LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Graduate School of
Quantitative Biosciences Munich

Biophysics Primer

Lecture Notes

Dr. Markus Hohle

Graduate School of Quantitative Bioscience Munich (QBM)
Feodor–Lynen–Str. 25, 81377 Munich;
hohle@genzentrum.lmu.de;

Version IV; 2017

Contents

1 Preface

In the last decades biology and biochemistry became more engineered. Be it that we apply advanced experimental methods in structure biology like cryo-electron microscopy or using sophisticated optical methods in synthetic biology or other subjects. We use thermodynamic models to understand transcription factor binding, stochastic differential equations to model gene expression or well elaborated Hidden Markov Models for motif finding. Minimizing entropy is a common concept in order to perform proper multiple sequence alignments in bioinformatics. Understanding the formation of stable gradients and patterns in the process of cell differentiation and segmentation requires the understanding of diffusion reactions, hence differential equations. We need tools like the Taylor expansion to deal with correlation in data sets, to estimate statistical errors or to perform proper model approximations. Applying all these methods, understanding their results and interpreting the biological impact of these results require a profound knowledge and understanding of various mathematical tools. It is inevitable that life science develops further into this (more quantitative rather then qualitative) direction in the near future. Thus, as bioscience became more quantitative and data analysis more automatized, physicists and mathematicians risk to replace biochemists because more sophisticated math, physics and programming skills are required now. To stop this process, at least a basic understanding of mathematical and physical concepts is needed for young life scientists in order to be successful. The language of nature is physics, written in math – it is not our goal to speak this language as mother tongue, but you should understand a few words!

These lecture notes contain topics that are inspired by the projects of the QBM students provided by their PIs. However, before starting with the actual biophysics, we first have to repeat some mathematics and physics from high school and second study further tools of higher math that we will need in many occasions. In contrast to typical life science lectures, that are mostly a presentation of a giant bunch of facts that have be to memorized by the students, in math and physics we have to derive an idea, a model or an equation step by step along a line of logical argumentation in order to understand the method completely. Then – and only then – it is possible to apply basic principles to any given situation or to adapt a model to specific requirements. In fact, it is not possible to *learn* math – this is already a misleading approach. One has to *understand* the mathematical tools. The good news is that math is, in contrast to its reputation, not difficult. Everyone who understands e.g. the trading on a stocks marked or is able to write his/her annual tax declaration should have no problems with the mathematical level in this script.

I recommend to read this script parallel to the lectures and to do the exercises provided in the script and in the lectures. The first part of this script (section 2) mainly corresponds to the content of the crash course "introduction on mathematical methods and physics". It should be your aim to understand each step and everything in detail, because these are the very basics, it is the foundation of your further work.
Thereafter, you should be able to understand the subsequent sections (3 to 8). It is not a problem if you do not understand each mathematical step of a derivation to its very details, but please try to follow the scope of the main ideas.

I put some effort to write this script in a relatively short time during the semester breaks and I like to thank my QBM students, in particular Marta Bozek (PCR example), Beatrice Ramm (microfluidics section), Lukas Kater, Hanna Kratzat (both cryo-EM) and Victor Solis

(mass spectrometry) for helpful suggestions, useful hints and for providing me data.
I also thank Erwin Frey and his students Patrick Dendorfer, Tobias Hermann, Jacqueline Janssen, Maximilian Lechner, Chris Lund, Morgan Maher and Moritz Striebel for improving the script by proofreading, adding new ideas and suggestions, updating the figures and complementing the examples. I further like to thank Christophe Jung from the Gaul lab and Baccara-Jale Hizli for providing me images for the Super Resolution and the noise filtering parts, respectively (see references therein) and Thomas Becker for providing his material for cryo-EM.
I would appreciate if the QBM students send me further ideas for improvement and examples or giving me hints when finding typos or inaccuracies.

Have fun ☺

Munich, summer 2017

2 Introduction on Mathematical Tools

2.1 Basics

The aim of this chapter is to repeat the topics like calculus, vector analysis, approximation methods, statistics and complex numbers that have been taught in high school and to introduce some additional mathematical tools and concepts like Fourier transformation or Bayesian statistics that are necessary for the subsequent chapters. For example, I expect the reader to be familiar with concepts like the dot product of vectors or the meaning of derivatives and integrals. Nevertheless, these mathematical tools are part of the subsequent sections.
Depending on how intensively we make use of the respective mathematical tools, some topics are elaborated in more or less detail. Since in some cases, the benefits are not immediately clear, I refer to later chapters where the particular methods are used. I strongly recommend to follow *all* examples and to do *all* exercises provided in this section.
In many cases I am not using the full mathematical rigorousness on purpose. For example I do not derive or proof the validity of the Taylor expansion or investigate the convergence of this series. It would be just too much algebra. However, in this case the correct application is far more important since we will widely use this particular tool. I also will not emphasize e. g. the difference between the gradient of a scalar field and a vector field too much since it would require the introduction of *tensors* – a mathematical object that you will rarely encounter. The drawback is that we have to accept some mathematical inaccuracies in particular when discussing the Navier-Stokes equation or coordinate transformations. The same applies for the surface element and its normal vector when introducing operators like *divergence* or avoiding co-variant and contra-variant vectors in X-ray crystallography. I hope that the conceptual explanation and the corresponding sketches help to compensate these problems.

2.1.1 Scalars and Vectors

In this course we will encounter many different quantities and mathematical constructions. It is useful to categorize them and one main property is the direction of a quantity.
Some quantities like mass, energy, temperature or density have no direction. A quantity with no direction is called a *scalar* (a "number"). A scalar is fully characterized by a number, e.g. $\text{mass} = 1\ \text{kg}$.
In contrast to scalars, some quantities have a direction, e.g. force, velocity or flux. Such quantities are characterized by an absolute value (e.g. the speed of a particle) **and** a direction (in which direction the particle moves). Such a quantity is represented by a *vector*. If we denote the three spatial coordinates with x_1, x_2 and x_3, then velocity has the coordinates $\vec{v} = (v_{x1}, v_{x2}, v_{x3})$. For example $v_{x1} = 5\ m/s$, $v_{x2} = -1\ m/s$ and $v_{x3} = 2\ m/s$ means, that an object having this velocity moves 5 meter towards x_1 direction, 1 meter against x_2 direction and 2 meter towards x_3 direction within one second.
In order to distinguish a vector from a scalar one writes an arrow above the variable: $\vec{v}$. This vector gives the direction of the motion, whereas the absolute value of the velocity (i.e. its length, the *speed*) can be calculated by adding up all the velocity "parts" from the different spatial coordinates according to the law of Pythagoras[1] (Figure 1), hence $|\vec{v}| = \sqrt{v_{x1}^2 + v_{x2}^2 + v_{x3}^2}$.
In a two dimensional coordinate system with the spatial coordinates (x_1, x_2) we write the

[1] Pythagoras of Samos, 570 - 495 BC

velocity as $\vec{v} = (v_{x_1}, v_{x_2})$. We may also write:

$$\vec{v} = \begin{pmatrix} v_{x1} \\ v_{x2} \end{pmatrix}. \tag{2.1}$$

Equation 2.1 can be easily generalized to the N-dimensional case. Then the absolute value

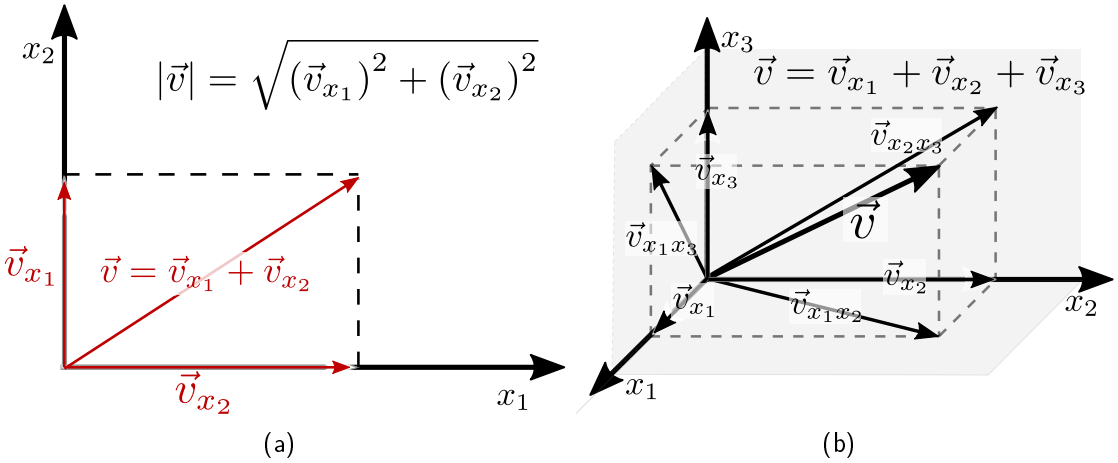

Figure 1: (a): A two dimensional vector is the sum of each of its one dimensional components. The absolute value of a vector (its length) follows from the law of Pythagoras.
(b): The same situation as left, but for a three dimensional vector. The projections in the $x_1\,x_2$-,$x_1\,x_3$- and $x_2\,x_3$-plane are labeled with $\vec{v}_{x_1\,x_2}$ etc.

reads

$$|\vec{v}| = \sqrt{\sum_i (v_{x_i})^2}. \tag{2.2}$$

that is just Pythagoras' law in N dimensions. A coordinate system where we can define vectors is called a *vector space*[2].
Addition/ subtraction of vectors and the multiplication by a scalar $\alpha \in \mathbb{R}$ of vectors are defined by

$$\vec{v} = \vec{u} \pm \vec{w} = \begin{pmatrix} u_{x_1} \pm w_{x_1} \\ u_{x_2} \pm w_{x_2} \\ u_{x_3} \pm w_{x_3} \end{pmatrix} \tag{2.3}$$

$$\alpha\,\vec{v} = \alpha \begin{pmatrix} v_{x_1} \\ v_{x_2} \\ v_{x_3} \end{pmatrix} = \begin{pmatrix} \alpha\, v_{x_1} \\ \alpha\, v_{x_2} \\ \alpha\, v_{x_3} \end{pmatrix} \tag{2.4}$$

Even a multiplication between two vectors $\vec{v}$ and $\vec{w}$ is defined: the so called *inner product* or *dot product* (often denoted as *scalar product*). In Cartesian coordinates, the dot product is written as

$$\vec{v} \cdot \vec{w} \equiv \langle \vec{v}, \vec{w} \rangle = (v_{x_1}, \cdots, v_{x_n}) \cdot \begin{pmatrix} w_{x_1} \\ \vdots \\ w_{x_n} \end{pmatrix} = \sum_i^n v_{x_i} w_{x_i}. \tag{2.5}$$

Note the different notations $\vec{v} \cdot \vec{w}$ and $\langle \vec{v}, \vec{w} \rangle$ that mean essentially the same (emphasized by the sign $\equiv$ that means "identical" in contrast to an assignment to a variable $=$). In

[2]Note that this is not the exact mathematical definition of a vector space, but it is sufficient for our purposes.

physics text books, the notation $\langle \vec{v}, \vec{w} \rangle$ is preferred since there is a more general approach that is, however, not in the scope of this script.
As the name *scalar product* implies, the result of the inner product is a scalar. Comparing Equation 2.2 and Equation 2.5 yields $\vec{v} \cdot \vec{v} = |\vec{v}|^2$. For the inner product the following relation holds:

$$\vec{v} \cdot \vec{w} = |\vec{v}|\,|\vec{w}| \cos(\phi) \tag{2.6}$$

where ϕ is the angle between $\vec{v}$ and $\vec{w}$ (see Figure 2(a)). Therefore, two orthogonal vectors have an inner product which is equivalent to zero, since $\cos(\pi/2) = 0$. It is actually the definition of orthogonality in an abstract sense: *If vectors have a inner product that equals zero, these vectors are orthogonal to each other.*
Another useful property of the inner product is that it is a projection of one vector onto another vector. This is illustrated in Figure 2: we know that $cos\,(\phi) = \frac{|\vec{\gamma}|}{|\vec{w}|}$ and also Equation 2.6 is valid, that leads to

$$|\vec{\gamma}| = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}|} = \vec{n}_v \cdot \vec{w} \tag{2.7}$$

where $\vec{n}_v$ denotes the *normal vector* of $\vec{v}$. $\vec{n}_v$ is called normal vector since it results from the normalization by its own length: $\frac{\vec{v}}{|\vec{v}|}$.

Exercise I:

Show that any normal vector has the length 1.

Thus, $\vec{\gamma}$ is the projection of $\vec{w}$ onto $\vec{v}$, similar to e. g. a stick in the ground casting a shadow (see Equation 2.7).

The axioms (i. e. "the rules of how to do math with them") for the inner product on a real vector space $\mathcal{V}$ are given (for three vectors $\vec{u}, \vec{v}, \vec{w} \in \mathcal{V}$ and a scalar $\alpha \in \mathbb{R}$) by the following relations:

$$\begin{aligned}
\langle \vec{v}, \vec{v} \rangle &\geq 0 && \text{(2.8a)}\\
\langle \vec{v}, \vec{v} \rangle &= 0 \Leftrightarrow \vec{v} = 0 && \text{(2.8b)}\\
\langle \vec{v}, \vec{w} \rangle &= \langle \vec{w}, \vec{v} \rangle && \text{(2.8c)}\\
\langle \alpha\vec{v}, \vec{w} \rangle &= \alpha\langle \vec{v}, \vec{w} \rangle && \text{(2.8d)}\\
\langle \vec{u} + \vec{v}, \vec{w} \rangle &= \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle && \text{(2.8e)}
\end{aligned}$$

Exercise II:

Show for the subcase of vectors in a three dimensional Cartesian coordinate system that these relations hold.

As mentioned, the notation $\langle \cdot, \cdot \rangle$ is more general and includes the subcase of vector spaces with finite dimensions (1-D, 2-D, etc) like the well known Cartesian coordinates. Writing a

inner product in the $\langle \cdot, \cdot \rangle$ form is also often used in the regard of the scalar product on an **infinite** vector space (the other subcase). There, the inner product is defined as

$$\langle v(x), w(x) \rangle = \int v(x) w(x) \, dx \; . \tag{2.9}$$

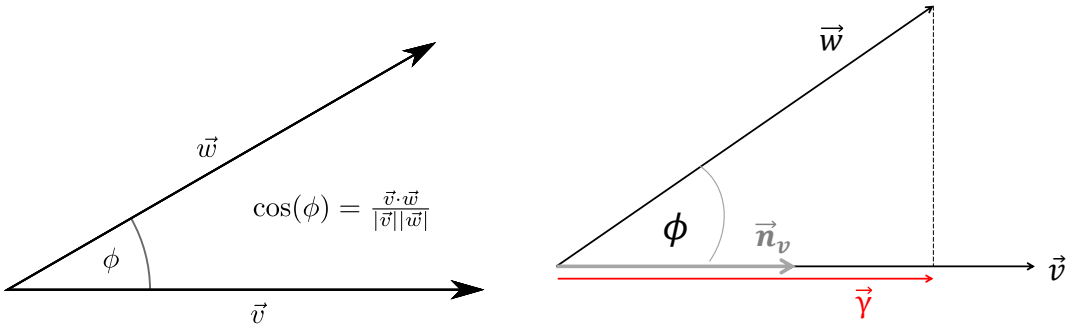

Figure 2: (a): Inner product between two vectors $\vec{v}$ and $\vec{w}$.
(b): The vector $\vec{\gamma}$ is the projection of the vector $\vec{w}$ onto $\vec{v}$ realized by the inner product between $\vec{w}$ and the normal vector of $\vec{v}$, $\vec{n}_v$.

Equation 2.8 are still valid in that case.
One may think about the scalar product given in Equation 2.9 as the case where we go from the discrete case (Equation 2.5) to the continuous, i.e. the summation in Equation 2.5 leads to an integration (that will be subject to Section 2.1.7)[3]. We will need the inner product in the form of Equation 2.9 for Fourier transformation (Section 2.4).

2.1.2 Matrices

Particularly in Section 4.3, Section 4.4.2, Section 8 and Section 9.3 we will need the concept of *matrices*. Suppose we had a set of m linear equations like

$$\begin{aligned} a_1 &= b_1 x_{1,1} + b_2 x_{1,2} + \cdots + b_n x_{1,n} \\ a_2 &= b_1 x_{2,1} + b_2 x_{2,2} + \cdots + b_n x_{2,n} \\ &\vdots \\ a_m &= b_1 x_{m,1} + b_2 x_{m,2} + \cdots + b_n x_{m,n} \end{aligned} \tag{2.10}$$

it would be more convenient to treat the left hand side of Equation 2.10 as a vector $\vec{a}$ of length m. On the right hand side of Equation 2.10 each line shows a resemblance to the inner product (c.f. Equation 2.5). One can treat the first line on the right hand side of Equation 2.10 as an inner product of vector $\vec{b}$ and vector $\vec{x}_1$, both having length n. The second line can be seen as an inner product of $\vec{b}$ and $\vec{x}_2$ and so on.
Therefore, it is convenient to define a matrix X with m rows and n columns that is a

[3] Discrete means that one can count something, like $1, 2, \ldots$ like the dimensions in Equation 2.5. But if one likes to sum up something that is continuous one needs an integral that is the limit of a summation (see Section 2.1.7).

mathematical object containing the values of $\vec{x}_1$ to $\vec{x}_m$. The rows of X are enumerated with $1, 2, \ldots, i, \ldots m$ and the columns with $1, 2, \ldots, j, \ldots n$. Therefore, X is called a $(m \times n)$ matrix "$(m$ *by* $n)$". An element x located in the i^{th} row and j^{th} column is then referred to as $x_{i,j}$. The enumeration of elements in a matrix according to their position wrt[4] the rows and columns is similar to the assignment of the location of an object in a coordinate system according to its coordinates. Using this notation Equation 2.15 can be written as

$$\vec{a} = X\ \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{2.11}$$

Like for vectors in the previous section, we can ask now for the properties of matrices and for rules to add or multiply them.

If we consider a vector as a matrix with one row and n columns (or m rows and one column) we can infer the addition of a matrix. Adding two matrices A and B is performed by just adding element wise:

$$X \pm Y = \begin{pmatrix} x_{1,1} \pm y_{1,1} & x_{1,2} \pm y_{1,2} & \cdots & x_{1,n} \pm y_{1,n} \\ x_{2,1} \pm y_{2,1} & x_{2,2} \pm y_{2,2} & \cdots & x_{2,n} \pm y_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} \pm y_{m,1} & x_{m,2} \pm y_{m,2} & \cdots & x_{m,n} \pm y_{m,n} \end{pmatrix} \tag{2.12}$$

and the multiplication by a scalar $\alpha \in \mathbb{R}$ is

$$\alpha X = \begin{pmatrix} \alpha x_{1,1} & \alpha x_{1,2} & \cdots & \alpha x_{1,n} \\ \alpha x_{2,1} & \alpha x_{2,2} & \cdots & \alpha x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha x_{m,1} & \alpha x_{m,2} & \cdots & \alpha x_{m,n} \end{pmatrix} . \tag{2.13}$$

The rule of multiplying a matrix X with another matrix Y, i. e. $Z = X \cdot Y$ is actually already given by definition if we compare Equation 2.10 to Equation 2.11. We see in Equation 2.11 *that a matrix* Y *can only be multiplied to a matrix* X *if the number of rows in* Y *equals the number of columns in* X, yielding a matrix Z with the same number of rows as X and the same number of columns as Y. Or in a more general way: if X is an n by m matrix (n rows and m columns) and Y is a m by p matrix then the resulting matrix $Z = X \cdot Y$ is an n by p matrix.
Following that, an element of the matrix Z is defined by

$$z_{i,j} = \sum_{k=1}^{m} x_{i,k}\, y_{k,j}\ , \tag{2.14}$$

i. e. the component $z_{i,j}$ of Z is given by the inner product of the i^{th} row vector of X with the k^{th} column vector of Y. For example $z_{1,2} = x_{1,1}\, y_{1,2} + x_{1,2}\, y_{2,2} + x_{1,2}\, y_{3,2} + \cdots + x_{1,m}\, y_{m,2}$.
In order to get familiar with Equation 2.14 we perform a few examples:

$$\begin{pmatrix} \boxed{1} & \boxed{3} \\ 2 & 4 \\ 0 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \boxed{-1} & -5 & 0 \\ \boxed{3} & 2 & -3 \end{pmatrix} = \begin{pmatrix} \boxed{8} & 1 & -9 \\ 10 & -2 & -12 \\ 3 & 2 & -3 \\ 10 & 11 & -9 \end{pmatrix} \tag{2.15}$$

[4] The abbreviation "wrt" stands for "with respect to"

where the first entry in the first row of the product is $(-1)1+3\cdot 3=8$ (framed numbers), the second entry in the first row is $(-5)1+2\cdot 3=1$ and so on. Hence, multiplying a $(4\times \mathbf{2})$ matrix to a $(\mathbf{2}\times 3)$ matrix yields a (4×3) matrix.
Note, that the multiplication between the two matrices would not be defined if we changed the order since the number of columns of the first matrix would not fit the number of rows of the second matrix. This property is called **non-commutative**. Thus, in general $X\cdot Y\neq Y\cdot X$! The next example illustrates the non-commutativity a bit further: suppose the multiplication of a 3×1 matrix (hence, a row vector) with a 1×3 matrix (a column vector)

$$\begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix}=\begin{pmatrix}\alpha a+\beta b+\gamma c\end{pmatrix}\ . \tag{2.16}$$

that is just the inner product of these two vectors.
However, changing the order would result into (c. f. Equation 2.14)

$$\begin{pmatrix}a\\ b\\ c\end{pmatrix}\begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix}=\begin{pmatrix}a\alpha & a\beta & a\gamma\\ b\alpha & b\beta & b\gamma\\ c\alpha & c\beta & c\gamma\end{pmatrix}\ . \tag{2.17}$$

that yields a 3×3 matrix.
To turn a row vector

$$\begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix} \tag{2.18}$$

into a column vector

$$\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix} \tag{2.19}$$

we define the operation *transposing* of a matrix X by

$$x^T_{i,j}=x_{j,i} \tag{2.20}$$

for each element $x_{i,j}$ that is intended by the T in Equation 2.20, hence the rows just change with the columns.
For two matrices $X,Y\in\mathbb{R}^{n\times n}$ and a scalar $\alpha\in\mathbb{R}$ following operations are valid:

$$\begin{aligned}(X+Y)^T &= X^T+Y^T && \text{(2.21a)}\\ (XY)^T &= Y^TX^T && \text{(2.21b)}\\ (\alpha X)^T &= \alpha X^T && \text{(2.21c)}\\ (X^T)^T &= X\ . && \text{(2.21d)}\end{aligned}$$

When performing operations like multiplication and addition with matrices we need rules for rearranging a matrix equation. For example we can ask for expressing the equation $X\cdot Y=Z$ in terms of X. For numbers we would just divide by Y on both sides on the equation. What we actually do is *multiplying* with the inverse of Y, Y^{-1}. As for numbers, the inverse is defined via the *neutral element* e. For numbers and multiplication the neutral element is just "1" since a number n times its inverse n^{-1} equals one and any number n multiplied with the neutral element yields the number itself; $n\times e=n$.
If we are dealing with matrix equations the neutral element of a matrix multiplication is not the number "1", but the *identity matrix* I. And, as required from the definition of a neutral

element and as the name implies, $X \cdot I = X$. Following this requirement, the identity matrix is given by

$$\mathrm{I} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} . \tag{2.22}$$

Given the identity matrix and a matrix X we can ask with which matrix we have to multiply X such that we get the identity matrix. This matrix is called the *inverse matrix* of X written as X^{-1}. Therefore we get

$$X = X \cdot I \quad \Leftrightarrow \quad X^{-1} \cdot X = I \text{ where } X^{-1} \text{ is the inverse element .} \tag{2.23}$$

Using this definition we are able to rearrange an equation like $XY = Z$ by multiplying it with the inverse matrix of Y (note, that any multiplication occurs only from the **right** since matrices are non commutative, as we know already). Thus we get $XYY^{-1} = XI = ZY^{-1}$ and therefore $X = Z\,Y^{-1}$.
As for transposing we can find some rules[5] for the inverse operation. For a matrix X and a scalar α they are

$$(X^{-1})^{-1} = X \tag{2.24a}$$
$$(\alpha X)^{-1} = \alpha^{-1} X^{-1} \tag{2.24b}$$
$$(X^T)^{-1} = (X^{-1})^T \tag{2.24c}$$
$$(XY)^{-1} = Y^{-1} X^{-1} \quad \text{for quadratic matrices } X, Y \tag{2.24d}$$

Another quantity of matrices which is often required is the determinant Δ of a matrix X (also often denoted as $det(X)$). The following mathematical operations might seem a bit arbitrary to you and I will not go into detail about the full mathematical justification of the determinant. The meaning of the determinant becomes clearer in Section 4 and when we discuss principle component analysis in the statistics primer. However, the determinant is a useful tool for solving a set of linear equations (that gave rise to the introduction of matrices, e.g. Equation 2.10), hence for solving matrix equations. Indeed, determinants are needed to e.g. calculate the inverse of a matrix explicitly. A determinant of a matrix can also be seen as the volume that is spanned by its column vectors. Therefore, determinants can only be calculated from square matrices[6]. For example, the determinant of a 2×2 matrix is given by

$$\Delta = \det(X) = \left| \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \right| = x_{11}\, x_{22} - x_{12}\, x_{21} \tag{2.25}$$

For the case of calculating the determinant of a 3×3 matrix there is another rule, the *rule of Sarrus*[7]. The general rule for deriving the determinant of an $n \times n$ matrix is more complicated and not required for our purpose.
One property of the determinant which we will need in the following is that if

$$\Delta = 0 \;\Leftrightarrow \textit{ there are column vectors which are linearly dependent.}$$

[5] I like to mention that all these rules that we take for granted here can be confirmed by executing the operations explicitly element by element. A motivated student might check this.

