Communications in Information and Systems
Volume 20, Number 3, 253–281, 2020

Extrema without convexity and stability without Lyapunov

Brian D. O. Anderson* and Mengbin Ye†,‡

The great majority of optimization problems where there is a global minimum are convex, and a great variety of demonstrations of equilibrium point stability of nonlinear systems involve Lyapunov functions. This work illustrates alternative techniques which may allow dispensing with a convexity assumption, or dispensing with use of a Lyapunov function. The techniques are grounded in topology, in particular Morse Theory, and results of Lefschetz-Hopf and Poincaré-Hopf. Illustrations are provided from the literature.

1. Introduction

Modern system theory probably has its roots in the commencement in the 19th century of the process of providing rigorous underpinnings to control, and the development in the 20th century of circuit theory and its associated mathematical foundations. The clear separation of system theory from these streams started to become evident with the appearance of one of the first books dealing with system theory, viz. [1]. Despite the word 'linear' in the title, this book ranged well beyond the theory of linear systems. From that time on the system theory field grew and spread vigorously, not necessarily because of, but no doubt aided by, the book. During this growth phase, *linear* system theory in the true sense of the words was heavily developed, with books such as [2], replete with its discussion of multivariate canonical forms among other things, showing the zenith of many of these developments. At the same time, the branching away from linear systems, in part motivated by challenging applications for which theory was often uncomfortably limited at best, gave rise to many tributaries. A great number of these fall under

*B. D. O. Anderson (corresponding author) was supported by the Australian Research Council under grant DP-160104500 and DP190100887, and by Data61-CSIRO.

†M. Ye was supported by the European Research Council (ERC-CoG-771687), and by Optus Business.

‡ORCID: 0000-0003-1698-0173.

the heading of nonlinear systems and a great many others fall under the heading of optimization.

In this paper, we discuss some recent developments, essentially grounded in differential topology, which bypass some of what we now see as the traditional foundations of nonlinear systems and optimization. In particular, we consider minimization problems where convexity is *not* assumed, and we consider stability problems, including global stability problems, where Lyapunov theory is not the foundation of the analysis. The set of developments we summarize are not that well known, although there are individual examples of their use in the system literature and in essence this paper in part is to proselytize, by using the opportunity to expose the ideas more widely.

Whatever the commonalities of the mathematical underpinnings, the material we present falls naturally into three distinct parts, presented over the following three sections of the paper. In the first of these, we discuss the application of *Morse Theory* [3, 4], which at its heart is concerned with counting the number of critical points of a scalar function, and classifying them as maxima, minima, saddle points etc. We note how such counting may lead one to infer the existence of a unique minimum (or maximum) of a function, without a convexity assumption. In the second part, we consider (continuous time) nonlinear dynamical systems, and use Poincaré-Hopf theory [5, 6] to identify situations in which there will be precisely one stable equilibrium. Further, for a large class of nonlinear systems known as monotone systems (which are widely occurring in many fields of application) [7, 8], we can often go further to establish global asymptotic stability, i.e. we rule out existence of chaotic or limit cycle behavior. The third part is a sort of discrete-time equivalent to the second part, in that we are able to identify situations where a mapping has precisely one fixed point. This development rests on results due to Lefschetz-Hopf [5, 9, 10].

A common feature of the mathematics behind the three parts is that *properties of a local nature are shown to imply properties of a global nature.* Thus for example, in connection with the Morse Theory development, it is not relevant whether a function is convex everywhere (i.e. has a positive definite Hessian matrix everywhere), but only that it be locally convex at each critical point. With additional rather general conditions, this is enough to yield the global property that there is a single critical point which is a minimum.

A different common feature over the three parts is our provision of applications examples illustrating the results. The examples are of a sort which are more attuned to engineering journals than mathematics journals, dealing with nuclear source detection, epidemics, and social networks. In each section of the paper, the general mathematical/system theory idea is first presented, and then it is followed by the example.

Morse Theory and the major results of Lefschetz-Hopf and Poincaré-Hopf are unknown to many, perhaps the majority, of systems researchers with engineering background. Isolated works have of course appeared in the engineering literature utilizing these tools, but not as far as we are aware with any sense of unification of the concepts as attempted here. We aim to present just enough ideas to allow engineers to see the general framework, and provide motivating systems examples at the same time to illustrate how one may seek to apply the framework.

1.1. Preliminaries and the Euler characteristic

We conclude the section with some brief notation. For two vectors $x, y \in \mathbb{R}^n$, we write $x \geq y$ and $x > y$ if for all $i = 1, \ldots, n$, $x_i \geq y_i$ and $x_i > y_i$, respectively. The $n \times n$ identity matrix is I_n, and the i^{th} canonical basis vector is $\mathbf{e}_i$. The Euclidean norm in $\mathbb{R}^n$ is $|\cdot|$.

For a smooth manifold M with boundary, we denote the boundary as ∂M and the interior as $\mathrm{Int}(M) = M \setminus \partial M$. Central to our treatment is the Euler characteristic of the manifold M, denoted as $\chi(M)$ [11, 5], which is an integer number associated with M. A key property is that $\chi(M)$ is invariant with respect to a homotopy of M (roughly speaking, distortion or bending of M). The Euler characteristics for a great number of manifolds are known. We remark that the Euler characteristic can certainly be associated with topological spaces that are not smooth manifolds, e.g. an n-dimensional hyperrectangle in $\mathbb{R}^n$. Other similar concepts and results we introduce may have extensions to general topological spaces, but for simplicity of exposition, we consider only smooth manifolds with boundary, or sometimes $\mathbb{R}^n$ itself.

We will focus on M which are contractible, and for which $\chi(M) = 1$ [5]. We say that M is contractible if it is homotopy equivalent to a single point (roughly, if M can be continuously shrunk and deformed to a point). Any convex and compact space in $\mathbb{R}^n$ is contractible, e.g. the unit disc, a hyperrectangle, or a ball. However, contractible spaces need not be convex; M is contractible if there exists an $x_0 \in M$ such that for all $x \in M$ and $t \in [0, 1]$, the point $(1 - t)x_0 + tx \in M$. Such an M is called a star domain.

2. Morse theory

Morse Theory is in part concerned with counting the number of minima, maxima and saddle points of a real function, and noting relations between the counts. For rigorous and detailed introduction, see e.g. [4, 3].

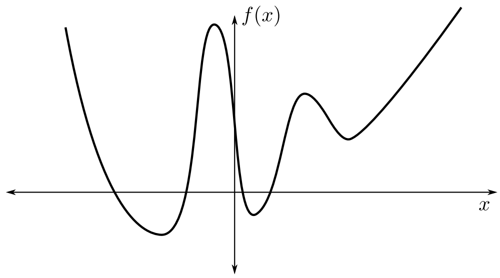

Figure 1: An illustration of a scalar function $f(x)$ that is a Morse function. Notice that the extreme points are isolated, $f(x) \to \infty$ as $|x| \to \infty$, and there is one more minimum than maximum.

We introduce the idea with a simple example, suitable for consumption by an engineering audience. Consider a (scalar) smooth function $f : \mathbb{R} \to \mathbb{R}$ with the properties that (a) all extreme points are isolated, (b) the function is bounded below, and (c) $f(x) \to \infty$ when $|x| \to \infty$. See Figure 1 for an example. The observation evident from this figure is that the number of minima exceeds the number of maxima by precisely 1. This is not just a property of the particular function illustrated, but in fact a property of any function f satisfying the stated conditions. It is this observation which Morse Theory generalizes. In particular, the generalization is to functions $f : M \to \mathbb{R}$ where M is a (smooth) manifold. In this work, we shall identify M with $\mathbb{R}^n$ for some positive integer n, or with a box sitting in $\mathbb{R}^n$. In the latter case, because there is a boundary to M, it is necessary to consider behavior on the boundary, and technically to assure the box boundary is smooth–thus corners for example would need to be rounded, but we skip these smoothing details, which are minor. In the more elementary forms of the theory, as we work with here, it is assumed that (a) f is smooth and bounded below, (b) at the critical points of f, i.e. the points where its gradient is zero, the Hessian matrix $\nabla^2 f$ is nonsingular, and (c) there are a finite number of critical points. (In the case of a box, this last property is automatic given the first two.) In the case of $M = \mathbb{R}^n$, it is most easily assured by requiring that $f \to \infty$ when $|x| \to \infty$. When M has a boundary, as with a box in $\mathbb{R}^n$, the additional requirements are (d) there are no critical points of f on the boundary and (e) the gradient of f points outwards from the boundary.

