Modular Stability Tools for
Distributed Computation and Control

Jean-Jacques E. Slotine

Nonlinear Systems Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, 02139, USA
jjs@mit.edu

Abstract

Much recent functional modelling of the central nervous system, beyond traditional "neural net" approaches, focuses on its distributed computational architecture. This paper discusses extensions of our recent work aimed at understanding this architecture from an overall nonlinear stability and convergence point of view, and at constructing artificial devices exploiting similar modularity. Applications to synchronisation and to schooling are also described. The development makes extensive use of nonlinear contraction theory.

Keywords: Contraction theory, modularity, synchronisation, networks.

1   Introduction

Any biological object, and specifically the brain, is the result of evolution. Evolution proceeds by accumulation and combination of stable intermediate states: Darwin's survival of the fittest really means survival of the stable (Simon, 1962; Dawkins, 1976). In the process, systems of increasing complexity are created (Kirschner and Gerhart, 1998; Ridley, 2000). Simple examples abound: for instance, motion control architecture in vertebrates is believed to involve combinations of simple motor primitives (Bernstein, 1967; Bizzi, et al., 1995); human emotional response involves both fast archaic loops bypassing the cortex, and slower cortical loops (Ledoux, 1996; Damasio, 2001). However, in themselves, accumulations and combinations of stable elements have no reason to be stable. Hence our hypothesis in (Slotine and Lohmiller,

2001) that evolution will favor a particular form of state stability, which automatically guarantees stability in combinations, such as parallel, series, feedback, or multiresolution. Such a form of nonlinear stability, which we refer to as "contraction" (Lohmiller and Slotine, 1998), can be characterized mathematically.

Conceptually, accumulations of stable dynamics have also been a recurrent theme in cybernetics and AI history (Walter, 1950, 1951; Chomsky, 1957; Wiener, 1961; Simon, 1962, 1981; Ashby, 1966; Braitenberg, 1984; Minsky, 1986) under various guises (Brooks, 1986, 1999). They also form the basis of several recent theories of brain function (Tononi, et al., 1998; Dehaene, et al., 1998; Crick and Koch, 1998; Edelman and Tononi, 2000; Grossberg, 2000) and of biological motor control (Bernstein, 1967; Bizzi, et al., 1995; Mussa-Ivaldi, 1997; Wolpert and Kawato, 1998; Tresch, et al., 1999; Jordan and Wolpert, 1999; Thoroughman and Shadmer, 2000; Giszter, et al., 2000). In control theory, passivity (Popov, 1973) was originally motivated by similar concerns.

This suggests that contraction theory may both guide functional modelling of the central nervous system beyond basic "neural net" approaches, and provide a systematic method to build arbitrarily complex robots out of simpler elements. Furthermore, it may shed light on the problem of perceptual unity (binding problem) by providing simple models and conditions for the overall convergence of a large number of specialized processing elements connected through networks of feedback loops (Slotine and Lohmiller, 2001). This paper discusses further examples along these lines.

Section 2 defines nonlinear contraction, reviews its basic properties (Lohmiller and Slotine, 1998), and presents some elementary applications. Section 3 further examines combinations properties of contracting systems and some of their implications. Section 4 discusses applications to oscillator synchronisation, another type of computation occuring in the central nervous system, and to schooling and flocking. Section 5 offers brief concluding remarks.

2 Modularity and Stability

Basically, a nonlinear time-varying dynamic system will be called contracting if initial conditions or temporary disturbances are forgotten exponentially fast, i.e., if trajectories of the perturbed system return to their nominal behavior with an exponential convergence rate. It turns out that relatively simple conditions can be given for this stability-like property to be verified, and furthermore that this property is preserved through basic system combinations, such as parallel combinations, feedback combinations, and series or hierarchies.

Incidentally, such a definition fits rather naturally with known data on biological motion perturbation, e.g. perturbation of arm movement (Soechting and Lacquaniti, 1988; Won and Hogan, 1995). Furthermore, it is intrinsic, in the sense that the system's "nominal" behavior needs not be known. Finally, such a form of stability, at least in a local sense, is also a basic prerequisite for any learning, since it guarantees the consistency of the system's behavior in the presence of small disturbances or variations in initial conditions.

2.1 Contraction Analysis

In this section, we summarize the basic results of (Lohmiller and Slotine, 1998), to which the reader is referred for more details. We consider general time-varying deterministic systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{1}$$

where $\mathbf{f}$ is an $n \times 1$ nonlinear vector function and $\mathbf{x}$ is the $n \times 1$ state vector. The above equation may also represent the closed-loop dynamics of a controlled system with state feedback $\mathbf{u}(\mathbf{x}, t)$. All quantities are assumed to be real and smooth, by which it is meant that any required derivative or partial derivative exists and is continuous. The basic result of (Lohmiller and Slotine, 1998) can then be stated as

Theorem 1 *Consider the system (1). If theres exist a uniformly positive definite metric*

$$\mathbf{M}(\mathbf{x}, t) = \mathbf{\Theta}(\mathbf{x}, t)^T \, \mathbf{\Theta}(\mathbf{x}, t)$$

such that the associated generalized Jacobian

$$\mathbf{F} = \left(\dot{\mathbf{\Theta}} + \mathbf{\Theta} \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \mathbf{\Theta}^{-1}$$

is uniformly negative definite, then all system trajectories then converge exponentially to a single trajectory, with convergence rate $|\lambda_{max}|$, where λ_{max} is the largest eigenvalue of the symmetric part of $\mathbf{F}$. The system is said to be contracting.

It can be shown conversely that the existence of a uniformly positive definite metric with respect to which the system is contracting is also a necessary condition for global exponential convergence of trajectories. In the linear time-invariant case, a system is globally contracting if and only if it is strictly stable, with $\mathbf{F}$ simply being a normal Jordan form of the system and $\mathbf{\Theta}$ the coordinate transformation to that form.

In this paper, for simplicity we shall concentrate on the global convergence result above. In the case that $\mathbf{F}$ is uniformly negative definite only in a finite region, then the result can be shown to hold for all trajectories starting in the largest ball (with respect to the metric $\mathbf{M}$) contained in that region. Note that since $\mathbf{\Theta}$ is invertible, requiring that $\mathbf{F}$ be uniformly negative definite is equivalent to requiring that $\dot{\mathbf{M}} + 2\,\mathbf{M}\,\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ be uniformly negative definite. Furthermore, since

$$\mathbf{\Theta}^{-1}\,\mathbf{F}_s\,\mathbf{\Theta} = \frac{1}{2}\,\mathbf{M}^{-1}\,(\dot{\mathbf{M}} + \mathbf{M}\,\frac{\partial \mathbf{f}}{\partial \mathbf{x}} + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^T\,\mathbf{M})$$

where is $\mathbf{F}_s$ the symmetric part of $\mathbf{F}$, all transformations $\mathbf{\Theta}$ corresponding to the same $\mathbf{M}$ lead to the same eigenvalues for $\mathbf{F}_s$, and therefore to the same contraction rate $|\lambda_{max}|$. In particular one may always define or redefine the transformation $\mathbf{\Theta}$ to be lower triangular, using a Cholesky decomposition of $\mathbf{M}$.

2.2 Examples

Example 2.1: Time-invariant contracting systems converge to a unique equilibrium point. Indeed, consider the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, contracting with respect to $\boldsymbol{\Theta}$. One can easily verify that

$$\frac{d}{dt}(\boldsymbol{\Theta}\mathbf{f}) = \mathbf{F}(\boldsymbol{\Theta}\mathbf{f})$$

which implies that $\boldsymbol{\Theta}\mathbf{f}$ and thus $\mathbf{f} = \dot{\mathbf{x}}$ converge exponentially to zero, and therefore that $\mathbf{x}$ converges exponentially to a constant vector.

Consider now an autonomous contracting system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, with constant $\boldsymbol{\Theta}$, and a constant vector $\mathbf{c}$ of the same dimension. The above implies that the dynamics

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) - \mathbf{c}$$

converges exponentially to the unique solution of the algebraic system $\mathbf{f}(\mathbf{x}) = \mathbf{c}$. $\square$

Similarly, contracting systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}(t))$$

where the input $\mathbf{u}(t)$ is periodic in time, can be shown to converge towards a periodic state of the same period as the input.

