ComPlexUs
ORIGINAL RESEARCH PAPER
ComPlexUs 2003;1:123–133; DOI: 22988476100
Received: February 17, 2003
Accepted after revision: June 30, 2003

Graphic Requirements for Multistationarity

Christophe Soulé

CNRS et Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France

Key Words
Interaction graph · Multistationarity · Jacobian matrix · Global univalence

Abstract
We discuss properties which must be satisfied by a genetic network in order for it to allow differentiation. These conditions are expressed as follows in mathematical terms. Let F be a differentiable mapping from a finite dimensional real vector space to itself. The signs of the entries of the Jacobian matrix of F at a given point *a* define an interaction graph, i.e. a finite oriented finite graph $G(a)$ where each edge is equipped with a sign. René Thomas conjectured 20 years ago that if F has at least two nondegenerate zeroes, there exists a such that $G(a)$ contains a positive circuit. Different authors proved this in special cases, and we give here a general proof of the conjecture. In particular, in this way we get a necessary condition for genetic networks to lead to multistationarity, and therefore to differentiation. We use for our proof the mathematical literature on global univalence, and we show how to derive from it several variants of Thomas' rule, some of which had been anticipated by Kaufman and Thomas.

Synopsis

Despite their highly specialized characteristics, each of the more than 250 distinct kinds of cells in the human body carries an identical copy of the DNA. Hence, a fundamental task of biology is to understand cell differentiation – what it is that makes a neuron so different from a blood cell, and how these differences come to be. As early as 1948, Max Delbrück suggested that since identical DNA does not imply identical patterns of gene expression, cell differentiation in humans or in any other organism might be associated with the existence of distinct states of expression in the genetic regulatory networks of the cell. Each type of cell would then correspond to a distinct 'attractor' in the dynamics of the network of interacting genes and proteins, and the process of differentiation would involve the transition of the cell from one such state to another.

Following this picture, a principal aim of modern molecular biology is to elucidate the nature of development by exploring the dynamics of gene regulatory networks at the systems level. A gene, when expressed, leads to the production of proteins that can act to either activate or inhibit the expression of other genes. Within networks of such interacting genes and proteins, feedback loops play a central role in controlling the dynamics, and biologists have for many years recognized that negative feedback loops frequently help to stabilize gene expression. In the simplest case, for example, the protein product of a gene may act to inhibit the synthesis of that very gene. In this case, as the concentration of the protein increases, it will ultimately turn off its own synthesis.

What about positive feedback loops? In a positive feedback loop, the expression of some gene leads to a cascade of cause and effect that eventually feeds back to trigger a further increase in its expression, which then, in turn, triggers a still further increase. In the absence of mitigating factors, positive feedback of this kind would

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com
KARGER

Christophe Soulé
IHÉS 35 Route de Chartres
FR–91440 Bures-sur-Yvette (France)
Tel. +33 1 60 92 66 00, Fax +33 1 60 92 66 09
E-Mail soule@ihes.fr

When studying interactions in a system of biochemical compounds, it is quite rare that one obtains quantitative results. One can show that, in a given tissue (the product of) a gene A is an activator (or a repressor) of the expression of a gene B, but the strength of this interaction, the concentrations and their kinetics are usually unknown. The resulting information is essentially summarized by an *interaction graph* G, by which one means a finite oriented graph together with a sign for every edge. The vertices correspond to the members of the network, and there is a positive (respectively negative) edge from A to B when A activates (respectively represses) the synthesis of B. Note that there can exist both a positive and a negative edge from A to B, since A can activate B at a certain concentration and repress it at another one.

These interaction graphs can be quite complicated. It is therefore very desirable to find ways to use them to restrict the possible behavior of the network they represent. In this paper, we shall address the question of when the network is susceptible to have several stationary states. A beautiful conjecture of Thomas [1] asserts that a *necessary* condition for multistationarity is that G has an (oriented) circuit C which is *positive*, i.e. such that the product of the signs of the edges of C is positive, at least in part of the phase space. This was already proved in several cases [2–5], and we shall present a proof in the general case (Theorem 1).

To formulate the question in mathematical terms, we choose a continuous model. Let $n \geq 1$ be an integer, and let F be a differentiable map form $\mathbb{R}^n$ to itself. The evolution of a network of n compounds can be modelled by the differential system

$$\frac{dx}{dt} = F(x),$$

where x is a differential path in $\mathbb{R}^n$: the components of $x(t)$ are the different concentrations at time t. Asking whether this system has several stationary states amounts to asking whether F has several zeroes. Given any $a \in \mathbb{R}^n$, the interaction graph $G(a)$ is defined from the signs of the partial derivatives $(\partial f_i/\partial x_j)(a)$ of the components of $F = (f_i)$ at a (see 2.1 below for a precise definition): a positive (respectively negative) sign indicates that j is an activator (respectively a repressor) of i. The precise formulation of Thomas' rule is then the following: if F has at least two (nondegenerate) zeroes in $\mathbb{R}^n$, there exists $a \in \mathbb{R}^n$ such that $G(a)$ contains a positive circuit.

The main remark leading to a proof of this assertion is the following. If F has several zeroes, it cannot be univalent (i.e. one to one). Therefore, if we know *sufficient* conditions for F to be univalent, their negation will give necessary conditions for F to have several zeroes. We might then deduce properties of $G(a)$ from these necessary conditions.

It turns out that finding sufficient conditions for F to be univalent is a classical issue in mathematical economy, when one wants to know that the factor prices are uniquely determined by the prices for goods. Several results have been obtained with this application in mind. Gale and Nikaido [6] gave an elegant criterion of univalence in terms of the Jacobian matrix $J(x) = (\partial f_i/\partial x_j)(x)$: assume that for every $x \in \mathbb{R}^n$, the principal minors (respectively the determinant) of $J(x)$ are nonnegative (respectively is positive), then F is univalent. As we shall see, this result leads precisely to a proof of Thomas' conjecture (Theorem 1).

Following the same line of arguments, we can also apply variants of the theorem of Gale-Nikaido, which are discussed for instance in the book of Parthasarathy [7]. A motivation for doing so is that one may want to restrict the domain of F to be the closed positive quadrant (since a concentration cannot be negative). In general, F will be defined on a product Ω of n intervals, open or not. However, in this generality, there exist counterexamples to the condition of Thomas (see 3.5). But, when F has at least two zeroes, we can still find lead to an explosive and unbounded increase in the concentrations of various proteins. In reality, of course, other factors always come into play, and positive feedback tends to be associated with dynamics that switch the cell from one stable condition to another.

For example, upon entering an environment that is sufficiently rich in the sugar lactose, an *Escherichia coli* bacterium will begin expressing the *lac* operon, thereby producing a handful of enzymes that allow lactose digestion. The onset of expression takes place through positive feedback. A small molecule related to lactose and known as the 'inducer' is the real stimulator of *lac* expression. Significantly, one of the enzymes that *lac* expresses, a permease, acts to pump this inducer into the cell. Hence, inside the cell, inducer tends to lead to more permease and permease to more inducer. So, if a few inducer molecules make it into the cell, which happens naturally when lactose concentration reaches a certain threshold on the cell's exterior, this kicks off a process of positive feedback that leads rapidly to a change in gene expression from one state to another.

