# Information Geometry in Epidemiology and Population Dynamics

Stefan Hohenegger

(IP2I Lyon, Università di Napoli Federico II)

QTC Journal Club

(19/April/2024)

Based on: 2402.16390

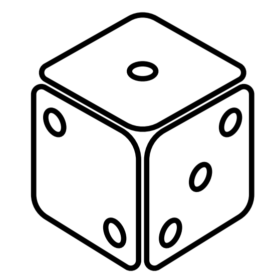
with B. Filoche and F. Sannino

# Information Geometry

Probability distribution on a discrete set $\mathbb{V}$

$$p : \mathbb{V} \longrightarrow [0,1] \qquad \text{such that} \qquad \sum_{X \in \mathbb{V}} p(X) = 1$$

Example: (faire) dice



$$\mathbb{V} = \left\{ \text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅} \right\} \cong \{1,2,3,4,5,6\}$$

with the map

$$p(\text{⚀}) = \frac{1}{6}\,, \qquad p(\text{⚁}) = \frac{1}{6}\,, \qquad p(\text{⚂}) = \frac{1}{6}\,,$$

$$p(\text{⚃}) = \frac{1}{6}\,, \qquad p(\text{⚄}) = \frac{1}{6}\,, \qquad p(\text{⚅}) = \frac{1}{6}$$

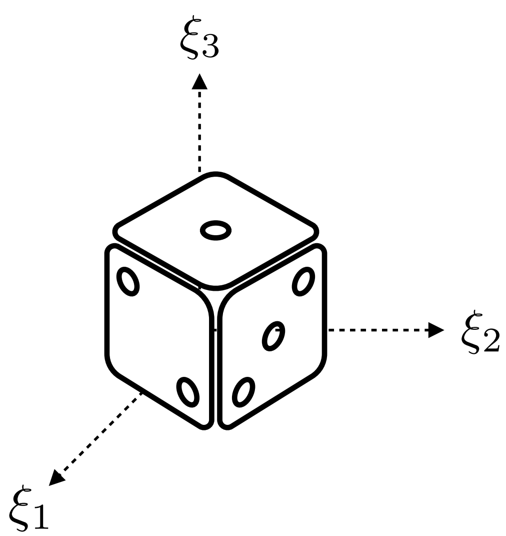
# Allow probabilities to depend on (cont.) parameters $(\xi^1, \dots, \xi^d) \in \Xi \subset \mathbb{R}^d$

$$p : \mathbb{V} \times \Xi \longrightarrow [0,1]$$

such that

$$\sum_{X \in \mathbb{V}} p(X, \xi) = 1 \ \ \forall \xi \in \Xi$$

## Example: loaded dice




shift center of mass of the dice by a vector

$$\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$$

Changes probabilities in a non-trivial fashion:

$$p(\text{⚀}, \xi) = p_1(\xi)\,, \qquad p(\text{⚁}, \xi) = p_2(\xi)\,, \qquad p(\text{⚂}, \xi) = p_3(\xi)\,,$$

$$p(\text{⚃}, \xi) = p_4(\xi)\,, \qquad p(\text{⚄}, \xi) = p_5(\xi)\,, \qquad p(\text{⚅}, \xi) = p_6(\xi)$$

which still satisfy $p_1(\xi) + p_2(\xi) + p_3(\xi) + p_4(\xi) + p_5(\xi) + p_6(\xi) = 1$

# Fisher information metric

[Fisher 1922]
[Rao 1945]
[Jeffreys 1946]

$$g_{ij}(\xi) := \sum_{X \in \mathbb{V}} \left( \frac{\partial}{\partial \xi^i} \log p(X, \xi) \right) \left( \frac{\partial}{\partial \xi^j} \log p(X, \xi) \right) p(X, \xi) \qquad \forall i, j \in \{1, \ldots, d\}$$

defines the metric of a Riemannian manifold

[Amari, Nagaoka 2020]

Fisher-Rao distance: $d(\xi_1, \xi_2) = \int_0^1 d\tau \sqrt{\dot{\xi}_c^i \, \dot{\xi}_c^j \, g_{ij}(\xi_c)}$

geodesic $\xi_c : [0, 1] \to \Xi$
with $\xi_c(0) = \xi_1$ and $\xi_c(1) = \xi_2$

special case: 1 dimension and $\Xi = \mathbb{R}$

$$g_{tt}(t) = \sum_{X \in \mathbb{V}} \left( \frac{\partial \log p(X, t)}{\partial t} \right)^2 p(X, t) \quad \text{and} \quad d(t_1, t_2) = \left| \int_{t_1}^{t_2} \sqrt{g_{tt}(t')} \, dt' \right|$$

# Towards Dynamics

$$g_{tt}(t) = \sum_{X\in\mathbb{V}} \left(\frac{\partial \log p(X,t)}{\partial t}\right)^2 p(X,t)$$

interpret $t \in \mathbb{R}$ as time variable

Resolve normalisation condition (denote cardinality of $\mathbb{V}$ by $N+1$)

$$p_i(t) = x_i(t) \qquad \forall i \in \{1,\ldots,N\}$$

$$p_{N+1}(t) = 1 - \sum_{i=1}^{N} x_i(t)$$

Fisher metric depends on time derivatives $\dot{x}_i(t)$

$$g_{tt} = \sum_{i=1}^{N} \frac{\left(1 - \sum_{j\neq i}^{N} x_j\right)\dot{x}_i^2}{x_i\left(1 - \sum_{j=1}^{N} x_j\right)} + 2\sum_{1\leq i<j\leq N} \frac{\dot{x}_i\dot{x}_j}{1 - \sum_{k=1}^{N} x_k}$$

$$g_{tt} = \sum_{i=1}^{N} \frac{\left(1 - \sum_{j\neq i}^{N} x_j\right) \dot{x}_i^2}{x_i \left(1 - \sum_{j=1}^{N} x_j\right)} + 2 \sum_{1\leq i<j\leq N} \frac{\dot{x}_i \dot{x}_j}{1 - \sum_{k=1}^{N} x_k}$$

Can be simplified in certain cases, e.g.:

$$|\dot{x}_1| \gg |\dot{x}_i| \sim 0 \qquad \forall i \in \{2, \ldots, N\}$$

'one probability is strongly growing at the expense of another'

define: $\sum_{i=2}^{N} x_i(t) \sim 1 - \alpha = \text{const.}$

In this case the expression for the metric simplifies

$$g_{tt} \sim \frac{\left(1 - \sum_{j=2}^{N} x_j\right)}{x_1 \left(1 - \sum_{j=1}^{N} x_j\right)} \dot{x}_1^2 + \mathcal{O}(\dot{x}_1) = \frac{\alpha\, \dot{x}_1^2}{x_1(\alpha - x_1)} + \mathcal{O}(\dot{x}_1)$$

$$g_{tt} = \frac{\alpha\, \dot{x}_1^2}{x_1(\alpha - x_1)} + \mathcal{O}(\dot{x}_1)$$

Can be resolved to give a flow equation for $\dot{x}_1(t)$

$$\frac{dx_1}{dt} \sim \pm\sqrt{g_{tt}\, x_1 \left(1 - \frac{x_1}{\alpha}\right)} \qquad \text{for} \qquad x_1 \in [0, \alpha]$$

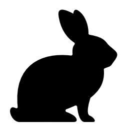
A priori only a re-writing of the definition of $g_{tt}$

Time dependence of metric in principle encodes dynamics of $x_1$

Is it possible to re-write $g_{tt}$ as a function of $x_1$?

# First Example: Lotka-Volterra Equation

Describes the time-evolution of a population of predators and prey

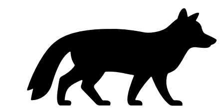


prey

$\xi_1(t)$

- have unlimited resources
- limited only through the pressure by the predators
  (exponential growth in absence of predators)

Predators eat prey



predators

$\xi_2(t)$

- require prey for survival
  (exponential decline in absence of prey)

Coupled differential equations:

[Lotka, 1909]
[Volterra 1926]

$$\frac{d\xi_1}{dt} = a_1\,\xi_1 - b_1\,\xi_1\,\xi_2\,,$$

$$\frac{d\xi_2}{dt} = -a_2\,\xi_2 + b_2\,\xi_1\,\xi_2\,,$$

with $a_{1,2}\,, b_{1,2} \in \mathbb{R}_+$

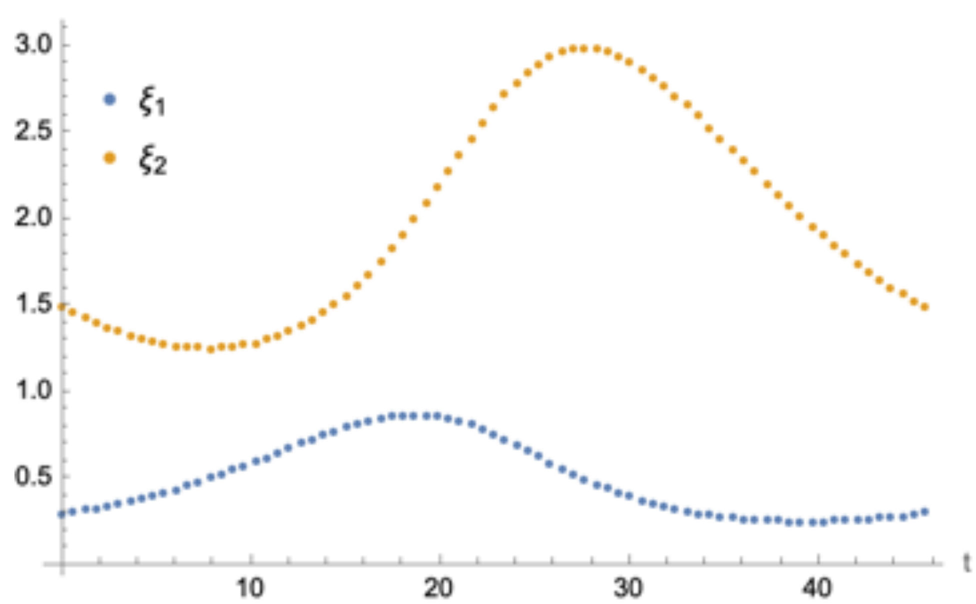
(and initial conditions $\xi_{1,2}(t=0)$)