Example 2.2: Consider a time-varying cost function $V(\mathbf{x} - \mathbf{r}(t))$, strictly convex in $\mathbf{x}$, and the dynamic system

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{r}(t)$$

$$\dot{\bar{\mathbf{x}}} = -\operatorname{grad} \mathbf{V}(\mathbf{x} - \mathbf{r}(\mathbf{t}))$$

One has

$$\frac{d}{dt}(\mathbf{x} - \mathbf{r}(t)) = -\operatorname{grad} V(\mathbf{x} - \mathbf{r}(t))$$

This dynamics is contracting, since its Jacobian $-\frac{\partial^2 V}{\partial \mathbf{x}^2}$ is uniformly negative definite Thus, $\mathbf{x}$ tracks the minimum of V after an exponential transient. $\square$

Example 2.3: In biological modelling, one frequently encounters (Tresch, 2002) so-called activation functions $x(t)$, driven by an electrical signal $u(t)$ according to a dynamics of the form

$$\tau\dot{x} + [\beta + (1 - \beta)u(t)]\,x = u(t) \qquad 0 \leq u(t) \leq 1$$

with β a constant, $0 < \beta \leq 1$, and constant $\tau > 0$. Intuitively, such systems, found e.g. in muscle models (Zajac, 1989) and in dynamic synapse models (Fuhrmann, et al., 2001), respond faster when $u(t)$ smoothly transitions from 0 to 1, and slower on the way back. This dynamics is contracting since its scalar Jacobian is upper bounded by $-\beta/\tau$. $\square$

Example 2.4: Contraction can also be defined in open connected subsets of the state space. For instance, the system

$$\dot{x} = x(u^2 - x^2)$$

with u a constant, is contracting for $x > \epsilon$ and $x < -\epsilon$ for arbitrary $0 < \epsilon < |u|$, with metric $1/x^2$, and thus converges accordingly to $\pm u$.

Partial contraction may also be defined, corresponding a metric of partial rank. Consider for instance the classical limit cycle dynamics (Luenberger, 1979)

$$\dot{x} = y - x\left(x^2 + y^2 - 1\right)$$
$$\dot{y} = -x - y\left(x^2 + y^2 - 1\right)$$

Letting $r = \sqrt{x^2 + y^2}$, one has

$$\dot{r} = r\left(r^2 - 1\right)$$

implying exponential convergence of r to 1 in any region $r > \epsilon$, with arbitrary $0 < \epsilon < 1$. The corresponding metric $\Theta^T\Theta$ for the original system is given by $\Theta = (\,x \quad y\,)/(x^2 + y^2)$. $\qquad\square$

Example 2.5: Chaotic synchronisation

Starting with (Pecora and Carroll, 1990), much attention has been devoted to synchronisation in "chaotic" systems. Consider the Lorenz system

$$\dot{x} = \sigma\,(y - x)$$
$$\dot{y} = \rho\,x - y - x\,z$$
$$\dot{z} = -\beta\,z + x\,y$$

with strictly positive constants σ, ρ, β, and, as in (Pecora and Carroll, 1990; Nijmejier, 2001), a reduced-order identity observer for this system based on an available measurement of the variable x,

$$\dot{\hat{y}} = \rho\,x - \hat{y} - x\,\hat{z}$$
$$\dot{\hat{z}} = -\beta\,\hat{z} + x\,\hat{y}$$

The symmetric part of the observer's Jacobian is $-\operatorname{diag}(1, \beta)$, and thus the observer is contracting with an identity metric. Since by construction $(\hat{y}, \hat{z}) = (y, z)$ is a particular solution, the estimated state converges exponentially to the actual state. $\qquad\square$

3 Combinations of contracting systems

As a form of stability, one of the main features of nonlinear contraction is that it is automatically preserved through a variety of system combinations (Lohmiller and Slotine, 1998; 2000a), a state-space property reminiscent of input-output passivity (Popov, 1973). In this section we illustrate these properties with simple applications to modular design. An interesting recent discussion of the application of passivity tools to recursive refinement of the control of movement can be found in (Arimoto and Naniwa, 2002).

Formally, some of the combination properties of contracting systems are most easily stated used the notion of a virtual displacement, from classical physics. A virtual displacement $\delta\mathbf{x}$ is an infinitesimal displacement *at fixed time* $-$ formally, if we view the position of the system at time t as a smooth function of the initial condition $\mathbf{x}_o$ and of time, $\mathbf{x} = \mathbf{x}(\mathbf{x}_o, t)$, then $\delta\mathbf{x} = \frac{\partial\mathbf{x}}{\partial\mathbf{x}_o} d\mathbf{x}_o$. Coordinate transformations of the form $\delta\mathbf{z} = \Theta(\mathbf{x}, t)\delta\mathbf{x}$ can be performed on virtual displacements. These are much more general than simple coordinate changes, since an explicit $\mathbf{z}$ need not need exist, i.e., the transformation need not be integrable. $\Theta(\mathbf{x}, t)^T \Theta(\mathbf{x}, t)$ defines the metric in Theorem 1.

By contrast with linear theory, contraction analysis does not use differentiation as local approximation, but rather as a formal tool to yield global results on the nonlinear system.

3.1 Parallel combination

Consider two systems of the same dimension, contracting *in the same metric*,

$$\dot{\mathbf{x}} = \mathbf{f}_i(\mathbf{x}, t) \qquad i = 1, 2$$

Assume further that the metric depends only the state $\mathbf{x}$ and not explicitly on time. Then, any uniformly positive superposition (where $\exists\, \alpha > 0,\ \forall t \geq 0,\ \exists\, i,\ \alpha_i(t) \geq \alpha$)

$$\dot{\mathbf{x}} = \alpha_1(t)\, \mathbf{f}_1(\mathbf{x}, t) + \alpha_2(t)\, \mathbf{f}_2(\mathbf{x}, t)$$

is contracting in the same metric. By recursion, this property can be extended to any number of systems.

Example 3.1: Control primitives

Recently, there has been considerable interest in analyzing feedback controllers for biological motor control systems as combinations of simpler elements, or primitives (Bizzi, et al., 1995; Mussa-Ivaldi, 1997; d'Avella and Bizzi, 1998; Tresch, et al., 1999; Giszter and Kargo, 2000; Kargo and Giszter, 2000). Besides being biological plausible, such a structure is intuitively appealing, as it may yield considerable dimensionality reduction in learning and planning. Similar goals motivate e.g. (Atkeson et al., 1997; Schaal, 1999; Fod et al., 2000). The structure is also reminiscent of potential fields in robotics (Khatib, 1986), although typically the control primitives or their modulating coefficients are time-varying and thus do not lend themselves easily to an energy-based analysis.

Consider more generally the system
$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{B}(\mathbf{x}, t)\, \mathbf{u}$$

and assume that there exist control primitives $\mathbf{u} = \mathbf{p}_i(\mathbf{x}, t)$ which, for any i, make the closed-loop system contracting in some common metric. Multiplying each equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{B}(\mathbf{x}, t)\, \mathbf{p}_i(\mathbf{x}, t)$$

by a positive coefficient $\alpha_i(t)$, and summing, shows that any convex combination of the control primitives $\mathbf{p}_i(\mathbf{x}, t)$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{B}(\mathbf{x}, t) \sum_i \alpha_i(t)\, \mathbf{p}_i(\mathbf{x}, t) \qquad \forall i,\ \alpha_i(t) \geq 0 \qquad \sum_i \alpha_i(t) = 1$$

also leads to a contracting dynamics in the same metric.

The time-varying convex combination may correspond, for instance, to smoothly turning on and off primitives, as e.g. in (D'Avella, et al. 2003). $\square$

3.2 Hierarchical Combination

Consider two contracting systems, of possibly different dimensions and metrics, and connect them in series, leading to a smooth virtual dynamics of the form

$$\frac{d}{dt}\begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{0} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{pmatrix} \begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix}$$

Then the overall system is contracting, as long as $\mathbf{F}_{21}$ is bounded, as can be seen for instance by using

$$\Theta = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{O} \\ \mathbf{O} & \epsilon\,\mathbf{I}_{n_2} \end{pmatrix}$$

for $\epsilon > 0$ sufficiently small. By recursion, the result extends to hierarchies or cascades of contracting systems of arbitrary depths.

Example 3.2: Composite variables

Composite signals, i.e., signals representing mixtures of more obvious physical quantities such as position or velocity, are pervasive in the nervous system (Berthoz, 1999), and also exploited in methodologies such as sliding control (see e.g., (Slotine and Li, 1991)). Using such combinations can often be interpreted as enforcing a hierarchy of contracting systems, so as to reduce the complexity of control or estimation problems.

