Chapter 2
Discrete Structures

Abstract
Questions:

- How can the cells of an organism which all share the same genes can fulfill so many different functions?
- Are there good mathematical tools to identify the important features in all those networks that modern biological data collection produces?
- How long ago did the last common ancestor of two species or two individuals live?

A model of combinatorial gene regulation shows the power of combinatorics. Graphs are useful tools for network analysis, and their spectral theory is developed. Phylogenetic relationships between species are modeled by particular types of graphs, the trees. Descendence relations between individuals involve two parents and lead to genealogies. Coalescents treat the question of common ancestors. Such structures also naturally lead to the stochastic processes treated in the Chap. 3.

2.1 Introductory Example: Gene Regulation and the Power of Combinatorics

In this section, I present an example of a combinatorial scheme in molecular biology. This is meant to show that even elementary mathematical reasoning can help us to clearly understand a biological situation that may initially look rather complicated. First, however, I shall sketch the most basic principles of molecular biology. More details can be found in standard textbooks, like [1] or [93].
Metabolism and other fundamental functions of the cell are essentially carried out by proteins. The building blocks of proteins are polypeptides, sequences of typically a few hundred amino acids that fold into particular three-dimensional shapes according to attractive forces between different amino acids and interactions with water molecules in the cell. A protein consists of one or several such polypeptides, and

J. Jost, *Mathematical Methods in Biology and Neurobiology*, Universitext,
DOI: 10.1007/978-1-4471-6353-4_2, © Springer-Verlag London 2014

its three-dimensional shape determines its function. The information for the particular sequence of each polypeptide is contained in the DNA of the cell. The DNA is a sequence itself, consisting of nucleotides instead of amino acids, and the DNA is inherited by the daughter cells under cell division and the germ cells in sexual recombination. This will now be described with some more details and precision.
The fundamental process of molecular biology then is gene expression, that is, the production of polypeptides, the building blocks of proteins, according to the genetic information contained in the DNA of a cell. The DNA (desoxyribonucleic acid) is a long string of base pairs, arranged in the shape of a double helix, as discovered by Watson and Crick. There are four different nucleotide bases, labelled A, C, G, and T (we are not concerned here with their precise chemical identity, and so, these letters may suffice for our purposes). Thus, each of the two strands of the double helix is a long sequence composed of these 4 "letters". Each strand determines the identity of the complementary strand, because C in one strand is paired with G in the other, and A with T. Therefore, when the double helix is split apart, each strand contains the complete information for assembling a new such double helix. This is the principle underlying genetic inheritance. Here, however, we are not concerned with inheritance, but rather with gene expression. The first step of gene expression, called transcription, then consists in copying the information in a segment of one of the strands into another macromolecule, RNA (ribonucleic acid), which is chemically more active and flexible. It also consists of sequences composed of 4 letters, A, C, G as in the DNA and a new letter U taking the place of T. Again, this copying works according to the above complementarity principle. Which segments of the DNA are thus copied under a given cellular condition is controlled by certain proteins, the transcription factors that typically bind to locations in the DNA nearby those to be copied and that can then trigger, enhance or block the transcription process [26]. Of course, one and the same stretch of DNA can be repeatedly transcribed, and the regulation of the number of such transcripts is essential, but we shall not emphasize this aspect in the sequel. The resulting RNA is then further processed, through interactions with itself or with other RNAs or with certain proteins again. The final mRNA (m standing for "messenger") can then be translated into a polypeptide, in a certain complex called the ribosome, with the help of some other auxiliary RNA, the rRNA (r standing for "ribosomal"). The principle of the translation is that the unit of translation in the mRNA is a triplet of nucleotides, like *ACG* or *UAA*, also called a codon. Each such triplet is translated into a specific amino acid, and the resulting polypeptide thus is a sequence of amino acids. Since there are 64 possible triplets, but only 20 amino acids, several different triplets can correspond to the same amino acid. This fact is called the degeneracy of the genetic code, although redundancy might be the more accurate word. (Actually, the triplet *UGA* has a special role: It serves as the stop codon, that is, when this triplet is encountered in the ribosomal complex, the polypeptide is released, and a new translation can start.) In fact, the relation between such triplets and amino acids is mediated by another type of RNA, called tRNA (t for "transfer"). Chemically, this relation, called the genetic code, that is, which triplet is translated into which amino acid, is arbitrary, and so the question emerges why the translation rules are as they are, instead of being different. That is, why is for instance

GCC translated into the amino acid alanine, instead of, say, cysteine? Is that simply a historical accident, an arbitrary rule that all living creatures have inherited from their common ancestor who had adopted these translation rules by chance? Or are there some chemical or formal principles behind this, like symmetry considerations or coding efficiency? There have been many different speculations about this issue, but none so far has met with general approval.

One or more polypeptides then are combined into a protein. An important point is that a protein is not simply an amino acid sequence, but that for its molecular function, it assumes a specific three-dimensional shape. This shape, is determined by chemical attraction and repulsion between different pieces, but the details are very intricate, and the problem of computing the three-dimensional shape of a protein, or better, the process, called protein folding, by which it acquires this shape from its constituting amino acid sequence is not yet fully solved, despite considerable attempts by many mathematicians and physicists.

The fundamental question for a cell then is which genes to express when, under which circumstances. The mechanism of the cell for answering this question is gene regulation. I have already described that specific proteins, the transcription factors, trigger or inhibit the transcription of DNA segments. In eukaryotic cells (the cells that we are made of, those containing a nucleus, in contrast to prokaryotic cells, without nucleus, like bacteria), the most important part of gene regulation, however, seems to take place at the level of RNA rather than DNA. First of all, the transcribed RNA, called pre-mRNA, is reassembled in a process called splicing into mRNA. Here, on one hand, certain segments, the so-called introns, are cut out whereas the remaining ones, the exons, can then be assembled possibly in different ways, so as to produce different results from one and the same stretch of DNA [13], or pieces of different origins can be put together or interact in other ways. The processing on one hand is based on the spatial configuration assumed by an RNA molecule, on the basis of bindings between complementary nucleotides (*A* with *U* or *C* with *G*), no longer between different strands as in the DNA, but now between bases in one and the same RNA sequence [59]. On the other hand, it results from interactions with certain other small RNAs, the so-called miRNAs (mi for "micro") or siRNAs (si for "small interfering" or "silencing") or with specific proteins. These proteins bind to RNA molecules to form so-called RNP complexes (where P stands for "protein") [119]. Much of this RNA regulation works as repression, that is, preventing the mRNA from being translated. The biological rationale for this is that on one hand, the production of RNA is energetically cheap, and on the other hand, with mRNA already around, it is much faster to produce the corresponding proteins than if the process had to start anew from the DNA level. Thus, the cell can respond much quicker to new circumstances. (For a systematic analysis, see [103, 104] and the subsequent discussion in the journal Theory in Biosciences, see [105].)

After the genome of humans (and several other species) has been sequenced, that is, the the identity of all the 3 billion letters in the DNA sequence has been established [63], now the ENCODE project systematically records and catalogues all the different RNA molecules that can be present in human (and other) cells [34, 38, 49]. The genetic sequence contains both coding information that can be potentially

activated and utilized in a cell with the assistance of specific proteins, and important structural elements. But we need to identify all the different RNAs and understand their interactions with other RNAs and proteins in order to understand the regulation of gene expression in the active cell.

Now, obviously, the scheme described offers many possibilities for combinatorial reasoning as a formal description of the rules governing those processes. Here, as an example I shall discuss a model that arises from my work with the molecular biologist Klaus Scherrer, see [76]. The important point here is that the nucleotides in an mRNA can assume two different roles simultaneously. On one hand, they are parts of coding triplets (except for certain portions at the beginning or end of an mRNA sequence). On the other hand, stretches of about 30 nucleotides can function as binding sites for specific proteins which then regulate the fate of the mRNA, as explained (see [103, 104]). We call such a regulatory stretch of nucleotides an oligomotif. In the basic version of the model, there then is a one-to-one correspondence between such oligomotives and mRNA binding proteins. That is, there is a second, regulatory, code superimposed upon the first code, the genetic code governing translation. In both cases, however, the chemical identity of the nucleotides involved is crucial. An average mRNA may then possess about 20 such oligomotives. The ground state then is when the corresponding proteins are attached to all those 20 oligomotives. In this state, the mRNA is repressed and not translated. It only becomes available for translation when at least 3 of those proteins are removed. (We shall call such a set of 3 oligomotives, or equivalently, of 3 mRNA binding proteins, a triple, not to be confused with the triplet of the genetic code.) That is, when a signal arrives in the cell that causes the release of 3 such binding proteins, the corresponding mRNA gets translated, and a specific polypeptide is produced. Now, however, in a given situation, a cell needs not only one type of polypeptide, but a suitable combination of perhaps hundreds of polypeptides. The preceding structure now offers an elegant scheme for the coordinated expression of groups of genes, that is, the coordinated production of specific combinations of polypeptides and proteins. First of all, there are then $\binom{20}{3} = 1,140$ different possibilities for such triples of oligomotives. The key point now is that different mRNAs will share some, but not all of their oligomotives. That is, whenever we identify 3 proteins for removal, that is, select 3 oligomotives, we then get a specific set of mRNAs that contain those 3 oligomotives and that will then get translated, whereas the remaining ones will stay repressed. And when we select a different set of 3 oligomotives, we obtain a different combination of mRNAs to be translated, hence a different combination of proteins in the cell. This set may partially overlap with the preceding one, depending on the distribution of oligomotives across the different RNAs. In fact, one estimates that there are about 3,000 different mRNA binding proteins, hence also about 3,000 different oligomotives according to the model. We thus have $\binom{3,000}{20}$ different possibilities to distribute the oligomotives across the mRNAs (there are perhaps around 10,000 different mRNAs in a typical mammalian cell).

Let us now look into this scheme in more numerical detail. As explained, in order that several mRNAs participate in the same condition, they need to share at least 3 oligomotives. And when some mRNAs share m oligomotives ($3 \leq m \leq 20$), they can

simultaneously participate in $\binom{m}{3}$ conditions. This number varies from 1 (for $m = 3$) to 1,140 (for $m = 20$). However, when $m = 20$, that is, when the mRNAs share all their oligomotives, they can no longer be distinguished in this scheme. Let us consider some numerical examples, on the basis of the general scheme. For K oligomotives, there are $\binom{K}{20}$ different possibilities to choose 20 among them. This means that we can distinguish that many mRNAs through their different endowments with 20 out of these K oligomotives. As explained, a condition for translation is achieved by the selection of 3 (or more) out of these K oligomotives. Every choice of $\binom{K}{3}$ yields a different condition. Precisely those mRNAs will participate in such a condition that carry all those 3 oligomotives. Thus, 3 out of their 20 oligomotives are fixed, and 17 remain for free choice. That is, we have $\binom{K-3}{17}$ different possibilities. Thus, assuming that all the above $\binom{K}{20}$ possibilities are realized, by selecting 3 oligomotives, we select $\binom{K-3}{17}$ different mRNAs. Here are simple numerical examples.

- Distribute 21 oligomotives among 21 mRNAs (20 oligomotives/mRNA) so that each mRNA is identified by which oligo it does not contain. By specifying 3 oligomotives, any of the possible $\binom{21}{3} = \binom{21}{18} = 1,330$ combinations of 18 mRNAs can then be selected. Here, we have only relatively few different mRNAs.
- Distribute 23 oligomotives among $\binom{23}{3} = 1,771$ mRNAs (20 oligomotives/mRNA) so that each mRNA is identified by which 3 oligomotives it does not contain. By specifying 3 oligomotives, any of the possible $\binom{23}{3} = \binom{23}{20} = 1,771$ combinations of $\binom{20}{3} = 1,140$ mRNAS can be selected. Here, we obtain a large collection of selected mRNAs.
- Distribute 22 oligomotives among $\binom{22}{2} = 231$ mRNAs (20 oligomotives/mRNA) so that each mRNA is identified by which 2 oligomotives it does not contain. By specifying 3 oligomotives, any of the possible $\binom{22}{3} = \binom{22}{19} = 1,440$ combinations of $\binom{19}{2} = 171$ mRNAs can be selected. This is a biologically reasonable number.

Obviously, the number 3,000 of different mRNA binding proteins, that is, of different oligomotives is far larger than needed in our model. This indicates that, in reality, gene regulation at mRNA level is more complex than captured by the model. Nevertheless, the model should describe a core principle of regulation. Moreover, there is an interesting combinatorial problem suggested by this model: How to distribute K labels among N units so that each unit receives k of them so that by selecting $\kappa < k$ of them (for which we have $\binom{k}{\kappa}$ different possibilities), we identify the maximal number of different subsets of those N units? We may here wish to constrain those subsets to be of some fixed size n, or to be within a certain size range, say between n_1 and n_2.

In order to understand the mathematical structure of this problem better, it is helpful to translate it into a combinatorial design problem. We consider an $N \times K$ matrix with entries 1 or 0 where each of the N rows has precisely k 1s, and hence $K - k$ 0s. For $\kappa < k$, we then want to find collections of rows that have (at least) κ 1s in common. The question then is how to distribute the 1s in the rows so as to find as many such collections as possible within a given size range.

No full solution seems to be known for this problem. In any case, the example is meant to show that by elementary mathematical reasoning, we can come up with clever ways of how a cell could regulate its genes so that in one situation, in a single stroke, it can co-activate specific groups of genes, and in another situation, again in a single stroke, it can activate another set of genes, perhaps partly overlapping with the first one, without having to address all these genes individually. This is the power of combinatorics.

2.2 Graphs and Networks

2.2.1 Graphs in Biology

A graph is the mathematical structure representing binary relationships between discrete elements. These elements are the vertices of the graph, and the relationships are encoded as connections or edges between vertices. Such a graph can then be a network, that is, the substrate of dynamical interactions carried by the edges between processes located at the vertices. Biological applications abound.

In neural networks, the vertices stand for neurons, and the edges for synaptic connections between them. The interaction is the electrochemical transmission of pulsed dynamical activity, the spikes generated in the neurons. This activity is considered to be the carrier of information, enabling cognitive processes, but the precise identification of the information inside that dynamical activity remains unclear at present.

At smaller scales, the vertices can represent molecules like proteins, and the edges again interactions between them. The vertices can also stand for genes, and the edges for correlations in expression patterns indicating functional interactions.

At larger scales, the vertices can be the members of a population, and the edges social or other interactions, like mating. For a population with separate sexes, we then have a bipartite graph, that is, one with two distinct classes of elements such that edges exist only between members of opposite classes, but not inside one class.

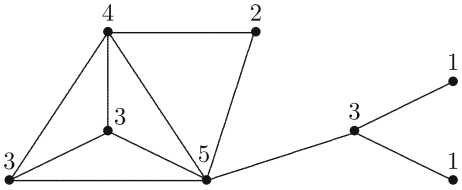
At the still larger scale of ecosystems, the vertices can represent species, and the edges stand for trophic interactions. The graph then encodes a food web.

Another important class of biological graphs are the phylogenetic trees that turn genetic or other similarities between species into descendence relations from common ancestors. For individual descendence relations inside a sexually recombining species we rather have pedigrees because each individual then has two parents which in turn may have more than one offspring.

For detailed studies of biological networks and their properties, the reader can consult [94] and [111] and the many references therein.

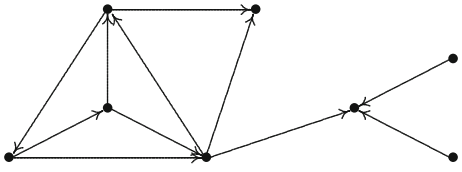
2.2.2 Definitions and Qualitative Properties

We now display some formal definitions and start with the simplest situation. A graph Γ is a pair (V, E) of a finite set V of vertices or nodes and a set E of unordered pairs, called edges or links, of different elements of V (and we assume $E \neq \emptyset$ to make the graph nontrivial). Thus, when there is an edge $e = (i, j)$ for $i, j \in V$, we say that i and j are connected by the edge e and that they are neighbors, $i \sim j$. Defining edges as unordered pairs of vertices means that we consider (i, j) and (j, i) as the same pair. Thus, the neighborhood relation is symmetric. Requiring that the vertices connected by an edge be different then means that there are no edges connecting a vertex to itself. Thus, the neighborhood relation is not reflexive. In general, it is not transitive either, that is, $i \sim j$ and $j \sim k$ need not imply $i \sim k$. The degree n_i of the vertex i is the number of its neighbors. Also, the order $|\Gamma|$ is the number of vertices in Γ, i.e., the cardinality of the vertex set V.