[6] Please try to understand how this is connected to the interpretation as volume.

[7] Pierre Frédéric Sarrus, 1798 – 1861

Meaning that if the determinant equals zero we can create a row of zeros in the matrix by rearranging the corresponding set of linear equations (matrices with $det(X) = 0$ have no defined inverse X^{-1}). This is helpful if we want to compute another important quantity of matrices namely the *eigenvalues* λ and *eigenvectors* $\vec{v}$ of a matrix, which are defined via the following equation

$$X \cdot \vec{v} = \lambda \, \vec{v} \quad \Leftrightarrow \quad (X - \lambda \, I) \, \vec{v} = 0 \, . \tag{2.26}$$

This equation has only a non-trivial ($\vec{v} \neq \vec{0}$) solution for $\vec{v}$ if

$$\det (X - \lambda \, I) = 0 \, , \tag{2.27}$$

in that case we are able to generate a row of zeros and have therefore the opportunity to chose one component of $\vec{v}$ free. Especially in the section about principle component analysis in the statistics primer the geometrical meaning of eigenvectors and eigenvalues becomes clear. In this script eigenvectors and eigenvalues are needed to investigate the stability of feedback loops and regulatory mechanisms in biological systems (see Section 4).
For a diagonal matrix of any size the determinant is given by

$$\Delta = \prod_i x_{ii} \, . \tag{2.28}$$

i. e. the product of the diagonal elements.
Furthermore,

$$\sum_i x_{ii} = \sum_i \lambda_i \, , \tag{2.29}$$

i. e. the sum of the eigenvalues is equal to the sum of the main diagonal elements of the matrix X. And

$$\Delta = \prod_i \lambda_i \tag{2.30}$$

i. e. the determinant is equal to the product of the eigenvalues.
There are way more matrix operations and properties (for example if a matrix is complex), but a full discussion would go beyond the scope of this script.

2.1.3 Derivatives of Functions With One Variable

Derivatives and integrals are the main topics in calculus and were elaborated independently from Newton and Leibniz[8] in the 17th century. Calculus is essential for any natural science. Not knowing derivatives or integrals in natural science is like not knowing letters or grammar in order to study literature or to write a poem. Moreover, it is extremely important to understand what derivatives and integrals actually *are* (especially for numerical reasons), rather than performing derivatives and integrals (since we have textbooks and mathematicians for tricky integrals). Thus, we will start from the very basic idea of derivatives and integrals in order to derive the rules of derivation and integration later on.

Suppose we want to know how a quantity changes with respect to another quantity. For example we want to know how the concentration of a reactant A in a chemical reaction changes with time t. Therefore the concentration can be measured at a time t_1 and at a

[8] Isaac Newton, 1642 – 1727 & Gottfried Wilhelm Leibniz, 1646 - 1716. Note, that ancient Greeks (Eudoxus of Cnidus, 408 – 355 BC & Archimedes of Syracuse, 287 – 212 BC) probably discovered calculus 2000 years earlier.

subsequent time point t_2. The change in the concentration $\bar{A}$ is determined by the ratio $\bar{A} = \frac{A(t_2)-A(t_1)}{t_2-t_1} \equiv \frac{\Delta A}{\Delta t}$.
The quantity $\bar{A}$ itself changes with time (i. e. is a function of time). Thus, the intervals Δt in which we measure ΔA have to be chosen carefully. It is clear that it is very unsuitable if Δt is too large because $\bar{A}$ would become too inaccurate. On the other hand, if Δt is very small, we have to consider many time steps that could be very inconvenient (for example for numerical reasons), although we could improve the accuracy of $\bar{A}$. How small must Δt be in order to satisfy our requirements? Maybe there are times when A does not change a lot (hence, we need a large Δt), or there are time intervals where A changes rapidly (Δt has to be very small in order to monitor $\bar{A}$ properly). Do we have to calculate $\bar{A}$ for each time interval Δt separately or is $\bar{A}$ a function that we could derive explicitly and therefore can give $\bar{A}$ for any t?
Let us consider any function f that depends on a quantity x, hence $f(x)$ (equivalent to $A(t)$). The goal is to find the change of $f(x)$, $f'(x)$, with respect to x as a function. We obtain the particular value $f(x_0)$ if $f(x)$ is measured at a given point x_0. Increasing x_0 by the interval Δx changes f accordingly to $f(x_0 + \Delta x)$ hence the change of f at x_0 can be estimated by

$$\frac{\Delta f\,(x)^+}{\Delta x} = \frac{f\,(x_0 + \Delta x) - f\,(x_0)}{\Delta x}. \tag{2.31}$$

Equation 2.31 is called *forward difference*, since the function of a variable x_0 is compared to the function at the position $x_0 + \Delta x$. In the same manner, we can compare $f(x_0)$ to $f(x_0 - \Delta x)$ which yields the *backward difference*

$$\frac{\Delta f\,(x)^-}{\Delta x} = \frac{f\,(x_0) - f\,(x_0 - \Delta x)}{\Delta x}\,. \tag{2.32}$$

Both differences are of equal validity and are identical for sufficiently small intervals Δx (what the term "sufficiently" means in this regard becomes clear in a few lines). The differences Δx and Δf are called *finite differences*. The idea of finite differences is illustrated in Figure 3.

In order to be more accurate we take the average of these two different quantities and obtain

$$\frac{1}{2}\left[\frac{\Delta f\,(x)^+}{\Delta x} + \frac{\Delta f\,(x)^-}{\Delta x}\right]_{x=x_0} = \left.\frac{\Delta f\,(x)}{\Delta x}\right|_{x=x_0} = \frac{f\,(x_0 + \Delta x) - f\,(x_0 - \Delta x)}{2\Delta x}\,. \tag{2.33}$$

Now it is clear that Equation 2.33 becomes more accurate if Δx decreases and we therefore obtain the slope of f at position $x = x_0$ (see also Figure 3). Thus, we have to find the limit for $\Delta x \to 0$. To distinguish the case when $\Delta x \to 0$ from the actual difference quotient $\frac{\Delta f}{\Delta x}$ we write $\frac{df}{dx}$ instead. Thus, we write

$$\left.\frac{d\,f(x)}{dx}\right|_{x=x_0} = \lim_{\Delta x \to 0} \left.\frac{\Delta f}{\Delta x}\right|_{x=x_0} = \lim_{\Delta x \to 0} \frac{f\,(x_0 + \Delta x) - f\,(x_0 - \Delta x)}{2\Delta x} \tag{2.34}$$

Equation 2.34 gives the definition of the first derivative $\frac{df}{dx}$ at the position $x = x_0$.
Often the notation $f'(x)$ is used instead of $\frac{df(x)}{dx}$ for the first derivative. However, physicists and mathematicians rather use the latter in order to refer to the origin from the difference quotient.
Equation 2.34 seems to be somehow technical, so let us perform two simple examples:

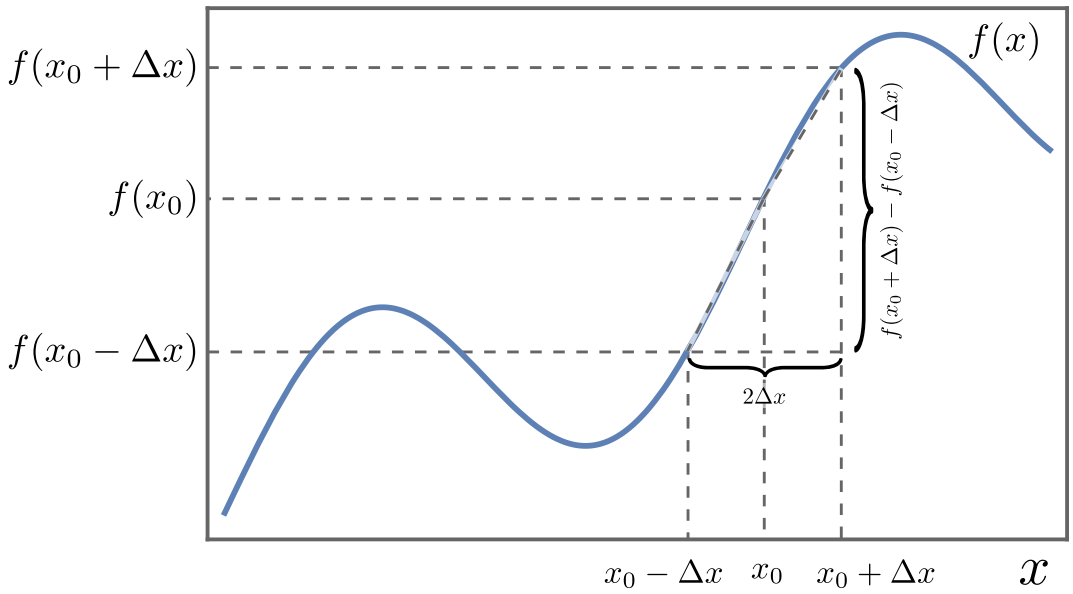

Figure 3: A small change of x, Δx, at the position $x = x_0$ causes a small change in $f(x)$. If Δx is infinitely small, the ratio $\Delta f(x)/\Delta x$ approaches the slope of the function, hence its tangent at position $x = x_0$.

Example I:

$f(x) = x^2$

$$\frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x} = \frac{(x_0 + \Delta x)^2 - (x_0 - \Delta x)^2}{2\Delta x} = 2x_0 \qquad (2.35)$$

Thus the first derivative, i. e. the function that describes how f changes wrt x, is given by $df/dx = 2x$.

Example II:

$\mathrm{f(x)} = \sqrt{\mathrm{x}}$

$$\begin{aligned}
\frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x} &= \frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0 - \Delta x}}{2\Delta x} \\
&= \frac{x + \Delta x - x + \Delta x}{2\Delta x\left(\sqrt{x_0 + \Delta x} + \sqrt{x_0 - \Delta x}\right)} \\
&= \frac{1}{\sqrt{x_0 + \Delta x} + \sqrt{x_0 - \Delta x}} \\
\Rightarrow \left.\frac{df}{dx}\right|_{x=x_0} &= \lim_{\Delta x \to 0} \frac{1}{\sqrt{x_0 + \Delta x} + \sqrt{x_0 - \Delta x}} = \frac{1}{2\sqrt{x_0}}
\end{aligned}$$

Thus, the first derivative of $\sqrt{x}$ wrt x is $1/(2\sqrt{x})$. Motivated students might think about the problems we encounter for $x_0 = 0$ (*hint: have a look at the backward difference*)

Of course it is inconvenient to perform the transformation to $\Delta x \to 0$ for every individual

function $f(x)$. Therefore, we rather use rules for a general treatment of $f(x)$, that can be derived from the previous considerations. These rules (for two differentiable functions $f(x), g(x)$ and $a, b \in \mathbb{R}$) are given by:

$$\frac{d}{dx}\left(af(x) + bg(x)\right) = a\ f'(x) + b\ g'(x) \tag{2.36a}$$

$$\frac{d}{dx}\left(f(x) \cdot g(x)\right) = f'(x)g(x) + f(x)g'(x) \tag{2.36b}$$

$$\frac{d}{dx}\left(f\left[g(x)\right]\right) = \frac{\partial f\left[g(x)\right]}{\partial g(x)}\frac{d\,g(x)}{dx} \tag{2.36c}$$

$$\frac{d}{dx}\left(f\left[g\left(x\right), x\right]\right) = \frac{\partial f\left[g\left(x\right), x\right]}{\partial g(x)}\frac{d\,g(x)}{dx} + \frac{\partial f[g(x), x]}{\partial x}\,. \tag{2.36d}$$

The "$\partial/\partial g$" in Equation 2.36c and Equation 2.36d denotes a *partial derivative* of g wrt x meaning that the derivative is only performed with respect to this particular quantity, here x. We will come back to partial derivatives in the next chapter where its intention becomes clear. There are some important derivatives which are worth knowing:

$$\frac{d}{dx}\mathrm{e}^x = \mathrm{e}^x \tag{2.37a}$$

$$\frac{d}{dx}\sin(x) = \cos(x) \tag{2.37b}$$

$$\frac{d}{dx}\cos(x) = -\sin(x) \tag{2.37c}$$

$$\frac{d}{dx}x^n = nx^{n-1} \tag{2.37d}$$

$$\frac{d}{dx}\ln(x) = \frac{1}{x}\,. \tag{2.37e}$$

Indeed, the exponential function and e are just defined by Equation 2.37a.

Exercise:

Write down the first derivatives of the following functions:

1. $$\left(1 + 2\sqrt{x}\right)^{3/2} \tag{2.38a}$$

2. $$x + x\sin(3x) \tag{2.38b}$$

3. $$x^3 + 2x\tan\left(x^3\right) + 1 \tag{2.38c}$$

4. $$\tan\left(\sin\left(\sqrt{x}\right)\right) \tag{2.38d}$$

5. $$e^{x^3} \tag{2.38e}$$

6.

$$(e^x)^3 \tag{2.38f}$$

7.

$$xe^{(x+2)^3} \tag{2.38g}$$

The second derivative can be derived by treating the first derivative as the *function* and following the procedure we used for deriving the first derivative since *the second derivative is the first derivative of the first derivative*. Therefore we compute the finite difference between the forward and backward derivative and divide it by Δx. This yields

$$\frac{1}{\Delta x}\left[\frac{f(x)^+}{\Delta x} - \frac{f(x)^-}{\Delta x}\right] = \frac{f(x_0+\Delta x) + f(x_0-\Delta x) - 2f(x_0)}{\Delta x^2}\,, \tag{2.39}$$

and finally we obtain

$$\left.\frac{d^2f(x)}{dx^2}\right|_{x=x_0} \equiv f''(x)|_{x=x_0} = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x) + f(x_0-\Delta x) - 2f(x_0)}{\Delta x^2}\,. \tag{2.40}$$

In principle the n^{th} derivative $f^n(x) \equiv \frac{d^n f(x)}{dx^n}$ can be obtained by repeating the procedure of finite differences n times, always treating the $(n-1)^{st}$ derivative as a new function from where we want to compute the first derivative. *This kind of approach will be essential when we will model diffusion reactions in Section 7.*

2.1.4 Derivatives of Functions with N Variables

At the point where the derivative of a function with one variable is understood it is only a small step to generalize the concept to functions of N variables. For our purposes it is often sufficient to treat the case $N = 4$ since we are dealing with models in three spatial and one temporal dimension (Section 4.3, Section 6 and Section 8).
An example is the concentration c of a molecule in a cell, which depends on the spatial coordinates x, y and z and the time t (the molecule can be depleted or generated by chemical reactions). Moreover, the spatial coordinates x, y and z can be time dependent, meaning that there is a flow of molecules within the cell and/or that sources and sinks of the molecule change their location with time. Therefore, the concentration is a function of time and the three spatial coordinates (that are functions of time themselves) and we have to write $c = c\,[x(t), y(t), z(t), t]$. How do we calculate the derivative of c wrt time now?
As stated in the previous section, we want to investigate the change of a quantity with respect to another quantity. If we kept y, z and t fixed and only x is free, then we would have $\frac{dc}{dx}$, or if we kept x, z and t fixed and only y is free, we would have $\frac{dc}{dy}$. To express that c does depend on other (now fixed) variables, we write $\frac{\partial c}{\partial y}$, where the operation ∂ is called the *partial derivative*. For example the partial derivative of c wrt time t is $\frac{\partial c}{\partial t}$ while keeping all other variables (x, y and z) fixed. If we are interested in the total change of the concentration wrt time we have to perform the *total derivative* $\frac{dc}{dt}$. Taken into account that x,y and z are functions of t itself, they can be treated like the outer derivative and (e.g. $\frac{\partial c}{\partial y}$) multiplied with the corresponding inner derivative $\frac{dy}{dt}$. This is just the chain rule and according to Equation 2.36d the total derivative of c is than given by

$$\frac{d}{dt}\left(c\,[x(t), y(t), z(t), t]\right) = \frac{\partial c}{\partial x}\frac{dx}{dt} + \frac{\partial c}{\partial y}\frac{dy}{dt} + \frac{\partial c}{\partial z}\frac{dz}{dt} + \frac{\partial c}{\partial t} \tag{2.41}$$

If we compare the first three addends in Equation 2.41 to the definition of the inner product (c.f. Equation 2.5) we see that we can rewrite them as

$$\frac{\partial c}{\partial x}\frac{dx}{dt} + \frac{\partial c}{\partial y}\frac{dy}{dt} + \frac{\partial c}{\partial z}\frac{dz}{dt} = \begin{pmatrix} \frac{\partial c}{\partial x} \\ \frac{\partial c}{\partial y} \\ \frac{\partial c}{\partial z} \end{pmatrix} \cdot \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix}, \tag{2.42}$$

where

$$\begin{pmatrix} \frac{\partial c}{\partial x} \\ \frac{\partial c}{\partial y} \\ \frac{\partial c}{\partial z} \end{pmatrix} =: grad\, c \tag{2.43}$$

is called the *gradient* ($grad$) of c (see more in Section 2.1.6). The gradient is just a short cut for the first derivatives of the spatial directions. Often, the gradient is written as ∇, i.e. $\nabla c \equiv grad\, c$. The gradient turns a scalar (here the concentration c) into a vector. Furthermore, we can see that

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \vec{v} \tag{2.44}$$

is just the flow velocity of the concentration. Thus, Equation 2.41 can be written as

$$\frac{d}{dt}\left(c\left[x(t), y(t), z(t), t\right]\right) = grad\, c \cdot \vec{v} + \frac{\partial c}{\partial t}. \tag{2.45}$$

The deeper meaning of Equation 2.45 will be revealed in Section 2.1.6 and Section 6.
A further example of this type of equation can be derived from Newton's law $\vec{F} = m\,\vec{a}$ that can be written as $\vec{F} = m\frac{d\vec{v}}{dt}$ or

$$\vec{F} = m\left(\frac{\partial \vec{v}}{\partial \vec{x}}\frac{d\vec{x}}{dt} + \frac{\partial \vec{v}}{\partial \vec{y}}\frac{d\vec{y}}{dt} + \frac{\partial \vec{v}}{\partial \vec{z}}\frac{d\vec{z}}{dt} + \frac{\partial \vec{v}}{\partial t}\right) \tag{2.46}$$

Again, we obtain the same structure as in Equation 2.41 and can use the gradient of the velocity (that is the only difference now since the velocity is a vector itself). Hence we obtain the form

$$\vec{F} = m\left(\vec{v}\,\nabla\vec{v} + \frac{\partial v}{\partial t}\right) \tag{2.47}$$

that is called *Navier – Stokes*[9] *– equation*. It will play an important role in micro fluidics (see Section 8).
As mentioned in the previous chapter it is often necessary to perform derivatives numerically when investigating and modeling biological systems. Therefore, we need to consider derivatives of functions of N variables in the same manner as we did for derivatives of functions with one variable. For the sake of better visualization let us discuss a function f that depends only on two variables x and y (Figure 4). Then $f(x, y)$ represents a three dimensional surface. If x is held constant (say, $x = x_0$), $f(x_0, y)$ equals a line on the surface of $f(x, y)$. The same applies if y is held constant ($y = y_0$) – we obtain the line $f(x, y_0)$ along the surface $f(x, y)$ in x direction. Therefore, we can define two first derivatives: $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ which are defined in a equivalent fashion as for the one dimensional case

$$\left.\frac{\partial f(x,y)}{\partial x}\right|_{x=x_0} = \lim_{\Delta x \to 0} \frac{f\left(x_0 + \Delta x, y\right) - f\left(x_0 - \Delta x, y\right)}{2\Delta x} \tag{2.48a}$$

[9] Claude Louis Marie Henri Navier, 1785 – 1836 and Sir George Gabriel Stokes, 1. Baronet PRS, 1819 – 1903

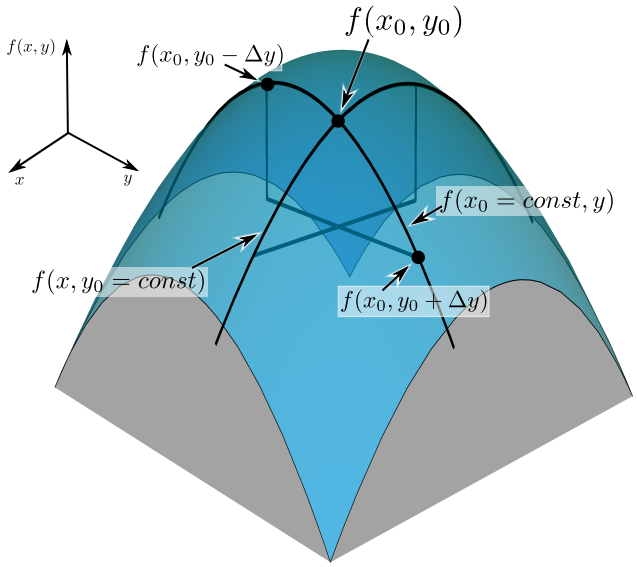

Figure 4: The derivatives of functions with more than one variable work similar to that shown in Figure 3. But now, since all other variables are held constant, the first derivatives are not a tangent, but a tangent plane with the slopes that correspond to the slopes of the point (x_0, y_0) etc. Thus, to find an extremum, the first derivatives wrt **all** variables must be zero.

$$\left.\frac{\partial f(x,y)}{\partial y}\right|_{y=y_0} = \lim_{\Delta y \to 0} \frac{f\left(x, y_0 + \Delta y\right) - f\left(x, y_0 - \Delta y\right)}{2\Delta y}\,. \tag{2.48b}$$

Generally, if a function depends on N variables we have N first derivatives. When searching for an extremum *all* first derivatives have to be set to zero. For $f(x,y)$ we obtain a tangent in x direction *and* y direction that spans a tangent plane. A point is located at the extremum if the tangent plane has no slope in any direction.
What happens with the second derivatives? In the case of $f(x,y)$ we have *four* second derivatives:

$$\left.\frac{\partial^2 f(x,y)}{\partial x^2}\right|_{x=x_0} = \lim_{\Delta x \to 0} \frac{f\left(x_0 + \Delta x, y\right) + f\left(x_0 - \Delta x, y\right) - 2f\left(x_0, y\right)}{\Delta x^2} \tag{2.49a}$$

$$\left.\frac{\partial^2 f(x,y)}{\partial y^2}\right|_{y=y_0} = \lim_{\Delta y \to 0} \frac{f\left(x, y_0 + \Delta y\right) + f\left(x, y_0 - \Delta y\right) - 2f\left(x, y_0\right)}{\Delta y^2} \tag{2.49b}$$

and the *mixed* second derivatives $\frac{\partial^2 f}{\partial x \partial y}$, $\frac{\partial^2 f}{\partial y \partial x}$

$$\begin{aligned}\left.\frac{\partial^2 f(x,y)}{\partial y \partial x}\right|_{x=x_0,y=y_0} &= \left.\frac{\partial^2 f(x,y)}{\partial x \partial y}\right|_{x=x_0,y=y_0} \\ &= \lim_{\Delta x, \Delta y \to 0} \left\{ \frac{f\left(x_0 + \Delta x, y_0 + \Delta y\right) - f\left(x_0 - \Delta x, y_0 + \Delta y\right)}{4\Delta y \Delta x} \right. \\ &\quad \left. - \frac{f\left(x_0 + \Delta x, y_0 - \Delta y\right) + f\left(x_0 - \Delta x, y_0 - \Delta y\right)}{4\Delta y \Delta x} \right\}.\end{aligned} \tag{2.50}$$

This reveals that the mixed derivatives are identical (changing x with y would not change Equation 2.50). The above equations can be derived in the same manner as the derivatives for functions with one variable (c.f. Equation 2.31 and Equation 2.32).