In a second-order mechanical system, for instance, choosing a sliding variable $s = \dot{\tilde{x}} + \lambda\tilde{x}$, where $\tilde{x}(t)$ is the tracking error and λ is a strictly positive constant, simply corresponds to creating a hierarchy of contracting systems

$$\dot{s} = \phi(s, t) \qquad \text{nominally contracting by choice of control law}$$
$$\dot{\tilde{x}} + \lambda\tilde{x} = s \qquad \text{contracting by definition of } s$$

where "nominally" refers to the uncertainty-free case. The actual choice of the composite variable can be shaped according to the desired qualitative behavior of the contracting system. For instance, a system using instead

$$\dot{\tilde{x}} + (\lambda_1 + \lambda_2|\tilde{x}|)\tilde{x} = s$$

(with constants $\lambda_i > 0$) reacts faster to larger errors, since the corresponding scalar Jacobian is $-(\lambda_1 + 2\lambda_2|\tilde{x}|)$.

Similarly, selecting transfer functions $G(p)$ and $L(p)$, where p is the Laplace variable, and defining

$$\dot{x} - G(p)\,\dot{x}_d + L(p)\,(x - x_d) = s$$

with $x_d(t)$ an external input, leads to the hierarchical layer

$$x = \frac{1}{p + L(p)}\,s + \frac{p\,G(p) + L(p)}{p + L(p)}\,x_d$$

This definition of s corresponds to the classical problem of selecting, for a plant $1/p$, a feedback compensator $L(p)$ and a feedforward compensator $G(p)$ to obtain appropriate frequency shaping of x and "disturbance attenuation" of s (Young and Ozguner, 1993; Zuo and Slotine, 2003). Note that in this case the target dynamics is of order $n_L + 1$, where n_L is the order of $L(p)$. □

Example 3.3: In Example 2.5, consider constructing a full-state identity observer by augmenting the dynamics with

$$\dot{\hat{x}} \;=\; \sigma\,(\hat{y} - \hat{x})$$

Then, the observer corresponds to a hierarchy of contracting systems, and thus the full estimated state converges exponentially to the actual state. □

3.3   Static nonlinear maps

Consider a smooth map

$$\mathbf{y} = \mathbf{f}(\mathbf{x}_1,\; \ldots\;,\mathbf{x}_n, t)$$

where the $\mathbf{x}_i$ are states of contracting systems (of possibly different metrics). Then $\delta\mathbf{y} \to \mathbf{0}$ exponentially as long as all $\frac{\partial \mathbf{y}}{\partial \mathbf{x}_i}$ are bounded.

Similarly, given a contracting system of state $\mathbf{x}$, boundedness of the Jacobian and its time-derivatives up to order j implies exponential convergence to zero of any time-derivative of $\delta\mathbf{x}$ up to order $j + 1$. Thus the nonlinear maps above may depend accordingly on time-derivatives of the $\mathbf{x}_i$.

Example 3.4: Full-order target dynamics

Extending Example 3.2, for an uncertain nonlinear system $x^{(n)} \;=\; f(x, \dot{x}, ..., x^{(n-1)}, u, t)$ to be controlled, the desired dynamics $s = 0$ can be selected to correspond to any contracting system $x^{(n)} \;=\; g(x, \dot{x}, ..., x^{(n-1)}, t)$ *of the same order*. This may be a natural choice, e.g., when the system is controlled intermitently, when it must exhibit specific disturbance responses (Zuo and Slotine, 2003), when robot impedances must be accurately controlled based on sensed interaction forces, or when control is used to simplify a subsystem dynamics as seen from other subsystems. It can be achieved by defining, based on state measurements alone

$$s \;=\; x^{(n-1)} \;-\; \bar{s}$$
$$\dot{\bar{s}} \;=\; -\,k\,(\bar{s} \,-\, x^{(n-1)}) \;+\; g(x, \dot{x}, ..., x^{(n-1)}, t)$$

with constant $k > 0$. Indeed one then has

$$x^{(n)} \;-\; g(x, \dot{x}, ..., x^{(n-1)}, t) \;=\; \dot{s} \;+\; k\,s$$

Note that this also suggests that the nominal (uncertainty free) dynamics of s, one layer above, should be best selected as $\dot{s} \;=\; -\,k\,s$, although other choices are possible as long as nominally $\dot{s} \;+\; k\,s \;\to\; 0$.   □

Example 3.5: If $\mathbf{x}$ is the state of a contracting system, then so is $\mathbf{x}^T\mathbf{x}$. Such positive scalars may be used hierarchically to define the coefficients $\alpha_i(t)$ in control primitives, for instance. □

Example 3.6: Multi-stream processing

Combining the above properties, contraction can be preserved if one first "splits" a signal into parallel streams through linear or nonlinear maps, processes each stream through contracting dynamics (e.g., sequences of non-linear filters and maps), and then "recombine" the streams.

By analogy, recall that visual information flow is processed by 10^8 neurons on the retina, compressed (or partially discarded) as input to 10^6 neurons in the optic nerve, analyzed by 10^{10} neurons within several specialized areas in the cortex, before being shared between areas and with the whole brain (Keat, et al., 2001). Similar compression and distributed filtering occurs for auditory processing (Watts, 2002). □

3.4   Feedback Combination

Consider two contracting systems, of possibly different dimensions and metrics, and connect them in feedback, in such a way that the overall virtual dynamics is of the form

$$\frac{d}{dt}\begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{G} \\ -\mathbf{G}^T & \mathbf{F}_2 \end{pmatrix}\begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix}$$

where the matrix $\mathbf{G}(\mathbf{x}_1, \mathbf{x}_2, t)$ is arbitrary (other than being a matrix of partial derivatives). Then the overall system is contracting. The result can of course be extended to any number of systems: with obvious notations, overall contraction is achieved if $\mathbf{G}_{ij} + \mathbf{G}_{ji}^T = \mathbf{0}, \ \forall i, j, \ i \neq j$.

Example 3.7: A system of the form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, t) - \sum_i \lambda_i \nabla\mathbf{g}_i \\ \dot{\lambda} &= \mathbf{g}(\mathbf{x}, t) + \mathbf{h}(\lambda, t) \end{aligned}$$

with $\mathbf{f} - \sum_i \lambda_i \nabla\mathbf{g}_i$ contracting in $\mathbf{x}$, and $\mathbf{h}$ contracting in λ, is contracting. Similar dynamics occur in constrained optimization. □

Example 3.8: Feedback as above is a very efficient way to share information and distribute computation among specialized contracting subdynamics, since the overall convergence rate is simply the slowest of the individual convergence rates.

Similar feedback configurations may be involved when binding information from different specialized areas in the visual cortex (processing edges, others shape, motion, depth, color, and so on (Kandel, et al., 2000)) or in the auditory system (time-frequency analysis, spectral edges, interaural time difference, and so on (Watts, 2002)), as well as in merging information from different senses. □

More generally, consider two contracting systems and an arbitrary feedback connection between them (Wang and Slotine, 2001). The overall virtual dynamics can be written

$$\frac{d}{dt}\begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix} = \mathbf{F}\begin{pmatrix} \delta\mathbf{z}_1 \\ \delta\mathbf{z}_2 \end{pmatrix}$$

Compute the symmetric part of $\mathbf{F}$, in the form

$$\frac{1}{2} \left(\mathbf{F} + \mathbf{F}^T \right) = \left(\begin{array}{cc} \mathbf{F}_{1s} & \mathbf{G}_s \\ \mathbf{G}_s^T & \mathbf{F}_{2s} \end{array} \right)$$

where by hypothesis the matrices $\mathbf{F}_{is}$ are uniformly negative definite. Then $\mathbf{F}$ is uniformly negative definite if and only if $\mathbf{F}_{2s} < \mathbf{G}_s^T \mathbf{F}_{1s}^{-1} \mathbf{G}_s$, a standard result from matrix algebra (Horn and Johnson, 1985). Thus, a sufficient condition for contraction of the overall system is that

$$\sigma^2(\mathbf{G}_s) < \lambda(\mathbf{F}_1) \lambda(\mathbf{F}_2) \qquad \text{uniformly } \forall \mathbf{x}, \; \forall t \geq 0$$

where $\lambda(\mathbf{F}_i)$ is the contraction rate of $\mathbf{F}_i$ and $\sigma(\mathbf{G}_s)$ is the largest singular value of $\mathbf{G}_s$.