A graph $\Gamma$of order 8, with vertex degrees indicated (2.2.1)

So far, we are assuming that the edges are undirected, that is, the edge (i, j) is the same as (j, i). One may, naturally, also consider directed graphs, that is, where an edge $e = (i, j)$ is considered to go from i to j rather than connect i and j in a symmetric manner. For example, this is appropriate for formalize neurobiological networks because synapses between neurons are directed, starting at the presynaptic neuron and going to the postsynaptic one. In addition, synapses have strengths or weights, and so, we can also consider weighted graphs where each edge e carries a weight or label w_e that indicates its strength. In fact, we may then also allow that some of the weights are negative. In a neural network, an edge with a negative weight would represent an inhibitory synapse.



The graph Γ from (2.2.1) turned into a directed graph (2.2.2)

Of course, every unweighted graph becomes a weighted one by assigning the weight 1 to every edge. An undirected graph with positive weights becomes a metric space by identifying each edge e with the interval of length $(w_e)^{-1}$. In particular, an unweighted graph then is a metric space where each edge is isometric to the unit interval. The distance between vertices then equals the length of the shortest path joining them. In particular, neighbors in the graph have distance 1.

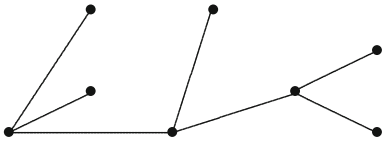
We shall start with undirected and unweighted graphs as the simplest case. In the definition, we require that our graphs Γ be finite, a biologically directly plausible assumption. Moreover, we shall assume, unless stated to the contrary, that they are connected. That means that for every pair of distinct vertices i, j in Γ, there exists a path between them, that is, a sequence $i = i_0, i_1, \ldots, i_m = j$ of distinct vertices such that $i_{\nu-1} \sim i_\nu$ for $\nu = 1, \ldots, m$. Since we can decompose graphs that are not connected into their connected components, the connectivity assumption is no serious restriction.

An obvious way of representing a graph Γ with vertices $i = 1, \ldots, N$ is provided by its adjacency matrix $A = (a_{ij})$. In the unweighted case, we put $a_{ij} = 1$ when there is an edge from i to j and $= 0$ else. We have $a_{ii} = 0$ because we exclude self-loops of vertices, and Γ is undirected iff $a_{ij} = a_{ji}$ for all i, j. In the weighted case, we simply put $a_{ij} = w_{ij}$, the weight of the edge from i to j. Of course, most large graphs arising in applications are sparse, that is, between most pairs i, j, there is no edge. This means that most of the entries of the adjacency matrix are 0. Therefore, that matrix does not provide a very efficient way of encoding the graph. A more efficient way is provided by simply listing for each i those vertices that send links to i, together with the corresponding weights in the weighted case.

An isomorphism between graphs $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ is a bijection $\Phi : V_1 \to V_2$ that preserves neighborhood relations, that is, $i \sim j$ iff $\Phi(i) \sim \Phi(j)$. In other words, i and j are connected by an edge precisely if their images under Φ are. Isomorphisms preserve the degrees of vertices, that is, $n_i = n_{\Phi(i)}$ for every vertex i. An automorphism of Γ is an isomorphism from Γ onto itself. The identity map of the vertex set of Γ is obviously an automorphism, but there may or may not be others, depending on the structure of Γ. The automorphisms of Γ form a group under composition. We can then quantify the symmetry of Γ as the order of its automorphism group.

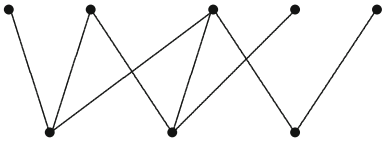
The number of graphs of order k grows very fast as a function of k, and therefore, it becomes unwieldy already for rather small k to list all graphs of order k. Therefore, it is of interest to develop constructions for particular classes or types of graphs. There exist deterministic and stochastic construction schemes. We shall discuss stochastic constructions below in 3.5 in the chapter on stochastic processes. Deterministic constructions typically produce rather regular graphs, that is ones with high degrees of symmetries whereas the stochastic constructions can produce typical representatives of larger classes of graphs. A paradigm of a symmetric graph is a complete graph, meaning that any two different vertices are connected by an edge. For a complete graph, every bijection of its vertices yields an automorphism, and therefore, it is maximally symmetric.

A cycle in Γ is a closed path $i_0, i_1, \ldots, i_m = i_0$ for which all the vertices $i_1, \ldots, i_m$ are distinct. For $m = 3$, we speak of a triangle. A cycle that contains all the vertices of Γ is called a Hamiltonian cycle (and such a cycle need not exist for a given graph). A graph without cycles is called a tree. A maximal tree contained in a graph Γ is called a spanning tree. A spanning tree is obtained by eliminating all cycles from a graph, that is, by cutting an edge in each cycle.



A spanning tree for the graph of (2.2.1) (2.2.3)

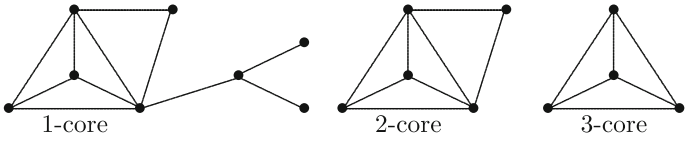
A graph is called k-regular if all vertices have the same degree k. As already mentioned, a graph is bipartite if its vertex set can be decomposed into two disjoint components V_1, V_2 such that whenever $i \sim j$, then i and j are in different components.



A bipartite graph (2.2.4)

It is not hard to see that a graph is bipartite iff it does not contain cycles of odd length. In particular, it cannot contain any triangles.

Another useful concept for analyzing graphs is the k-core. For $k \in \mathbb{N}$, the k-core of a graph Γ is the not necessarily connected maximal subgraph H of Γ with the property that every vertex of H has at least k neighbors in H, that is, its degree in H is at least k. When we exclude the trivial case of an isolated vertex, then Γ itself coincides with its 1-core. When Γ is a tree, already its 2-core is empty. Every cycle of Γ is contained in its 2-core. The core decomposition of Γ, that is, the successive determination of its k-cores for increasing k, is a computationally simple way of decomposing the graph.



of the graph of (2.2.1) (2.2.5)

There exist other parameters that describe certain—more or less—important qualitative properties of graphs. One set of such parameters arises from the metric on the graph generated by the above assignment of length 1 to every edge. The diameter of the graph is the maximal distance between any two of its nodes. As an example how such a parameter can distinguish between typical and non-typical, special graphs, we record that there exists a constant c with the property that the fraction of all graphs with N nodes having diameter exceeding $c \log N$ tends to 0 for $N \to \infty$. Informally expressed, most graphs of N nodes have a diameter of order $\log N$. Thus, graphs with large diameters, like a chain $i_1 \sim i_2 \sim \cdots \sim i_N$ with no other edges, are rare. In the other direction, that is, considering graphs with very small diameters, of course, a fully connected graph has diameter 1. However, one can realize a small diameter already with much fewer edges; namely, one selects one central node to which every other node is connected. In that manner, one obtains a graph of N nodes with $N - 1$ edges and diameter 2. (This graph is called the $(N - 1)$-star, and it will be discussed further below.) Of course, the central node then has a very large degree, namely $N - 1$. It is a big hub. Similarly, one can construct graphs with a few hubs, so that none of them has to be quite that big, efficiently distributed so that the diameter is still rather small. Such graphs can be realized as so-called scale free graphs to be discussed below. Another useful quantity is the average distance between nodes in the graph. The property of having a small diameter or average distance has been called the small-world effect.

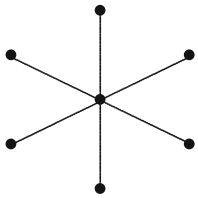
A rather different quantity is the clustering coefficient that measures how many connections there exist between the neighbors of nodes. For this purpose, a triple is a set of three connected vertices, that is, a path with three vertices. As mentioned, a cycle of length 3 is called a triangle

open triple and closed triple = triangle

Note that the triangle contains three triples, *ABC*, *BCA* and *CAB*.
The (global) clustering coefficient then is defined as

$$C := \frac{3 \times \text{number of triangles}}{\text{number of connected triples of nodes}}. \tag{2.2.6}$$

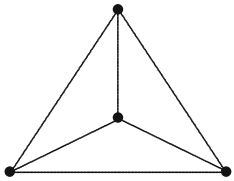
The normalization is that C becomes one for a fully connected graph. It vanishes for trees and other bipartite graphs.
As already mentioned, a k-star is a graph consisting of one central vertex connected to k peripheral vertices, with no connections between those other vertices.



The 6-star (2.2.7)

In particular, each k-star is a tree. A k-star has $\binom{k}{2}$ connected triples of nodes, obtained by connecting the central node with any two peripheral ones. Thus, when we want to compute the clustering index of a graph, we count $\sum_{i\in V}\binom{n_i}{2}$ connected triples of vertices. Thus, the graph Γ of (2.2.1) has 5 triangles and 26 connected triples of nodes, and hence its clustering coefficient is $\frac{15}{26}$.

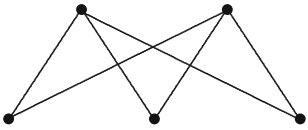
A triangle is a cycle of length 3. One may then also count the number of cycles of length k, for integers >3. A different generalization consists in considering complete subgraphs of order k. Here, the complete k-graph K_k is the graph with k vertices and links between all $i \neq j$. A k-clique in a graph Γ is a subgraph that is a complete k-graph. For example, for $k = 4$, we would have a subset of 4 nodes that are all mutually connected.



The complete graph K_4, the only 4-clique in the graph of (2.2.1) (2.2.8)

One may then associate a simplicial complex to our graph by assigning a k-simplex to every such complete subgraph, with obvious incidence relations. For example, two such k-simplices share a $(k-1)$-dimensional face and are called adjacent when the two corresponding complete k-subgraphs have a complete $(k-1)$-graph in common. This is the basis of topological combinatorics, enabling one to apply tools from simplicial topology to graph theory. See for instance [65].

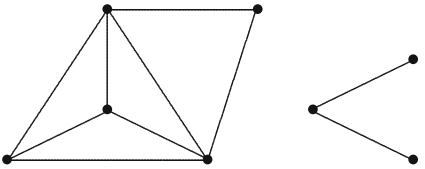
Besides the complete graphs K_k, one also frequently encounters the complete bipartite graphs $K_{m,n}$ consisting of two classes of m and n, resp., vertices such that every vertex in the first class is connected with every vertex in the second class.



The complete bipartite graph $K_{2,3}$ (2.2.9)

The graph $K_{1,n}$ is of course simply the n-star.

A basic question in the analysis of graphs is the cluster decomposition. That means that we try to find subgraphs, the clusters, that are densely connected inside, but only sparsely connected to the rest of the graph. For example, one can try to disconnect the graph by cutting as few edges as possible, to obtain two large (super) clusters,



A decomposition of the graph Γ from (2.2.1) (2.2.10)

and then perhaps iterate the process inside these superclusters to find a finer decomposition. Conversely, one can try to build up the clusters from inside, for example by identifying maximal sets of adjacent k-cliques, or, equivalently, in the simplicial complex defined above, finding maximal sets of k-simplices that are connected by $(k-1)$-dimensional faces. Here, the clusters found are typically not disjoint, in contrast to those obtained by the edge-cutting methods. Of course, one may then analyze the overlap between those clusters.

Concerning the number of edges needed to disconnect a graph, some insight is provided by the following result of Menger:

Lemma 2.2.1. *Let V_1 and V_2 be disjoint subsets of the vertex set of a graph $\Gamma = (V, E)$. The minimal number of edges that need to be deleted from Γ in order to disconnect it in such a manner that V_1 and V_2 are in different components is equal to the maximum number of edge-disjoint paths (that is no two paths are allowed to have an edge in common, even though they may well pass through the same vertex) with one endpoint in V_1 and the other in V_2.*

Another general question is to identify the most important "core" of the graph. The k-core defined above is one useful concept for that. The idea there is that a node is important when it is connected with other important nodes. Thus, one finds the core by successively deleting the less important nodes. That procedure might make some nodes that have originally been highly connected, that is, have a large degree, less relevant, because they had only been connected to other nodes of low degrees. Therefore, in particular, the degree of a node in general is not a good measure of its importance. One can also quantify the importance of a vertex or an edge by counting how many shortest connections between pairs of nodes pass through them. Again, one should be a bit cautious here because in some cases, there exist alternatives to shortest paths that are not substantially longer but that avoid the vertex or edge in question. In other words, sometimes vertices or edges can easily be replaced as parts of short connections while in other cases that may not possible. When one decides

the importance according to such considerations, this effect should also be taken into account.

2.2.3 The Graph Laplacian and its Spectrum

As before, Γ is a finite and connected graph. Probably the most powerful and comprehensive set of invariants comes from the spectrum of the graph Laplacian of Γ to which we now turn. (In general terms, this means that, in order to analyze a graph Γ, we shall study functions defined on Γ. These functions will then be decomposed in terms of a particular set of basis functions, as in Fourier analysis. From those basis functions, we shall obtain spectral values that incorporate the characteristic properties of Γ.)

There are several non-equivalent definitions of the graph Laplacian employed in the literature. In order to clarify this issue, we assign weights $b_i\,(>0)$ to the vertices[1] and introduce an L^2-product for (complex-valued) functions on Γ:

$$(u, v) := \sum_{i\in V} b_i u(i)\overline{v(i)}. \tag{2.2.11}$$

(Since we shall only consider real operators below, it suffices to consider real valued functions, and then the complex conjugate in (2.2.11) is not relevant.)

The most natural choices are $b_i = 1$ or $b_i = n_i$ where n_i is the degree of the vertex i.[2] We may then choose an orthonormal base of that space $L^2(\Gamma)$. In order to find such a basis that is also well adapted to dynamical aspects, we study the graph Laplacian

$$\Delta : L^2(\Gamma) \to L^2(\Gamma)$$

$$\Delta v(i) := \frac{1}{b_i}\Big(\sum_{j,j\sim i} v(j) - n_i v(i)\Big) \tag{2.2.12}$$

where $j \sim i$ means that j is a neighbor of i.[3]

---

[1] These vertex weights should not be confused with the edge weights discussed above; in other words, here, we are *not* considering weighted graphs in the sense defined above.

[2] For purposes of normalization, one might wish to put an additional factor N in front of the product where N is the number of vertices of the graph or, equivalently, divide all the vertex weights by N, but we have decided to omit that factor in our conventions.

[3] There are several different definitions of the graph Laplacian in the literature. Some of them are equivalent to ours inasmuch as they yield the same spectrum, but others are not. The reason is simply that the weights in the underlying product are chosen differently. The operator $Lv(i) := n_i v(i) - \sum_{j,j\sim i} v(j)$ that is often employed in the literature corresponds to the weights $b_i = 1$ (up to the minus sign, of course). The operator $\mathcal{L}v(i) := v(i) - \sum_{j,j\sim i} \frac{1}{\sqrt{n_i}\sqrt{n_j}} v(j)$ employed in the

We, in contrast to much of the literature on graph theory (see e.g. [50]), but in accordance with [28], prefer the weights $b_i = n_i$ over $b_i = 1$ because the former are well adapted to random walks and conservation laws. (When we have a particle randomly moving on a graph with step size 1 then when it is at vertex i it can choose each of the neighbors of i with probability $1/n_i$ for its next move, and this leads to the corresponding factor in the Laplace operator underlying that random walk. This process will be investigated in detail in Sect. 4.2.1.)