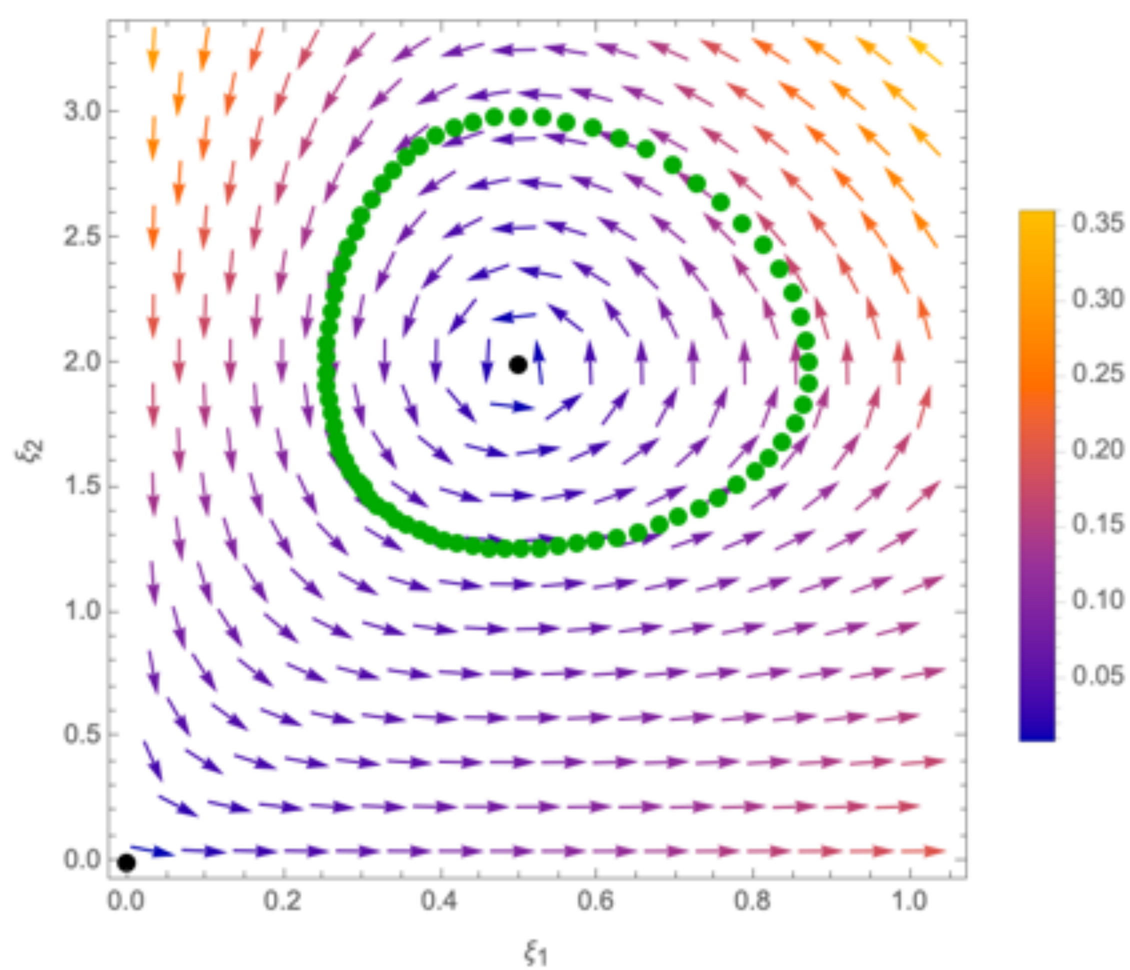
Total population in general not conserved: $\frac{d}{dt}(\xi_1 + \xi_2) \neq 0$

Two fixed points such that $\frac{d\xi_{1,2}}{dt} = 0$

$$(\xi_1, \xi_2) = (0,0) \quad \text{and} \quad (\xi_1, \xi_2) = \left(\frac{a_2}{b_2}, \frac{a_1}{b_1}\right)$$

Away from these fixed points: periodic solutions



# Introduce probability distribution

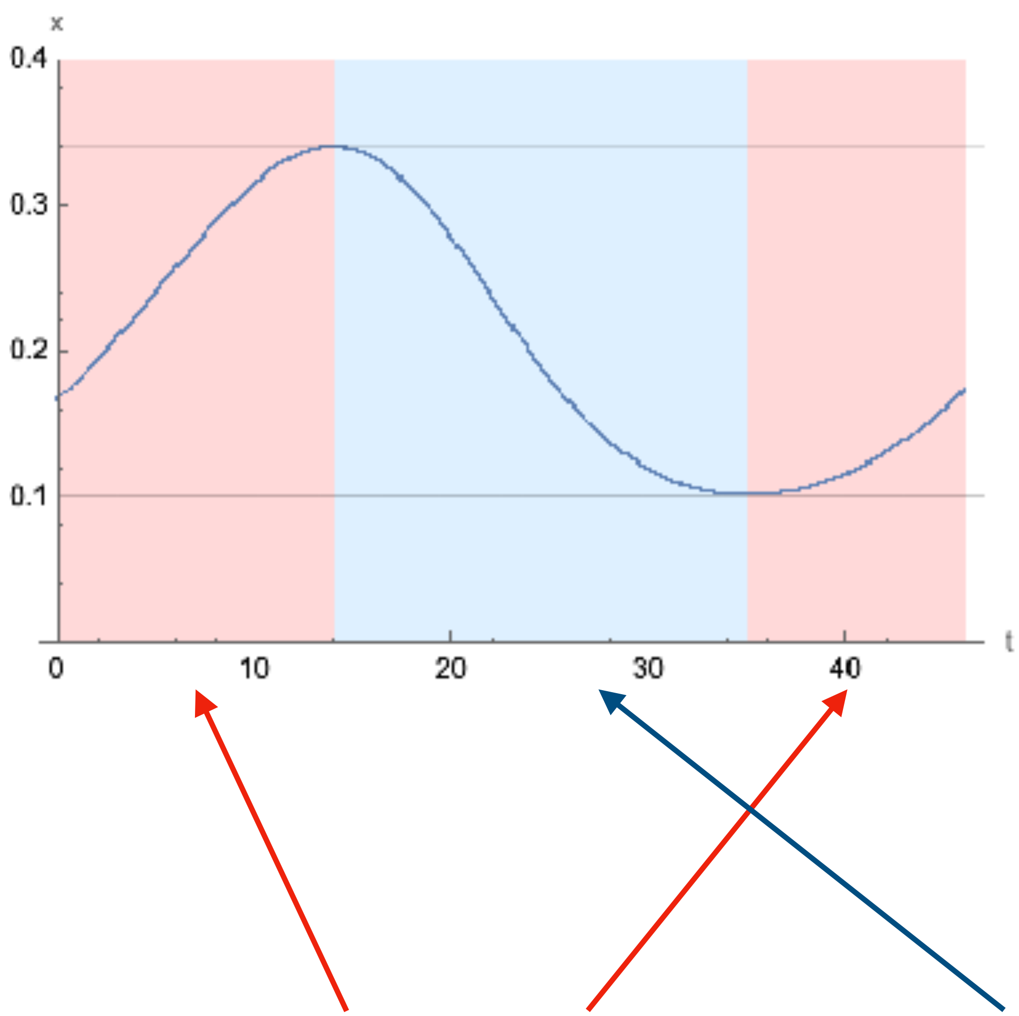
$$p_1 = \frac{\xi_1}{\xi_1 + \xi_2} =: x$$

