Converging towards the optimal path to extinction - supplement

By Ira B. Schwartz[1], Eric Forgoston[2], Simone Bianco[3], and Leah B. Shaw[4]

[1] *Nonlinear Systems Dynamics Section, Plasma Physics Division, U.S. Naval Research Laboratory, Code 6792, Washington, DC 20375, USA*

[2] *Department of Mathematical Sciences, 1 Normal Avenue, Montclair State University, Montclair, NJ 07043, USA*

[3] *Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, 1700 4th Street, San Francisco, CA 94158-2330, USA*

[4] *Department of Applied Science, The College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187-8795, USA*

1. SIR epidemic model

In the Susceptible-Infectious-Recovered (SIR) model, we let $X_1 = S$, $X_2 = I$, and $X_3 = R$ denote the numbers of individuals who are susceptible, infectious, and recovered, respectively, and $\boldsymbol{r}$, a 3-dimensional vector with components having integer increments. The transition rates, $W(\boldsymbol{X};\boldsymbol{r})$, are given by the following:

$$
\begin{aligned}
W(\boldsymbol{X};(1,0,0)) = \mu N \qquad & W(\boldsymbol{X};(-1,0,0)) = \mu X_1 \\
W(\boldsymbol{X};(-1,1,0)) = \beta X_1 X_2 \qquad & W(\boldsymbol{X};(0,-1,1)) = \gamma X_2 \\
W(\boldsymbol{X};(0,-1,0)) = \mu X_2 \qquad & W(\boldsymbol{X};(0,0,-1)) = \mu X_3. \qquad (1.1)
\end{aligned}
$$

The social parameters in equation (1.1) are the birth (death) rate, μ, and contact rate, β, while γ represents the epidemiological recovery rate.

Using the definition of the master equation and the ansatz given by equation (2.1) in the main article, the Hamiltonian is given by

$$
\begin{aligned}
H(\boldsymbol{X},\boldsymbol{p}) &= \mu N(e^{p_1}-1) + \beta X_1 X_2 (e^{-p_1+p_2}-1) + \gamma X_2(e^{-p_2+p_3}-1) \\
& \mu X_1(e^{-p_1}-1) + \mu X_2(e^{-p_2}-1) + \mu X_3(e^{-p_3}-1). \qquad (1.2)
\end{aligned}
$$

The equations of motion describing the optimal path are then given by the ODEs

$$
\begin{aligned}
\dot{\boldsymbol{X}} &= \frac{\partial H}{\partial \boldsymbol{p}}(\boldsymbol{X},\boldsymbol{p}) \\
\dot{\boldsymbol{p}} &= -\frac{\partial H}{\partial \boldsymbol{X}}(\boldsymbol{X},\boldsymbol{p}), \qquad (1.3)
\end{aligned}
$$

where the trajectory satisfies the asymptotic boundary conditions

$$
\begin{aligned}
\lim_{t\to-\infty}(\boldsymbol{X}(t),\boldsymbol{p}(t)) &= (\boldsymbol{X_a},\boldsymbol{p_a}) \\
\lim_{t\to\infty}(\boldsymbol{X}(t),\boldsymbol{p}(t)) &= (\boldsymbol{X_s},\boldsymbol{p_s}), \qquad (1.4)
\end{aligned}
$$

Here the starting point is the endemic zero fluctuational state, given by $\boldsymbol{X_a}/N = [\frac{\mu+\gamma}{\beta}, \frac{\mu(\beta-\mu-\gamma)}{\beta(\mu+\gamma)}, 1 - \frac{\mu+\gamma}{\beta} - \frac{\mu(\beta-\mu-\gamma)}{\beta(\mu+\gamma)}]$ and $\boldsymbol{p}_a = (0, 0, 0)$, and the end point is the stochastic extinct state, given by $\boldsymbol{X}_s/N = [1, 0, 0]$ and $\boldsymbol{p_s} = [0, \ln(\frac{\mu+\gamma}{\beta}), 0]$.