2.1.5 Lagrangian Multipliers

We know already that the extremum of a function can be calculated by setting the first derivative to zero. For example the function $z(x,y) = -(x-1)^2 - (y-2)^2 + 10$ has a (global) maximum at the point $x = 1$ and $y = 2$. In some cases, the task is that one has to find an *extremum that is subject to a constrain*. In our example, a given constrain might be the relation $g(x,y) = x - y = 0$. Thus, we need to find the maximum of z for which $x = y$ holds. The answer to this particular problem is that $x = 3/2$ and (since $x - y = 0$ holds) $y = 3/2$. There might be even a further constrain in another situation. Such a problem seems abstract, but many biological systems face such a situation. For example a population on a small island wants to have as numerous individuals as possible in order to increase diversity and the gene pool. However, since resources are limited, the size of an individual is constrained (i. e. after some generations this population has evolved to smaller individuals compared to their relatives on the main land), but on the other hand a larger size might save for being attacked by predators. Another related problem might be that locally, biological systems like to decrease entropy on molecular scale subject to the constrain that energy is conserved and finally in bioinformatics alignment algorithms or algorithms for motif finding like to optimize a particular objective function subject to the constrain that probability is conserved (see bioinformatics primer).
Although all these problems seem to be very different, they are mathematically identical and there is just one recipe that solves them all. Let us return to our function z and the corresponding constrain in order to derive this recipe: The function z is just a parabola in three dimensions that has the shape of a cone with the top at $x = 1$, $y = 2$ and $z = 10$. The constrain $x - y = 0$ is just the diagonal in the $x - y$ plane that cuts a piece out of the cone and the extremum subject to the constrain is the extreme (here the maximum) of the cutting edge. The situation is illustrated in Figure 5. The extrmum in z that is subject to the constrain $g(x,y) = x - y = 0$ equals the peak (that I like to denote as $\bar{z}$) of the red parabola (the cutting edge) in Figure 5.

The point at $\bar{z}$ has to satisfy two conditions: the total derivative of z

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \tag{2.51}$$

has to be zero and the total derivative of g has to be zero

$$dg = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy \tag{2.52}$$

since we aim on finding the extremum under the aforementioned conditions. I like to emphasize that the conditions $dz = 0$ and $dg = 0$ are different from requiring $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$ and $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y} = 0$ (that we used for finding e. g. the peak of the particular function). Demanding $dz = 0$ and $dg = 0$ means that we search for solutions along constant z or constant g, respectively, hence, a location along one particular level line. In such a case, the solution of our problem, $\bar{z}$, is located at this line of constant z, $dz = 0$, where the constrain g equals the tangent at the section of the cone defined by z along this particular level line (red dashed line in Figure 6). This (and only this) allows us to rearrange Equation 2.51 and Equation 2.52 and express both in terms of $\frac{dy}{dx}$ (we know that we can treat dx and dy like variables since they originated from finite differences) that leads to the relation

$$\frac{dy}{dx} = -\frac{\partial z/\partial x}{\partial z/\partial y} = -\frac{\partial g/\partial x}{\partial g/\partial y}\,. \tag{2.53}$$

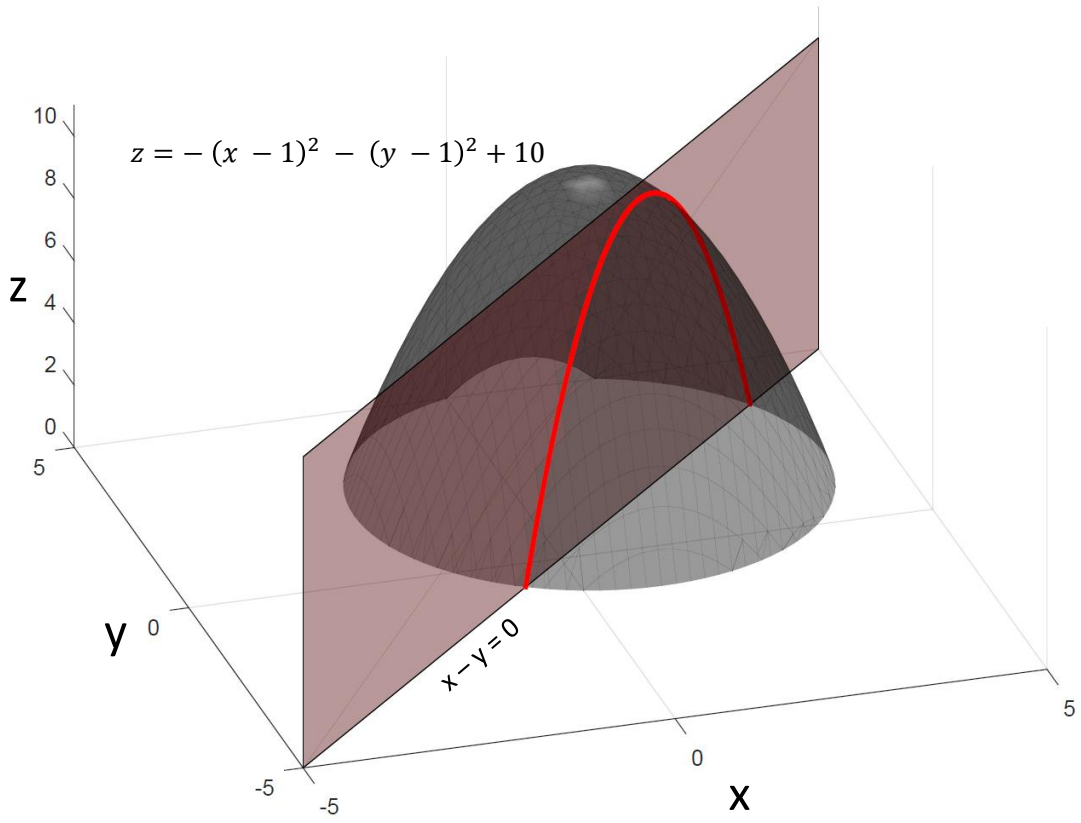

Figure 5: The function $z(x,y) = -(x-1)^2 - (y-2)^2 + 10$ (dark cone) with the constrain $g(x,y) = x - y = 0$ (dark red plane). The intersection of these two functions equals the red parabola that peaks at $z = \bar{z}$ that is the maximum of z, subject to the constrain g.

This relation only holds, since $\frac{dy}{dx}$ is the tangent of the level line at the extreme.
Equation 2.53 is just a ratio and multiplying both sides with a constant, say λ, does not change the mathematical meaning. Thus, we obtain

$$\frac{\partial z}{\partial x} = \lambda \frac{\partial g}{\partial x} \tag{2.54a}$$

$$\frac{\partial z}{\partial y} = \lambda \frac{\partial g}{\partial y} \tag{2.54b}$$

where the constants λ are called *Lagrangian multipliers*[10].

We now have the recipe we were aiming on. The first relation in Equation 2.54 yields $x = -\frac{\lambda}{2} + 1$ and from the second equation we obtain $y = \frac{\lambda}{2} + 1$. Since the constrain $x - y = 0$ holds, we obtain the solution $x = y = 3/2$ that gives us $\bar{z} = 9.5$ from $z(x,y)$ (see also Figure 6).
Note, that this particular problem could have been solved without using Lagrangian multipliers, but the idea is a very general concept. For example in N dimensions (like almost every problem in thermodynamics and related tasks in bioinformatics) we may have a function $f(x_1, x_2, \ldots, x_N)$ with the constrain $g(x_1, x_2, \ldots, x_N)$ that is not possible to visualize anymore. Fortunately, we know now how to derive the solution when applying the concept

[10] Joseph-Louis de Lagrange, 1736 – 1813

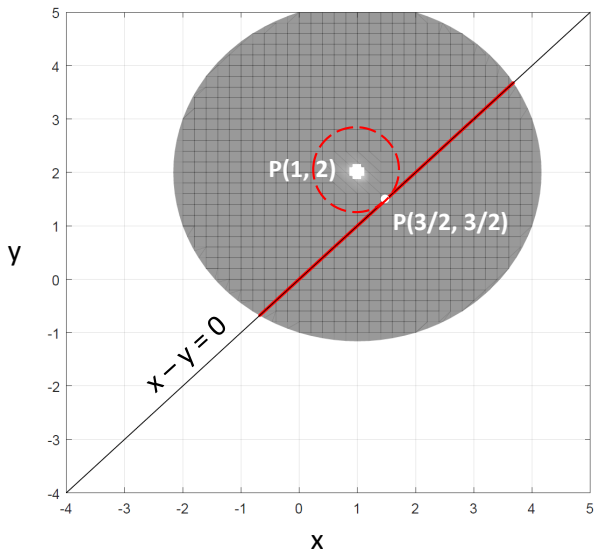

Figure 6: View from the top of the problem that is visualized in Figure 5. The solution of this problem is graphically indicated by the red dashed circle that corresponds to the level line $dz = 0$ (Equation 2.51), where the constrain $g(x, y) = x - y = 0$ equals the tangent.

of Lagrangian multipliers:

$$\frac{\partial f}{\partial x_1} = \lambda \frac{\partial g}{\partial x_1} \tag{2.55a}$$

$$\frac{\partial f}{\partial x_2} = \lambda \frac{\partial g}{\partial x_2} \tag{2.55b}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \lambda \frac{\partial g}{\partial x_n}. \tag{2.55c}$$

For many constrains $g_1, g_2, \ldots g_M$, one can generalize the concept even further to

$$\frac{\partial f}{\partial x_1} = \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} + \cdots + \lambda_M \frac{\partial g_M}{\partial x_1} \tag{2.56a}$$

$$\frac{\partial f}{\partial x_2} = \lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2} + \cdots + \lambda_M \frac{\partial g_M}{\partial x_2} \tag{2.56b}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \lambda_1 \frac{\partial g_1}{\partial x_n} + \lambda_2 \frac{\partial g_2}{\partial x_3} + \cdots + \lambda_M \frac{\partial g_M}{\partial x_3}. \tag{2.56c}$$

In principle one can have as many constrains as free variables ($M = N$), but not more since then the problem is not solvable in principle. We will utilize the concept of Lagrangian multipliers in Section 3.

2.1.6 Gradient and Divergence

With Equation 2.43 we introduced the operation *gradient* that is a vector containing the first derivatives in every spatial direction. For example the gradient of the concentration $c(x, y, z)$, that is a scalar, yields the vector $grad\, c = (\frac{\partial c}{\partial x}, \frac{\partial c}{\partial y}, \frac{\partial c}{\partial z}) = \vec{f}$ that is also frequently written as $grad\, c = \nabla\, c$. The symbol "∇" is called *"nabla"* and therefore the gradient is

often called *nabla operator*.
The gradient $\vec{f}$ of a scalar has then the form

$$\vec{f} = \begin{pmatrix} f_x(x,y,z) \\ f_y(x,y,z) \\ f_z(x,y,z) \end{pmatrix} \tag{2.57}$$

or in general (for n dimensions)

$$\vec{f} = \begin{pmatrix} f_{x_1}(x_1,x_2,\ldots,x_n) \\ f_{x_2}(x_1,x_2,\ldots,x_n) \\ \cdots \\ f_{x_n}(x_1,x_2,\ldots,x_n) \end{pmatrix} . \tag{2.58}$$

One can show that the gradient of a scalar function (for example $c(x,y,z)$) at a given point $\vec{r}_0 = (x_0, y_0, z_0)$ points in the direction of the greatest rate of increase at the point $\vec{r}_0$ and the magnitude is the slope of the function in that point. Thus, the gradient is perpendicular to the level lines i.e. the lines where the function is constant. In this regard, gradients are used for optimization or fitting. A useful and very common application for the gradient are for example optimization and fitting algorithms (*gradient descent*).
Let us briefly investigate the scalar function $F(x,y) = x^2 + y^2$, that looks a bit like a cone (c. f. Figure 5). For a given set of x and y, $F(x,y)$ yields the corresponding z value. A line at constant level means $F(x,y) = const$, i.e. $const = x^2 + y^2$ that is the equation of a circle. Thus, when looking from above onto the top of the cone, the level lines would look like circles with radius $\sqrt{x^2+y^2}$. According to Equation 2.57, the gradient of F, $grad\, F = \vec{f}$, is then

$$\vec{f} = \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix} \tag{2.59}$$

that is illustrated in Figure 7.

Another useful concept is the divergence *div* of a vector field $\vec{f}$. I like to introduce this concept by investigating a flux Ψ though a volume element $\Delta V = \Delta x \Delta y \Delta z$. A volume element is the three dimensional equivalent of the finite differences we had in Section 2.1.3, hence an infinitely small cube. The quantity flux is defined as "something" per time, for example number of particles n per time and the flux density ϕ is defined as flux per area.
Let us consider the flux density into the volume element in y direction, hence $\phi_y(y)$. The total flux into the volume equals then the flux density times the area of the penetrated surface element $\Delta x \Delta z$, hence $\Psi_y(y) = \phi_y(y) \Delta x \Delta z$. The flux *out of* the volume box (of length Δy) is then $\Psi_y(y + \Delta y)$. If the flux *into* the volume $\Psi_y(y)$ is less then the flux *out* of the volume $\Psi_y(y + \Delta y)$, the number of particles n within the volume *decreases* with time, or

$$\frac{\Delta n(t)}{\Delta t} = -\left[\phi_y(y + \Delta y) - \phi_y(y)\right] \Delta x \Delta z. \tag{2.60}$$

It is important to understand the appearance of the minus in Equation 2.60 since the number of particles is decreasing if the net flux out of the volume is positive.
Since the flux is a vector that could point into any direction it is useful to decompose it into the contributions parallel to the coordinate axes. Hence, like we investigated the flux Ψ_y, we can include the flux into the volume increment ΔV from any direction (Figure 8) and therefore obtain the three dimensional equivalent of Equation 2.60

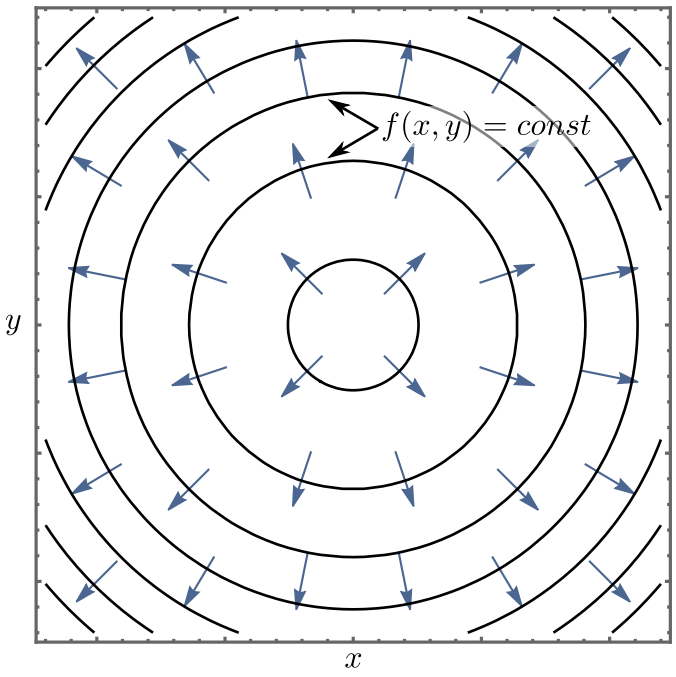

Figure 7: The level lines of the function $F(x,y) = x^2+y^2$ (solid black circles) and the corresponding vector field $grad\, F(x,y) = (2x,\, 2y)$ (blue arrows), which is perpendicular to the level lines $F(x,y) = const$.

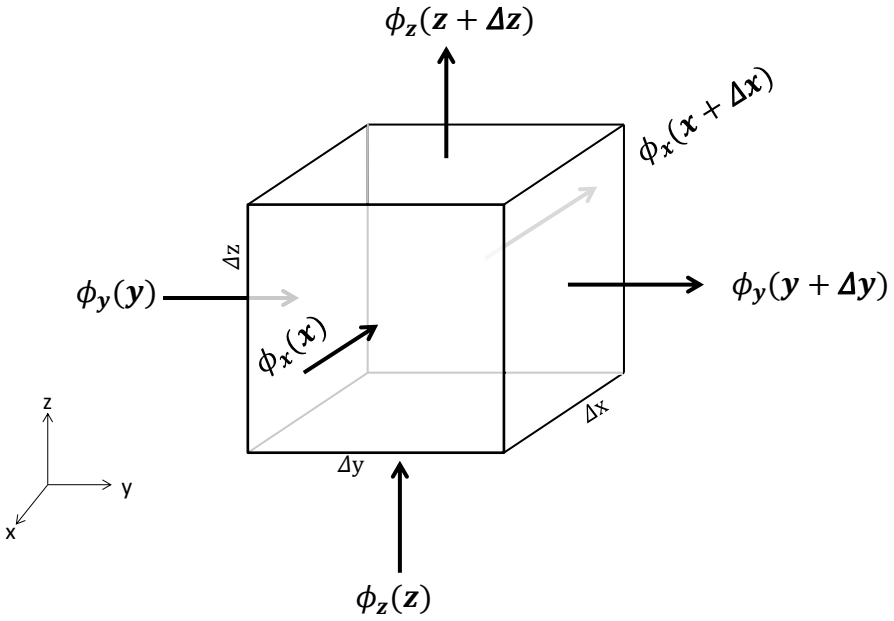

Figure 8: The flux of particles into and out of the volume element causes the change of the number of particles within ΔV.

$$\begin{aligned}\frac{\Delta n}{\Delta t} = -\{&[\phi_x(x+\Delta x,y,z,t) - \phi_x(x,y,z,t)]\,\Delta y\Delta z \\ &+ [\phi_y(x,y+\Delta y,z,t) - \phi_y(x,y,z,t)]\,\Delta x\Delta z \\ &+ [\phi_z(x,y,z+\Delta z,t) - \phi_z(x,y,z,t)]\,\Delta x\Delta y\}.\end{aligned} \tag{2.61}$$

The number of particles divided by volume equals the concentration $c(x,y,z,t)$, thus $n(x,y,z,t) = c(x,y,z,t)\Delta V$ and we therefore can rewrite Equation 2.61 by dividing by the volume increment $\Delta V = \Delta x \Delta y \Delta z$ that leads to

$$\begin{aligned}\frac{\Delta c}{\Delta t} = -\Bigg[&\frac{\phi_x(x+\Delta x,y,z,t)-\phi_x(x,y,z,t)}{\Delta x}\\ &+\frac{\phi_y(x,y+\Delta y,z,t)-\phi_y(x,y,z,t)}{\Delta y}\\ &+\frac{\phi_z(x,y,z+\Delta z,t)-\phi_z(x,y,z,t)}{\Delta z}\Bigg]\,.\end{aligned} \tag{2.62}$$

Equation 2.62 exhibits the same mathematical structure as Equation 2.34 and we therefore know already that we can take the limit $(\Delta x, \Delta y, \Delta z, \Delta t) \to 0$ which results in

$$\frac{\partial c}{\partial t} = -\left(\frac{\partial \phi_x}{\partial x} + \frac{\partial \phi_y}{\partial y} + \frac{\partial \phi_z}{\partial z}\right)\,. \tag{2.63}$$

The right hand side of Equation 2.63 reminds us on the gradient in Equation 2.43. However, in contrast to the gradient which generates a vector from a scalar function, Equation 2.63 yields a scalar (change of concentration wrt time) generated from a vector (the flux density). This structure is called *divergence*, written as $div\,\vec{\phi}$. This operation can be viewed as the scalar product (or dot product) between the operation"nabla" (i.e. ∇) and the vector field $\vec{\phi}$:

$$div\,\vec{\phi} = \nabla\cdot\vec{\phi} = \begin{pmatrix}\frac{\partial}{\partial x}\\ \frac{\partial}{\partial y}\\ \frac{\partial}{\partial z}\end{pmatrix}\cdot\begin{pmatrix}\phi_x\\ \phi_y\\ \phi_z\end{pmatrix} = \frac{\partial\phi_x}{\partial x}+\frac{\partial\phi_y}{\partial y}+\frac{\partial\phi_z}{\partial z}. \tag{2.64}$$

So that Equation 2.63 can be written as

$$\boxed{\frac{\partial c}{\partial t} = -div\,\vec{\phi}}\,, \tag{2.65}$$

which is called *continuity equation*. This equation is very important and we will come back to it in Section 6 and Section 8.
Comparing Equation 2.65 to Equation 2.45 we find that we can combine both to

$$\frac{dc}{dt} = \vec{v}\,grad\,c - div\,\vec{\phi}\,. \tag{2.66}$$

The unit of ϕ is a number per time per area and the product of velocity with concentration has the same unit (see both addends in Equation 2.66) that leads to the conclusion that the flux density can be written as $\vec{\phi} = c\,\vec{v}$ so that $div\ \vec{\phi} = div\ \ (c\,\vec{v})$. One can show by explicit writing (product rule) that $div\ \ (c\,\vec{v}) = (grad\ c)\,\vec{v} + (div\ \vec{v})\,c$ and thus we can infer that

$$\frac{\partial c}{\partial t} = -(grad\ c)\ \vec{v} - (div\ \vec{v})\ c \ \Leftrightarrow \frac{dc}{dt} = -c\ div\ \vec{v}\,. \tag{2.67}$$

Equation 2.67 appears frequently in physics e.g. in electrodynamics, we will encounter it when dealing with diffusion (Section 6) and pattern formation (morphogenisis, Section 7.2) or in hydrodynamics (hence, microfluidics as a biological application, see Section 8).
Equation 2.65 is called homogeneous continuity equation meaning that there are neither sinks nor sources (mass and/or number of particles is conserved). More generally the continuity equation can be written with an additional term σ, if e.g. n is not conserved i.e.