$$p_2 = \frac{\xi_2}{\xi_1 + \xi_2} = 1 - x$$

with

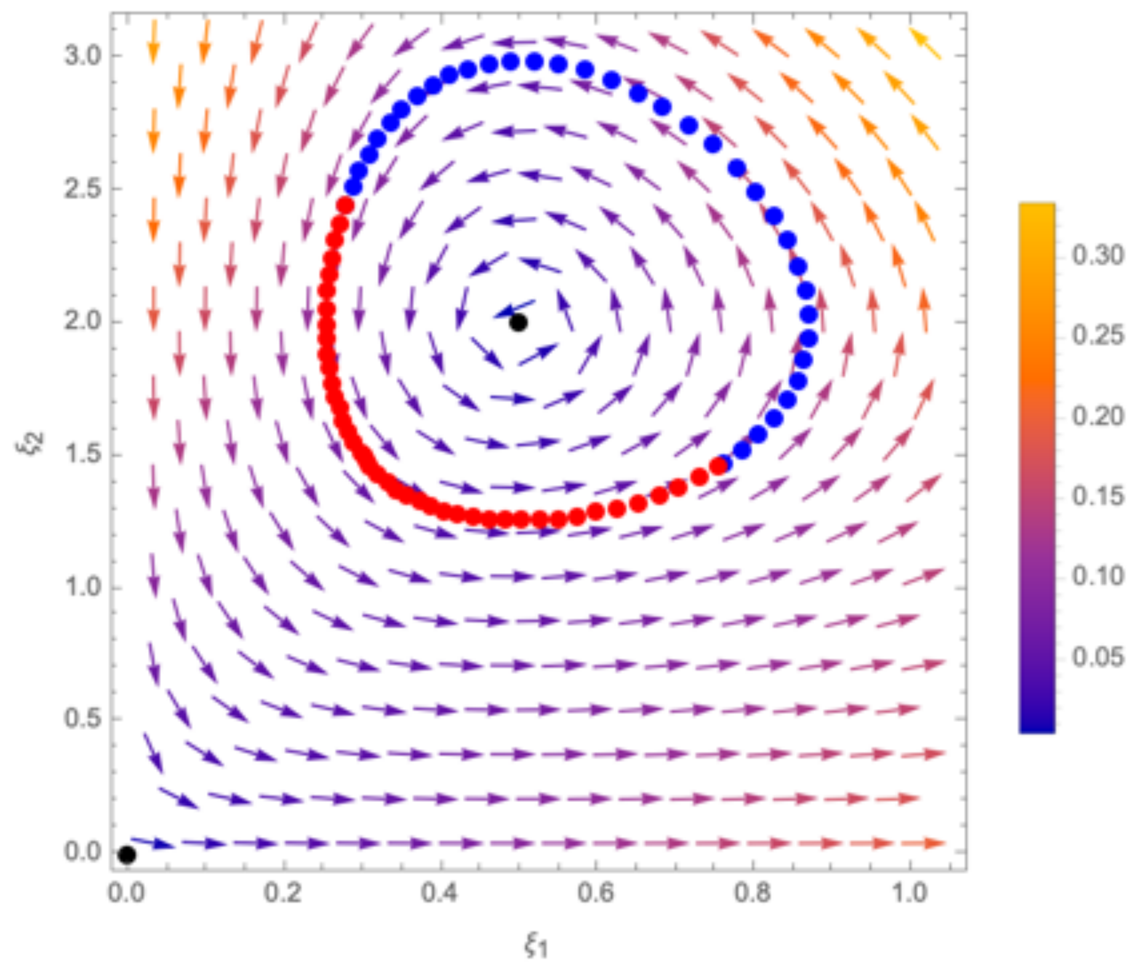
$$x : \mathbb{R} \longrightarrow [0, 1]$$

single degree of freedom

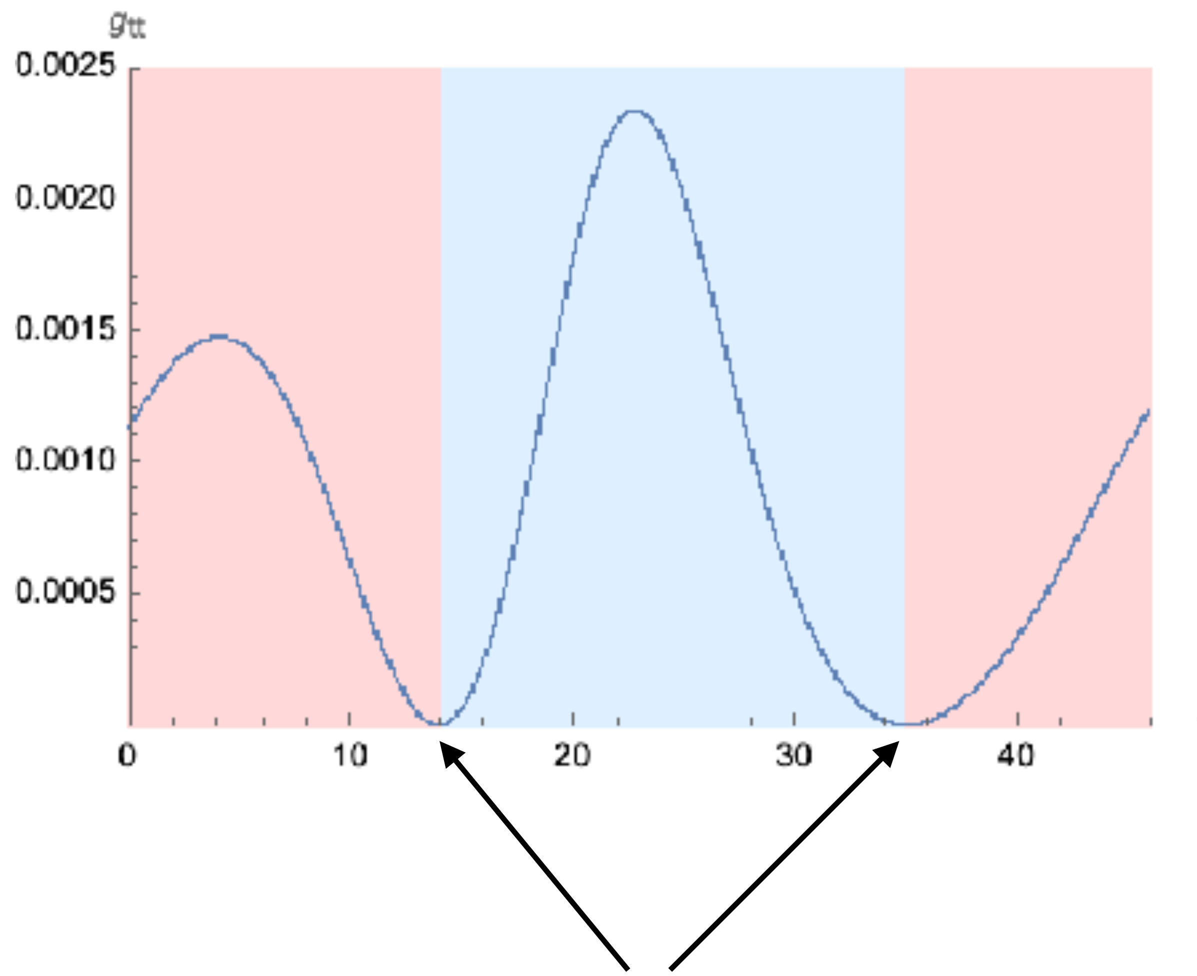



fraction of prey increasing

fraction of prey decreasing



$$g_{tt} = \left(\partial_t \log(p_1)\right)^2 p_1 + \left(\partial_t \log(p_2)\right)^2 p_2 = \frac{\xi_1 \xi_2 (a_1 + a_2 - b_2\xi_1 - b_1\xi_2)^2}{(\xi_1+\xi_2)^2}$$




Zeroes of the metric correspond to extrema of $x$

Metric is a function of time: re-write it in terms of $x$:

$$g_{tt} = \frac{1-x}{x}\,\left(a_1 x + a_2 x - (b_1(1-x) + b_2 x)\xi_1\right)^2$$

requires solving the dynamics of the system

$$\frac{d\xi_1}{dt} = a_1\,\xi_1 - b_1\,\xi_1\,\xi_2\,,$$

$$\frac{d\xi_2}{dt} = -a_2\,\xi_2 + b_2\,\xi_1\,\xi_2\,,$$

Integrating the equations gives an algebraic equation:

$$\left(b_2 + b_1\frac{1-x}{x}\right)\xi_1 - (a_1+a_2)\,\log\xi_1 - a_1\log\left(\frac{1-x}{x}\right) = L$$

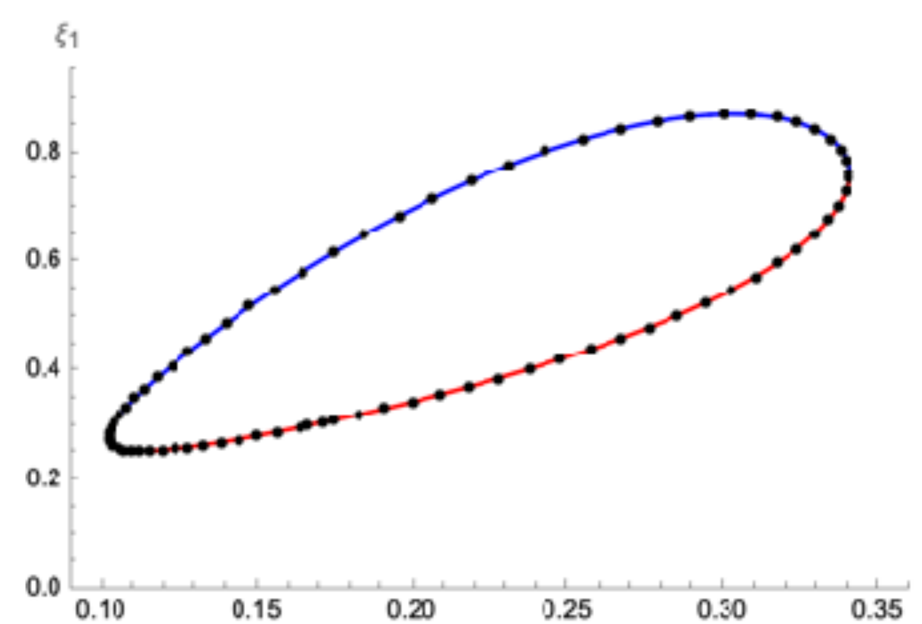
fixed by initial conditions:

