

CONSTRAINTS ON THE FORM OF NEURONAL CODES: A DYNAMICAL PERSPECTIVE

Karl J. Friston

The Wellcome Dept. of Cognitive Neurology, University College London, 12, Queen Square,
London UK WC1N 3BG

k.friston@fil.ion.ucl.ac.uk

<http://www.fil.ion.ucl.ac.uk>

1. Codes and coupling

The brain can be regarded as an ensemble of connected dynamical systems and as such conforms to some simple principles relating the inputs and outputs of its constituent parts. The ensuing implications, for the way we think about, and measure, neuronal interactions can be quite profound. These range from: (i) Implications for which aspects of neuronal activity are important to measure and how to characterise coupling among neuronal populations. (ii) Implications for understanding the emergence of dynamic receptive fields and functionally specialised brain architectures, and (iii) teleological implications pertaining the genesis of dynamic instability and complexity, that is necessary for adaptive self-organisation.

This presentation focuses on the first set of implications (Friston 2000) by looking at neuronal interactions, coupling and implicit neuronal codes from a dynamical perspective. By considering the brain in this light one can show that a sufficient description of neuronal activity must comprise activity at the current time *and its recent history*. This history constitutes a neuronal transient. As such transients represent an essential metric of neuronal interactions and, implicitly, a code employed in the functional integration of brain systems. The nature of transients, expressed conjointly in different neuronal populations, reflects the underlying coupling among brain systems. A complete description of this coupling, or *effective connectivity*, can be expressed in terms of generalised convolution kernels [Volterra kernels] that embody high-order or nonlinear interactions. This coupling may be *synchronous*, and possibly oscillatory, or *asynchronous*. A critical distinction between synchronous and asynchronous coupling is that the former is essentially linear and the latter is nonlinear. The nonlinear nature of asynchronous coupling enables the rich, context-sensitive interactions that characterise real brain dynamics, suggesting that it plays a role in functional integration that may be as important as synchronous interactions.

2. Neuronal Transients

The assertion that meaningful measures of brain dynamics have a temporal domain is neither new nor contentious (*e.g.* von der Malsburg 1983, Engel *et al* 1991, Aertsen *et al* 1994, Freeman and Barrie 1994, Abeles *et al* 1995, deCharms and Merzenich 1996). A straightforward analysis demonstrates its veracity: Suppose that one wanted to posit some variables x that represented a complete and self-consistent description of brain activity. In other words, everything needed to determine the evolution of the brain's state, at a particular place and time, was embodied in these measurements. Consider a component of the brain

(e.g. a neuron or neuronal population). If such a set of variables existed for this component system they would satisfy some immensely complicated nonlinear equations

$$\frac{\partial x(t)}{\partial t} = f(x(t), u(t)) \quad 1$$

where x is a huge vector of state variables which range from depolarisation at every point in the dendritic tree to the phosphorylation status of every relevant enzyme, from the biochemical status of every glial cell compartment to every aspect of gene expression.. $u(t)$ represents a set of inputs conveyed by afferent from other regions. Eq(1) simply says that the changes in the state variables are nonlinear functions of the variables themselves and some inputs. The vast majority of these variables are hidden and not measurable directly. However, there are a small number of derived measurements y that can be made,

$$y(t) = \lambda(x(t)) \quad 2$$

such as activities of whole cells or populations. These activities could be measured in many ways, for example firing at the initial segment of an axon or local field potentials. The problem is that a complete and sufficient description appears unattainable, given that the underlying state variables cannot be observed directly. This is not the case. The resolution of this apparent impasse rests upon two things. (i) Firstly, a fundamental mathematical equivalence relating the inputs and outputs of a dynamical system and (ii) the fact that these measurable outputs constitute the inputs to other cells or populations.

3. A fundamental equivalence

Assume that every neuron in the brain is modelled by a nonlinear dynamical system of the sort described by Eq(1). Under this assumption it can be shown that *the output is a function of the recent history of its inputs*.

$$y(t) = h(u(t - \sigma)) \quad 3$$

where $u(t - \tau)$ represents the inputs in the recent past. Furthermore this relationship can be expressed as a Volterra series of the inputs (see Appendix). The critical thing here is that we never need to know the underlying and 'hidden' variables that describe the details of each cell's electrochemical and biochemical status. We only need to know the history of its inputs, which, of course, are the outputs of other cells. Because the inputs and outputs are homologous we can rewrite Eq(3) so that all connected brain systems are considered together.

$$y(t) = h(y(t - \sigma)) \quad 4$$

Eq(4) is, in principle, a sufficient description of brain dynamics and involves the variables $y(t - \sigma)$ that represent activity at all times σ preceding the moment in question. These are simply neuronal transients. The degree of transience depends on how far back in time it is necessary to go to fully capture the brain's dynamics. The sensible nature of Eq(4) can be seen readily.. For example, we wanted to determine the behaviour of a cell in V1 (primary visual cortex) then we would need to know the activity of all connected cells in the immediate vicinity over the last millisecond or so to account for propagation delays down afferent axons. We would also need to know the activity in distant sources, like the lateral

geniculate nucleus and higher cortical areas that send afferents, some ten or more milliseconds ago. In short we need the recent history of all inputs.

