# Complex Network Dynamics

# CONTENTS

# Basics

## Network basics (1VL)

**Further Reading**
- Newman, M. E. J. *Networks: An Introduction.* (Oxford University Press, 2010).
- Jungnickel, D. *Graphs, networks and algorithms.* (Springer, 2005).
- `https://www.maa.org/press/periodicals/convergence/leonard-eulers-solution-to-the-konigsberg-bridge-problem`

## Examples of Networks

Networks are like Graphs only with meaning

Under these links you can find some rather artistic visualisations of complex network dynamics:

- fireflies:
  `https://ncase.me/fireflies/`
- power grid:
  `https://bodograumann.github.io/powergrid-dynamics-simulation/`
- flight patterns:
  `https://vimeo.com/5368967`
- cell behaviour:
  `http://cancerres.aacrjournals.org/content/74/21/5963.`

The roots of network theory lie in small puzzles. It is commonly believed that one such puzzle, the Koenigsberg bridge problem, stimulated the mathematical approach to networks aka graph theory. In the 18th century, Euler solved the following puzzle: "Is it possible to find a route through Koenigsberg, crossing each of its seven bridges exactly once?" (Fig. 0.1).
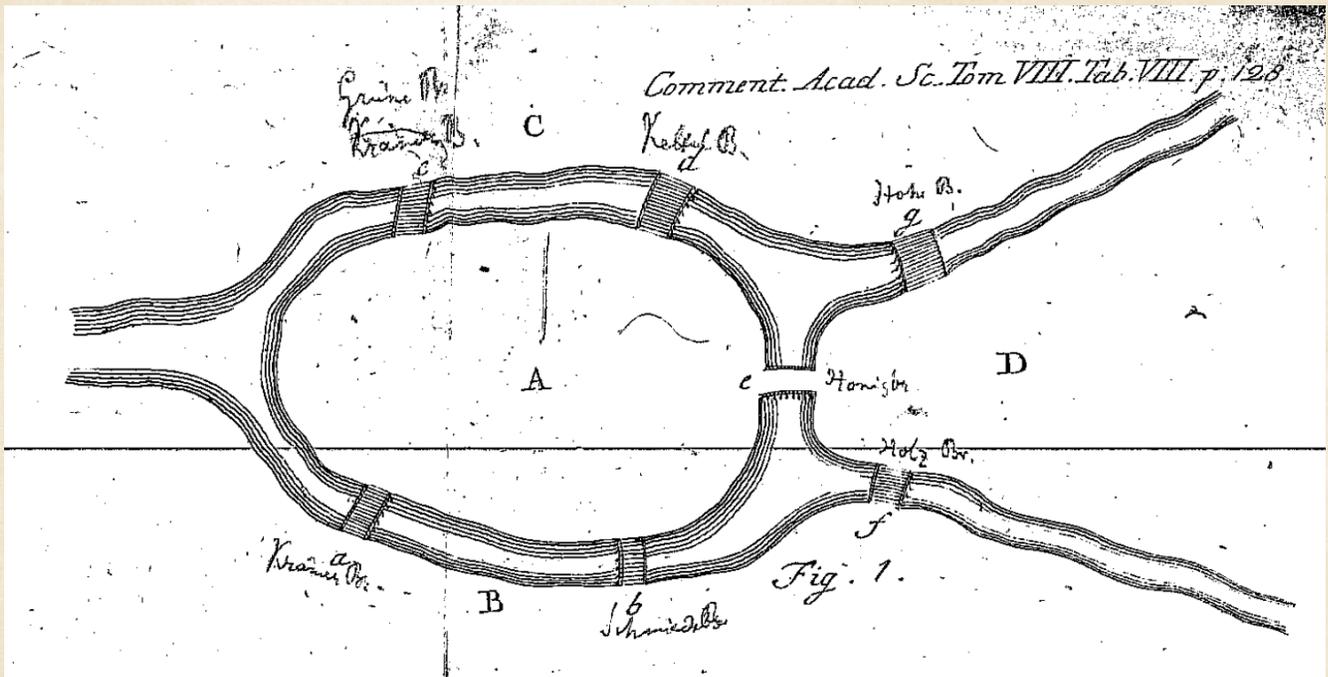
Figure 0.1: Source: Euler, L. *Solutio problematis ad geometrian situs pertinentis.* Comentarii academiae scientarum Petropolitanae 8, 128–140 (1736).

## MATHEMATICAL GRAPHS

**Graphs** $\mathcal{G} = (N, L)$ are sets of nodes $N$ and links $L$. Links $(u, v)$ are ordered or unordered pairs of nodes $u, v \in N$. The former give directed graphs, the latter undirected graphs. We denote by $n = |N|$ and $m = |L|$ the number of nodes and links in a graph throughout the lecture.

The links define a subset of the space of ordered pairs of nodes: $L \subset N \times N$. A natural way to encode this structure is by using the indicator function of this subset on the set of ordered pairs of nodes:

$$A : (i, j) \in N \times N \to \{0, 1\} \tag{0.1}$$

> This definition easily generalizes to other subsets of the power set of N with higher cardinality, leading to the concept of hyper-graphs.

An algebraic representation of $A$ is given by the **adjacency matrix**, which we also denote by $A$ for simplicity. $A$ is an $n \times n$-matrix with entries $1$ if $(i, j) \in L$ and $0$ otherwise. We exclude here the possibility for (self) loops and multiple links.

There is an important generalization of graphs. Assume a function $w : L \to \mathbb{R}_+ \setminus \{0\}$, assigning a positive number to each link. We call $\mathcal{G} = (N, L, w)$ a **weighted graph**. It is represented by a weighted adjacency with entries $w((i, j))$ if $(i, j) \in L$ and $0$ otherwise.

Mathematical graphs can be given a large number of structures in a natural way. Most naturally they are simplicial complexes.

If there can be several links between nodes, we need a different mathematical structure to capture this. One option is the incidence matrix:

$$B_{il} = \begin{cases} 1 & \text{if } i \in l \\ 0 & \text{otherwise} \end{cases} \tag{0.2}$$

You can check that if the graph is simple we have:

$$A = B \cdot B^T - D \tag{0.3}$$

where $D_{ij} = \delta_{ij}\, d_i$ is the diagonal matrix with the degree sequence on the diagonal.

## COUNTING FUNCTIONS ON GRAPHS



We can now translate Eulers problem into Graph terms. Landmasses are nodes, bridges are links. A route involving all nodes and using each link only once is called an **Eulerian path**.

Let us denote by $d_k$ the number of bridges connecting a landmass $k$. In graph terms, the degree.
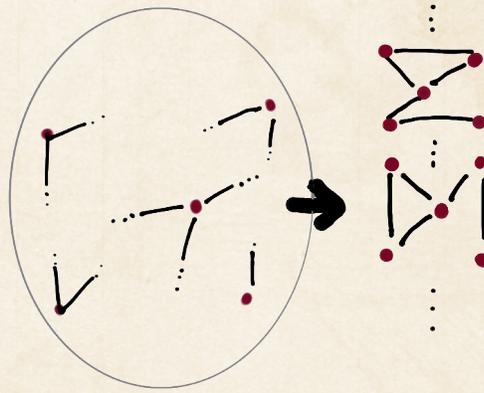
Figure 0.2: Creating graphs from a set of nodes with attached half-links.

The total number of bridges is $m = 7$.

Euler observed the so-called **handshaking lemma**

$$\sum_k d_k = 2m ,$$ 
(0.4)

where $m$ is the total number of links (bridges) in the graph. This can be immediately seen from Fig. 0.2, where we assign a number of $d_k$ half-links to each node $k$ (landmass). As all links are cut in halves, there a $2m$ half-links in total. Likewise, summing over all $k$ we count each link twice. The immediate consequence is, that the number of *odd* $d_k$ is even.

## EXERCISE: EULER'S PROOF
*Try to reproduce Euler's findings.*

Show that:

- If there are more than two landmasses with an odd $d_k$, no Eulerian path exists.
- If the number is exactly two, an Eulerian path exists and has to start in one of the two landmasses with odd $d_k$.
- If all $d_k$ are even, an Eulerian path exists and can start in any landmass.

### HINTS
.

Many interesting aspects of graphs are combinatorial in nature. That is, we simply count how many structural elements of a particular type we have.

***Nodewise-Neighbourhood.*** : Degree, Clustering

$$d_j = \sum_{k=1}^{n} A_{jk} \qquad (0.5)$$

The set of degrees $d_j$ is called the degree sequence of the graph.

Now we can restate Euler's **handshaking lemma** $\sum_k d_k = 2m$ in more mathematical terms: *Every finite undirected graph has an even number of vertices with odd degree.*

For unweighted graphs, the clustering coefficient is defined as:

$$C_j = \frac{\sum_{i,k=1;i\neq k}^{n} A_{ji} A_{jk} A_{ik}}{d_j(d_j - 1)} , \qquad (0.6)$$

i.e. it counts how many neighbours of a node are again linked. If they are, there will exist a triangle involving $i, j, k$ and $A_{ji} A_{jk} A_{ik} \neq 0$.

***Nodewise-Graph.*** : Betweeness, Centrality measures

A cycle-free path $\pi_{jk}$ in $\mathcal{G}$ is defined as an alternating sequence of nodes and links, starting at node $j$ and ending in $k$, such that no node is traversed more than once and no edge appears more than once in the sequence. We denote by $\pi_{jk}^s$ a **shortest path** between $j$ and $k$ containing the minimal number of links. Then, $\ell_{jk}^s$ is the **shortest-path length**.

There are two commonly used characteristics based on shortest paths.

Take the fraction of shortest paths between a pair of nodes $s, t$ that involve node $i$. The avarage over all possible pairs $s, t$ defines the betweenness of $i$.

$$b_i = \frac{2}{(n-1)(n-2)} \sum_{s,t=1;s\neq t\neq i}^{n} \frac{|\{\pi_{st}^s : i \in \pi_{st}^s\}|}{|\{\pi_{st}^s\}|} \qquad (0.7)$$

Complementary to the notion of a node's betweenness, it is useful to consider the centrality. A particular example is the so-called closeness centrality:

$$c_i = \frac{n-1}{\sum_{j=1;j\neq i}^{n} \ell_{ij}^s} \,'' \qquad (0.8)$$

i.e. the inverse average distance of node $i$ to the all other nodes.

Both betweenness and centrality have been defined firstly in sociology to characterise the roles of individuals in social networks of acquaintances. Caution: You'll often find the term 'betweenness centrality' in the literature as a common misconception. Betweeness is an independent concept. It might often be correlated to a node's centrality but it is easy to construct counterexamples of central nodes with vanishing betweenness.

**_Global._** : Density, Mean degree

A global characteristic is for instance the graph density $\rho = \frac{2m}{n(n-1)}$. Besides, many global measures are just moments of the distribution of nodewise characteristics. Examples are the mean degree and average clustering.

## SPECIAL GRAPHS
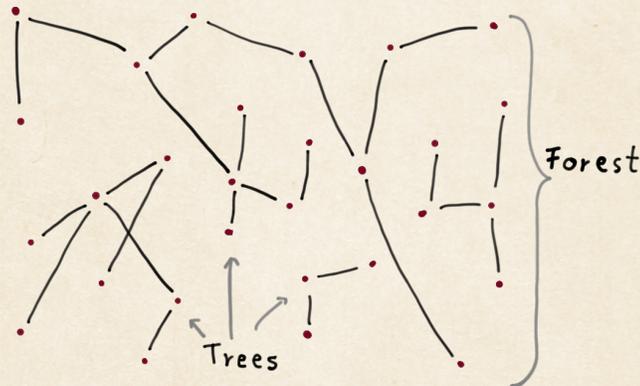


Figure 0.3: **Forest, tree**: A forest is an undirected graph, whose components are trees. A tree is an acyclic connected graph, i.e. every pair of nodes is connected by exactly one path.
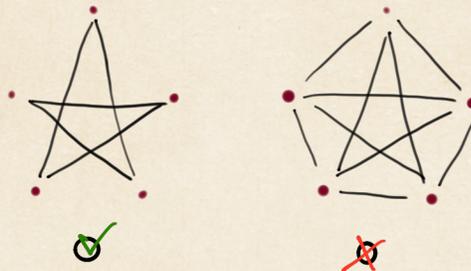


Figure 0.4: **Planar graphs**: A planar graph can be embedded in the plane without links crossing each other. The left graph is planar, the right one is not.