Furthermore, applying this result to the matrix $(\mathbf{F} + \lambda \mathbf{I})$ for an arbitrary $\lambda \geq 0$ yields an explicit lower bound on the overall contraction rate

$$|\lambda_{\text{max}}| \geq \frac{\lambda(\mathbf{F}_1) + \lambda(\mathbf{F}_2)}{2} - \sqrt{\left(\frac{\lambda(\mathbf{F}_1) - \lambda(\mathbf{F}_2)}{2} \right)^2 + \sigma^2(\mathbf{G}_s)}$$

The results can be applied recursively to combinations of arbitrary size.

3.5 Translation and scaling

In space

It is straightforward to show that if $\mathbf{f}(\mathbf{x}, t)$ defines a contracting dynamics with respect to a *constant* metric, so does any scaled and translated version $\mathbf{f}(a(t)\mathbf{x} - \mathbf{b}(t), t)$, where $a(t)$ and $\mathbf{b}(t)$ are arbitrary differentiable functions and $a(t)$ is uniformly positive definite. This property, combined with the parallel combination property above, can allow contracting dynamics to be used as wavelet-like basis functions in problems of dynamic approximation, estimation, and adaptive control,

$$\dot{\mathbf{x}} = \sum_i \alpha_i(t) \, \mathbf{f}(a_i(t)\mathbf{x} - \mathbf{b}_i(t), t)$$

and thus can provide practical tools for progressive refinement and learning.

In time

One immediately sees that if $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ is contracting, so is

$$\epsilon(t) \, \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \beta(t))$$

where $\epsilon(t) > 0$ and $\beta(t) \geq 0$ are arbitrary functions of time. Note that $\beta(t)$ may be discontinuous.

Of course, contraction is also preserved by *any* combination of all the above. External inputs can be provided through any subsytem dynamics. Overall contraction also implies that the system will recover exponentially fast from temporary disturbances in any subsystem.

Note that, essentially, the convergence rate of parallel combinations is the weighted sum of the individual convergence rates, the convergence rate of skew-symmetric feedback combinations is the slowest of the individual convergence rates, and the convergence time-constant of hierarchical combinations is the sum of the individual layer time-constants.

Also note that contraction is preserved even though the network's topology and/or connections strengths can be time-varying, possibly in a discontinuous fashion. In particular, contraction can be preserved through system growth and pruning, and similarly different aggregates may be recruited according to the task. By analogy, subdynamics in the central nervous system are constantly re-used and recombined as part of different overall circuits and behaviors, plausibly as an efficient result of evolutionary "tinkering". This reutilisation or nesting principle (Tononi, et al., 1998; Berthoz, 1999; Damasio, 2003) may also account in part for the pervasiveness of inhibition mechanisms in the central nervous system.

Finally, although for simplicity we concentrate on systems described by ordinary differential equations, the discussion extends to large classes of partial differential equations (Lohmiller and Slotine, 1999, 2001). Ultimately, networks of massively parallel dynamic "neurons" may be best represented by p.d.e.'s.

3.6 Adaptive combinations

It is straightforward to incorporate adaptive techniques in contraction-based designs if part of the system's uncertainty consists of a vector $\mathbf{a}$ of unknown but constant (or slowly-varying) parameters (Lohmiller and Slotine, 2000b). For instance, consider a closed-loop plant dynamics with state $\mathbf{z}(t)$, desired state vector $\mathbf{z}_d(t)$, parameter estimate vector $\hat{\mathbf{a}}(t)$, in the form

$$\dot{\mathbf{z}} = \dot{\mathbf{z}}_d + \mathbf{f}(\mathbf{z}, t) - \mathbf{f}(\mathbf{z}_d, t) - \mathbf{W}(\mathbf{z}, t)\tilde{\mathbf{a}} \tag{2}$$

where $\tilde{\mathbf{a}}(t) = \hat{\mathbf{a}}(t) - \mathbf{a}$. Letting $\tilde{\mathbf{z}} = \mathbf{z} - \mathbf{z}_d$ and choosing the parameter adaptation law

$$\dot{\hat{\mathbf{a}}} \; = \; \mathbf{PW^T}(\mathbf{z}, \mathbf{t}) \, \tilde{\mathbf{z}}$$

where $\mathbf{P}$ is a constant symmetric positive definite (s.p.d.) gain matrix, Barbalat's lemma (Slotine and Li, 1991) and the Lyapunov-like analysis

$$\dot{V} = \frac{d}{dt} \left(\tilde{\mathbf{z}}^T \tilde{\mathbf{z}} + \tilde{\mathbf{a}}^T \mathbf{P}^{-1} \tilde{\mathbf{a}} \right) = 2 \, \tilde{\mathbf{z}}^T \int_o^1 \frac{\partial \mathbf{f}}{\partial \mathbf{z}}(\mathbf{z}_d + \lambda \tilde{\mathbf{z}})d\lambda \, \tilde{\mathbf{z}}$$

show asymptotic convergence of $\tilde{\mathbf{z}}$ to zero for uniformly negative definite $\frac{\partial \mathbf{f}}{\partial \mathbf{z}}$ and bounded $\ddot{V}$.

Assume now that $\mathbf{f}$ is indeed contracting, and further that some or all of the terms $\mathbf{W}(\mathbf{z}, t)\mathbf{a}$ are contracting in the same metric. Faster convergence can then be achieved by exploiting these terms, similarly

to exploiting viscous friction in mechanical systems (Slotine and Li, 1991). Specifically, let

$$\mathbf{W}(\mathbf{z}, t)\mathbf{a} = \mathbf{W}_c(\mathbf{z}, t)\mathbf{a}_c + \mathbf{W}_o(\mathbf{z}, t)\mathbf{a}_o$$

where $\mathbf{W}_c(\mathbf{z}, t)\mathbf{a}_c$ is contracting in the same metric as $\mathbf{f}$, and let us use $\mathbf{W}_c(\mathbf{z}_d, t)$ instead of $\mathbf{W}_c(\mathbf{z}, t)$ in the adaptive cancellation, leading to

$$\dot{\mathbf{z}} = \dot{\mathbf{z}}_d + \mathbf{f}(\mathbf{z}, t) - \mathbf{f}(\mathbf{z}_d, t) + \mathbf{W}(\mathbf{z}, t)\mathbf{a} - \mathbf{W}_c(\mathbf{z}_d, t)\hat{\mathbf{a}}_c - \mathbf{W}_o(\mathbf{z}, t)\hat{\mathbf{a}}_o$$

Choosing the parameter adaptation

$$\dot{\hat{\mathbf{a}}}_{\mathbf{c}} = \mathbf{P_c}\, \mathbf{W_c^T}(\mathbf{z_d}, \mathbf{t})\, \tilde{\mathbf{z}}$$
$$\dot{\hat{\mathbf{a}}}_{\mathbf{o}} = \mathbf{P_o}\, \mathbf{W_o^T}(\mathbf{z}, \mathbf{t})\, \tilde{\mathbf{z}}$$

leads to

$$\dot{V} = \frac{d}{dt}\left(\tilde{\mathbf{z}}^T \tilde{\mathbf{z}} + \tilde{\mathbf{a}}^T \mathbf{P}^{-1}\tilde{\mathbf{a}}\right) = 2\,\tilde{\mathbf{z}}^T \int_o^1 \frac{\partial \mathbf{f}_c}{\partial \mathbf{z}}(\mathbf{z}_d + \lambda\tilde{\mathbf{z}})d\lambda\,\tilde{\mathbf{z}}$$

where the matrix $\mathbf{P}$ is block-composed of the constant s.p.d. matrices $\mathbf{P}_c$ and $\mathbf{P}_o$, and

$$\mathbf{f}_c(\mathbf{z}, t) = \mathbf{f}(\mathbf{z}, t) + \mathbf{W}_c(\mathbf{z}, t)\mathbf{a}_c$$

and thus to a faster convergence rate.

The discussion extends immediately to the case where $\mathbf{z}$ is a composite variable in a hierarchy, similarly to adaptive sliding control for instance. Also, in the case that $\mathbf{a}_o$ and $\mathbf{a}_c$ contain some parameters in common, the parameter estimates can be updated by simply adding the corresponding adaptation rates. Convergence of $\tilde{\mathbf{z}}$ can be further accelerated by imposing known convex constraints on the parameter estimates.