$$L = b_2\xi_1(0) - a_2\log\xi_1(0) + b_1\xi_2(0) - a_1\log\xi_2(0)$$

leads to

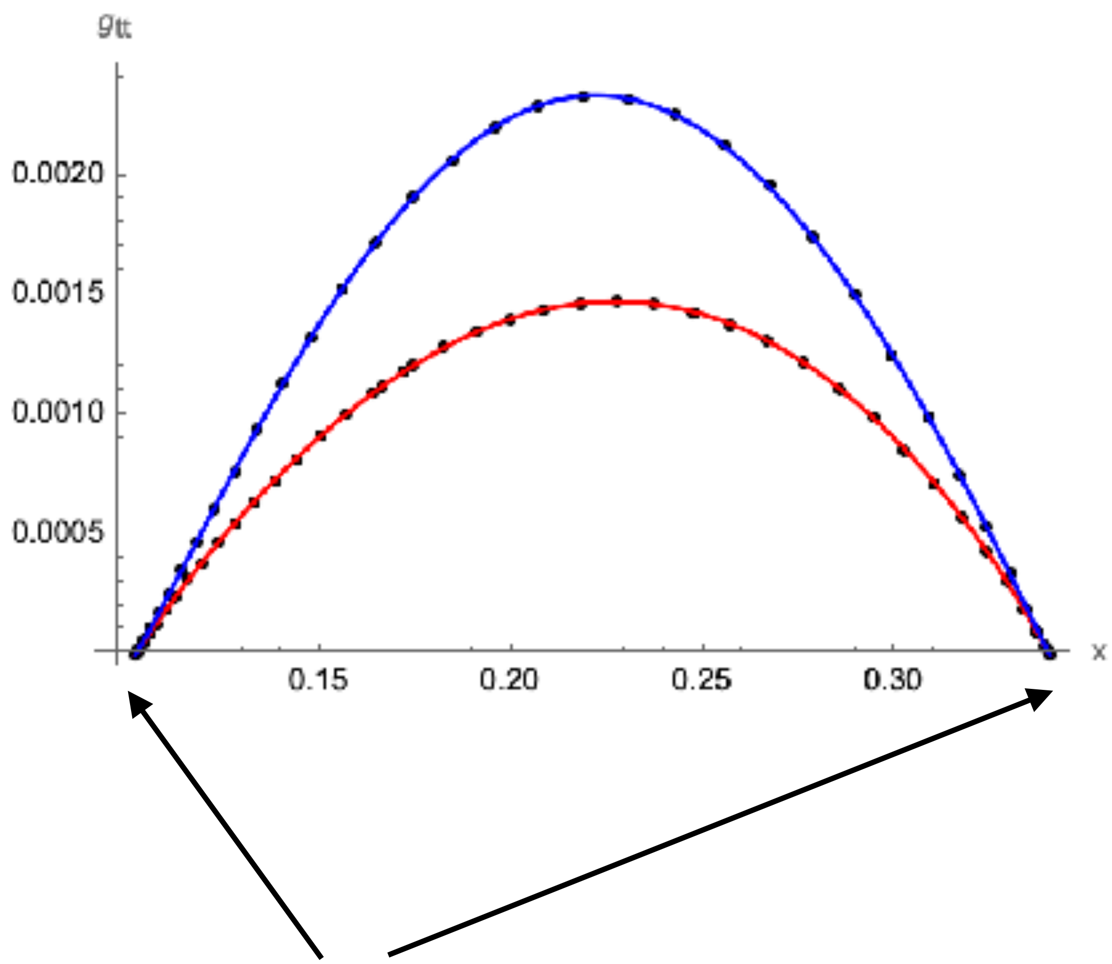
$$\xi_1 = -\frac{(a_1+a_2)x}{b_1(1-x)+b_2x}\,W\left(-\frac{e^{-\frac{L}{a_1+a_2}}\left(\frac{1-x}{x}\right)^{-\frac{a_1}{a_1+a_2}}(b_1(1-x)+b_2x)}{x(a_1+a_2)}\right)$$

Lambert function with 2 real branches



# Fisher information metric as function of $x$

$$g_{tt} = (a_1+a_2)^2\, x\,(1-x)\left[1+W\left(-\frac{e^{-\frac{L}{a_1+a_2}}\left(\frac{1-x}{x}\right)^{-\frac{a_1}{a_1+a_2}}(b_1(1-x)+b_2 x)}{x(a_1+a_2)}\right)\right]^2$$




Two branches:

$x$ growing: $\frac{dx}{dt} > 0$

$x$ declining: $\frac{dx}{dt} < 0$

both branches meet at the same zeroes $x_{1,2}$
which are solutions of

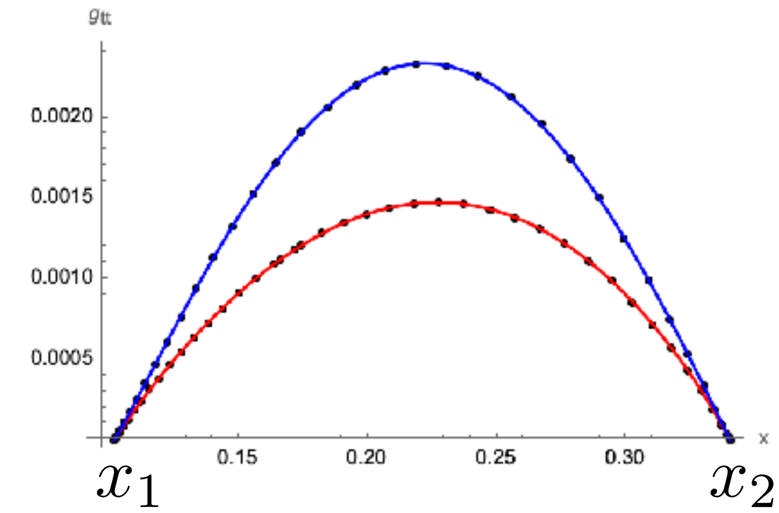
$$b_1\left(\frac{1-x}{x}\right)^{\frac{a_2}{a_1+a_2}} + b_2\left(\frac{1-x}{x}\right)^{-\frac{a_1}{a_1+a_2}} = (a_1+a_2)\, e^{\frac{L}{a_1+a_2}-1}$$

periodicity

Metric can be very well approximated by

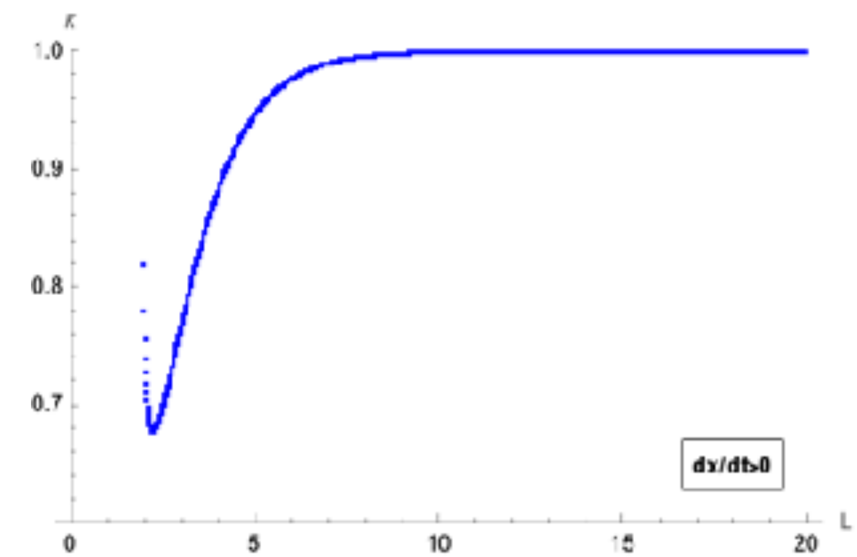
$$g_{tt}(x) = a\,(x - x_1)^b\,(x_2 - x)^c \qquad \forall x \in [x_1, x_2]$$