This is a striking example of cell differentiation – a change in the state of a cell that is triggered by a historical event, in this case, by the temporary exposure to an adequate concentration of lactose. Once induced into *lac* expression, the bacterium will remain in this state even if the concentration of inducer subsequently falls below the threshold required to initiate *lac* expression in the first place. So the bacterium can exist in either of two stable states – induced or uninduced, corresponding to expression or lack of expression of the *lac* operon. The cell's condition depends on not only its genes, but its history.[1]

1 For a more detailed discussion of this example, see R. Thomas, *Laws for the Dynamics of Regulatory Networks* [Int J Dev Biol 1998;42:479–485]. As Thomas notes, this is also one of the clearest and most striking

properties of its interaction graphs (Theorem 5), as a consequence of a univalence theorem of Gale-Nikaido [6, 7]. Furthermore, it is often the case that, when defining the interaction graph of several biochemical compounds, the degradation of each of them is not taken into account. We show in Theorem 6 that this makes Thomas' rule valid in full generality.

Another refinement of the theorem of Gale-Nikaido, by Garcia and Zangwill [8], applies to the case where Ω is closed and bounded. This leads to Theorem 7 which, in turn, for general Ω, gives some information on the location of the zeroes of F (Theorem 8).

The economist P.A. Samuelson imagined a stronger univalence criterion than those above. It is not true for general F, but L.A. Campbell proved it when $\Omega = \mathbb{R}^n$ and F is *algebraic,* e.g. when each component of F is the quotient of two real polynomials. In Theorem 9 we translate his result in terms of properties of the interaction graph.

One can expect more graphical requirements for multistationarity. One of them is a conjecture of Thomas and Kaufman [9], which is stronger than the original Thomas' conjecture. We have been unable to prove this assertion, but we obtained some evidence for it in Theorems 3, 4 and 5.

The paper is organized as follows. Section 1 gives definitions about graphs, matrices and determinants. It shows basic lemmas, which are standard knowledge in the literature on interaction graphs. In Section 2 we define the interaction graphs $G(a)$ and we state the conjectures of Thomas and Kaufman-Thomas. In Section 3 we prove the conjecture of Thomas and we give some (counter)examples. In Section 4 we discuss the conjecture of Kaufman-Thomas. Next, we give results when the domain Ω is not necessarily open, and we refine them in Section 6. Finally, we discuss the case of an algebraic map in the last section.

The literature on the question studied here is scattered in several different journals, and it has been crucial for me to receive papers from various authors (see Acknowledgment).

1 Graphs and Matrices

1.1

An *interaction graph* $G = (V, E, \mathrm{sgn})$ is a finite oriented graph (V, E) together with a *sign* map $\mathrm{sgn}: E \to \{\pm 1\}$. In other words, V (the vertices) and E (the edges) are two finite sets and each edge $e \in E$ has an origin $o(e) \in V$ and an endpoint $t(e) \in V$ (it may happen that $o(e) = t(e)$).

A *circuit* in G is a sequence of edges $e_1, \ldots, e_k$ such that $o(e_{i+1}) = t(e_i)$ for all $i = 1, \ldots, k-1$ and $t(e_k) = o(e_1)$.

A *hooping* is a collection $C = \{C_1, \ldots, C_k\}$ of circuits such that, for all $i \neq j$, C_i and C_j do not have a common vertex. A circuit is thus a special case of hooping. We let

$$V(C) = \coprod_{i=1}^{k} V(C_i)$$

be the (unordered) set of vertices of C. This set $V(C)$ will also be called the *support* of C. A hooping is called *Hamiltonian* when its set of vertices is maximal, i.e. $V(C) = V$. Note that hoopings are called 'generalized circuits', or 'g-circuits' *i* [10] and Hamiltonian ones are called 'full circuits' *o* [9].

The sign of a circuit C is

$$\mathrm{sgn}(C) = \prod_{e \in C} \mathrm{sgn}(e) \in \{\pm 1\}.$$

When $\mathrm{sgn}(C) = +1$ (respectively -1) we say that C is positive (respectively negative). The sign of a hooping C is

$$\mathrm{sgn}(C) = (-1)^{p+1}, \qquad (1)$$

where p is the number of positive circuits in C [see 10].

Given any subset $I \subset V$ we let $\tau_I G$ be the interaction graph obtained from G by changing the sign of every edge $e \in E$ such that $t(e) \in I$. Given any permutation $\sigma \in \mathrm{Aut}(V)$ of the vertices, we let σG be the interaction graph obtained from G by replacing each edge $j \to i$ by an edge $j \to \sigma(i)$, with the same sign.

In this example, the existence of multiple stationary states is clearly associated with a positive feedback loop. From a more fundamental point of view, one might wonder if this is always the case. Would it be possible to have multiple stationary states within a genetic regulatory network and yet have no positive feedback loops? Or are these an absolutely necessary ingredient? In 1984, René Thomas conjectured that *E. coli* is indeed no isolated case: that multiple stationary states can exist only in the presence of at least one positive feedback loop. Over the past decade, various researchers have proven Thomas's conjecture under a series of progressively more general conditions, and here, in the present paper, Christophe Soulé offers further progress along these lines. Most biologists seem to accept that positive feedback is a frequent component of gene regulatory networks, and strongly associated with cellular multi-stability. While the present work cannot be said to prove that positive feedback is absolutely necessary for cellular differentiation, or for the existence of multiple stable states in a general sense, it does tend to suggest that biologists' intuition is probably correct.

To begin with, the paper sets out a mathematical framework for the proof with a few preliminaries.

The dynamical evolution of the chemical concentrations of a network of N compounds can be expressed in the form $dx/dt = F(x)$, where $F(x) = (f_1(x), f_2(x), \ldots f_N(x))$ is a vector-valued function of the vector $x = (x_1, x_2, \ldots x_N)$ which represents the concentrations of the various compounds. For any set of concentrations $x = a$, the function $F(x)$ gives the rate of increase of each compound, and leads in a simple way to an interaction graph

examples of epigenetic inheritance. Experimentally, a culture of genetically identical *E. coli* bacteria can be manipulated to produce two distinct sub-cultures of induced and un-induced organisms that will transmit these phenotypic differences faithfully to their progeny.

1.2

Let $n \geq 1$ be an integer and $A = (a_{ij})$ an n by n real matrix. We can attach to A an interaction graph G as follows. The set of vertices of G is $\{1, ..., n\}$. There is an edge e with $o(e) = j$ and $t(e) = i$ if and only if $a_{ij} \neq 0$. The sign of e is the sign of a_{ij}.

Given any subset $I \subset \{1, ..., n\}$, the *principal minor* of A with support I is the real number $\det(A_I)$, where A_I is the square matrix $(a_{ij})_{i,j \in I}$. By definition

$$\det(A_I) = \sum_{\sigma \in \Sigma_I} \epsilon(\sigma) \prod_{i \in I} a_{i\sigma(i)}, \qquad (2)$$

where Σ_I is the group of permutations of I and $\epsilon(\sigma)$ is the signature of σ. (Recall that ϵ is defined by the equalities $\epsilon(\sigma\sigma') = \epsilon(\sigma)\epsilon(\sigma')$ for all $\sigma, \sigma' \in \Sigma_I$ and $\epsilon(\sigma) = -1$ when $\sigma \in \Sigma_I$ is a transposition).

For any $\sigma \in \Sigma_I$ we let

$$a(\sigma) = \epsilon(\sigma) \prod_{i \in I} a_{i\sigma(i)} \qquad (3)$$

so that

$$\det(A_I) = \sum_{\sigma \in \Sigma_I} a(\sigma).$$

When $a(\sigma) \neq 0$, we let $\mathrm{sgn}(a(\sigma)) = \pm 1$ be its sign.