In system combinations such as those described in the previous subsections, local adaptation loops may be added both to achieve some desired behavior, and to ensure contraction of the local dynamics itself. The overall sytem can then be viewed as the nominal contracting network (i.e., the contracting network obtained in the case all parameters are known), driven by terms of the type $\mathbf{W}(\mathbf{z}, t)\tilde{\mathbf{a}}$, each of which tends asymptotically to zero thanks to the local adaptation process. Asymptotic convergence of the overall system to a unique trajectory is thus achieved.

4   Oscillator Synchronisation

Rythmic phenomena are pervasive in physiology. These include, for instance, the rhythmic motor behaviors used in locomotion, as in walking, swimming, or flying, automatic mechanisms such as breathing and heart cycles, and intrinsic pacemakers in the brain (Kandel, et al., 2000; Dowling, 1992; Kopell and Ermentrout, 2002). Although in general individual oscillator dynamics are not contracting, contraction theory can provide convenient tools to analyze synchronisation and coupling phenomena between oscillators. The results of this section are based on (Wang and Slotine, 2002).

4.1 Basic examples

Example 4.1: Consider two identical Van der Pol oscillators, coupled as

$$\ddot{x}_1 + \alpha(x_1^2 - 1)\dot{x}_1 + \omega^2 x_1 = \alpha k_1(\dot{x}_2 - \dot{x}_1)$$
$$\ddot{x}_2 + \alpha(x_2^2 - 1)\dot{x}_2 + \omega^2 x_2 = \alpha k_2(\dot{x}_1 - \dot{x}_2)$$

where $\alpha > 0$, $\omega > 0$, k_1 and k_2 are arbitrary constants. Let $\mathbf{x}_i = (x_i \ \dot{x}_i)^T$ for $i = 1, 2$. Grouping terms in x_i leads to

$$\ddot{x}_1 + \alpha(x_1^2 + k_1 + k_2 - 1)\dot{x}_1 + \omega^2 x_1 = \ddot{x}_2 + \alpha(x_2^2 + k_1 + k_2 - 1)\dot{x}_2 + \omega^2 x_2 \qquad (3)$$

Now the system

$$\ddot{x} + \alpha(x^2 + k)\dot{x} + \omega^2 x = \alpha\, u(t)$$

can be shown to be contracting for $k > 0$ (Combescot and Slotine, 2000), using

$$\boldsymbol{\Theta} = \begin{pmatrix} \omega & 0 \\ \alpha(x^2 + k) & 1 \end{pmatrix} \qquad (4)$$

Thus, with $k_1 + k_2 > 1$, equation (3) implies that $\mathbf{x}_1$ tends to $\mathbf{x}_2$ as $t \to +\infty$, and thus that the oscillators synchronize (Wang and Slotine, 2001).

Note that the above system may be viewed as a two-way observer, where each sub-system uses the other's velocity as a measurement.

The synchronisation result immediately extends to systems of n coupled oscillators of the form

$$\ddot{x}_1 + \alpha(x_1^2 - 1)\dot{x}_1 + \omega^2 x_1 = \alpha k \left(-\dot{x}_1 + \sum_{i \neq 1} \dot{x}_i\right)$$
$$\ddot{x}_2 + \alpha(x_2^2 - 1)\dot{x}_2 + \omega^2 x_2 = \alpha k \left(-\dot{x}_2 + \sum_{i \neq 2} \dot{x}_i\right)$$
$$\ldots\ldots$$
$$\ddot{x}_n + \alpha(x_n^2 - 1)\dot{x}_n + \omega^2 x_n = \alpha k \left(-\dot{x}_n + \sum_{i \neq n} \dot{x}_i\right)$$

where an identical coupling term is "broadcast" from each oscillator to the others.

Similar mechanisms may occur in collective synchronisation phenomena (Murray, 1993; Strogatz, 1994, 2001), as we shall further discuss. Also note that, comparing (2) with (3), implementing adaptation mechanisms is straightforward. Conversely, two uncertain second-order nonlinear systems of states $(x_i \ \dot{x}_i)^T$ can be controlled to mimic specific coupled oscillators as above, by defining as in Example 3.4

$$s_1 = \dot{x}_1 - \bar{s}_1 \qquad\qquad \dot{\bar{s}}_1 = -k\,(\bar{s}_1 - \dot{x}_1) - \alpha(x_1^2 + k_1 - 1)\dot{x}_1 - \omega^2 x_1 + \alpha k_1 \dot{x}_2$$
$$s_2 = \dot{x}_2 - \bar{s}_2 \qquad\qquad \dot{\bar{s}}_2 = -k\,(\bar{s}_2 - \dot{x}_2) - \alpha(x_2^2 + k_2 - 1)\dot{x}_2 - \omega^2 x_2 + \alpha k_2 \dot{x}_1$$

with constant $k > 0$. The overall target dynamics corresponds to the coupled oscillators. $\qquad\square$

The very elementary model above is also suggestive of a reutilisation mechanism, as we now illustrate.

Example 4.2: Switch connection signs, and consider instead the system

$$\ddot{x}_1 + \alpha(x_1^2 - 1)\dot{x}_1 + \omega^2 x_1 = -\alpha k_1(\dot{x}_1 + \dot{x}_2)$$

13

$$\ddot{x}_2 + \alpha(x_2^2 - 1)\dot{x}_2 + \omega^2 x_2 = -\alpha k_2(\dot{x}_1 + \dot{x}_2)$$

Then x_1 synchronizes exponentially to $-x_2$ (180° phase shift), as can be seen by applying the result of Example 4.1 to x_1 and $x_3 = -x_2$.

Switch signs again, now according to

$$\ddot{x}_1 + \alpha(x_1^2 - 1)\dot{x}_1 + \omega^2 x_1 = \alpha k_1(\dot{x}_2 - \dot{x}_1)$$
$$\ddot{x}_2 + \alpha(x_2^2 - 1)\dot{x}_2 + \omega^2 x_2 = \alpha k_2(-\dot{x}_1 - \dot{x}_2)$$

Then, for $k_1 = k_2 > 1$, the states $\mathbf{x}_1$ and $\mathbf{x}_2$ both tend to zero. Indeed, one has

$$\ddot{x}_1 + \alpha(x_1^2 + k_1 - 1)\dot{x}_1 + \omega^2 x_1 = \alpha k_1 \dot{x}_2$$
$$\ddot{x}_2 + \alpha(x_2^2 + k_1 - 1)\dot{x}_2 + \omega^2 x_2 = -\alpha k_1 \dot{x}_1$$

and furthermore, since Θ in (4) varies only with x, the feedback terms on the right-hand side only add a skew-symmetric term to the overall $\dot{\mathbf{M}} + 2\,\mathbf{M}\,\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$, thus preserving contraction. Alternatively, the result can also be shown using the invariant set theorem with the Lyapunov-like function $V = \sum_i (\dot{x}_i^2 + \omega^2 x_i)$, since this particular coupling is "energy" conserving.

Similar mechanisms may also occur in the control of animal gaits in walking or swimming (Winfree, 2001). $\square$

Example 4.3: The previous results can be generalized. Consider the system

$$\ddot{x}_1 + \alpha(x_1^2 - 1)\dot{x}_1 + \omega^2 x_1 = \alpha\,(\gamma\dot{x}_2 - \kappa\dot{x}_1)$$
$$\ddot{x}_2 + \alpha(x_2^2 - 1)\dot{x}_2 + \omega^2 x_2 = \alpha\,(\gamma\dot{x}_1 - \kappa\dot{x}_2)$$

where α and κ are strictly positive constants and γ is an arbitrary coupling gain. Using the same reasoning as before, under the condition

$$|\gamma| + \kappa > 1$$

x_1 converges to x_2 for all $\gamma \geq 0$ (excitatory coupling), and to $-x_2$ for all $\gamma \leq 0$ (inhibitory coupling). Note that if $\gamma = 0$ (no coupling), x_1 and x_2 both converge to zero, consistent with the previous cases. Whether the systems actually tend to synchronized oscillations or to the origin depends on the value of γ. If $\gamma > 0$, the limit behavior is

$$\ddot{x}_i + \alpha(x_i^2 + \kappa - \gamma - 1)\dot{x}_i + \omega^2 x_i = 0 \qquad i = 1, 2$$

which tends (for non-zero initial conditions) to a stable limit cycle if $\gamma > \kappa - 1$ and to a stable equilibrium at the origin otherwise. Similarly, if $\gamma < 0$, x_1 and x_2 reach anti-synchrony if $\gamma < 1 - \kappa$ and tend to zero otherwise.

Note that if $\kappa = 1$, the convergence to the origin as the system transitions between synchrony and antisynchrony is limited to $\gamma = 0$. $\square$

Conversely, the same mathematics shows that two nonlinear systems initially at rest can be made to oscillate by simple coupling.