The idea behind the definition of Δ is of course that one compares the value of a function v at a vertex i with the average of the values at the neighbors of i. When that average is larger than the value at i, we have $(\Delta v)(i) > 0$.
The important properties of Δ are the following ones:

1. Δ is selfadjoint w.r.t. $(.,.)$:

$$(u, \Delta v) = (\Delta u, v) \tag{2.2.13}$$

   for all $u, v \in L^2(\Gamma)$.[4] This holds because the neighborhood relation is symmetric.
2. Δ is nonpositive:

$$(\Delta u, u) \leq 0 \tag{2.2.14}$$

   for all u. This follows from the Cauchy-Schwarz inequality.
3. $\Delta u = 0$ precisely when u is constant. This one sees by observing that, when $\Delta u = 0$, there can neither be a vertex i with $u(i) \geq u(j)$ for all $j \sim i$ with strict inequality for at least one such j, that is, a nontrivial local maximum, nor a nontrivial local minimum, as this would contradict the fact that $\Delta u(i) = 0$ means that the value $u(i)$ is the average of the values at the neighbors of i. Since Γ is connected, u then has to be a constant (when Γ is not connected, a solution of $\Delta u = 0$ is constant on every connected component of Γ.)

The preceding properties have consequences for the eigenvalues of Δ:

- By 1, the eigenvalues are real.
- By 2, they are nonpositive. We write them as $-\lambda_k$ so that the eigenvalue equation becomes[5]

$$\Delta u_k + \lambda_k u_k = 0. \tag{2.2.15}$$

(Footnote 3 continued)
monograph [28], apart from the minus sign, has the same eigenvalues as Δ for the weights $b_i = n_i$: if $\Delta v(i) = \mu v(i)$, then $w(i) = \sqrt{n_i} v(i)$ satisfies $\mathcal{L} w(i) = -\mu w(i)$.

[4] An operator $A = (A_{ij})$ is symmetric w.r.t. a product $\langle v, w \rangle := \sum_i b_i v(i) \overline{w(i)}$, that is, $\langle Av, w \rangle = \langle v, Aw \rangle$ if $b_i A_{ij} = b_j \bar{A}_{ji}$ for all indices i, j. The b_i are often called multipliers in the literature.

[5] Subsequently, we shall thus call the λ_k instead of the $-\lambda_k$ the eigenvalues, in order to avoid negative quantities. The sign problem comes from the—traditional—definition of the Laplacian (2.2.12) as a nonpositive operator.

- By 3, the smallest eigenvalue is $\lambda_0 = 0$. Since we assume that Γ is connected, this eigenvalue is simple, that is

$$\lambda_k > 0 \tag{2.2.16}$$

for $k > 0$ where we order the eigenvalues as

$$\lambda_0 = 0 < \lambda_1 \leq ... \leq \lambda_K$$

where we put $K := N - 1$.

We next consider, for neighbors i, j,

$$Du(i, j) := u(i) - u(j). \tag{2.2.17}$$

D can be considered as a map from functions on the vertices of Γ to functions on the edges of Γ. In order to make the latter space also an L^2-space, we introduce the product

$$(Du, Dv) := \sum_{e=(i,j)} (u(i) - u(j))(v(i) - v(j)). \tag{2.2.18}$$

Note that we are summing here over edges, and not over vertices. If we did the latter, we would need to put in a factor $1/2$ because each edge would then be counted twice. We also point out that in contrast to the product of (2.2.11), $(u, v) = \sum_i b_i u(i)v(i)$, we do not include weights here. The reason is that here the sum should be considered as a sum of edges and not one over vertices, and since we are considering unweighted graphs at this point, the edges do not carry any natural weights.
The product (2.2.18) encodes more information about the graph than the product (2.2.11). The latter only depends on the weights, but not on the connection structure of the graph. There exist many structurally quite diverse graphs with the same weight sequence, and given a graph, one can rewire it by a cross exchange of edges without changing the degrees of the nodes. Namely, given vertices $i_1 \sim j_1$ and $i_2 \sim j_2$, but without edges between i_1 and i_2, nor between j_1 and j_2, we create a new graph by deleting the edges between i_1 and j_1 and between i_2 and j_2 and inserting new edges between i_1 and i_2 and between j_1 and j_2. That operation preserves the degrees of all vertices, and therefore also the product (2.2.11) for any functions u, v on the graph. (2.2.18), in contrast, is affected because the edge set is changed.

We have

$$\begin{aligned}(Du, Dv) &= \frac{1}{2}\sum_i (n_i u(i)v(i) + \sum_j n_j u(j)v(j) - 2\sum_{j\sim i} u(i)v(j)) \\ &= -\sum_i u(i)\sum_{j\sim i}(v(j) - v(i)) \\ &= -(u, \Delta v). \end{aligned} \tag{2.2.19}$$

Thus, our product (2.2.18) is naturally related to the Laplacian Δ.
We may find an orthonormal basis of $L^2(\Gamma)$ consisting of eigenfunctions of Δ,

$$u_k,\ k = 0, ..., K$$

($K = N - 1$). This is achieved as follows. We iteratively define, with $H_0 := H := L^2(\Gamma)$ being the Hilbert space of all real-valued functions on Γ with the scalar product (., .),

$$H_k := \{v \in H : (v, u_i) = 0 \text{ for } i \leq k - 1\}, \tag{2.2.20}$$

starting with a constant function u_0 as the eigenfunction for the eigenvalue $\lambda_0 = 0$. Also

$$\lambda_k := \inf_{u\in H_k - \{0\}} \frac{(Du, Du)}{(u, u)}, \tag{2.2.21}$$

that is, we claim that the eigenvalues can be obtained as those infima. First of all, since $H_k \subset H_{k-1}$, we have

$$\lambda_k \geq \lambda_{k-1}. \tag{2.2.22}$$

Secondly, since the expression in (2.2.21) remains unchanged when a function u is multiplied by a nonzero constant, it suffices to consider those functions that satisfy the normalization

$$(u, u) = 1 \tag{2.2.23}$$

whenever convenient.
We may find a function u_k that realizes the infimum in (2.2.21), that is

$$\lambda_k = \frac{(Du_k, Du_k)}{(u_k, u_k)}. \tag{2.2.24}$$

Since then for every $\varphi \in H_k, t \in \mathbb{R}$

$$\frac{(D(u_k + t\varphi), D(u_k + t\varphi))}{(u_k + t\varphi, u_k + t\varphi)} \geq \lambda_k, \tag{2.2.25}$$

the derivative of that expression w.r.t. t vanishes at $t = 0$, and we obtain, using (2.2.19)

$$0 = (Du_k, D\varphi) - \lambda_k(u_k, \varphi) = -(\Delta u_k, \varphi) - \lambda_k(u_k, \varphi) \tag{2.2.26}$$

for all $\varphi \in H_k$; in fact, this even holds for all $\varphi \in H$, and not only for those in the subspace H_k, since for $i \leq k - 1$

$$(u_k, u_i) = 0 \tag{2.2.27}$$

and

$$(Du_k, Du_i) = (Du_i, Du_k) = -(\Delta u_i, u_k) = \lambda_i(u_i, u_k) = 0 \tag{2.2.28}$$

since $u_k \in H_k$. Thus, if we also recall (2.2.19),

$$(\Delta u_k, \varphi) + \lambda_k(u_k, \varphi) = 0 \tag{2.2.29}$$

for all $\varphi \in H$ whence

$$\Delta u_k + \lambda_k u_k = 0. \tag{2.2.30}$$

Since, as noted in (2.2.23), we may require

$$(u_k, u_k) = 1 \tag{2.2.31}$$

for $k = 0, 1, ..., K$ and since the u_k are mutually orthogonal by construction, we have constructed an orthonormal basis of H consisting of eigenfunctions of Δ. Thus we may expand any function f on Γ as

$$f(i) = \sum_k (f, u_k) u_k(i). \tag{2.2.32}$$

We then also have

$$(f, f) = \sum_k (f, u_k)^2 \tag{2.2.33}$$

since the u_k satisfy

$$(u_j, u_k) = \delta_{jk}, \tag{2.2.34}$$

the condition for being an orthonormal basis. Finally, using (2.2.33) and (2.2.19), we obtain

$$(Df, Df) = \sum_k \lambda_k (f, u_k)^2. \tag{2.2.35}$$

We next state **Courant's minimax principle**:
Let P^k be the collection of all k-dimensional linear subspaces of H. We have

$$\lambda_k = \max_{L \in P^k} \min \{ \frac{(Du, Du)}{(u, u)} : u \neq 0, \ (u, v) = 0 \text{ for all } v \in L \} \tag{2.2.36}$$

and dually

$$\lambda_k = \min_{L \in P^{k+1}} \max \{ \frac{(Du, Du)}{(u, u)} : u \in L \backslash \{0\} \}. \tag{2.2.37}$$

In words: In (2.2.36), we consider the minimal Rayleigh quotient under k constraints, and we maximize that w.r.t. the constraints. In (2.2.37), we consider the maximal Rayleigh quotient for $k + 1$ degrees of freedom, and we minimize that w.r.t. those degrees of freedom.
To verify these relations, we recall (2.2.21)

$$\lambda_k = \min \{ \frac{(Du, Du)}{(u, u)} : \ u \neq 0, (u, u_j) = 0 \text{ for } j = 0, ..., k-1 \}. \tag{2.2.38}$$

Dually, we have

$$\lambda_k = \max \{ \frac{(Du, Du)}{(u, u)} : \ u \neq 0 \text{ linear combination of } u_j \text{ with } j \leq k \}. \tag{2.2.39}$$

The latter maximum is realized when u is a multiple of the kth eigenfunction, and so is the minimum in (2.2.38). If now L is any $k+1$-dimensional subspace, we may find some v in L that satisfies the k conditions

$$(v, u_j) = 0 \text{ for } j = 0, ..., k-1. \tag{2.2.40}$$

From (2.2.33) and (2.2.35), we then obtain

$$\frac{(Dv, Dv)}{(v, v)} = \frac{\sum_{j \geq k} \lambda_j (v, u_j)^2}{\sum_{j \geq k} (v, u_j)^2} \geq \lambda_k. \tag{2.2.41}$$

This implies

$$\max_{v \in L \backslash \{0\}} \frac{(Dv, Dv)}{(v, v)} \geq \lambda_k. \tag{2.2.42}$$

We then obtain (2.2.37). Equation (2.2.36) follows in a dual manner. In particular, for any eigenfunction u for some eigenvalue $\lambda \neq 0$, we then have

$$\lambda = \frac{(Du, Du)}{(u, u)} \tag{2.2.43}$$

For a fully connected graph,[6] when all the weights b_i are equal, also all the nontrivial eigenvalues are equal. For our preferred choice of weights, $b_i = n_i (= N - 1$ for a fully connected graph of N vertices), we have

$$\lambda_1 = ... = \lambda_K = \frac{N}{N-1} \tag{2.2.44}$$

since

$$\Delta v = -\frac{N}{N-1} v \tag{2.2.45}$$

for any v that is orthogonal to the constants, that is

$$\frac{1}{N} \sum_{i \in V} n_i v(i) = 0. \tag{2.2.46}$$

In more detail, for a fully connected graph of N vertices, for v satisfying (2.2.46),

$$\begin{aligned} \Delta v(i) &= \frac{1}{n_i} \sum_{j, j \sim i} v(j) - v(i) \\ &= \frac{1}{N-1} \sum_{j \neq i} v(j) - v(i) \\ &= (-\frac{1}{N-1} - 1) v(i) \text{ since by (2.2.46) } v(i) = -\sum_{j \neq i} v(j) \\ &= -\frac{N}{N-1} v(i). \end{aligned}$$

We also recall that since Γ is connected, the trivial eigenvalue $\lambda_0 = 0$ is simple. If Γ had two components, then the next eigenvalue λ_1 would also become 0. A corresponding eigenfunction would be equal to a constant on each component, the two values chosen such (2.2.46) is satisfied; in particular, one of the two would be positive, the other one negative. We therefore expect that for graphs with a pronounced community structure, that is, for ones that can be broken up into two large components by deleting only few edges as discussed above, the eigenvalue λ_1 should be close to 0. Formally, this is easily seen from the variational characterization

[6] A fully connected graph is a complete graph K_N, possibly with vertex weights.

$$\lambda_1 = \min\{\frac{\sum_{e=(i,j)\in E}(v(i)-v(j))^2}{\sum_i b_i v(i)^2} : \sum_i b_i v(i) = 0\} \tag{2.2.47}$$

(see (2.2.21) and observe that $\sum_i b_i v(i) = 0$ is equivalent to $(v, u_0) = 0$ as the eigenfunction u_0 is constant). Namely, if two large components of Γ are only connected by few edges, then one can make v constant on either side, with opposite signs so as to respect the normalization (2.2.46) with only a small contribution from the numerator.

More generally, when Γ consists of several clusters with only very few connections between them, one should find several eigenvalues close to 0.

The strategy for obtaining an eigenfunction for the first eigenvalue λ_1 is, according to (2.2.47), to do the same as one's neighbors. Because of the constraint $\sum_i b_i v(i) = 0$, this is not globally possible, however. The first eigenfunction thus exhibits oscillations with the lowest possible frequency . Thus, if we take such a first eigenfunction u_1 and consider the connected components that remain after deleting all edges at whose endpoints u_1 has different signs, then there are precisely two such components, one on which u_1 is positive and one on which it is negative. More generally, the number of connected components of Γ where an eigenfunction for the kth eigenvalue has a fixed sign is at most $k+1$ when the eigenvalues are ordered in increasing order and appropriately when they are not simple, according to a version of Courant's nodal domain theorem proved by Gladwell-Davies-Leydold-Stadler [48].

We once more consider the case $b_i = n_i$. As noted, for a complete graph, we have $\lambda_1 = \frac{N}{N-1}$, see (2.2.44). For any other graph, that is, for any graph that is not complete, we have

$$\lambda_1 \leq 1. \tag{2.2.48}$$

This follows from (2.2.47), by taking two vertices i_1, i_2 that are not connected by an edge and by assigning values of u to those points satisfying $n_{i_1}u(i_1) + n_{i_2}u(i_2) = 0$ and 0 to all other vertices. The quotient in (2.2.47) then becomes 1, and therefore, the infimum characterizing λ_1 has to be ≤ 1.

By way of contrast, according to (2.2.37), the highest eigenvalue is given by

$$\lambda_K = \max_{u\neq 0} \frac{(Du, Du)}{(u, u)}. \tag{2.2.49}$$

Thus, the strategy for obtaining an eigenfunction for the highest eigenvalue is to do the opposite what one's neighbors are doing, for example to assume the value 1 when the neighbors have the value -1. Thus, the corresponding eigenfunction will exhibit oscillations with the highest possible frequency. Here, the obstacle can be local. Namely, any triangle, that is, a triple of three mutually connected nodes, presents such an obstacle. More generally, any cycle of odd length makes an alternation of the values 1 and -1 impossible. The optimal situation here is represented by a bipartite graph, that is, a graph that consists of two sets Γ_+, Γ_- of nodes without any links

between nodes in the same such subset. Thus, one can put $u_K = \pm 1$ on $\Gamma_\pm$. For our choice $b_i = n_i$, which we shall now adopt for the subsequent discussion, one then finds

$$\lambda_K = 2 \tag{2.2.50}$$

for a bipartite graph.
In contrast, the highest eigenvalue λ_K becomes smallest on a fully connected graph, namely

$$\lambda_K = \frac{N}{N-1} \tag{2.2.51}$$

according to (2.2.46). For graphs that are neither bipartite nor fully connected, this eigenvalue lies strictly between those two extremal possibilities.
Perhaps the following caricature can summarize the preceding: For minimizing λ_1—the minimal value being 0—one needs two subsets that can internally be arbitrarily connected, but that do not admit any connection between each other. For maximizing λ_K—the maximal value being 2—one needs two subsets without any internal connections, but allowing arbitrary connections between them. In either situation, the worst case—that is, a maximal value for λ_1 and a minimal value for λ_K—is represented by a fully connected graph. In fact, in that case, λ_1 and λ_K coincide.
Let us consider bipartite graphs in some more detail. We already noted above that on a bipartite graph, we can determine the highest eigenfunction u_K explicitly, as ± 1, being $+1$ on one set, -1 on the other set of vertices defining the bipartition. In fact, it is clear from that construction that this property is equivalent to the bipartiteness of the graph. Actually, if the graph is bipartite, then even more is true: Whenever λ_k is an eigenvalue, then so is $2 - \lambda_k$. Since 0 is an eigenvalue for any graph, this criterion implies our observation that 2 is an eigenvalue. The general statement is not difficult to see: Let G_1, G_2 be the two vertex sets defining the bipartition. When u_k is an eigenfunction for the eigenvalue λ_k, then

$$\tilde{u}_k(i) := \begin{cases} u_k(i) \text{ for } i \in G_1 \\ -u_k(i) \text{ for } i \in G_2 \end{cases} \tag{2.2.52}$$

is an eigenfunction with eigenvalue $2 - \lambda_k$ as is readily verified.
We now present some results from [15] about controlling the highest eigenvalue. In order to understand the significance of the highest eigenvalue λ_K better, we now derive some general identity first, for a function u on the vertex set of Γ.