Definition 1 (Morse function). *A function $f : M \to \mathbb{R}$ satisfying the just listed properties (a), (b), (c) (and if M has a boundary then also (d) and (e)), is termed a Morse function.*

In $\mathbb{R}^n$, there can be minima, maxima and saddle points of indices $1, 2, \ldots, n-1$, corresponding to there being $1, 2, \ldots, n-1$ negative eigenvalues of the Hessian $\nabla^2 f$ computed at the critical point in question. Suppose that the number of such critical points is finite, a feature which will follow if f has no critical points for suitably large values of the argument, and f is outwardly pointing on the boundary of any large box. Let m_i denote the number of critical points where the Hessian has i negative eigenvalues. Then Morse Theory establishes certain inequality relations among the m_i and indeed one equality relation. These relations are as follows:

$$
\begin{aligned}
m_0 &\geq 1 \\
m_1 - m_0 &\geq -1 \\
m_2 - m_1 + m_0 &\geq 1 \\
&\vdots \\
m_{n-1} - m_{n-2} + \cdots + (-1)^{n-1} m_0 &\geq (-1)^{n-1} \\
m_n - m_{n-1} + \cdots + (-1)^n m_0 &= (-1)^n
\end{aligned} \tag{1}
$$

For the case $n = 2$, these relations reduce simply to $m_0 - m_1 = 1$, as observed above with the example in Fig. 1. If the manifold in question is other than $\mathbb{R}^n$ (or a box within $\mathbb{R}^n$), the numbers on the right hand side of the inequalities are varied to include what are known as the *Betti numbers*, see e.g. [4], of the manifold M; these are integer numbers determined just by the manifold and certainly numbers which are independent of the function. The number on the right hand side of the equality is $(-1)^n$ times the previously mentioned *Euler characteristic* of the manifold M, denoted $\chi(M)$. The Euler characteristic of $\mathbb{R}^n$ is 1, and recall that any contractible manifold (e.g. a box in $\mathbb{R}^n$) also has Euler characteristic of 1.

There is a crucial consequence of these equations, easily obtained, which is not widely known. Mention can however be found in textbooks, see e.g. [12].

Theorem 1 (Morse Theory Corollary for System Theory). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Morse function. Suppose there is a finite number of critical points and every critical point is a minimum. Then the function has a unique minimum.*

The proof is an immediate consequence of setting $m_i = 0$ for all $i \neq 0$ in the Morse equality above.

Remark 1 (Convexity not needed). *Obviously, there is no requirement for a smooth function possessing a unique minimum to be convex–consider for*

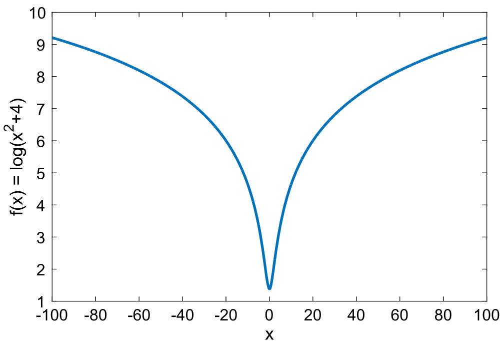

Figure 2: An example of a function that is not convex, and possesses a unique minimum: $f(x) = log(x^2 + a^2)$, with $a = 2$.

example the function $f(x) = log(x^2 + a^2)$, an example of which is given in Fig. 2 with $a = 2$. As seen in the figure, and may be checked by computing the first and second derivatives of f, there is a single critical point at $x = 0$, f is not convex and $f \to \infty$ as $|x| \to \infty$.

Remark 2 (Relaxing the critical point condition)**.** *If the hypothesis in the above theorem is changed to require that every critical point has a Hessian matrix with positive determinant, the same conclusion follows. This observation (with inessential replacement of f by $-f$ and minima by maxima) can be found in [13]. At a critical point of odd index $1, 3, \dots,$ the number of negative eigenvalues of the Hessian matrix is odd, and so its determinant will be negative. The hypothesis that the Hessian has a positive determinant at every critical point means then that $m_1 = m_3 = \cdots = 0$. The Morse equality then becomes $m_0 + m_2 + \cdots = 1$, and taken with the inequality $m_0 \geq 1$, it follows that $m_0 = 1$ and all other m_i are zero. As suggested by the title of [13], the proposed application is to potential games. In game theory, a potential game has the property that there is a function $P : X \to \mathbb{R}$ with X the strategy set (satisfying appropriate assumptions), and P has the further property that at a pure strategy Nash equilibrium $x^* \in X$, there holds $\nabla P(x^*) = 0$. The Hessian property then gives a tool for concluding uniqueness of the Nash equilibrium. Reference [13] however gives no example of a potential game where this property is actually used.*

Remark 3. *Manifolds other than $\mathbb{R}^n$ or a box within $\mathbb{R}^n$ do arise in some system theory problems. Multi-agent formation shape control problems are examined in [14] and [15], where the Morse function is used as a potential* *function, and distributed gradient descent controllers are used to drive the agents to the desired formation shape, corresponding to a minimum of the Morse function. The relevant manifold in particular in some instances has an Euler characteristic unequal to 1. This immediately guarantees that any Morse function cannot have a single critical point which is a minimum, and in some cases, the existence of multiple local minima can be a conclusion. This then has important implications regarding the possible final formation shape.*

2.1. Applications

There have been several applications of Theorem 1 in the literature. An early example, but only hinting at the way the theorem should be applied, is [16], while [17] indicates how the result above can be applied more clearly. Both these references deal with finding a coordinate basis change for a linear system to optimize performance of the system when calculations are implemented using the state-variable description but in an impaired way, due to constraints on word length of the real numbers. Convexity of the relevant performance index is *not* assured in either case.

We now present a much more recent example, associated with detection of a stationary nuclear source by a moving sensor by gradient-ascent-based maximum likelihood estimation, see [18]. A nuclear source of unknown strength is located at an unknown position (x_0, y_0) in the plane and a detector moves in the same plane at known speed ν along the x-axis in the positive direction. With reasonable assumptions, the particle arrivals detected constitute an inhomogeneous Poisson process with mean arrival rate

$$\lambda(t) = \frac{A_0}{(x_0 - \nu t)^2 + y_0^2} \tag{2}$$

where A_0 is a source strength parameter determined by the detector characteristics and the source itself. Given its partial dependence on the source, A_0 is necessarily assumed to be unknown. The observed particle arrival times over an interval $[T_1, T_2]$, call them $t_1, t_2, \ldots, t_n$, are independent. A limiting argument not provided here allowing $T_1 \to -\infty, T_2 \to \infty$ establishes that maximization of the associated likelihood function (with respect to the three scalar unknowns x_0, y_0 and A_0 and to obtain a maximum likelihood estimate of these quantities) is equivalent to the maximization of

$$L(A_0, x_0, y_0) = -\frac{\pi A_0}{\nu |y_0|} + n \log A_0 + \sum_{i=1}^{n} \log\left(\frac{1}{(x_0 - \nu t_i)^2 + y_0^2}\right) \tag{3}$$

or minimization of $-L$.

Observe for future reference that for very large values of $x_0^2+y_0^2$, L will be very large and negative, being approximated by $-nlog(x_0^2+y_0^2)$, and evidently it will be monotone in $x_0^2+y_0^2$.

At the outset, one must assume knowledge of the half plane, either $y>0$ or $y<0$, in which the source is located. The path of the detector cannot intersect the source, and the invariance of L with respect to the sign of y_0 is evident both mathematically and from the physical arrangement. To be definite, we will assume $y_0>0$ henceforth. Also, a normalization is possible to allow without loss of generality the assumption $\nu=1$, which we will also adopt henceforth.

Setting equal to zero the derivative of L with respect to each argument yields the critical point equations

$$\begin{aligned} A_0 &= \frac{ny_0}{\pi} \\ \sum_{i=1}^{n}\frac{x_0-t_i}{(x_0-t_i)^2+y_0^2} &= 0 \\ \sum_{i=1}^{n}\frac{1}{(x_0-t_i)^2+y_0^2} &= \frac{n}{2y_0^2} \end{aligned} \tag{4}$$

Evidently, A_0 is immediately known in terms of y_0 and the real issue is to determine (x_0,y_0). Our task is then to maximize

$$J(x_0,y_0)=L(\frac{ny_0}{\pi},x_0,y_0) \tag{5}$$

(or to minimize $-J$). It is readily verified that the critical points of J satisfy the second and third equations of (4). The associated Hessian of J can be calculated to be

$$H_J(x_0,y_0)=\begin{bmatrix} -2\sum_{i=1}^{n}\frac{y_0^2-(x_0-t_i)^2}{(y_0^2+(x_0-t_i)^2)^2} & 4y\frac{x_0-t_i}{(y_0^2+(x_0-t_i)^2)^2} \\ 4y\frac{x_0-t_i}{(y_0^2+(x_0-t_i)^2)^2} & -\frac{2n}{y_0^2}+2\sum_{i=1}^{n}\frac{y_0^2-(x_0-t_i)^2}{(y_0^2+(x_0-t_i)^2)^2} \end{bmatrix} \tag{6}$$

While the Hessian is not everywhere sign definite, it is negative definite at any point at which the critical point equations hold. (This last property can most easily be seen by defining $\theta_i=\arctan(y_0/(x_0-t_i))$ and then expressing the critical point equations and the Hessian using trigonometric functions of the θ_i). For details see [18]. This observation, the smoothness of J, its behavior for large values of its arguments as recorded below (3) and Theorem 1 guarantee uniqueness of the maximum of J, which can then be readily computed by a gradient ascent algorithm.