Example 4.4: Consider two independent contracting nonlinear systems, initially at equilibrium,

$$\ddot{x}_i + \alpha(x_i^2 + 2k - 1)\dot{x}_i + \omega^2 x_i = 0 \qquad i = 1, 2$$

and couple them according to

$$\ddot{x}_1 + \alpha(x_1^2 + 2k - 1)\dot{x}_1 + \omega^2 x_1 = \alpha k(\dot{x}_1 - \dot{x}_2)$$
$$\ddot{x}_2 + \alpha(x_2^2 + 2k - 1)\dot{x}_2 + \omega^2 x_2 = \alpha k(\dot{x}_2 - \dot{x}_1)$$

Then, from Example 4.2, the system will tend to anti-synchronized (180° phase shift) oscillations. Strictly speaking, of course, initially the coupled system should be disturbed slightly away from its unstable equilibrium at the origin.

In robotic locomotion, for instance, the above could serve as elementary models of pattern generators, or of the legs themselves. $\qquad\square$

Example 4.5: In a seminal paper, (Smale, 1976) showed that two nonlinear systems initially at rest could actually be made to oscillate by *diffusion* coupling, a property suggestive of biological behavior along the lines of (Turing, 1952). This can be easily analyzed using contraction theory (Wang and Slotine, 2002). Using results from (Leonov, et al., 1996), (Pogromsky and Nijmeijer, 2001; Pogromski, et al. 2002) give another instance of such systems, which can also be studied using contraction theory. $\qquad\square$

4.2 Extensions

As discussed extensively in (Wang and Slotine, 2002) and as we now summarize, the results extend to oscillator networks of various coupling topologies through basic matrix algebra calculations. Two cases are of particular interest.

The first is modelling locomotion, a subject with a considerable literature, where simple variations of coupling gains within small networks of oscillators can give rise to typical gaits in vertebrates and insects, extending Examples 4.2 and 4.3.

The second involves very large networks of oscillators, a topic of active current research (Strogatz, 2003), and a possible model of pacemakers (Kandel, et al., 2000) and binding mechanisms (Llinas, 2001) in the brain. For identical oscillators with possibly distinct coupling strengths, one result is that, under simple conditions on the individual Jacobians and the couplings, synchronization will always occurs for strong enough coupling. An explicit upper bound on the corresponding threshold can be computed through eigenvalue analysis, and the effect of adding or removing specific links or nodes can be quantified and related to recent results on "small world" and network robustness (Strogatz, 2001; Barabasi, 2002; Watts, 2003). Another result is that synchronised networks of increasing complexity can be generated through combinations of smaller networks, similarly to the aggregation properties of contracting systems.

The development is based on a simple extension of the techniques above, which we illustrate here on a particular example.

Example 4.6: Consider a unidirectional closed ring of the form

$$\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1, t) + \mathbf{K}(\mathbf{x}_4 - \mathbf{x}_1)$$

$$\dot{\mathbf{x}}_2 = \mathbf{f}(\mathbf{x}_2, t) + \mathbf{K}(\mathbf{x}_1 - \mathbf{x}_2)$$
$$\dot{\mathbf{x}}_3 = \mathbf{f}(\mathbf{x}_3, t) + \mathbf{K}(\mathbf{x}_2 - \mathbf{x}_3)$$
$$\dot{\mathbf{x}}_4 = \mathbf{f}(\mathbf{x}_4, t) + \mathbf{K}(\mathbf{x}_3 - \mathbf{x}_4)$$

with $\mathbf{K} = \mathbf{K}^T > 0$. We can write

$$\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1, t) - \mathbf{K}(2\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{x}}_2 = \mathbf{f}(\mathbf{x}_2, t) - \mathbf{K}(2\mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{x}}_3 = \mathbf{f}(\mathbf{x}_3, t) - \mathbf{K}(2\mathbf{x}_3 + \mathbf{x}_4 + \mathbf{x}_1) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{x}}_4 = \mathbf{f}(\mathbf{x}_4, t) - \mathbf{K}(2\mathbf{x}_4 + \mathbf{x}_1 + \mathbf{x}_2) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$

Consider now the auxiliary system, driven by the $\mathbf{x}_i$,

$$\dot{\mathbf{y}}_1 = \mathbf{f}(\mathbf{y}_1, t) - \mathbf{K}(2\mathbf{y}_1 + \mathbf{y}_2 + \mathbf{y}_3) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{y}}_2 = \mathbf{f}(\mathbf{y}_2, t) - \mathbf{K}(2\mathbf{y}_2 + \mathbf{y}_3 + \mathbf{y}_4) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{y}}_3 = \mathbf{f}(\mathbf{y}_3, t) - \mathbf{K}(2\mathbf{y}_3 + \mathbf{y}_4 + \mathbf{y}_1) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$
$$\dot{\mathbf{y}}_4 = \mathbf{f}(\mathbf{y}_4, t) - \mathbf{K}(2\mathbf{y}_4 + \mathbf{y}_1 + \mathbf{y}_2) + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i$$

and assume that it can be shown that this system is contracting. Then $\forall i$, $\mathbf{y}_i(t)$ tends to $\mathbf{x}_i(t)$ (since for the overall systems one can write as usual $\dot{\mathbf{y}} - \mathbf{g}(\mathbf{y}, t) = \dot{\mathbf{x}} - \mathbf{g}(\mathbf{x}, t)$, with $\mathbf{g}$ a contracting dynamics). Furthermore, the system tends to the particular solution $\forall i$, $\mathbf{y}_i(t) = \mathbf{y}_\infty(t)$ with

$$\dot{\mathbf{y}}_\infty = \mathbf{f}(\mathbf{y}_\infty, t) - 4\,\mathbf{K}\mathbf{y}_\infty + \mathbf{K}\sum_{i=1}^{4}\mathbf{x}_i \tag{5}$$

Hence all $\mathbf{x}_i(t)$ tend towards $\mathbf{y}_\infty(t)$ and thus synchronise.

Thus, synchronisation can be determined by examining the negative definiteness of the generalized Jacobian of the overall $\mathbf{y}$ system in some appropriate metric. Consider for simplicity the identity metric, and thus the matrix

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 - 2\mathbf{K} & -\mathbf{K} & -\mathbf{K} & 0 \\ 0 & \mathbf{J}_2 - 2\mathbf{K} & -\mathbf{K} & -\mathbf{K} \\ -\mathbf{K} & 0 & \mathbf{J}_3 - 2\mathbf{K} & -\mathbf{K} \\ -\mathbf{K} & -\mathbf{K} & 0 & \mathbf{J}_4 - 2\mathbf{K} \end{bmatrix}$$

where $\mathbf{J}_i = \partial \mathbf{f}(\mathbf{x}_i, t)/\partial \mathbf{x}_i$ for $i = 1, 2, 3, 4$. The symmetric part of $\mathbf{J}$ can be written

$$\frac{1}{2}\,(\,\mathbf{J} + \mathbf{J}^T\,) = \mathrm{diag}(\mathbf{J}_{is} - \mathbf{K}) - \frac{1}{2}\,\mathbf{K}\begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \end{bmatrix} - \frac{1}{2}\,\mathbf{K}\begin{bmatrix} \mathbf{I} & 0 & \mathbf{I} & 0 \\ 0 & \mathbf{I} & 0 & \mathbf{I} \\ \mathbf{I} & 0 & \mathbf{I} & 0 \\ 0 & \mathbf{I} & 0 & \mathbf{I} \end{bmatrix}$$

where $\mathbf{J}_{is}$ is the symmetric part of $\mathbf{J}_i$, and $\mathbf{I}$ is the identity matrix in the dimension of the subsystems. Assuming $\mathbf{K} \geq 0$, the last two matrices are negative semi-definite, and thus the system is contracting if the $(\mathbf{J}_i - \mathbf{K})$ are uniformly negative definite. Note that this implies that dynamics (5) is contracting as well (as it should, since synchronization defines both a linear constraint and an invariant set). All the $\mathbf{x}_i$ will then synchronize.

16

Specializing this computation to a unidirectional ring of identical van der Pol oscillators shows that the oscillators synchronise for $k \geq 1$.

This type of manipulation, extending the principle of Example 4.1 by exhibiting an auxiliary contracting system, can be made very general. Furthermore, although coupling symmetries simplify calculations, they are not of fundamental importance for synchronisation. $\square$

It should be clear from the above example that the development is not restricted to oscillators, but applies to general networks of "partially" contracting subsystems.