$$\sum_i \frac{1}{n_i} \sum_{j,k,j\sim i,k\sim i} (u(j) - u(k))^2$$
$$= \sum_i \sum_{k,k\sim i} \frac{1}{n_i} \sum_{j,j\sim i} (u(j) - u(k))^2$$

$$
\begin{aligned}
&= \sum_i \left(\sum_{k,k\sim i} \left(\frac{1}{n_i} \sum_{j,j\sim i} u(j)^2 - \frac{2}{n_i} \sum_{j,j\sim i} u(j)u(k) + u(k)^2 \right) \right) \\
&= \sum_i \left(2 \sum_{j,j\sim i} u(j)^2 - \frac{2}{n_i} (\sum_{j,j\sim i} u(j))^2 \right) \\
&= 2 \sum_i \sum_{j,j\sim i} u(j)^2 - \sum_i 2n_i \left(\frac{1}{n_i} \sum_{j,j\sim i} u(j) \right)^2 .
\end{aligned}
$$

We now observe that we can replace u by $u - u(i)$ in the first and hence also in all subsequent lines. This yields

$$
\begin{aligned}
&\sum_i \frac{1}{n_i} \sum_{j,k,j\sim i,k\sim i} (u(j) - u(k))^2 \\
&\quad = 2 \sum_i \sum_{j,j\sim i} (u(j) - u(i))^2 - \sum_i 2n_i \left(\frac{1}{n_i} \sum_{j,j\sim i} (u(j) - u(i)) \right)^2 \\
&\quad = 2 \sum_i \sum_{j,j\sim i} (u(j) - u(i))^2 - \sum_i 2n_i (\Delta u(i))^2 .
\end{aligned}
$$

When u now is an eigenfunction, $\Delta u + \lambda u = 0$ for some eigenvalue λ, then, recalling (2.2.43), we obtain

$$
\sum_i \frac{1}{n_i} \sum_{j,k,j\sim i,k\sim i} (u(j) - u(k))^2 = 2\lambda(2 - \lambda) \sum_i n_i u(i)^2 . \tag{2.2.53}
$$

Using (2.2.43) again, we can also reformulate this as

$$
2 - \lambda = \frac{\sum_i \frac{1}{n_i} \sum_{j,k,j\sim i,k\sim i} (u(j) - u(k))^2}{\sum_i \sum_{j,j\sim i} (u(j) - u(i))^2} . \tag{2.2.54}
$$

We now want to employ (2.2.54) to interpret $2 - \lambda_K$ (λ_K being the largest eigenvalue of our graph) as quantifying how much Γ is locally different from being bipartite, recalling that this quantity is 0 precisely if Γ happens to be bipartite.
In order to develop some intuition, we start with a bipartite graph Γ_0 with M vertices. We consider a highest eigenfunction $\bar{u}$ that is $+1$ on one class and -1 on the other class of vertices, as described above. In particular,

$$
\frac{\frac{1}{2} \sum_{j\sim k} (\bar{u}(j) - \bar{u}(k))^2}{\sum_i n_i \bar{u}(i)^2} = 2. \tag{2.2.55}
$$

We add another vertex i_0 and connect it to one of the edges of Γ_0. Of course, this new graph Γ_1 then is again bipartite, but we extend $\bar{u}$ by $\bar{u}(i_0) = 0$ to Γ_1. Thus, the numerator and the denominator of (2.2.55) are both increased by 1. Given any small $\epsilon > 0$, by assuming that Γ_0 is sufficiently large, that is, $\sum_i n_i$ is sufficiently large, we can therefore achieve that, for Γ_1,

$$\frac{\frac{1}{2}\sum_{j\sim k}(\bar{u}(j) - \bar{u}(k))^2}{\sum_i n_i \bar{u}(i)^2} > 2 - \epsilon. \tag{2.2.56}$$

Now, this is not affected when we construct a graph Γ by attaching another graph Γ_2 at i_0 and extend $\bar{u}$ by 0 to all of Γ_2. For instance, Γ_2 could be a complete graph of N vertices, for any N. In particular, the difference $2 - \lambda_K$ (where λ_K is the largest eigenvalue of Γ) which has to be larger than $2 - \epsilon$ by (2.2.51), is not very sensitive to the shape of Γ_2. This implies, for instance, that $2 - \lambda_K$ cannot reflect a global quantity like the clustering coefficient C of (2.2.6) that expresses an averaged difference from a graph being bipartite. In fact, our construction of attaching a complete graph K_N to a bipartite graph Γ_0 through a connecting node produces a graph with C arbitrarily close to its maximal value 1 when N is sufficiently large. By extending this example, we can also see that we should have many eigenvalues λ for which $2 - \lambda$ is small when the graph possesses several relatively large bipartite or almost bipartite parts that are only loosely connected with the rest. This is analogous to the fact that a graph possesses several small eigenvalues when it has many relatively large components that are only loosely connected to the rest, that is, when the graph can be easily decomposed into several large clusters. Of course, for a nonconnected graph, that is, one with several components without links between them, the spectrum simply is the union of the spectra of the components. Therefore, by the continuity principle, a graph consisting of clusters that are only loosely connected to each other has its spectrum approximated (in a sense not made completely precise here) by the spectra of these clusters, that is, by that of the graph resulting from deleting the few links between the clusters.

We can use (2.2.54), however, in order to control $2 - \lambda_K$ by the following local clustering measure

$$C_0(\Gamma) := \max\{\alpha : \text{ for each } i \in \Gamma, \text{ at least } \alpha n_i \text{ of its edges are contained in some triangle}\}. \tag{2.2.57}$$

Again, $C_0 = 0$ for a bipartite and $C_0 = 1$ for a complete graph. Thus, let us analyze (2.2.54) with this quantity in mind. We want to control $2 - \lambda_K$ from below in terms of C_0. This means that we need to match any term $(u(i) - u(j))^2$ in the denominator by some term in the numerator of comparable magnitude. Now, given such a term, we have two possibilities. Either we can at least find $\frac{\alpha n_i}{2}$ neighbors k of i for which $(u(k) - u(j))^2 \geq \frac{1}{2}(u(i) - u(j))^2$. Then $(u(i) - u(j))^2$ is matched in the numerator. Or, for at least $\frac{\alpha n_i}{2}$ neighbors k of i for which the edge $e = (i, k)$ is contained in some triangle (i, k, ℓ), we have $(u(i) - u(k))^2 \geq \frac{1}{12}(u(i) - u(j))^2$ for at least αn_i neighbors

k that are contained in some triangle (i, k, ℓ). Therefore, taking one of those k and one such triangle (i, k, ℓ), for every other vertex m of ℓ, either $(u(i) - u(m))^2$ or $(u(k) - u(m))^2$ has to be sufficiently large. Since we have the choice between at least αn_i such vertices k, all the edges $e = (i, j)$ can thus be matched with a controlled amount of duplication. Thus, $2 - \lambda_K$ can be controlled from below in terms of C_0, or conversely, C_0 can be controlled from above in terms of $2 - \lambda_K$. The control in the other direction does not work quite, because $2 - \lambda_K$ can still be made relatively large in a graph like that obtained from the complete graph K_N by attaching another node i_0 with a single connection to one of the vertices of K_N. Here, $C_0 = 0$, because the only edge from i_0 is not contained in a triangle. Perhaps more a more important example is a graph with many cycles of odd length, but all of them of length at least 5. Here, $C_0 = 0$ as there are no triangles, but $2 - \lambda_K \neq 0$ because the graph is not bipartite as bipartite graphs can only have cycles of even length.
In passing, we also observe that by a reasoning similar to that for (2.2.53), we can also show

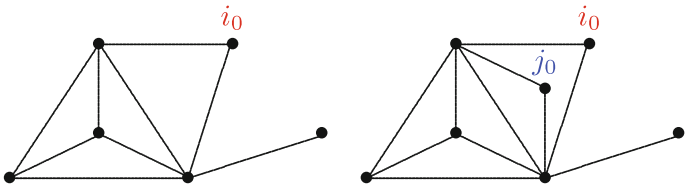
$$\sum_i \sum_{k,k\sim i} \left(\frac{1}{n_i} \sum_{j,j\sim i} (u(j) - u(k)) \right)^2 = \lambda(2 - \lambda) \sum_i n_i u(i)^2. \tag{2.2.58}$$

Again, this can be used to estimate the local difference from being bipartite in terms of $2 - \lambda_K$.
In fact, the preceding constructions can be understood and significantly extended through the concept of a neighborhood graph, see [15].

Having looked at the smallest and largest eigenvalues, we now take a look at the one in the middle, $\lambda = 1$ (we again fix the weights in (2.2.12) to be n_i). In order to see that this eigenvalue is special, we rewrite the eigenvalue equation for $\lambda = 1$ as

$$0 = \Delta v(i) + v(i) = \frac{1}{n_i} \left(\sum_{j,j\sim i} v(j) - n_i v(i) \right) + v(i) = \frac{1}{n_i} \sum_{j,j\sim i} v(j). \tag{2.2.59}$$

Thus, an eigenfunction for the eigenvalue 1 is *balanced* in the sense that for every node, the average of the values of its neighbors vanishes.
There is a simple way to generate the eigenvalue 1: node duplication. That means that we take some graph Γ_0 and some node $i_0 \in \Gamma_0$ and create a new graph Γ by adjoining an additional node j_0 to Γ_0 by the prescription that j_0 gets connected to the same nodes as i_0. That is, whenever a node $i \in \Gamma_0$ is connected to i_0, then we also connect it to j_0. Note that i_0 and j_0 are not directly connected by this rule. The node j_0 can then be considered as the double of i_0 because it shares the same neighbors in Γ.

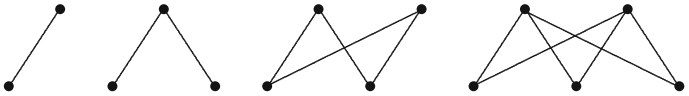


j_0 is the duplicate of i_0 (2.2.60)

We now simply observe that the function

$$u_1(i) := \begin{cases} 1 \text{ for } i = i_0 \\ -1 \text{ for } i = j_0 \\ 0 \text{ else} \end{cases} \tag{2.2.61}$$

satisfies the eigenvalue equation (2.2.59). By repeated node duplication, we can thus generate a graph with an arbitrary high multiplicity of the eigenvalue 1. In particular, a complete bipartite graph $K_{m,n}$, that is, a bipartite graph consisting of one class of size m and another of size n, such that any node in the first class is connected with every node in the second class, can be obtained by successive node duplications from the graph $K_{1,1}$ consisting of two nodes connected by an edge.



$K_{2,3}$ is generated by three vertex duplications from $K_{1,1}$ (2.2.62)

Therefore, we can deduce the spectrum of any such graph $K_{m,n}$: it has the eigenvalues 0 and 2 with multiplicity 1 each, and the eigenvalue 1 with multiplicity $m + n - 2$. Conversely, any graph with this spectrum is a complete bipartite graph $K_{m,n}$. In particular, all the graphs $K_{m,n}$ with the same value of $m + n$ of nodes have the same spectrum, that is, they are *isospectral*. In particular, the spectrum of Δ does not completely determine a graph. We also observe the following fact. For a complete graph K_N, for $N \to \infty$, all the eigenvalues converge to 1, see (2.2.44), except for $\lambda_0 = 0$. Therefore, in this limit, the difference between the spectra of a complete graph K_N and a complete bipartite graph $K_{m,n}$ with $m + n = N$ is only reflected by a single eigenvalue, the highest eigenvalue λ_K which remains 2 for $K_{m,n}$, but goes to 1 for K_N. To appreciate this phenomenon, we observe that K_N is the graph with the maximal number of 3-cycles, i.e., triangles, because every edge is contained in $N-2$ triangles and every vertex is a vertex of $\binom{N-1}{2}$ triangles. $K_{m,n}$ does not contain any triangles, but otherwise is the graph with the maximal number of 4-cycles, in the sense that every vertex of the second class, that with n vertices, is a vertex of $\binom{m}{2}(n-1)$ 4-cycles, and analogously for the first class. From this observation, we

also see that different such $K_{m,n}$ with the same sum $m+n$ are distinguished by their numbers of cycles. Therefore, this number, together with the spectrum, can uniquely identify a graph $K_{m,n}$.
These issues are further developed in [7, 8], and biological applications are given in [6, 9].

We now return to the issue of decomposing a graph by cutting edges. There exists an important relationship of this issue with the first eigenvalue λ_1 which we shall now describe. This is based on a quantity that is analogous to one introduced by Cheeger in Riemannian geometry, but had already been considered earlier in graph theory by Polya. We therefore call it the Polya-Cheeger constant. Letting $|E|$ denote the number of edges contained in an edge set E, the Polya-Cheeger constant is

$$h(\Gamma) := \inf_{E_0} \{ \frac{|E_0|}{\min(\sum_{i\in V_1} b_i, \sum_{i\in V_2} b_i)} \} \tag{2.2.63}$$

where removing E_0 disconnects Γ into the components V_1, V_2. Thus, we try to break the graph up into two large components by removing only few edges. We may then repeat the process within those components to break them further up until we are no longer able to realize a small value of h.
We now derive elementary estimates for λ_1 from above and below in terms of the constant $h(\Gamma)$. Our reference here is [28] (that monograph also contains many other spectral estimates for graphs, as well as the original references; the analogy between the Cheeger estimate in Riemannian geometry and in graph theory was discovered in [35]). We start with the estimate from above and use the variational characterization (2.2.47). Let the edge set E_0 divide the graph into the two disjoint sets V_1, V_2 of nodes, and let V_1 be that with the smaller vertex sum $\sum b_i$. We consider a function v that is =1 on all the nodes in V_1 and $= -\alpha$ for some positive α on V_2. α is chosen so that the normalization $\sum_V b_i v(i) = 0$ holds, that is, $\sum_{i\in V_1} b_i - \sum_{i\in V_2} b_i \alpha = 0$. Since V_2 is the subset with the larger $\sum b_i$, we have $\alpha \le 1$. Thus, for our choice of v, the quotient in (2.2.47) becomes $\le \frac{(1+\alpha)^2|E_0|}{\sum_{i\in V_1} b_i + \sum_{i\in V_2} b_i \alpha^2} = \frac{(\alpha+1)|E_0|}{\sum_{V_1} b_i} \le 2\frac{|E_0|}{\sum_{V_1} b_i}$. Since this holds for all such splittings of our graph Γ, we obtain from (2.2.63) and (2.2.47)

$$\lambda_1 \le 2h(\Gamma). \tag{2.2.64}$$

The estimate from below is slightly more subtle, and the estimate presented here works only for the choice

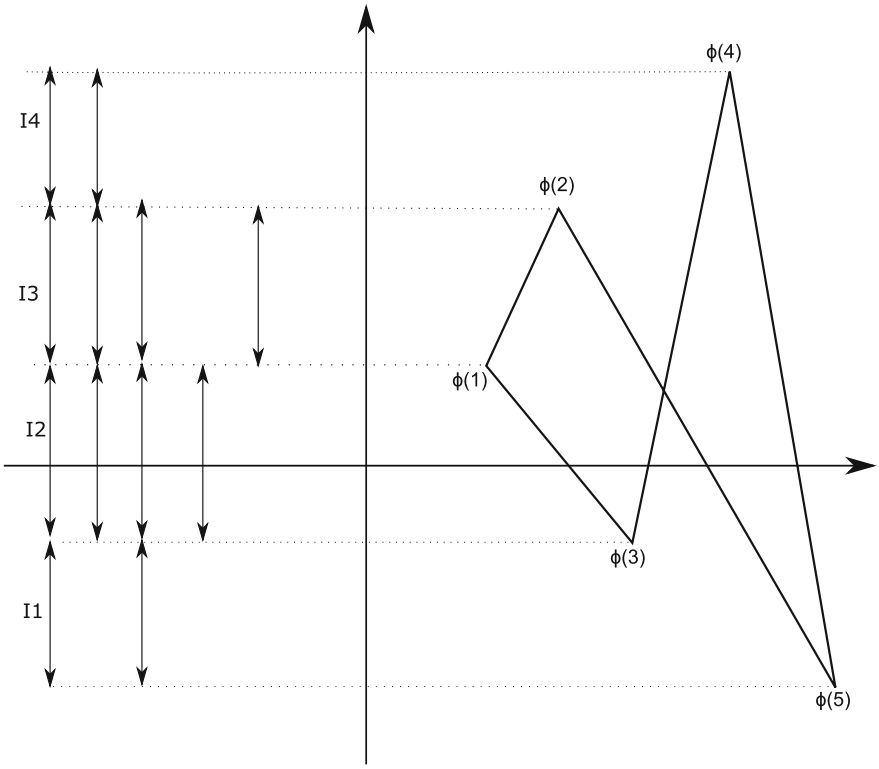
$$b_i = n_i. \tag{2.2.65}$$

We consider the first eigenfunction u_1. Like all functions on our graph, we consider it to be defined on the nodes. We then interpolate it linearly (or monotonically) on the edges of Γ. Since u_1 is orthogonal to the constants (recall $\sum_i n_i u(i) = 0$), it has to change sign, and the zero set of our extension then divides Γ into two parts