Let $D = \mathrm{diag}(d_1, ..., d_n)$ be a diagonal n by n real matrix and $I \subset \Sigma$ any subset. It follows from the definition (2) that

$$\det((A + D)_I) = \sum_{J \subset I} \det(A_J) \prod_{i \in I-J} d_i. \qquad (4)$$

Given $I \subset V$ any subset, we let $\tau_I A$ be the matrix obtained from A by replacing a_{ij} by $-a_{ij}$ whenever $i \in I$. Given any $\sigma \in \Sigma_n = \mathrm{Aut}(V)$ we let σA be the product of A with the permutation matrix defined by σ. Clearly

$$G(\tau_I A) = \tau_I G(A)$$

and

$$G(\sigma A) = \sigma G(A).$$

1.3

We keep the notation of the preceding paragraph. Note that, given any permutation $\sigma \in \Sigma_I$, there is a unique decomposition

$$I = I_1 \amalg I_2 \amalg ... \amalg I_k$$

of I into a disjoint union of nonempty subsets such that the restriction σ_α of σ to I_α is a *cyclic permutation* for all $\alpha = 1, ..., k$. Let $C(\sigma_\alpha)$ be the circuit of G with edges $(i, \sigma_\alpha(i))$, $i \in I_\alpha$ (note that $a_{i\sigma_\alpha(i)} \neq 0$ since $a(\sigma) \neq 0$). We denote by $C(\sigma)$ the hooping of G which is the disjoint union of the circuits $C(\sigma_\alpha)$, $\alpha = 1, ..., k$.

When X is a finite set, we let $\#(X)$ be its cardinality. The following lemma is due to Eisenfeld and DeLisi [10, Appendix, Lemma 2], and is probably at the origin of the definition (1).

Lemma 1. *Let $I \subset \{1, ..., n\}$ be any subset and let $\sigma \in \Sigma_I$ be such that $a(\sigma) \neq 0$. Then the following identity holds*

$$\mathrm{sgn}(C(\sigma)) = \mathrm{sgn}(a(\sigma))\,(-1)^{\#(I)+1}.$$

Proof. Since $\epsilon(\sigma) = \sum_{\alpha=1}^{k} \epsilon(\sigma_\alpha)$ and $I = \coprod_\alpha I_\alpha$, we get from (3) that

$$a(\sigma) = \prod_\alpha a(\sigma_\alpha). \qquad (5)$$

For each $\alpha = 1, ..., k$, since σ_α is cyclic we have

$$\epsilon(\sigma_\alpha) = (-1)^{\#(I_\alpha)+1}.$$

The circuit $C(\sigma_\alpha)$ is positive if and only if $\prod_{i \in I_\alpha} a_{i\sigma_\alpha(i)}$ is positive, in which case we get from (3) that

$$\mathrm{sgn}(a(\sigma_\alpha)) = (-1)^{\#(I_\alpha)+1}. \qquad (6)$$

When $C(\sigma_\alpha)$ is negative we get

$$\mathrm{sgn}(a(\sigma_\alpha)) = (-1)^{\#(I_\alpha)}. \qquad (7)$$

Therefore, according to (5), (6), (7), we get

$$\mathrm{sgn}(a(\sigma)) = (-1)^p\,(-1)^{\sum_\alpha \#(I_\alpha)} = (-1)^p\,(-1)^{\#(I)},$$

where p is the number of positive circuits in $C(\sigma)$. Since, according to (1), $\mathrm{sgn}(C(\sigma)) = (-1)^{p+1}$, the lemma follows.

1.4

From Lemma 1 we get the following:

Lemma 2. *Let A be an n by n real matrix and let $I \subset \{1, ..., n\}$ be any subset.*

(i) *Assume that* $\det(-A_I)$ *is negative (respectively positive). Then there exists a*

$G(a)$ – a network of nodes and links that offers a logical representation of the various interactions (see diagram 1). Each element (or vertex) in the graph represents one compound (a protein, for example). If the partial derivative $(\partial f_i/\partial x_j)(a)$ is nonzero, then compound j influences the concentration of compound i, and the graph contains a link (or edge) running from element j to element i. This defines an oriented graph, as each edge has a definite direction. The sign of the partial derivative (+ or –) defines the sign of the edge $j \rightarrow i$, and reveals whether compound j tends to increase or decrease the concentration of compound i.

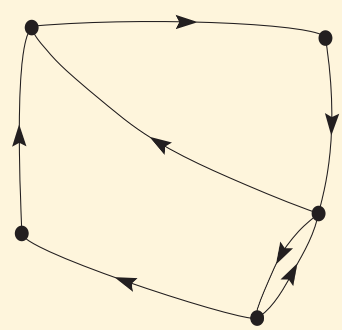

Diagram 1

The author also reviews a few technical definitions that provide convenient tools for describing a graph of this kind. A ‘circuit’ within the graph – if any exist – would be a closed path that travels along a set of oriented links. By definition, the sign of a circuit is just the product of the signs of the edges it involves. More generally, a ‘hooping’ is defined as a collection of circuits, all of which follow entirely distinct routes (with no element of the graph belonging to more than one circuit). A hooping can also be given a sign, and it would seem natural to define this as the product of the signs of the circuits that make it up, in which case the sign would depend on the number of negative circuits. Soulé, however, adopts a different definition (introduced originally by Eisenfeld and De Lisi) in which the sign of *positive (respectively negative) hooping in $G(A)$ with support I.*

(ii) *Assume $det(-A_I) = 0$. Then either there exist two hoopings in $G(A)$ with opposite signs and support I, or there is no hooping in $G(A)$ with support I.*

Proof. From (2) we get

$$\det(-A_I) = (-1)^{\#(I)}\det(A_I).$$

Therefore, according to Lemma 1,

$$\det(-A_I) = -\sum_{\sigma\in\Sigma_I} \mathrm{sgn}(C(\sigma))|a(\sigma)|. \qquad (8)$$

Assume $\det(-A_I)$ is negative (respectively positive). Then, according to (8), there exists $\sigma \in \Sigma_I$ such that $C(\sigma)$ is positive (respectively negative). This proves (i).

If $\det(-A_I) = 0$, either there exist two summands with opposite signs on the right hand side of (8), or $a(\sigma)$ is zero for all $\sigma \in \Sigma_I$. This proves (ii), since every hooping of $G(A)$ with support I is of the form $C(\sigma)$ for some $\sigma \in \Sigma_I$.

2 Conjectures on Multistability

2.1

Let Ω_i be a nonempty interval in $\mathbb{R}$:

$$\Omega_I =]a_i, b_i[, [a_i, b_i[,]a_i, b_i] \text{ or } [a_i, b_i],$$

with $a_i \geq -\infty$ and $b_i \leq +\infty$. Denote by $\Omega \subset \mathbb{R}^n$ the product $\Omega = \prod_{i=1}^{n} \Omega_i$.
Consider a map

$$F: \Omega \to \mathbb{R}^n$$

which is *differentiable*, i.e. such that, for each $i,j \in \{1, ..., n\}$ and any $a \in \Omega$, the i-th component f_i of F has a partial derivative $\partial f_i/\partial x_i\,(a)$ at the point a and

$$f_i(x) = f_i(a) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(a)(x_j - a_j) + o(||x - a||),$$

where $||x - a||$ is the norm of $x - a$ and o is the Landau o-symbol (we could also use a weaker notion of differentiability, see [11]). The map F is called C^1 when every partial derivative $\partial f_i/\partial x_j$ is continuous on Ω.

