, for any differentiable function F and any continuous function α. To prove this general fact, let us call J the range $\{\alpha(u), u \in \mathbb{R}_{>0}\}$ of α. Since α is continuous, J is an interval. There are two possibilities: (a) J has only one point or (b) J has interior. Case (a) means that α is a constant function, which obviously implies that $F'(\alpha(u))\alpha(u)$ is constant. If, instead, case (b) holds, then F' must vanish identically on the interval J, which implies that $F'(\alpha(u))\alpha(u) = 0$ for all u, and thus again this expression is constant.

Examples of Generalized FCD Systems. ***Log-linear systems.*** **FCD properties for example shown in Fig. 4C.** The system depicted in Fig. 4*C* satisfies the general FCD conditions (Eqs. **S13** and **S14**)

$$f(\varphi(p,\ x),\ y,\ pu) = \partial_x\varphi(p,\ x) f(x,\ u,\ y) \quad \textbf{[S13]}$$

$$g(\varphi(p,\ x),\ y,\ pu) = g(x,\ u,\ y) \quad \textbf{[S14]}$$

using the transformation (Eq. **S15**)

$$\varphi(p,\ x) = \log(p) + x. \quad \textbf{[S15]}$$

The above conditions are a slight generalization of the basic conditions (Eqs. **5** and **6**) in *Text*. One can prove them directly using the same methodology. In addition, they are a subset of the generalized conditions discussed in *Sufficient conditions for FCD*.

General Analysis of Log-Linear Systems. An interesting class of perfectly adapting systems with the FCD property is that of linear systems with logarithmic memory-free input transformations or more generally, nonlinear functions of such log-linear vector fields:

$$\dot{x} = F(Ax + B\log u)$$
$$y = G(Cx + D\log u)$$

where A, B, C, and D are matrices of sizes $n \times n$, $n \times m$, $q \times n$, and $q \times m$, respectively, and F and G are differentiable maps, possibly nonlinear, that vanish only at 0. For example, F and G might be the identity mappings. We interpret $\log u$ as $\log(u_1, \ldots, \log u_m)^T$ if $u = (u_1, \ldots, u_m)^T$ is a vector.

Lemma 4: assume that $\mathcal{P}$ consists of scalings $\pi u = (p_1 u_1, \ldots, p_m u_m)^T$ and that the system perfectly adapts. Then, the system has the FCD property.

Proof: given a constant input u, the corresponding steady states x satisfy $F(Ax + B\log u) = 0$, which, because of the property that F vanishes only at 0, means that $Ax + B\log u = 0$. Thus, uniqueness of steady-states property is equivalent to the assumption that A is invertible, and

$$\sigma(u) = -A^{-1}B\log u.$$

Because $h(\sigma(u), u) = G(C\sigma(u) + D\log u) = G((D - CA^{-1}B)\log u)$, perfect adaptation, the property that this expression must be independent of u, amounts to the following condition (Eq. **S16**):

$$D - CA^{-1}B = 0. \quad \textbf{[S16]}$$

Given any $\pi = (p_1, \ldots, p_n) \in \mathcal{P}$, we define the equivariance $\rho(x) = x - A^{-1}B\log p$. We must verify (Eq. **S5**):

$$F(A\rho(x) + B\log pu) = F\left(A\left(x - A^{-1}B\log p\right) + B\log pu\right)$$
$$= F(Ax + B\log u)$$
$$= \rho*(x)\, F(Ax + B\log u)$$

(because $\log pu = \log p + \log u$ and $\rho_*(x)$ is the identity matrix) and also need to have (Eq. **S6**):

$$G\left(C\left(x - A^{-1}B\log p\right) + D\log pu\right) = G(Cx + D\log u),$$

which holds because of Eq. **S16**.