Γ' and Γ'' by sign. W.l.o.g., Γ' is the part with fewer nodes. The points where (the extension of) $u_1 = 0$ are called boundary points. We now consider any function φ that is linear on the edges, 0 on the boundary, and positive elsewhere on the nodes and edges of Γ'. We also put $h'(\Gamma') := \inf_{E_1}\{\frac{|E_1|}{\sum_{i\in\Omega} n_i}\}$ where removing the edges in E_1 cuts out a subset Ω of the vertex set of Γ' that is disjoint from the boundary. We also use the identity

$$\sum_{e=(i,j)} |\varphi(i) - \varphi(j)| = \int_\sigma \sharp_e(\varphi = \sigma) d\sigma \tag{2.2.66}$$

where $\sharp_e(\varphi = \sigma)$ denotes the number of edges on which φ attains the value σ. This is illustrated by the following figure for the case of a cyclic graph with 5 vertices



From (2.2.66), we proceed to

$$\begin{aligned}\sum_{e=(i,j)} |\varphi(i) - \varphi(j)| &= \int_\sigma \frac{\sharp_e(\varphi = \sigma)}{\sum_{i:\varphi(i)\geq\sigma} n_i} \sum_{i:\varphi(i)\geq\sigma} n_i \, d\sigma \\ &\geq \inf_\sigma \frac{\sharp_e(\varphi = \sigma)}{\sum_{i:\varphi(i)\geq\sigma} n_i} \int_s \sum_{i:\varphi(i)\geq s} n_i \, ds\end{aligned}$$

$$= \inf_{\sigma} \frac{\sharp_e(\varphi = \sigma)}{\sum_{i:\varphi(i)\geq\sigma} n_i} \sum_i n_i |\varphi(i)|$$
$$\geq h'(\Gamma') \sum_i n_i |\varphi(i)|$$

when the sets $\varphi = \sigma$ and $\varphi \geq \sigma$ satisfy the conditions in the definition of $h'(\Gamma)$; that is, the infimum has to be taken over those $\sigma < \max \varphi$. Applying this to $\varphi = v^2$ for some function v on Γ' that vanishes on the boundary, we obtain

$$\begin{aligned} h(\Gamma') \sum_i n_i |v(i)|^2 &\leq \sum_{e=(i,j)} |v(i)^2 - v(j)^2| \\ &\leq \sum_{e=(i,j)} (|v(i)| + |v(j)|)|v(i) - v(j)| \\ &\leq \sqrt{2} (\sum_i n_i |v(i)|^2)^{1/2} (\sum_{e=(i,j)} |v(i) - v(j)|^2)^{1/2} \end{aligned}$$

from which

$$\frac{1}{2} h(\Gamma')^2 \sum_i n_i |v(i)|^2 \leq \sum_{e=(i,j)} |v(i) - v(j)|^2. \tag{2.2.67}$$

We now apply this to $v = u_1$, the first eigenfunction of our graph Γ. We have $h'(\Gamma') \geq h(\Gamma)$, since Γ' is the component with fewer nodes. We also have[7]

$$\lambda_1 \sum_{i\in\Gamma'} n_i u_1(i)^2 = \frac{1}{2} \sum_{i\in\Gamma'} \sum_{j\sim i} (u_1(i) - u_1(j))^2, \tag{2.2.68}$$

cf. (2.2.24) (this relation holds on both Γ' and Γ'' because u_1 vanishes on their common boundary).[8] Equation (2.2.67) and (2.2.68) yield the desired estimate (under the assumption (2.2.66))

$$\lambda_1 \geq \frac{1}{2} h(\Gamma)^2. \tag{2.2.69}$$

From (2.2.64) and (2.2.69), we also observe the inequality

$$h(\Gamma) \leq 4 \tag{2.2.70}$$

[7] We obtain the factor $1/2$ because we are now summing over vertices so that each edge gets counted twice.

[8] To see this, one adds nodes at the points where the edges have been cut, and extends functions by 0 on those nodes. These extended functions then satisfy the analogue of (2.2.19) on either part, as one sees by looking at the derivation of that relation and using the fact that the functions under consideration vanish at those new "boundary" nodes.

for any connected graph, when the weights b_i are the vertex degrees n_i. In fact, we can obtain a better estimate from (2.2.69). Since, as noted above in (2.2.44), (2.2.48), we always have $\lambda_1 \leq \frac{N}{N-1}$, we see directly that

$$h(\Gamma) \leq \sqrt{\frac{2N}{N-1}}. \tag{2.2.71}$$

Also, unless the graph is complete, we have $\lambda_1 \leq 1$, see (2.2.48), and therefore, for non-complete graphs, we have the estimate

$$h(\Gamma) \leq \sqrt{2}. \tag{2.2.72}$$

One can also think about the decomposition of a graph by removing vertices instead of edges. This issue is amenable to a similar treatment, and one can define a quantity analogous to $h(\Gamma)$ that has the number of vertices whose elimination is needed to disconnect the graph in the numerator; see [28] for details. Moreover, we can also define a dual Cheeger constant that can be utilized to control the largest eigenvalue, see [15].

The spectrum of the graph Laplacian is a useful tool to analyze biological networks, see [6, 8, 9, 10]. Whereas most computational problems on graphs are NP-hard or even NP-complete, and hence require a number of steps that grows exponentially with the number of vertices, the computation of the spectrum proceeds by linear algebra. Therefore, there exist algorithms that grow only like a polynomial of low order in the number of vertices. With current methods, one can determine the spectrum of graphs with about half a million nodes, and if one exploits the particular structures that can typically be found in empirical networks, one can handle even larger ones.

Therefore, one computes spectra of graphs in order to compare or distinguish biological and other networks by their spectral properties. We should remark at this point that the spectrum of its Laplacian does not always determine a graph uniquely. For instance, all complete bipartite graphs $K_{m,n}$ with the same total number $m + n$ of vertices have the same spectrum, i.e., they are isospectral, as we have already observed in the discussion after (2.2.61), with their spectrum consisting of 0 and 2 with multiplicity 1 each, and the eigenvalue 1 with multiplicity $m + n - 2$.

Nevertheless, as we have seen above, the spectrum reflects many important structural properties of a graph, like its decomposability.

Networks from specific domains, for instance protein-protein interaction networks (see e.g. [6]) usually share specific properties that distinguish them from networks from other domains. By investigating these specific spectral properties, one can then gain insight about the structure of such networks. For instance, a high multiplicity of the eigenvalue 1 in molecular networks may indicate gene duplications underlying the evolutionary history of such networks.

Many biological networks are, in fact, directed. In metabolic networks, there is the distinction between inputs and outputs of reactions, and one is interested in flows through such a directed network. In neuronal networks, information is transmitted

from a presynaptic to a postsynaptic neuron, and one wants to understand the resulting dynamics. In food webs, trophic interactions ("who eats whom") are naturally unsymmetric. Therefore, the spectral theory of directed graphs has been systematically developed by F.Bauer, see [14]. Applications to biological networks are explored in [11].

2.3 Descendence Relations

2.3.1 Trees and Phylogenies

Trees are the formal tool for representing ancestor-descendent relations in biology and other fields. At first sight, the concept of a tree as defined below seems not appropriate for that task, however, when one thinks of parent-offspring relationships in sexually recombining species. There, the relationship graph, the so-called pedigree is branching in the backward direction because each individual has two parents, as well as in the forward direction because individuals on average have more than one offspring if the population is not going extinct. When one considers asexual reproduction, however, the situation becomes simpler because each individual then has only one parent, and branching can occur only forward in time when one considers the descendents over the generations of a single ancestor. This, perhaps, is not such an exciting problem, and, in fact, biologists are rather interested in trees for describing phylogenetic relationships between species instead of individuals. The endpoints of a tree, the so-called leaves (see below for the formal definitions), then correspond to a collection of recent species, and one tries to construct a tree in which the internal vertices represent ancestral species that are the common ancestors of all the species below them. Here, one usually assumes that speciation events are binary branchings, that is, one species splits into two daughter species. (In order to make this consistent, at least some biological taxonomists, the cladists, adopt the convention that whenever a new species branches off from an existing one, the remaining part of the latter then is also classified as a new species.) Traditionally, the similarities between species were gauged on the basis of morphological features, and paleontologists tried to identify the hypothetical ancestral species with ones documented in the fossil record. (In practice, this encounters many problems, but that is not our concern here.) Today, there exists a powerful alternative to that classical method, the comparison on the basis of genetic data. The idea is obvious, to take DNA samples from members of different species and count the differences so as to determine the genetic distances between the species. On the basis of those distances, a hierarchical grouping should be possible that can be represented by a tree. Of course, in practice, this is not so simple. First of all, the genetic samples need to be comparable. For that, one needs to identify DNA segments in the species representatives that are homologous to each other, that is, derived from the same ancestral sequence through a process of accumulation of mutations. Since besides point mutations in the DNA, there can

also occur rearrangements like insertions, deletions, inversions, first the problem of sequence alignment needs to be addressed and solved for the samples at hand. This is usually done with the BLAST algorithm [2]. Next, one assumes that mutations occurred at the same rate in the different lineages, the hypothesis of the molecular clock. Otherwise, the number of genetic differences would not be a uniform measure of the time since branching from a common ancestor. Moreover, one needs to find genetic regions that have not been under selective pressure, but rather where there is a uniform probability of the retention of any mutation. Under stabilizing selection, most mutations are eliminated, and this would lead to an underestimate for the time since branching. For directed selection, in contrast, adaptive pressure leads to a more rapid accumulation of mutations and then to an overestimate of the time since branching.

Even if one can align the sequences successfully and eliminate selection effects, there still remain substantial problems. Often, the genetic distances vary with the genomic regions considered. Thus, depending on the DNA region considered, one might get a different tree. In that case, one might try to find some kind of compromise tree. That will depend on the criterion adopted, however, as we shall discuss a little more below. Sometimes, the data even do not fit into a tree because distances on a tree need to satisfy some necessary conditions discussed below. The question then is what substitute to choose for a tree, an issue that we shall also address below. Also, a species is not entirely homogeneous, and there are also genetic differences between the members of the same species (otherwise, evolution could not work by differential selection). Therefore, one needs to gauge intraspecies differences against interspecies ones. Finally, speciation is not an event that takes place at one clearly identifiable point in time, but rather is a gradual process of the accumulation of differences between different populations until reproductive barriers emerge that prevent further genetic mixing between those populations. Here, we need to invoke the species concept of modern biology. A species is defined as a population of organisms that can sexually produce viable and fertile offspring among them. In practice, however, sometimes that relationship is not necessarily transitive. That is, there can exist subpopulations $A_1, \ldots, A_k$ such that individuals from A_i can reproduce with those of A_{i+1} for all i, but those from A_1 are no longer able to reproduce with those from A_k. An example are the races of domestic dogs that range from rather large to very small ones. More generally, for the assembly of phylogenetic trees, species are considered as static ensembles, while in reality speciation is a temporally extended dynamic process inside groups of indidivuals (see the discussion in [20]). (As an aside, some of those population dynamics can be reconstructed on the basis of a statistical analysis of the distribution of alleles in recent populations, in particular from their deviations from equilibria defined by independence hypotheses.)

In spite of all these problems, phylogenetic tree reconstructions are a useful tool for many biologists. There is one issue, however, that calls for a generalization of the representation of phylogenies by trees. As L. Margulis emphasized, many genetic changes are not caused by mutations in inherited genomes, but rather by horizontal gene transfer through virusses and other processes [88]. That, of course, cannot be represented in a tree. Therefore, the tree formalism has recently been extended in

[112] to allow for horizontal gene transfer. On the other hand, over the course of evolution, organisms seem to have developed some protective mechanisms against such horizontal gene insertions, and the relative efficiency of those provides some justification for attempting to represent genetic data in a tree. In the light of all the difficulties mentioned above, it is then necessary to develop methods for finding trees that contain as few hypotheses as possible not supported by the available data.

We now start with the mathematical formalism as pioneered by Andreas Dress; we treat a particular class of graphs, the so-called trees. Our basic references are [107] and [37]. We shall not provide the proofs of the mathematical results discussed, but rather refer the reader to the literature.
We recall that a **tree** $T = (V, E)$ is a graph without cycles.

Lemma 2.3.1. *For a graph $\Gamma = (V, E)$, the following statements are equivalent:*

1. *Γ is a tree, that is, has no cycles.*
2. *For any two distinct vertices i, j, there exists a unique path of distinct vertices joining them (we shall call that path a "shortest path" even though we do not yet have specified a metric at this point—it will, however, turn out to be a shortest path for any metric on the tree).*
3. *$|V| = |E| + 1$.*
4. *The deletion of any edge disconnects Γ.*

The *proof* of this lemma is an easy exercise. Since for any graph $\Gamma = (V, E)$, we have $|V| \leq |E| + 1$, a tree thus is a graph with the minimal number of edges needed to connect a vertex set V.
The vertices of a tree that have degree 1 are called leaves. The other vertices are called interior vertices. Sometimes, it is convenient to exclude vertices of degree 2.
A rooted tree is a tree with one distinguished vertex i_0, the root.
Rooted trees are the formal tool to represent hierarchical relationships between individual entities. We say that the vertex i_1 is above the vertex i_2, or in the phylogenetic interpretation to follow that i_1 is an ancestor of i_2, and i_2 a descendent of i_1, when the shortest path from i_0 to i_2 passes through i_1.
In phylogenies, the aim is the comparison between extant species. Those species then are represented as the leaves of some tree, and the rest of the tree then is built with the purpose that the interior vertices represent common ancestors of all those below some. Thus, the interior vertices may correspond to hypothetical species on which no data need to be available. Of course, paleontologists try to identify those interior vertices with fossil species, but the modern data usually consist of genetic data like pieces of DNA sequences for which one rarely has fossil samples. Thus, in paleontology, it is natural to allow for degree 2 vertices, representing ancestors of a single extant species that are documented in the fossil record. In molecular sequence analysis, however, one would exclude degree 2 vertices because all interior vertices represent hypothetical reconstructions of common ancestors of several descendent species.

In order to proceed with this formalization, we consider X-trees where X is some set. In applications, X of course is a or the data set. An X-tree is a tree $T = (V, E)$ together with a map $\phi : X \to V$ whose image contains all vertices of degrees 1 and 2. (In the rooted case, we do not require that the root be in the image of ϕ even though it may have degree ≤ 2.) The map need not be injective. For a phylogenetic (X-) tree, however, we require that ϕ be a bijection onto the leaves of T. In particular, such a phylogenetic tree has no vertices of degree 2. When every interior vertex has degree 3, we speak of a binary phylogenetic tree. This is a natural assumption in biology because, in evolution, a species can split into two daughter species, and each of those can then split again, and so on, but one does not see the emergence of three or more daughter species at the same time. In fact, much of phylogenetic tree reconstruction is about resolving the question in which temporal order the various splits into daughter species took place.

An X-split $A|B$ is a partition of X into two non-empty subsets A, B.[9] Thus, in biological applications, A might represent those members of X where a certain feature is present, and B those where that feature is absent.—Two such splits $A_1|B_1$ and $A_2|B_2$ are called compatible when at least one of the intersections $A_1 \cap A_2$, $A_1 \cap B_2$, $B_1 \cap A_2$, $B_1 \cap B_2$ is empty. If, say, $A_1 \cap B_2 = \emptyset$ then $A_1 \subset A_2$ and $B_2 \subset B_1$, and vice versa, and so, there is an alternative way of expressing compatibility of splits.