σ counts for the generation or consumption of n per unit volume and time. In this case Equation 2.65 is called non-homogeneous and reads

$$\boxed{\frac{\partial c}{\partial t} + div\ \vec{\phi} = \sigma}. \tag{2.68}$$

2.1.7 Integrals

Suppose the flux Ψ in Section 2.1.6 through a membrane is measured and that Ψ is a function of time $\Psi = \Psi(t)$. For some reason, we want to know the total flow F through the membrane between the time point a and the time point b, i.e. within a time span $b-a$. If $\Psi(t)$ would be constant, we would multiply it with the time interval in order to get the flow: $F = \Psi \cdot (b-a)$. However, in most of the cases, $\Psi(t)$ is not constant wrt time and such a simple attempt would not help.
There are numerous problems of the same class and therefore let us switch to the notation that is commonly used in text books. Let us denote the function of intrest (here the flux) as f and the free variable time t as x. If $f(x)$ is not constant wrt x we might separate the interval $b-a$ into N subintervals Δx, where

$$N = \frac{b-a}{\Delta x}. \tag{2.69}$$

For the first subinterval from a to $a+\Delta x$ we would get the flow $f(a)\Delta x$, for the second interval we would obtain $f\left(a+\Delta x\right)\Delta x$ and so on. Finally, we would add up all the flows in the different intervals and therefore obtain an approximation for the total flow F. This procedure corresponds to calculating the area under the curve $f = f(x)$, defined by the interval $b-a$ and the x axis (Figure 9).
If the size of Δx is decreasing, we divide the surface under the curve $f(x)$ into thinner bars

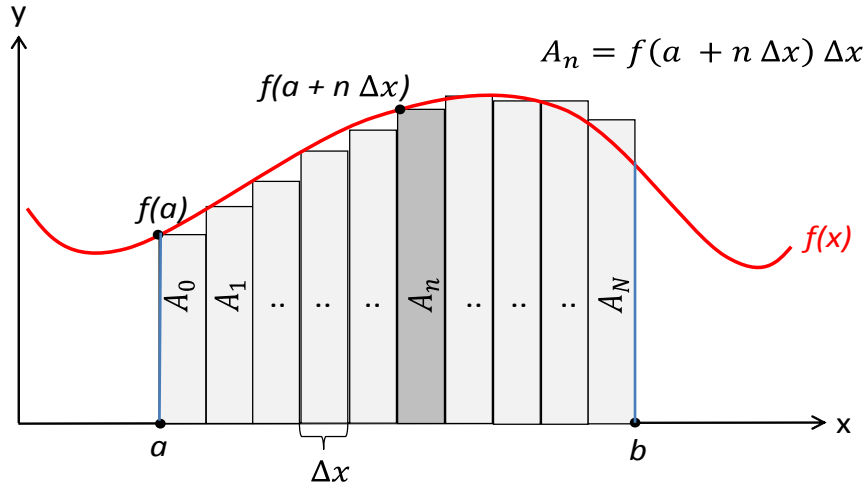

Figure 9: The integral of a function f within the interval $[a;b]$ equals the area under the curve $f(x)$ between the limits $x=a$ and $x=b$ and the axis of the free variable (here x).

and therefore obtain a more accurate approximation for F. Let us denote the area of the first bar as $A_0 = f(a)\Delta x$. The second measurement starts at $x = a+\Delta x$ and ends at $x = a+2\Delta x$ so that the second bar has the area $A_1 = f(a+\Delta x)\Delta x$. Therefore the total flux

within the time span $b-a$ is approximately given by $f(a)\Delta x+f(a+\Delta x)\Delta x+\cdots+f(b)\Delta x$. The n^{th} measurement is labeled as $A_n = f(a+n\Delta x)\Delta x$ and we get a sequence of measurements given by

$$\begin{aligned} A_0 &= f(a)\cdot\Delta x \\ A_1 &= f\left(a+\Delta x\right)\cdot\Delta x \\ A_2 &= f\left(a+2\Delta x\right)\cdot\Delta x \\ &\vdots \\ A_n &= f\left(a+n\Delta x\right)\cdot\Delta x \\ &\vdots \\ A_{N-1} &= f\left(a+(N-1)\,\Delta x\right)\cdot\Delta x\,. \end{aligned} \tag{2.70}$$

Note, that the last bar is denoted as A_{N-1}, since $A_N = f(b)\cdot\Delta x$ would reach *outside* the interval $[a;b]$.
The sum of the areas of all bars yields an approximation of the total flow within the time span $b-a$:

$$A_{\rm tot} \approx \sum_{n=0}^{N-1} A_n \approx \sum_{n=0}^{N-1} f\left(a+n\Delta x\right)\ \Delta x. \tag{2.71}$$

As discussed, the limit $\Delta x \to 0$ leads to the exact result by

$$A_{\rm tot} = \lim_{\Delta x\to 0}\sum_{n=0}^{N-1} f\left(a+n\Delta x\right)\Delta x := \int_a^b f(x)\ dx\ . \tag{2.72}$$

that is the definition of the *integral* $\int_a^b f(x)dx$. Hence, *an integral is nothing but the limit of a sum.*

Let us examine two examples so that the idea becomes clear:

Example I:

$f(x) = x$

$$\begin{aligned} A_{tot} &= \lim_{\Delta x\to 0}\sum_{n=0}^{N-1} f\left(a+n\Delta x\right)\Delta x \\ &= \lim_{\Delta x\to 0}\sum_{n=0}^{N-1}\underbrace{\left(a+n\Delta x\right)}_{f(x)=x}\Delta x = \lim_{\Delta x\to 0}\sum_{n=0}^{N-1}\left(a\Delta x\ +n\Delta x^2\right)\ . \end{aligned} \tag{2.73}$$

Using $\sum_{n=0}^{N-1} n = \frac{N(N-1)}{2}$, Equation 2.73 turns into

$$\begin{aligned} A_{tot} &= \lim_{\Delta x\to 0}\sum_{n=0}^{N-1}\left(a\Delta x\ +n\Delta x^2\right) \\ &= \lim_{\Delta x\to 0}\left(N\ a\ \Delta x + \frac{N^2}{2}\Delta x^2 - \frac{N}{2}\Delta x^2\right)\ . \end{aligned} \tag{2.74}$$

Inserting N from Equation 2.69 into Equation 2.74 yields

$$\begin{aligned} A_{tot} &= \lim_{\Delta x \to 0} \left(N\ a\ \Delta x + \frac{N^2}{2}\Delta x^2 - \frac{N}{2}\Delta x^2 \right) \\ &= \lim_{\Delta x \to 0} \left[a(b-a) + \frac{1}{2}(b-a)^2 - \frac{1}{2}(b-a)\Delta x \right] = \frac{1}{2}(b^2 - a^2)\ . \end{aligned} \quad (2.75)$$

Hence,

$$\int_a^b x\ dx = \frac{1}{2}(b^2 - a^2). \quad (2.76)$$

More general, the integral of $f(x)$ can be written as

$$\int x\ dx = \frac{1}{2}x^2 + C\ . \quad (2.77)$$

Note, that calculating the integral of a function is the reverse process of calculating the derivative of a function. Therefore, the general form of the integral of a function (like Equation 2.77) is called *anti-derivative*.
Since the derivative of a constant vanishes, for example $\frac{d}{dx}\left(x^2 + C\right) = 2\,x \quad \forall\, C = const$, we obtain always the same result for any $C = const$. Therefore, one has to add the constant C to the anti-derivative $\int 2x\,dx = x^2 + C$ for the reverse process. When calculating the integral, the constant C is unknown and thus, there is an infinite number of integrals for each function, unless C is specified. Often, if the free variable x is time, C sets the initial conditions of a process (e.g. number of individuals at $t = 0$ for population growth).

Example II:

$f(x) = x^2$

$$\begin{aligned} A_{tot} &= \lim_{\Delta x \to 0} \sum_{n=0}^{N-1} f\left(a + n\Delta x\right) \Delta x = \lim_{\Delta x \to 0} \sum_{n=0}^{N-1} \underbrace{(a + n\Delta x)^2}_{f(x)=x^2} \Delta x \\ &= \lim_{\Delta x \to 0} \sum_{n=0}^{N-1} \left(a^2\Delta x\ + 2a\ n\Delta x^2 + n^2\Delta x^3 \right) \quad (2.78) \\ &= \lim_{\Delta x \to 0} \left[N\ a^2\Delta x + a\ N(N-1)\Delta x^2 + \frac{N(N-1)(2N-1)}{6} \right] \Delta x^3 \\ &= \frac{1}{3}(b^3 - a^3)\ . \end{aligned}$$

Hence,

$$\int_a^b x^2\ dx = \frac{1}{3}(b^3 - a^3), \quad (2.79)$$

or more general

$$\int x^2\ dx = \frac{1}{3}x^3 + C. \quad (2.80)$$

Like for derivatives (Section 2.1.3) there are rules to perform certain types of integrals

$$\int a\ f(x) + b\ g(x)\ \ dx = a \int f(x)\ dx + b \int g(x)\ dx \tag{2.81a}$$

$$f(x)\ g(x) = \int f'(x) g(x)\ dx + \int f(x) g'(x)\ dx\ . \tag{2.81b}$$

Moreover there are some important integrals

$$\int e^x\ dx = e^x + C \tag{2.82a}$$

$$\int \sin(x)\ dx = -\cos(x) + C \tag{2.82b}$$

$$\int \cos(x)\ dx = \sin(x) + C \tag{2.82c}$$

$$\int C_1\ x^n\ dx = \frac{C_1}{n+1} x^{n+1} + C_2\ \ \forall n \neq -1 \tag{2.82d}$$

$$\int \frac{1}{x}\ dx = \ln(x) + C\ . \tag{2.82e}$$

Exercise:

Write down the integrals of the following functions:

1. $$a^x \tag{2.83a}$$
2. $$\sinh\ (x) \tag{2.83b}$$
3. $$\frac{1}{\cos^2(x)} \tag{2.83c}$$
4. $$\frac{1}{\sqrt{1-x^2}} \tag{2.83d}$$
5. $$\frac{1}{\sqrt{1+x^2}} \tag{2.83e}$$
6. *For our motivated students*
$$cot(x) \tag{2.83f}$$
7. *For our motivated students*
$$arctan(x) \tag{2.83g}$$

2.1.8 Potential and Exact Differentials

A further important quantity is the *potential* Ψ (not to change with the notation for the flux in Section 2.1.6). For life scientists, the most important potentials appear in thermodynamics, since thermodynamical potentials are the driving forces for chemical reactions and

thus for life itself (see Section 3 or Section 4.3). But what is behind the notation *potential* and what does it mean?
First of all, a potential Ψ is an abstract (but very fruitful) mathematical concept. Let us consider the scalar function Ψ that depends on the spatial directions, so that $\Psi = \Psi(x, y, z)$. We can calculate the spatial derivatives using the chain rule (Section 2.1.4) and obtain

$$d\Psi = \frac{\partial \Psi}{\partial x} dx + \frac{\partial \Psi}{\partial y} dy + \frac{\partial \Psi}{\partial z} dz \ , \tag{2.84}$$

and use the definition of the dot product (Equation 2.5) to rewrite this equation to

$$d\Psi = \begin{pmatrix} \frac{\partial \psi}{\partial x} \\ \frac{\partial \psi}{\partial y} \\ \frac{\partial \psi}{\partial z} \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \ . \tag{2.85}$$

The first parenthesis in Equation 2.85 is a vector derived from the spatial derivatives of a scalar. Such a structure is known as gradient $grad\,\Psi$ (Section 2.1.6). The second parenthesis denote also a vector. This vector contains the finite differences in all three spatial directions and we can summarize it to the vector

$$d\vec{s} = \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \tag{2.86}$$

so that Equation 2.84 reads $d\Psi = grad\,\Psi\ \vec{ds}$.
In nature, the quantity Ψ is identified as a scalar field (since it has no direction) called *potential* and the gradient of the potential

$$\boxed{\vec{f} = -grad\ \Psi} \tag{2.87}$$

is identified via a proportionality constant q as its force $\vec{f}_{\times} = q\,\vec{f} = q\,grad\ \Psi$. The proportionality constant is a "charge", like the electrical charge or mass. The potential *energy* in such a system is a product of Ψ times the charge $E_{\times} = q\,\Psi$. I like to give two examples you probably know from high school.
The potential gravitational energy E_G between two point masses m_1 and m_2 of separation r is

$$E_G \sim \frac{m_1 m_2}{r} \tag{2.88}$$

where it turns out that the proportionality constant is Newtons constant G. In contrast to the charge as proportionality constant, the constant G appears here, since we measure the physical quantities in meters, kilograms and seconds. Indeed, theoreticians set them to one since it does not change the overall physics. Thus, the gravitational potential equals

$$\Psi = G\frac{m_2}{r} \tag{2.89}$$

and the potential energy can be written as $E_G = m_1\Psi$, where the mass m_1 acts as "charge". The gradient is then

$$grad\,\Psi = grad\left(G\frac{m_2}{r}\right) = G\,m_2\ grad\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) = -G\,m_2\frac{1}{r^3}\vec{r}. \tag{2.90}$$

We used the product rule in the last step of Equation 2.90[11]. Thus, according to the definition in Equation 2.87, the gravitational force equals

$$\vec{f}_G = -G\,\frac{m_1 m_2}{r^3}\vec{r}. \tag{2.91}$$

Equation 2.91 should look familiar to you.
The same approach works with the electrical potential Ψ_E between two charges q_1 and q_2 (see also Section 9.4) that yields the electrical energy E_E

$$E_E \sim \frac{q_1 q_2}{r} \sim q_1 \Psi_E \tag{2.92}$$

where it turns out that the proportionality constant (again, required by the unit system) is $\frac{1}{4\pi\mu_0\mu_r\epsilon_0\epsilon_r}$ (Section 9.4). The electrical potential Ψ_E equals the voltage U so that the potential electrical energy equals $E_E = q_1\,U = q_1\,\Psi_E$.
The forces, hence the gradients of a potential, determines the equation of motion (e.g. $f_G = m_1\,a$ for gravity) that describes the system. However, since a force is derived from a potential by performing a derivative, there is some freedom of how to choose the potential Ψ. For example, we can add a constant C to the potential ($\Psi + C = \bar{\Psi}$) and will derive the same force (i. e. the same behavior of the system). For example in a one dimensional system (in order to keep it simple), we find that

$$\frac{\partial\bar{\Psi}(x)}{\partial x} = \frac{\partial}{\partial x}\,[\Psi(x) + C] = \frac{\partial\Psi(x)}{\partial x} = f(x). \tag{2.93}$$

We have some freedom for choosing C that physicist call *gauge freedom* of a system.
Therefore, a potential itself is not uniquely defined and we are only able to measure a *potential difference* $\Delta\Psi$. For example the voltage U is actually a difference ΔU. Another prominent example for a potential in nature is, as mentioned earlier, the gravitational potential and its gradient the gravitational force, we all feel when e. g. climbing a mountain. Climbing a mountain costs energy since we have to work against the potential. The difference of the potential between the bottom and the top $\Delta\Psi = \Psi(top) - \Psi(bottom)$ of the mountain is proportional to the energy required to climb the mountain, where the "charge" equals the mass of the mountaineer m_1 (c. f. $E_G = m_1\Delta\Psi$). On the other hand we gain (kinetic) energy when we are jumping off a cliff i. e. the potential energy we gained by climbing the mountain is transferred into kinetic energy.

Let us return to Equation 2.84, the derivative of Ψ. The reverse operation of Equation 2.84 is integrating $d\Psi$ along a path with the start point A and end point B

$$\int_A^B d\Psi = \int \vec{f}\,d\vec{s}_{A\to B} = \Psi(B) - \Psi(A)\ . \tag{2.94}$$

We are integrating along a path because $d\vec{s}$ is a vector containing the finite differences in all three spatial directions. Equation 2.94 might look trivial, but it tells us that the result of the integral only depends on the start point A and end point B, but not on the path itself (see the examples in this section later on). That is the reason why e. g. climbing a mountain always costs the same amount of energy, no matter which path is chosen, as long as start point (say base camp) and end point (summit) do not change. Equation 2.94 also implies that if start point and end point are identical, hence, when integrating along

[11] Please try to verify the last step as exercise!

a *closed path*, the integral vanishes. To indicate an integration along a closed path, we introduce the integral sign $\oint$ and write

$$\oint d\Psi = \oint \vec{f}\, d\vec{s} = 0\,. \tag{2.95}$$

Potential and force are connected via Equation 2.87 and one might wonder if a force

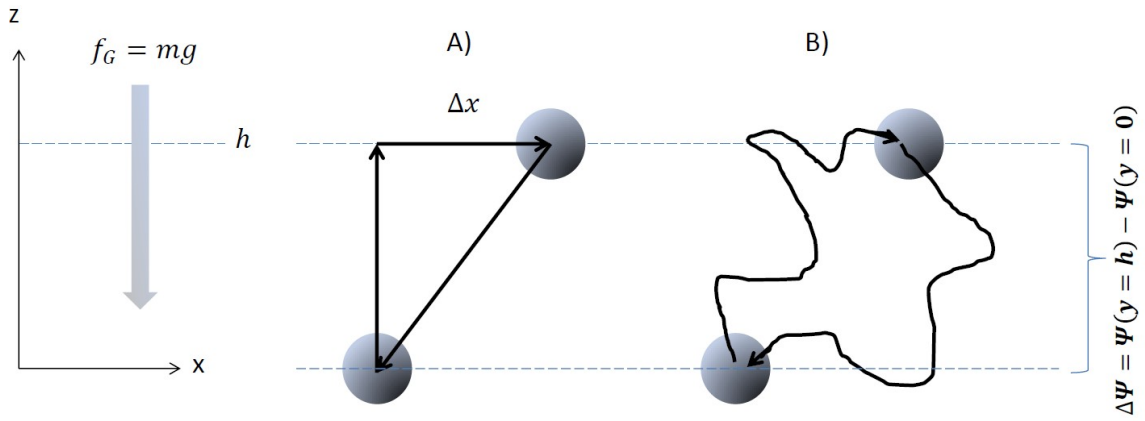

Figure 10: The integral along the motion of a ball in a potential (here: gravity) is independent from the shape of the path (A versus B). Only the difference in height h between start point and end point is important.

always implies the existence of a potential. The answer is no. There are forces that have no potential so that Equation 2.95 is not valid, but every potential implies a force. An example of a force that does not imply a potential is friction. If we encounter friction, for example while scuffing our feed, we always lose energy. There is no balance of energy if we walk in a circle (closed path) in contrast to the conversion of potential energy to kinetic energy in a gravitational field (see also Figure 10). Forces which have a potential and obey Equation 2.95 are called *conservative* since they conserve energy. Forces that have no potential are called *dissipative* forces.
But how can we distinguish between conservative and dissipative forces? Returning to Equation 2.84 and Equation 2.87 we see that the spatial components of the force are $f_x = \frac{\partial \Psi}{\partial x}$, $f_y = \frac{\partial \Psi}{\partial y}$ and $f_z = \frac{\partial \Psi}{\partial z}$. We inferred in Section 2.1.4 that the mixed second derivatives $\frac{\partial^2 \Psi}{\partial x \partial y}$ and $\frac{\partial^2 \Psi}{\partial y \partial x}$ are identical. Thus, $\frac{\partial f_x}{\partial y} = \frac{\partial f_y}{\partial x}$ or in general

$$\boxed{\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}}\,. \tag{2.96}$$

that is equivalent to

$$\boxed{\frac{\partial^2 \Psi_i}{\partial x_j \partial x_i} = \frac{\partial^2 \Psi_j}{\partial x_i \partial x_j}}\,. \tag{2.97}$$

Hence, a force f has a potential, if it obeys Equation 2.96. For three spatial coordinates, Equation 2.96 reads

$$\frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x} = 0 \, , \frac{\partial f_y}{\partial z} - \frac{\partial f_z}{\partial y} = 0 \, , \frac{\partial f_x}{\partial y} - \frac{\partial f_y}{\partial x} = 0 \, . \tag{2.98}$$

The structure of Equation 2.98 is called *curl* of a vector field, often denoted as $rot\vec{f}$ or $\nabla \times \vec{f}$. We will not need this operator any further and I therefore leave it with the definition. Thus, we now can summarize all the properties of a potential to

- a force $\vec{f}$ has a potential if integration along a *closed path* yields zero (Equation 2.95)
- a force has a potential if the mixed derivatives are identical (Equation 2.96)
- a force has a potential if $\nabla \times \vec{f} = 0$ (Equation 2.98)

Each of these statements is equivalent. If only one of the above statements is proven, it means that the others are fulfilled too and the force is conservative. Let us now consider some examples:

Example I:

Is g a potential if $dg = xy^2 \, dx + yx^2 \, dy$? According to Equation 2.84 $f_x = xy^2$ and $f_y = yx^2$. Therefore, $\frac{\partial f_x}{\partial y} = 2xy$ and $\frac{\partial f_y}{\partial x} = 2xy$ (Equation 2.96). Hence, g is a potential. dg is called **exact differential** *since it obeys Equation 2.96.*

Example II:

Is h a potential if $dh = \frac{\alpha}{x} \, dx + \beta x \, dy$? $f_x = \frac{\alpha}{x}$ and $f_y = \beta x$. Therefore, $\frac{\partial f_x}{\partial y} = 0$ and $\frac{\partial f_y}{\partial x} = \beta$. Hence, h is **not** *a potential unless $\beta = 0$. Since h is no potential, dh is called* **inexact differential**. *In order to distinguish dh from an exact differentials, inexact differentials are written δh.*

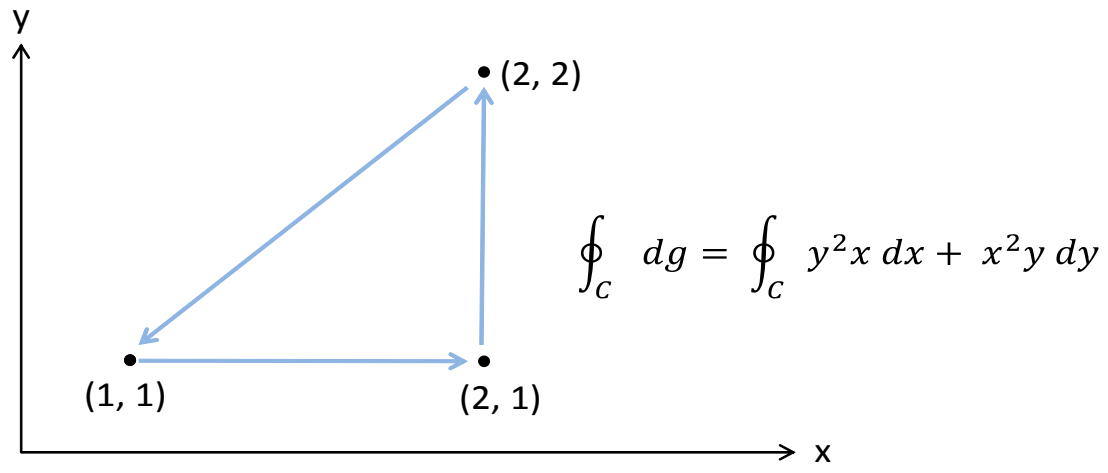

Figure 11: Path of integrating dg.

Example III:

Show that the integral of $dg = xy^2dx + yx^2dy$ *along a closed path is zero (Equation 2.95).*
Since we can use **any** *path, we take a very simple one: first integrating from (1,1) to (2,1) and then from (2,1) to (2,2) and finally back to (1,1) via (2,2), see Figure 11. The first part of the path leads from (1,1) to (2,1) that is parallel to the* x*-axis and therefore* $dy = 0$*. Thus, we integrate*

$$\int_{(1,1)}^{(2,1)} xy^2\ dx + 0 = \left.\frac{x^2y^2}{2}\right|_{(1,1)}^{(2,1)} = 4/2 - 1/2 = 3/2. \tag{2.99}$$

The second part of the path integrates dg *from (2,1) to (2,2) that is parallel to the* y*-axis and therefore* $dx = 0$*:*

$$\int_{(2,1)}^{(2,2)} 0 + x^2y\ dy = \left.\frac{x^2y^2}{2}\right|_{(2,1)}^{(2,2)} = 6. \tag{2.100}$$

Adding the contributions of both integrals gives the result of the path from (1,1) to (2,2) via (2,1) that yields $15/2$*.*
Now we integrate the last part from (2,2) to (1,1) that describes the function $y = x$ *and we can therefore substitute* $dy = dx$ *and obtain*

$$\int_{(2,2)}^{(1,1)} xy^2\ dx + yx^2\ dy = \int_{(2,2)}^{(1,1)} xx^2\ dx + xx^2\ dx = \left.\frac{2}{4}x^4\right|_{(2,2)}^{(1,1)} = -15/2. \tag{2.101}$$

Altogether the integral sums up to $3/2 + 6 - 15/2 = 0$*. Indeed,* g *is a potential since Equation 2.95 is valid.*

Exercise I:

Perform the integral of $dg = xy\ dx + yx^2\ dy$ *along the same closed path as before. What is the result now?*

Exercise II (advanced):

If $\delta g = xy\ dx + yx^2\ dy$ *is not an exact differential - what is the function* $\lambda(x, y)$ *you have to multiply* δg *with in order to turn it into an exact differential* $dg = \delta g\lambda(x, y)$*?*

2.2 Mathematical Approximation Methods

In many occasions exact equations are actually not required, be it if we are only interested in an order of magnitude estimate, or if the exact equation is too complicate and a simpler, but sufficiently accurate, estimate exists. For example we will often use an approximation for large N (meaning number of atoms/molecules or other particles, or number of measurements etc), especially in Section 3 and Section 5.2.2. Other reasons for approximations can be numerical treatments in programming and simulating (end of Section 7.3) and also for estimating the error propagation of measurements.