The solution to the boundary value problem given by equations (1.3)-(1.4) is accompanied by the constraint that $H(\boldsymbol{X}, \boldsymbol{p}) = 0$; i.e., the optimal path lies on a surface where the Hamiltonian is zero. Since there is no analytical solution to the problem due its nonlinearities, we solve the problem by employing multi-point shooting methods [Keller, 1976, Schwartz, 1983].

2. Finite-time Lyapunov exponents

We consider a velocity field $\boldsymbol{v} : \mathbb{R}^n \times I \to \mathbb{R}^n$ which is defined over the time interval $I = [t_i, t_f] \subset \mathbb{R}$ and the following system of equations:

$$\dot{\boldsymbol{x}}(t; t_i, \boldsymbol{x}_0) = \boldsymbol{v}(\boldsymbol{x}(t; t_i, \boldsymbol{x}_0), t), \tag{2.1a}$$

$$\boldsymbol{x}(t_i; t_i, \boldsymbol{x}_0) = \boldsymbol{x}_0, \tag{2.1b}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}_0 \in \mathbb{R}^n$, and $t \in I$.

Such a continuous time dynamical system has quantities, known as Lyapunov exponents, which are associated with the trajectory of the system in an infinite time limit. The Lyapunov exponents measure the growth rates of the linearized dynamics about the trajectory. To find the finite-time Lyapunov exponents (FTLE), one computes the Lyapunov exponents on a restricted finite time interval. For the purpose of completeness, we briefly recapitulate the derivation of the FTLE. Details regarding the derivation along with the appropriate smoothness assumptions can be found in Haller [2000, 2001, 2002], Shadden *et al.* [2005], Lekien *et al.* [2007], and Branicki & Wiggins [2010].

The integration of equations (2.1a)-(2.1b) from the initial time t_i to the final time $t_i + T$ yields the flow map $\phi_{t_i}^{t_i+T}$ which is defined as

$$\phi_{t_i}^{t_i+T} : \boldsymbol{x}_0 \mapsto \phi_{t_i}^{t_i+T}(\boldsymbol{x}_0) = \boldsymbol{x}(t_i + T; t_i, \boldsymbol{x}_0). \tag{2.2}$$

Then the FTLE can be defined as

$$\sigma(\boldsymbol{x}, t_i, T) = \frac{1}{|T|} \ln \sqrt{\lambda_{\max}(\Delta)}, \tag{2.3}$$

where $\lambda_{\max}(\Delta)$ is the maximum eigenvalue of the right Cauchy-Green deformation tensor Δ, which is given as

$$\Delta(\boldsymbol{x}, t_i, T) = \left(\frac{d\phi_{t_i}^{t_i+T}(\boldsymbol{x}(t))}{d\boldsymbol{x}(t)} \right)^* \left(\frac{d\phi_{t_i}^{t_i+T}(\boldsymbol{x}(t))}{d\boldsymbol{x}(t)} \right), \tag{2.4}$$

with * denoting the adjoint.

For a given $\boldsymbol{x} \in \mathbb{R}^n$ at an initial time t_i, equation (2.3) gives the maximum finite-time Lyapunov exponent for some finite integration time T (forward or backward) and provides a measure of the sensitivity of a trajectory to small perturbations.

For such FTLE fields $\sigma(\boldsymbol{x}, t_i, T)$, the "ridges" of the field indicate the location of attracting (backward time FTLE field) and repelling (forward time FTLE field) structures. In 2D, these structures are curves which locally maximize the FTLE field so that transverse to these curves one finds the FTLE to be a local maximum.

3. Local linear variation near the optimal path

The main heuristic argument of this section is to show the equivalence between the path which optimizes the probability of extinction and the ridge upon which the finite time Lyapunov exponent (FTLE) is optimized. Although we consider the dynamical system in one dimension, the arguments carry over to arbitrary dimensions.