## ADVANCED FUNCTIONS ON GRAPHS

**FURTHER READING**
- Modularity: `https://arxiv.org/abs/physics/0602124`
- Fortunato, S. Community detection in graphs. Phys. Rep. 486, 75–174 (2010).

Figure 0.5: **Bipartite graph**: A graph where nodes form two groups and every link connects nodes from both groups.

How can we understand the macroscopic organisation of networks? Many examples show a structural organisation in 'losely-connected' components. This high-level information about the underlying complex system can be grasped by more advanced functions. We will study a popular example in more detail: the **modularity** function.
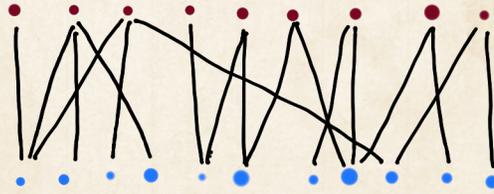
Define $Q_{ij} := A_{ij} - \frac{d_i d_j}{2m}$. The second term is the expectation value for a link $(i,j)$ to be present in a random graph. The reasoning is as follows. We start with the degree distribution of the original graph and assign randomly to each node a corresponding number of $2m$ half-links (see Fig. 0.2). The joint probability of drawing a half-link from $i$ and from $j$ is then just given by $\frac{d_i d_j}{2m}$. Constructing a random graph in such a way from a given degree sequence is called a **configuration model**. In the formulation of $Q$, the configuration model serves as a **null model** for testing the hypothesis that $\mathcal{G}$ is partitioned into components, i.e. we expect a random graph to have no macroscopic structure.

Now let $s$ be a vector with values $s_i = \pm 1$ such that $(s_i s_j + 1)/2$ is 1 when $i$ and $j$ are in the same component and 0 otherwise. $s$ is an indicator vector of a bipartition of $\mathcal{G}$. For any guess of $s$, we can calculate the expectation value $\mathcal{Q}$ of $Q$ over all links falling into the same component as:

$$\mathcal{Q} = \frac{1}{2m} \sum_{i,j=1}^{n} Q_{ij} \frac{s_i s_j + 1}{2} = \frac{1}{4m} \sum_{i,j=1}^{n} Q_{ij} s_i s_j + 0 = \frac{1}{4m} \langle s, Qs \rangle \tag{0.9}$$

We call $\mathcal{Q}$ the modularity of $\mathcal{G}$ given a guess $s$ for a bipartition. Newman writes: "*The modularity is, up to a multiplicative constant, the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random.*"

In general, one looks for an $s$ that maximizes the modularity. Let's look at a special case to better understand what this means. Sort $s$ such that we have $\nu$ 1s followed by $n - \nu$ $-1$s and two partitions $\alpha$ and $\beta$. The superscript $\cdot^{(\alpha)}$ means a quantity is restricted to links within $\alpha$.

$$Qs = \sum_{j=1}^{\nu} A_{ij} - \frac{d_i}{2m} \sum_{j=1}^{\nu} d_j - \sum_{j=\nu+1}^{n} A_{ij} + \frac{d_i}{2m} \sum_{j=\nu+1}^{n} d_j \tag{0.10}$$

$$= d_i^{(\alpha)} - d_i^{(\beta)} - \frac{d_i}{m}(m^{(\alpha)} - m^{(\beta)}) \tag{0.11}$$

$$\langle s, Qs \rangle = (m^{(\alpha)} - m^{(\beta)}) - \frac{(m^{(\alpha)} - m^{(\beta)})^2}{m} - \left( \sum_{i=\nu+1}^{n} d_i^{(\alpha)} + \sum_{i=1}^{\nu} d_i^{(\beta)} \right) \tag{0.12}$$

For $m^{(\alpha)} = m^{(\beta)}$, an equal partition, the maximum is obtained when there are no links between the communities. Of course, we do not have a perfect bipartition in general, but the intuition carries over that we want to maximize link density inside the communities and simultaneously minimize the cross-community links.

> Finding optimal partitions of graphs is generally NP hard which we can only approach using heuristic optimisations. Also, there is no precise definition of a community as such.

In undirected networks, $Q$ is symmetric and hence diagonalisable. We can express $s$ in an eigenvector base of $Q$ as $s = \sum_j \langle s, u_j \rangle u_j$.

$$\mathcal{Q} = \langle s, Q \sum_j \langle s, u_j \rangle u_j \rangle \tag{0.13}$$

$$= \sum_j \langle s, u_j \rangle \langle s, Q u_j \rangle \tag{0.14}$$

$$= \sum_j \lambda_j \langle s, u_j \rangle^2 \tag{0.15}$$

There are different ways how this can be maximized. Assume the largest eigenvalue $\lambda_{max}$ is positive. Then $\mathcal{Q}$ is proportional to the overlap of $s$ with $u_{max}$ and maximal when they are parallel. The latter, however is generally impossible due to the nature of $s$. Choose e.g. $s = sign(u_{max})$. Then:

$$\mathcal{Q} = \sum_j \lambda_j \left( \sum_k |u_k| \right)^2 \tag{0.16}$$

with contributions from many (also negative) eigenvalues.

In the other case, $\lambda_{max} = 0$ and $u_{max}$ is a constant vector. Hence, there is no meaningful bipartition of the network.

Note also that the summands attribute a non-negative membership weight to each node, i.e. $\mathcal{Q}$ typically depends on some nodes with higher weight compared to the rest.

So far, we only considered bipartitions. We can continue partitioning the obtained communities until the modularity does not increase anymore. For this, however, we have to generalize our $Q$. As we do not consider edges leaving community $\alpha$, we need to correct the degrees. Otherwise, the total modularity would change.

$$Q_{ij}^{(\alpha)} := A_{ij} - \frac{d_i d_j}{2m} - \delta_{ij} \left( d_i^{(\alpha)} - d_i \frac{m^{(\alpha)}}{m} \right) \tag{0.17}$$

This yields

$$\mathcal{Q}^{(\alpha)} = \langle s, Q^{(\alpha)} s \rangle \tag{0.18}$$

as the contribution of partitioning community $\alpha$ to the initial modularity $\mathcal{Q}$. We continue the consecutive bipartitions as long as the resulting $\mathcal{Q}^{(\alpha)}$ are positive.

## GRAPHS AS OBSERVABLES: FUNCTIONAL NETWORKS

**FURTHER READING**
- Marwan, N., Carmen Romano, M., Thiel, M. & Kurths, J. Recurrence plots for the analysis of complex systems. Phys. Rep. 438, 237–329 (2007).
- Donges, J. F., Heitzig, J., Donner, R. V. & Kurths, J. Analytical framework for recurrence network analysis of time series. Phys. Rev. E 85, 046105 (2012).

In many examples there is a natural network structure of the data. Examples are social networks, road networks, etc. It can also be highly interesting to construct networks from data that is not intrinsically network-like (climate, proteins, time-series). This means identifying what we mean by nodes and what we mean by nodes to be connected.

Examples include

|  | Nodes: | Links: |
|---|---|---|
| Correlation networks | measurements | correlations |
| Contact networks | amino acids | closeness in space |
| Recurrence networks | points in time | closeness in phase space |

Let us assume we have a time series $s = (s_0, \ldots, s_N$ of $N$ discrete measurements (at equal intervals). Then we can construct the so-called **recurrence plot** as follows.

$$R_{ij} := \Theta \left( \epsilon - c(s_i, s_j) \right) \tag{0.19}$$

$R$ is a binary matrix with entries $0$ and $1$. The function $c(\cdot)$ is a distance function. Here, however, $s$ needs to be embedded in some metric space where we can define a

dstance, e.g. euclidean distance or supremum norm. A common approach is delay embedding

$$\vec{x}_i := \sum_{j=1}^{m} s_{i-(j-1)\tau}\vec{e}_j \qquad (0.20)$$

where $m$ denotes the embedding dimension and $\tau$ the delay. This is especially necessary as typically not all varibles of a system are measured but only one ($s$).

There are various possibilities to choose $c$ and to select an $\epsilon$, this is an active area of research. The resulting $R_{ij}$ then looks like in Fig. 0.6:
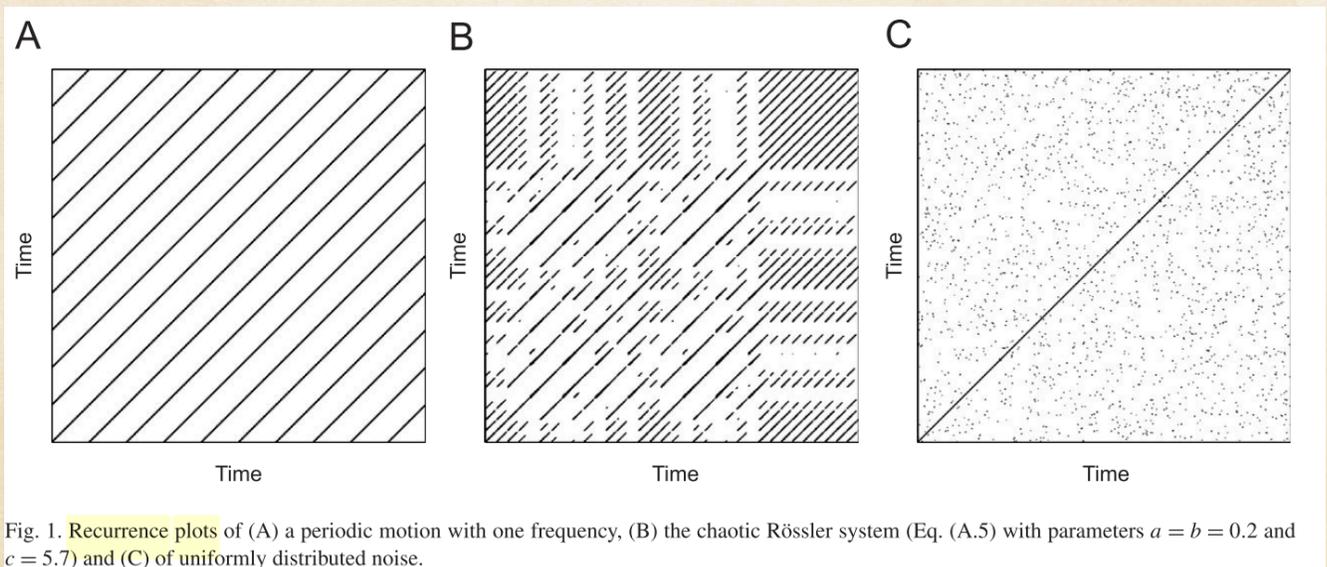


Fig. 1. Recurrence plots of (A) a periodic motion with one frequency, (B) the chaotic Rössler system (Eq. (A.5) with parameters $a = b = 0.2$ and $c = 5.7$) and (C) of uniformly distributed noise.

Figure 0.6: Recurrence plots for different dynamical regimes.

As you might already guess, a **recurrence network** is defined by identifying $R$ with the adjacency matrix of a graph (see Fig. 0.7).