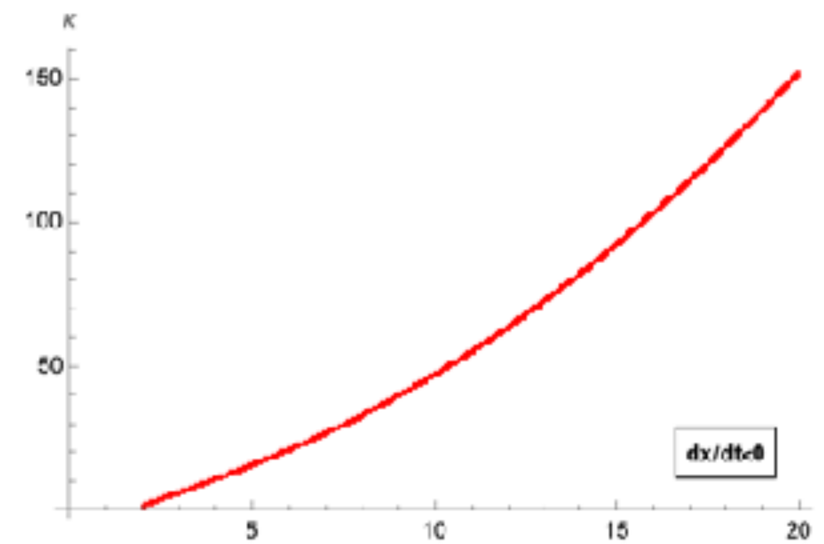
suitable fitting parameters

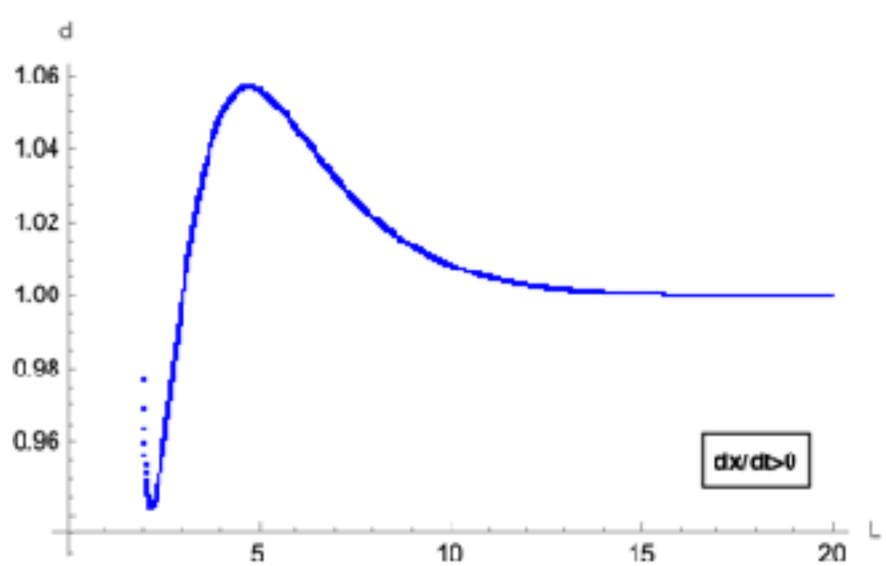


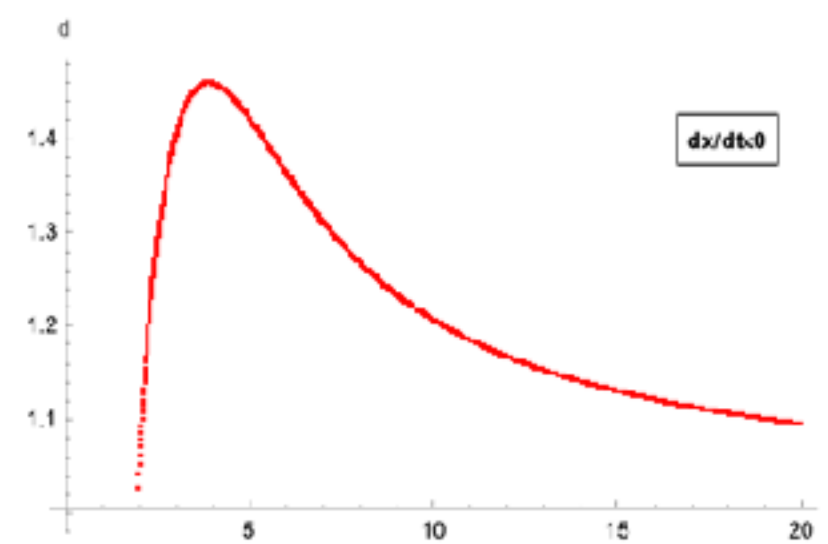

Simple example: $a_1 = a_2 = b_1 = b_2 = 1$

$$g_{tt} \sim 4\kappa\,(x - x_1)^d\,(x_2 - x)^d \qquad x_{1,2} = \frac{1}{2}\left(1 \mp \sqrt{1 - e^{2-L}}\right) \qquad \forall L > 2$$



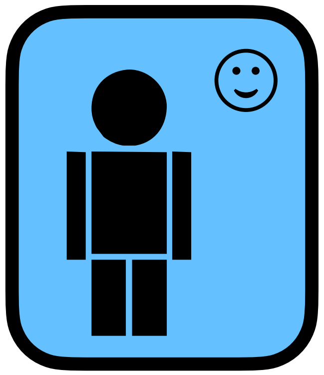



# Second Example: SIR(S) Model

Oldest mathematical models in epidemiology

[Ross, 1916]
[Ross, Hudson 1917]
[Kermack, McKendrick 1927]

Idea: group members of the population into compartments

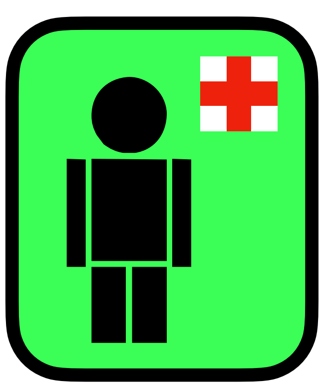


Susceptible: not infectious; can become infectious if in contact with the pathogen

$S(t)$

Infectious: infected with the pathogen; can actively transmit the pathogen to a susceptible individual

$I(t)$

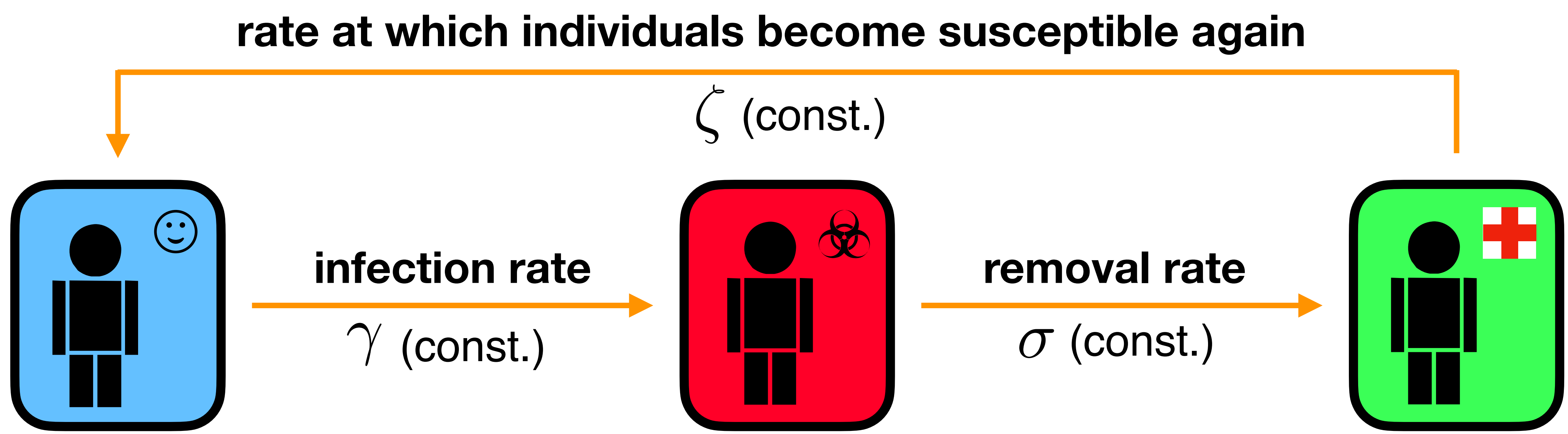


Removed: can neither be infected themselves, nor infect other individuals

$R(t)$

Population conserved: $(S + I + R)(t) = 1 \qquad \forall t \in \mathbb{R}$

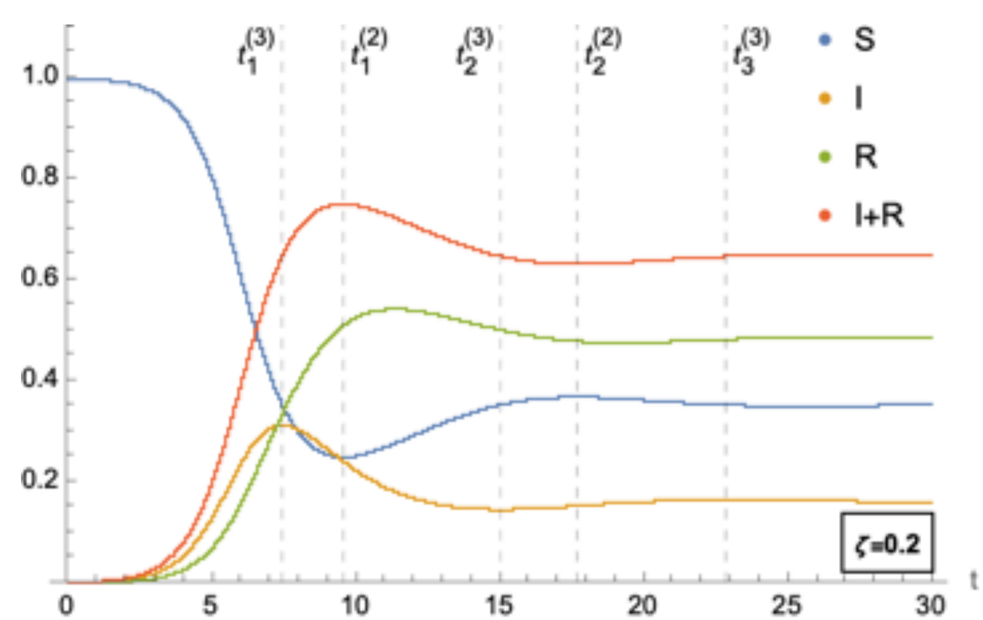
# Equations for the time evolution of the number of individuals:

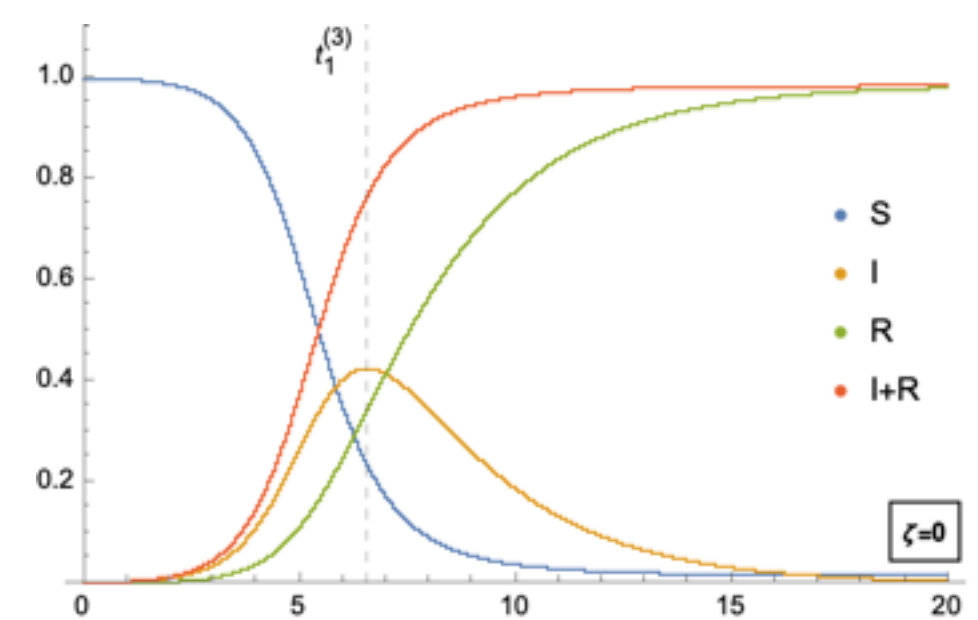



$$\frac{dS}{dt} = -\gamma\, S\, I + \zeta\, R$$

$$\frac{dI}{dt} = \gamma\, S\, I - \sigma I$$

$$\frac{dR}{dt} = \sigma\, I - \zeta R$$



# Different ways to define probability distributions:

$$p^{(1)} : \{1,2,3\} \times \mathbb{R} \to [0,1] \quad \text{with} \quad p_1^{(1)}(t) = S(t), \quad p_2^{(1)}(t) = I(t), \quad p_3^{(1)}(t) = R(t)$$

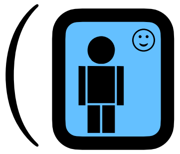
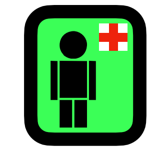
$$p^{(2)} : \{1,2\} \times \mathbb{R} \longrightarrow [0,1] \quad \text{with} \quad p_1^{(2)}(t) = S(t), \quad p_2^{(2)}(t) = (I+R)(t)$$

$$p^{(3)} : \{1,2\} \times \mathbb{R} \longrightarrow [0,1] \quad \text{with} \quad p_1^{(3)}(t) = (S+R)(t), \quad p_2^{(3)}(t) = I(t)$$

Probability distributions with only one degree of freedom

$\zeta = 0$

$p^{(2)} = ($  ,  $+$  $)$

$p^{(3)} = ($  $+$  ,  $)$

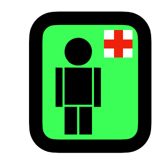
## Fisher information metric:

$$g_{tt}^{(2)}(t) = \frac{\gamma^2 I^2 S}{(I+R)}$$

$$g_{tt}^{(3)}(t) = \frac{I(\gamma S - \sigma)^2}{R+S}$$

## Single degree of freedom:

$x(t) = (I+R)(t) =$  $+$ 

$x(t) = (S+R)(t) =$  $+$ 

## Resolve dynamics:

initial conditions
$x_0 = x(t=0)$

$$I = x + \frac{\sigma}{\gamma} \log\left(\frac{1-x}{1-x_0}\right)$$

$$S = -\frac{\sigma}{\gamma} W\left(-\frac{x_0\gamma}{\sigma}\, e^{-\frac{x\gamma}{\sigma}}\right)$$

two real branches (two solutions)

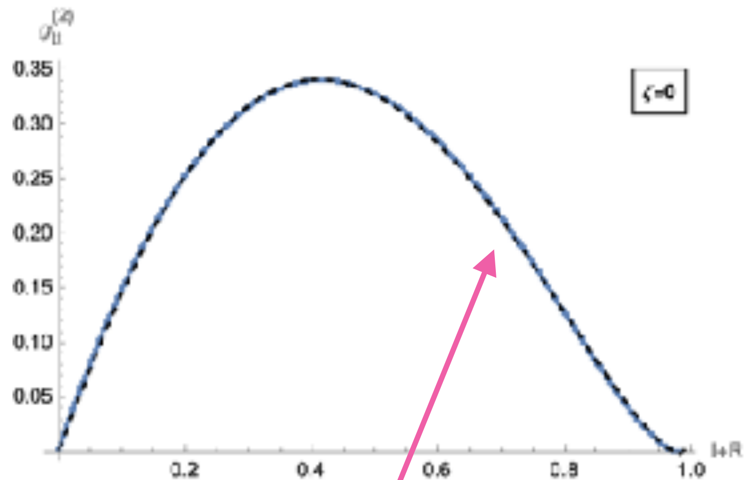
## Metric as a function of $x$

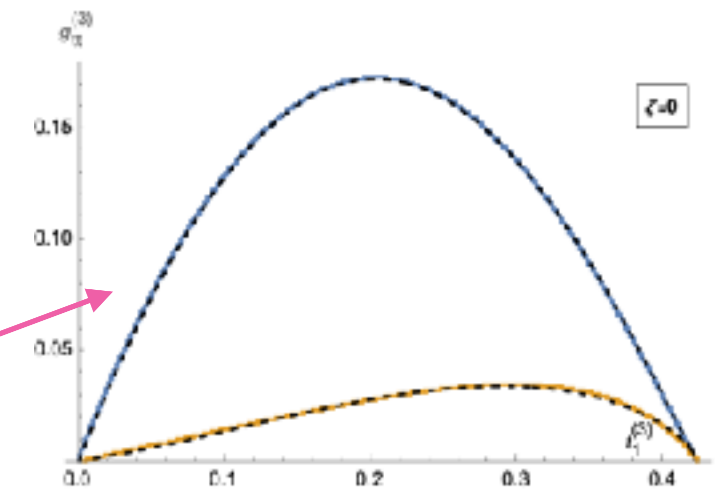
$$g_{tt}^{(2)} = \frac{1-x}{x}\left(x\gamma + \sigma \log\left(\frac{1-x}{1-x_0}\right)\right)^2$$

$$g_{tt}^{(3)}(x) = \frac{1-x}{x}\,\sigma^2\left(1 + W\left(-\frac{x_0\gamma}{\sigma}\, e^{-\frac{x\gamma}{\sigma}}\right)\right)^2$$

$$g_{tt}^{(2)} = \frac{1-x}{x} \left( x\gamma + \sigma \ \log\left(\frac{1-x}{1-x_0}\right)\right)^2$$

$$g_{tt}^{(3)}(x) = \frac{1-x}{x}\, \sigma^2 \left(1 + W\left(-\frac{x_0 \gamma}{\sigma}\, e^{-\frac{x\gamma}{\sigma}}\right)\right)^2$$



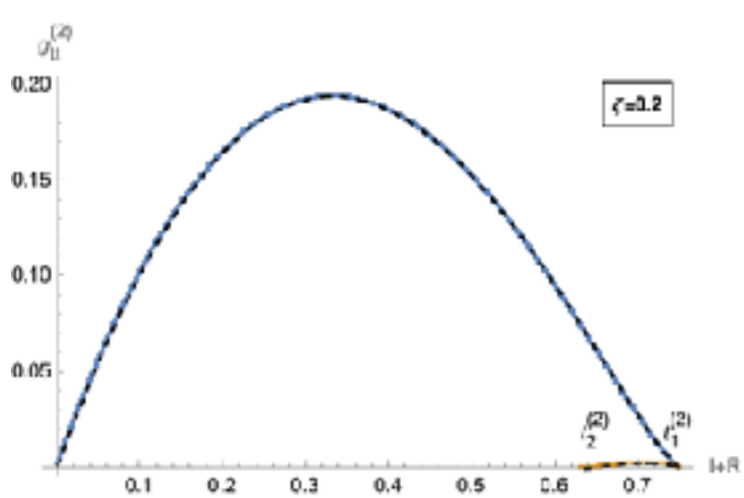

two real branches corresponding to

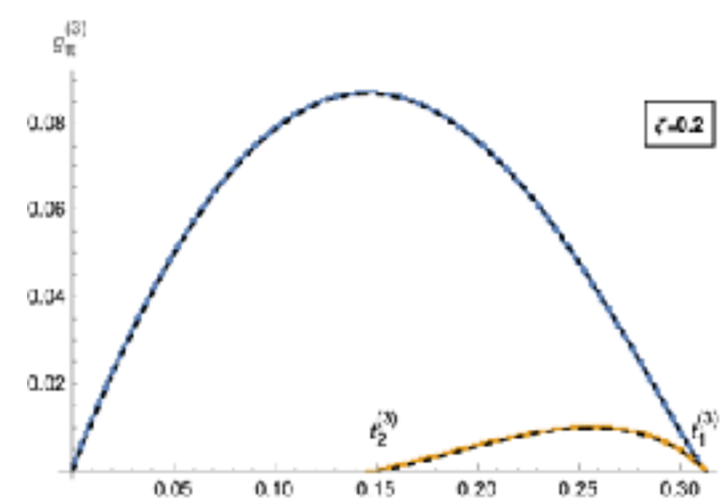
$\frac{dx}{dt} > 0$ and $\frac{dx}{dt} < 0$

dashed lines correspond to approximations

$$g_{tt}(x) = a\,(x - x_1)^b\,(x_2 - x)^c$$

# Generalisation to $\zeta > 0$



# Common Features of all Examples

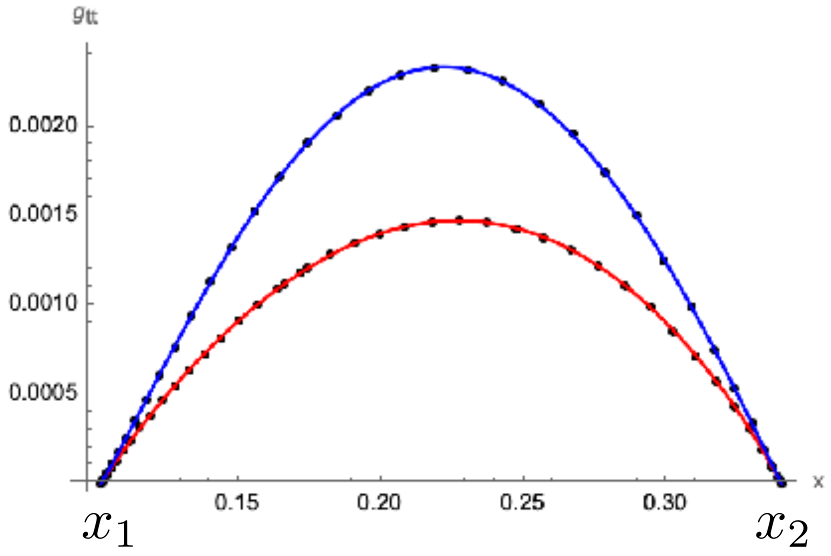
For systems with single degree of freedom $x(t)$