Recasting of log-linear systems. Log-linear systems can be recast in the following way, after a change of variables. Let us introduce variables $z_i = e^{x_i}$. Then, $\dot{z} = \mathrm{diag}(z)F(A\log z + B\log u)$, where diag ($z$) is the diagonal matrix whose diagonal entries are $z_1, \ldots, z_n$. The ith row of $A\log z + B\log u$ is:

$$\sum_{j=1}^{n} a_{ij}\log z_j + \sum_{j=1}^{m} b_{ij}\log u_j = \log z^{a_i}u^{b_i},$$

where the notation z^{a_i} means $z_1^{a_{i1}} \ldots, z_n^{a_{in}}$ (analogously for u). A similar rewriting may be done for the output function. Let us define $M(z) := F(\log z)$ and $N(z) := G(\log z)$. We have shown that a log-linear system can also be written as

$$\dot{z} = \mathrm{diag}(z)M\left(z^A u^B\right)$$
$$\dot{y} = N\left(z^C u^D\right)$$

where the variables x_i are positive. The monomials appearing in the above expression represent the entries $z_1^{a_{i1}} \ldots z_n^{a_{in}} u_1^{b_{i1}} \ldots u_m^{b_{im}}$ (analogously for outputs). Furthermore, if N is invertible, one may redefine the output as $N^{-1}(y)$, so that no N is required.

For example, consider this 1D log-linear system:

$$\dot{x} = F(-x + \log u)$$
$$y = G(-x + \log u)$$

(F and G are two scalar nonlinear maps). We let $z = e^x$. Then, with $M = F(\log z)$ and $N(z) = G(\log z)$,

$$\dot{z} = zM(u/z)$$
$$y = N(u/z).$$

Let us redefine the output to be $w = N^{-1}(y)$ (assuming that N is invertible). We arrive to the following system:

$$\dot{z} = zM(w)$$
$$w = u/z.$$

Coordinate projection. Another interesting general subclass is that in which the output $y(t)$ is one coordinate (or, more generally, a subset of coordinates). That is to say, the state space can be written as a Cartesian product $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2$, and using the obvious block notation $x = (x_1, x_2)$ (Eq. **S17** ***A–C***),

$$\dot{x}_2 = f_2(x_1, x_2, u) \quad \textbf{[S17}\textbf{\textit{A}}\textbf{]}$$

$$\dot{x}_1 = f_1(x_1, x_2, u) \quad \textbf{[S17}\textbf{\textit{B}}\textbf{]}$$

$$y = x_2. \quad \textbf{[S17}\textbf{\textit{C}}\textbf{]}$$

Most of the examples in the main part of this paper are of this form. Suppose that for each $\pi\in$, there is some differentiable map $\rho_1 : \mathbb{X}_1 \rightarrow \mathbb{X}_1$ with the following properties (Eq. **S18** ***A*** **and** ***B***):

$$f_1(\rho_1(x_1),\ x_2, \pi u) = (\rho 1)*(x)f_1(x_1, x_2, u) \quad \textbf{[S18}\textbf{\textit{A}}\textbf{]}$$

$$f_2(\rho_1(x_1), x_2, \pi u) = f_2(x_1, x_2, u) \quad \textbf{[S18}\textbf{\textit{B}}\textbf{]}$$

Lemma 5: FCD holds for the system (Eq. **S17**), provided that (Eq. **S18**) holds.

Proof: we observe that the map $\rho(x_1, x_2) = (\rho_1(x_1), x_2)$ is an equivariance. Indeed, its Jacobian has the block form diag $[(r_1)_*(x_1), I]$; therefore, Eq. **S18** is equivalent to Eq. **S5**, and Eq. **S6** is true because $h(x, u) = x_2$ is independent of x_1 and u.

A special case of this setup is when the x_1 subsystem is linear and independent of x_2 (feed-forward connection), $\mathbb{U} = \mathbb{R}_{>0}$(scalar positive inputs), and $\mathcal{P} = \mathbb{R}_{>0}$ acts by scalings $u \mapsto pu$. We write (Eq. **S19**)

$$f_1(x_1,\ u) = Ax_1 + bu \quad \textbf{[S19]}$$

(because u is scalar, $B = b$ is a column vector). Let us suppose that the following property is satisfied (Eq. **S20**):

$$f_2(px_1, x_2, pu) = f_2(x_1, x_2, u) \text{ for all } x_1, x_2, u, p. \quad \textbf{[S20]}$$

Then, FCD holds, because we may use $\rho_1(x_1) = px_1$ for $\pi = p$, in which case $(\rho_1)_*(x_1) = p$ and therefore,

$$f_1(\rho_1(x_1), \pi u) = A(px_1) + b(pu) = p[Ax_1 + bu]$$
$$= (\rho_1)*(x)\, f_1(x_1, u)$$

and

$$f_2(\rho_1(x_1), x_2, \pi u) = f_2(px_1, x_2, pu) = f_2(x_1, x_2, u)$$

Therefore, Eq. **S18** holds.