What are the advantages of recurrence networks in contrast to other methods? Recurrences provide a lot of information about dynamical systems, which we recapitulate in the next section. They can be used to

- uncover complex bifurcation scenarios, e.g. reliably distinguish between chaotic and nonchaotic dynamics.
- analyse significantly shorter time series (100+) than required by classical techniques such as estimating the maximum Lyapunov exponent from data.
- cope with problematic data commonly found in the geosciences or in astrophysics, i.e. irregularly sampled data, uncertain timing of observations.
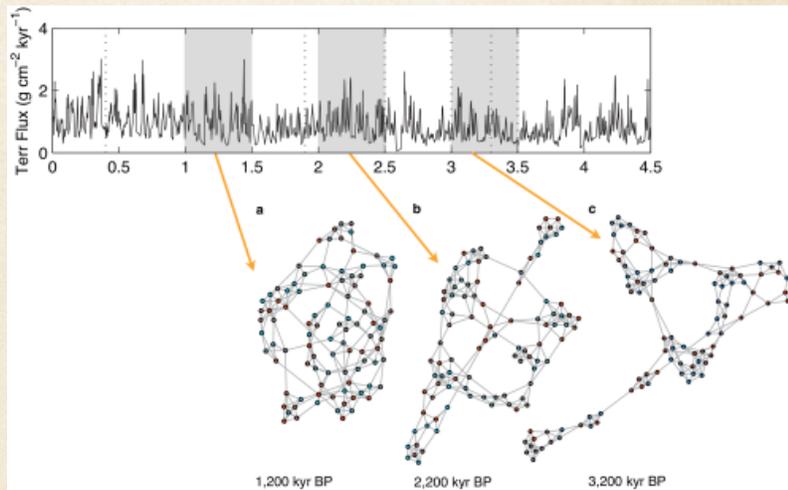
Figure 0.7: Recurrence networks extracted from windows in a time series. Source: `http://econs.pik-potsdam.de/index.php?a=methods`

# A RECAP ON DYNAMICAL SYSTEMS (1VL)

> **FURTHER READING**
> - Guckenheimer, J. & Holmes, P. J. Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Applied Mathematical Sciences, Volume 42 (Springer-Verlag, 2002).

> Dynamical Systems are like a box of chocolate. You never know what you get.

We define a dynamical system as the time evolution of a set of variables in the context of a given model. A set of values associated to these variables uniquely defines the system's **state** $x$. The space given by the range of all state variables, i.e. the space of all possible states the dynamical system might obtain, is called **phase space** $X$ with points $x \in X$. Typically, we will consider $X \subseteq \mathbb{R}^D$ if not stated otherwise.

We can typically formulate the time evolution as a (system of) first-order nonlinear differential equation(s)

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \tag{0.21}$$

with a suitable smooth function $f : X \to X$. The vector field $f$ gives the 'velocity' in every point, i.e. it tells you how fast $x$ is changing and in which direction.

Alternatively, the evolution can be given by a discrete map:

$$x_{n+1} = g(x_n) \tag{0.22}$$

which we do not discuss further for now.

When $f$ does not explicitly depend on time $t$, we say the dynamical system is **autonomous** as it is usually the case in the absence of driving forces, control or stochastic terms. Otherwise, **non-autonomous** systems can be made autonomous by including $t$ as an artificial variable with dynamics $\dot{t} = 1$.

Furthermore, a model is said to be **deterministic** when the time evolution always yields the same output for identical input. Consequently, the evolution only depends on the initial state $x_0$ and the evolution laws do not change over time.

Solutions (= integral curves) $x(t; x_0)$ to Eq.0.21 are everywhere tangent to the vector field $f(x)$ and are uniquely characterised by the initial state $x_0$ from where we start the time evolution. In phase space, solutions curves are called **trajectories**.

> For nice enough $f$ there are theorems yielding **existence and uniqueness** of solutions $x(x_0, t)$ (check the Guckenheimer book for details). The important consequences are:
>
> (i)  It is not possible for two trajectories to intersect or to be tangential in any point.
>
> (ii)  A trajectory cannot intersect itself, except for cycles.
>
> (iii)  Fixed points cannot be reached by a trajectory in finite time.

To get a better intuition for dynamical systems, let us start with deriving classic equations of motion from Newton's second law:

$$\frac{d}{dt}(m\dot{x}) = F = F(x, \dot{x}, t) \tag{0.23}$$

Typically, the mass $m$ is constant and we obtain the second law, i.e. a (net) force $F$ acting on a mass leads to an acceleration.

In physics, we often we deal with a **gradient system**

$$m\ddot{x} = F = -\nabla V(x) \tag{0.24}$$

where $V$ is a potential function. Such systems are termed **conservative**, containing the important class of Hamiltonian systems.

As an example, take the mathematical pendulum of fixed length $l$, i.e. an oscillator. We have the gravitational potential $V(\phi) = -mg\cos\phi$ leading to a tangential acceleration as

$$ml\ddot{\phi} = -mg\sin\phi \,, \tag{0.25}$$

where $\phi$ is the displacement angle. This is our first case of a **nonlinear** dynamical equation. Solutions are given by elliptic integrals, for small angles we obtain a linear harmonic oscillator. The potential is sketched in Fig. 0.8. To make the story more
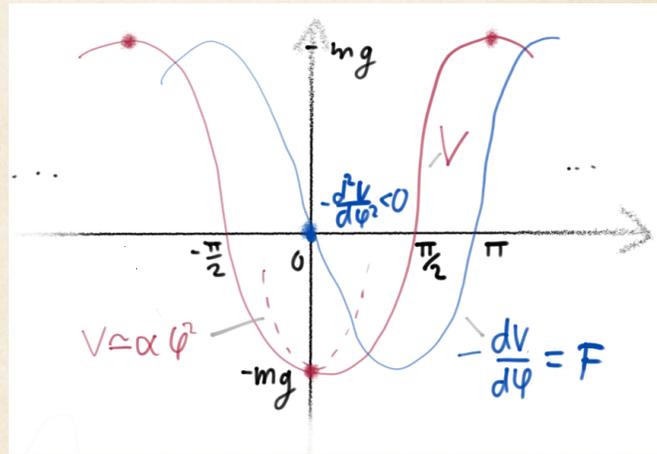
Figure 0.8: The potential $V$ is periodic with minima in $0+k2\pi$ and maxima in $\pi+k2\pi$ for $k \in \mathbb{N}$. A quadratic approximation of the potential around the minimum yields the harmonic oscillator.

interesting, we add a motor exerting a constant driving force $M$ on the pendulum and account for damping by adding a term $-d\dot{\phi}$.

$$ml\ddot{\phi} = M - d\dot{\phi} - mg\sin\phi \tag{0.26}$$

which we can rewrite as

$$\ddot{\phi} = a - b\dot{\phi} - c\sin\phi \tag{0.27}$$

By adding a dissipative force (damping) the pendulum becomes non-conservative. We generally call this a **dissipative** dynamical system.

This pendulum has two equilibrium positions, given by $\phi_1 = \arcsin\left(\frac{a}{c}\right)$, $\dot{\phi}_1 = 0$ and $\phi_2 = \pi - \arcsin\left(\frac{a}{c}\right)$, $\dot{\phi}_2 = 0$. There is, however, more to discover from this equation of a damped-driven pendulum as we will see below.

## ASYMPTOTIC BEHAVIOUR

In the theory of dynamical systems we aim for understanding the behaviour of a model without necessarily solving its equations. On the one hand, solutions are typically not known for non-linear $f$. On the other hand, we can read a lot of information from the structure of the equations.

Consider again the conservative pendulum Eq.0.25 and the potential $V$ around $\phi = 0$ in Fig.0.8. The extreme values are equilibrium points where the gradient vanishes and the resulting force $F$ is 0. We immediately see that the derivative $dF/d\phi|_{\phi=0}$ is negative, hence small displacements around the minimum experience a reversing force. The origin thus is a **stable equilibrium** while the maxima are **unstable**.

Gradient systems are fully characterised by their potential function. Initialising the pendulum at some state, we can expect it to oscillate around the stable equilibrium with a certain period. It will, however, never stop there unless we start at $\phi = 0$.

Only in a dissipative system, the trajectory would asymptotically approach a fixed point. This is the case for the damped-driven pendulum Eq.0.27. We need a re-placement of the potential landscape for analysing our system. **The** approach for general non-linear dynamical systems is given by **phase portraits**, i.e. sketches of the system's phase space structure.

In Fig. 0.9 we see the phase portrait of the damped-driven pendulum. We can construct it as follows.

1. For each variable $x_k$ in $x$, sketch the curve $\dot{x}_k = 0$. These are called **nullclines**, their intersections are the fixed points.

2. Sketch the vector field. We can read of the directions of the $\dot{x}_k$ in the equations. Regions of equal sign are separated by the nullclines.

At this point, it is typically already possible to infer the asymptotic behaviour qualitatively.

In a dissipative system each asymptotic state $\mathcal{A}$ (e.g. fixed points) attracts trajectories from a certain set $\mathcal{B}(\mathcal{A}) \subseteq X$ which we call **basin of attraction**. We call $\mathcal{A}$ an **attractor**. Note that due to the theorem of existence and uniqueness, attractors are not reached by trajectories in finite time but they come arbitrarily close.

Dynamical systems show various types of asymptotic behaviour:

- equilibrium position: **fixed point** of $f$
- periodic process, oscillation: **closed orbit** in $X$
- oscillations with $k$ different frequencies, beats: surface of $k$-**torus** in $X$
- irregular, non-periodic oscillation: **chaotic motion**

In non-linear systems, several attractors can coexist with separate basins. The asymptotic behaviour hence depends on the initial state of the system and we say the system is **multistable**.

An example is again our damped-driven pendulum. Besides the fixed points that we already identified, there is also a periodic orbit as another attractor (see Fig.0.9). By numerically integrating Eq.0.27, we obtained two trajectories from nearby initial conditions. While the blue trajectory starts in the basin of the fixed point $(\phi_1, \dot{\phi}_1)$,
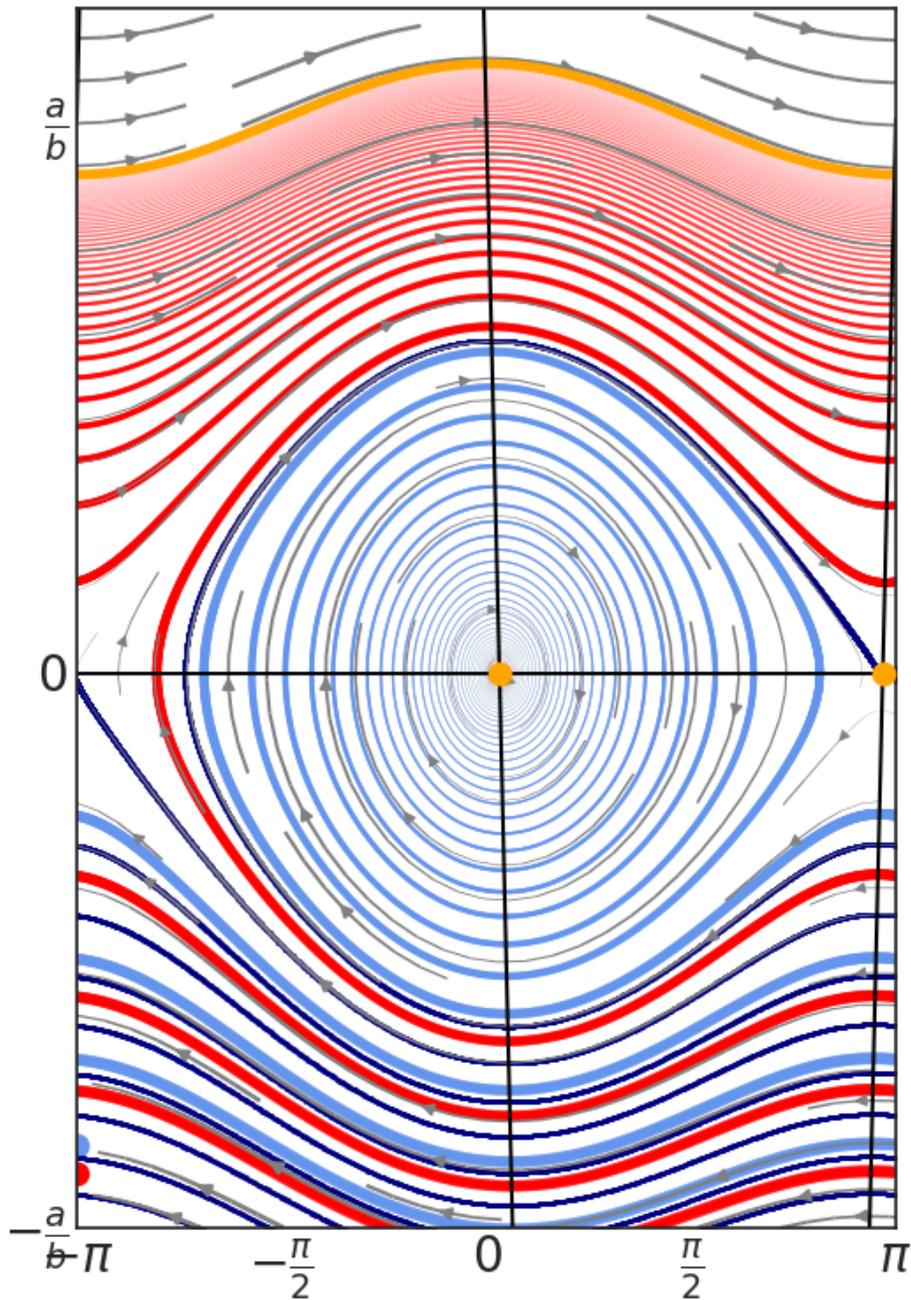
Figure 0.9: Phase space of the damped-driven pendulum $\omega$ vs. $\phi$. **light blue** : trajectory towards stable fixed point, **red**: traj. towards limit cycle, **orange curve**: periodic orbit, **orange dots**: fixed points, **dark blue curves**: stable manifolds of the saddle, **black curves**: null clines.

the red one approaches a periodic orbit (The orbit is closed because the system is $2\pi$-periodic.).