2.2.1 Error Estimation

In any experiment we can only measure a variable within a certain range of accuracy. The degree of accuracy depends on the used facilities and on the experimental set up itself. There is even a natural limit of accuracy given by the laws of physics. Once, a quantity is measured within a certain accuracy, i. e. within its *erros* (where the term "error" refers to the limited accuracy and not to an erroneous of false measurement or invalid data), we often want to derive a further quantity from that and therefore are also interested in the *error propagation*. For example, we want to measure the volume of a cell via its diameter. Once, the error of the diameter (that might have been measured with a microscope) is estimated (say 1%) we ask for the error (accuracy) of the volume. Does the error of the volume also equal 1%?
The volume V is a function f of the diameter d or the radius $r = d/2$. Thus, r equals the independent variable x. The error of the measured variable is assigned with a Δ, like Δx here. The aim is to derive the error of $V(r) = f(x)$, hence ΔV or Δf. We can easily imagine that Δf depends on at *which* value x we measure the error Δx, as illustrated in Figure 12. If the slope of $f(x)$ at position x is steep, the error Δf is large and if the slope

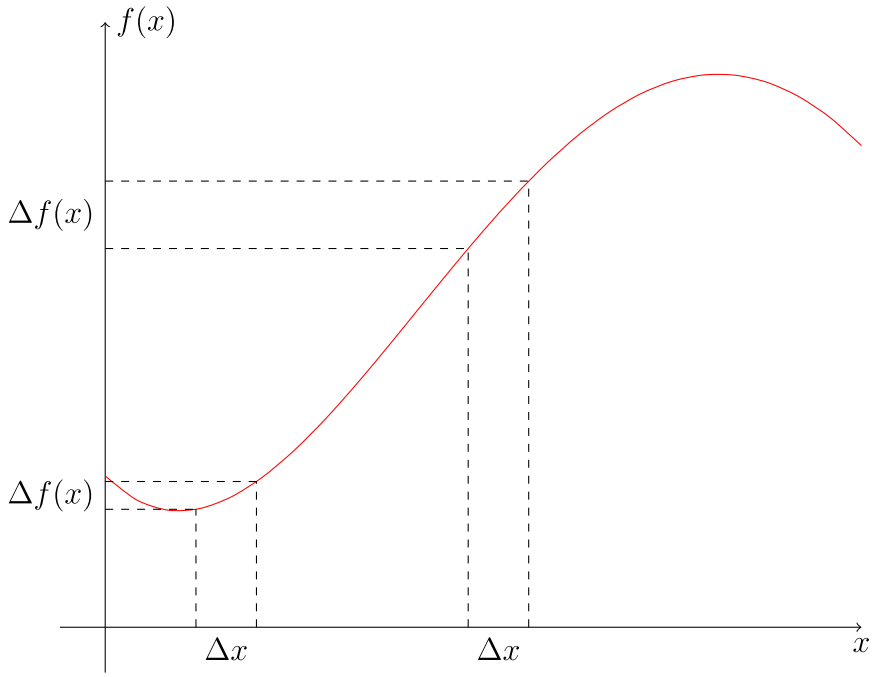

Figure 12: An error Δx causes errors in $f(x)$. The value of the error in $f(x)$ depends on the slope of $f(x)$ at position x.

is flat (meaning that f does not really change with x, hence a little inaccuracy does not

matter a lot), the error is small. Furthermore, we are only interested in the absolute value of the error, i.e. whether the slope is steep or not, but not in which direction it points to. Obviously, since the slope is important, we need the first derivative of the function $f(x)$. If Δx is sufficiently small we have seen in Section 2.1.3 that we can approximate the first derivative by

$$\frac{df}{dx} \approx \frac{\Delta f}{\Delta x}\,. \tag{2.102}$$

Hence, the error of f depending on Δx is

$$\Delta f = \left|\frac{df}{dx}\right| \Delta x\,. \tag{2.103}$$

Thus, the error of f can be estimated by the derivative of f with respect to x multiplied with the error Δx. Since the error is always positive we use the absolute value of the first derivative (the slope). As mentioned, the error has to be small in order to be allowed to use this approximation. As a rule of thumb, the relative error of x should be in the order of 1% and if the error of x is not symmetric (for example $r = 500^{+22}_{-38}\ \mu m$ instead of $r = 500 \pm 10\ \mu m$), it can be treated as too large and Equation 2.103 is not applicable.
Let us now return to our spherical cell and answer the question about the accuracy of its volume:

Example I:

The volume of a sphere is $V = \frac{4}{3}\pi r^3$ and r is measured with 1% accuracy. What is the relative error of V?
Here, r corresponds to x and thus V to f. According to Equation 2.103 we need the first derivative of $V(r)$ wrt r and therefore

$$\frac{dV}{dr} = 4\pi r^2 \tag{2.104}$$

and according to Equation 2.103, the error of V equals

$$\Delta V = \left|4\pi r^2\right| \Delta r\,. \tag{2.105}$$

The relative error is $\frac{\Delta V}{V}$ yielding

$$\frac{\Delta V}{V} = 3\frac{\Delta r}{r}\,. \tag{2.106}$$

Thus, if the value of r is measured with 1% accuracy, the relative error of V is 3%.

In the above example, the function depends only on one variable, but f can also be a function of n independent variables $f = f(x_1, x_2, \ldots, x_n)$. Then, according to Section 2.1.4 we have to sum up the contributions from all variables. This leads to the general form of Equation 2.103

$$\boxed{\Delta f = \sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right| \Delta x_i}\,, \tag{2.107}$$

that is illustrated in the next example.

Example II:

You try to quantify the amount of certain DNA fragments in different samples using a PCR (polymerase chain reaction). A fluorescent dye binds to your products of amplification. The strength of the signal is proportional to the amount of DNA in each tube.
The DNA doubles in each cycle, hence after c cycles the signal is proportional to 2^c. You measure until you reach a given threshold of DNA. Suppose you like to check whether the DNA really doubles and the reaction works properly. Therefore, you measure different samples $a, b, c, \ldots$ of different starting amounts of DNA and you have to calculate the factor

$$E = \frac{2^{c_a}}{2^{c_b}} = 2^{c_a - c_b} \tag{2.108}$$

after the threshold is reached. What is the error of E if Δc_a and Δc_b are known? Following Equation 2.107 we have to calculate

$$\Delta E = \left|\frac{\partial E}{\partial c_a}\right| \Delta c_a + \left|\frac{\partial E}{\partial c_b}\right| \Delta c_b \tag{2.109}$$

that is

$$\Delta E = \left|\ln(2)\ 2^{c_a - c_b}\right| \Delta c_a + \left|-\ln(2)\ 2^{c_a - c_b}\right| \Delta c_b \tag{2.110}$$

and finally

$$\Delta E = \ln(2)\ 2^{c_a - c_b} \left(\Delta c_a + \Delta c_b\right). \tag{2.111}$$

Since we add up the different error contributions, that in principle could (at least partly) cancel each other out, we estimate the **maximum** error. Nevertheless, a further rule of thumb is that if the errors are sufficiently small (hence symmetric), they correspond to the typical 1σ errorbar (see also Section 2.6.7 and Section 2.6.8).
If the quantities x_i are independent to each other, i. e. **are uncorrelated**, one can show that Equation 2.107 turns into

$$\boxed{\Delta f^2 = \sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right|^2 (\Delta x_i)^2}. \tag{2.112}$$

A mathematical justification for this expression is given in the statistics primer. However, one can infer already by the structure of Equation 2.112 that it is the length of the vector Δf which is constructed as linear combination of the Δx_i, weighted by the respective derivatives. Hence, the x_i are linearly independent, whereas Equation 2.107 would contain the mixed terms (identified as covariance terms in the statistics primer). Therefore, Equation 2.112 leads to **smaller errors** Δf as if calculated from Equation 2.107.

Exercise:

A simple model of gene expression can be described by the following reactions (c. f. Section 5.2.4):

$$\frac{dr(t)}{dt} = k_R - \gamma_R\, r(t)\,, \tag{2.113}$$

$$\frac{dp(t)}{dt} = k_p\, r(t) - \gamma_P\, p(t)\,, \tag{2.114}$$

where γ_R and γ_P are the degradation rates of mRNA and a particular protein, respectively, and $r(t)$ and $p(t)$ denote the respective concentrations. The production rates are denoted as k_R and k_P, respectively.
Express k_R in terms of all the other variables for steady state $\left(\frac{dr(t)}{dt} = 0;\ \frac{dp(t)}{dt} = 0\right)$.
Find an equation for the relative error of k_R if the errors of $p(t)$ and k_P are known once for maximum error estimation and by assuming no correlation between any of the quantities.

2.2.2 Taylor Series

If the errors are too large, Equation 2.107 and Equation 2.112 are not applicable. Therefore, there exists another, even more general, approximation method that is called *Taylor*[12] *approximation*. We will use this approximation many times in this script (Section 2.6.5, Section 2.6.7, Section 4.4.1, Section 5.2, Section 5.3, Section 6.2, Section 7.3.1 and Section 9.4), in the lecture and it is a very important tool appearing frequently in quantitative bioscience. Hence, we have to get familiar with it and you should be able to apply it from scratch in any situation.
The mathematical proof for the Taylor approximation (also called Taylor series) is based on induction and is a bunch of lengthy, but basic, algebra. Since the proof is not important for our purposes, I like to give the final equation that reads

$$\begin{aligned} f(x) =& \sum_{k=0}^{\infty} \frac{f^{(k)}(x)|_{x_0}}{k!}(x-x_0)^k \\ =& f(x_0) + f'(x)|_{x_0}(x-x_0) + \frac{f''(x)|_{x_0}}{2}(x-x_0)^2 + \dots\,. \end{aligned} \tag{2.115}$$

The Taylor series states, that a function $f(x)$ can be represented as an infinite sum of the k^{th} derivatives at position x_0, presupposed that these derivatives exist. For $n < \infty$, i. e. truncating terms of order higher than n, we obtain the approximation

$$f(x) \approx \sum_{k=0}^{n} \frac{f^{(k)}(x)|_{x_0}}{k!}(x-x_0)^k\,. \tag{2.116}$$

Frequently, when writing the full infinite series but expressing only the terms up to order n, the higher orders are joined in the expression "$\mathcal{O}$", so that you will occasionally find an expression like

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \mathcal{O}(x^{n+1})\,. \tag{2.117}$$

Let us now investigate Equation 2.115 a bit further and write down the first two orders, i.e. the zeroth order ($k = 0$) and the first order ($k = 1$). According to Equation 2.115, the zeroth order equals the function itself at position $x = x_0$, i.e. $f(x_0)$ and the first order

[12] James Gregory, 1638 - 1675 & Brook Taylor, 1685-1731

equals the first derivative at $x = x_0$, so that we obtain

$$\begin{aligned} f(x) \approx & \frac{f(x)|_{x_0}}{0!}(x - x_0)^0 + \frac{f'|_{x_0}}{1!}(x - x_0)^1 \\ = & f(x_0) + \frac{df}{dx}|_{x_0}(x - x_0)\,. \end{aligned} \tag{2.118}$$

Denoting $(x - x_0)$ as Δx and rearranging the above equation, we can recover the approximation of the first derivative (c.f. Equation 2.34):

$$\frac{df}{dx}|_{x_0} \approx \frac{f(x) - f(x_0)}{\Delta x}\,. \tag{2.119}$$

It is important to note that the point x_0 from where the series is expanded from does not

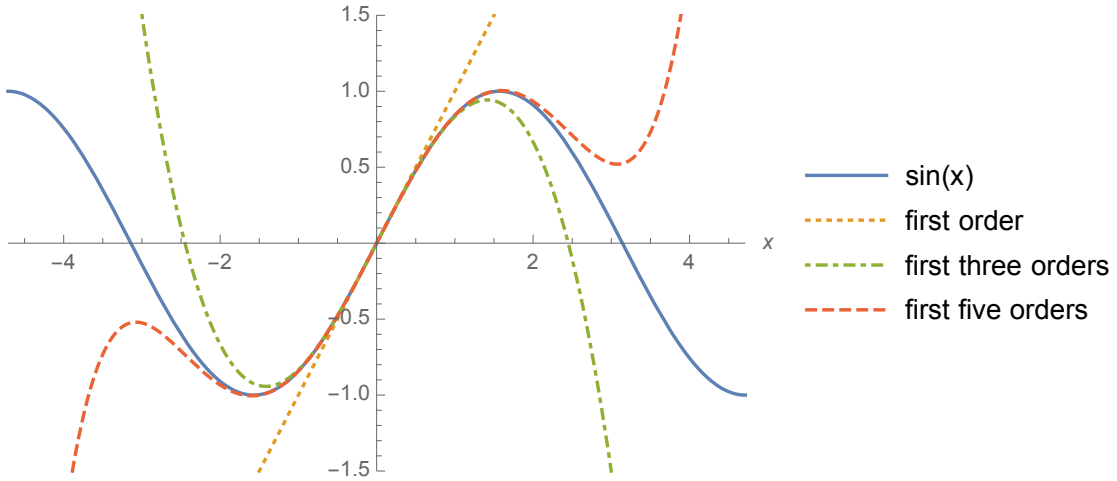

Figure 13: The first five orders of the Taylor series of the function $\sin(x)$ at $x_0 = 0$. As more orders are joined in the series (Equation 2.115), as more accurate the approximation of the function $f(x)$ becomes.

appear on the lhs of the Tayler series (Equation 2.115). Thus, the approximation of $f(x)$ is independent from the point x_0 and we therefore choose a value, where the derivatives become very simple. Let us examine the example of the sin function. The first three derivatives are

$$f(x) = \sin(x) \qquad f'(x) = \cos(x) \qquad f''(x) = -\sin(x) \qquad f'''(x) = -\cos(x).$$

We now choose an x_0 where these derivatives become very simple, for example $x_0 = 0$. Therefore, we obtain

$$f(0) = 0 \qquad f'(0) = 1 \qquad f''(0) = 0 \qquad f'''(0) = -1$$

Inserting these results into Equation 2.115, we find the series for the sine that reads

$$\boxed{\sin(x) = 0 + x - 0 - \frac{x^3}{6} + 0 + \frac{x^5}{120} \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}}\,. \tag{2.120}$$

The sine can be expressed as a power series with odd powers.
While adding more orders, the approximation improves. For a small number of orders, the validity of the approximation holds only for small Δx (like in Equation 2.119). If the number of orders increases, the approximation holds for even larger Δx. The improvement of the approximation by including higher orders in the Taylor series is illustrated in Figure 13.
Before proceeding any further, I like to include an example of biological relevance:

Example:

How does the energy E_{strain} for the deformation of macromolecular assemblies depend on the displacement Δx? For example, we want to calculate the amount of energy required to stretch an actin filament. However, the energy $E(x)$ as function of length x is completely unknown. What we can deduce is that the filament is in its energetic minimum (the equilibrium state) when it is neither stretched nor compressed. No matter how the exact energy landscape looks like, the energy reaches a minimum at the equilibrium length $x = x_{eq}$. For small displacements Δx we can perform the Taylor expansion of $E_{strain}(x)$. The function we want to approximate is $E(x_{eq} + \Delta x)$ around its equilibrium length x_{eq}, so that $(x - x_0)$ in Equation 2.115 corresponds to $(x_{eq} + \Delta x - x_{eq})$ and x in Equation 2.115 equals $(x_{eq} + \Delta x)$. Thus, a second order expansion yields the approximation

$$E\left(x_{eq} + \Delta x\right) \approx E\left(x_{eq}\right) + \frac{dE}{dx} \mid_{x_{eq}} \Delta x + \frac{1}{2} \frac{d^2E}{dx^2} \mid_{x_{eq}} \left(\Delta x\right)^2 . \tag{2.121}$$

$E_{strain}(x)$ reaches a minimum at the equilibrium length, i.e. the first derivative of $E(x)$ at $x = x_{eq}$ must equal zero $\frac{dE}{dx} \mid_{x=x_{eq}} = 0$. This cancels the second term in Equation 2.121.
The second derivative at $x = x_{eq}$ is just a constant, usually denoted as k, so that Equation 2.121 reads

$$E\left(x_{eq} + \Delta x\right) \approx E\left(x_{eq}\right) + 0 + \frac{1}{2} k \left(\Delta x\right)^2 . \tag{2.122}$$

The change of energy from the non-stretched to the stretched state can be denoted as $\Delta E = E\left(x_{eq} + \Delta x\right) - E\left(x_{eq}\right)$, so that we finally obtain the approximation

$$\Delta E \approx \frac{1}{2} k \left(\Delta x\right)^2 . \tag{2.123}$$

The above equation is called Hooke's[13] law and can also be used to model a spring (therefore, the constant k is called spring constant*) or the vibrations of an AFM tip (Section 9.4).*

Often, a function $f(x)$ is approximated around its extreme by the second order Taylor series leading to a quadratic expression like in Equation 2.123. Such an approximation is called *harmonic* and is visualized in Figure 14.

Let us now return to the Taylor series of known functions. Equation 2.120 expresses the sine function as a sum of **odd** powers of x. The same procedure can be performed with the cosine function and we obtain for the first three derivatives

$$f(x) = \cos(x) \qquad f'(x) = -\sin(x) \qquad f''(x) = -\cos(x) \qquad f'''(x) = \sin(x) .$$

Again, we insert a particular x_0 where the derivatives become simple, that is again $x_0 = 0$ and we find analogously that

$$f(0) = 1 \quad f'(0) = 0 \quad f''(0) = -1 \quad f'''(0) = 0$$

[13] Robert Hooke, 1635 - 1703

and the cosine can be written as the sum of **even** powers of its argument

$$\boxed{\cos(x) = 1 - 0 - \frac{x^2}{2} + 0 + \frac{x^4}{24} - 0 \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}}. \tag{2.124}$$

Finally, we derive the Taylor series of the exponential function e^x around $x_0 = 0$ and find that

$$\begin{array}{llll} f(x) = e^x & f'(x) = e^x & f''(x) = e^x & f'''(x) = e^x \\ f(0) = 1 & f'(0) = 1 & f''(0) = 1 & f'''(0) = 1 \end{array},$$

$$\boxed{e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}}. \tag{2.125}$$

Hence, the Taylor series of the exponential function e^x contains both, **even and odd** powers of x. Therefore, one might suspect, that we can join sine and cosine in order to obtain the exponential function. This *almost* works, but doesn't fit with the signs. If we introduce a constant i with the yet strange property $i^2 = -1$, we can indeed write the relation

$$\boxed{e^{ix} = cos\, x + i\, sin\, x}, \tag{2.126}$$

that is called *Euler's*[14] *relation*. Eulers relation is of fundamental relevance and its purpose, and the meaning of the mysterious constant i, becomes more apparent in Section 2.3.

Finally, I like to mention, that the Taylor series does not always work. For example, one or more derivatives at x_0 might not exist (like at $x_0 = 0$ for the derivatives of $\ln x$) or the Taylor series does not converge in some rare cases, like e.g. for the function $f(x) = e^{-\frac{1}{x^2}}$.

Exercise:

Proof Euler's relation (Equation 2.126) by examining the Taylor series of e^{ix}, $cos\, x$ *and* $i\, sin\, x$.

[14] Leonhard Euler, 1707 - 1783

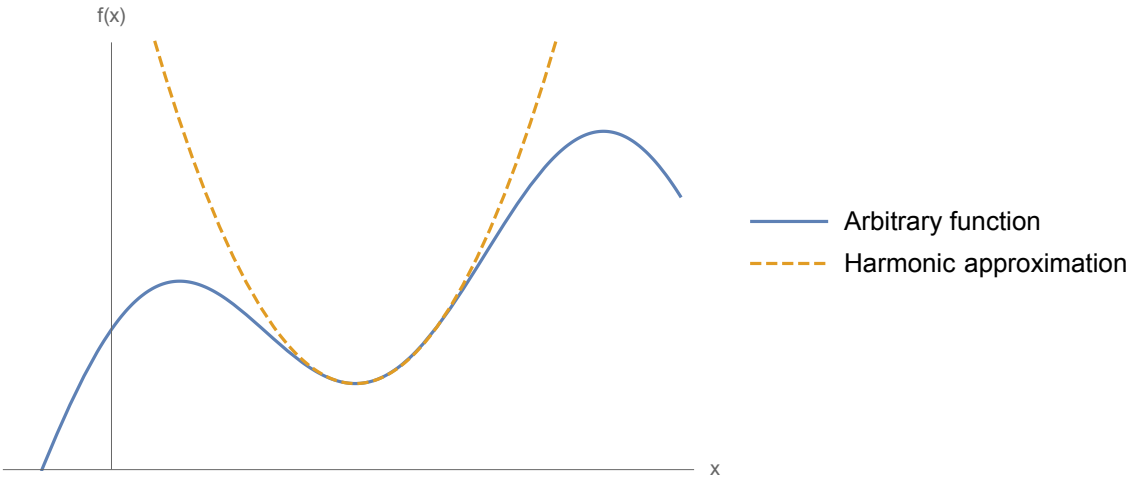

Figure 14: A harmonic (second order Taylor series) approximation of an unknown function $f(x)$ around its extreme.

2.2.3 Newton's Method

Suppose you want to infer the solution of an equation of the type $\cos(x) = x$ or $x = e^x$, etc., i. e. when it is not possible to rearrange the equation in order to derive an explicit expression for x. One way to solve the problem is to plot the two functions $y_1 = x$ and $y_2 = \cos(x)$ (or $y_1 = x$ and $y_2 = e^x$, respectively) and to derive x from the intersection point of the two functions graphically. The accuracy of this method depends on the accuracy of the plot and it is not applicable as a programming task and therefore not suited for an automated approach.
A better way of deriving a solution of (in principle) infinite accuracy is the mathematical approach called *Newton's method*. Finding a solution for the this problem equals finding the zeros of a function like $f(x) = \cos x - x$ or $f(x) = e^x - x$, respectively. If we guess a starting point x_n, we would obtain the corresponding value $f(x_n)$. The tangent at this point is a linear function of the general from $y = m\,x + c$. Since m is the slope of the function f at $x = x_n$, i. e. the derivative, we can write the equation for the tangent as

$$f(x_n) = f'(x_n)\,x_n + c\,. \tag{2.127}$$

Rearranging this equations leads to the intercept

$$c = f(x_n) - f'(x_n)\,x_n \tag{2.128}$$

and we can therefore express the tangent as

$$y = (x - x_n)f'(x_n) + f(x_n)\,. \tag{2.129}$$

Of course, $f(x_n)$ does not equal zero since the first guess $x = x_n$ is almost certainly not the correct x value. But with Equation 2.129 we can determine that tangent, that corresponds to the correct solution: it must yield $y = 0$ for a new guess $x = x_{n+1}$. Therefore, we find the solution

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}\,. \tag{2.130}$$

This equation describes an iterative process of n steps approaching the exact solution $x_{n=\infty} := x_*$. The method is illustrated in Figure 15.

If we take the example $f(x) = \cos(x) - x$ and guess $x_1 = 0.5$ $(n = 1)$ as the possible zero we have to calculate $f'(x) = -\sin(x) - 1$ and Equation 2.130 yields $x_2 = 0.755$ $(n = 2)$. Inserting this value on the rhs of Equation 2.130 as the updated x_n we obtain $x_3 = 0.739\ldots$ and so on. After already four steps $(x_4 = 0.73906\ldots)$, we would end up with $\cos(0.73906) = 0.739102\ldots$, that corresponds to an accuracy of less than 0.1%, although the initial guess $(x_1 = 0.5)$ was completely off. Therefore, Newton's method is very efficient and converges fast.