Example 4.7: Consider for instance a simplified model of schooling or flocking similar to (Jadbabaie, et al., 2003), in continuous-time

$$\dot{\mathbf{x}}_i = -\mathbf{K} \sum_{j \in \mathcal{N}_i(t)} (\mathbf{x}_i - \mathbf{x}_j) \qquad i = 1, ..., n$$

where the $\mathbf{x}_i$ denote heading vectors of the subsystems, $\mathcal{N}_i(t)$ the set of indices of the nearest neighbours of subsystem i at current time t (defined for instance as the subsystems within a certain distance of subsystem i), and $\mathbf{K} > \mathbf{0}$ is a constant matrix. The sets $\mathcal{N}_i(t)$ can change abruptly and *asynchronously*. We assume for notational simplicity that n is constant and that the group remains connected.

Construct as in the previous example the auxiliary system (with $\mathcal{N}_i(t)$ unchanged)

$$\dot{\mathbf{y}}_i = -\mathbf{K} \sum_{j \in \mathcal{N}_i(t)} (\mathbf{y}_i - \mathbf{y}_j) - \mathbf{K} \sum_{j=1}^{n} \mathbf{y}_j + \mathbf{K} \sum_{j=1}^{n} \mathbf{x}_j \qquad i = 1, ..., n$$

Although it can be discontinuous in time, the corresponding Jacobian $\mathbf{J}(t)$ is symmetric and uniformly negative definite

$$\forall t \geq 0, \forall \mathbf{v}, \quad \mathbf{v}^T \mathbf{J}(t)\mathbf{v} = - \sum_{\text{active links}} (\mathbf{v}_j - \mathbf{v}_i)^T \mathbf{K} (\mathbf{v}_j - \mathbf{v}_i) - \left(\sum_{i=1}^{n} \mathbf{v}_i\right)^T \mathbf{K} \left(\sum_{i=1}^{n} \mathbf{v}_i\right)$$

This implies that $\delta \mathbf{y}^T \delta \mathbf{y}$ (which is continuous in time, but whose time-derivative can be discontinuous at discrete instants) is upper bounded by a vanishing exponential. Thus $\mathbf{y}$ converges exponentially to the particular solution $\forall i, \mathbf{y}_i(t) = \mathbf{y}_\infty(t)$, with

$$\dot{\mathbf{y}}_\infty + n \mathbf{K} \mathbf{y}_\infty = \mathbf{K} \sum_{j=1}^{n} \mathbf{x}_j$$

Since by construction $\mathbf{y}$ tends to $\mathbf{x}$ exponentially, this implies in turn that the headings $\mathbf{x}_i$ tend exponentially towards a constant common value.

Note that the analysis carries on straightforwardly to nonlinear couplings, for instance

$$\dot{\mathbf{x}}_i = - \sum_{j \in \mathcal{N}_i(t)} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{K} + \mathbf{R} |\mathbf{x}_i - \mathbf{x}_j|)$$

17

with constant $\mathbf{K} > \mathbf{0}$, $\mathbf{R} > \mathbf{0}$, corresponding to subsystems reacting faster to larger deviations. It also extend immediately to time-varying or link-dependent coupling gains

$$\dot{\mathbf{x}}_i = - \sum_{j \in \mathcal{N}_i(t)} (\mathbf{x}_i - \mathbf{x}_j) \, (\, \mathbf{K}_{ij}(t) \, + \, \mathbf{R}_{ij}(t) \, |\mathbf{x}_i - \mathbf{x}_j| \,)$$

as long as these gains are uniformly strictly positive and symmetric in i, j. For instance the gains may depend smoothly on distance.

While in this particular case the final behavior is actually known (since $\sum_i x_i$ is constant), the reasoning applies just as well to asynchronous oscillator synchronisation for large enough interaction gain, in a way reminiscent e.g. of Bose-Einstein condensation. $\square$

Finally, one can easily verify that the above results can be viewed as particular instances of the following very general principle (Wang and Slotine, 2002). Consider a nonlinear system of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{x}, t)$$

and the auxiliary, observer-like contracting system

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{x}, t)$$

If a particular solution of the auxiliary system verifies a smooth specific property, then all trajectories of the original system verify this property exponentially − indeed, another particular solution of the auxiliary contracting system is $\mathbf{y}(t) = \mathbf{x}(t), \forall t \geq 0$.

5   Concluding remarks

Extending (Slotine and Lohmiller, 2001), this paper surveys further examples of biologically motivated stable distributed computations. Further work will study applications of locally contracting systems with singular metrics, as can occur in hierarchies of dynamic classifiers or in phase-locking of diverse oscillators, as well as p.d.e. versions of the results.

REFERENCES

Arnold, V.I. (1978). Mathematical Methods of Classical Mechanics, *Springer Verlag*.
Arimoto, S., and Naniwa, T. (2002). Learnability and Adaptability from the Viewpoint of Passivity Analysis, *Intelligent Automation and Soft Computing, Vol. 8, No. 2 2002*.
Ashby, W.R. (1966). Design for a Brain : the Origin of Adaptive Behavior, 2nd ed.,*Science Paperbacks*

Atkeson, C. G., Moore, A. W., and Schaal, S. (1997). Locally Weighted Learning for Control, *Artificial Intelligence Review, 11*.

Barabasi, A.L. (2002). Linked: The New Science of Networks, *Perseus*.

Bernstein, N. (1967). The Coordination and Regulation of Movements, *Pergamon*.

Berthoz, A. (1999). The Sense of Movement, *Harvard University Press*.

Bertsekas, D., and Tsitsiklis, J. (1989). Parallel and Distributed Computation, *Prentice-Hall*.

Bizzi E., Giszter S.F., Loeb E., Mussa-Ivaldi F.A., Saltiel P (1995). Modular organization of motor behavior in the frog's spinal cord, *Trends in Neurosciences. Review 18:442-445*.

Burridge R., Rizzi, A., and Koditschek, D. (1999). Sequential Composition of Dynamically Dexterous Robot Behaviors, *Int. J. Robotics Res. 18(6).*

Braitenberg, V. (1984). Vehicles: Experiments in Synthetic Psychology, *MIT Press*.

Brooks, R. (1986). A Robust Layered Control System for a Mobile Robot, *I.E.E.E. J. Robotics and Automation, 2(1).*

Brooks, R. (1999). Cambrian Intelligence, *M.I.T. Press*.

Chomsky, N. (1957). Syntactic Structures.*Mouton, The Hague*.

Combescot, C., and Slotine, J.J.E. (2000). A Study of Coupled Oscillators Using Contraction Theory, *MIT Nonlinear Systems Laboratory, Report MIT-NSL-000801*.

Crick, F. (1994). The Astonishing Hypothesis, *Scribner*

Crick, F., and Koch, C. (1998). Constraints on Cortical and Thalamic projections: the No-Strong-Loops Hypothesis, *Nature, 491*.

Damasio, A. (2001). Fundamental feelings, *Nature, 413*.

Damasio, A. (2003). Looking for Spinoza, *Harcourt*.

D'Avella, A., and Bizzi, E. (1998). Low Dimensionality of Supraspinally Induced Force Fields, *Proc. Natl. Acad. Sci. USA 95*.

D'Avella, A., Saltiel, P., Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior, *Nature Neuroscience*.

Dawkins, R. (1976)., The Selfish Gene, *Oxford University Press*.

Dennett, D. (1995). Darwin's Dangerous Idea, *Simon and Schuster*.

Dehaene, S., Kerzberg, M., and Changeux, J.-P. (1998), *Proc. Natl. Acad. Sci. USA 95, pp. 14529-14534*.

Dickmanns, E. (1998). Dynamic Vision for Intelligent Vehicles, *M.I.T. Course Notes*.

Dowling, J.E. (1992). Neurons and Networks, *Belknap*.

Droulez, J., et al. (1983). Motor Control, *7th International Symposium of the International Society of Posturography, Houston, Karger, 1985*.

Edelman, G., and Tononi, G. (2000). A Universe of Consciousness, *Basic Books*.

Floreano, D., and Nolfi, S. (2000). Evolutionary Robotics, *MIT Press*.

Fod, A., Mataric, M., Jenkins, O.C. (2000). Automated Extraction of Primitives for Movement Classification, *I.E.E.E. Int. Conf. Humanoid Robots, Cambridge, MA*.