3. The Poincaré-Hopf Theorem

We introduce the application of this theorem in system theory by making an important observation drawn from the discussion on Morse Theory. Suppose that f is a Morse function. Consider the gradient descent system $\dot{x} = -\nabla f(x)$. Theorem 1 then establishes that if there are no equilibrium points other than stable ones, there can be but one such equilibrium point. The Poincaré-Hopf Theorem, or more properly the extension thereof presented below, effectively extends this result to more general systems of the form $\dot{x} = F(x)$. Very roughly speaking, if all equilibrium points are stable, there can only be one equilibrium point.

Of course, this theorem rests on ideas of differential topology, see e.g. [5], and we only exploit an elementary version of it, suited to motion in $\mathbb{R}^n$ or a box thereof. Consider a smooth map $F : X \to Y$, where X and Y are manifolds. Associated with F and any point $x \in X$ is a linear derivative mapping $dF_x : T_xX \to T_{F(x)}Y$, where T_xX and $T_{F(x)}Y$ denote respectively the tangent space of X at x and Y at $y = F(x) \in Y$. Of course, the manifold X locally at x looks like $\mathbb{R}^n$ for some n, and in an associated local coordinate basis dF_x is simply the Jacobian of F evaluated at x. Suppose now X and Y have the same dimension. A point $x \in X$ is called a regular point if dF_x is nonsingular, and a point $y \in Y$ is called a regular value if $F^{-1}(y)$ contains only regular points.

Suppose further that both X and Y are manifolds, with X compact and Y connected. The (Brouwer) degree of F at a regular value $y \in Y$ is given by [11]

$$\deg(F, y) = \sum_{x \in F^{-1}(y)} \operatorname{sign} \det(dF_x) \tag{7}$$

(It can be argued that the sum can only have a finite number of terms due to the compactness of X.) Evidently, sign $\det(dF_x)$ assumes the value $+1$ or -1 according as dF_x preserves or reverses orientation. A major result is that $\deg(F, y)$ is the same for all regular values y, see [11], and so the left side of (7) can be written simply as $\deg(F)$.

A point $x \in X$ is a zero of F if $F(x) = 0$, and it is an isolated zero if there exists an open ball around x in which there are no other zeros. When a zero x has nonsingular dF_x it is termed nondegenerate, and nonsingularity of dF_x is a sufficient condition for x to be isolated. The *index* of an isolated zero $x \in \mathbb{R}^n$, denoted $\operatorname{ind}_x(F)$ is the degree of the map

$$u : \partial\mathcal{D} \to \mathcal{S}^{n-1}$$

$$z \mapsto \frac{F(z)}{|F(z)|},$$

where $\mathcal{D}$ is a closed ball centred at x containing no other zero. A slightly simplified version of the Poincaré-Hopf theorem, but allowing manifolds with boundary, states:

Theorem 2 (Simplified Poincaré-Hopf Theorem)**.** *Consider a smooth vector field F on a compact n-dimensional manifold $\mathcal{M}$, i.e. a smooth map $F : \mathcal{M} \to T\mathcal{M}$. If $\mathcal{M}$ has a boundary $\partial\mathcal{M}$, then F must point outwards at every point on the boundary. Suppose that every zero $x^i \in \mathcal{M}$ is nondegenerate. Then*

$$\sum_i \operatorname{ind}_{x^i}(F) = \sum_i \operatorname{sign}\det(dF_{x^i}) = \chi(\mathcal{M}) \tag{8}$$

where $\chi(\mathcal{M})$ is the Euler characteristic of $\mathcal{M}$.

Remark 4. *The notion of a vector field F "pointing outwards" on the boundary $\partial\mathcal{M}$, and similarly of "pointing inwards" as introduced in the sequel, has both an intuitive geometric interpretation and a consistent rigorous technical definition. To keep this treatment illustrative in nature, we omit the rigorous technical definition, which relies on the concept of a tangent cone. Interested readers are referred to [19, Section III and Appendix D].*

For our purposes, we are interested in a specialization to compact, contractible and smooth m-dimensional manifolds embedded in $\mathbb{R}^n$, where $m \le n$. Recall that for such a manifold, the Euler characteristic is $\chi(\mathcal{M}) = 1$. Further, we suppose that dF_x is Hurwitz matrix at every zero x of F. The key result is as follows.

Theorem 3 (Poincaré-Hopf Corollary for System Theory)**.** *Consider the autonomous system*

$$\dot{x} = F(x) \tag{9}$$

where F is smooth, and $x \in \mathbb{R}^n$. Suppose that $\mathcal{M} \subset \mathbb{R}^m$ is an m-dimensional compact, contractible and smooth manifold with boundary $\partial\mathcal{M}$, such that F points inward to $\mathcal{M}$ at every point on $\partial\mathcal{M}$. Then there exists at least one equilibrium in Int($\mathcal{M}$). If dF_x is a Hurwitz square matrix (i.e. its eigenvalues all have negative real part) for every equilibrium point $\bar{x} \in \mathcal{M}$ (i.e. for which $F(\bar{x}) = 0$), then (9) has a unique equilibrium $x^* \in$ Int($\mathcal{M}$), and x^* is locally exponentially stable.

That there exists at least one equilibrium follows from the fact that $\chi(M) = 1$, implying that the summation on the left in (8) (which is over the equilibrium points) is nonempty. The proof follows by observing that the requirement that dF_x is Hurwitz for all zeros of F ensures that the signs of the determinants of dF_{x_i} must be the same for all equilibrium points. Since also $\chi(\mathcal{M}) = 1$, applying the main theorem to $G := -F$ establishes the result. At this point, the reader can verify that, as earlier observed, this result encompasses the Theorem 1, through identification of $F(x)$ here with the gradient of $-f(x)$ in Theorem 1.

There is however a significant distinction in the applicability of the two theorems. The generalization from gradient flows to general vector fields may be welcome, but it brings with it the possibility of limit cycles and chaotic behavior which simply cannot occur in a gradient flow, at least if the function f is real analytic [20]. We therefore introduce another tool which helps us eliminate the possibilities when paired with Theorem 3, and is applicable in many situations of interest–*monotone systems.*

3.1. Monotone systems

We provide a brief introduction to monotone systems, and again, one lesser known or possibly unknown extension of the theory of particular relevance in applications. Details (apart from Theorem 4 below) are available in e.g. [7, 21]. We continue to study the system $\dot{x} = F(x)$ and we focus on behavior of the system in particular orthants of $\mathbb{R}^n$, including but not limited to the positive orthant $x_i \geq 0 \; \forall i$.

To be more precise, consider a sequence $m_i, i = 1, 2, \ldots, n$ with $m_i \in \{0, 1\}$. Then the sequence defines a particular orthant K_m by

$$K_m = \{x \in \mathbb{R}^n : (-1)^{m_i} x_i \geq 0, \; \forall i \in \{1, \ldots, n\}\} \tag{10}$$

For a given orthant K_m, we write $x \leq_{K_m} y$ and $x <_{K_m} y$ if $y - x \in K_m$ and $y - x \in \text{Int } K_m$ respectively.

We consider the system (9) on a convex, open set $U \subseteq \mathbb{R}^n$, with F smooth such that dF_x exists for all $x \in U$, and the solution of (9) exists and is unique for every $x(0) \in U$. Let $\phi_t(x_0)$ denote the solution $x(t)$ when $x(0) = x_0$.

Definition 2 (Type K_m monotone system)**.** *The system (9) is a type K_m monotone system if whenever there exists $x_0, y_0 \in U$ with $x_0 \leq_{K_m} y_0$, then $\phi_t(x_0) \leq_{K_m} \phi_t(y_0)$; then ϕ_t is said to preserve the partial ordering $\leq_{K_m}$ for $t \geq 0$.*

There is a simple necessary and sufficient condition for the monotone property involving the Jacobian (Kamke–Müller condition) [7, Lemma 2.1]:

Lemma 1 (Condition for monotone property)**.** *With notation as above, suppose that F is of class C^1 in U, where U is open and convex in $\mathbb{R}^n$. Then $\phi_t(x_0)$ preserves the partial ordering $\leq_{K_m}$ for $t \geq 0$ if and only if $P_m dF_x P_m$ has all off-diagonal entries nonnegative for every $x \in U$, where $P_m = \mathrm{diag}((-1)^{m_1}, \dots, (-1)^{m_n})$.*

Given this characterization, the following definition makes sense.