2.2.4 Stirling Approximation

Especially in thermodynamics and statistics (Section 3) we frequently encounter factorials and their derivatives, in particular cases with $N!$ (c. f. Section 2.6.4), where N refers to the number of particles. Unfortunately, derivatives of factorials are not trivial and we would need a so called *Gamma function* for that. However, in classical thermodynamics, we usually deal with **large numbers** of particles so that we can apply an approximation

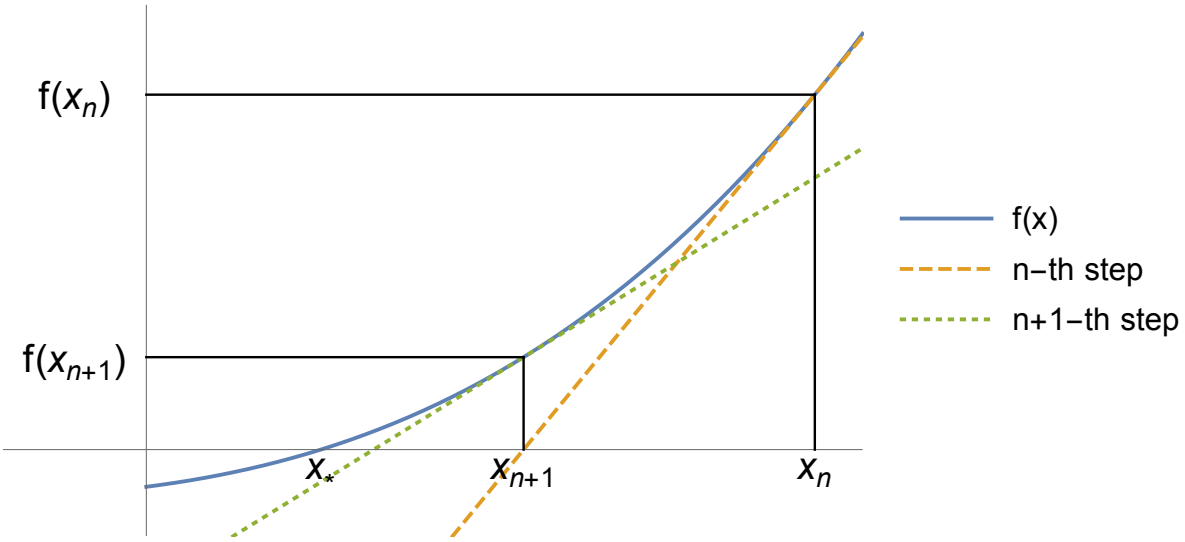

Figure 15: Visualization of Newton's method. The tangent lines are converging towards the correct value x_* for increasing number of steps n.

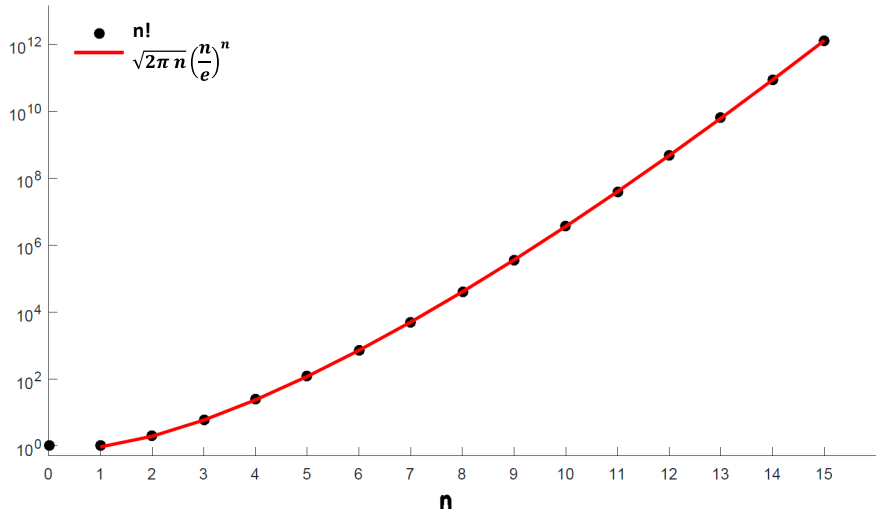

Figure 16: Comparison of factorial $N!$ (black, filled dots) to Stirling's approximation (red, solid line).

for it. This so called *Stirling's*[15] *approximation*

$$\boxed{N! \approx \sqrt{2\pi N}\left(\frac{N}{e}\right)^N} \tag{2.131}$$

is given here without a proof.
Often, Stirling's approximation is also written in $\ln$

$$\ln(N!) = \frac{1}{2}\ln 2\pi + \frac{1}{2}\ln N + N\ln N - N\underbrace{\ln e}_{\ln e=1} \tag{2.132}$$

and approximated even further by omitting non leading terms, so that it reads

$$\boxed{\ln(N!) \approx N\ln N - N}\,. \tag{2.133}$$

[15] James Stirling, 1692 - 1770

Figure 16 illustrates Stirling's approximation compared to the exact values of $N!$. N counts as sufficiently large, if $N \gtrsim 100$ since Stirling's approximation is accurate by $< 0.1\%$ above this value. In classical thermodynamics $N \approx 10^{23}$ (one mol) and even in the nano molar regime, N is by far large enough to apply Stirling's approximation.

2.3 Complex Numbers

So far we just have worked with real numbers $x \in \mathbb{R}$. In many cases (Section 2.4 and Section 7.1), it will turn out that the set of real numbers is not sufficient and we have to extent it by *imaginary numbers*. The joint set of real numbers and imaginary numbers is called *complex numbers* $z \in \mathbb{C}$.
The term "imaginary number" sounds somewhat artificial, but they are actually as "real" as real numbers and they help to simplify many calculations that would be quite complicated otherwise (see Section 2.4, Section 4.3 and Section 7.1). The resistance in an electromagnetic coil, the absorption of light that helps to measure the amount of cells in a Petri dish and the Fourier image of a crystallized protein - all these things cannot be understood without accepting the existence of complex numbers and their imaginary part. Extending the set of real numbers to complex numbers equals logically the step of extending natural numbers (1, 2, 3 ...) to integers by including negative numbers in order to explain operations like $1 - 3 = -2$ etc.
In Section 2.2.2 we introduced the *imaginary unit* i by its property

$$i^2 = -1\,. \tag{2.134}$$

Taking the square root of this equation leads to

$$i = \sqrt{-1}\,, \tag{2.135}$$

hence, complex numbers enable us to solve roots of negative numbers. Therefore, we find solutions for equations like $x^2 = -1$, namely $x = i$ and $x = -i$.
Every complex number $z \in \mathbb{C}$ can be written as the sum of a real part $Re(z) = a$ and an imaginary part $Im(z) = b$ by

$$z = a + ib, \quad Re(z) = a, \quad Im(z) = b, \quad a, b \in \mathbb{R}\,. \tag{2.136}$$

While real numbers can be visualized on a line running from $-\infty$ to ∞, complex numbers are visualized by a *complex plane*, where both, the real part and the imaginary part, running from $-\infty$ to ∞ (Figure 17). A point in the complex plane equals the complex number

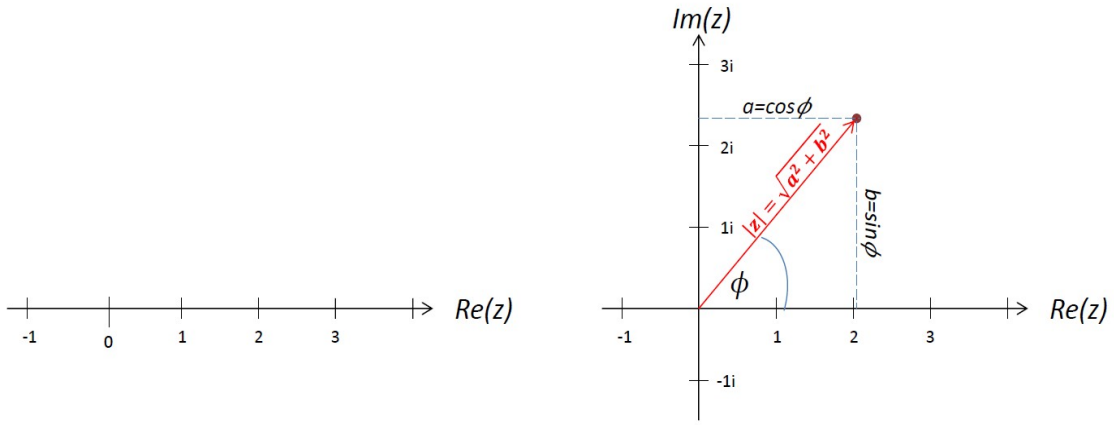

Figure 17: Real numbers $x \in \mathbb{R}$ are located on a line (left), whereas complex numbers $z \in \mathbb{C}$ are described by a complex plane (right).

z and the complex plane can be treated as a coordinate system containing the points $\vec{z}$

with their coordinates $\begin{pmatrix} a \\ b \end{pmatrix}$. Therefore, the absolute value of a complex number equals $|z| = \sqrt{a^2 + b^2}$ with its *phase angle* $\tan\phi = \frac{b}{a}$. To describe a complex number completely, we need both: $|z|$ **and** ϕ (Figure 17, right).
We can find to each complex number $z = a + ib$ its *complex conjugate* $\bar{z} = a - ib$. This definition corresponds to a reflection at the real axis (x-axis). Any calculation with complex numbers works as with real numbers, the only difference is that we have to be aware of the property $i^2 = -1$:

$$z_1 \pm z_2 = (a_1 + ib_1) \pm (a_2 + ib_2) = (a_1 \pm a_2) + i \cdot (b_1 \pm b_2) \tag{2.137}$$
$$z_1 \cdot z_2 = (a_1 + ib_1) \cdot (a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i \cdot (a_1 b_2 + a_2 b_1)\,. \tag{2.138}$$

We will often use some useful relations like

$$|z|^2 := z\bar{z} = a^2 + b^2 \tag{2.139a}$$
$$Re(z) = \frac{z + \bar{z}}{2} \tag{2.139b}$$
$$Im(z) = \frac{z - \bar{z}}{2i} \tag{2.139c}$$
$$\frac{z_1}{z_2} = \frac{1}{|z_2|^2}(a_1 a_2 + b_1 b_2) + \frac{i}{|z_2|^2}(b_1 a_2 - a_1 b_2) \tag{2.139d}$$

and one should consider the powers of i:

$$\begin{array}{llll} i^0 = 1, & i^1 = i, & i^2 = -1, & i^3 = -i, \quad i^4 = 1 \\ i^0 = 1, & i^{-1} = -i, & i^{-2} = -1, & i^{-3} = i, \quad i^{-4} = 1\,. \end{array} \tag{2.140}$$

These relations are repetitive after the fourth power (positive or negative) and we can therefore summarize

$$i^{4n} = 1, \quad i^{4n+1} = i, \quad i^{4n+2} = -1, \quad i^{4n+3} = -i, \quad n \in \mathbb{Z}\,. \tag{2.141}$$

2.3.1 Eulers relation

Following the vector representation of z, we can move from the Cartesian coordinates $z = z(a, b,)$ to polar coordinates $z = z(\phi, r)$ by (c. f. Figure 17 and Section 2.5)

$$r = \sqrt{a^2 + b^2} = |z|, \qquad a = r\,\cos\phi, \qquad b = r\,\sin\phi, \qquad z = r\,(\cos\phi + i\sin\phi) \tag{2.142}$$

and use *Euler's equation* (Equation 2.126), to obtain the important relations

$$e^{i\phi} = \cos\phi + i\sin\phi \tag{2.143}$$
$$z = a + ib \Longleftrightarrow z = r \cdot e^{i\phi}\,. \tag{2.144}$$

Euler's famous relation reveals that $\cos(x)$ and $\sin(x)$ are actually the real and imaginary part, respectively, of the exponential function e^{ix}. This becomes apparent, when we like to solve equations of the type

$$\frac{d^2}{dt^2}x = kx\,. \tag{2.145}$$

We can try $x(t) = x_0\, sin\,\omega t$ and see that

$$x(t) = x_0\, sin\,\omega t$$

$$\begin{aligned}\frac{d}{dt}x(t) =& \omega\, x_0\, cos\, \omega t \quad \text{and} \\ \frac{d^2}{dt^2}x(t) =& -\omega^2\, x_0\, sin\, \omega t\,,\end{aligned}$$

so that Equation 2.145 holds for $x(t) = x_0\, sin\, \omega t$ and $k = -\omega^2$.
This also works with $x(t) = x_0\, cos\, \omega t$ and $x(t) = x_0\, e^{\omega t}$:

$$\begin{aligned}x(t) =& x_0\, cos\, \omega t \\ \frac{d}{dt}x(t) =& -\omega\, x_0\, sin\, \omega t \quad \text{and} \\ \frac{d^2}{dt^2}x(t) =& -\omega^2\, x_0\, cos\, \omega t\end{aligned}$$

and

$$\begin{aligned}x(t) =& x_0\, e^{\omega t} \\ \frac{d}{dt}x(t) =& \omega\, x_0\, e^{\omega t} \quad \text{and} \\ \frac{d^2}{dt^2}x(t) =& \omega^2\, x_0\, e^{\omega t}\,,\end{aligned}$$

with the difference that $k = +\omega^2$ for the exponential. All three functions, $\sin$, $\cos$ and the exponential solve Equation 2.145, but the more general solution is the combination of all three: $x(t) = x_0\, e^{i\omega t}$, the complex solution that combines these sub cases. The same applies for so-called wave equations having the form

$$\frac{d^2}{dt^2}f(x,t) = k\frac{d^2}{dx^2}f(x,t)\,. \tag{2.146}$$

Such relations are extensively used in Section 2.4, Section 4.3 and Section 9.

Another strength of complex numbers is, that Euler's equation can be used to derive relations between the trigonometric functions with a relatively small effort. For example, we know that $\left(e^{i\phi}\right)^2 = e^{2i\phi}$ and using Euler's equation, we find

$$\begin{aligned}\left(\cos\phi + i\, \sin\phi\right)^2 =& \cos\, 2\phi + i\, \sin\, 2\phi \\ \cos^2\phi + 2i\, \cos\phi\, \sin\phi - sin^2\phi =& \cos\, 2\phi + i\, \sin\, 2\phi\end{aligned} \tag{2.147}$$

Equation 2.147 is valid, if both, the real and imaginary parts on both sides of the equation are independently equal and we therefore find the identities

$$\begin{aligned}\cos\, 2\phi &= \cos^2\phi - \sin^2\phi \\ \sin\, 2\phi &= 2\, \cos\phi\, \sin\phi\,.\end{aligned} \tag{2.148}$$

Another useful relation can be derived from $\frac{e^{i\phi_1}}{e^{i\phi_2}} = e^{i(\phi_1-\phi_2)}$, that we can write as (c. f. Equation 2.126)

$$\frac{\cos\phi_1 + i\, \sin\phi_1}{\cos\phi_2 + i\, \sin\phi_2} = \cos\left(\phi_1 - \phi_2\right) + i\, \sin\left(\phi_1 - \phi_2\right)\,. \tag{2.149}$$

Using Equation 2.139 leads to the identities

$$\begin{aligned}\cos\left(\phi_1 - \phi_2\right) &= \cos\phi_1\, \cos\phi_2 + \sin\phi_1\, \sin\phi_2 \\ \sin\left(\phi_1 - \phi_2\right) &= \cos\phi_2\, \sin\phi_1 - \cos\phi_1\, \sin\phi_2\,.\end{aligned} \tag{2.150}$$

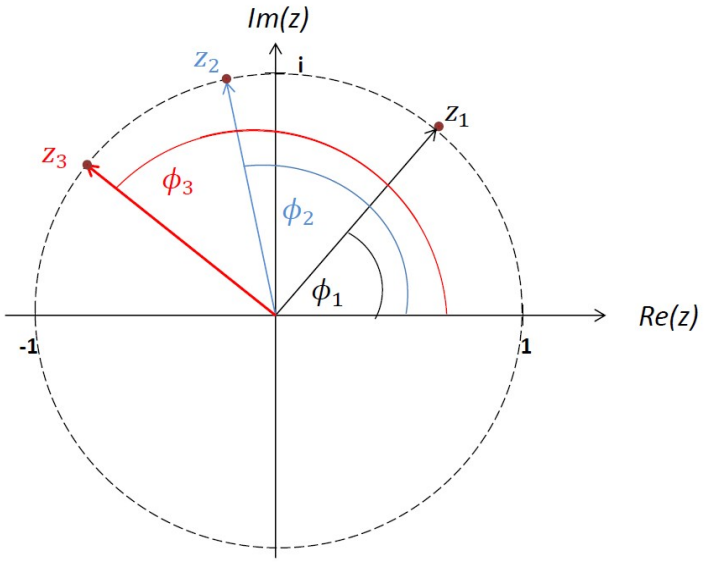

Figure 18: The multiplication of two complex numbers $z_1\, z_2 = z_3$ equals a rotation in the complex plane, since the phase angles are added ($\phi_3 = \phi_1 + \phi_2$).

All trigonometrical relations can be derived in such a manner. Without complex numbers, we would have to proof these identities by relatively complicated geometrical constructions. Euler's relation also helps to multiply complex numbers and calculating their roots. When multiplying two complex numbers z_1 and z_2, we perform the operation

$$z_1 \cdot z_2 = r_1 e^{i\phi_1} \cdot r_2 e^{i\phi_2} = r_1 r_2\, e^{i(\phi_1+\phi_2)}\,. \tag{2.151}$$

We see, that we add the two angles $\phi_1 + \phi_2$ in the exponential, so that the product z_3 has the new phase angle $\phi_3 = \phi_1 + \phi_2$, that equals a **rotation** in the complex plane. A multiplication of complex numbers with $r = 1$ for example, equals a rotation along the unit circle (Figure 18). Multiplying a complex number n times by itself leads according to Equation 2.151 to the result $z^n = r^n\, e^{n\,i\phi}$. Therefore, the n^{th} root of a complex number, hence the inverse operation, yields a **division** of the phase angle by n and we find that

$$\left(r\, e^{i\phi}\right)^{1/n} = r^{1/n}\, e^{i\phi/n} = r^{1/n}\left[\cos\left(\phi/n\right) + i\, \sin\left(\phi/n\right)\right]\,. \tag{2.152}$$

The result of the root is again a complex number that we can denote $\xi = \rho\,(\cos\alpha + i\, \sin\alpha)$, where $\rho = r^{1/n}$ and $\alpha = \frac{\phi}{n}$. The solution for $\alpha = \frac{\phi}{n}$ is only **one** solution, since trigonometric functions are periodic for a shift of 2π times an integer number k and therefore we obtain n different solutions

$$\alpha_k = \frac{\phi}{n} + \frac{2\pi(k-1)}{n} \quad \text{for } k = 1, 2, \ldots, n\,. \tag{2.153}$$

Example I:

What is the square root ($n = 2$) of $z = i$?
For $z = i$, the phase angle equals $\phi = 90°$ (Figure 17) and $\alpha_1 = 45°$. According to Equation 2.153, the second solution has the phase angle $\alpha_2 = 45° + 180° = 225°$. Since $|z| = 1$, also the absolute value of the two solutions $1^{1/2} = \rho_{1;2}$ equals one.

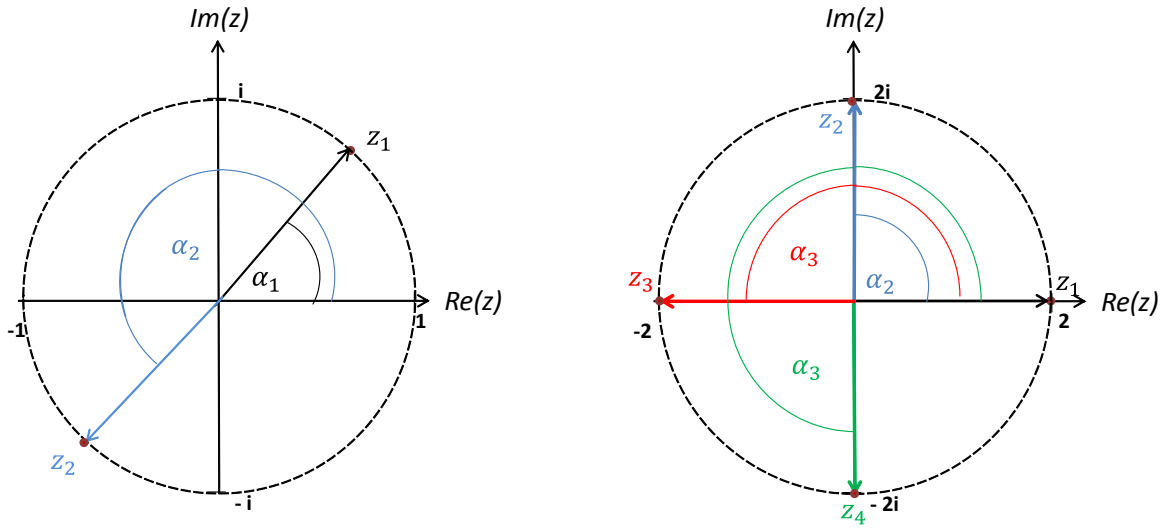

Figure 19: Solutions of $\sqrt{i}$ (left) and $\sqrt[4]{16}$ (right).

Thus, the two solutions are

$$z_1 = \cos 45° + i \sin 45° = \frac{1}{2}\sqrt{2} + i\frac{1}{2}\sqrt{2} \quad \textit{and}$$
$$z_2 = \cos 225° + i \sin 225° = -\frac{1}{2}\sqrt{2} - i\frac{1}{2}\sqrt{2}\,.$$

Example II:

What is the fourth root $(n = 4)$ *of* $z = 16$*?*
For real numbers, we have only two solutions $z_1 = 2$*, since* $2^4 = 16$ *and* $z_2 = -2$*, since* $(-2)^4 = 16$*, but we know now that there must be two other solutions. The phase angle of* z *equals* $0°$ *or* $360°$*, so that we find* $\alpha_1 = 0°$*,* $\alpha_2 = 90°$*,* $\alpha_3 = 180°$ *and* $\alpha_4 = 270°$ *(Equation 2.153). Since* $|z| = 16$*, the absolute value of the four solutions is* $16^{1/4} = \rho_{1;2;3;4} = 2$*. Therefore, the solutions are*

$$z_1 = 2\,(\cos 0° + i \sin 0°) = 2$$
$$z_2 = 2\,(\cos 90° + i \sin 90°) = 2i$$
$$z_3 = 2\,(\cos 180° + i \sin 180°) = -2 \quad \textit{and}$$
$$z_4 = 2\,(\cos 270° + i \sin 270°) = -2i\,.$$

The solutions of the two examples are illustrated in Figure 19.

Example III:

What is the third root of $z = -1 + i$*?*
First, we find that $|z| = \sqrt{(-1)^2 + (1)^2} = \sqrt{2}$ *(because* $a = -1$ *and* $b = 1$*, Equation 2.136 and Figure 17) and that* $\phi = arctan\left(\frac{1}{-1}\right) = 3\pi/4$*. According to Equation 2.153, we have three solutions with* $\alpha_1 = \frac{3\pi}{4}\frac{1}{3} = \frac{\pi}{4}$*,* $\alpha_2 = \frac{3\pi}{4}\frac{1}{3} + \frac{2\pi 1}{3} = \frac{11\pi}{12}$ *and* $\alpha_3 = \frac{3\pi}{4}\frac{1}{3} + \frac{2\pi 2}{3} = \frac{19\pi}{12}$*, so that the complete solutions are*

$$z_1 = (2)^{1/6} e^{i\frac{\pi}{4}}$$
$$z_2 = (2)^{1/6} e^{i\frac{11\pi}{12}} \quad \textit{and}$$
$$z_4 = (2)^{1/6} e^{i\frac{19\pi}{12}}\,.$$

A further benefit of complex numbers is that we are now able to calculate negative logarithms. For real numbers, solutions for e.g. $\ln(-1)$ did not exist. Thanks to Euler's relation, we can now write $e^{-i\pi} = -1$ and therefore obtain $\ln(-1) = -i\pi$, a solution in the complex plane.