For any $a \in \Omega$, the *Jacobian* of F at a is the n by n real matrix

$$J(a) = J(F)(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right).$$

For any $a \in \Omega$ we let

$$G(a) = G(J(a))$$

be the interaction graph of $J(a)$, defined as in 1.2. We also let $G(F)$ be the interaction graph defined as follows. Its set of vertices is $V = \{1, ..., n\}$. Given i and j in V, there is at most one positive (respectively negative) edge from j to i; it exists if and only if there is a positive (respectively negative) path in $G(a)$ for some $a \in \Omega$. In other words, $G(F)$ is the 'superposition' of all the interaction graphs $G(a)$.

Given any subset $I \subset V$, we let $\tau_I F$ be the map obtained from $F = (f_i)$ by replacing f_i by $-f_i$ when $i \in I$. Given $\sigma \in \Sigma_n$, we let $\sigma F = (f_{\sigma^{-1}(i)})$. The interaction graphs of $\sigma\tau_I F$ are $\sigma\tau_I\, G(a)$, $a \in \Omega$, and $\sigma\tau_I\, G(F)$.

2.2

We shall be interested in the set of zeroes of F, i.e. the points $a \in \Omega$ such that $F(a) = 0$. They can be viewed as the *stationary states* of the system of differential equations

$$\frac{dx(t)}{dt} = F(x),$$

where $x(t)$ is a differentiable mapping from a real interval to Ω.

We say that a zero a of F is *nondegenerate* when $\det(J(a)) \neq 0$.

2.3 Conjecture 1 (Thomas [1])

Assume that Ω is open and that F has at least two nondegenerate zeroes in Ω. Then there exists $a \in \Omega$ such that $G(a)$ contains a *positive* circuit.

Remark. This conjecture has already been proved in several cases. First, it is known to hold when the signs of the entries in $J(a)$ are independent of $a \in \Omega$ [3, 4, see also 5] (also the remark in 5.2 below). It was also shown [in 2] for stable stationary states when Ω contains the positive quadrant and $f_i(x) > 0$ whenever $x_i = 0$.

a hooping depends on the number of positive circuits. Specifically, the sign is positive if the hooping contains an odd number of positive circuits, and negative otherwise. This particular definition, peculiar though it may seem, turns out to be fruitful.

In terms of this graph, the Thomas conjecture claims that the existence of multiple stationary states for the network requires the existence of at least one positive circuit. In terms of the definitions given above, Soulé's basic approach to the Thomas conjecture follows a straightforward though at times somewhat delicate logic.

The author begins by interpreting 'multiple stationary states' to mean that there must be at least two values of x for which the function $dx/dt = F(x) = 0$. These are fixed points, representing unchanging states of the dynamics. The existence of at least two such points implies that the function $F(x)$ cannot represent a one-to-one mapping, for it maps at least two points into the same value. In the context of this interpretation, Soulé's strategy is to explore what properties of the graph $G(a)$ can be inferred from the statement that $F(x)$ is not one-to-one. This can be done by identifying conditions C that would guarantee that $F(x)$ is one-to-one, for the failure of $F(x)$ to be one-to-one would then imply that conditions C do not hold (or that the negation of C is true). Hence, a starting point for the investigation is the identification of suitable conditions C, which Soulé takes from an earlier mathematical result of Gale and Nikaido.

As these authors established, a function $F(x)$ is guaranteed to be one-to-one if its Jacobian matrix $J(x)$, defined as $J_{i,j}(x) = (\partial f_i/\partial x_j)(x)$, satisfies certain properties. A 'principal minor' of a matrix A is defined to be the determinant of the sub-matrix that is obtained from A by restricting the indices to some subset I. (For $n = 5$, for example, one possible subset would be $I = \{1,3,5\}$). Gale and Nikaido established that

2.4 Conjecture 2 (Kaufman-Thomas [9])

Under the same assumption as Conjecture 1,

(i) either there exist $a \in \Omega$ such that $G(a)$ has a positive Hamiltonian hooping and a negative Hamiltonian hooping,

(ii) or there is a cyclic permutation σ of a subset of $\{1, ..., n\}$ and there exist $a, b \in \Omega$ such that the circuit $C(\sigma)$ in $G(a)$ (respectively in $G(b)$) is positive (respectively negative).

Note that Conjecture 2 for F implies Conjecture 1 for F. It is quite different though, since positive and negative circuits play a symmetric role in it. In case (i) we shall say that 'G has two Hamiltonian hoopings of opposite signs' and in case (ii) we shall say that 'G has an ambiguity'. In the latter case, $G(F)$ contains two circuits with the same ordered set of vertices and opposite signs [but (ii) is stronger than that statement].

3 A proof of Thomas' Conjecture 1

3.1

Theorem 1. *Assume Ω is open and F has at least two nondegenerate zeroes in Ω. Then, for every $I \subset V$ and every $\sigma \in \Sigma_n$, there exists $a \in \Omega$ such that $\sigma\tau_I\, G(a)$ has a positive circuit. In particular, Thomas' Conjecture 1 holds true.*

Since $\sigma\tau_I\, G(x) = G(\sigma\tau_I\, J(x))$ and since $\sigma\tau_I(F)$ satisfies the hypotheses of Theorem 1 if and only if $-F$ does, we just have to check that, for some $a \in \Omega$, the interaction graph $G(-J(a))$ has a positive circuit. According to Lemma 2 (i), it will be enough to show that, for some $a \in \Omega$, a principal minor of $-J(a)$ is negative. In other words:

Theorem 1′. *Assume that, for every $\sigma \in \Omega$, all the principal minors of $-J(a)$ are nonnegative. Then $-F$ can have at most one nondegenerate zero.*

3.2

For any positive real number λ let

$$\phi_\lambda(x) = -F(x) + \lambda x.$$

Since

$$J(\phi_\lambda) = -J(F) + \mathrm{diag}(\lambda, ..., \lambda),$$

it follows from (4) that, under the hypotheses of Theorem 1′, each principal minor of $J(\phi_\lambda)$ is positive on Ω. According to Gale and Nikaido [6, Theorem 4], this implies that ϕ_λ is univalent. The following proposition ends the proof of Theorem 1′:

Proposition 1. *Let $\phi = \Omega \longrightarrow \mathbb{R}^n$ be a differentiable map defined on an open set $\Omega \subset \mathbb{R}^n$. Assume that, for all $\lambda > 0$, the map*

$$\phi_\lambda(x) = \phi(x) + \lambda x$$

is univalent. Then ϕ can have at most one nondegenerate zero in Ω.

3.3

The proof of proposition 1 proceeds as Theorem 4′ [6] and IV, Theorem 4 [7, p. 35]. Assume a and b are two nondegenerate zeroes of ϕ in Ω. According to Alexandroff and Hopf [12, XII, § 2.9, p. 477], we can choose open neighborhoods U_a and U_b of a and b, respectively, such that $\bar{U}_a \cap \bar{U}_b = \varnothing, \bar{U}_a \subset \Omega, \bar{U}_b \subset \Omega$, a (respectively b) is the unique zero of ϕ in $\bar{U}_a$ (respectively $\bar{U}_b$), and the degrees $\deg(\phi, \bar{U}_a, a)$ and $\deg(\phi, U_b, 0)$ are equal to ± 1. Arguing as in reference 7 [or 6, loc. cit.], we get

$$1 = \deg(\phi, \bar{U}_a \cup \bar{U}_b, 0) = \deg(\phi, \bar{U}_a, 0) + \deg(\phi, \bar{U}_b, 0),$$

hence a contradiction. q.e.d.