Definition 3 (Irreducible monotone system)**.** *An irreducible type K_m monotone system is one for which dF_x is irreducible for all $x \in U$.*

There are many results establishing convergence for type K_m monotone systems. Here is one from [7, Theorem 2.6] which (among other things) rules out the occurrence of chaos, and limits substantially the role played by limit cycles. In the lemma, we use $B(x^i)$ to denote the basin of attraction of an equilibrium x^i.

Lemma 2 (Monotone system equilibria and limit cycles)**.** *Let $\mathcal{M}$ be an open, bounded and positively invariant set for an irreducible type K_m monotone system (9). Suppose the closure of $\mathcal{M}$, call it $\bar{\mathcal{M}}$, contains a finite number of equilibria x^i. Then*

$$\cup_{x^i} \mathit{Int}\ (B(x^i)) \cap \bar{\mathcal{M}}$$

is open and dense in $\mathcal{M}$, and any point not in this set lies on a nonattractive limit cycle.

The above result restricts the nature of allowed limit cycles substantially. More again is true however if we are willing to assume (or prove with e.g. Theorem 3) that there is but a single equilibrium in $\bar{\mathcal{M}}$ and it is actually contained in $\mathcal{M}$. The result is given below, first reported in [22], and with a greatly simplified proof in [19].

Theorem 4 (Global convergence for monotone systems with a single equilibrium point)**.** *Let $\mathcal{M}$ be an open, bounded, convex and positively invariant set for an irreducible type K_m monotone system (9). Suppose there is a unique equilibrium $x^* \in \mathcal{M}$ and no equilibrium in $\bar{\mathcal{M}} \setminus \mathcal{M}$. Then convergence to x^* occurs for every initial condition in $\mathcal{M}$.*

Proof. In light of Lemma 2, we must prove there does not exist a limit cycle. The argument is by contradiction. Let a be a point on a limit cycle. Then one can pick points $\underline{a} \in \mathcal{M}$ and $\bar{a} \in \mathcal{M}$ such that $\underline{a} <_{K_m} a <_{K_m} \bar{a}$;

construct two small balls $\mathcal{B}_1$ and $\mathcal{B}_2$ surrounding $\underline{a}$ and $\bar{a}$ respectively, which do not intersect the boundary of $\mathcal{M}$ nor contain a, with every $x \in \mathcal{B}_1$ and $y \in \mathcal{B}_2$ obeying $x <_{K_m} < a <_{K_m} y$. Without loss of generality also, $\underline{a}$ and $\bar{a}$ can be assumed not to lie on any limit cycle. Then $\lim_{t\to\infty} \phi_t(\underline{a}) = x^*$ and $\lim_{t\to\infty} \phi_t(\bar{a}) = x^*$ according to Lemma 2. Using the monotone property easily gives $\phi_t(\underline{a}) \leq_{K_m} \phi_t(a) \leq_{K_m} \phi_t(\bar{a})$ and combining these gives that $\lim_{t\to\infty} \phi_t(a) = x^*$, a contradiction, since a is a point on a limit cycle. □

Evidently then, the combination of the Poincaré-Hopf and the monotone system ideas gives a powerful result (roughly speaking) on global convergence to a unique equilibrium, for monotone systems for which any equilibrium is known to be locally exponentially stable.

3.2. Applications

We present here one application of the above ideas, in the analysis of a deterministic networked Susceptible-Infected-Susceptible (SIS) epidemic model. These models are well analyzed in the literature, in the main using Lyapunov theory to study equilibria and convergence, [23, 24, 25]. However, use of the methods just presented allows us to go further, in that we consider classes of systems not yet handled in the literature apart from our very recent work under review, which is available on arXiv [19]. In the extended class of systems, the construction of a Lyapunov function has not been presented, and in fact may not be straightforward.

We now present the basic SIS model, and refer the reader to e.g. [23, 26, 25] for more details on motivation and derivation. Each individual resides in one of n different populations, of fixed and large size. Each individual is either infected (I) with or susceptible (S) to some disease of interest, being capable of transitioning in either direction between the two states. The collection of n populations forms a metapopulation, with linkages described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, B)$. Here, $\mathcal{V} = \{1, \ldots, n\}$ is the set of nodes, corresponding to each of the populations, with $n \geq 2$ nodes. The set of ordered edges is $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, and the nonnegative adjacency matrix B is defined so that $b_{ij} > 0$ if and only if an edge $e_{ji} = (v_j, v_i)$ is in $\mathcal{E}$, and $b_{ij} = 0$ otherwise. The neighbor set of node i is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} : e_{ji} \in \mathcal{E}\}$. A graph is strongly connected if there is a path from every node to every other node and the results below will invoke an explicit assumption of strong connectivity.

Associated with node $i \in \mathcal{V}$ is the variable $x_i(t)$ representing the *fraction* of population i that is Infected. The dynamic equation for node i is:

$$\dot{x}_i(t) = -d_i x_i(t) + (1 - x_i(t)) \sum_{j \in \mathcal{N}_i} b_{ij} x_j(t) \quad i = 1, 2, \ldots, n \tag{11}$$

where d_i is the *recovery rate* of node i, and for a node j that is a neighbor of node i in $\mathcal{G}$, $b_{ij} > 0$ is the *infection rate* from node j to node i; if $j \notin \mathcal{N}_i$, then $b_{ij} = 0$. With $x = [x_1, \ldots, x_n]^\top, D = \text{diag}(d_1, \ldots, d_n)$ and $X(t) = \text{diag}(x_1(t), \ldots, x_n(t))$, we can also write the networked SIS dynamical equation as

$$\dot{x}(t) = (-D + B - X(t)B)x(t) \tag{12}$$

Define the hypercube

$$\Xi_n = \{x \in \mathbb{R}^n_{\geq 0} : 0 \leq x_i \leq 1, i \in \{1, \ldots, n\}\}.$$

Unsurprisingly, it follows from (12) that Ξ_n is an invariant set for the motion, which is a physical requirement ensuring that the x_i variables retain their important meaning in the model context.

It is immediate that $x = \mathbf{0}_n$ is an equilibrium for (12); for obvious reasons, it is termed the healthy equilibrium, and it may or may not be stable, as explored further below. Any other equilibrium in Ξ_n is termed an *endemic equilibrium* since the disease is persistent in a nonzero proportion of at least one population. Equilibrium properties in the literature are discussed with the aid of the spectral radius of $D^{-1}B$:

$$\mathcal{R}_0 = \rho(D^{-1}B), \tag{13}$$

and $\mathcal{R}_0$ is often termed the *effective reproduction number* of the disease for reasons that will become apparent below. The following is a key result derived in [23, 24, 27] with various mild adjustments to the assumptions and with the references containing different proofs.

Theorem 5 (SIS equilibria). *With notation as above, consider the system (12) and suppose that the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, B)$ is strongly connected, or equivalently, that B is an irreducible matrix. Then the following hold:*

- *If $\mathcal{R}_0 \leq 1$, the healthy equilibrium $x = \mathbf{0}_n$ is the unique equilibrium, and is globally attractive for all $x(0) \in \Xi_n$;*
- *If $\mathcal{R}_0 > 1$, there is in addition to the equilibrium $x = \mathbf{0}_n$, a single endemic equilibrium $x^* > \mathbf{0}_n$, which is attractive for all $x(0) \in \Xi_n \setminus 0$.*

We remark that the persistence or elimination of the disease depends solely on the effective reproduction number $\mathcal{R}_0$, which can intuitively be considered as the number of new infections transmitted by a single infected individual before becoming cured.