A universal definition of an attractor has not yet been found. We commonly work with the one given by Milnor.

**Definition 1** (Attractor). *An **attractor** $\mathcal{A}$ is defined by:*

1. *There is a minimal compact invariant set $\mathcal{A} \subset X$ whose basin of attraction $\mathcal{B}(\mathcal{A})$ has positive Lebesgue measure.*

2. *There is no strictly smaller set $\mathcal{A}'$ whose basin of attraction coincides with that of $\mathcal{A}$ up to a zero-measure set. (Minimality)*

You can think of the following interpretation. Draw a random state $s$ from $X$. We only want to consider $\mathcal{A}$ an attractor if the probability that $s \in \mathcal{B}(\mathcal{A})$ – i.e. that it is attracted by $\mathcal{A}$ – is positive. Then there is a significant chance to observe the system in $\mathcal{A}$ e.g. in an experiment.

Given our observations so far, the fixed point $(\phi_2, \dot{\phi}_2)$ does not seem to be an attractor. Why is that? This question is closely linked to the stability of possible asymptotic states. For instance, unstable equilibria are fixed points but they do not attract any trajectories, i.e they are no attractors according to our definition above.

## STABILITY AND BIFURCATIONS OF FIXED POINTS

Going back to the pioneering ideas of Aleksandr Michajlovič Lyapunov in the 19th century, certain behaviour (e.g. fixed points) of dynamical systems are considered to be stable if small perturbations to the initial conditions lead to small reactions of the system, i.e. small perturbations cannot substantially alter the system's asymptotic behaviour.

To quantify what 'small' means in this context, let us define a distance measure between a point $x$ and an arbitrary set $C \subset X$ as

$$d(x, C) := \inf\{\|x - y\| \mid y \in C\}\,, \tag{0.28}$$

where $\|\cdot\|$ is a suitable norm (the choice is not important). For fixed points, $d(\cdot, x^\star)$ is just the regular distance according to the norm, while for e.g. a periodic orbit $\gamma$ it reduces to the minimal distance to a point on $\gamma$.

Now, we can formalise Lyapunov's intuition in the following three definitions.

**Definition 2** (Locally attracting). *A set $C$ is said to be* locally attracting *if there exists a neighbourhood*

$$U_\epsilon = \{x_0 \in X \mid d(x_0, C) < \epsilon\}$$

*with $\epsilon > 0$ for which*

$$\lim_{t \to \infty} d(x(t; x_0), C) = 0 \ ,$$

*i.e. all trajectories from $U_\epsilon$ converge to $C$.*

This definition basically says that small perturbations return to $C$ asymptotically. Note that this definition alone does still allow for large transient deviations of a trajectory before it returns.

**Definition 3** (Lyapunov stable). *Consider a $\delta > 0$, $\delta < \epsilon$ defining a set*

$$U_\delta = \{x_0 \in X \mid d(x_0, C) < \delta\} \ .$$

*A set $C$ is said to be* Lyapunov stable *if for each neighbourhood $U_\epsilon$ as above there exists a $U_\delta$ such that for all $x_0 \in U_\delta$*

$$\forall_{t \geq 0} : \ x(t; x_0) \in U_\epsilon \ .$$

*Trajectories starting sufficiently close to $C$ remain bounded in a finite neighbourhood.*

This definition encodes the notion that trajectories emanating from small perturbations stay close-by.

**Definition 4** (Asymptotically stable). *When a set is both locally attracting and Lyapunov stable – i.e. small deviations asymptotically converge back to the set and are bounded – $C$ is said to be* asymptotically stable.

As we are dealing with small perturbations around a solution $x^*$ of Eqn. 0.21, we can introduce new variables $y := x - x^*$:

$$\dot{y} = \dot{x} - 0 = f(x^* + y) \tag{0.29}$$
$$\simeq f(x^*) + J(x^*)y + \mathcal{O}(\|y\|^2) \tag{0.30}$$

We observe that for small displacements, the dynamics is linear and determined by the **Jacobian matrix** $J(x^*)$ of partial derivatives $J_{jk} = \partial f_j / \partial y_k$. Linear ODEs with constant coeffients have exponential solutions $y(t) \sim \exp(\lambda t)$ where $\lambda \in \mathbb{C}$ is an eigenvalue of $J(x^*)$. For asymptotic stability, it is sufficient that $\Re \lambda_k < 0$ for all $k$ eigenvalues, such that small perturbations decay exponentially in all directions.

The fixed point $(\phi_2, \dot{\phi}_2)$ of Eqn. 0.27 is not asymptotically stable. This is because $J(\phi_2, \dot{\phi}_2)$ has an unstable direction with $\Re \lambda_1 > 0$ and a stable direction with $\Re \lambda_2 < 0$. Such fixed points with mixed stability are termed **saddle points**. The other two cases are **stable** and **unstable** fixed points. In fact, fixed points are charecterised by there stable

$$W_s(x^*) := \{x_0 \in X \mid \lim_{t \to \infty} d(x(t; x_0), x^*) = 0\} \tag{0.31}$$

and unstable

$$W_u(x^*) := \{x_0 \in X \,|\, \lim_{t \to -\infty} d(x(t;x_0), x^*) = 0\} \tag{0.32}$$

invariant manifolds. Close to a fixed point, $W_s$ and $W_u$ are tangential to the eigenvectors of $J(x^*)$. An example are the invariant manifolds of the saddle point in Fig. 0.10.
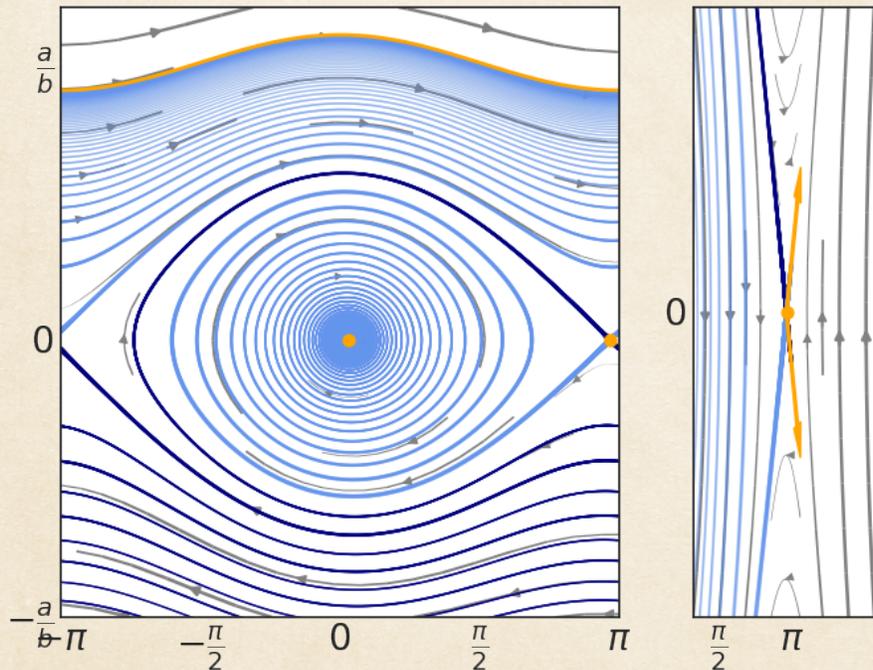


Figure 0.10: Invariant manifolds of the saddle in the damped-driven pendulum. **light blue** : unstable manifold $W_u$, **dark blue**: stable manifold $W_s$. **orange curve**: periodic orbit, **orange dots**: fixed points. The right panel shows a zoom-in around the saddle and the corresponding Jacobian's eigenvectors.

In the case of $\Re\lambda = 0$ for some $\lambda$, the fixed point has a neutral direction. Then, a linearisation is not sufficient to infer stability and higher orders need to be considered.

The notions presented here can be transfered to periodic trajectories in a straightforward way. This leads to the so-called **Floquet theory**, which we do not discuss here for breveity.

## CHAOS

**Chaos** is another qualitatively different type of asymptotic behaviour. It is characterised by:

- irregular, non-periodic *deterministic* motion

$$\dot{x} = x(\mu + x)$$

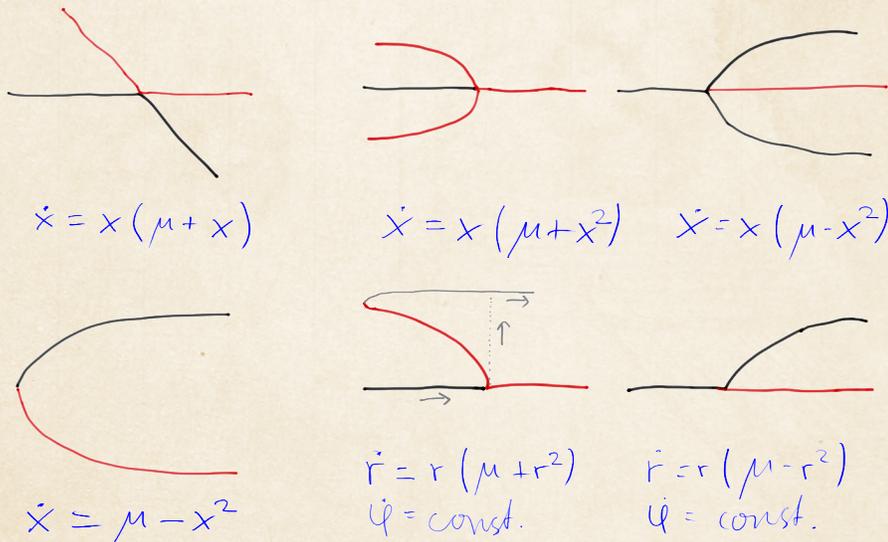$$\dot{x} = x(\mu + x^2) \qquad \dot{x} = x(\mu - x^2)$$

$$\dot{x} = \mu - x^2 \qquad \dot{r} = r(\mu + r^2) \qquad \dot{r} = r(\mu - r^2)$$
$$\dot{\varphi} = \text{const.} \qquad \dot{\varphi} = \text{const.}$$

Figure 0.11: Schematic bifurcation diagrammes $x^*$ vs. $\mu$ for selected bifurcations. Black: stable branch, red: unstable branch. **Top row**: transcritical, subcritical pitchfork, supercritical pitchfork. **Bottom row**: saddle-node (fold), subcritical Andronov-Hopf, supercritical Andronov-Hopf. These are examples for local codimension-1 bifurcations (top, saddle-node) and global codimension-2 bifurcations (Hopf).

- sensitive dependence on initial conditions
- so-called strange attractors with fractal (non-integer) dimension.