For these special systems for which Eq. **S19** holds, Eq. **S20** is not merely sufficient, but it is also necessary for FCD to hold (still assuming $\mathbb{U} = \mathcal{P} = \mathbb{R}_{>\prime}$ and an action by scalings $u \mapsto pu$.) We prove this next.

More precisely, we will assume that the system (Eq. **S1**) is controllable from steady states, meaning that for each state $\zeta \in \mathbb{X}$, there is some steady state $\xi = \sigma(u)$ (for some constant input u), some input $v(t)$, and some finite time $T \geq 0$ such that $\zeta = \varphi(T, \xi, v)$. There are control theory tools for checking controllability of linear and nonlinear systems (1). Without loss of generality, one may assume that v is continuous at T and has an arbitrary prespecified value v_0 there. Proof: for any desired value v_0, consider a solution $z(t)$ of Eq. **S1** backward in time, starting from ζ and using the constant input v_0. Let us pick some $\zeta' = z(-t_0)$, $t_0 > 0$. Now, find a $v(t)$ that sends ξ to ζ' in time T'. The concatenation of v and the constant v_0 is an input so that at time $T := T' + t_0$, the state ζ is reached and its value is v_0 at time T.

Lemma 6: suppose that the system (Eq. **S1**) is controllable from steady states and has the form (Eq. **S17**) with Eq. **S19**. Then, the system satisfies FCD for the scaling action $u \mapsto pu$ if and only if Eq. **S20** holds.

Proof: sufficiency was already proved, and therefore, we show necessity. Pick some $\xi = (\xi_1, \xi_2) \in \mathbb{X}$ and $p, v_0 \in \mathbb{R}_{>0}$. We need to show that Eq. **S20** holds (Eq. **S21**) (i.e., that

$$f_2(p\xi_1, \xi_2, pv_0) = f2(\xi_1, \xi_2, v_0)). \qquad \textbf{[S21]}$$

Pick a constant input u and some input $v(t)$ such that $\xi = \varphi(T, \sigma(u), v)$ and $v(T) = v_0$. The assumption is that FCD holds, which means, in particular, that $x_2(t) = \widehat{x}_2(t)$ for all $t \geq 0$, where $x(t) = \varphi(t, \sigma(u), v)$ and $\widehat{x}(t) = \phi(t, \sigma(pu), pv)$. Because also the derivatives of x_2 and $\widehat{x}_2$ must coincide (at the points of differentiability of these functions), it follows, in particular, that

$$\begin{aligned} f_2(\xi_1, \xi_2, v_0) &= f_2(x_1(T), x_2(T), v(T)) = f_2(\widehat{x}_1(T), \widehat{x}_2(T), pv(T)) \\ &= f_2(\widehat{x}_1(T), \widehat{x}_2(T), pv_0) = f_2(\widehat{x}_1(T), \xi_2, pv_0) \end{aligned}$$

(the last equality because $\widehat{x}_2(t) = x_2(t)$, again using FCD). To conclude, observe that $\widehat{x}_1(t) = px_1(t)$ (by linearity of the equation for x_1) and therefore, evaluating at $t = T$, $\widehat{x}_1(T) = p\xi_1$; thus, we have proven that Eq. **S21** is satisfied.

Relationship between the incoherent feed-forward loop and integral feedback. Here, we show the relationship between the incoherent feed-forward loop and integral feedback.

A system is said to be affine in inputs if the vector field has degree 1 on u. Using control-theory notations, one writes the differential equations for the system as follows (assuming, for notational simplicity, that the input u is scalar):

$$\dot{x} = f(x) + ug(x)$$

where f and g are two vector fields. That is, the $f(x, u)$ in the general form $\dot{x} = f(x, u)$ is written as $f(x) + ug(x)$.