We begin by demonstrating that given an optimal path, then the FTLE attains its maximum values on the optimal path. From the Hamiltonian or Lagrangian equations of motion, the process that leads to extinction consists of a trajectory that emanates from an endemic steady state x_a and approaches the extinct state x_s. Since the endemic and extinct states are both regular saddles (or unstable foci) in the variational formulations, they both have hyperbolic structure. Moreover, every point along the trajectory connecting the the two states as $t \to \pm\infty$ also possesses a local hyperbolic structure. As an example, consider the Langevin problem having a scalar vector field of position $V(x)$, which has Lagrangian $L(x, \dot{x}) = (\dot{x} - V(x))^2/2$ to describe the action. Converting to a Hamiltonian formulation leads one to the following equations of motion:

$$\dot{x} = p + V(x), \tag{3.1a}$$

$$\dot{p} = -pV'(x), \tag{3.1b}$$

$$H(x, p) = \frac{p^2}{2} + pV(x). \tag{3.1c}$$

It is immediate from equations (3.1a)-(3.1c) that $p = 0$ is an invariant manifold. In addition, the optimal path must lie along the $H = 0$ surface, which means that in addition to the $p = 0$ manifold, the zero surface includes $p = -2V(x)$.

To clarify the direction along the optimal path, we make the following assumptions regarding $V(x)$:

1. $V(x)$ is smooth,

2. $V(x_a) = V(x_s) = 0$,

3. $V'(x_a) < 0$, and $V'(x_s) > 0$.

Items 2 and 3 imply that x_a is an attracting steady state and x_s is a repelling steady state in the 1D deterministic dynamical system. We now assume that the optimal path must satisfy $\lim_{t\to+\infty}(x(t), p(t)) = (x_s, 0)$, while $\lim_{t\to-\infty}(x(t), p(t)) = (x_a, 0)$. Since $H(x(t), p(t)) = 0$, the limits provide direction along the optimal path.

The optimal path lies on the curve $p = -2V(x)$, and $p = 0$ corresponds to the zero fluctuation case. We shift the optimal path to the origin by using the following transformation:

$$u = x, \tag{3.2a}$$

$$w = p + 2V(x), \tag{3.2b}$$

$$\hat{H}(u, w) = \frac{w^2}{2} - wV(u). \tag{3.2c}$$

The new equations of motion are now:

$$\dot{u} = \partial\hat{H}/\partial w = w - V(u), \tag{3.3a}$$

$$\dot{w} = -\partial\hat{H}/\partial u = wV'(u). \tag{3.3b}$$

The optimal path now corresponds to $w = 0$, and the zero fluctuation case, $p = 0$, corresponds to $w = 2V(u)$.

The linearized variation along the optimal path is given by the following matrix initial value problem from equations (3.3a)-(3.3b):

$$\dot{\boldsymbol{X}} = \begin{bmatrix} -V'(u(t)) & 1 \\ 0 & V'(u(t)) \end{bmatrix} \boldsymbol{X}, \quad \boldsymbol{X}(0) = \boldsymbol{I}. \tag{3.4}$$

Equation (3.4) is hyperbolic as long as $V'(u(t) \neq 0$ since the local eigenvalues are given by $\pm V'(u(t))$. Therefore, there exists a time-dependent transformation which recasts equation (3.4) into a diagonal system. Assuming the solution changes slowly and the domain about the saddle structure is small, we illustrate our point that the FTLE takes its maximum on the optimal path locally by considering the following linear system:

$$\dot{\boldsymbol{X}} = \begin{bmatrix} -\alpha & 0 \\ 0 & \alpha \end{bmatrix} \boldsymbol{X}, \quad \boldsymbol{X}(0) = \boldsymbol{I}, \tag{3.5}$$

and for any initial value, the particular solution is $\boldsymbol{x}_p(t; \boldsymbol{x}_0) = (x_1(t), x_2(t)) = (e^{-\alpha t} x_{10}, e^{\alpha t} x_{20})$.