3.4

The restriction to nondegenerate zeroes of F in Theorem 1 is necessary. For example, if $n = 2, \Omega = \mathbb{R}^n$ and

$$F(x, y) = (-xy^2, -y)$$

we have

$$-JF(x, y) = \begin{pmatrix} y^2 & 2xy \\ 0 & 1 \end{pmatrix}.$$

a function $F(x)$ will be one-to-one if all of the principal minors of $J(x)$ are non-negative, and if the determinant of $J(x)$ (the principal minor when I is the entire set of indices) is positive for all x. These are the conditions C.

Since the principal minors play an important role in this theorem, Soulé's next step is to express their values in a convenient fashion. A permutation σ represents a re-ordering of a set of indices. (In the case $n = 3$, for example, one permutation would represent the mapping $(1, 2, 3) \longrightarrow (2, 1, 3)$, while another would represent $(1, 2, 3) \longrightarrow 3, 1, 2)$). As Soulé notes in section 1.2, any principal minor is by definition a sum over all the permutations of the various indices contained in a subset I. Consider the full determinant, for example, when I is the entire set of n indices. In this case, each term in the sum is the quantity

$$a(\sigma) = \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} ... a_{n,\sigma(n)},$$

where $\sigma(i)$ gives the index into which the permutation maps i, and $\epsilon(\sigma)$ gives the sign with which this term appears in the sum, either $+1$ or -1. (As a technical point, $\epsilon(\sigma)$ is defined to be 1 when σ is the identity permutation, which represents no re-arrangement of the indices at all. The signs of all other terms in a principal minor follow from the fact that any permutation can be realized as a sequence of 'transpositions' – permutations that simply switch two indices while leaving all others unchanged. By definition, the sign $\epsilon(\sigma)$ is positive if, starting from the identity, the permutation σ can be achieved by an even number of transpositions, and is negative if it requires an odd number of transpositions.)

These are standard formalities in linear algebra. To this point, the paper has simply noted that the principal minors of the Jacobian play a central role in the conditions C (sufficient for $F(x)$ to be one-to-one), and has expressed these principal minors as sums over permutations. The

Clearly G has no positive circuit. However, $F(x, 0) = 0$ for any $x \in \mathbb{R}$.

3.5

It is also essential that Ω is open. Let Ω be the set of $(x, y) \in \mathbb{R}^2$ such that $x \geq 0$ and $y \geq 0$. Consider the map $F: \Omega \longrightarrow \mathbb{R}^2$ defined by

$$F(x,y) = ((y-2)^2 - x^2 - 1, 4x - 2xy).$$

We get

$$-JF(x,y) = \begin{pmatrix} 2x & 4-2y \\ 2y-4 & 2x \end{pmatrix}.$$

Therefore, according to (8), $G(F)$ does not have any positive circuit. On the other hand, both $(0, 3)$ and $(0, 1)$ are nondegenerate zeroes of F in Ω. We shall discuss the case of an arbitrary Ω in Sections 5 and 6 below.

4 On the Conjecture of Kaufman-Thomas

4.1

We first assume that $n = 2$. Let $\Omega \subset \mathbb{R}^2$ be as in 2.1 and open, with coordinates x and y. For any $h: \Omega \longrightarrow \mathbb{R}$ we write $h \equiv 0$ to mean that $h(a) = 0$ for every $a \in \Omega$. If h is differentiable, we let h_x (respectively h_y) be its partial derivative with respect to the first (respectively second) variable.

Let

$$F = (f, g): \Omega \longrightarrow \mathbb{R}^2$$

be a differentiable mapping.

Theorem 2. *Assume that F has at least two nondegenerate zeroes in Ω. Then, one of the following conditions holds:*

(i) *G has two Hamiltonian hoopings of opposite signs,*

(ii) *G has an ambiguity, and*

(iii) *$f_x g_y \equiv 0$ but $f_x \not\equiv 0$ and $g_y \not\equiv 0$.*

Proof. Assume that G has no ambiguity, and that its Hamiltonian hoopings have all the same sign. If, in addition, $f_x g_y \not\equiv 0$, since the sign of $f_x(a)$ (respectively $g_y(a)$) is independent of a, we can multiply f and g by ± 1 to get to the case where $f_x(a) \geq 0$ and $g_y(a) \geq 0$ for all $a \in \Omega$. It follows that $f_x(a) g_y(a) \geq 0$ and, since all Hamiltonian hoopings have the same sign, $f_y(a) g_x(a) \leq 0$ (by Lemma 1). Therefore, Theorem 1′ applies to F and we conclude that F has at most one nondegenerate zero.

Assume now that $g_y \equiv 0$. After multiplying f and g by ± 1 we can assume that $f_x \geq 0$ and $f_y g_x \leq 0$. Once again, Theorem 1' implies that F has at most one nondegenerate zero. q.e.d.

4.2

With an additional assumption, Conjecture 2 is true for all $n \geq 2$:

Theorem 3. *Let $F: \Omega \longrightarrow \mathbb{R}^n$ be a differentiable mapping such that Ω is open and F has at least two nondegenerate zeroes. Then, one of the following conditions holds:*

(i) *G has two Hamiltonian hoopings with opposite signs,*

(ii) *G has an ambiguity,*

(iii) *for any point $a \in \Omega$ there exists $i \in \{1, ..., n\}$ such that $G(a)$ does not contain an edge from i to itself.*

Proof. Assume that all Hamiltonian hoopings in G have the same sign, that G has no ambiguity and that, for some $a \in \Omega$, and for any $i \in \{1, ..., n\}$, there is an edge in $G(a)$ from i to itself. The last condition means that all diagonal entries in $J(a)$ are nonzero. After multiplying each component of F by ± 1 we can assume that all diagonal entries of $J(a)$ are positive. Let C be any hooping in $J(a)$, and let I be its set of vertices. The disjoint union of C with all the circuits $i \longrightarrow i$, $i \in \{1, ..., n\} - I$, is a Hamiltonian hooping of $G(a)$. Its sign must be the sign of the Hamiltonian hooping which is the disjoint union of all the positive circuits $i \longrightarrow i, i \in \{1, ..., n\}$, namely $(-1)^{n+1}$. We conclude that

$$\mathrm{sgn}(C) = (-1)^{\#(I)}.$$

According to Lemma 1, this implies that $C = C(\sigma)$ with $\mathrm{sgn}(a(\sigma)) = +1$, $\sigma \in \Sigma_I$. Since G has no ambiguity, for any $\sigma \in \Sigma_I$ we have $\mathrm{sgn}(a(\sigma)) \geq 0$ in Ω. Therefore, for any $x \in \Omega$, all the principal minors of $J(x)$ are nonnegative. Applying Theorem 1′, we key to approaching the Thomas conjecture lies next in establishing an interesting link between these principal minors and the circuits of the interaction graph $G(x)$.

Soulé establishes this important connection in Section 1.3. As he points out, while each term in a principal minor corresponds to a permutation σ, each such permutation also defines a natural 'hooping' of the graph $G(x)$ – that is, a set of disjoint circuits within the graph. To see this, note that any permutation produces a decomposition of the network into a set of disjoint 'sectors'. For example, consider for $n = 5$ the permutation $(1, 2, 3, 4, 5) \longrightarrow (3, 1, 2, 5, 4)$. A visualization of this mapping reveals that the permutation acts independently within two sectors, one comprising elements 1, 2 and 3 and the other elements 4 and 5 (see diagram 2). The permutation produces a natural circuit within each sector, and these circuits together make up a hooping. By considering how the permutations relate to both the principal minors of the Jacobian matrix and to the graph $G(x)$, it becomes clear that each non-zero term in any principal minor corresponds uniquely to some hooping of the graph.