2.4 Fourier Transformation

A direct mathematical application of complex numbers is the *Fourier transformation*[16]. This mathematical tool appears frequently in biophysics and biochemistry and is usually a mystery to the students. This is natural, since it is one of the most abstract tools we will use. Applications of the Fourier transformation reach from diffusion reaction equations (Section 7.1), to noise filtering (Section 9.1.8) and image processing in cryo - electron microscopy (Section 9.3), super resolution microscopy (Section 9.1.4) and image reconstruction in X-Ray crystallography (Section 9.2).
Although Fourier transformation is very abstract (in particular if treated with the full mathematical rigorousness), its idea can be easily visualized. I try to explain Fourier transformation in two ways in this section. One way is the visual approach and the other way is the underlying math that will be discussed to some extend. In fact, it is actually more important for a life scientist to understand the idea and the concept of Fourier transformation (but this exhaustively) rather than the entire mathematical set up (that is therefore not complete here on purpose[17])

2.4.1 Think Inverse!

The main idea of Fourier transformation (FT) is the transformation from a quantity in the actual space into the reciprocal or inverse space and back. For example a single tone is a sound wave of a particular frequency ω. This sound wave can be modeled as $x(t) = x_0 \sin \omega t$ or $x(t) = x_0 \cos \omega t$, where x_0 refers to the amplitude. The sine or cosine in the *time domain* equals the classical wave that is shown in Figure 20, upper left.
Since this cosine wave has only one frequency, it equals only one (infinitely) narrow peak if we would use not time but frequency as x-axis (Figure 20, upper right). The frequency $\omega = \frac{2\pi}{T}$ (with period T) has the unit $1/time$, i. e. is the inverse of time. Therefore, the *frequency domain* is the *inverse or reciprocal* of the time domain. The set of frequencies is therefore called *inverse space* of the time domain. A large period corresponds to a small frequency and vice verse. The sine wave in the time domain with its amplitude and period hosts the same information as the corresponding peak in the inverse domain with its frequency and the power (height of the peak). Therefore, it is possible, to reconstruct the cosine wave from the peak and vice verse. The mathematical recipe that performs the transformation is the FT.
If our signal in the time domain would not be just a single tone, but a signal of the form

$$f(t) = \cos \omega t + \frac{1}{2} \cos\, 2\omega t + \frac{1}{3} \cos\, 3\omega t + \frac{1}{4} \cos\, 4\omega t + \frac{1}{5} \cos\, 5\omega t\,, \tag{2.154}$$

we would obtain a plot like in Figure 20, lower left. In the inverse domain, we would find five peaks at the corresponding frequencies with decreasing powers (since the amplitudes are decreasing with increasing frequency in Equation 2.154), see Figure 20, lower right. In principle, we can have any signal $f(t)$ in the time domain (not necessarily a periodic one) that can be transformed into a corresponding signal in the frequency domain $F(\omega)$. The signal in both domains contains the same information.
What works in the time coordinate also works with spatial coordinates. For simplification, let us only consider the two spatial coordinates x and y. The corresponding frequencies, the *spatial frequencies* are usually denoted as k_x and k_y, respectively. Since the spatial coordinates have the units $length$, the spatial frequencies have the unit $1/length$. The

[16] Jean Baptiste Joseph Fourier, 1768 - 1830
[17] However, I recommend [1] for further reading for devoted students.

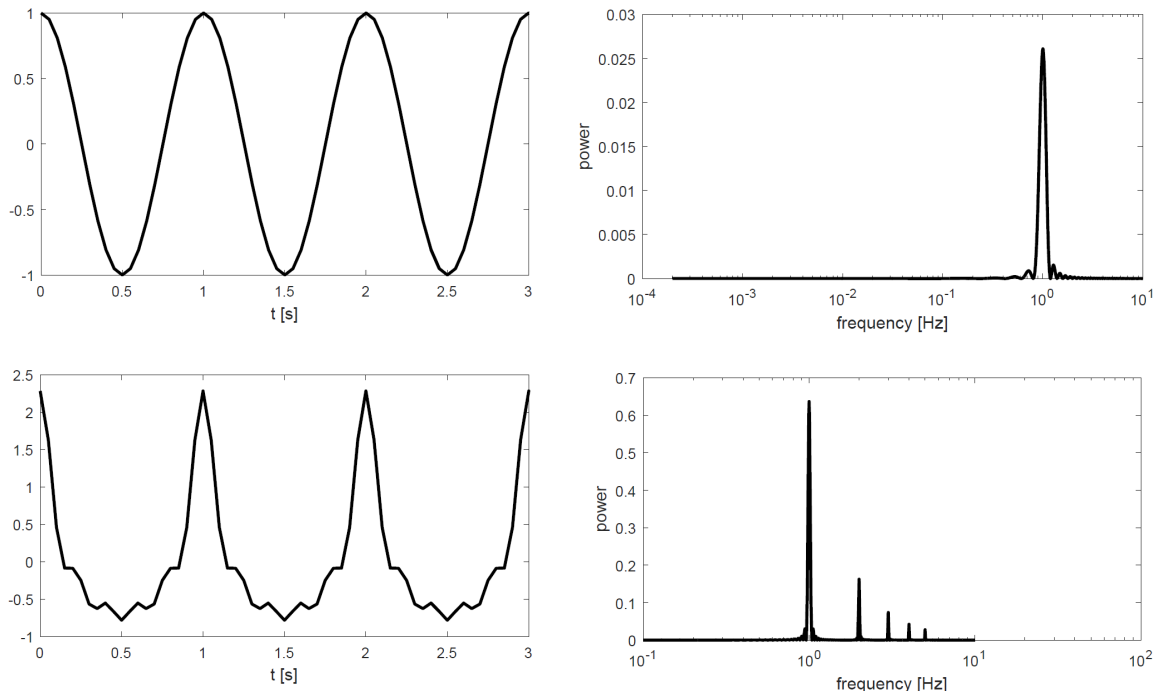

Figure 20: A cosine wave in the time domain (upper left) and its equivalent in the corresponding inverse space (here the frequency domain; upper right). A signal composed out of five superimposed cosine waves in the time domain is shown in the lower left with the corresponding plot in the inverse domain (lower right) exhibiting the five frequency peaks.

corresponding cosine signal would be something like a corrugated sheet. If we would look from the top, we would see a wavy structure in one particular direction (say x) whereas the the surface remains unchanged in the other direction (y, for any constant x). The amplitude would correspond to the height of the corrugates. The corrugated sheet as viewed from above is shown in Figure 21, top left.

Since there is a periodic change of constant frequency of the surface features along x direction, it corresponds to one peak along the k_x axis in the inverse space. There is also a second peak in negative direction and a third peak in the coordinate origin (that we will discuss later). Since there is no change along the y direction, there is no corresponding feature on the k_y axis. These three peaks in the $k_x - k_y$ plane are three dots when viewed from above (Figure 21, top right). If the corrugates repeat with a smaller period, the frequency is higher and therefore, the dots in the $k_x - k_y$ plane appear at higher frequency (Figure 21, b)). We could also turn the corrugate sheet in any direction that would lead to a rotation of the dots in the $k_x - k_y$ plane (Figure 21, c) and d)).

In principle, we could reconstruct the images a) - d) on the left in Figure 21 from the images on the right in this figure. But there is one information that is missing in these examples: it is the **phase**. If we would shift (not turn!) the images on the left in Figure 21 by a certain angle ϕ_0 (for example having a sine instead of a cosine, corresponding to a shift of $\phi_0 = 90°$), we would still get the same images on the right. Vice verse, for a given image on the right in Figure 21, we would not obtain an unique corresponding image on the left. How the phase information is in fact recovered is explained in Section 2.4.2.

Let us discuss further examples before exploring the mathematical background. Imagine an image of spherical cells, all of approximately the same diameter (Figure 22, upper left). These cells would all represent the same spatial frequency, since they are equal in size. Since the cells are spherically, the spatial frequency is independent from the direction,

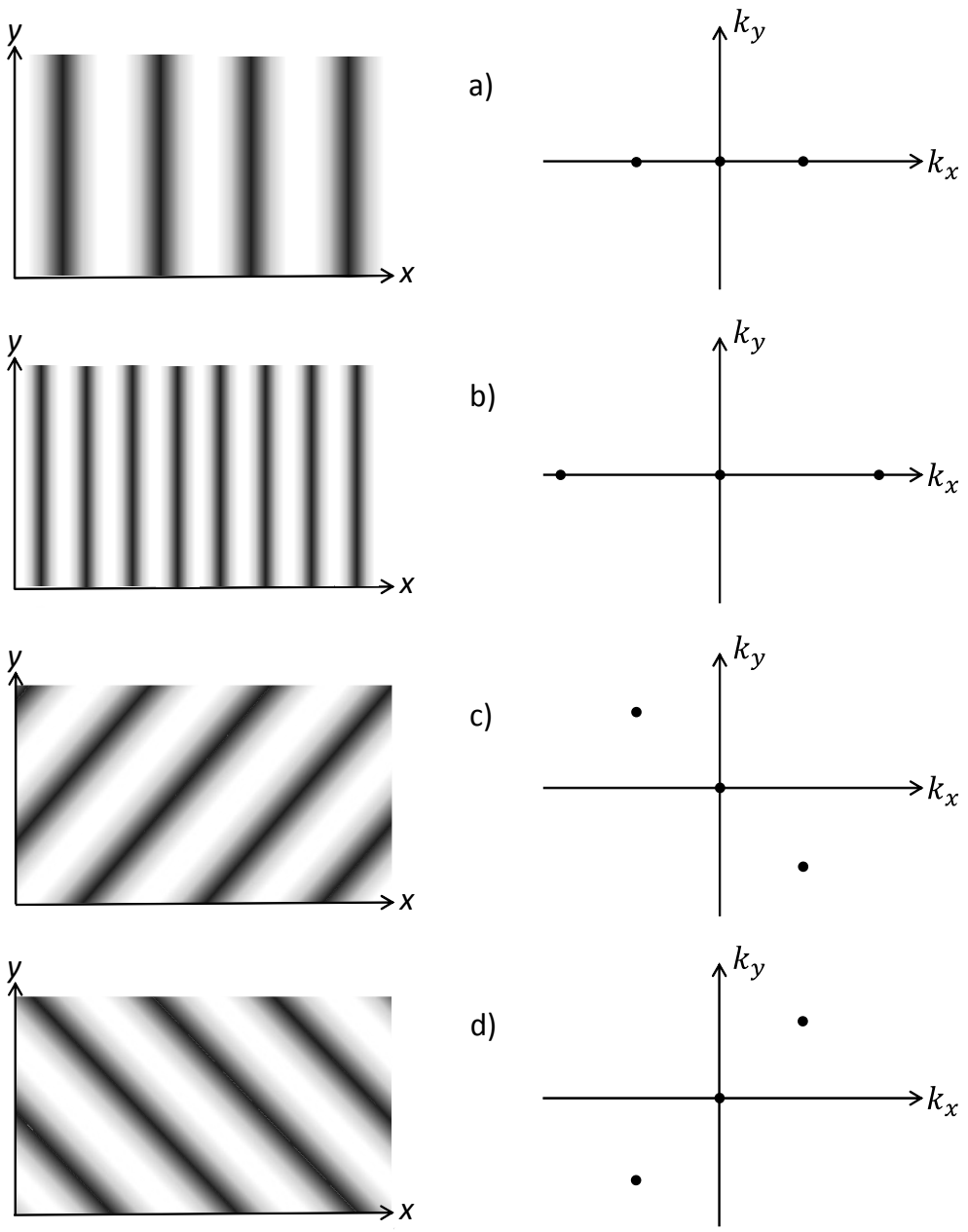

Figure 21: Periodic structures like a corrugated sheet in the spatial domain (left) and the same object shown in the inverse domain of spatial frequencies (right). The colour indicates the amplitude (left) and the power of the frequencies (right), respectively. Please try to understand qualitatively how we get from the plots on the left column to the corresponding plots on the right column (see also in the text).

i.e. the absolute value of their spatial frequency k is constant for all k_x and k_y, i. e. $k = \sqrt{k_x^2 + k_y^2} = const.$ This is just the equation of a circle of radius k centered at the coordinate origin and we therefore see one single circle in the inverse domain (Figure 22, upper right). If the cells had two distinctive diameter (Figure 22, middle left), the image in the inverse domain would exhibit two circles of the corresponding k (Figure 22, middle right). Since the conservation of the phase information is not included yet, any spatial arrangement of these cells would yield the same circle in the inverse space.
Of course, a given image is not that simple, but would look more like the one in Figure 22,

lower left. Therefore, also the image in the inverse domain is more complicated since many spatial frequencies are present in the original image.

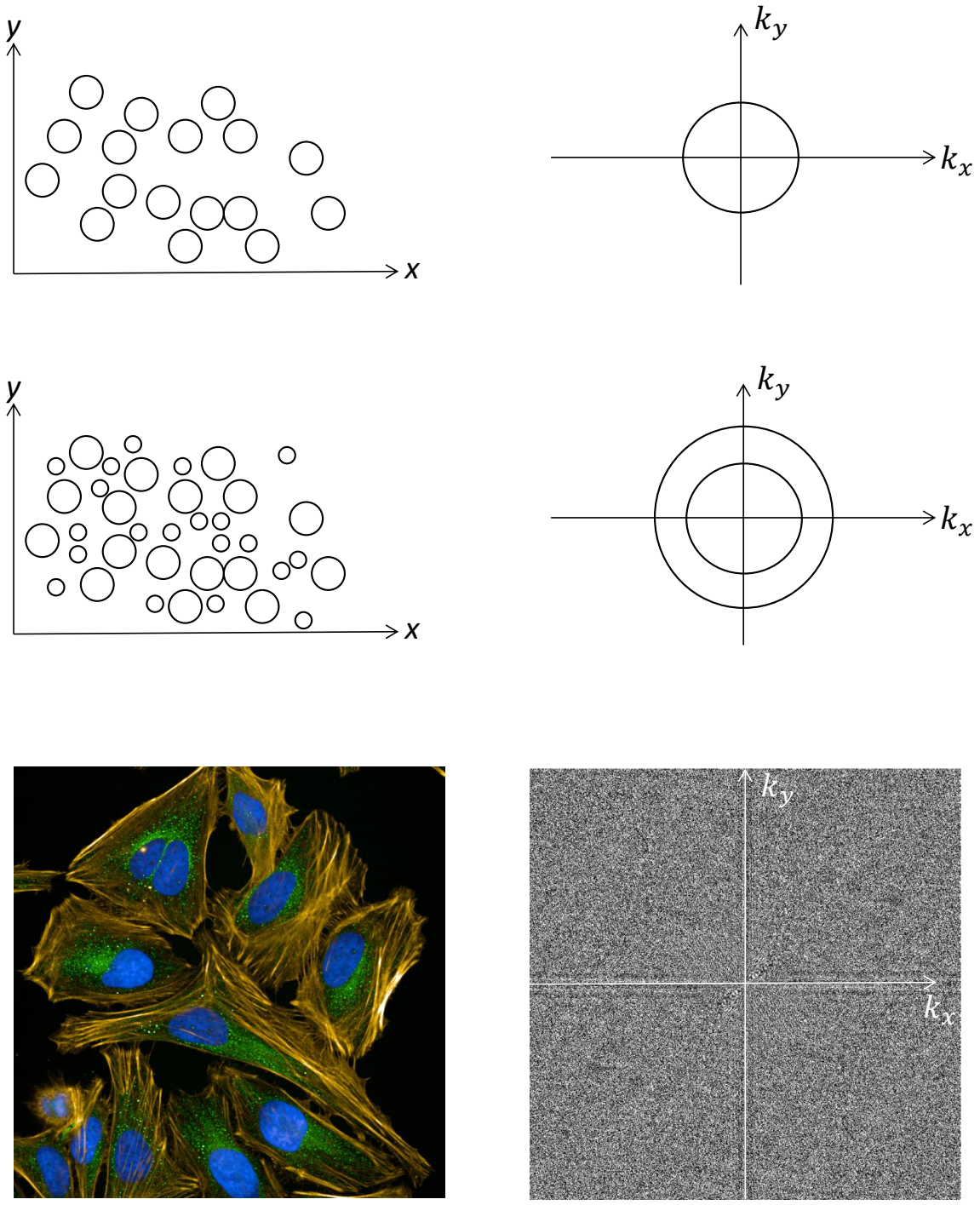

Figure 22: Images in the spatial domain (left) and the corresponding images in the inverse domain (right). Courtesy of the image in the lower left: Christophe Jung.

2.4.2 The Math

The qualitative visualization in the previous subsection is helpful for understanding the intensities/powers of the frequencies in the inverse space. However, as already mentioned,

we have to conserve the phase in order to fully reconstruct the images. This problem is known in practice e. g. as "phase problem" in x-ray crystallography (Section 9.2) that is solved by staining of the specimen with heavy metal atoms. In order to understand the problem, we have to understand the underlying mathematical concept to a certain extend. Suppose, we only know the signal in the time domain $f(t)$ but want to infer the frequencies and their individual contributions (hence their power) to the signal. If we assume, that $f(t)$ is composed out of the frequencies, we can express it similar to Equation 2.154 as

$$f(t) = \frac{a_0}{2} + \sum_{\omega=0}^{\infty} \left[a_\omega \cos(\omega t) + b_\omega \sin(\omega t)\right] \tag{2.155}$$

with a possible offset $\frac{a_0}{2}$. We do not know the values of the a_ω and b_ω yet. If one or more of these coefficients equal zero, the particular frequencies ω have no contribution to the entire signal $f(t)$. As a simplification and in order to illustrate the method, we only allow integer products of a fundamental frequency ω_0 (similar to the frequencies in Equation 2.154) as possible frequencies $\omega = \{1\omega_0, 2\omega_0, \ldots, n\omega_0, \ldots\}$ for now, but we will extend the ansatz to a continuous spectrum of frequencies later on.
We know from Section 2.3.1 that the sine and cosine functions are actually only two parts of the more general complex exponential. Therefore, we express Equation 2.155 in terms of $e^{i\phi}$ with the phase $\phi = n\omega_0 t$ using the identities (Equation 2.139b and Equation 2.139c)

$$\begin{aligned} \cos\phi &= \frac{e^{i\phi} + e^{-i\phi}}{2} \\ \sin\phi &= \frac{e^{i\phi} - e^{-i\phi}}{2i}\,. \end{aligned} \tag{2.156}$$

Thus, Equation 2.155 reads

$$\begin{aligned} f(t) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \left(\frac{e^{i\omega_0 nt} + e^{-i\omega_0 nt}}{2}\right) + b_n \left(\frac{e^{i\omega_0 nt} - e^{-i\omega_0 nt}}{2i}\right) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \underbrace{\left(\frac{a_n}{2} + i\frac{b_n}{2}\right)}_{c_n} e^{i\omega_0 nt} + \underbrace{\left(\frac{a_n}{2} - i\frac{b_n}{2}\right)}_{\bar{c}_n} e^{-i\omega_0 nt} \\ &= \sum_{n=-\infty}^{\infty} c_n e^{i\omega_0 nt}\,. \end{aligned} \tag{2.157}$$

This equation is exactly the same as Equation 2.155, but now expressed with complex numbers. Compared to Equation 2.155, the index n in the sum now does not run from 0 to ∞, but from $-\infty$ to $+\infty$, hence we also allow *negative frequencies*. Technically, this is introduced by the indexing when counting two addends in an expression like $e^{i\omega_0 nt} + e^{-i\omega_0 nt}$, where the index runs from 0 to ∞, at once by running the index from $-\infty$ to $+\infty$. Physically, a negative index can be understood for example as a negative spatial frequency k_x and/or k_y, like in Figure 21 and Figure 22.
The coefficients c_n are complex numbers now and are a measure of the contribution of the frequency $\omega_n = \omega_0\, n$ to the entire signal. Since the c_n are complex, they host the phase information. The c_n of the same absolute value (hence, amplitude or power, i.e. that what we see in the images Figure 20 to Figure 22, right) can have different phase angles in the complex plane. If we knew the c_n, we would be able to reconstruct the signal in the inverse space. But how can we derive the c_n?

Fortunately, there exists the mathematical trick found by Fourier: Suppose, we want to derive the amplitude c_m of a particular frequency m from the signal $f(t)$. The trick is to first multiply Equation 2.157 by $e^{-i\omega_0 mt}$:

$$\begin{aligned} e^{-i\omega_0 mt} f(t) &= e^{-i\omega_0 mt} \sum_{n=-\infty}^{\infty} c_n e^{i\omega_0 nt} = \sum_{n=-\infty}^{\infty} c_n e^{i\omega_0 (n-m)t} \\ &= \cdots + c_{n=m-2} \cdot e^{i(-2\omega_0)t} + c_{n=n-1} \cdot e^{i(-1\omega_0)t} \\ &+ c_{n=m} \cdot 1 \\ &+ c_{n=m+1} \cdot e^{i(+1\omega_0)t} + c_{n=m+2} \cdot e^{i(+2\omega_0)t} + \ldots \\ &= c_{n=m} + \sum_{n=-\infty;\, n\neq m}^{\infty} c_n e^{i\omega_0 (n-m)t} \end{aligned} \tag{2.158}$$

and second to integrate over the period $T = \frac{2\pi}{\omega_0}$:

$$\int_0^T f(t)\, e^{-i\omega_0 mt}\, dt = \int_0^T c_{n=m}\, dt + \sum_{n=-\infty;\, n\neq m}^{\infty} c_n \int_0^T e^{i\omega_0 (n-m)t}\, dt\,. \tag{2.159}$$

The first addend on the rhs in this equation yields $c_m\, T$ since the coefficients are no functions of time. The second addend contains the integral over the exponential wrt time that can be performed with basic algebra

$$\int_0^T e^{i\omega_0 (n-m)t} dt = \frac{1}{i\omega_0\,(n-m)} \left[e^{i\omega_0 (n-m)T} - 1 \right]\,. \tag{2.160}$$

Both, n and m were introduced as integers, hence also their difference $n-m$ is an integer b. Thus, by including the definition of the period T, the exponential reads

$$e^{i\omega_0 (n-m)T} = e^{2\pi i\, b}\,. \tag{2.161}$$

On the other hand according to Euler's relation, $e^{2\pi i\, b} = \cos(2\pi\, b) + i \sin(2\pi\, b) = 1$ for any integer b. Thus, the second addend in Equation 2.159 vanishes, so that we finally find

$$\boxed{c_m = \frac{1}{T} \int_0^T f(t)\, e^{-i\omega_0 mt}\, dt}\,. \tag{2.162}$$

Given a signal $f(t)$, we can derive the power c_m of any frequency $\omega_0\, m$ by performing the above integral. This operation is called *Fourier analysis*, since we analyse the signal $f(t)$ by its frequencies. The reverse process, hence composing (i.e. synthesizing) a signal $f(t)$ from its frequencies by a sum

$$\boxed{f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\omega_0 nt}} \tag{2.163}$$

is called *Fourier synthesis*. Both, Fourier analysis and Fourier synthesis are together the *Fourier transformation*. The Fourier analysis maps the signal into the inverse domain (from time to frequency) and therefore the inverse domain is often called *Fourier space*. The Fourier synthesis maps the signal in frequency space back to the signal in the time domain. Let us now investigate the case $m = 0$. Equation 2.162 yields then

$$c_0 = \frac{1}{T} \int_0^T f(t)\, dt\,, \tag{2.164}$$

that is the average of the signal over the period T. The sub case $m = 0$ equals zero frequency, hence an infinitely large period. Since any signal is periodic for an infinite time span, we always get a solution for $m = 0$ or for the spatial equivalents $k_x = k_y = 0$. That is the reason why we always obtain a dot in the coordinate origin in the frequency space (e. g. Figure 21 and Figure 22, right). On the other hand, for a signal in the time domain, we would have to measure an infinite time span to obtain all (lower) frequencies. In fact, for measuring over a time span Δt, we can only find periods up to $T_{max} < \frac{\Delta t}{2}$. Interrupting a measurement mimics the end of the signal and therefore causes artificial peaks (so called "aliases") in the frequency domain, see Figure 23. In practice, the signal $f(t)$ is not measured continuously, but with a certain sampling rate ρ (e. g. one measurement every two seconds etc.). Thus, the maximum frequency f_{max} that can be measured is half of the sampling rate $f_{max} < \frac{\rho}{2}$. For images, the sampling rate corresponds to the spatial resolution given by the pixel size of the image and therefore the largest spatial frequency (hence the smallest feature) that can be resolved is larger than two pixel (see Section 9.1.8 and Section 9.3.3). Thus, the features resolved in an image are located within a ring centered in the coordinate origin in the k_x - k_y plane. Removing information within this ring, close to the edge in the Fourier image results in a blur of the original image (Figure 24).