- real zeroes of the metric correspond to extrema of $x(t)$

- $g_{tt}$ described piecewise between two consecutive extrema of $x(t)$

$$g_{tt} = a(x - x_1)^b (x_2 - x)^c$$

for $0 \leq x_1 < x < x_2 \leq 1$




(imaginary zeroes possible in more complicated systems,
e.g. SIR(S) model with multiple pathogens)

- branches of the metric can be glued together along the zeroes

# Inverse Problem

Assume that the Fisher metric for a 1-dimensional system is

$$g_{tt}(x) = a\,(x - x_1)(x_2 - x)$$

for $a \in \mathbb{R}_+$, $0 \leq x_1 \leq x \leq x_2 \leq 1$

sign $\pm 1$

and solve the flow equation $\frac{dx}{dt} = \kappa \sqrt{g_{tt}\, x\, (1 - x)}$

For $0 < x_1 < x_2 < 1$ direct integration (initial condition: $x(t_0) = x_0$ )

$$t - t_0 = \frac{2i\kappa\,(F(u(x), m) - F(u(x_0), m))}{\sqrt{a(1 - x_1)x_2}}$$

inverted for $x(t)$

$$x(t) = \frac{x_1\,\mathrm{sn}\left(\frac{i\kappa(t-t_0)}{2}\,\sqrt{a(1-x_1)x_2} - F(u(x_0), m), m\right)}{1 - x_1 + x_1 \mathrm{sn}\left(\frac{i\kappa(t-t_0)}{2}\,\sqrt{a(1-x_1)x_2} - F(u(x_0), m), m\right)}$$

**Definition:**

Jacobi elliptic function

$$\mathrm{sn}(u, k^2) = \frac{\theta_3(0,\tau)}{\theta_2(0,\tau)}\,\frac{\theta_1(u/\theta_3^2(0,\tau),\tau)}{\theta_4(u/\theta_3^2(0,\tau),\tau)}$$

with argument $k = \frac{\theta_2^2(0,\tau)}{\theta_3^2(0,\tau)}$

first kind

$\ldots \sin^2 z$

$m = \frac{\ldots}{x_2(1 - x_1)}$ and

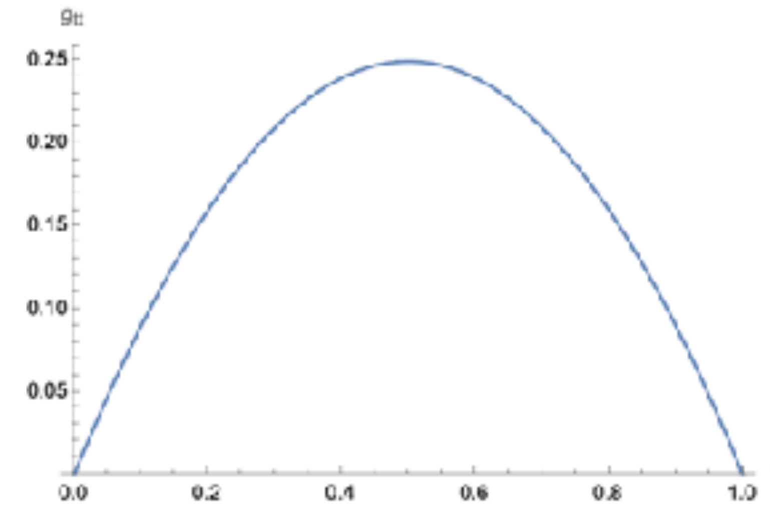
$u(x) = \arcsin\left(\sqrt{\frac{x(1-x_1)}{x_1(1-x)}}\right)$

use the solution to describe piecewise time evolution:

$$x(t) = \frac{x_1\,\mathrm{sn}\left(\frac{i\kappa(t-t_0)}{2}\,\sqrt{a(1-x_1)x_2} - F(u(x_0),m),m\right)^2}{1 - x_1 + x_1\mathrm{sn}\left(\frac{i\kappa(t-t_0)}{2}\,\sqrt{a(1-x_1)x_2} - F(u(x_0),m),m\right)^2}$$

(i) monotonic time evolution:

take $x_1 = 0$ and $x_2 = 1$

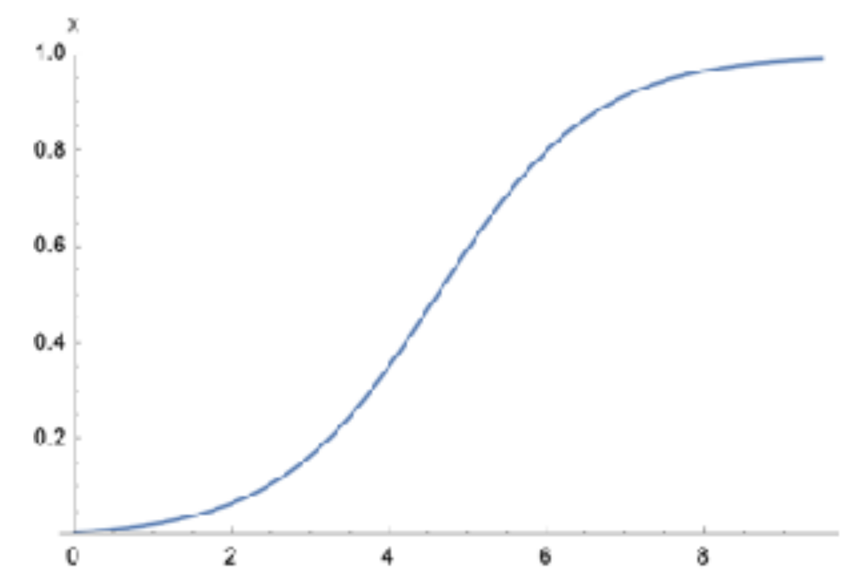



$$t - t_0 = \frac{\kappa}{\sqrt{a}} \log\left(\frac{x(1-x_0)}{x_0(1-x)}\right)$$

takes $t \to \infty$ for $x \to 0$ or $x \to 1$

logistic function

$$x(t) = \frac{x_0\, e^{\sqrt{a}\,\kappa(t-t_0)}}{1 - x_0 + x_0\, e^{\sqrt{a}\kappa(t-t_0)}}$$



## (ii) periodic time evolution

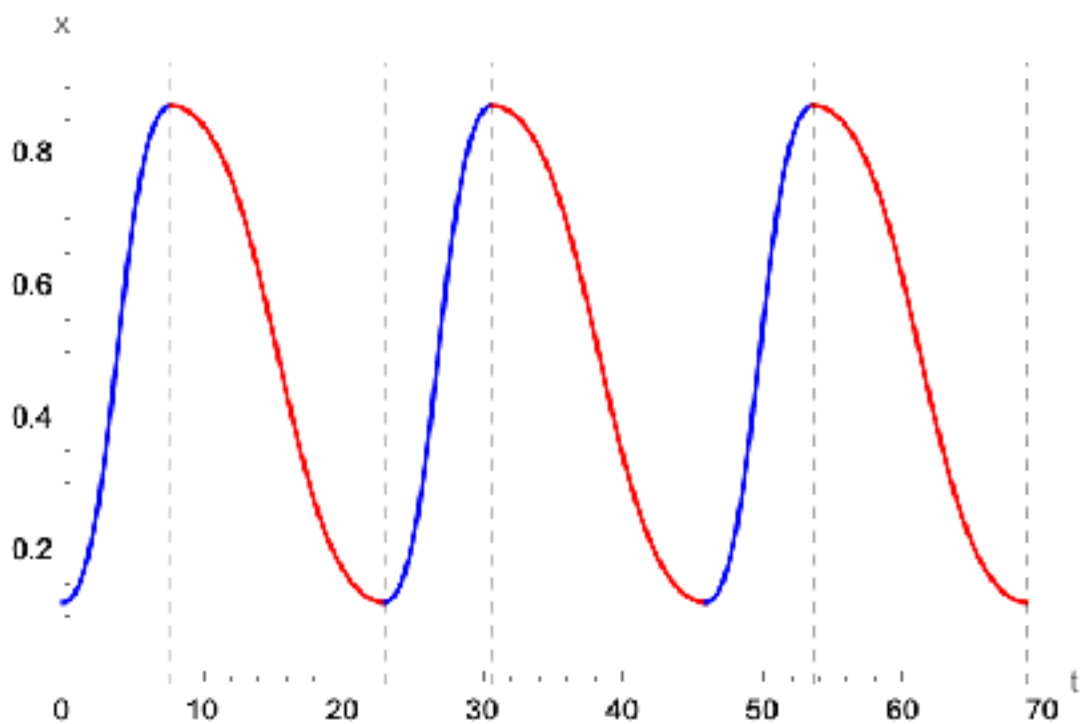
$$g_{tt} = \begin{cases} a_1\,(x - x_1)(x_2 - x) & \text{for} \quad \frac{dx}{dt} > 0\,, \\ a_2\,(x - x_1)(x_2 - x) & \text{for} \quad \frac{dx}{dt} < 0 \end{cases}$$