The most famous example is Lorenz' model of convective turbulence. The equations are:

$$\dot{x} = \sigma(y - x) \tag{0.33}$$
$$\dot{y} = x(r - z) - y \tag{0.34}$$
$$\dot{z} = xy - bz \tag{0.35}$$

For $\sigma = 10$, $b = 8/3$ and $r = 28$, the Lorenz model exhibits chaotic motion as depicted in Fig. 0.12. As is visualised on the left, the **sensitive dependence** causes an exponential separation of nearby initial conditions over time. This has often been referenced as the butterfly effect, i.e. the slightest change in a system causes a large reaction. Still, note that the two trajectories remain in the same bounded region of the phase space. They infinitely wind around a fractal attractor, 'randomly' switching between two modes of oscillation (right panel). Such fractal sets of non-integer dimension are termed **strange attractors**.
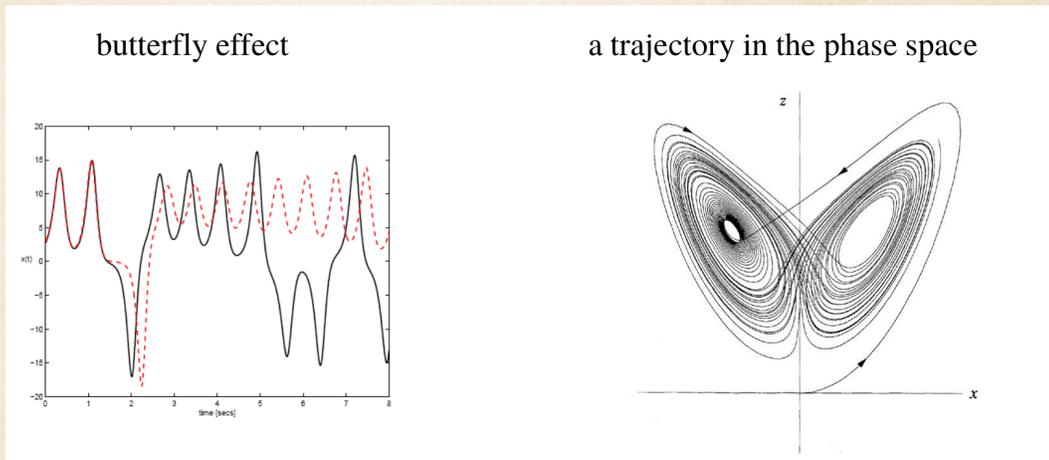
butterfly effect            a trajectory in the phase space

Figure 0.12: **Left panel**: Time series of $x$. For two near initial conditions. **Right panel**: Projection of the three-dimensional phase space on the $xz$-plane. Source: Yuri Maistrenko: *Complex Network Dynamics: Theory & Application.* Lecture Notes TU Darmstadt (2014)

### EXERCISE: LORENZ MODEL

*Analyze Lorenz' model of convective turbulence.*

Find the fixed points and bifurcations in the Lorenz model.

#### HINTS

***The discrete derivative.*** Use $r$ as the bifurcation parameter and fix $\sigma = 10$, $b = 8/3$.

When nearby points get separated exponentially but the whole motion still is confined to the vicinity of the strange attractor, how can we characterise its stability?

Consider a small volume $V = \prod_k \Delta x_k$ with length $\Delta x_k$ in every direction. We can calculate the change of $V$'s shape under the dynamics (Eqn. 0.21):

$$\dot{V} = V \sum_k \frac{\dot{\Delta x_k}}{\Delta x_k} = V \sum_k \frac{f(\Delta x_k)}{\Delta x_k}$$
$$\simeq V \sum_k \frac{\partial f_k(\Delta x_k)}{\partial x_k}$$
$$\simeq V \operatorname{div} f$$

For the Lorenz model we find $\operatorname{div} f = -(1 + \sigma + b) < 0$, i.e. every volume element shrinks over time. This is actually the case for all dissipative systems (you can show that gradient systems have $\operatorname{div} f = 0$) and the chaotic attractor is an attractor in the sense we considered above.

Let $u(t)$ and $u(t) + \delta(t)$ be two neighbouring trajectories with $\|\delta\| \ll 1$. In general, we

find :

$$\dot{u} + \dot{\delta} = f(u) + f(u + \delta) \tag{0.36}$$
$$\Rightarrow \dot{\delta} \simeq f(u) + J(u(t))\delta + \mathcal{O}(\|\delta\|^2) \tag{0.37}$$

or in componentwise form:

$$\dot{\delta}_k \simeq f_k(u) + \sum_{l=1}^{D} J_{kl}(u(t))\delta_l \tag{0.38}$$

Then we define the $n$ **Lyapunov exponents** for each direction as

$$\Lambda_k := \limsup_{t\to\infty} \frac{1}{t} \ln \frac{\|\delta_k(t)\|}{\|\delta_k(0)\|} \tag{0.39}$$

At a fixed point, i.e. $u(t) \equiv u^*$, the $\Lambda_k$ just equal the real part of the Jacobian matrices' spectrum. For periodic or chaotic attractors, Lyapunov exponents need to be calculated via numerical simulations. Sort them by magnitude as $\Lambda_1 \leq \Lambda_2 \leq \cdots \leq \Lambda_D$, then $\Lambda_D$ determines the stability. For $\Lambda_D \leq 0$, we observe asymptotic stability as nearby trajectories approach each other with an exponential rate less or qual to $\Lambda_D$ (In autonomous systems, there's always one zero Lyapunov exponent corresponding to longitudinal perturbations along a trajectory $u(t)$.). Correspondingly, $\Lambda_D > 0$ indicates divergent directions.

> How important is the initial condition $\delta(0)$? The 'multiplicative ergodic theorem' by Oseledec (1968) implies that almost all solutions have the same set of Lyapunov exponents because almost all trajectories are dense on the chaotic attractor and any group of them has the same time average. (Here we consider a smooth $f$ with only one attractor.)

Based upon the dynamics of a volume element, it can be shown that the Lyapunov exponents relate to the time average of the divergence as

$$\sum_k \Lambda_k = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \, \mathrm{div} f \tag{0.40}$$

This means that the sum of Lyapunov exponents is negative in dissipative systems. In the Lorenz model, we have $\Lambda_1 < \Lambda_2 = 0 < \Lambda_D$. Eqn. 0.40 implies $|\Lambda_1| > |\Lambda_D|$, i.e. although a deviation $\delta(t)$ is divergent in one direction, it is subject to an even stronger contraction in the other. This interplay of *stretching* and *folding* is characteristic for any chaotic systems. In case of several positive Lyapunov exponents, one speaks of **hyperchaos**.

# Dynamics on Networks I

**Further Reading**
- Merris, R. Laplacian matrices of graphs: a survey. Linear Algebra Appl. 197–198, 143–176 (1994).
- Lovász, L. Random walks on graphs: A survey. Comb. Paul Erdos is Eighty 2, 1–46 (1993).

## Diffusion processes (1VL)

Random walks are diffusion processes. They are described by the graph Laplacian:

## The Graph Laplacian

### Exercise: Graph Laplacian

*Show that the Graph Laplacian is the discrete version of the usual Laplacian.*

The graph Laplacian is the discrete version of the usual Laplacian operator. Write down the Graph Laplacian on a lattice and show that in the limit of the lattice spacing going to zero an appropriately scaled version of the Laplacian converges to the continuous Laplacian.

#### Hints

**The discrete derivative.** Work from the definition of the derivative:

$$\frac{df}{dx}\bigg|_{x=x^*} = \lim_{\epsilon \to 0} = \frac{f(x^* + \epsilon) - f(x^*)}{\epsilon}$$

Derive the second derivative.

## Spectral properties

If needed, Linear Algebra Recap.

SOURCES

## PROPERTIES OF DIFFUSION ON NETWORKS

(More detailed subsections?)

Graph partitioning by eigenvectors. Eigenvalues.

Weighted directed graph Laplacians generate stochastic matrices.

## MASTER STABILITY FUNCTION (1VL)

Frank

## EPIDEMIC SPREADING – SIRS MODEL (1VL)

> **FURTHER READING**
> This section losely follows the discussion in Newman's book, ch. 17.6 and 17.11.

Consider a popularion of $n$ individuals subject to a disease outbreak. We assume that each individual is susceptible, i.e. they all can receive the disease. Typically, a disease shows several stages. The most simple model distinguishes susceptible ($S$), infected ($X$) and recovering ($R$) individuals. The transition between the stages happens with fixed rates $\beta$, $\gamma$ and $\delta$. The recovery rate $\delta$, for instance, encodes the statistical recovery probability per unit time.
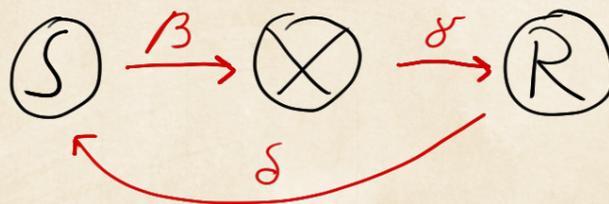


Figure 0.1: Disease mechanism

## MACROSCOPIC DYNAMICS

We start looking aat the SIRS model from a macroscopic point of view and are interested in the share of the infected part of the population. Define the densities $s = S/n$, $x$ and $r$. We can read the transition dynamics from Fig. 0.1:

$$\dot{s} = \delta r - \beta s x \tag{0.1}$$
$$\dot{x} = \beta s x - \gamma x \tag{0.2}$$
$$\dot{r} = \gamma x - \delta r \tag{0.3}$$
$$0 = s + x + r - 1 \tag{0.4}$$

These rate equations have two fix points. The first, $(s^*, x^*, r^*) = (1, 0, 0)$ corresponds to a healthy population where each individual either fully recovered or never had the disease in first place. We are going to be interested in its stability in the following. The alternative fix point is $(s^*, x^*, r^*) = (\frac{\gamma}{\beta}, (1 - \frac{\gamma}{\beta})\frac{\delta}{\gamma+\delta}, (1 - \frac{\gamma}{\beta})\frac{\gamma}{\gamma+\delta})$ where the disease is not going to vanish but rather becomes endemic.

## MICROSCOPIC DYNAMICS

We can further zoom into the SIRS model and consider the fate of individuals. Each one has an independent probability to be in stage $S$, $X$ or $R$ at a particular point in time. Let us denote this as $s_i := P(i \in S)$ and $x_i$, $r_i$ analogously. A disease is transmitted when an infected individual meets a susceptible one. Depending on the type of disease, the spreading process is mediated by some form of contact network (e.g. acquaintances, travel patterns, ...) that we assume to be connected and not time-dependent for the moment. (This can be relaxed to studying the largest connected component as we are not interested in isolated outbreaks.)

$$\dot{s}_i = \delta r_i - \beta s_i \sum_k A_{ik} x_k \tag{0.5}$$

$$\dot{x}_i = \beta s_i \sum_k A_{ik} x_k - \gamma x_i \tag{0.6}$$

$$\dot{r}_i = \gamma x_i - \delta r_i \tag{0.7}$$
$$0 = s_i + x_i + r_i - 1 \tag{0.8}$$

We can further simplify the system by replacing $x_i$ using the constraint:

$$\dot{s}_i = \delta r_i - \beta d_i s_i + \beta s_i \sum_k A_{ik}(s_k + r_k) \tag{0.9}$$

$$\dot{r}_i = \gamma - \gamma s_i - (\delta + \gamma) r_i \tag{0.10}$$

$$x_i = 1 - s_i - r_i \, , \tag{0.11}$$

where $d_i$ denotes the degree of node $i$. Our healthy fix point from above is in the microscopic SIRS model given by $(s_i^*, x_i^*, r_i^*) = (1, 0, 0)$ for all $i$. We can redefine the variables relative to the fix poin as $s_i \mapsto s_i^* + s_i$, $r_i \mapsto r_i^* + r_i$ and linearise the system to obtain:

$$\dot{s}_i \simeq \beta \left( \sum_k A_{ik}(s_k^* + r_k^*) - d_i \right) s_i + \delta r_i + \beta s_i^* \sum_k A_{ik}(s_k + r_k) \tag{0.12}$$

$$\dot{r}_i \simeq -\gamma s_i - (\delta + \gamma) r_i \tag{0.13}$$

$$x_i = 1 - s_i - r_i \, , \tag{0.14}$$

For our particular fix point, the system takes the simple form of:

$$\begin{pmatrix} \dot{s}_i \\ \dot{r}_i \end{pmatrix} \simeq \underbrace{\begin{pmatrix} 0 & \delta \\ -\gamma & -(\gamma + \delta) \end{pmatrix}}_{=:\partial f} \begin{pmatrix} s_i \\ r_i \end{pmatrix} + \beta \sum_k A_{ik} \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}}_{=:\partial g} \begin{pmatrix} s_k \\ r_k \end{pmatrix} \tag{0.15}$$

This has already a form we have seen in the previous lecture, suggesting to apply the master stability function approach. Diverting to the discussion above, you find in he literature that the dynamical system is modified in a further step.