Fuhrmann, G., Segev, I., Markram, H. and Tsodyks, M. (2001). Coding of Temporal Information by Activity Dependent Synapses. *Journal of Neurophysiology*.

Giszter, S., and Kargo, W. (2000). Conserved Temporal Dynamics and Vector Superposition of Primitives in Frog Wiping Reflexes During Spontaneous Extensor Deletions, *Neurocomputing 32-33*.

Grossberg, S. (2000). The complementary brain: A unifying view of brain specialization and modularity. *Trends in Cognitive Sciences, in press*.

Hartmann, P. (1982). Ordinary differential equations, second ed., *Birkhauser*.

Horn, R.A., and Johnson C.R. (1985). Matrix Analysis, *Cambridge University Press*.

Jadbabaie A., Lin J., and Morse A.S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbour rules, *I.E.E.E. Trans. Autom. Control, 14(5)*.

Jordan, M., and Wolpert D. (1999). Computational Motion Control, *The Cognitive Neurosciences, Gazzaniga (Ed.), M.I.T. Press*.

Kandel, E.R., Schwartz, J.H., and Jessel, T.M. (2000). Principles of Neural Science, 4th ed.*McGraw-Hill*.

Kargo, W., and Giszter, S. (2000). Rapid Correction of Aimed Movements by Summation oof Force-Field Primitives, *J. Neuroscience 2(1)*.

Keat, J., P. Reinagel, R. C. Reid, and M. Meister (2001). Predicting every spike: a model for the responses of visual neurons, *Neuron. 30:803-17*.

Khatib, O. (1986) Real-time obstacle avoidance for manipulators and mobile robots. *International Journal of Robotics Research, 5(1)*.

Kirschner, M., and Gerhart, J. (1998). Evolvability. *Proc. Nat. Acad. Sci. 95:8420-8427*.

Kopell N., and Ermentrout, G.B. (2002). Mechanisms of phase-locking and frequency control in pairs of coupled neural oscillators, *Handbook on Dynamical Systems, vol. 2: Toward applications. Ed. B. Fiedler, Elsevier*.

Ledoux, J. (1996). The Emotional Brain. *Simon and Schuster*.

Leonov, G., Burkin I., and Shepeljavyi A. (1996). Frequency methods in oscillation theory, *Kluwer*.

Llinas, R. (2001). I of the Vortex: From Neurons to Self, *MIT Press*.

Lohmiller, W., and Slotine, J.J.E. (1998). On Contraction Analysis for Nonlinear Systems, *Automatica 34(6)*.

Lohmiller, W., and Slotine, J.J.E. (1999). Contraction Analysis and Observer Design for Nonlinear Partial Differential Equations, *New Directions in Nonlinear Observer Design, Niejmeijer H. and Fossen T., Eds., LNCIS 244, Springer-Verlag*.

Lohmiller, W., and Slotine, J.J.E. (2000a). Control System Design for Mechanical Systems Using Contraction Theory, *I.E.E.E. Transactions on Automatic Control 45(5)*.

Lohmiller, W., and Slotine, J.J.E. (2000b). Nonlinear Process Control Using Contraction Theory, *A.I.Ch.E. Journal*.

Lohmiller, W., and Slotine, J.J.E. (2001). Contraction Analysis of Nonlinear Partial Differential Systems, *I.E.E.E. M.M.A.R., Poland*.

Luenberger, D.G. (1979). Introduction to dynamic systems, *Wiley*.

Minsky, M. (1986). The Society of Mind, *M.I.T. Press*.

Murray, J.D. (1993). Mathematical Biology, 2nd Ed., *Springer Verlag*.

Mussa-Ivaldi, F.A. (1997). *1997 I.E.E.E. International Symposium on Computational Intelligence in Robotics and Automation, pp. 84-90*.

Nijmeijer, H. (2001). A dynamical control view on synchronisation, *Physica D, 154*, 219-228.

Pecora, L.M., and Carroll, T.L. (1990). Synchronisation in chaotic systems, *Phys. Rev. Letter, 64*, 821-824.

Pecora, L.M., Carroll, T.L. , Johnson, G.A., and Mar, D.J. (1997). Fundamentals of synchronisation in chaotic systems, *Chaos 7(4)*.

Pogromsky, A., and Nijmeijer, H. (2001). Cooperative oscillatory behavior of mutually coupled dynamical systems, *I.E.E.E. Trans. Circuits and Systems, 48(2)*, 152-162.

Pogromsky, A. Santoboni, G, and Nijmeijer H. (2002) Partial synchronization: from symmetry towards stability. *Physica D, 172, 65-87*.

Popov, V.M. (1973). Hyperstability of Control Systems, *Springer-Verlag*.

Raibert, M. (1986). Legged robots that balance, *M.I.T. Press*.

Raibert, M., et al. (1983). 3-D balance using 2-D algorithms, *Int. Symp. Rob. Res., M.I.T. Press*.

Ridley, M. (2000). Mendel's demon, *Free Press*.

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences, 3(6)*.

Soechting J., and Lacquaniti F. (1988). Quantitative evaluation of the electromyographic responses to multidirectional load perturbations of the human arm, *J Neurophysiol. 59(4)*.

Simon, H.A. (1962). The Architecture of Complexity, *Proc. Am. Philo. Soc. 106*.

Simon, H.A. (1981). The Sciences of the Artificial, 2nd edition, *M.I.T. Press*.

Schwartz, L. (1993). Analyse, *Hermann, Paris*.

Slotine, J.J.E., and Li, W. (1991). Applied Nonlinear Control, *Prentice-Hall*.

Slotine, J.J.E., and Lohmiller, W. (2001). Modularity, Evolution, and the Binding Problem: A View from Stability Theory, *Neural Networks, 14*.

Smale, S. (1976). A mathematical model of two cells via Turing's equation, *The Hopf Bifurcation and its Applications, Mardsen, J. and Mc Cracken M. Eds, Springer-Verlag*.

Strogatz, S. (1994). Nonlinear dynamics and chaos, *Addison Wesley*.

Strogatz, S. (2001). Exploring complex networks, *Nature*, 410.

Strogatz, S. (2003). Sync, The Emerging Science of Spontaneous Order, *Hyperion*.

Thoroughman, K.A., and Shadmer, R. (2000). Learning of action through adaptive combination of motor primitives, *Nature*.

Tononi, G., et al. (1998). *Proc. Natl. Acad. Sci. USA 95, pp. 3198-3203*.

Tresch, M., Saltiel, P. , and Bizzi, E. (1999). The Construction of Movement by the Spinal Cord, *Nature Neuroscience 2*.

Tresch, M. (2002). Personal communication.

Turing, A. (1952). The Chemical Basis of Morphogenesis, *Phil. Trans. Roy. Soc. Lond., B237*.

Walter, W.G. (1950). An Imitation of Life, *Scientific American*.

Walter, W.G. (1951). A Machine that Learns, *Scientific American*.

Wang, W., and Slotine, J.J.E. (2001). Synchronisation of bilaterally coupled oscillators, *MIT Nonlinear Systems Laboratory, Report MIT-NSL-010201*.

Wang, W., and Slotine, J.J.E. (2002). A Study of Synchronisation and Phase Locking using Contraction Theory, *MIT Nonlinear Systems Laboratory, Report MIT-NSL-021201*, submitted.

Watts, D.J.(2003). Six Degrees: The Science of a Connected Age, *W.W. Norton & Company*.

Watts, L. (2002). http://www.lloydwatts.com/neuroscience.shtml.

Wiener, N. (1961). Cybernetics, *M.I.T. Press*.

Williamson, M. (1999). Robotic Oscillators, *Doctoral Thesis, M.I.T. Dept. of Electrical Engineering and Computer Science*.

Winfree, A. (2001). The Geometry of Biological Time, *Springer Verlag*.

Wolpert, D., and Kawato, M. (1999). Multiple Paired Forward and Inverse Models for Motor Control, *Neural Networks, 11.*.

Won, J., and Hogan, N. (1995). Stability properties of human reaching movements, *Exp Brain Res. 107(1)*.

Young, K.D., and Ozguner, U. (1993). Frequency shaping compensator design for sliding mode, *Int. J. Control, 57(5), 1005-1019*.

Zajac (1989). Muscle and Tendon Properties, *Crit. Rev. Biomed. Eng., 17*.

Zuo, L., and Slotine, J.J.E. (2003). Robust vibration isolation using frequency shaped sliding variables, *MIT Nonlinear Systems Laboratory, Report MIT-NSL-030201*, submitted.