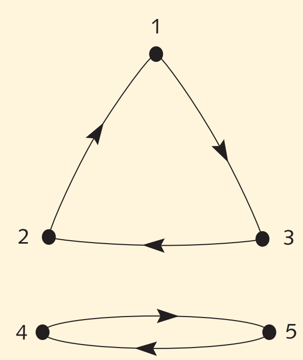

Diagram 2

This elegant connection is useful because it links properties of the principal minors to properties of the interaction graph. Specifically, the sign of each term $a(\sigma)$ in any principal minor of $J(x)$ can be linked directly to the sign of the corresponding hooping $C(\sigma)$, and hence, to the conclude that F has at most one nondegenerate zero.

4.3

Concerning the graph $G(F)$ (see 2.1) we can prove the following:

Theorem 4. *Let $F: \Omega \longrightarrow \mathbb{R}^n$ be a differentiable mapping such that Ω is open and F has at least two nondegenerate zeroes. Then $G(F)$ has two Hamiltonian hoopings with opposite signs.*

Proof. We first remark that we can find $a \in \Omega$ and $\sigma \in \Sigma_n$ such that none of the diagonal entries in $J(\sigma F)(a)$ is zero. Indeed these entries are, by definition, $\frac{\partial f_{\sigma^{-1}(j)}}{\partial x_j}(a)$, $j = 1, ..., n$, and, when a is a nondegenerate zero of F, it follows from (1) that, for some $\sigma \in \Sigma_n$,

$$\prod_{i=1}^{n} \frac{\partial f_i}{\partial x_{\sigma(i)}}(a), (a) \neq 0.$$

This proves the claim.

We may then choose I such that all the diagonal entries in $J(\tau_I \sigma F)(a)$ are negative. In other words, for each $i \in \{1, ..., n\}$, the interaction graph $\tau_I \sigma G(a)$ contains a negative edge $i \longrightarrow i$. On the other hand, according to Theorem 1, there exists $b \in \Omega$ such that $\tau_I \sigma G(b)$ has a positive circuit C. In $\tau_I \sigma G(F)$ the disjoint union of C with the negative edges $i \longrightarrow i$, $i \notin V(C)$, is a positive Hamiltonian hooping, when the union of all the negative edges $i \longrightarrow i$, $i \in \{1, ..., n\}$, is a negative one.

It is thus enough to show that if $\tau_I G(F)$ or $\sigma G(F)$ has two Hamiltonian hoopings with opposite signs, the same is true for $G(F)$. This is clear for $\tau_I G(F)$ and for $\sigma G(F)$, it follows from Lemma 3 below.

4.4.

Lemma 3. *Let G be any interaction graph, $\sigma \in \mathrm{Aut}(V)$ a permutation of its vertices, and C a Hamiltonian hooping in G. The image of C in the interaction graph σG is then a Hamiltonian hooping with sign $\epsilon(\sigma)\mathrm{sgn}(G)$.*

Proof. Recall from 1.1 that σG is obtained from G by replacing each edge $j \longrightarrow i$ by an edge $j \longrightarrow \sigma(i)$, with the same sign. As a collection of edges, C has a well-defined image σC in σG. To check Lemma 3 we may assume that σ is the transposition of two vertices i and j (transpositions span $\mathrm{Aut}(V)$).

Assume first that i and j are in the same circuit C_1 of C. Then all the circuits in C other than C_1 are fixed by σ. The image of C_1 consists of two disjoint circuits D_1 and D_2. More precisely, if the vertices of C_1 are $1\ 2 \ldots i \ldots j \ldots m$ (as we can assume), we get

$$\sigma C_1 = D_1 \coprod D_2$$

where the sequence of vertices in D_1 (respectively D_2) is $1\ 2 \ldots i-1\ j\ j+1 \ldots k1$ (respectively $i\ i+1 \ldots j-1i$). Furthermore, if C_1 has an even (respectively odd) number of negative edges, D_1 and D_2 will have the same (respectively a different) number of edges modulo two. From this it follows that σC is a Hamiltonian hooping such that with the definition (1),

$$\mathrm{sgn}(\sigma C) = -\mathrm{sgn}(C),$$

as was to be shown.

Note that we also have

$$C_1 = \sigma D_1 \coprod \sigma D_2,$$

therefore, by exchanging the roles of C and σC, the previous discussion applies also to the case where i and j lie in two different circuits of C. q.e.d.

5 The Case of a Domain Which Is Not Open

5.1

We keep the notation of 2.1, where Ω is an arbitrary product of intervals and $F: \Omega \longrightarrow \mathbb{R}^n$ is differentiable.

Theorem 5. *Assume that F is not univalent. Then,*

(1) *For every $\sigma \in \Sigma_n$,*

(i) *either there exists $a \in \Omega$ and $i \in V$ such that $\sigma G(a)$ does not contain any edge from i to itself;*

(ii) *or, for any subset $I \subset V$, there exists $a \in \Omega$ such that $\tau_I \sigma G(a)$ has a positive circuit.*

signs of the various circuits involved in it. In his Lemma 1, Soulé establishes by straightforward calculation that the relationship between these signs is, in the general case,

$$\mathrm{sgn}(C(\sigma)) = \mathrm{sgn}(a(\sigma))(-1)^{\#(I)+1},$$

where $\#(I)$ is the number of indices in the subset I of indices. (The simplicity of this relationship finally explains the apparently odd definition of the sign of a hooping given in equation (1)).

This result leads immediately to the important result of Lemma 2. For any matrix A, consider the principal minor of $-A$ relative to some set of indices I. Using the formula produced in Lemma 1, Soulé arrives at his equation (8) for this principal minor. It is then straightforward to deduce that if $\det(-A_I) < 0$, at least one term within the summation on the right hand side has to be positive. Hence, the sign of at least one hooping has to be positive, which can only be the case if the graph contains a positive circuit. Likewise, if $\det(-A_I) > 0$, then at least one term in the sum has to be negative and there must be a negative hooping. This establishes the first part of Lemma 2.

Next, if $\det(-A_I) = 0$, there are two possibilities. It might be that $a(\sigma) = 0$ for all terms on the right-hand side of equation (8), in which case there simply is no hooping (the graph does not present any circuits whatsoever). Alternatively, if any term on the right side is non-zero, there must also be another term of opposite sign, for the sum could not otherwise turn out to be zero. There would then be a pair of hoopings of opposite sign. This establishes the second part of Lemma 2.

These results take the paper to the brink of the Thomas conjecture. In accordance with Soulé's interpretation of the term 'multiple stationary states', Section 2.3 of the paper states this conjecture in a particular and precise form: 'Assume that Ω is an open set and that the function $F(x)$ has two non-degenerate zeros within Ω. Then for

(2) *When the condition* (1(i)) *above is not satisfied,* $G(F)$ *contains both a positive and a negative Hamiltonian hooping.*

Proof. The map F is univalent if and only if $\tau_I\sigma F$ is. So, to prove (1), we can restrict our attention to $-F$. Note also that (i) is equivalent to the assertion that there exists $I \subset V$ such that $\sigma G(a)$ does not contain any hooping with support I. According to Lemma 2, if none of the conclusions in (1) is true, all the principal minors of the Jacobian matrix of $-F$ are positive on Ω. By the Gale-Nikaido theorem [6], Theorem 4, this implies that $-F$ is univalent. This proves (1).

To prove (2), by replacing F by $\tau_I F$ for an appropriate choice of I, we can assume that, for each vertex i in $G(F)$, there is a negative edge from i to itself. Since $G(F)$ also contains a positive circuit by (1), we get, as in 4.3 above, that $G(F)$ contains both a positive and a negative Hamiltonian hooping.