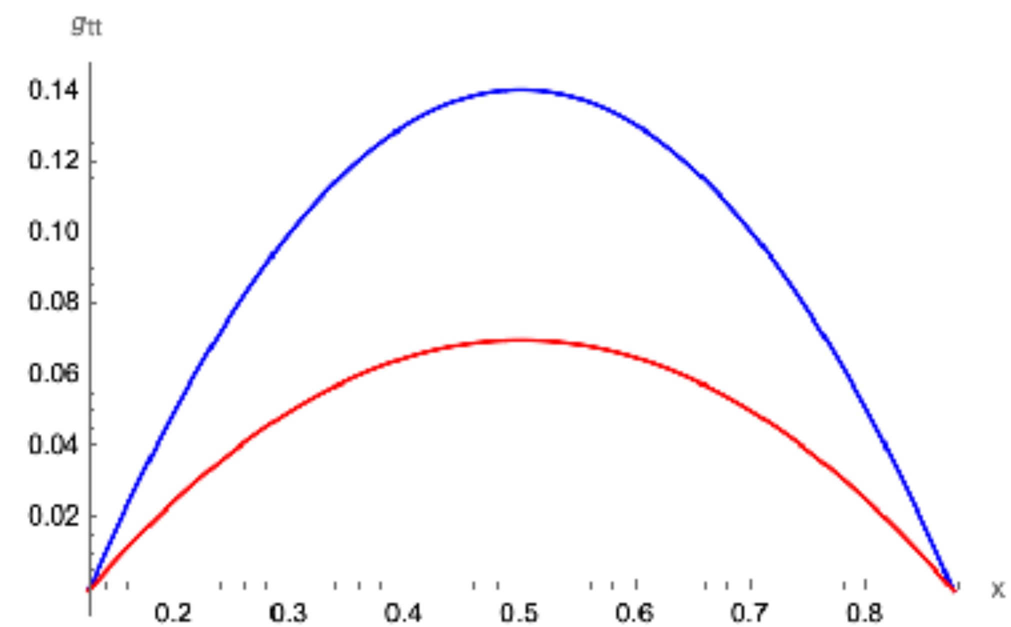
with

$$a_1, a_2 \in \mathbb{R}_+$$

$$0 < x_1 \leq x \leq x_2 < 1$$

periodic solution for $x(t)$




periodicity: $$T = \frac{2(\sqrt{a_1} + \sqrt{a_2})\,K\left(\frac{x_2 - x_1}{(1-x_1)x_2}\right)}{\sqrt{a_1 a_2 (1 - x_1) x_2}}$$

**Definition:**

complete elliptic integral of first kind

$$K(m) = \frac{\pi}{2} \sum_{n=0}^{\infty} \left( \frac{(2n-1)!!}{(2n)!!} \right)^2 m^n$$

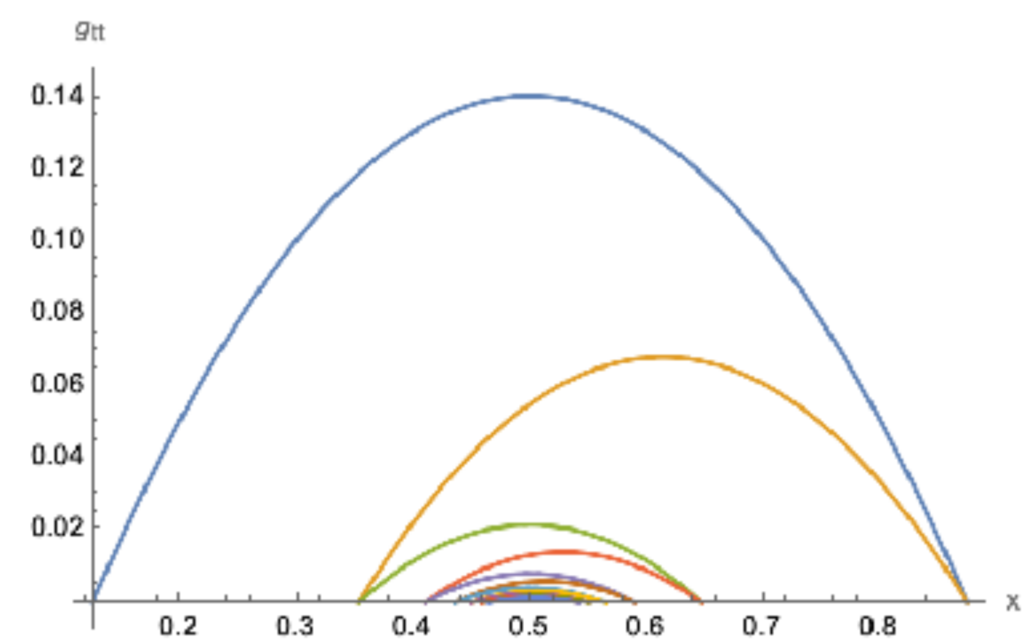
## (iii) oscillating time evolution

$\{x_1(n)\}_{n\in\mathbb{N}^*}$ and $\{x_2(n)\}_{n\in\mathbb{N}^*}$ convergent series with

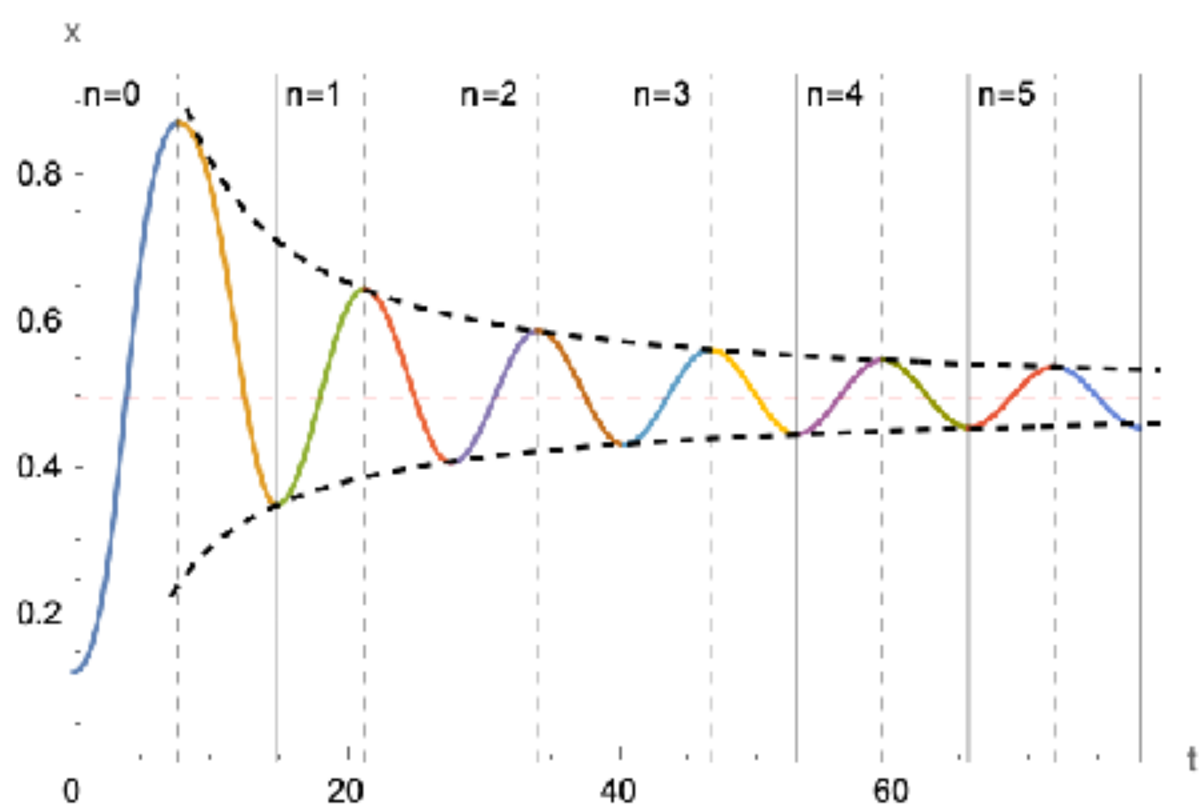
$0 < x_1(n) < x_1(n+1) \leq x_2(n+1) < x_2(n) < 1$ then

$$g_{tt}(n) = \begin{cases} a_1\,(x - x_1(n))(x_2(n) - x) & \text{for} \quad \frac{dx}{dt} > 0\,, \\ a_2\,(x - x_1(n))(x_2(n) - x) & \text{for} \quad \frac{dx}{dt} < 0 \end{cases}$$

with $a_1, a_2 \in \mathbb{R}_+$, $0 < x_1 \leq x \leq x_2 < 1$




oscillating solution of the form



# Conclusions

- simple models in epidemiology/population dynamics in terms of probability distributions
- in certain scenarios dynamics re-formulated in terms of the Fisher information metric
- universal features across different models (Lotka-Volterra, compartmental models)
- re-organisation around zeroes of the metric

# Outlook

- more general models with more degrees of freedom
- include further elements of information theory/geometry