---

The **Kronecker product** of a $m_1 \times n_1$ matrix $A$ and a $m_2 \times n_2$ matrix $B$ gives a $m_1 m_2 \times n_1 n_2$ matrix as

$$A \otimes B := \begin{pmatrix} a_{11}B & \ldots & a_{1n_1}B \\ & & \\ a_{m_11}B & \ldots & a_{m_1n_1}B \end{pmatrix} \, .$$

It is bilinear and non-commutative. The following three relations hold:
- $(A \otimes B)(C \otimes D) = (AC \otimes BD)$
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
- $(A \otimes B)^\top = A^\top \otimes B^\top$

---

Using the Kronecker product, we can write the overall linear system dynamics for a state $y := (s_1, r_1, \ldots, s_n, r_n)^\top$ as:

$$\dot{y} \simeq (\mathbb{I}_n \otimes \partial f + \beta(A \otimes \partial g)) \, y \tag{0.16}$$

The important observation here is that we can independently diagonalise $A$. Assume a transformation $T$ with $T^{-1}AT = \Lambda$, $\Lambda = \mathrm{diag}(\mu_1, \ldots, \mu_n)$ and $Av_k = \mu_k v_k$. The transformed variables are $\tilde{y} := (T \otimes \mathbb{I}_2)y$.

$$\dot{\tilde{y}} \simeq (T \otimes \mathbb{I}_2)^{-1} \left( \mathbb{I}_n \otimes \partial f + \beta(A \otimes \partial g) \right) (T \otimes \mathbb{I}_2)\tilde{y} \tag{0.17}$$

$$\simeq \left( \mathbb{I}_n \otimes \partial f + \beta(\Lambda \otimes \partial g) \right) \tilde{y} \tag{0.18}$$

Hence, the system separates in $n$ independent $2 \times 2$ blocks with system matrix $\partial f + \beta\mu_k \partial g$. The latter has two distinct eigenvalues $\lambda_1 = -\delta$ and $\lambda_2 = \beta\mu_k - \gamma$. Consequently, all eigenvalues have negative real part when the following master stability condition is fulfilled:

$$\mu_{max} < \frac{\gamma}{\beta} \tag{0.19}$$

This realtes the stability of the healthy fix point to parameters of disease on the one hand and a network property on the other hand. We can combine this with bounds of the adjacency spectrum we discussed before, e.g. $\sqrt{d_{max}} \leq \mu_{max}$ (see Newman, Eq. 18.88).

## PAIR APPROXIMATION

In the formulation of the microscopic SIRS model, we used the interaction term $A_{ik}s_ix_j = A_{ik}P(i \in S)P(j \in X)$ as the probability that the disease can be transmitted from node $k$ to $i$. Technically, we need the joint probability $s_i \wedge x_k := P(i \in S \wedge k \in X)$ that $i$ is susceptible and has an infected neighbour at the same time. Generally, we can write the joint probability using the conditional as $s_i \wedge x_k = P(i \in S | k \in X)$. This highlights that our interaction term can only be true if $s_i$ and $x_k$ are independent random variables, which is typically not true as they are connected via the spreading process.

The correct model hence reads as

$$\dot{s}_i = \delta r_i - \beta \sum_k A_{ik}s_i \wedge x_k \tag{0.20}$$

$$\dot{x}_i = \beta \sum_k A_{ik}s_i \wedge x_k - \gamma x_i \tag{0.21}$$

$$\dot{r}_i = \gamma x_i - \delta r_i \tag{0.22}$$

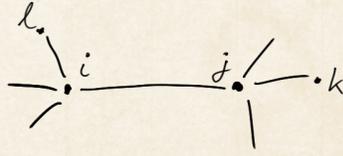$$0 = s_i + x_i + r_i - 1 \tag{0.23}$$

Figure 0.2: Sketch of the local neighbourhood

The problem now is that the system is not closed anymore. We need additional equations for all $s_i \wedge x_k$ which in turn depend on quantities like $s_i \wedge s_k \wedge x_l$ and so on . . .

Still, we can do better than our initial model by performing a so-called pair approximation to obtain a firs-order correction to the model. For the sake of simplicity, let us demonstrate this using a simplified model with $\delta = \gamma = 0$ where the only transition is between $S$ and $I$. The model reads as

$$\dot{s}_i = -\beta d_i s_i + \beta \sum_k A_{ik} s_i \wedge s_k \tag{0.24}$$

$$0 = s_i + x_i - 1 \tag{0.25}$$

We can continue be constructing the time evolution of the joint probability (see Fig. 0.2):

$$\frac{d}{dt} s_i \wedge s_j = -\beta \sum_{l \neq j} A_{il} \underbrace{x_l \wedge s_i \wedge s_j}_{\text{infection of } i} - \beta \sum_{k \neq i} A_{jk} \underbrace{s_i \wedge s_j \wedge x_k}_{\text{infection of } j} \tag{0.26}$$

$$\tag{0.27}$$

Factoring out the conditional, we see that $x_l \wedge s_i \wedge s_j = P(l \in X | i, j \in S) \cdot s_i \wedge s_j$. The pair approximation now assumes that all two-step and more distant interactions can be neglected, e.g. $P(l \in X | i, j \in S) \approx P(l \in X | i \in S)$. Inserting the approximation, we get

$$x_l \wedge s_i \wedge s_j \approx P(l \in X | i \in S) \cdot s_i \wedge s_j = \frac{x_l \wedge s_i}{s_i} s_i \wedge s_j \tag{0.28}$$

and the model becomes:

$$\dot{s}_i = -\beta d_i s_i + \beta \sum_k A_{ik} s_i \wedge s_k \tag{0.29}$$

$$\frac{d}{dt} s_i \wedge s_j \approx -\beta(d_i + d_j - 2) s_i \wedge s_j + \beta s_i \wedge s_j \left( \sum_{l \neq j} A_{il} \frac{s_l \wedge s_i}{s_i} + \sum_{k \neq i} A_{jk} \frac{s_k \wedge s_j}{s_j} \right) \tag{0.30}$$

$$0 = s_i + x_i - 1 \tag{0.31}$$

In summary, the pair approximation circumvents the recursion of joint probabilities for the cost of an additional equation for each edge in the network. This is, however, still much less then the full model (as long as the network is finite we could theoretically close he equations). In general, pair approximation is an example for what is called a *moment closure* in probability theory.

# Dynamics of Networks and Network Ensembles I

## Erdös-Renyi Random graphs (1VL)

## Configuration model (1VL)

> **Further Reading**
> This section follows the discussion in Newman's book, ch. 13.

In the lat lecture, we've learned an important lesson: When we attempt to create a random network by placing edges randomly between nodes with some fixed probability, this will always result in a binomial (Poisson for $n \to \infty$) degree distribution. This way, it is not possible to create the huge variety of network characteristics that we observe in natural, social or technological networks. The question is: How can we obtain random graphs with arbitrary degree distribution? The answer is given by the so-called **configuration model**.

### Configuration model algorithm

- Specify a degree sequence $\{d_i\}$ (or draw a sequence from a given degree distribution).
- Assign $d_i$ stubs (half-edges) to each node $i$.
- Out of all possible matchings of stubs, pick one uniformly at random (i.e. no configuration is prefered).

All possible matchings of stubs constitute a network ensemble parametrised by $\{d_i\}$. Up to permutation symmetries, many matchings yield the same graph, though. Hence, graphs with many symmetries have a higher probability to be realised as we draw uniformly from the ensemble. We denote by $p_k$ the values of the empirical degree distribution in a network realisation, denoting the fraction of nodes with degree value $k$.

Note that we have seen this approch already in Fig. 0.2 as the null model of the modularity function.

## Basic properties

***What is the prob.*** $p_{ij}$ ***of having a single edge between two nodes*** $i,j$ ***with degrees*** $d_i, d_j$***?.*** Pick one stub at node $i$. Out of the $2m - 1$ remaining stubs, there are $d_j$ associated to $j$, hence the prob. to connect one stub at $i$ with node $j$ is $d_j/(2m-1)$. Each stub at $i$ has the same independent prob., hence they add up to

$$p_{ij} = d_i \frac{d_j}{2m-1} \overset{m \gg 1}{\approx} \frac{d_i d_j}{2m} \tag{0.1}$$

Again, this should remind you of the modularity function. The latter constructs the configuration model ensemble associated to the degree sequence of a given network and compares the observed edges with the expectation value over the ensemble.

---

## Exercise: Multiedges and loops

Derive the probability to observe multiple edges between nodes $i$ and $j$ as well as the probability for loops ($i = j$) to occur.

### Hints

The considerations are analogous to the derivation of $p_{ij}$.

---

***What is the expected number*** $n_{ij}$ ***of common neighbours?.*** The calculation is fairly similar to the previous one, the counting is more subtle. For $m \gg 1$ we find:

$$n_{ij} = \sum_{k \neq i,j} p_{ik} \frac{d_j(d_k - 1)}{2m} = p_{ij} \sum_{k \neq i,j} \frac{d_k(d_k - 1)}{2m} \tag{0.2}$$

$$= p_{ij} \frac{\sum_{k \neq i,j} d_k(d_k - 1)}{n \langle d \rangle} \overset{n \gg 1}{\approx} p_{ij} \frac{\langle d^2 \rangle - \langle d \rangle}{\langle d \rangle} \tag{0.3}$$

Here and in the following discussion, $\langle d \rangle$ denotes the average w.r.t. the given degree sequence.

## Excess degree distribution

Take some random node $i$. A common quantity of interest in calculations is the joint probability $P(d_j = k \wedge A_{ij})$ of having an edge to a neighbour $j$ with degree being exactly $k$. Obviously, $P(d_j = 0 \wedge A_{ij}) = 0 \neq P(d_j = 0) = p_0$, i.e. it is not possible to reach a node with $d_j = 0$ by following an edge. The probability to find a certain degree $k$ at the end of an edge is not directly related to the probability to observe the degree value $k$ at all. The total probability that any node at the end of an edge has degree $k$ can be found to be

$$\sum_{j=1}^{n} \delta_{d_j,k} P(d_j \wedge A_{ij}) = \sum_{j=1}^{n} \delta_{d_j,k} P(A_{ij}|d_j) \cdot p_{d_j} = \sum_{j=1}^{n} \delta_{d_j,k} \frac{k}{2m-1} p_{d_j} = \frac{k}{2m-1} n p_k \approx \frac{k p_k}{\langle d \rangle} \quad (0.4)$$

Note that this relation has an important consequence. The average degree of a neighbour in general, calculated as $\sum_k k \frac{k p_k}{\langle d \rangle} = \frac{\langle d^2 \rangle}{\langle d \rangle}$ is generally different from the averge degree $\langle d \rangle$ itself. How can this be? The reason is that the counting of neigbours introduces a bias towards high-degree nodes. First of all, isolated nodes contribute to $\langle d \rangle$ but are not part of any neighbourhood. Secondly, in the neighbour-degree average, each node is counted multiple times, according to its degree. The higher the value $k$, the more a node contributes to the average.