To show that the FTLE takes it maximum along the path, we notice that any point along the path is hyperbolic with a saddle structure. Therefore, we consider an arbitrary initial condition lying within a small domain containing the origin. Since almost any initial condition hits the boundary of the domain in finite time due to the saddle structure of the origin, we use the escape time as the final time for the FTLE. The definition we use of the FTLE is the direct comparison of the distance between two close trajectories as follows:

$$\sigma(t; \boldsymbol{x}_0) = \frac{1}{2t} \log\left(||\boldsymbol{x}_p(t; \boldsymbol{x}_0 + \boldsymbol{\epsilon}) - \boldsymbol{x}_p(t; \boldsymbol{x}_0)||\right), \tag{3.6}$$

where $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2)$.

Rescaling the domain to be $D = [-1, 1] \times [-1, 1]$, clearly any point not on the invariant manifolds $x_1 = 0$, and $x_2 = 0$ will escape in the x_2 direction. We assume $\boldsymbol{x}_0 = (x_{10}, \delta)$, where $0 < \delta << 1$. Then the time to escape for an arbitrary non-zero initial condition is given by

$$t_f = -\frac{\log(\delta)}{\alpha}, \tag{3.7}$$

and the FTLE using equation (3.6) yields

$$\sigma(t, \boldsymbol{x}_0) = \frac{-\alpha \log(\delta^2 \epsilon_1^2 + \epsilon_2^2/\delta^2)}{2 \log \delta}. \tag{3.8}$$

For δ small and positive and $t = t_f$, we also find that

$$\frac{\partial \sigma(t_f; \boldsymbol{x}_0(\delta))}{\partial \delta} = \frac{\alpha \log\left(\epsilon_2^2\right)}{2\left(\log(\delta)\right)^2 \delta} + O(\delta^3), \tag{3.9}$$

which can be shown to be negative assuming $\epsilon_2 < 1$. A similar argument holds for $\delta < 0$. Therefore, the FTLE as a function of distance to the stable invariant manifold is a decreasing function, and thus takes it maximum values on the manifold.

To complete the argument, we now demonstrate that a FTLE ridge of (local) maximal values corresponds to the optimal path. We begin with the mild assumption that the FTLE ridge connects the endemic state to the extinct state. The endemic and extinct states are equilibria of the Hamiltonian, so that $H = 0$. In addition, the FTLE ridge is invariant. Therefore, the ridge is a zero-energy orbit of the Hamiltonian H, and thus determines the optimal path to extinction.

4. SIS epidemic model

In the Susceptible-Infectious-Susceptible (SIS) epidemic model, the population consists of susceptible individuals and infectious individuals. The population is driven via the birth rate μ, which is also equal to the death rate. The changes in increments are given by $\boldsymbol{r} = (r_1, r_2)$, where r_1 and r_2 can take the values of -1, 0, or 1. The transition rates $W(\boldsymbol{X}, \boldsymbol{r})$ for $\boldsymbol{X} = (s, i)$, where s and i are respectively the total number of susceptible and infectious individuals, are given as: $W\big(\mathbf{X}; (1, 0)\big) = N\mu$, $W\big(\mathbf{X}; (-1, 0)\big) = \mu X_1$, $W\big(\mathbf{X}; (0, -1)\big) = \mu X_2$, $W\big(\mathbf{X}; (1, -1)\big) = \gamma X_2$, and $W\big(\mathbf{X}; (-1, 1)\big) = \beta X_1 X_2 / N$, where β is the mass action contact rate, γ is the recovery rate, and N is a parameter for the average size of the population. The transition rates give the probability of change in one iterate in time of the system. For example, $W(\boldsymbol{X},(1,0)$ gives an increase in susceptible individuals due to births at a rate $N\mu$, while $W\big(\mathbf{X};(-1,1)\big)$ generates an increase in infectious individuals and a decrease in susceptible individuals due to an effective infectious contact.