The proofs of the above theorem in [23, 24, 27] all rely on Lyapunov theory *after* using algebraic arguments to establish the precise number of equilibria, for both $\mathcal{R}_0 \leq 1$ and $\mathcal{R}_0 > 1$. For example, when $\mathcal{R}_0 > 1$, the existing proofs use complicated algebraic arguments centered on proof by contradiction to establish the uniqueness of the endemic equilibrium x^*, and then construct an appropriate Lyapunov function driven by knowledge that x^* is unique. The paper [24] for example establishes uniqueness of any equilibrium point using a series of inequalities, and then with x^* denoting the equilibrium point, shows that a differential equation for the difference $y := x - x^*$ takes the form $\dot{y} = A(y)y$. where the entries of A obey certain sign constraints associated with $A(0)$. (The relevant class of matrices are called Metzler-matrices, and we briefly discuss this class in the sequel.) This allows a Lyapunov function $V = \sum_i c_i|y_i|$ to be constructed for certain positive constants c_i, with somewhat nontrivial arguments being used to establish that $\dot{V} < 0 \ \forall \ y \neq 0$, given $\bar{x} + y(0) \in \Xi_n \setminus 0$. The paper also suggests that an alternative Lyapunov function

$$V = \sum_i c_i(x_i - x_i^* \ln x_i)$$

with certain positive c_i can be used to establish stability, but the calculation is not straightforward. It is interesting to note that the time between [23] and [24] is more than three decades, with both using Lyapunov theory to establish very similar results, but with the latter paper doing so more simply than the first, though not entirely painlessly. Subsequent to both [23, 24], and four decades after the first of these references, in [27] a quadratic Lyapunov function was put forward for the $\dot{y} = A(y)y$ system introduced in [24], again relying on results concerning Metzler matrices.

It turns out that Theorem 5 can also be derived using the methods outlined earlier in this section. We omit such a treatment however, since it becomes a special case of a further theorem below where these methods are used to obtain the result, and where it is not clear one could even hope to generalise the Lyapunov methods mentioned above.

Those with responsibility for public health are obviously interested in understanding how to limit epidemic spreading, and several approaches can also be reflected in changes to the above model. We consider one of those possibilities here. We postulate that the recovery rate $d_i > 0$ is modified, e.g. by reflecting increased medical resources being applied, with more resources corresponding to larger values of x_i. More specifically, we suppose that d_i is replaced by $\bar{d}_i(t)$, being the sum of a base level constant recovery rate d_i and a control term $u_i(t) = h_i(x_i(t))$:

$$\bar{d}_i(t) = d_i + h_i(x_i(t)) \tag{14}$$

where for all i, $h_i : [0,1] \to \mathbb{R}_{\geq 0}$ is bounded, $h_i(0) = 0$, and h_i is smooth and monotone nondecreasing. The vector equation then becomes

$$\dot{x}(t) = (-D - H(x(t)) + B - X(t)B)x(t) \tag{15}$$

where $H(x(t)) = \mathrm{diag}(h_1(x_1(t), \ldots, h_n(x_n(t))$. It is immediate to establish that $\mathbf{0}_n$ remains a (healthy) equilibrium of the system, and not hard to establish that Ξ_n remains an invariant set for this system. One can for example appeal to Nagumo's Theorem [28], checking that for all i, $\dot{x}_i \geq 0$ if $x_i = 0$ and $\dot{x}_i \leq 0$ if $x_i = 1$. Without feedback, the condition $\mathcal{R}_0 \leq 1$ ensured the healthy equilibrium is attractive. Intuitively, one would expect that applying feedback is not going to change that conclusion, and indeed the methods of e.g. [23] for the nonfeedback problem continue to be applicable. Our interest is in considering the effect of feedback when there is an endemic equilibrium in the nonfeedback case, i.e. $\mathcal{R}_0 > 1$.

Before proceeding further however, we note a small linear algebra result. A real square matrix A is termed a *Metzler* matrix when it has all off-diagonal entries nonnegative [29]. Then the following result holds [30]:

Lemma 3 (Metzler matrix property)**.** *Let A be an $n \times n$ irreducible Metzler matrix. Let $s(A)$ denote the real part of the eigenvalue of A with the largest real part. Then $s(A)$ is a simple eigenvalue of A and there exists a unique (up to scalar multiple) $x > \mathbf{0}_n$ such that $Ax = s(A)x$, and there is no other linearly independent eigenvector with positive entries.*

To use the theoretical tools presented earlier in Theorem 3, we must first identify a contractible manifold $\mathcal{M}$ for the system (15) with the property that at all points on $\partial\mathcal{M}$, $F(x)$ points inward. In principle, this calculation is straightforward, but slightly messy. We indicate it only in outline. Because $\mathcal{R}_0 > 1$ and the matrix $-D + B$ has nonnegative off-diagonal entries, the matrix can be shown to have the property that there exists $y > \mathbf{0}_n$ for which $(-D + B)y = \phi y$ for the simple eigenvalue $s(-D + B) \triangleq \phi > 0$ (see e.g. [31, Proposition 1]). Without loss of generality, let $\max_i y_i = 1$. Given $0 < \epsilon < 1$, define the set

$$\mathcal{M}_\epsilon = \{x : \epsilon y_i \le x_i \le 1, i = 1, 2, \ldots, n\} \subset \Xi_n. \tag{16}$$

Then, as may be checked, there exist sufficiently small $\epsilon_u \in (0, 1)$ and $\epsilon_v \in (0, \epsilon_u]$ such that for all $\epsilon \in (0, \epsilon_u]$, trajectories of (15) intersecting the boundary of $\mathcal{M}_\epsilon$ are inward pointing at the point of intersection and for any $x(0) \in \partial\Xi_n \setminus \mathbf{0}_n$, there holds $x(\bar{t}) \in \mathcal{M}_{\epsilon_v}$ for some finite $\bar{t} > 0$. This latter fact implies that any endemic equilibrium $\bar{x}$ necessarily satisfies $\bar{x} \in \mathcal{M}_{\epsilon_v}$. For details of the calculations and precise results, see [19].

In order to apply one of the Poincaré-Hopf theorems, we must smooth each edge and corner of $\mathcal{M}_\epsilon$ to create an arbitrarily close manifold $\tilde{\mathcal{M}}_\epsilon$. This essentially technical adjustment preserves the inward-pointing property of trajectories at the boundary. And now we can record the key result.

Theorem 6 (No elimination of endemic equilibrium via feedback)**.** *With notation as above, consider the system* (15) *and suppose that the graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, B)$ *is strongly connected, or equivalently, that* B *is an irreducible matrix. Suppose that* $\mathcal{R}_0 > 1$*, and the* h_i *satisfy the conditions given below* (14)*. Then there is a single endemic equilibrium* $\tilde{x}^* > \mathbf{0}_n$*, which is attractive for all* $x(0) \in \Xi_n \setminus \mathbf{0}_n$*, with convergence to the equilibrium being exponentially fast.*

A compact version of the proof is now presented to allow the reader an appreciation of how one can apply Theorem 3. Further details, including specific calculations and lengthened arguments, may be found in [19].

Proof. We shall first establish that there is a unique endemic equilibrium, by studying the Jacobian matrix properties computed at any possible equilibrium. Suppose that $\tilde{x}$ is one such equilibrium (the existence of which is derived from Theorem 3). Below (15), we established that any such equilibrium satisfies $\mathbf{0}_n < \tilde{x} < \mathbf{1}_n$, and $\tilde{x}$ must further obey

$$\mathbf{0}_n = (-D + H(\tilde{x}) + (I - \tilde{X})B)\tilde{x} \tag{17}$$

Recall that B is irreducible and nonnegative; since $I - \tilde{X}$ is nonsingular and nonnegative, $(I - \tilde{X})B$ is also irreducible and nonnegative. Hence $P(\tilde{x}) \triangleq -D - H(\tilde{x}) + (I - \tilde{X})B$ is an irreducible Metzler matrix. Since $\tilde{x} > 0$, Equation (17) yields $s(P(\tilde{x})) = 0$ from Lemma 3. Now the Jacobian matrix for (15) can be checked to be

$$dF_x = P(x) - \sum_{i=1}^{n}\Big(\sum_{j=1}^{n} b_{ij}x_j\Big)\mathbf{e}_i\mathbf{e}_i^\top - \operatorname{diag}\big(\frac{\partial h_1}{\partial x_1}x_1, \dots, \frac{\partial h_n}{\partial x_n}x_n\big) \tag{18}$$

One can check that the irreducibility of B ensures that $\sum_{i=1}^{n}(\sum_{j=1}^{n} b_{ij}x_j)\mathbf{e}_i\mathbf{e}_i^\top$ is a positive diagonal matrix for all $x \in \tilde{\mathcal{M}}_{\epsilon_v}$ while $\operatorname{diag}(\frac{\partial h_1}{\partial x_1}x_1, \dots, \frac{\partial h_n}{\partial x_n}x_n)$ is obviously a nonnegative diagonal matrix. In summary then, dF_x is obtained from $P(x)$ through the subtraction of a negative definite diagonal matrix. Now specialize to $x = \tilde{x}$. Since $s(P(\tilde{x})) = 0$ is the eigenvalue of $P(\tilde{x})$ with the greatest real part, the subtraction ensures that $dF_{\tilde{x}}$ is a Hurwitz matrix (in fact $s(dF_{\tilde{x}}) < 0$). The specific arguments are laid out using standard results from M-matrix theory[1] [29, 32]. Since $\tilde{x}$ can be any endemic equilibrium, by Theorem 3 we conclude it must be unique and locally exponentially stable. Call it henceforth $\tilde{x}^*$.