For convenience, we will not work with the neighbour degree in the following but define the **excess degree** $q_k$ that substracts the arriving edge:

$$q_k := \frac{p_{k+1}(k+1)}{\langle d \rangle} \quad (0.5)$$

## CLUSTERING COEFFICIENT

Two neighbours $i, j$ of a node with degrees $d_i + 1$ and $d_j + 1$ have $d_i$ and $d_j$ available stubs to be again connected. The probability to observe these values is given by the excess degree distribution. Hence, the **clustering coefficient**, i.e. the prob. that two neighbours of a node are neighbours themselves, is given by:

$$C = \sum_{d_i=0}^{\infty} \sum_{d_j=0}^{\infty} q_{d_i} q_{d_j} p_{ij} \approx \sum_{d_i=0}^{\infty} \sum_{d_j=0}^{\infty} \frac{p_{d_i+1}(d_i+1)}{\langle d \rangle} \frac{p_{d_j+1}(d_j+1)}{\langle d \rangle} \frac{d_i d_j}{2m} \quad (0.6)$$

$$= \frac{1}{2m} \left[ \sum_{k=0}^{\infty} \frac{p_{k+1} k(k+1)}{\langle d \rangle} \right]^2 = \frac{1}{n} \frac{1}{\langle d \rangle^3} \left[ \sum_{l=0}^{\infty} p_l(l-1)l \right]^2 \quad (0.7)$$

$$= \frac{1}{n} \frac{(\langle d^2 \rangle - \langle d \rangle)^2}{\langle d \rangle^3} \sim \frac{1}{n} \quad (0.8)$$

Similar to Erdös-Renyi random graphs, networks in the configuration model have a clustering coefficient that vanishes for large system size.

## POLYNOMIAL GENERATING FUNCTIONS

Let's say we have a distribution $p_k$ of a variable with discrete values $k$ e.g. the degree or excess distribution. Calculations often involve expressions for the $p_k$ that are

difficult to be evaluated directly. Here, **generating functions** provide a useful tool that might help. They are defined as:

$$g(z) := \sum_{k=0}^{\infty} p_k z^k = p_0 + p_1 z + p_2 z^2 + \ldots \tag{0.9}$$

As they are polynomial, it is straight-forward to find some useful properties:

$$g(1) = \sum_{k=0}^{\infty} p_k = 1 \tag{0.10}$$

$$g(0) = p_0 \tag{0.11}$$

$$\Rightarrow p_k = \left[ \frac{1}{k!} \frac{d^k}{dz^k} g(z) \right]_{z=0} \tag{0.12}$$

$$\frac{d}{dz} g(1) = \sum_{k=0}^{\infty} k p_k = \langle k \rangle \tag{0.13}$$

$$z \frac{d}{dz} g(z) = z \sum_{k=0}^{\infty} k p_k z^{k-1} = \sum_{k=0}^{\infty} k p_k z^k \tag{0.14}$$

$$\Rightarrow \langle k^m \rangle = \left[ \left( z \frac{d}{dz} \right)^m g(z) \right]_{z=1} \tag{0.15}$$

Essentially, knowing only the generating function $g(z)$, it is possible to recover the full distribution and all of its moments!

Furhtermore, generating functions share the so-called **power property**. Pick a sample of $m$ nodes uniformly at random. There degrees are independent outcomes, hence the prob. that the sample has a certain degree sequence $\{d_i\}$ is $P(\{d_i\}) = \prod_{i=1}^{m} p_{d_i}$. What is the probability $\pi_s$ that the total number of stubs in the sample is exactly $s = \sum_{l=1}^{m} d_l$?

$$\pi_s = \sum_{d_1=0}^{\infty} \cdots \sum_{d_m=0}^{\infty} \prod_{i=1}^{m} p_{d_i} \delta(s, \sum_{l=1}^{m} d_l) \tag{0.16}$$

$$h(z) = \sum_{s=0}^{\infty} \pi_s z^s = \sum_{s=0}^{\infty} z^s \sum_{d_1=0}^{\infty} \cdots \sum_{d_m=0}^{\infty} \prod_{i=1}^{m} p_{d_i} \delta(s, \sum_{l=1}^{m} d_l) \tag{0.17}$$

$$= \sum_{d_1=0}^{\infty} \cdots \sum_{d_m=0}^{\infty} \prod_{i=1}^{m} p_{d_i} z^{\sum_{l=1}^{m} d_l} = \sum_{d_1=0}^{\infty} p_{d_1} z^{d_1} \cdots \sum_{d_m=0}^{\infty} p_{d_m} z^{d_m} \tag{0.18}$$

$$= (g(z))^m \tag{0.19}$$

Here, $g(z)$ denotes the generating function of the degree distribution. This result is known as the power property of generating functions and applies generrly to sums of independent random variables. It is now easy to obtain $\pi_s$ directly from $g(z)$ as

$$\pi_s = \left[ \frac{1}{s!} \frac{d^s}{dz^s} g^m(z) \right]_{z=0} \tag{0.20}$$

As an example, consider the Erdös-Renyi random graph with Poisson degree distribution $p_k = e^{-c} \frac{c^k}{k!}$ from the previous lecture. The degree generating function is

$$g(z) = e^{-c} \sum_{k=0}^{\infty} \frac{c^k}{k!} z^k = e^{-c} e^{cz} = e^{c(z-1)} \tag{0.21}$$

and the generating function of $\pi_s$ is directly determined via the power property to be

$$h(z) = g^m(z) = e^{mc(z-1)} \ . \tag{0.22}$$

Hence, we find

$$\pi_s = \left[ \frac{1}{s!} \frac{d^s}{dz^s} e^{mc(z-1)} \right]_{z=0} = e^{-mc} \frac{(mc)^s}{s!} \tag{0.23}$$

to be Poisson-distributed as well.

Finally, for ease of notation, we introduce the following two quantities. The degree generating function is denoted as

$$g_0(z) := \sum_{k=0}^{\infty} p_k z^k \ . \tag{0.24}$$

Furthermore, we introduce the excess degree generating function

$$g_1(z) := \sum_{k=0}^{\infty} q_k z^k = \sum_{k=0}^{\infty} \frac{(k+1) p_{k+1}}{\langle d \rangle} z^k = \frac{1}{\langle d \rangle} \sum_{l=0}^{\infty} l p_l z^{l-1} = \frac{g_0'(z)}{\langle d \rangle} = \frac{g_0'(z)}{g_0'(1)} \tag{0.25}$$

which is by definition closely related to $g_0(z)$.

## GIANT COMPONENT THRESHOLD

As for the Erdös-Renyi ensemble, we are interested to find the transition (if it exists) towards the appearance of a giant connected component. To start, let us first consider the number of second neighbours of a node. Having $k$ second neighbours introduces the constraint that the excess degrees $e_r$ of the first neighbours sum up to $k$. Note that again we neglect links between neighbours and assume a locally tree-like network as earlier. See Newman's book for a more detailed discussion. We can define the probability of having $k$ second neighbours, given the excess degree sequence $\{e_r\}$ of the $m$ neighbours, as:
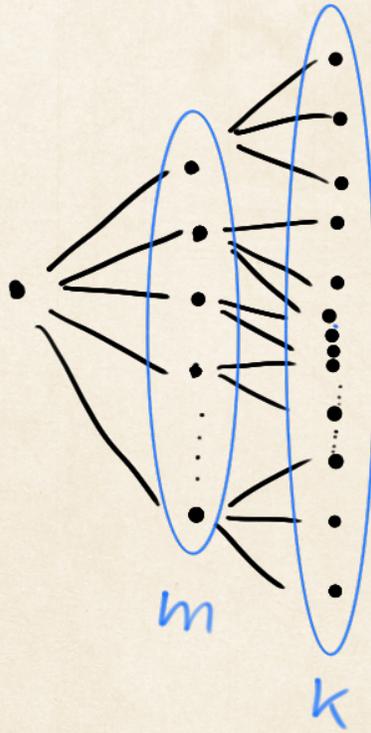
Figure 0.1: Illustration of neighbourhood sets.

$$p_k^{(2)} := \sum_{m=0}^{\infty} P(k|m \text{ neighbours}) \cdot p_m \tag{0.26}$$

$$= \sum_{m=0}^{\infty} p_m \cdot \left( \sum_{e_1} \cdots \sum_{e_m} \prod_{r=1}^{m} q_{e_r} \delta(k, \sum_{l=1}^{m} e_l) \right) \tag{0.27}$$

To simplify this expression, we use again the power property and construct the generating function of $p_k^{(2)}$.

$$g^{(2)}(z) := \sum_{k=0}^{\infty} p_k^{(2)} z^k = \sum_{m=0}^{\infty} p_m \left( \sum_e q_e z^e \right)^m = \sum_{m=0}^{\infty} p_m \left( g_1(z) \right)^m = g_0 \left( g_1(z) \right) \tag{0.28}$$

You can check easily that

$$g^{(2)}(z) = g_0 \left( g_1 \left( g_1(z) \right) \right) \tag{0.29}$$

$$\dots \tag{0.30}$$

$$g^{(d)}(z) = g^{(d-1)} \left( g_1(z) \right) \tag{0.31}$$

Although we have found this recursion relation, actually calculating $p_k^{(2)}$ is still a

hard problem for general degree distributions. Nevertheless, it is possible to obtain its expectation value, the average number of second neighbours $c^{(2)}$ as:

$$c^{(2)} = \left[ z \frac{d}{dz} g^{(2)}(z) \right]_{z=1} = [z g_0'(g_1(z)) g_1'(z)]_{z=1} = g_0'(g_1(1)) g_1'(1) = g_0'(1) g_1'(1) \tag{0.32}$$

$$= g_0''(1) = \sum_{k=0}^{\infty} k(k-1) p_k = \langle k^2 \rangle - \langle k \rangle \tag{0.33}$$

However, we can also use the recursion relation to calculate the average number of neighbours $d$ steps away as

$$c^{(d)} = g^{(d)'}(1) = g^{(d-1)'}(g_1(1)) g_1'(1) = c^{(d-1)} g_1'(1) = c^{(d-1)} \frac{c^{(2)}}{g_0'(1)} = c^{(d-1)} \frac{c^{(2)}}{c} \tag{0.34}$$

$$\Rightarrow c^{(d)} = \left( \frac{c^{(2)}}{c} \right)^{d-1} c \tag{0.35}$$

where $c$, the average number of neighbours, is just the mean degree $\langle k \rangle$. Consequently, we need to distinguish two cases. Firstly, when $c^{(2)} > c$, the average number of neighbours grows exponentially from step to step until the component (to which the arbitrary starting node belongs) reaches the finite system size. In this case a giant component can exist. Secondly, if $c^{(2)} < c$ the expected component size is exponentially suppressed for any starting node such that there can be no giant component. Hence, $c^{(2)} = c$ is the existence threshold for a giant connected component.

Inserting our result for $c^{(2)}$, we finally find the following condition:

$$\langle k^2 \rangle - 2\langle k \rangle \overset{!}{>} 0 . \tag{0.36}$$

Remember that the generating function of the Poisson distribution is $g(z) = e^{c(z-1)}$. It is straight-forward to calculate the second moment as $\langle k^2 \rangle = \langle k \rangle^2 + \langle k \rangle$. It follows that the giant component threhold of Poisson-distributed random graphs is just $\langle k \rangle > 1$ as we have seen in the previous lecture on Erdös-Renyi ensembles.

# CHRISTMAS: SOCIAL NETWORK GROWTH (1VL)

Just read the paper. ;-)

# DYNAMICS ON NETWORKS II: EMERGENT PHENOMENA

## FLOW NETWORKS (1 VL)

The first dynamical phenomenon on networks discussed in the first part were diffusion and (epidemic) spreading processes. Before we continue with emergent nonlinear processes, **flow networks** deserve closer attention. They are closely linked to diffusion and to steady states of nonlinear processes in general. In flow networks, nodes are typically sinks or sources of some quantity which is transported along the edges of the graph, which can have a finite capacity. There are many examples, ranging from transport, trade or financial networks to food webs or functional climate networks. The respective quantities propagated by the "flow" are vehicles, goods, money, nutrients or information.