When we have an X-tree (T, ϕ), then every edge e of T induces an X-split because it decomposes T into two subgraphs T_1, T_2 (which might include the degenerate case where one of them consists of a single vertex and no edges), and their preimages under ϕ then constitute a split of X. When we assume that the tree has no vertices of degree 2—which we shall henceforth do—different edges lead to subgraphs with different leaf sets, and therefore different edges induce different splits of X. Those splits then are compatible. We denote the splits of X induced by the X-tree (T, ϕ) by $\Sigma(T, \phi)$, or simply by $\Sigma(T)$ when the map ϕ is implicitly understood.

The converse question of what classes of splits of X come from X-trees is answered by the following result of Buneman

Theorem 2.3.1. *Given a collection Σ of X-splits, there exists an X-tree (T, ϕ) (which then is unique—up to isomorphism, of course) for which $\Sigma = \Sigma(T, \phi)$ precisely if all the splits in Σ are pairwise compatible.*

A tree carries an obvious metric, in the sense that we can quantify the distance between vertices i_1 and i_2 by counting the number of edges in the shortest path between them. More generally, we can assign positive weights $w(e)$ to the edges e and then take the sum of the weights of the edges in such a path as the distance $d(i_1, i_2)$.

When we consider a set X, there may already exist some distance function on X, and the question then emerges whether that distance is compatible with the metric on some X-tree. The answer is pretty simple, and in fact, we can even take something more general than a metric on X, namely a so-called dissimilarity map, that is, a non-negative map $\delta : X \times X \to \mathbb{R}$ with $\delta(x, x) = 0$ and otherwise positive, and

[9] That A and B yield a partition of X means that $A \cup B = X$ and $A \cap B = \emptyset$.

$\delta(x, y) = \delta(y, x)$ for all $x, y \in X$. For example, $\delta(x, y)$ could just count in how many characters (see below for a formal definition) the elements x and y differ.
The question then is whether we can find an X-tree (T, ϕ) with weights $w(e)$ on its edges and associated distance function $d(.,.)$ such that

$$\delta(x, y) = d(\phi(x), \phi(y)) \tag{2.3.1}$$

for all $x, y \in X$. In that case, we call δ a tree metric. The answer is

Theorem 2.3.2. *A dissimilarity map δ on X is a tree metric precisely if it satisfies the 4-point condition*

$$\delta(x, y) + \delta(z, w) \leq \max(\delta(x, z) + \delta(y, w), \delta(x, w) + \delta(y, z)) \tag{2.3.2}$$

for all $x, y, z, w \in X$.

In the sequel, (2.3.2) will give rise to two different issues. One is whether it holds or not for all points, and this issue is exemplified in the case where δ is the metric coming from a quadrilateral graph where x, w, y, z are arranged in cyclic order, for example $\delta(x, w) = \delta(w, y) = \delta(y, z) = \delta(z, x) = 1$ and $\delta(x, y) = \delta(z, w) = 2$. Thus, (2.3.2) is not satisfied here. The other issue arises when (2.3.2) is satisfied for all quadruples and consists in the question under which conditions we have even strict inequality for certain quadruples.
Since every edge e of an X-tree corresponds to a split σ of X, we can write a tree metric as

$$d = \sum_{\sigma \in \Sigma(T)} w(e_\sigma)\delta_\sigma \tag{2.3.3}$$

where e_σ is the edge inducing the split σ and

$$\delta_\sigma(i, j) = \begin{cases} 1 \text{ if } i, j \text{ are in different components of } T - e \\ 0 \text{ otherwise.} \end{cases}$$

The point here is that the edges e_σ occurring for $d(x, y)$ in (2.3.3) with $\delta_\sigma(x, y) = 1$ are precisely those contained in the shortest path from x to y.
This will now lead us to the decomposition theorem of Bandelt and Dress [12]. Let δ be a dissimilarity map on X. For a split $\sigma = A|B$ of X, we consider

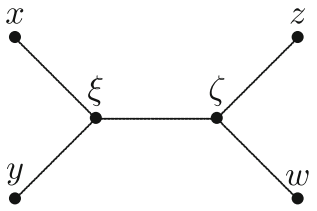
$$\begin{aligned} i_\delta(\sigma) :=& \frac{1}{2} \min_{a_1, a_2 \in A, b_1, b_2 \in B} (\max(\delta(a_1, b_1) + \delta(a_2, b_2), \delta(a_1, b_2) + \delta(a_2, b_1)) \\ & - (\delta(a_1, a_2) + \delta(b_1, b_2))). \end{aligned} \tag{2.3.4}$$

It is not required that the points a_1 and a_2 or b_1 and b_2 be different. For example, this expression can become negative when δ does not satisfy the triangle inequality:

take $a_1 = a_2 =: a$ and b_1, b_2 with $\delta(b_1, b_2) > \delta(b_1, a) + \delta(a, b_2)$.—In order to understand the significance of $i_\delta(\sigma)$ better, we consider some examples. These examples will be graphically displayed in the figure below. We first take a space $X = \{x, y, z, w\}$ consisting of 4 points, with the condition

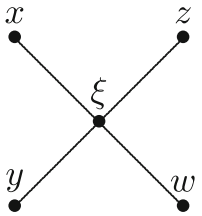
$$\delta(x, y) + \delta(z, w) < \max(\delta(x, z) + \delta(y, w), \delta(x, w) + \delta(y, z)). \tag{2.3.5}$$

If $\delta(x, y) = \delta(z, w) = 2$, $\delta(x, z) = \delta(x, w) = \delta(y, z) = \delta(y, w) = 3$, the split $\{x, y\}|\{z, w\}$ has index $i_\delta = 1/2$ and is induced from a tree with leaves x, y, z, w and interior nodes ξ, ζ with $\delta(x, \xi) = \delta(y, \xi) = \delta(\xi, \zeta) = \delta(z, \zeta) = \delta(w, \zeta) = 1$, see the following figure



(2.3.6)

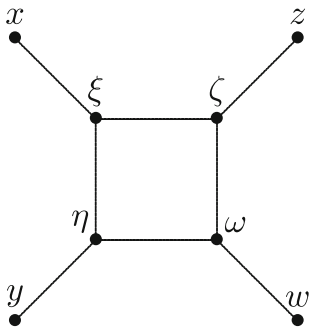
The splits $\{x, z\}|\{y, w\}$ or $\{x, w\}|\{y, z\}$, however, have $i_\delta(\sigma) = 0$ and are not induced by that tree metric. When we have equality in (2.3.5), say, $\delta(x, y) = \delta(z, w) = 2$, $\delta(x, z) = \delta(z, w) = \delta(y, z) = \delta(y, w) = 2$, the metric can still be represented by a tree metric, this time with a single interior vertex ξ that has distance 1 from all leaves. Thus, we have a 4-star



(2.3.7)

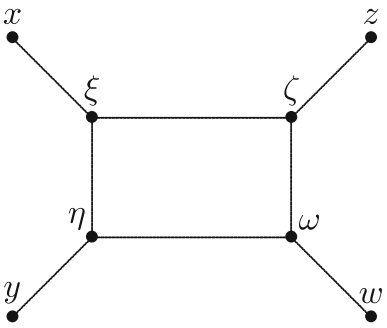
Here, there is no longer a natural grouping of the vertices into two pairs.—When instead $\delta(x, y) = \delta(z, w) = \delta(x, z) = \delta(y, w) = 3$, and $\delta(x, w) = \delta(y, z) = 4$, then (2.3.5) holds again. This time, we can represent the metric by a graph with 4 interior vertices ξ, η, ζ, ω that is not a tree. ξ, η, ζ, ω form a rectangle with $\delta(\xi, \eta) = \delta(\xi, \zeta) = \delta(\omega, \zeta) = \delta(\eta, \omega) = 1$, the other nontrivial distances between them being

equal to 2, and with x connected to ξ, y to η, z to ζ, w to ω, all with distance 1. Thus, we need to insert an interior rectangle in order to represent the metric on a graph,



(2.3.8)

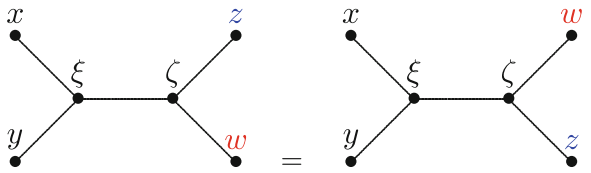
That rectangle then expresses the ambiguity in the dissimilarity map for a hierarchical grouping. Of course, the rectangle is in fact a square, and so there is some special symmetry. We therefore also consider the case where $\delta(x, y) = \delta(z, w) = 3,\ \delta(x, z) = \delta(y, w) = 4,\ \delta(x, w) = \delta(y, z) = 5$. In that case, we again insert 4 interior vertices ξ, η, ζ, ω that form a rectangle, this time with $\delta(\xi, \eta) = \delta(\zeta, \omega) = 1,\ \delta(\xi, \zeta) = \delta(\omega, \eta) = 2$,



(2.3.9)

In any case, when we have such a rectangle, we produce splits by cutting pairs of parallel edges. Cutting the edges between ξ and ζ and between η and ω, for example, produces the split $\{x, y\}|\{z, w\}$. Cutting the edges between ξ and η and between ζ and ω instead produces the split $\{x, z\}|\{y, w\}$. Now, in contrast to the tree case, both these splits have $i_\delta(\sigma) > 0$. The split $\{x, w\}|\{y, z\}$, however, has $i_\delta(\sigma) < 0$. In the

tree case, interchanging x with y or z with w would not have made any difference for the distances between those 4 vertices,



(2.3.10)

but this is no longer so in the rectangle case.

After this example, let us return to the general case. When δ is a tree metric from an X-tree (T, ϕ), the split σ of X is induced by that X-tree precisely if $i_\delta(\sigma) > 0$. In that case, we then have $i_\delta(\sigma) = w(e_\sigma)$ for the weight of the edge inducing the split. And we can rewrite (2.3.3) then as

$$\delta = \sum_{\sigma \ X\text{-split with } i_\delta(\sigma)>0} i_\delta(\sigma)\delta_\sigma. \tag{2.3.11}$$

The split decomposition theorem of Bandelt and Dress [12] then says that every dissimilarity map can be written as a sum over such tree metrics plus a remainder that has no splits with $i_\delta(\sigma) > 0$:

Theorem 2.3.3. *Let δ be any dissimilarity map on X. We then have a decomposition*

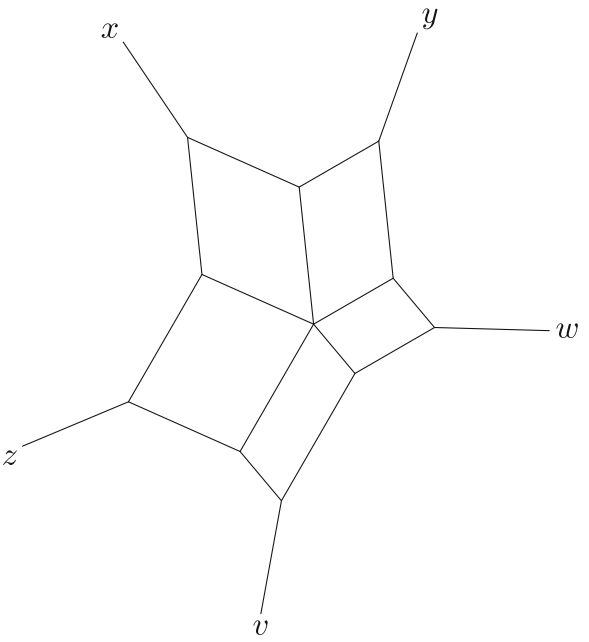
$$\delta = \delta_0 + \sum_{\sigma \ X\text{-split with } i_\delta(\sigma)>0} i_\delta(\sigma)\delta_\sigma \tag{2.3.12}$$

where δ_0 admits no splits with $i_\delta(\sigma) > 0$.

The star in our above example (2.3.7) admits no splits into pairs of points with $i_\delta(\sigma) > 0$. This is an undesirable situation in phylogenetic tree reconstruction because the grouping of the four vertices into pairs is ambiguous. However, when we split off a single point from the remaining three, we get $i_\delta(\sigma) > 0$. The simplest example of a metric space admitting no splits at all with $i_\delta(\sigma) > 0$ is given by 5 points x, y, z, w, v with $d(x, v) = d(y, z) = d(z, w) = d(y, w) = 2$ and the other distances between different points all being one. To describe this metric space somewhat differently, we take the two sets $A := \{x, v\}$, $B := \{y, z, w\}$ and connect each point in A with every point in B by an edge of length 1. Thus, we see that the graph constructed in this way is the bipartite graph $K_{2,3}$ of (2.2.9).—For this example, then δ_0 is nontrivial, and moreover, $d = \delta_0$.

When, conversely, δ_0 vanishes, the dissimilarity map δ is called totally decomposable. We recall that in the above example with the interior rectangle, the splits $\{x, y\}|\{z, w\}$ and $\{x, z\}|\{y, w\}$ both have positive $i_\delta(\sigma)$, and they decompose the metric. Thus, a

totally decomposable metric need not be a tree metric. The problem of this example for phylogenetic tree reconstruction is that there is no unique split that decomposes the dissimilarity map, that is, on the basis of the dissimilarity map, we do not know how to group the elements. Another, larger, example of this type, that is, of a totally decomposable metric that is not a tree metric and where therefore the groupings of the elements are not unique, is displayed in the next figure.

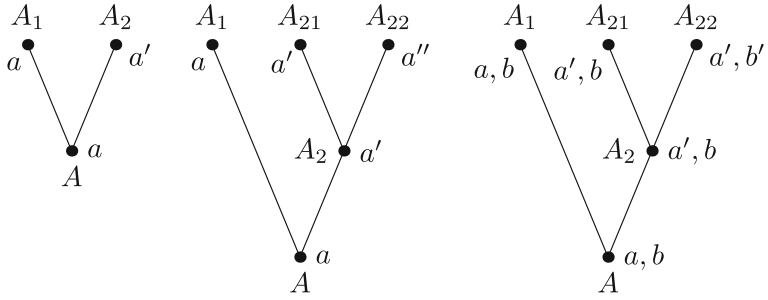


Bandelt and Dress[12] proved that a dissimilarity map δ is totally decomposable iff for all $x, y, z, v, w \in X$,

$$i_\delta(\{x, y\}|\{z, v\}) \leq i_\delta(\{x, w\}|\{z, v\}) + i_\delta(\{x, y\}|\{z, w\}). \tag{2.3.13}$$

Splits are decompositions of X into two subsets. More generally, we can consider characters, that is, functions $\chi : X_0 \to S$ where $\emptyset \neq X_0 \subset X$ and S is a finite set, the set of characters states. χ is called non-trivial if there are at least two character states that are each assumed by more than one element of X_0. We say that the character χ factors through the X-tree (T, ϕ) when there exists $\chi' : T \to S$ (here, we mean by a function on the tree T a function that is defined on the vertices of T) with $\chi = \chi' \circ \phi_{|X_0}$. Such a character χ that factors through the X-tree (T, ϕ) is called convex on T if for each $a \in S$, the subgraph with vertex set $(\chi')^{-1}(a)$ is connected. This is equivalent to the existence, for every pair a, b of different character states, of an X-split $A|B$ of T with $(\chi')^{-1}(a) \subset A$ and $(\chi')^{-1}(b) \subset B$.
The concept of character convexity is fundamental for the phylogenetic systematics developed by W.Hennig, the so-called cladism [58]. There, one wants to identify

monophyletic groups, that is rooted subtrees of phylogenetic trees that contain all the descendents of that vertex that is declared to be the common ancestor and made the root of the subtree. For example, in standard zoological systematics, vertebrates constitute a monophyletic group while fish don't because the other vertebrate groups (amphibians, reptiles, birds, and mammals) are also descendents of fish; in fact, here only birds and mammals are monophyletic in the sense of cladism. We consider an ancestral species A with daughter species A_1 and A_2.[10] Of course, this can be represented by a tree with root A and leaves A_1, A_2. Suppose that a character state a in A is preserved in A_1, but changed into a' in A_2. Now suppose that that species A_2 further splits into two daughter species A_{21} and A_{22}. We then get a new tree with leaves A_1, A_{21}, A_{22} if there are no further splittings while A_2 now is an interior node of degree 3. We consider two cases as displayed in the following figure.