For large N, random internal fluctuations of s and i are small on average. If these fluctuations are disregarded, one arrives at the following well-known deterministic mean-field equations for the SIS model:

$$\dot{X}_1 = N\mu - \mu X_1 + \gamma X_2 - \beta X_1 X_2/N, \tag{4.1a}$$

$$\dot{X}_2 = -(\mu+\gamma)X_2 + \beta X_1 X_2/N. \tag{4.1b}$$

For the reproductive rate of infection defined by parameter $R_0 = \beta/(\mu+\gamma) > 1$, equations (4.1a)-(4.1b) have a solution $\boldsymbol{X}_A = N\boldsymbol{x}_A$ with $x_{1A} = R_0^{-1}$ (susceptible), and $x_{2A} = 1 - R_0^{-1}$ (infectious), which describes the endemic disease. In addition, equations (4.1a)-(4.1b) have a stationary state given by $\boldsymbol{X}_S = N\boldsymbol{x}_S$ with $x_{1S} = 1$, and $x_{2S} = 0$, which corresponds to the extinct, or disease-free state. Without any fluctuations, extinction cannot occur since the epidemic state is stable and the extinct state is unstable. The inclusion of noise will drive components of the population to extinction. This is due to an effective mechanistic force $\boldsymbol{p}$, which is found by solving the equations of motion for the Hamiltonian.

5. Derivation of scaling law for SIS model with internal fluctuations

The Hamiltonian governing the probability density function (see Sec. 4(c) of the main article) is given by

$$H(I,p) = (\mu+\gamma)I(e^{-p}-1) + \beta I(1-I)(e^{p}-1). \tag{5.1}$$

For parameters in which the endemic and extinct states are close together, we assume that both I and p are small, and that $0 < R_0 - 1 = \epsilon << 1$, where $R_0 = \beta/(\mu+\gamma)$. We can therefore scale I and p by ϵ and consider a new Hamiltonian in terms of the scaled variables. Expanding this new Hamiltonian in a Taylor series about $\epsilon = 0$ to third-order yields

$$H(\epsilon x, \epsilon y) = xy(\mu+\gamma)\left(R_0-1\right)\epsilon^2 + xy(\mu+\gamma)\left(\frac{y}{2} + \frac{R_0 y}{2} - R_0 x\right)\epsilon^3 + \mathcal{O}(\epsilon^4). \tag{5.2}$$

Hamilton's equations are therefore given as

$$\dot{x} =(\mu+\gamma)x\left(R_0 - 1 - R_0 x\epsilon + y\epsilon + R_0 y\epsilon\right), \tag{5.3a}$$

$$\dot{y} =-\frac{1}{2}y(\mu+\gamma)\left(2R_0 - 2 - 4R_0 x\epsilon + y\epsilon + R_0 y\epsilon\right). \tag{5.3b}$$

The action as a function of R_0 can then be found to be

$$S(R_0) = \int_{-\infty}^{\infty} y(t)\dot{x}(t)dt = \frac{\left(R_0-1\right)^2}{R_0\ \left(1+R_0\right)}. \tag{5.4}$$

References

Branicki, M. & Wiggins, S. 2010 Finite-time Lagrangian transport analysis: Stable and unstable manifolds of hyperbolic trajectories and finite-time Lyapunov exponents. *Nonlinear Proc. Geoph.*, **17**, 1–36.

Haller, G. 2000 Finding finite-time invariant manifolds in two-dimensional velocity fields. *Chaos*, **10**(1), 99–108.

Haller, G. 2001 Distinguished material surfaces and coherent structures in three-dimensional fluid flows. *Physica D*, **149**, 248–277.

Haller, G. 2002 Lagrangian coherent structures from approximate velocity data. *Phys. Fluids*, **14**(6), 1851–1861.

Lekien, F., Shadden, S. C. & Marsden, J. E. 2007 Lagrangian coherent structures in n-dimensional systems. *J. Math. Phys.*, **48**, 065 404.

Shadden, S. C., Lekien, F. & Marsden, J. E. 2005 Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D*, **212**, 271–304.

Keller, H. 1976 *Numerical Solution of Two Point Boundary Problems*. CBMS-NSF.

Schwartz, I.B. 1983 Estimating regions of existence of unstable periodic orbits using computer-based techniques . *SIAM J. Num. Anal.* **20**, 106–120.