A peculiar phenomenon in flow networks is that typically, the flow pattern responds with non-local changes to local changes of the capacity (e.g adding/removing an edge). A paramount example is that of **Braess' paradox** in traffic networks. Assume a road network populated by cars, each choosing the fastest route between start and destination. When the capacity of an existing road is improved to diminish congestions or a new one is opened to diversify traffic routes, there are certain cases where the travel time is actually increased for all cars. In other words, improving the capacity of the flow network causes a decrease in performance. The converse is true for closing a road. Braess' paradox highlights how a non-local change of flow in response to a local modification alters the flow pattern in a way that is difficult to predict and that it is insufficient to consider only the local situation in simulations. This is supported by various empirical measurements gathered since Braess' discovery in 1968. Note that Braess' paradox is a rather general phenomenon which can also be observed, for instance, in power grids.

### DEFINITIONS

Define a **flow network** as the pair $\mathcal{F}(\mathcal{G}, C)$ of a Graph $\mathcal{G} = (V, E)$ with node set $V$ and edge set $E$, as well as the **capacity** $C : E \to \mathbb{R}_+$ s.t. $e \in E \mapsto C(e) \geq 0$. A **flow** is a map $f : E \to \mathbb{R}_+$ s.t. $e \in E \mapsto f(e)$ with the following properties:

- $f_{ij}$ is a shorthand for $f(e)$, where the head of $e$ is node $i$ and the tail is node $j$.
- $f$ is anti-symmetry (introduces an orientation): $f_{ij} = -f_{ji}$.

- $f$ is capacity-constrained: $f(e) \leq C(e) \; \forall e \in E$
- The flow is conserved.

The flow conservation requires that the net flow $I_k$ across each node vanishes, except for sinks and sources. Assume without loss of generality there is one source $s$ and one sink $t$:

$$I_k = \sum_{l=1}^{n} A_{kl} f_{kl} = I_{in}\delta_{ks} - I_{out}\delta_{kt} \qquad (0.1)$$

As a consequences, we have $I_{in} = -I_{out} =: I$ and the total in-flow has to equal the total out-flow.

## THE LINK TO DIFFUSION

Often, flows are a function of the differences in the value of a potential that takes values at each node. For instance, in an electrical network the flow $f_{ij}$ on an edge is the current induced by the difference in the node voltages $v_i$ and given by Ohm's law.

$$f_{ij} = \frac{v_i - v_j}{r_{ij}}, \quad r_{ij} > 0 \qquad (0.2)$$

Flow conservation implies

$$I(\delta_{ks} - \delta_{kt}) = \sum_{l=1}^{n} A_{kl} \frac{v_k - v_l}{r_{kl}} = \sum_{l=1}^{n} L_{kl} v_l \; . \qquad (0.3)$$

Here, $L$ is a weighted Laplacian matrix where the resistances $r_{ij}$ define the edge weights. This expression of flow conservation is known as **Kirchhoff's current law** in electrostatics. Moreover, this is exactly the Laplace problem obtained earlier for diffusion. It appears generically in flow networks whenever the flow is proportional to a difference. Hence, this functional form is commonly termed **diffusive coupling** in network dynamics.

Looking for solutions of Kirchhoff's equation in terms of the yet undetermined voltages poses a problem since the Laplacian is not invertible. We can however

construct a so-called **Moore-Penrose pseudoinverse** $R$ to obtain solutions which are unique up to an additive constant. A convenient definition is

$$\mathbf{R} := (\mathbf{L} + \mathbf{J})^{-1} - \frac{1}{n^2}\mathbf{J} \tag{0.4}$$

$$\tag{0.5}$$

where $\mathbf{J}$ is the unit matrix such that all $\mathbf{J}_{ij} \equiv 1$.

> Every pseudoinverse $R$ of $L$ needs to fulfill the Moore-Penrose condition:
>
> $$\mathbf{LRL} = \mathbf{L}, \quad \mathbf{RLR} = \mathbf{R}, \quad (\mathbf{RL})^\top = \mathbf{RL}, \quad (\mathbf{LR})^\top = \mathbf{LR} . \tag{0.6}$$

Finally, the solutions $v^*$ are obtained as:

$$v^* = \mathbf{R}I(\chi_s - \chi_t) + \text{const.} \tag{0.7}$$

where $\chi_s$ is an indicator vector with entries $(\chi_s)_k = \delta_{ks}$.

## EXERCISE: PSEUDOINVERSE

1. Derive the form of the pseudoinverse given above.
2. Check that $R$ fulfills the Moore-Penrose conditions.

### HINTS

***Ansatz.*** Evaluate the expression $(L + \alpha J)(R + \beta J)$ with some $\alpha, \beta \in \mathbb{R}$. Furthermore, require $LR = \mathbf{I}_n - \frac{1}{n}J$. (It's easy to see that this complies with the Moore-Penrose conditions.)

***Spectrum.*** You may use the fact that $R$ has a singular spectrum to evalute $RJ$.

## NEWMAN'S BETWEENNESS

> **FURTHER READING**
>
> Newman, Mark EJ. "A measure of betweenness centrality based on random walks." Social networks 27.1 (2005): 39-54.
>
> `arxiv.org/pdf/cond-mat/0309045.pdf`

Now that we solved the Laplace problem arising from the flow conservation, we can calculate the actual corresponding flow which we will also denote by $f$.
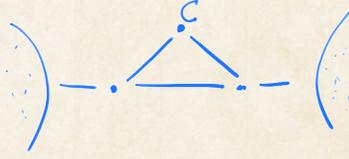
Figure 0.1: Difference between shortest-path and current-flow betweenness.

$$f_{ij} = \frac{v_i^* - v_j^*}{r_{ij}} = \frac{1}{r_{ij}}(\chi_i - \chi_j)^\top v^* = \frac{I}{r_{ij}}(\chi_i - \chi_j)^\top \mathbf{R}(\chi_s - \chi_t) \tag{0.8}$$

$$= \frac{I}{r_{ij}}(R_{is} + R_{jt} - R_{it} - R_{js}) \tag{0.9}$$

The share of the current between $s$ and $t$ that goes via node $i \neq s, t$, i.e. the through-current $c_i^{st}$, is given by halve the sum of the modulus as you can easily verify:

$$c_i^{st} = \frac{1}{2}\sum_{j=1}^n |A_{ij} f_{ij}| = \frac{I}{2}\sum_{j=1}^n \frac{A_{ij}}{r_{ij}}|R_{is} + R_{jt} - R_{it} - R_{js}|\,, \quad i \neq s, t\,, \tag{0.10}$$

with $c_s^{st} = I$ and $c_t^{st} = -I$. This quantity, averaged over all pairs $s/t$ of source/target nodes became known as the **current-flow** or **Newman's betweenness** $b_i^f$ associated to a node $i$ :

$$b_i^f = \frac{2}{(n-1)(n-2)}\sum_{s<t}\frac{c_i^{st}}{I}\,, \quad b_i^w \in [0;1]\,. \tag{0.11}$$

Newman's betweenness can be interpreted as a measure of a node's importance in a flow network. Contrary to the popular shortest-path betweenness, $b_i^f$ does not only take shortest-paths between $s$ and $t$ into account but all according to their weight $r_{ij}$. Fig. 0.1 highlights a case where both measures differ drastically, as there is no shortest path through the central node C.

# MAX-FLOW/MIN-CUT THEOREM

**FURTHER READING**

Edmonds, Jack, and Richard M. Karp. "Theoretical improvements in algorithmic efficiency for network flow problems." Journal of the ACM (JACM) 19.2 (1972): 248-264

Ford Jr, Lester R., and Delbert R. Fulkerson. A simple algorithm for finding maximal network flows and an application to the Hitchcock problem. No. RAND/P-743. RAND CORP SANTA MONICA CA, 1955.

Meyner, Felix. "Einführung in Netzwerk-Fluss-Probleme Das min-cut max-flow Theorem." Innovative Internet Technologies and Mobile Communications (IITM) 33 (2012).

Finally, we turn to the question: What is the maximum possible flow on $\mathcal{F}$ that is still capacity-constrained?

This is adressed by the so-called **max-flow/min-cut theorem**:

> The maximum flow from the sources to the sinks of a flow network is determined by the total capacity of the minimum cut.

The essential statement of this theorem is that maximum flow is determined by the bottlenecks of the flow network. This consequence should be intuitively clear, take for instance the case sketched in Fig. 0.1. The extent of the bottleneck, i.e. the minimum cut, however is interesting in itself as it measures the robustness of the flow network in some sense. Hence, it is often of interest to determine the minimum cut for a given network.

First, we need some definitions. Let $\mathcal{C} = (S, T)$ be a crisp partition of the node set $V$ in two parts $S$ and $T$, s.t. $s \in S$ and $t \in T$. The **cut set** is then defined as $X_\mathcal{C} := \{(l, k) \in E : l \in S \land k \in T\}$. It has the associated **cut capacity** $C(X_\mathcal{C}) = \sum_{e \in X_\mathcal{C}} C(e)$.

The **minimum cut** is given by $\mathcal{C}_{min} = \min_\mathcal{C} C(X_\mathcal{C})$.

The **maximum flow** is given by the total in-flow due to flow conservation as $|f|_{max} = \max_{(s,t)} \sum_{(s,k) \in E} f_{sk}$.

The statement of the theorem is then $\mathcal{C}_{min} = |f|_{max}$. There are various algorithms to determine $X_{\mathcal{C}_{min}}$ associated to $\mathcal{C}_{min}$ via maximising the flow. As an example, we consider the **Ford-Fulkerson algorithm** from 1956.

The algorithm uses the concept of an augmenting path $p$, i.e. a path with free capacaty on all edges $p = \{e_1, \ldots, e_m\} : \forall_{e \in p} f(e) < C(e)$ directed from the source $s$ to the sink $t$. The Ford-Fulkerson algorithm iteration then uses the following iteration procedure.

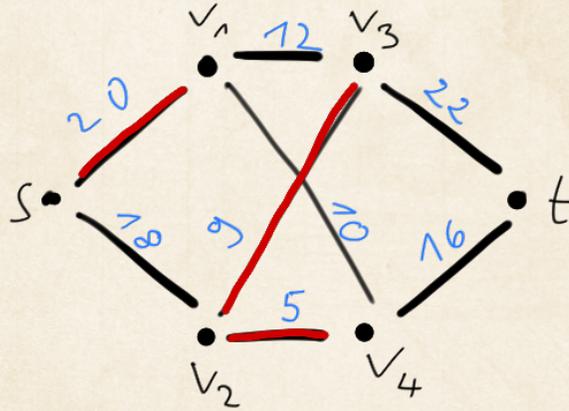1. Initially, all paths between $s$ and $t$ are augmented.

Figure 0.2: Example flow network.

2. Pick one path $p$ at random from all available independent augmented paths.

3. Determine the minimum residual capacity $c_p = \min_{e \in p} C(e)$ along $p$.

4. Add the edge associated to $c_p$ to the potential minimum cut set $\hat{X}_C$.

5. Substract $c_p$ from all $C(e)$, $e \in p$.

6. Repeat from step 2 until no augmented path is left.

An actual minimum cut $X_C \subseteq \hat{X}_C$ is determined as the subset of the potential minimum cut that is a proper cut set. However, as $\hat{X}_C$ usually contains only very few edges compard to the total size of $E$, this minimisation problem is much simpler than the original one. A variation of the algorithm that only considers shortest $st$-paths was given 1972 by Edmonds and Karp.

Consider the example in Fig. 0.2. A possible sequence of paths in the iteration scheme is:

1. $p_1 : s, v_1, v_3, t$ with $c_p = 12$ on $(v_1, v_3)$

2. $p_2 : s, v_2, v_4, t$ with $c_p = 5$ on $(v_2, v_4)$

3. $p_3 : s, v_1, v_4, t$ with $c_p = 8$ on $(s, v_1)$

4. $p_4 : s, v_2, v_3, t$ with $c_p = 9$ on $(v_2, v_3)$

Hence, the potential minimum cut is $\hat{X}_C = \{(s, v_1), (v_1, v_3), (v_2, v_3), (v_2, v_4)\}$. This is, however, not a a minimal cut set as one is redundant. By removing the edge $(v_1, v_3)$ we finally obtain a minimum cut set $X_{C_{min}} = \hat{X}_C \setminus \{(v_1, v_3)\}$ and the maximum flow is $|f|_{max} = 20 + 9 + 5 = 34$.