In the first case, A_{21} preserves the state a' while in A_{22} it is further transformed into a''. In the other case, both of them preserve a', but in A_{22} the state of some other character is transformed into the value b' from the common value b shared by A, A_1, A_2, A_{21}. In such a situation, the ancestral states a, b are called plesiomorph, the derived states a', a'', b' apomorph. These are relative concepts because a' is plesiomorph compared with a'', that is, when we only consider the subtree with root A_2 and leaves A_{21}, A_{22}. Two species sharing the same plesiomorph state of a character are called symplesiomorphic w.r.t. that character, those sharing an apomorphic state are called synapomorphic. In the last example, A_1 and A_{21} are symplesiomorphic for b while A_{21} and A_{22} are synapomorphic w.r.t. a'. In the preceding example, where A_{22} had the character state a'', the states a', a'' together constitute a synapomorphy between A_{21} and A_{22}. Only synapomorphy, but not symplesiomorphy, can be an indication of a monophyletic group. Here then enters the convexity assumption. Namely, in order to be able to use shared derived characters, that is, synapomorphies for identifying monophyletic groups, we must exclude the following two possibilities:

1. Reversion: In the last example, A_{22}, instead of assuming the new state a'', reverts to the ancestral state a.

[10] It is a basic principle of cladism that whenever a new species splits off from some line, the remaining part of that line is also classified as a new species. This makes the systematics amenable to tree representations. Moreover, from the morphological approach underlying cladism that is based on paleontological data, any two species differ in the state of at least one character.

2. Convergence: In the same example, A_1, instead of keeping the state b, assumes the same state b' that originated in the species A_{22} while A_{21} kept b.

Of course, there exist biological examples for either possibility. Snakes have lost the limbs that their ancestors had gained. Birds, bats, and insects have independently developed wings. In fact, the wings of birds and bats are plesiomorph when considered as forelimbs, but not as wings. Sometimes, the distinction between plesiomorphy and apomorphy is not clear or needs to be reconsidered in the light of genetic sequence data. For example, it had been thought for a long time that the eyes in arthropods, molluscs, and vertebrates are examples of a convergent evolution. It has been discovered, however, that eye formation in all these lineages is directed by the same master control gene, called *Pax6*, from the class of homeotic (Hox) genes [101]. An uncontroversial[11] example of convergence is mimikry where one species imitates the coloration or other pattern of an unrelated species that is avoided by predators. In any case, reversion and convergence are relatively rare in biological evolution, however. Both these possibilities are excluded by character convexity.

When one has several characters, one wants to find a single X-tree for which all of them are convex. When such a tree exists, these characters are called compatible. As for compatibility of splits, there exists a theorem characterizing the compatibility of characters, but since the formulation is more complicated we refer to [107].

When working with biological data, typically not all the characters are compatible, and one then wishes to quantify that non-compatibility and construct a tree that comes as close as possible to rendering all the characters convex. This is the idea of parsimony. More precisely, given a function f on the vertex set V of a graph Γ, the changing number of f is the number of those edges of Γ on whose endpoints f assumes different values, that is, the number of all edges $e = (i, j)$ with $f(i) \neq f(j)$. Let now $\chi : X_0 \to S$ be a character that factors through the X-tree (T, ϕ), with $\chi = \chi' \circ \phi_{|X_0}$ as above. Here, we are assuming that χ' is already defined on all the vertices of T. Of course, it is then arbitrary how to define χ' on those vertices of T that are not in the image of $\phi(X_0)$, in case ϕ is not surjective on X_0. For a character χ, we then define its parsimony score $s(\chi, T)$ for the X-tree (T, ϕ) as the minimal changing number of all those extensions χ' on T that factor χ. Given a set of characters, its parsimony score on an X-tree then is simply the sum of the individual parsimony scores. A maximal parsimony X-tree for that set of characters then is one that minimizes that parsimony score.

It is not difficult to see that the parsimony score is related to character convexity. In fact, given a character that assumes ν different states, the so-called homeoplasy of the character χ on T

$$h(\chi, T) := s(\chi, T) - \nu + 1 \geq 0 \tag{2.3.14}$$

[11] at least as long as one does not look at the underlying genetic mechanisms; in fact, it may well turn out in a given example that the imitation of a pattern is produced by the same kind of genetic regulatory mechanism as the imitated pattern, or at least the general framework of that genetic regulation might be derived from some common ancestor

with equality precisely if χ is convex on T. Thus, the total homeoplasy of a character set, the sum of the individual homeoplasies, is also non-negative and vanishes precisely when the characters are compatible.
The concept of maximum parsimony trees is not without difficulties, both conceptually and mathematically. The conceptual difficulties arise from the arbitrariness in the definition and choice of characters. It is a fundamental problem in paleontology and morphology to clearly state what a character is and to decide which characters are independent of each other. Of course, large sets of dependent characters would bias the parsimony concept. The mathematical problems become clear when one considers stochastic processes on trees and other graphs. One then realizes that any method of reconstructing a structure from a data set depends on a model for the underlying process that created the data.
A standard problem is to amalgate phylogenetic relationships between subsets of X as expressed in trees into an encompassing tree representing all of X. Of course, the issue of compatibility will arise again. The smallest meaningful subsets here consist of 4 elements and are called quartets, and trees with 4 leaves are called quartet trees. Also, if one has data about the relationships between the elements of X and wants to construct a tree or, more generally, find out whether these relationships fit into a tree, a natural strategy is to first construct all local quartet trees and then assemble those into a common tree. When we have a collection $\mathcal{Q}$ of quartet trees that contains exactly one quartet tree $\{a, b\}|\{c, d\}$ for every quartet $Y = \{a, b, c, d\}$ of X, then, as discovered by Colonius and Schulze [29], there exists a unique X-tree containing all these quartet tree iff the following two quartet rules hold for all $a, b, c, d, e \in X$:

1.
$$\text{If } \{a, b\}|\{c, d\}, \{a, b\}|\{d, e\} \in \mathcal{Q}, \text{ with } c \neq e, \text{ then } \{a, b\}|\{c, e\} \in \mathcal{Q}$$

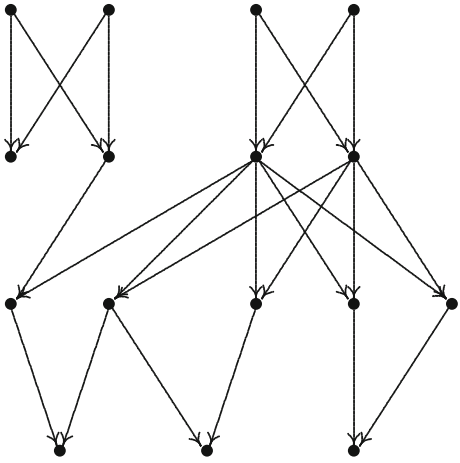
2.
$$\text{If } \{a, b\}|\{c, d\}, \{a, c\}|\{d, e\} \in \mathcal{Q}, \text{ then } \{a, b\}|\{c, e\} \in \mathcal{Q}.$$

In practice, of course, these rules will be violated for some quintuples of elements of X, and one therefore cannot construct a tree.
There are other, in fact infinitely many, quartet rules. If $\mathcal{Q}$ does not contain a quartet tree for every quartet in X, that is, if we only have a subcollection of quartet trees, then we need to invoke more of those rules to check for compatibility, see [107] for more on this topic. For an algorithm for the (re)construction of a tree from quartets, see [114].

2.3.2 Genealogies (Pedigrees)

While species can be considered as important biological entities in their own right, the ancestor-descendent relationships in phylogenetic trees can also be viewed as accumulated genealogies of the individuals constituting the populations underlying

the species. Thus, let us consider those genealogies a little, even though they in turn can be viewed as combinations of inheritance processes of genes passed on from parents to offspring. The latter, in fact, will lead us back to trees below.
The genealogy or pedigree of an individual in a sexually recombining population is a directed graph. Each individual has two incoming links from its parents while the number of outgoing links counts its offspring. Since no individual can be a descendent of its own offspring, or an ancestor of its own parents, the graph is acyclic (it has no directed cycles; the underlying undirected graph may well have cycles as the result of inbreeding in the population). The nodes without outgoing links represent those individuals that did not produce or have not yet produced offspring.

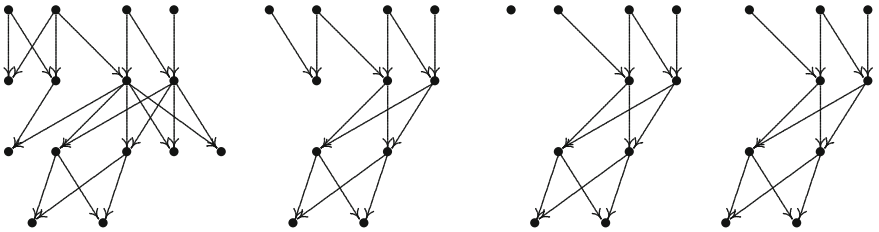


A pedigree graph; time runs downwards (2.3.15)

In the pedigree graph (2.3.15), in the ancestral generation 1, we have two pairs that produce two offspring each. In generation 2, one individual leaves no offspring whereas another one contributes to five of them. In contrast, in generation 3, every individual leaves one or two descendents in generation 4. In a bisexual population, we can also identify two subgraphs, one corresponding to the females and the other to the males. In those subgraphs, every node then has precisely one incoming arrow. We shall return to this issue in Sect. 2.3.3.

When the graph represents a population history, one can essentialize it by pruning all the vertices without outgoing links that correspond to individuals having died without leaving offspring. This will then be an iterative process because in the next step one would have to prune those vertices that have outgoing links only to vertices that have been pruned in the previous step. In that manner, one iteratively eliminates all vertices that do not have living descendents. Thus, one is left with the ancestral relationships leading to the present population.

Since one does not want to extend the pedigree to the infinite past, one starts with some ancestral population. The essentialized pedigree then contains only those members of the ancestral population that have descendents in the present generation. If one moves further to the next generation, then some of those ancestors may cease to have descendents and therefore will get eliminated. Some of those ancestors, called the lucky ones, however, will turn out to be ancestors of all members of the present population, and they will therefore also leave descendents in all future generations, until the entire population goes extinct.



A pedigree graph and its prunings (2.3.16)

In the pedigree graph of (2.3.16), there are three such lucky ancestors from whom the current population of two individual descends.

Often, one assumes that the different generations do not overlap, as in (2.3.15), (2.3.16). The generations can then be labelled by their distance from the ancestral one, and links always go from generation n to generation $n + 1$.

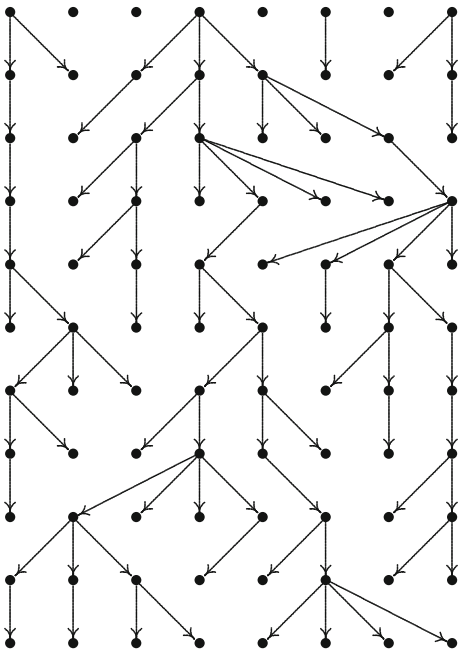
Also, from the pedigree graph, one can construct another graph expressing mating relationships. In that graph, there is an (undirected) edge between two individuals when they have produced offspring together. When the species is bisexual or dioecious, that is, has separate sexes, the mating graph is bipartite, the two classes corresponding to the females and the males. The mating graph usually is not connected, however, therefore, strictly speaking, violating our definition of a graph. When the population is strictly monogamous, the graph consists of disjoint pairs only, after we have essentialized it and eliminated all the bachelors and spinsters.

Of course, this is all rather simple. Later on, when we consider stochastic branching processes, pedigrees of sexually recombining populations become rather difficult, but for the moment we leave the subject and turn to

2.3.3 Gene Genealogies (Coalescents)

The pedigree just considered for a dioecious population, i.e., with two different sexes, contains two trees (more precisely, so called forests, that is, not necessarily connected unions of trees) as subgraphs, namely those corresponding to the male and the female individuals. Let us take one of them, say the female one. Thus, we only consider mother-daughter relationships. For two individuals, we can then ask when their lineages coalesce or merge back in the past, that is, how many generations back they had the same female ancestor. For two sisters, we need only go one generation back, as they have the same mother, while first (in the female line) cousins share a maternal grandmother, and so on. Once the lineages coalesce, they will stay together all the way back to the ancestral population. Of course, in principle, they may never merge, that is the two females under consideration may be descendents of different females in the ancestral population. When we go sufficiently many generations back into the past, however, with overwhelming probability, all presently living females in the populations will descend from the same ancestral female, the "Eve". All other females in that ancestral population will then have no descendents from an uninterrupted female line in the present populations; of course they may or may not have descendents from some lineages that include some males. As already described above, we can essentialize the graph by eliminating all females without female descendents in an iterative manner so that only those remain that have an uninterrupted line of female descendents down to the present sample. When we do coalescence theory, that is, follow the ancestry of the present sample back in time, then, in fact, those eliminated individuals will never occur in the consideration. This represents an enormous simplification in practice when compared with considering the forward branching process for the (female) descendents of an ancestral populations where all descendents will occur regardless of whether they contribute to future generations or not.

Let us consider this scenario in more detail in a simple example that will lead us to the Wright-Fisher model of population genetics. We consider a population with non-overlapping generations, and we assume that the size of the population remains constant $= 2N$ across generations. We also assume, for simplicity, that the sex ratio remains constant and equal so that we are dealing with a population of N females. The assumption of the Wright-Fisher model is that, given generation n, consisting of a population of N individuals, generation $n + 1$ is (mathematically) created by choosing N times randomly and independently an individual from generation n as mother.



A Wright-Fisher genealogy with 8 individuals and 11 generations; note that we have arranged the order of the individuals in each generation so as to render the scheme clearer. (2.3.17)

The 4th individual from generation 1 is the sole ancestors of all individuals in generation 10, and the 5th individual from generation 6 is the sole ancestors of all individuals in generation 11.

Here, it is assumed that the population is entirely homogeneous, or, in more biological terms, that all members are equally fit, so that at each selection step, each member has the same chance of being chosen. Also, creating daughters does not affect the fitness, and so, the chance to be chosen at a given step does not depend on how often one has already been chosen in previous steps. Putting it another way, each individual in generation $n+1$ picks individual j in generation n with probability $1/N$ as its mother, and this sampling is carried out N times with replacement. If d_j is the number of daughters of individual j, we thus have for the probability of having ν daughters

$$p(d_j = \nu) = \binom{N}{\nu} (\frac{1}{N})^{\nu} (1 - \frac{1}{N})^{N-\nu}. \tag{2.3.18}$$

This is a binomial distribution, $Bi(N, \frac{1}{N})$, and so, the number of daughters of a given female is binomially distributed. The binomial distribution will be introduced more systematically in Chap. 3.1, see (3.1.8). The expectation value is

$$E(d_j) = N\frac{1}{N} = 1 \tag{2.3.19}$$

which of course reflects the fact that the population size is constant, and the variance is

$$Var(d_j) = N\frac{1}{N}(1 - \frac{1}{N}) = 1 - \frac{1}{N} \tag{2.3.20}$$

(see (3.1.20) below). The correlation between the numbers of daughters of different females j, k is

$$Cor(d_j, d_k) = \frac{Cov(d_j, d_k)}{\sqrt{Var(d_j)Var(d_k)}} = -\frac{1}{N-1}. \tag{2.3.21}$$

The correlation is negative, again because the population size is constant, and therefore, when j has many daughters, there is less room for k to have many daughters as well (when we already know that an individual different from j has one daughter, then the expected number of daughters of j is reduced to $\frac{N-1}{N}$ in place of the value 1 of (2.3.19)). This effect is rather small in large populations.
For large N, the binomial distribution $Bi(N, \frac{1}{N})$ is approximated by a Poisson distribution

$$p(d_j = \nu) \approx \frac{1}{\nu!}e^{-1} \tag{2.3.22}$$

with mean and variance =1 (see (3.1.8), (3.1.9) in Chap. 3.1 below). In particular, the probability of having no daughters is

$$p(d_j = 0) \approx e^{-1} \approx .37 \tag{2.3.23}$$

while then the probability to have at least one daughter becomes

$$p(d_j > 0) \approx 1 - e^{-1} \approx .63 \tag{2.3.24}$$

Therefore, the present population descends from a fraction of about $.63^n$ females n generations ago. Of course, this eventually goes to 0 for large n which leads to the absurd result that the present females derive from fewer than one individual in the ancestral generation. Of course, the puzzle is resolved by observing that these approximations were only valid for large population sizes. For small populations, a more refined analysis is needed. This is the subject of coalescence theory, originally founded by J. Kingman [81].