It now remains to establish that the region of attraction for $\tilde{x}^*$ is $\Xi_n \setminus \mathbf{0}_n$. Observe that dF_x has all nonnegative off-diagonal terms and in fact since $(I - X)B$ is irreducible in $\tilde{M}_{\epsilon_v}$, (15) is a an irreducible monotone system where K_m coincides with the positive orthant. Then Theorem 4 yields the desired conclusion concerning the region of attraction of $\tilde{x}^*$. □

Note that the above theorem includes the original SIS model, viz. the no-feedback model where $H \equiv \mathbf{0}_{n\times n}$, as a special case. Note also that although feedback of the type defined cannot eliminate occurrence of an endemic equilibrium, it can move the endemic equilibrium, and arguments using the monotone property actually establish that the endemic equilibrium is always moved closer to the origin as a result of feedback. We state the result without proof here, but interested readers can see [19].

Theorem 7 (Feedback improves the endemic equilibrium)**.** *With notation as above, suppose that the condition $\mathcal{R}_0 > 1$ for existence of an endemic equilibrium holds. Let x^* and $\tilde{x}^*$ denote the unique equilibria for the uncontrolled and controlled systems,* (12) *and* (15)*, respectively. Then $\tilde{x}^* < x^*$.*

Remark 5. *We have shown how the combination of Poincaré–Hopf theory and monotone systems theory can be combined to analyse a large class of networked SIS models with distributed feedback control at each node, with the*

[1] An M-matrix is a matrix with all off-diagonal entries nonpositive, and with all eigenvalues having nonnegative real part. The specific arguments for the result of this theorem, including the specific M-matrix results, are detailed in [19], and rest fundamentally on the Perron–Frobenius theorem for nonnegative matrices.

original SIS model is a special case. Apart from the new theoretical conclusions themselves, a separate key takeaway is that the new analysis framework allows for conclusions to be drawn on a broad class of dynamics without tedious calculations (or at most doing the calculations once for the general system). This is in contrast to the algebraic and Lyapunov-based approach, in which each change to the $h_i(x_i(t))$ *term in* (14) *would likely require adjustments to the algebraic calculations and/or Lyapunov function.*

4. Lefschetz-Hopf theory

Lefschetz fixed point theory and the developments joined to the name Hopf apply to smooth maps $F : X \to X$ where X is a compact oriented manifold [9] or a compact triangulated space [10]. Of course, a fixed point of such a mapping is also an equilibrium point for a system defined by

$$x(k+1) = F(x(k)) \tag{19}$$

Evidently, $X = \mathbb{R}^n$ is excluded, though obviously compact contractible subsets of $\mathbb{R}^n$ are allowed. Also, if a map F is known to have no fixed points for large values of its argument, the theory we will present in this section can often be applied by considering the restriction of F to a compact subset X of $\mathbb{R}^n$ such as a ball of suitably large radius that ensures X is a positively invariant set of (19).

Lefschetz fixed-point theory involves derivatives. Because we are assuming that F is smooth, we can associate with every $x \in X$ a linear derivative mapping dF_x; much as was detailed in Section 3, dF_x is essentially the Jacobian of F evaluated at x in local coordinates of X. Our interest is mainly in maps which have a finite number of fixed points, even though maps with an infinite number of fixed points exist, including important ones like the identity map $F(x) = x$.

A fixed point is called a *Lefschetz fixed point* if the eigenvalues of dF_x are unequal to 1. (This is an analog of requiring for $\dot{x} = F(x)$ that at an equilibrium point x, dF_x is nonsingular). A fixed point is *isolated* if there are no other fixed points in a sufficiently small ball around it; this again parallels the definition of an isolated zero in Section 3. Smoothness of F means that any Lefschetz fixed point is necesssarily isolated, and since X is compact, there can only be a finite number of isolated fixed points.

If a particular fixed point, call it x_i, is a Lefschetz fixed point, then the mapping $I - dF_{x_i}$ is an isomorphism of the tangent space $T_{x_i}(X)$ at x_i. If the mapping preserves orientation, its determinant is positive while if it reverses orientation, its determinant is negative. The *local Lefschetz number* of F at a fixed point x_i, written $L_{x_i}(F)$ is defined as $+1$ or -1 according as the determinant of $I - dF_x$ is positive or negative.

The map F is termed a *Lefschetz map* if all the fixed points are Lefschetz fixed points, and the *Lefschetz number* of F is defined as

$$L(F) = \sum_{x_i : F(x_i) = x_i} L_{x_i}(F) \tag{20}$$

To this point, the development has roughly paralleled that of the Poincaré-Hopf theorem, but at this point it potentially ceases to be a parallel. In fact, we can still relate the Lefschetz number to an Euler characteristic, but only by making an additional assumption on the manifold X. The result is as follows, and is a straightforward consequence of standard ideas in Lefschetz-Hopf theory, see e.g. [5], [9], [10], and is recorded explicitly in [33]).

Theorem 8 (Specialization of Lefschetz-Hopf Theorem)**.** *Let X be a compact oriented manifold or a compact triangulable space, and suppose $F : X \to X$ is a Lefschetz map, i.e. there are a finite number of fixed points x_i at each of which $I - DF_{x_i}$ is an isomorphism. Suppose further that F is homotopically equivalent to the identity map. Then there holds*

$$L(F) = \sum_{x_i : F(x_i) = x_i} L_{x_i}(F) = \chi(X) \tag{21}$$

where $L_{x_i}(F)$ is $+1$ or -1 according as det$(I - dF_{x_i})$ has positive or negative sign, and $\chi(X)$ is the Euler characteristic of X.

How do we deal with the requirement that F should be homotopically equivalent to the identity map? Either one has to generate a family of maps $X \to X$, parameterized by λ say, and smooth in x and λ, such that the identity map and F are both in the family, or we can impose a sufficient condition on X that implies the property. Fortunately, X being contractible is a sufficient condition so that for a given F, there always exists a homotopy from F to the identity map, see e.g. [5].

Now we have the machinery in place to establish that systems of the form in (19) that are known to be positively invariant on a contractible X which have all asymptotically stable fixed points can only in fact have one such fixed point. A slightly more restrictive version of this corollary appeared in [33], where it was assumed that X was convex, rather than just contractible.

Theorem 9 (Lefschetz-Hopf Corollary for System Theory)**.** *Consider a smooth map $F : X \to X$ where X is a compact, oriented and contractible manifold of arbitrary dimension. Suppose that for all fixed points x_i the eigenvalues of dF_{x_i} have magnitude less than 1. Then F has a unique fixed point, and in a local neighborhood about the fixed point, Equation* (19) *converges to the fixed point exponentially fast.*

Proof. Note that the assumption on the fixed points guarantees that $I - dF_{x_i}$ is nonsingular and all of its eigenvalues have strictly positive real part. This implies that the determinant of $I - dF_{x_i}$, which is the product of its eigenvalues, is positive. The fixed points are then necessarily isolated and the compactness of X guarantees that there are a finite number. Theorem 8 (The Specialization Theorem) above applies, with $\chi(X) = 1$ due to the contractibility assumption on X. The left side of (21) evaluates to be the number of fixed points. Thus there is just one fixed point.

The exponentially fast character of the local convergence is a consequence of the fact that the linearized system around the fixed point is exponentially fast, as we have assumed that dF_{x_i} at any fixed point x_i has eigenvalues all with magnitude less than 1. ☐

In contrast to what was done in the previous section, we omit development of some kind of global convergence property. It seems likely that a development using the theory of discrete-time monotone systems should be possible, see e.g. [34]. We turn therefore to an illustrative application.

4.1. Applications

We now present an application of Lefschetz–Hopf theory to the DeGroot–Friedkin model, which is dynamical model of a social network. The application result first appeared in [33]. We give here a brief introduction to the model, and readers interested in the details on the modelling derivation and motivation are referred to [35, 36]. The model considers $n \geq 3$ individuals[2] in a strongly connected social network discussing their opinions on a sequence of topics $k = 0, 1, \ldots$. The individuals discuss a topic k until a consensus of opinions is reached, and during this discussion, individual i has a self-confidence $x_i(k)$ in her own opinion. After consensus is reached on topic k and before discussion begins on the next topic $k+1$, individual i's self-confidence, for all $i = 1, \ldots, n$, is updated to become $x_i(k+1)$ according to how much "social power" individual i had in the discussion of topic k (there is a mathematical definition of social power which we omit here).

[2]The dynamics for $n = 2$ are too simple to be of interest for this discussion.