A theorem is given in ref. 2 showing that, under appropriate technical assumptions, if a system perfectly adapts to constant signals, then there is a global transformation of coordinates that brings the system into an integral-feedback form. (More generally, the theorem considers adaptation to other, not necessarily constant, types of signals, and an analog of integral feedback, called an internal model, is shown to exist.)

The construction in ref. 2 is a bit involved because of the need to use Lie-theory concepts. Here, we limit ourselves to the following example. We consider a system of 2D, in which the output is the coordinate x_2 and for notational simplicity, write $x = x_1$ and $y = x_2$:

$$\begin{aligned} \dot{x} &= u - x \\ \dot{y} &= u/x - y \end{aligned}$$

evolving on positive variables. This system perfectly adapts, with $y_0 = 1$. We have:

$$f(x,y) = \begin{pmatrix} -x \\ -y \end{pmatrix}, g(x,y) = \begin{pmatrix} 1 \\ 1/x \end{pmatrix}.$$

The relative degree of this system (2) is $r = 1$. One can verify the assumptions of the main theorem in ref. 2 for this system.

The recipe for coordinate changes in ref. 2 (see also the Feedback Linearization Theorem, Theorem 15 in ref. 3) is to use $z_1 = y$ and $z_2 = \varphi(x, y)$ with the following conditions on the differentiable map φ:

1. The map $(x, y) \mapsto (y, \varphi(x, y))$ has a differentiable inverse (technically, is a diffeomorphism).
2. The Lie-derivative $L_g\varphi$ vanishes everywhere, which means $\nabla\varphi.g = 0$ (∇g is the gradient of φ).

The condition $\nabla\varphi.g = 0$ says, more explicitly, for this example:

$$\phi_x(x,y) + \frac{1}{x}\phi_y(x,y) = 0$$

where φ_x, φ_y are partial derivatives. This linear first-order partial differential equation on φ may be solved by the method of characteristics, but a solution can be seen by inspection:

$$\phi(x,y) = y - \log x \; .$$

Observe that $(x, y) \mapsto (y, y - \log x) = (z_1, z_2)$ is clearly invertible, with inverse $y = z_1$ and $x = e^{z_1 - z_2}$. In the new coordinates z_1, z_2, we have:

$$\begin{aligned} \dot{z}_1 &= ue^{z_2 - z_1} - z_1 \\ \dot{z}_2 &= 1 - z_1. \end{aligned}$$

Up to a change of coordinates $z_1 \mapsto 1 - z_1$ to bring the system into the form in ref. 2 (which normalized the adaptation value to 0; it is 1 in this example), we have that the variable z_2 implements the integral feedback ensured by theorem 1 in ref. 2.

The form in (z_1, z_2) coordinates is known in control theory as the feedback linearization normal form (3) and is a special case of a normal form for affine nonlinear systems.

Stability Result. We wish to show the global asymptotic stability (GAS) of the unique steady state $(x,y) = (\frac{\alpha u_0}{\beta y_0}, y_0)$ of the nonlinear integral feedback system (Eq. **S22 *A* and *B***):

$$\dot{x} = \gamma x(y - y_0) \qquad \textbf{[S22}\textbf{\textit{A}}\textbf{]}$$

$$\dot{y} = \alpha\frac{u_0}{x} - \beta y \qquad \textbf{[S22}\textbf{\textit{B}}\textbf{]}$$

where α, β, γ, u_0, and y_0 are positive constants and the integrator variable $x(t)$ is positive. We prove this as a consequence of a more general result.

Lemma 7: consider a 2D system of the following general form (Eq. **S23 *A* and *B***):

$$\dot{x} = g(y) \qquad \textbf{[S23}\textbf{\textit{A}}\textbf{]}$$

$$\dot{y} = -f(x) - k(y) \qquad \textbf{[S23}\textbf{\textit{B}}\textbf{]}$$

where f and g are functions with positive derivatives, $(y - y_0)k(y) > 0$ whenever $y \neq y_0$. Let (x_0, y_0) be so that $f(x_0) = g(y_0) = k(y_0) = 0$, which means that (x_0, y_0) is the unique steady state of the system. Then, (x_0, y_0) is a globally asymptotically stable state.

We provide a proof below but first remark how the stability of Eq. **S22** is a consequence of this Lemma.