Again, we stay with our simple example and ask for the distribution of the number T_2 of generations that we need to go back in time to find a common ancestor of two individuals from the present population. That is, we seek the time to the most recent common ancestor (MRCA) of the two individuals. The probability that the two individuals have the same mother, that is, that the MRCA is found already in the first generation from the past, is $\frac{1}{N}$ because once we have identified the mother of the first individual, the probability that the second one has the same mother is $\frac{1}{N}$. Thus, the two have different mothers with probability $1 - \frac{1}{N}$. Iteratively, the chance to find the MRCA n generations back then is

$$p(T_2 = n) = (1 - \frac{1}{N})^{n-1} \frac{1}{N} \tag{2.3.25}$$

because they then have different ancestors in $n - 1$ generations. This is a geometric distribution, and its mean is

$$E(T_2) = \frac{1}{\frac{1}{N}} = N \tag{2.3.26}$$

which is equal to the population size.
In a similar manner, we can consider the time to find the MRCA for M individuals. The probability that m individuals have all different mothers is

$$\frac{N-1}{N}\frac{N-2}{N}\cdots\frac{N-m+1}{N} = \prod_{\mu=1}^{m-1}(1 - \frac{\mu}{N}) = 1 - \binom{m}{2}\frac{1}{N} + O(\frac{1}{N^2}) \tag{2.3.27}$$

because when the mother of the first individual is determined, there are $N - 1$ possibilities for the second to have a different mother, and when that is also determined, there remain $N - 2$ possibilities for the third individual to have a mother different from the previous two, and so on. Thus, neglecting terms of order $\frac{1}{N^2}$ for a large population size N, a coalescence event occurs in a given generation with probability $\binom{m}{2}\frac{1}{N}$, while no coalescence event occurs with probability $1 - \binom{m}{2}\frac{1}{N}$. Thus, the probability distribution for the time T_m of a coalescence event that reduces the number of different ancestors from m to $m - 1$

$$p(T_m = n) = (1 - \binom{m}{2}\frac{1}{N})^{n-1}\binom{m}{2}\frac{1}{N}. \tag{2.3.28}$$

In analogy to (2.3.26), we have

$$E(T_m) = \frac{1}{\binom{m}{2}\frac{1}{N}} = \frac{2N}{(m-1)m}. \tag{2.3.29}$$

When we then want to go back from M individuals to a single ancestor, we need to consider all the coalescent events from m to $m-1$ for $m = 2, \ldots, M$. Since the times for these events are independent of each other, the expected number of generations back in the past for M female individuals to have a single female ancestor is

$$\sum_{m=2}^{M} E(T_m) = \sum_{m=2}^{M} \frac{1}{\binom{m}{2}\frac{1}{N}} = \sum_{m=2}^{M} \frac{2N}{(m-1)m} = 2N(1 - \frac{1}{M}). \qquad (2.3.30)$$

In other words, this is the expected height (measured in number of generations) of the tree starting with a single female ancestor and leading to the present ensemble of M females. When we compare (2.3.30) with (2.3.26), we see that the latter is less than 2 times the former. This means that the final step of reducing the number of ancestors from 2 to 1 typically takes at least half the time of the whole process. Thus, the long branches of the tree arise when there are only few females in the ancestry of the sample.

One can, of course, perform the same analysis with males in place of females. Let us insert a small variation, however, to account for the fact that in many animal species, like most mammals, and also in many human societies, the variance in the number of offspring for males is considerably higher than for females while obviously the expectation value is the same, assuming that the population is in gender equilibrium (that issue will be treated in Sect. 5.1). This higher variance is easy to achieve in our model. The simplest version just stipulates that in each generation only a certain fraction $0 < q < 1$ of the number of males is having offspring at all. When we then look for the father of an individual, each of those ones is taken with probability $1/qN$ while the other ones are simply discarded. Thus, two individuals now stand a chance of $1/qN$ of having the same father. Thus, N gets replaced by qN in all formulae. In particular, the expectation values for the waiting times in (2.3.26), (2.3.30) are shortened by a factor q, and we expect to find the MRCA in the male line correspondingly fewer generations ago than that in the female line. In other words, "Adam" lived many generations after "Eve". (In fact, it has recently been discovered that there is a small number of males of African origin that carry Y-chromosomes of different origin [90]. (The Y-chromosomes determine the male gender; a female possesses two X-chromosomes, a male one X- and one Y-chromosome; the latter are therefore passed on only the male line.) Thus "Adam" is not the male ancestor of all living humans, but only of the vast majority of them.)

Coalescence theory is mainly interested in describing the ancestry of genes, or more precisely, of DNA segments, rather than of individuals. Formally, the basic scenario is the same, however, and therefore, we have described the basic situation above for the more intuitive case of individuals. The basic scenario neglects the issues of mutation and recombination. In order to exclude recombination, there are two possibilities, one of significance for biological data, the other solely for modeling purposes. The first one consists in considering those DNA segments that do not recombine. One class is given by non-nuclear mitochondrial DNA that is only contained in egg cells, but not in sperm, and therefore is only passed on in the female line. This, in fact,

makes the above example of female lineages relevant for treating biological data. The other example is the Y-chromosome in humans and other mammals which is only carried by males (and determines the male gender) and therefore is only transmitted in the male lineage. The mathematically convenient solution, in contrast, is to simply consider the smallest DNA segments, the single nucleotides. The biological problem here is that even though each nucleotide is derived from a unique parent, this usually cannot be identified from genetic data because in a given species, at most positions, most members share the same nucleotide. Nevertheless, for so-called SNPs, single nucleotide polymorphisms, consideration of single nucleotide positions can contain some useful population biological information. Even when we consider single nucleotide positions, however, we only get rid of the problem of recombination while absence of mutations, and of other processes of genetic rearrangement, then is still a hypothesis imposed. For simplicity of the model, we also consider the haploid case where each individual has only one set of genes. Thus, each such DNA segment in an individual is derived from one of the parents. (In the diploid case, each individual has two sets of genes. The genes corresponding to each other in those sets are selected from different parents, that is, one is taken from the mother and the complementary one from the father. This imposes additional restrictions when compared with the haploid case, but typically their effect is not so prominent.) For single nucleotides in the haploid case which we shall now consider the situation is formally the same as that where one of the two parents of each individual is its mother. So, one might call that individual that gives the nucleotide in question the nucleotide parent for that particular position in the DNA. When in turn we consider only nucleotide parents, for a fixed position, then the situation is the same as before, with the only formal difference that two siblings can derive the nucleotide at that position from different parents. Therefore, the size of the population that has to be taken into account is $2N$ in place of N.

In any case, for each such nucleotide position, we can perform the coalescent analysis and find the expected number of generations for having a single ancestor. Of course, the ancestors for the different positions will in general be different individuals. We can then also ask the following questions:

1. What is the expected number of generations for finding a single ancestor for each position? That is, what is the expected maximal height of the coalescence trees for a given population?
2. In the corresponding ancestral population, how many individuals are ancestors for some position for the present sample? Those lucky ancestors then are ancestors of every individual in the present sample whereas the remaining members of the ancestral population then are not genetically represented at all in that present sample because for each position, there is only one ancestor by assumption.
3. Actually, the last issue is a little more subtle. In principle, an individual in the ancestral population can be a genealogical ancestor of a present one without being genetically represented in the latter. What are the chances for that? Here, one should essentially use some tree counting arguments. The pedigree graph contains many trees with a root in the ancestral population and leading down

to the present population, and each such root then represents a genealogical ancestor. Not all of these trees, however, arise from the coalescence processes just investigated.

So far, only some partial answers are known to these questions, see [57] for a brief discussion and references. Here is an example that exhibits some of the problems. Until relatively recently, modern man, Homo sapiens sapiens, was not the only human species. The best known other such species are the Neanderthals who became extinct less than 30,000 years ago. This raises the question whether these different human species lived just alongside each other for a certain period, and perhaps violently competed, or whether they also mixed and interbred to a certain degree. The question is who are the ancestors of present humans, only some group of individuals that originated in Africa and whose descendents then spread to the other continents, or whether other human species that had lived in Africa and Eurasia prior to the spread of this lineage also contributed to our genomes. This question has been addressed in recent years both via the analysis of the genomes of living people from various populations in the world and from the analysis of the DNA of the bone fossils of Neanderthals and other human species, to the extent that this is technically feasible. First, it was found that in the mitochondrial DNA, no evidence of an admixture can be found. This DNA is not contained in the nucleus of a cell, but only in some organelles. It is therefore passed on only along the female line, as male sperm does not contain those organelles, but only female egg cells do. However, the more recent sequencing of large parts of the Neanderthal genome [54] showed that Eurasians do carry some percentage of DNA inherited from Neanderthals. Thus, the coalescence trees for different parts of the human genome are significantly different.

Exercises for This Chapter

1. This exercise introduces the perhaps best known combinatorial design problem. A Hadamard matrix is an $n \times n$ matrix whose entries are 0 or 1,[12] with the property that any two rows share precisely $n/2$ entries. For $n = 2$, an example is

$$H := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

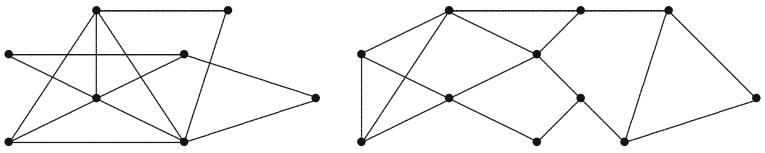
Here, the two rows share the first entry, but differ at the second position. For $n = 4$, an example is

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

[12] In the literature, more precisely, it is required that the entries be 1 or -1, but this leads to an equivalent problem.

Here, any two rows agree in precisely two positions, as required. For instance, the third and the fourth row share the first and the third entry. Construct a Hadamard matrix for $n = 2^k, k \in \mathbb{N}$. (By understanding how the example for $n = 4$ is constructed from the building block H for $n = 2$, you will probably rediscover a construction first found by Sylvester, a long time before Hadamard.) Try to find Hadamard matrices for other values of n.
It is conjectured that a Hadamard matrix exists if and only if $n = 2$ or $n = 4m$ for some $m \in \mathbb{N}$, but this is as yet unsettled.

2. Here is an easy exercise: Consider the graph whose vertices are the members of a population and whose edges represent matings between them (for simplicity, we do not consider the number of matings between the same pair, but only discuss the unweighted graph with an edge between two individuals that have mated at least once). What qualitative properties should this graph possess? How can you read off particular mating structures in the population, like polygamy or monogamy, polygyny (a male individual may have several mating partners, a female only one) or polyandry (the other way around)?
3. Another easy one: Argue that a trophic network, or put in simpler words, a food web, whose vertices are species in an ecosystem and an edge between two vertices expresses that one species feeds upon the other one, should be represented by a directed graph that does not contain directed cycles. Or should we admit exceptions? Estimate how long a directed path could maximally be (this is called the number of trophic levels—you may want to check the biological literature on this issue). Develop criteria in terms of the structure of this directed graph for assessing the importance of a particular species for an ecosystem.
4. List all non-isomorphic connected graphs with 5 vertices.
5. What is the smallest order for which there exists a graph without any nontrivial automorphism?
6. Determine the clustering coefficients and the k-cores of the following graphs and estimate their Polya-Cheeger constants,



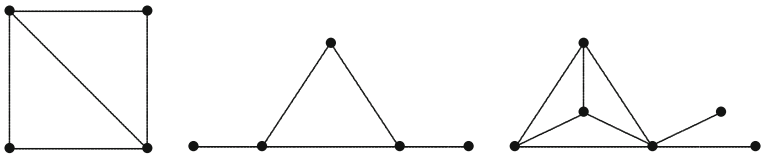
7. Let Γ be a k-regular graph with N vertices. Denote the eigenvalues of the adjacency matrix of Γ by $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_N$. What is the relationship between the μ_j and the eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1}$ of the normalized Laplacian of Γ?
8. (a) The N-cycle C_N is the graph of N vertices $\{i_1, \ldots, i_N\}$ where vertex i_k is connected with the vertices i_{k-1} and i_{k+1} mod N. Show that its eigenvalues are $1 - \cos \frac{2\pi j}{N}$, $j = 0, \ldots, N-1$.

(b) The N-path P_N is obtained from the N-cycle C_N by cutting the link between the vertices i_1 and i_N. Show that the eigenvalues of P_N are $1-\cos\frac{\pi j}{N-1}$, $j = 0, \ldots, N-1$.

(c) Show that the eigenvalues of the N-cube Q_N (which has 2^N vertices) are $\frac{2j}{N}$, $j = 0, \ldots, N$, with multiplicities $\binom{N}{j}$.

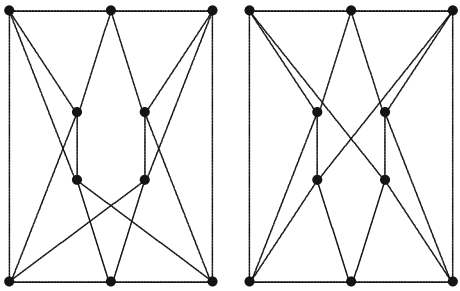
(d) The m-petal graph has one central vertex i_0 and $2m$ peripheral vertices $i_1, \ldots, i_{2m}$ such that i_0 is connected with every other vertex and in addition, vertex i_{2j-1} is connected with vertex i_{2j} for $j = 1, \ldots, m$. Show that its eigenvalues are 0, $\frac{1}{2}$ with multiplicity $m-1$, and $\frac{3}{2}$ with multiplicity $m+1$.

9. Determine the spectra of the following graphs,



by using symmetries and/or node duplications.

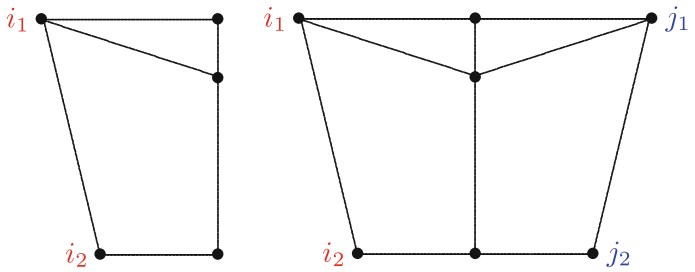
10. Here is a more difficult exercise. Show that the following two graphs have the same spectrum, i.e., they are isospectral.



(Note: This example is taken from [118], but you need an additional step to solve this exercise, because in that reference, the spectrum of the adjacency matrix is studied instead of that of the Laplacian—see a preceding exercise for the relationship between the two.)

11. Take a graph Γ with an edge $i_1 \sim i_2$ and create a graph Γ' by duplicating that edge, i.e., add two vertices connected by an edge, $j_1 \sim j_2$, to Γ and connect j_1

to all neighbors of i_1, j_2 to all neighbors of i_2. What can you say about the effect on the spectrum? Determine the spectrum of the following example.



Also, observe that the m-petal graph described in one of the preceding exercises is obtained from the complete graph K_3 by successive edge duplications. Use this observation to explain its spectrum as computed in that preceding exercise.

12. What constraints does a pedigree graph in a bisexual population have to satisfy?

http://www.springer.com/978-1-4471-6352-7

Mathematical Methods in Biology and Neurobiology
Jost, J.
2014, X, 226 p. 37 illus., 13 illus. in color., Softcover
ISBN: 978-1-4471-6352-7