Leaving aside the individual steps within the formulation of the dynamics, the final model is given as in (19) where $x = [x_1, \ldots, x_n]^\top$ is the vector of self-confidences, and the map F is given by

$$F(x) = \begin{cases} \mathbf{e}_i & \text{if } x = \mathbf{e}_i \text{ for any } i \\ \frac{1}{\sum_{i=1}^n \frac{\gamma_i}{1-x_i}} \begin{bmatrix} \frac{\gamma_1}{1-x_1} \\ \vdots \\ \frac{\gamma_n}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases} \tag{22}$$

Here, γ_i is the eigenvector centrality of individual i, and for generic network topologies,[3] it was shown in [35] that $\gamma_i \in (0, 0.5)$ for all i and $\sum_i^n \gamma_i = 1$.

A number of properties of the map F in (22) have been established in [35, 36], and we recall several important ones. First, the n-dimensional unit simplex (in fact an $n-1$-dimensional manifold embedded in $\mathbb{R}^n$) defined as

$$\Delta_n = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \, \forall \, i = 1, \ldots, n, \sum_{i=1}^n x_i = 1\}$$

is an invariant of the map F. That is, $F : \Delta_n \to \Delta_n$. Second, it is clear that the corners of Δ_n, $\mathbf{e}_1, \ldots, \mathbf{e}_n$, are fixed points of F. Third, and despite its appearance in (22), F is smooth on Δ_n (in fact of class $\mathcal{C}^\infty$) [36, Corollary 2]. Fourth, there exists a sufficiently small $\delta > 0$ such that

$$\tilde{\Delta}_n = \{x \in \mathbb{R}^n : \delta \leq x_i \leq 1 - \delta \, \forall \, i = 1, \ldots, n, \sum_{i=1}^n x_i = 1\} \tag{23}$$

is also an invariant of the map F, and no fixed points can be in found in $\Delta_n \setminus \tilde{\Delta}_n$ other than the corners of Δ_n. These four facts will yield from the Brouwer Fixed Point Theorem [37] that at least one fixed point $\bar{x}$ exists in $\tilde{\Delta}_n$. At such a fixed point, every individual i has a nonzero self-confidence $\bar{x}_i \in (0, 1)$.

To establish whether $\bar{x}$ is in fact unique or not, the original analysis in [35, Appendix F] used an extensive set of complicated algebraic manipulations on the specific functional form of F in (22) to obtain a proof by

[3]There is a special topology class called star topology which is not considered in this treatment, but topologies are examined in [35, 36]. For a star topology, the map remains the same as in (22), but there exists a k such that $\gamma_k = 0.5$, which has nontrivial consequences on the limiting dynamics.

contradiction. Here, we bypass this by analysing only the Jacobian at potential fixed points in $\tilde{\Delta}_n$, and show how to apply Theorem 9 to establish the uniqueness of the fixed point in the interior of Δ_n.

We shall not detail all the steps and calculations, which can be found in [33], but will provide enough for the reader to understand the analysis process. Different from the above two application examples, the space X in question is the $n-1$-dimensional manifold $\tilde{\Delta}_n$ embedded in $\mathbb{R}^n$, so we must analyze the Jacobian in local coordinates. To do so, we first compute the Jacobian $\frac{\partial F}{\partial x}$ in the coordinates of $\mathbb{R}^n$ in which $\tilde{\Delta}_n$ is embedded, and then introduce a coordinate transformation y and associated map G, so that Theorem 9 can applied to the Jacobian dG_y.

One can compute the ii^{th} entry of $\frac{\partial F}{\partial x}$ to be $\frac{\partial F_i}{\partial x_i} = F_i \frac{1-F_i}{1-x_i}$. Similarly, we obtain, for $j \neq i$, the ij^{th} entry of $\frac{\partial F}{\partial x}$ as $\frac{\partial F_i}{\partial x_j} = -\frac{F_i F_j}{1-x_j}$. It can be proved that for all values of $x \in \tilde{\Delta}_n$, $\frac{\partial F}{\partial x}$ has a single zero eigenvalue, and all of the other eigenvalues are positive real (this is achieved by showing that $\frac{\partial F}{\partial x}^\top$ is the Laplacian matrix of a strongly connected directed graph having strictly real eigenvalues). We remark that in general, the Laplacian matrix of a strongly connected directed graph has all eigenvalues with nonnegative real part, but may have complex eigenvalue pairs. Unusually in our case, a specific positive diagonal matrix A can be found such that $\frac{\partial F}{\partial x} A$ is the symmetric Laplacian matrix of a connected undirected graph, thus having real eigenvalues. A standard linear algebra result yields that $\frac{\partial F}{\partial x}$ has real eigenvalues. Details are given in [33, Proof of Theorem 4]. Now, at a fixed point $\bar{x} \in \tilde{\Delta}_n$, one has

$$\left.\frac{\partial F_i}{\partial x_i}\right|_{\bar{x}} = \bar{x}_i. \tag{24}$$

$$\left.\frac{\partial F_i}{\partial x_j}\right|_{\bar{x}} = -\frac{\bar{x}_i \bar{x}_j}{1-\bar{x}_j} \quad , j \neq i. \tag{25}$$

Since $\bar{x} \in \tilde{\Delta}_n$ implies $\sum_{i=1}^n \bar{x}_i = 1$, the trace of $\frac{\partial F}{\partial x}|_{\bar{x}}$ must be equal to 1. Since $n \geq 3$ and $\frac{\partial F}{\partial x}|_{\bar{x}}$ has $n-1$ positive real eigenvalues, it immediately follows that all eigenvalues of $\frac{\partial F}{\partial x}|_{\bar{x}}$ are less than 1 in absolute value.

Next, let us introduce a vector $y \in \mathbb{R}^{n-1}$, of lower dimension therefore than x, and obtained from x as $y_1 = x_1, y_2 = x_2, \ldots, y_{n-1} = x_{n-1}$. This means that *on the manifold* $\tilde{\Delta}_n$, there holds $x_n = 1 - \sum_{k=1}^{n-1} y_k$. On the manifold, and in the y coordinates, we define G as the map with $G_1(y) = F_1(x), \ldots, G_{n-1}(y) = F_{n-1}(x)$, which implies that $F_n = 1 - \sum_{k=1}^{n-1} G_k$. As explained above, the Jacobian of interest is the one on the manifold of $\tilde{\Delta}_n$, that being in fact dG_y. For any $G_i(y_1, \ldots, y_{n-1}) = F_i(y_1, \ldots, y_{n-1}, 1 - \sum_{k=1}^{n-1} y_k)$, the Chain rule yields

$$\frac{\partial G_i}{\partial y_j} = \sum_{k=1}^{n} \frac{\partial F_i}{\partial x_k}\frac{\partial x_k}{\partial y_j} = \frac{\partial F_i}{\partial x_j}\frac{\partial x_j}{\partial y_j} + \frac{\partial F_i}{\partial x_n}\frac{\partial x_n}{\partial y_j}$$

because $\partial x_k / \partial y_j = 0$ for $k \neq j, n$. Define the matrices

$$T = \begin{bmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ -\mathbf{1}_{n-1}^\top & 1 \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{1}_{n-1}^\top & 1 \end{bmatrix}, \tag{26}$$

where $\mathbf{0}_{n-1}$ and $\mathbf{1}_{n-1}$ are the $n-1$ dimensional vector of all zeros and all ones, respectively, and I_{n-1} is the $n-1$ dimensional identity matrix. One can compute (see [33] for the precise steps) that

$$\begin{bmatrix} dG_y & \frac{\partial F}{\partial x_n} \\ \mathbf{0}_{n-1}^\top & 0 \end{bmatrix} = T^{-1}\frac{\partial F}{\partial x}T, \tag{27}$$

where $\frac{\partial F}{\partial x_n}$ is a column vector with i^{th} element $\frac{\partial F_i}{\partial x_n}$. The similarity transform in (27) tells us that the matrix on the left of (27) has the same eigenvalues as $\frac{\partial F}{\partial x}$, and since the matrix is block triangular, it follows that dG_y has the same nonzero eigenvalues as $\frac{\partial F}{\partial x}$. From the above, and restricting x to being an arbitrary fixed point $\bar{x}$, with $\bar{y} \in \tilde{\Delta}_n$ corresponding, we immediately conclude that all eigenvalues of $dG_y|_{\bar{y}}$ are less than 1 in magnitude. Since $\tilde{\Delta}_n$ is contractible, applying Theorem 9 establishes that G has a unique fixed point $\bar{y}$, and it is locally exponentially stable. Equivalently, F has a unique fixed point in $\tilde{\Delta}_n$ which is locally exponentially stable for the system (19).