Remark. Since the condition 1(i) is often satisfied, when Ω is open Theorem 5 is much weaker than Theorems 1 and 4.

5.2

Here is a variant of Theorem 5.

Theorem 6. *Assume given, for every* $a \in \Omega$, *a diagonal matrix* $D(a)$ *with positive entries. Assume* F *is not univalent. Then, for some* $a \in \Omega$, *the interaction graph*

$$H(a) = G(J(a) + D(a))$$

has a positive circuit.

Proof. Assume that, for any $a \in \Omega$, none of the circuits of $H(a)$ is positive. We know from Lemma 2(i) that all the principal minors of $-J(a) - D(a)$ are nonnegative. From (4) this implies that all the principal minors of $-J(a)$ are positive, and, again according to reference 6, Theorem 4, F must be univalent.

Remark. Theorem 6 applies to the situation considered for instance in references 13 and 14, where $-D(a)$ comes from 'terms of decay', which are not taken into account when drawing the interaction graph. It was proven by Snoussi [5] when the signs of the entries of $J(a)$ are constant.

6 On the Location of Stationary States

6.1

Assume that $\Omega = \prod_{i=1}^{n} [a_i, b_i]$ is a closed bounded subset of $\mathbb{R}^n$. In this case, Garcia and Zangwill [8] got a stronger result than Gale and Nikaido [6] [see also 7, V, Theorem 1, p. 41]. For any $I \subset V$ define $I^c = V - I$ and, for any n by n real matrix A, let $m_I(A) = \det(A_{I^c})$ and $m_i(A) = m_{\{i\}}(A)$ for each $i \in V$. When x lies in Ω, we write $m_I(x)$ for $m_I(J(x))$ and $m_i(x) = m_i(J(x))$. Define $I(x) \subset V$ as the set of vertices i such that $x_i = a_i$ or $x_i = b_i$. The result of Garcia and Zangwill is the following:

Theorem 7′. *Assume* F *is* C^1 *and that, for every* $a \in \Omega$, *and every subset* $I \subset I(a)$,

$$m_I(a) \prod_{i \in I} m_i(a) > 0.$$

Then F *is univalent.*

In particular, when $a \in \mathring{\Omega} = \prod_{i \in I}^{n}]a_i, b_i[$, $I(a) = \varnothing$ and the only assertion made is that $\det J(a) > 0$ [when F is only differentiable, see 11, p. 930, Remark].

6.2

Theorem 7′ implies the following refinement of Theorem 5:

Theorem 7. *Assume that* Ω *is bounded and closed, and that* $F : \Omega \longrightarrow \mathbb{R}^n$ *is* C^1. *If* F *is not univalent, for every* $\sigma \in \Sigma_n$, *one of the following conditions holds true:*

(i) *There exists* $a \in \Omega$ *and* $I \subset I(a)$ *such that no hooping of* $\sigma G(a)$ *has support* I^c, *or* $\sigma G(a)$ *contains two hoopings with support* I^c *and opposite signs.*

(ii) *There exists* $a \in \partial\Omega = \Omega - \mathring{\Omega}$, $I \subset I(a)$ *and hoopings* C_I *and* C_i, *for each* $i \in I$, *in* $\sigma G(a)$ *such that* $\#(I) \geq 2$, *the support of* C_I (*respectively* C_i) *is* I^c (*respectively* $\{i\}^c$), *and*

$$\mathrm{sgn}(C_I) \prod_{i \in I} \mathrm{sgn}(C_i) > 0.$$

Furthermore, either $\#(I)$ *is even or there exist* $b \in \partial\Omega$, $I' \subset I(b)$, $C_{I'}$, C_j, $j \in I'$, *with* some value a belonging to Ω, the graph $G(a)$ has a positive circuit.' In section 3, Soulé proceeds to prove this result, which now only requires a minor extension of the theorem of Gale and Nikaido.

Recall that this theorem establishes that a function $F(x)$ is one-to-one over a set Ω if all the principal minors of the Jacobian $J(x)$ are non-negative and the determinant of $J(x)$ is positive everywhere in Ω. These are the conditions C referred to earlier. It would be easy to prove the Thomas conjecture if the word 'positive' in this statement could be changed to 'non-negative'; that is, if it were possible to prove an analog of the Gale-Nikaido theorem that would involve slightly less restrictive conditions on the determinant (specifically, which would allow it sometimes to be zero). We might refer to these as conditions C'. The author tackles this task through his Theorem 1′, which extends the Gale-Nikaido theorem in this way.

For convenience, the author here considers the function $-F(x)$ rather than $F(x)$, but this is of no consequence. This theorem establishes that if the principal minors of $-J(x)$ are non-negative at all points within Ω (conditions C'), then $-F(x)$ has at most one non-degenerate zero. It is important to note that in generalizing the Gale-Nikaido result to hold under the slightly broader set of conditions C', Soulé obtains a theorem that ultimately makes a slightly weaker claim. Whereas Gale-Nikaido established a function as being one-to-one, Soulé in Theorem 1′ only establishes that the function will have at most one non-degenerate zero. Strictly speaking, the function is not guaranteed to be one-to-one, for it may have degenerate zeros – that is, the function may be zero over some continuous set of points.

This is adequate, however, to prove the Thomas conjecture, which follows immediately from this theorem in conjunction with Lemma 2. For, if a function $F(x)$ has two non-degenerate zeros within an open set Ω (so that the regulatory network is multi-

similar properties as above and

$$\operatorname{sgn}(C_{I'})\prod_{j\in I'}\operatorname{sgn}(C_j)<0.$$

Proof. Again, it is enough to treat the case $\sigma = 1$. When F is not univalent we know from Theorem 7′ that there is $a \in \Omega$ and $I \subset I(a)$ such that

$$m_I(a)\prod_{i\in I} m_i(a)\le 0.$$

Assume $m_I(a) = 0$. Then, by Lemma 2(ii), the statement (i) must hold. Assume now that for all $J \subset I(a)$, $m_J(a) \neq 0$. Since

$$m_I(a)\prod_{i\in I} m_i(a)<0$$

we must have $\#(I) \geq 2$. Furthermore, for every vertex $k \in V$, if we replace F by $\tau_k F = \tau_{\{k\}}F$ the quantity $m_I(a)$ is multiplied by +1 (respectively −1) if $k \in I$ (respectively $k \in I^c$). Therefore, in all cases, $m_I(a) \prod_{i\in I} m_i(a)$ gets multiplied by $(-1)^{\#(I)}$. It is thus invariant if and only if $\#(I)$ is even. If this number is odd, we apply the same discussion to $-F$ and we get b and I' such that

$$m_{I'}(b)\prod_{j\in I'} m_j(b)>0.$$

Using (8), the statement (ii) follows.

6.3

Let $\Omega = \prod_{i=1}^{n} \Omega_i$ be an arbitrary product of intervals as in 2.1 and let $F: \Omega \longrightarrow \mathbb{R}^n$ be a C^1 map. From Theorem 7 one gets some information on where two zeroes of F can be:

Theorem 8. *Fix $\sigma \in \Sigma_n$. Assume that there exist two points a and b in Ω such that $F(a) = F(b)$. Assume furthermore that, when x lies in $\mathring{\Omega}$, all hoopings of $\sigma G(x)$ are negative. Then there exists $x \in \Omega$ such that, when $x_i \in \partial\Omega_i$, the i-th coordinate of a or b is equal to x_i, and a subset $I \subset I(x)$ such that no hooping in $\sigma G(x)$ has support equal to I^c.*

Proof. One can find a bounded closed product of intervals $\Omega' \subset \Omega$ containing a and b and such that, whenever $x \in \Omega'$ and $x_i \in \partial\Omega_i$, the i-th coordinate of a or b is equal to x_i. Since the restriction of F to Ω' is not univalent, we can apply Theorem 7. We are not in case (ii) because, $J(x)$ being continuous in x, for every $x \in \Omega'$ all the hoopings of $G(x)$ are nonpositive.