Corollary: consider a 2D system of the following general form (Eq. **S24 *A* and *B***):

$$\dot{x} = xg(y) \quad \textbf{[S24}\textbf{\textit{A}}\textbf{]}$$

$$\dot{y} = -f(x) - k(y) \quad \textbf{[S24}\textbf{\textit{B}}\textbf{]}$$

where f and g are functions with positive derivatives, $(y - y_0)k(y) > 0$ whenever $y \neq y_0$, the variable $x(t)$ is positive, and (x_0, y_0) is so that $f(x_0) = g(y_0) = k(y_0) = 0$, which means that (x_0, y_0) is the unique steady state of the system. Then, (x_0, y_0) is a globally asymptotically stable state.

This corollary is proved as follows. We let $z = \ln x$ and express the system in the variables (z, y). We have that (Eq. **S25 *A* and *B***):

$$\dot{z} = g(y) \quad \textbf{[S25}\textbf{\textit{A}}\textbf{]}$$

$$\dot{y} = -\tilde{f}(z) - k(y) \quad \textbf{[S25}\textbf{\textit{B}}\textbf{]}$$

where $\tilde{f}(z) := f(e^z)$ again has a positive derivative. Now the unique steady state is (z_0, y_0), where $z_0 := \ln x_0$. By the Lemma, this state is globally asymptotically stable, which implies that the system in original coordinates (Eq. **S24**) is also stable.

The system (Eq. **S22**) is the particular case of Eq. **S24** with $f(x) = \beta y_0 - \frac{\alpha u_0}{x}$, $g(y) = \gamma(y - y_0)$, and $k(y) = \beta(y - y_0)$.

We now prove Lemma 7. The proof is based on the LaSalle Invariance Principle (3). We must produce a function $V(x, y)$ of two variables with the following properties:

1. $V(x_0, y_0) = 0$.
2. $V(x, y) > 0$ for all $(x, y) \neq (x_0, y_0)$.
3. $V(x, y) \to \infty$ as $\|(x, y)\| \to \infty$ (properness or radial unboundedness).
4. $\dot{V}(x,y) := \frac{\partial V}{\partial x}(x,y)g(y) + \frac{\partial V}{\partial y}(x,y)[-f(x) - k(y)]$ is so that (*i*) $\dot{V}(x,y) \leq 0$ for all (x, y) and (*ii*) if a solution satisfies that $\dot{V}(x(t), y(t)) \equiv 0$, then $(x(t), y(t)) \equiv (x_0, y_0)$.

We define:**

$$V(x,y) := \int_{x_0}^{x} f(r)dr + \int_{y_0}^{y} g(r)dr.$$

Observe that properties 1 and 2 (positive definiteness) are satisfied by definition. Regarding property 3, we note that $\frac{\partial^2 V}{\partial^2 x} = f'(x) > 0$, $\frac{\partial^2 V}{\partial^2 y} = g'(y) > 0$ and mixed second derivatives are 0, and therefore, the Hessian matrix of V is positive definite everywhere. This implies that V is strictly convex, and principle 3 follows. Finally, we prove principle 4. Observe that

$$\dot{V}(x,y) = f(x)g(y) + g(y)[-f(x) - k(y)] = -g(y)k(y)$$

from which it follows that i holds, and moreover, $\dot{V}(x,y) = 0$ implies that $y = y_0$. Suppose that a solution satisfies that $\dot{V}(x(t),y(t)) \equiv 0$. Then, $y(t) \equiv y_0$, and therefore, $\dot{y}(t) \equiv 0$. Substituted into the second equation of Eq. **S23**, we have that $0 = -f(x(t)) - 0$, which implies that $x(t) \equiv x_0$; therefore, ii is true.

**This construction is based on the following idea: when k(y) is omitted, the vector field is Hamiltonian, with Hamiltonian function V; this provides an energy-conservation constraint, but k(y) then adds damping to the system.

1. Keener J, Sneyd J (2009) *Mathematical Physiology* (Springer, New York), 2nd Ed.
2. Sontag ED (2003) Adaptation and regulation with signal detection implies internal model. *Syst Control Lett* 50:119–126.
3. Sontag ED (1998) *Mathematical Control Theory. Deterministic Finite-Dimensional Systems* (Springer, New York), 2nd Ed.