4. Conclusion

The first conclusion is that neuronal transients are necessary to specify brain dynamics. The second conclusion is that the influence one neuronal population exerts over another can be formulated in terms of a Volterra series. This implies that a complete characterisation of these influences (*i.e.* effective connectivity – see Appendix) comprises the Volterra kernels that are applied to the inputs to yield the outputs. These kernels embody the invariant dynamical structure of neuronal transactions and constrain the form that codes can take. By considering the shape and order of the kernels, a natural taxonomy of neuronal codes emerges. This taxonomy will be presented and used to highlight the potential importance of nonlinear coupling and asynchronous interactions.

Appendix

A1. Input-state-output systems and Volterra series

Neuronal systems are inherently nonlinear and lend themselves to modelling by nonlinear dynamical systems. However due to the complexity of biological systems it is difficult to find analytic equations that describe them adequately. Even if these equations were known the state variables are often not observable. An alternative approach to identification is to adopt a very general model (Wray and Green 1994) and focus on the inputs and outputs. Consider the single input-single output (SISO) system

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= \lambda(x(t))\end{aligned}\tag{A.1}$$

The Fliess fundamental formula (Fliess *et al* 1983) describes the causal relationship between the outputs and the recent history of the inputs. This relationship can be expressed as a Volterra series, which expresses the output $y(t)$ as a nonlinear convolution of the inputs $u(t)$, critically without reference to the state variables $x(t)$. This series is simply a functional Taylor expansion of $y(t)$ in Eq(3)

$$\begin{aligned}y(t) = h(u(t - \sigma)) &= \kappa_0 + \sum_{i=1}^{\infty} \int_0^t \dots \int_0^t \kappa_i(\sigma_1, \dots, \sigma_i) u(t - \sigma_1) \dots u(t - \sigma_i) d\sigma_1 \dots d\sigma_i \\ \kappa_i(\sigma_1, \dots, \sigma_i) &= \frac{\partial^i y(t)}{\partial u(t - \sigma_1) \dots \partial u(t - \sigma_i)}\end{aligned}\tag{A.2}$$

where $\kappa_i(\sigma_1, \dots, \sigma_i)$ is the i th order kernel. Volterra series have been described as a 'power series with memory' and are generally thought of as a high-order or 'nonlinear convolution' of the inputs to provide an output. See Bendat (1990) for a fuller discussion.

A.2 Volterra kernels and effective connectivity

Volterra kernels are essential in characterising the effective connectivity or influences that one neuronal system exerts over another because they represent the causal characteristics of the system in question. Neurobiologically they have a simple and compelling interpretation – they are synonymous with effective connectivity: From A.2

$$\kappa_1(\sigma_1) = \frac{\partial y(t)}{\partial u(t - \sigma_1)}, \quad \kappa_2(\sigma_1, \sigma_2) = \frac{\partial^2 y(t)}{\partial u(t - \sigma_1) \partial u(t - \sigma_2)}, \quad \dots$$

It is evident that the first-order kernel embodies the response evoked by a change in input at $t - \sigma_1$. In other words it is a time-dependant measure of *driving* efficacy. Similarly the second order kernel reflects the *modulatory* influence of the input at $t - \sigma_1$ on the evoked response at $t - \sigma_2$. And so on for higher orders.

References

- Abeles M, Bergman H, Gat I, Meilijson I, Seidmann E, Tishby N and Vaadia E (1995) Cortical activity flips among quasi-stationary states. *Proc. Natl. Acad. Sci.* **92**:8616-8620
- Aertsen A, Erb M and Palm G (1994) Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation, *Physica D* **75**: 103 - 128
- Bendat JS (1990) *Nonlinear System Analysis and Identification from Random Data*. John Wiley and Sons, New York USA
- deCharms RC and Merzenich MM (1996) Primary cortical representation of sounds by the co-ordination of action potential timing. *Nature* **381**:610-613
- Engel A K, Konig P and Singer W (1991) Direct physiological evidence for scene segmentation by temporal coding. *Pro Natl Acad Sci USA* vol. **88**: 9136-40
- Fliess M, Lamnabhi M and Lamnabhi-Lagarrigue F (1983) An algebraic approach to nonlinear functional expansions. *IEEE Transactions on Circuits and Systems* **30**:554-570
- Freeman W and Barrie J (1994) Chaotic oscillations and the genesis of meaning in cerebral cortex. *In Temporal Coding in the Brain* Eds. Buzsaki, Llinas R, Singer W, Berthoz A and Christen T. Springer Verlag, Berlin. p13-38
- Friston KJ (2000) The Labile Brain I: Neuronal Transients and nonlinear coupling. *Phil. Trans. R. Soc. (Lond.)* **355**:215-236
- von der Malsburg C (1985) Nervous structures with dynamical links. *Ber Bunsenges. Phys. Chem.* **89**:703-710
- Wray J and Green GGR (1994) Calculation of the Volterra kernels of non-linear dynamic systems using an artificial neuronal network. *Biol. Cybern.* **71**:187-195