Therefore, Theorem 7(i) holds true for some $x \in \Omega'$, hence the conclusion. q.e.d.

7 The Algebraic Case

Assume now that $\Omega = \mathbb{R}^n$ and that F is C^1 and *algebraic*, by which we mean that its graph $\{(x, F(x))\} \subset \mathbb{R}^n \times \mathbb{R}^n$ is the set of zeroes of a family of real polynomials in $2n$ variables. This will be the case for instance when each component f_i of F is the quotient of two polynomials in n variables. In that case, a result of Campbell [15] leads to a stronger conclusion than Theorem 5(1).

Theorem 9. *Choose any ordering of V. Let $F: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a C^1 map which is not univalent. Then we can choose $k \leq n$ such that if I consists of the first k vertices,*

(i) *either G(F) has two hoopings of opposite signs and support equal to I,*

(ii) *or there exists $a \in \mathbb{R}^n$ such that none of the hoopings of $G(a)$ has support equal to I.*

Proof. For any $k \leq n$ and $a \in \mathbb{R}^n$ we let

$$d_k(a) = \det(J(a)_I).$$

According to reference 15, if $d_k(a) > 0$ for all $k = 1, ..., n$, the map F is univalent. Therefore, under our assumption, $d_k(a) \leq 0$ for some a and some k. When $d_k(a) = 0$, either (i) or (ii) is true (by Lemma 2(ii)). Given $i \in V$, when we replace F by $\tau_i F$, $d_k(a)$ gets multiplied by −1 (respectively +1) if $i \leq k$ (respectively $i > k$). So we can assume that, for every $I \subset V$, there exist k and a such that $\tau_I d_k(a) < 0$ (with obvious notation). By replacing F by $\tau_1 F$ we see that we can also assume that there exist k' and b such that $\tau_I d_{k'}(b) > 0$.

Unless (i) or(ii) holds, we can assume (using Lemma 2 again) that for every $I \subset V$ and every $k \leq n$ the sign of $\tau_I d_k(x)$ does not depend on $x \in \mathbb{R}^n$. Let then m be the maximum of all integers p such that if $i \leq k \leq p$, $d_k(x)$ and $d_1(x)$ have the same sign. When $m < n$, we can increase m to stable), then by Theorem 1′, the negation of the conditions C' must be true: the principal minors of $-J(x)$ cannot be non-negative at all points within Ω. Instead, at least one of these must be negative at some point. But, by Lemma 2, if a principal minor of $-J(x)$ is negative at some point a, there must exist a positive hooping in the graph $G(a)$, and therefore a positive circuit within the interaction graph. Within the context of the mathematical assumptions here proposed, the Thomas conjecture is true.

This is the central result of the paper, and it is important to note that the conclusion depends crucially on the assumptions that the set Ω over which the function $F(x)$ is defined is open. As Soulé demonstrates in section 3.5, it is easy to find counterexamples if Ω is a closed set, such as the positive quadrant $x \geq 0, y \geq 0$ in the case $n = 2$. The reasoning that led to the theorem collapses in this case because the zeros can occur on the boundaries of Ω. In the remainder of the paper, Soulé goes on to explore a generalization of the Thomas conjecture, as well as further results that can be obtained under different assumptions about the nature of the set Ω.

All of these results represent important, if limited, steps toward exploring the dynamical requirements within regulatory networks that would allow multi-stability, and differentiation, to occur. It is important to note, however, that these results are achieved within the context of a particular interpretation of the notion of 'multiple stationary states'. These are generally interpreted in this paper to be fixed points of the dynamical system $dx/dt = F(x)$; that is, points for which $F(x) = 0$. A more general (and more realistic) conception would interpret them as stable attractors, which would include fixed points, but also oscillatory limit cycles as well as chaotic attracting states. This limitation takes nothing away from the general proofs here achieved, but underscores the difficulties yet to be faced in exploring in full generality the dynamics of genetic regulation.

Mark Buchanan

$m+1$ by replacing F by $\tau_{m+1}F$. By repeating this process, we find I such that $\tau_I d_k(x)$ has a fixed sign for every $k \leq n$ and every $x \in \mathbb{R}^n$. As we saw in the previous paragraph, this cannot happen. q.e.d.

Acknowledgment

I would like to thank J. Aracena, O. Cinquin, J. Demongeot, J.L. Gouzé, M.S. Gowda, M. Kaufman, D. Thiéfry and R. Thomas. I am also grateful to the participants of the 'Séminaire d'initiation à la génomique fonctionnelle' of IHÉS, and especially to F. Képès, who explained the importance of genetic networks, and first mentioned the rule of Thomas to me. Finally, I am extremely thankful to M. Kaufman and R. Thomas for several discussions, where they patiently and carefully explained their ideas, provided examples and encouraged me to do this work.

References

1 Thomas R: On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations. Springer Ser Synergetics 1981;9:180–193.

2 Cinquin O, Demongeot J: Positive and negative feedback: Striking a balance between necessary antagonists. J Theor Biol 2002;216:229–241.

3 Gouzé J-L: Positive and negative circuits in dynamical systems. J Biol Syst 1998;6:11–15.

4 Plahte E, Mestl T, Omholt WS: Feedback circuits, stability and multistationarity in dynamical systems. J Biol Syst 1995;3:409–413.

5 Snoussi EH: Necessary conditions for multistationarity and stable periodicity. J Biol Syst 1998;6:3–9.

6 Gale D, Nikaido H: The Jacobian matrix and global univalence of mappings. Math Ann 1965;159:81–93.

7 Parthasarathy T: On Global Univalence Theorems, Lecture Notes in Mathematics. Berlin, Springer, 1983, p 977.

8 Garcia CB, Zangwill WI: On univalence and P-matrices. Linear Algebra Appl 1979;24:239–250.

9 Thomas R, Kaufman M: Multistationarity, the basis of cell differentiation and memory. I. Structural conditions of multistationarity and other nontrivial behaviour. Chaos 2001;11:170–179.

10 Eisenfeld J, DeLisi C: On conditions for qualitative instability of regulatory circuits with application to immunological control loops; in Eisenfeld J, DeLisi C (eds): Mathematics and Computers in Biomedical Applications. Amsterdam, Elsevier, 1985, pp 39–53.

11 Gowda MS, Ravindran G: Algebraic univalence theorems for nonsmooth functions. J Math Anal Appl 2000;252:917–935.

12 Alexandroff P, Hopf H: Topologie, vol 45: Berichtigter Reprint, Die Grundlehren der mathematischen Wissenschaften. Berlin, Springer, 1974, p 636.

13 Thomas R: Logical description, analysis, and feedback loops; in Nicolis G (ed): Aspects of Chemical Evolution. 17th Solvay Conference on Chemistry. Chichester, Wiley, 1980, pp 247–282.

14 Kaufman M, Thomas R: Model analysis of the bases of multistationarity in the humoral immune response. J Theor Biol 1987;129:141–162.

15 Campbell L.A: Rational Samuelson maps are univalent. J Pure Appl Algebra 1994;92/3:227–240.