Remark 6. *Having established the uniqueness of the interior fixed point $\bar{x}$, convergence of (19) with the map F in (22) for all initial conditions $x(0) \in \Delta_n \setminus \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is proved in [35] using a Lyapunov function that is constructed using the fact that $\bar{x}$ is unique. In contrast, [36] continues the theme of [33] and this treatment by analyzing the Jacobian $\frac{\partial F}{\partial x}$. However, [36] analyzes $\frac{\partial F}{\partial x}$ over the entire manifold $\tilde{\Delta}_n$, and by introducing a nonlinear differential coordinate transformation (a transformation of the tangent space), proves simultaneously the uniqueness of $\bar{x}$ and the exponential convergence for all $x(0) \in \Delta_n \setminus \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ using* ***nonlinear contraction theory*** *[38]. The differential transform, and the subsequent contraction analysis is nontrivial and not obvious. In comparison then, the analysis presented in this treatment to establish the uniqueness of $\bar{x}$ is much simpler than the original approach in [35], and also establishes local exponential convergence.*

However, the analysis presented in this treatment does not yield a global convergence result, for which more sophisticated arguments are required as in [36].

5. Conclusions

We have introduced three different results from topology, viz. Morse Theory, Poincaré–Hopf theory, and Lefschetz–Hopf theory, presenting them in such a way that they can be readily applied to systems theory analysis. We then showed how each result could be applied to models that are derived from real applications in localisation, epidemics, and social networks. A unifying theme of the three results is the drawing of global conclusions, viz. uniqueness of minima or equilibria or fixed points, from analysis of local quantities, viz. the Hessian or Jacobian of the relevant function, vector field, or map. Moreover, the domain of the function, or space of the vector field or map, need not be convex. When paired with other methods (gradient ascent algorithms in Section 2, and monotone systems theory in Section 3), global convergence results can be established without relying on Lyapunov theory.

We are in no way suggesting the techniques covered in this paper are generally superior to the fundamental Lyapunov methods that have underpinned a great number of developments of systems theory. For example, the methods of Section 3 and 4 cannot be easily adapted for analysis of non-smooth or non-autonomous systems. However, we do hope that we have laid out a convincing argument that a variety of other, sometimes non-standard, tools can be employed and just sometimes, the analyses and calculations are made easier.

References

[1] L. Zadeh and C. Desoer, *Linear System Theory: the State Space Approach*. McGraw-Hill, New York, 1963.

[2] T. Kailath, *Linear Systems*. Prentice-Hall Englewood Cliffs, NJ, 1980, vol. 156. MR0569473

[3] J. Milnor, *Morse Theory* (AM-51). Princeton University Press, 2016, vol. 51. MR0163331

[4] Y. Matsumoto, *An Introduction to Morse Theory*, vol. 208. American Mathematical Soc., 2002. MR1873233

[5] V. Guillemin and A. Pollack, *Differential Topology*, vol. 370. American Mathematical Soc., 2010. MR2680546

[6] J. Lafontaine, *An Introduction to Differential Manifolds*. Springer, 2015. MR3379533

[7] H. L. Smith, "Systems of ordinary differential equations which generate an order preserving flow. A survey of results," *SIAM Review*, vol. 30, no. 1, pp. 87–113, 1988. MR0931279

[8] D. Angeli and E. D. Sontag, "Monotone control systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1684–1698, 2003. MR2021417

[9] M. W. Hirsch, *Differential Topology*. Springer Science & Business Media, vol. 33, 2012. MR1336822

[10] M. A. Armstrong, *Basic Topology*. Springer Science & Business Media, 2013. MR0705632

[11] J. W. Milnor, *Topology from the Differentiable Viewpoint*. Princeton University Press, 1997. MR1487640

[12] U. Helmke and J. B. Moore, "l^2 sensitivity minimization of linear system representations by gradient flows," *Journal of Mathematical Systems, Estimation and Control*, vol. 5, no. 1, 1995. MR1304406

[13] F. Christensen, "A necessary and sufficient condition for a unique maximum with an application to potential games," *Economics Letters*, vol. 161, pp. 120–123, 2017. MR3723237

[14] M. A. Belabbas, "On global stability of planar formations," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 2148–2153, Aug 2013. MR3090050

[15] B. D. O. Anderson, "Morse theory and formation control," in: *2011 19th Mediterranean Conference on Control & Automation (MED)*. IEEE, 2011, pp. 656–661.

[16] J. Perkins, U. Helmke, and J. Moore, "Balanced realizations via gradient flow techniques," *Systems & Control Letters*, vol. 14, no. 5, pp. 369–379, 1990. MR1057154

[17] A. G. Madievski, B. D. O. Anderson, and M. Gevers, "Optimum realizations of sampled-data controllers for FWL sensitivity minimization," *Automatica*, vol. 31, no. 3, pp. 367–379, 1995. MR1321007

[18] B. D. O. Anderson, S. Dasgupta, H. Baidoo-Williams, M. Anjum, and R. Mudumbai, "Unique maximum likelihood localization of nuclear

sources," in: *58th IEEE Conference on Decision and Control*. IEEE, 2019, pp. 1–2.

[19] M. Ye, J. Liu, B. D. O. Anderson, and M. Cao, "Applications of the Poincaré–Hopf Theorem: Epidemic Models and Lotka–Volterra Systems." [Online]. Available: arXiv:1911.12985

[20] P. Absil and K. Kurdyka, "On the stable equilibrium points of gradient systems," *Systems and Control Letters*, vol. 55, no. 7, pp. 573–577, 2006. MR2225367

[21] H. L. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, vol. 41. American Mathematical Society: Providence, Rhode Island, 2008. MR1319817

[22] J. Ji-Fa, "On the global stability of cooperative systems," *Bulletin of the London Mathematical Society*, vol. 26, no. 5, pp. 455–458, 1994. MR1308362

[23] A. Lajmanovich and J. A. Yorke, "A deterministic model for gonorrhea in a nonhomogeneous population," *Mathematical Biosciences*, vol. 28, no. 3–4, pp. 221–236, 1976. MR0403726

[24] A. Fall, A. Iggidr, G. Sallet, and J.-J. Tewa, "Epidemiological models and Lyapunov functions," *Mathematical Modelling of Natural Phenomena*, vol. 2, no. 1, pp. 62–83, 2007. MR2434863

[25] P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, 2009.

[26] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: a survey of spreading processes on complex networks," *IEEE Control Systems*, vol. 36, no. 1, pp. 26–46, 2016. MR3458715

[27] A. Khanafer, T. Başar, and B. Gharesifard, "Stability of epidemic models over directed graphs: A positive systems approach," *Automatica*, vol. 74, pp. 126–134, 2016. MR3569379

[28] J.-P. Aubin, *Viability Theory*, Systems & Control: Foundations & Applications. Birkhauser, Boston, 1991. MR1134779

[29] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Computer Science and Applied Mathematics, W. Rheinboldt. Academic Press, London, 1979. MR0544666

[30] R. S. Varga, *Matrix Iterative Analysis*, vol. 27. Springer Science & Business Media, 2009. MR1753713

[31] J. Liu, P. E. Paré, A. Nedich, C. Y. Tang, C. L. Beck, and T. Başar, "Analysis and control of a continuous-time bi-virus model," *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 4891–4906, Dec. 2019. MR4044288

[32] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer Science & Business Media, 2009.

[33] B. D. O. Anderson and M. Ye, "Nonlinear mapping convergence and application to social networks," in: *European Control Conference, Limassol*, Cyprus, Jun. 2018, pp. 557–562.

[34] S. Dashkovskiy, B. S. Rüffer, and F. R. Wirth, "Discrete time monotone systems: Criteria for global asymptotic stability and applications," in: *Proc. 17th Inter. Symp. Math. Theory of Networks and Systems* (Kyoto, Japan), 2006, pp. 89–97.

[35] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, "Opinion dynamics and the evolution of social power in influence networks," *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015. MR3376762

[36] M. Ye, J. Liu, B. D. O. Anderson, C. Yu, and T. Başar, "Evolution of social power in social networks with dynamic topology," *IEEE Transaction on Automatic Control*, vol. 63, no. 11, pp. 3793–3808, Nov. 2018. MR3875378

[37] M. A. Khamsi and W. A. Kirk, *An Introduction to Metric Spaces and Fixed Point Theory*. John Wiley & Sons, 2011, vol. 53. MR1818603

[38] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998. MR1632102

Brian D. O. Anderson
Research School of Electrical, Energy and Material Engineering
Australian National University
Acton ACT 2601
Australia
Hangzhou Dianzi University
Hangzhou
China
Data61-CSIRO
Acton, ACT 2601
Australia
E-mail address: brian.anderson@anu.edu.au

Mengbin Ye
Optus–Curtin Centre of Excellence in Artificial Intelligence
Faculty of Science and Engineering
Curtin University
Perth, Western Australia
Australia
E-mail address: mengbin.ye@curtin.edu.au

Received December 6, 2019