Equation 2.162 and Equation 2.163 are the Fourier transformations between time domain

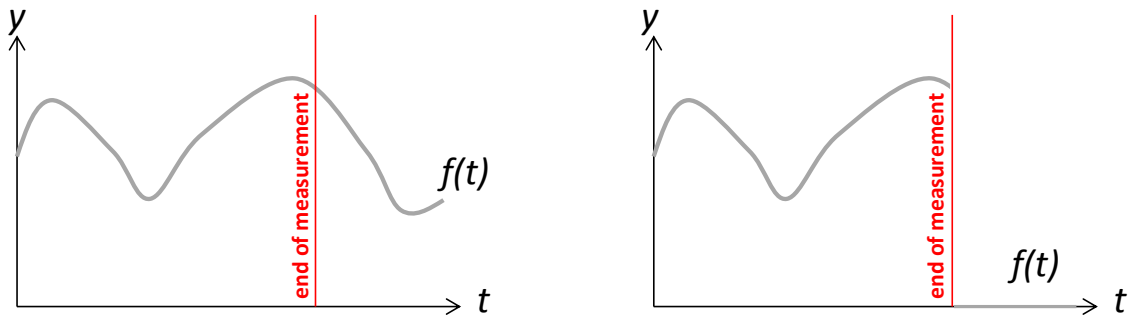

Figure 23: Interrupting the measurement of a signal mimics a sudden end of the signal itself and therefore causes artificial peaks (so called "aliases") in the frequency domain.

and frequency domain. Since it does not play a role if the frequency is a temporal frequency (unit *Hertz*) or a spatial frequency k (unit $1/length$), we can immediately write down the equivalent equations for a spatial Fourier transformation. The only differences are that we have three spatial directions $\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ (instead of one time direction) and therefore also three spatial frequencies $\vec{k} = \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix}$ and a volume V instead of the period T so that Equation 2.162 and Equation 2.163 read

$$\boxed{c_{\vec{k}} = \frac{1}{V} \int f(\vec{r})\, e^{-i\,\vec{k}\cdot\vec{r}}\, dV} \tag{2.165}$$

and

$$\boxed{f(\vec{r}) = \sum_{k_z=-\infty}^{\infty} \sum_{k_y=-\infty}^{\infty} \sum_{k_x=-\infty}^{\infty} c_{\vec{k}}\, e^{i\,\vec{k}\cdot\vec{r}}}\,. \tag{2.166}$$

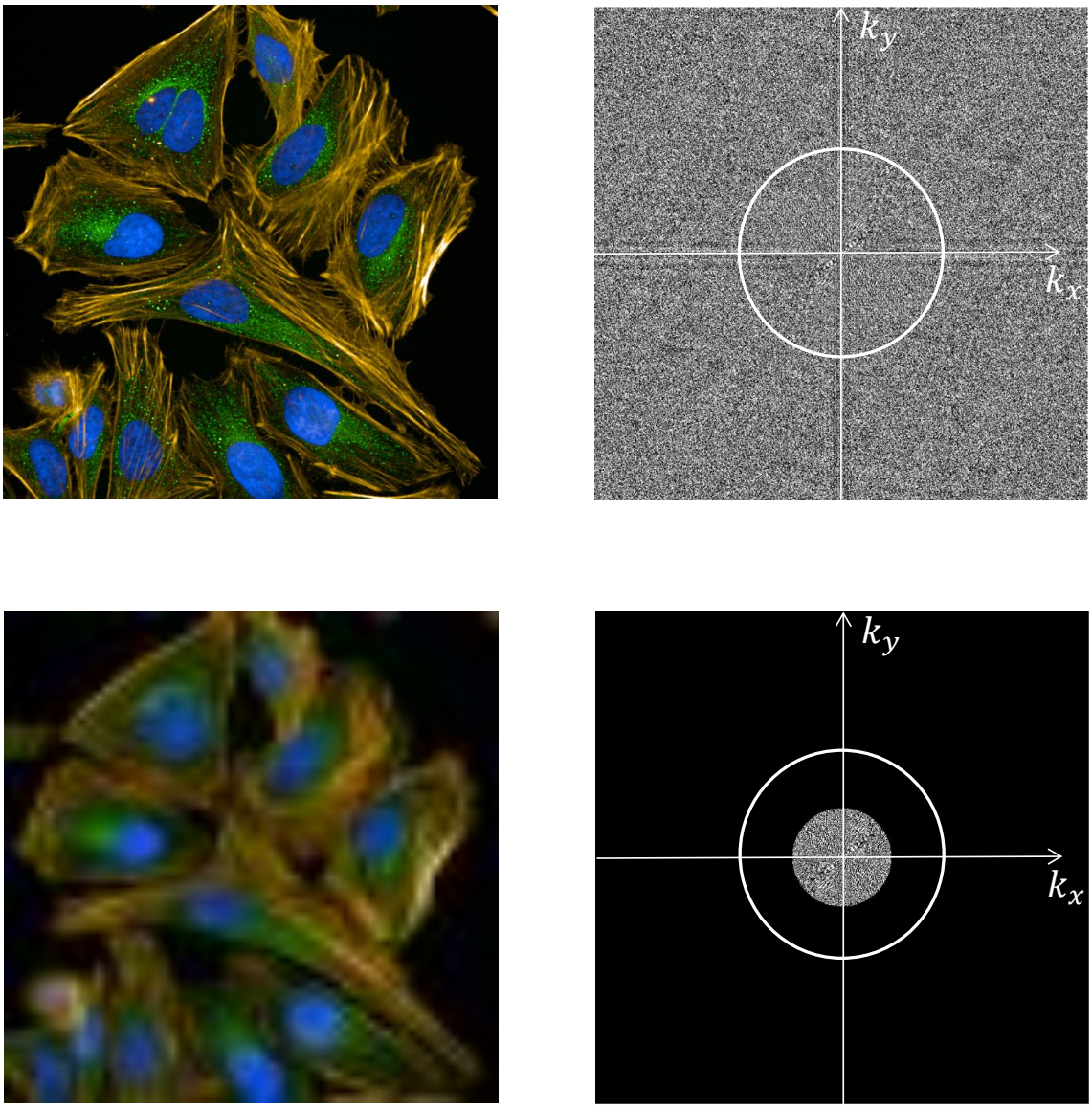

Figure 24: Clockwise from upper left: Original image and the corresponding image in inverse space (Fourier space). The white ring indicates the spatial frequency $k = \sqrt{k_x^2 + k_y^2}$ that corresponds to twice the pixel size in the original image (everything within the ring is resolved).
Removing higher spatial frequencies (setting pixel values to zero) in the Fourier image results in removing smaller features in the actual image that is visible as blur.

We will encounter these equations again in Section 9.2, where V corresponds to the volume of an unit cell and the equivalent of $c_{\vec{k}}$ is called *structure factor*.

The c_n and c_k are the coefficients of a discrete frequency spectrum (that for example would e. g. correspond to discrete pixels in an image). In general, the frequency spectrum can be also continuous so that c_k and c_n are continuous functions G and not a set of coefficients. In such a case, the sums in Equation 2.163 and Equation 2.166 become integrals. Therefore, the Fourier analysis for a frequency ξ reads

$$\boxed{G(\xi) = \int_{-\infty}^{\infty} g(x)\, e^{-i\,x\cdot\xi}\, dx} \tag{2.167}$$

and the Fourier synthesis has the form

$$\boxed{g(x) = \int_{-\infty}^{\infty} G(\xi)\, e^{i\,x\cdot\xi}\, d\xi}\,, \tag{2.168}$$

where x and ξ can be vectors and the integral is executed over the N dimensional volume element dx and $d\xi$, respectively.

2.4.3 The Saw Tooth

Let us now explicitly apply Equation 2.162 and Equation 2.163 to a simple mathematical example. Unfortunately, the examples of biological relevance are too complicated and should be only done on the computer (c. f. Section 7.2 for one of the simplest biological example). One of the cases where a straight forward calculation is feasible is the *saw tooth* function shown in Figure 25 or functions of a similar structure. According to Equation 2.163 these functions are signals in the time domain that can be represented as a sum of sines and cosines.

The saw tooth function in Figure 25 has a period T and an amplitude A and is repetitive

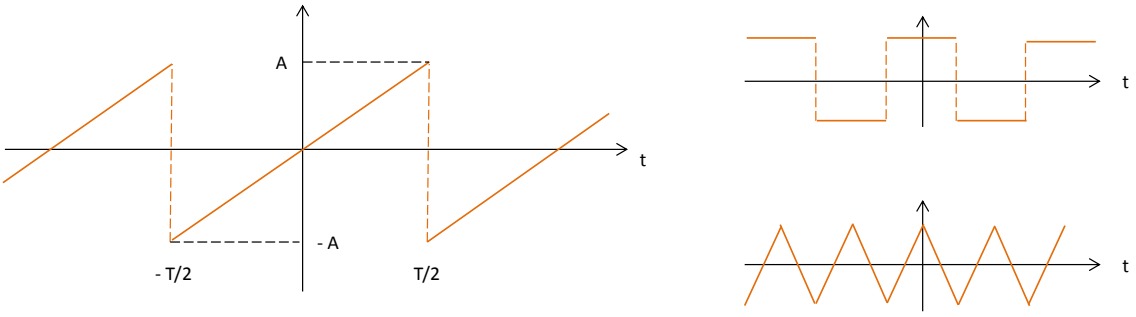

Figure 25: The saw tooth (left) of period T and amplitude A and two other typical periodical functions constructed by a series of trigonometric functions (right).

in the interval $[-T/2,\, T/2]$. Shifting this interval by an integer multiplied with T along the t axis reproduces the same function. Within the intervals, the saw tooth $f(t)$ is an ordinary linear function and we can write

$$f(t) = \frac{2A}{T} t\,. \tag{2.169}$$

According to Equation 2.162, the powers c_n of the frequencies n can be derived by integrating the Fourier transformed within the interval and therefore,

$$\begin{aligned} c_n &= \frac{1}{T}\int_{-T/2}^{T/2} f(t)\, e^{-i\omega t}\, dt = \frac{1}{T}\int_{-T/2}^{T/2} \left(\frac{2A}{T} t\right)\, e^{-i\omega t}\, dt \\ &= \frac{2A}{T^2}\int_{-T/2}^{T/2} t\, e^{-i\omega t}\, dt\,. \end{aligned} \tag{2.170}$$

The integral

$$I := \int_{-T/2}^{T/2} t\, e^{-i\omega t}\, dt \tag{2.171}$$

can be solved by using the product rule (Section 2.1.7) and we find

$$I = \frac{t}{-i\omega}\, e^{-i\omega t}\Big|_{-T/2}^{T/2} - \int_{-T/2}^{T/2} \left(\frac{1}{-i\omega}\right)\, e^{-i\omega t}\, dt$$

$$= \frac{iT}{2\omega}\left(e^{-i\omega T/2} + e^{i\omega T/2}\right) + \frac{1}{\omega^2}\left(e^{-i\omega T/2} - e^{i\omega T/2}\right)$$
$$= \frac{iT}{2\omega}\underbrace{2\cos\left(\frac{\omega T}{2}\right)}_{Equation\ 2.126} + \frac{1}{\omega^2}\underbrace{(-2)\,i\sin\left(\frac{\omega T}{2}\right)}_{Equation\ 2.126}. \tag{2.172}$$

The angular frequency itself is a function of n. As before, we assume a discrete spectrum where $\omega = \omega_0\, n$ with the ground frequency $\omega_0 = \frac{2\pi}{T}$. Thus, the argument in the trigonometric functions equals $\pi\, n$. Therefore, the $\sin$ terms disappear, whereas the $\cos$ terms equal -1 for odd n and $+1$ for even n and we finally find that

$$c_n = \frac{2Ai}{T\omega}\,(-1)^n = \frac{Ai}{\pi n}\,(-1)^n\,. \tag{2.173}$$

The coefficients c_n are all imaginary and decrease with $\frac{1}{n}$.
According to Equation 2.163, the signal must obey the equation

$$f(t) = -\sum_{n\neq 1}\frac{Ai}{\pi n}\,(-1)^n\; e^{i\omega\, t}\,, \tag{2.174}$$

where n runs from $-\infty$ to $+\infty$. We might worry about the case $n = 0$. According to Equation 2.170, $n = 0$ yields

$$c_0 = \frac{1}{T}\int_{-T/2}^{T/2}\left(\frac{2A}{T}t\right)\, e^0\, dt = 0\,, \tag{2.175}$$

hence, everything is consistent.
We use Eulers identity again for Equation 2.174 and use $i^2\,(-1)^n = (-1)^{n+1}$ and find that

$$f(t) = \frac{Ai}{\pi}\sum_{n\neq 1}\frac{(-1)^n}{n}\cos\left(\frac{2\pi t}{T}n\right) + \frac{A}{\pi}\sum_{n\neq 1}\frac{(-1)^{n+1}}{n}\sin\left(\frac{2\pi t}{T}n\right)\,. \tag{2.176}$$

The $\cos$ is an even function and for negative n we obtain terms of the from $-\cos(-x) = -\cos(x)$ whereas positive n will yield terms with $+\cos(x)$ and therefore the first sum in the above equation vanishes. The $\sin$ is an odd function ($-\sin(x) = sin(-x)$). Consequently, negative n *and* positive n yield the same terms and we are allowed to double the amplitudes if n runs only over positive integers:

$$\boxed{f(t) = \frac{A}{\pi}\sum_{n\neq 1}\frac{(-1)^{n+1}}{n}\sin\left(\frac{2\pi t}{T}n\right) = \frac{2A}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\left(\frac{2\pi t}{T}n\right)}\,. \tag{2.177}$$

The result seems surprising. The above equation states that the saw tooth is equal to a sum of $\sin$ functions. Indeed, plotting the sum of the first addends in Equation 2.177 leads to the saw tooth signal (Figure 26, left). The coefficients c_n (Equation 2.173) are shown in Figure 26, right.
We even recover that the amplitude of the saw tooth A is roughly three times larger than the amplitude of the $\sin$ (red line in Figure 26, left), caused by the factor $\frac{1}{\pi}$ in Equation 2.177.
The exercise we did here is exactly the same procedure performed by the computer when creating a Fourier image or when analyzing the frequencies of an arbitrary signal $f(t)$ - or when reconstructing the crystal structure of a protein from X-ray diffraction pattern[18]. It is

[18]When I was a student, we had to do this by hand for simple crystals like potassium chloride or iron sulfate in the lab courses. A tedious work, which however was a good exercise.

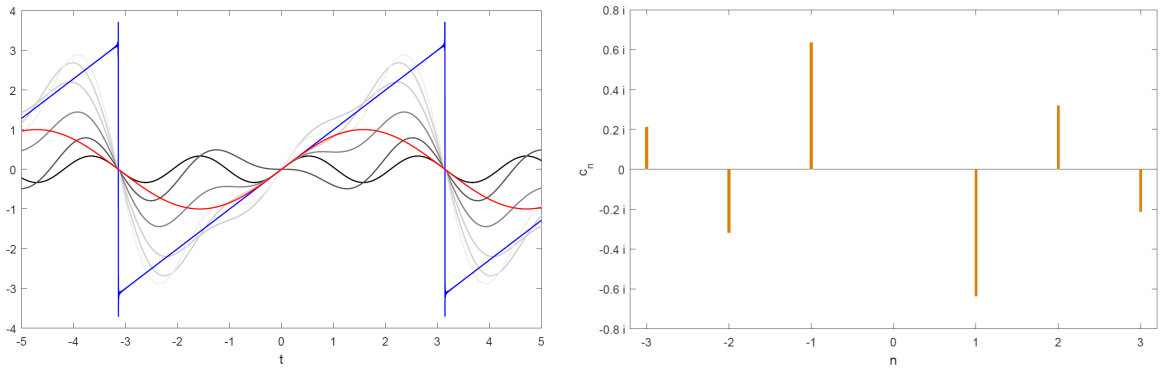

Figure 26: **Left:** The sine wave (red) and the first three positive and negative orders of the sum from Equation 2.177 (gray). The saw tooth clearly emerges after summing all orders from $n = -3\,000$ to $n = +3\,000$ (blue).
Right: The power spectrum c_n of the first three positive and negative orders. Note that all c_n are imaginary (Equation 2.173).

remarkable that this technique was developed by Fourier around the year 1800 when visualizations as used here (Figure 26) as a helpful tool were completely beyond any imagination.

References

[1] Brad Osgood, *"Lecture Notes: The Fourier Transformation and its Applications"*, Electrical Engineering Department, Stanford University.

2.5 From Carthesian to other Coordinate Systems

So far we did all our calculations in the Cartesian coordinate system with the orthogonal, non-curved coordinates $\vec{x}$, $\vec{y}$ and $\vec{z}$. This is most appropriate if we want to do calculations with quantities that are also orthogonal and non-curved. For example the volume of a cube $\int dV = \int\int\int dxdydz = x\,y\,z$ can be easily calculated. But sometimes the objects of interest are curved and the coordinates can be even non orthogonal. Already in Section 2.3 we switched to polar coordinates in order to calculate roots of complex numbers. In Section 6.3.1 we calculate the diffusion into a spherical cell in spherical coordinates and in Section 8.3 we use cylindrical coordinates to calculate the volume flow through a cylindrical micro channel and we use spherical coordinates in Section 9.3.5 for 3D reconstruction. Thus, it is more appropriate to adapt the coordinate system to the geometry of the problem. In most of the cases, we will need polar coordinates, spherical coordinates or cylindrical coordinates (Figure 27).

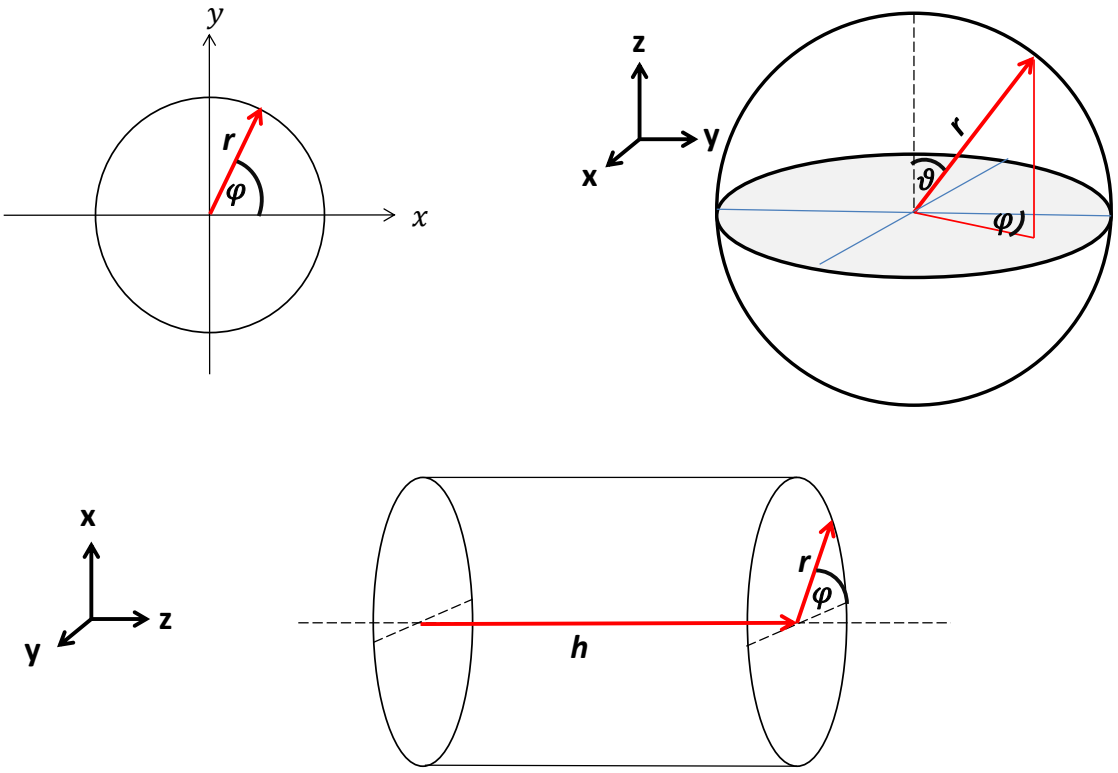

Figure 27: Transformation from Cartesian coordinates to polar coordinates (upper left), spherical coordinates (upper right) and cylindrical coordinates (lower panel).

2.5.1 Polar Coordinates

The two dimensional Cartesian coordinates are usually denoted as $\vec{x}$ and $\vec{y}$. These coordinates describe the location of a point $P(\vec{x}, \vec{y})$, the area $dA = dx\,dy$, the length (so called *line element*) $ds^2 = dx^2 + dy^2$ and also operators like gradient and divergence. We now ask for their equivalents in polar coordinates.
In Section 2.3 we found already that a point in the x - y plane can also be described by

the angle ϕ and radius r, so that we find (see also Figure 27) the relations

$$x = r\,\cos\phi \tag{2.178a}$$

$$y = r\,\sin\phi \tag{2.178b}$$

We even can express Equation 2.178 in terms of the differentials by applying the chain rule and obtain

$$dx = \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \phi}d\phi = \cos\phi\,dr - r\sin\phi\,d\phi \tag{2.179a}$$

$$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \phi}d\phi = \sin\phi\,dr + r\cos\phi\,d\phi \tag{2.179b}$$

that we can write as matrix equation

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} \cdot \begin{pmatrix} dr \\ d\phi \end{pmatrix} . \tag{2.180}$$

Now, we can ask for example whether or not we are allowed to write the length in these coordinates as $ds^2 = dr^2 + d\phi^2$. Let us start with the old Cartesian coordinates and according to Equation 2.179, we can write

$$\begin{aligned} ds^2 &= dx^2 + dy^2 \\ &= (\cos\phi\,dr - r\sin\phi\,d\phi)^2 + (\sin\phi\,dr + r\cos\phi\,d\phi)^2 \\ &= \underbrace{\left(\cos^2\phi + \sin^2\phi\right)\,dr^2}_{=dr^2} + \underbrace{\left(\sin^2\phi + \cos^2\phi\right)\,r^2 d\phi^2}_{=r^2 d\phi^2} \\ &= dr^2 + r^2 d\phi^2\,, \end{aligned} \tag{2.181}$$

that is not $ds^2 = dr^2 + d\phi^2$! Thus, the equivalents in the line element are actually

$$\begin{aligned} dx &\to dr \\ dy &\to r d\phi\,. \end{aligned} \tag{2.182}$$

The same applies for the volume element that consequently[19] turns into

$$dx\,dy = r dr\,d\phi\,. \tag{2.183}$$

I now like to illustrate the benefits we get from the adapted coordinate system by a simple example. Let us calculate the area of a circle of radius R, centered in the coordinate origin, first in Cartesian coordinates and then in polar coordinates. The radius equals $R = \sqrt{x^2 + y^2}$ and we have to calculate the surface under the curve $y = \sqrt{R^2 - x^2}$ from $x_1 = -R$ to $x_2 = R$. This is half the circle, so that we have to multiply the result with 2. Hence, we have to perform the following steps

$$\begin{aligned} A = \int dA &= \int\int dy dx \\ &= 2\int_{-R}^{R} \sqrt{R^2 - x^2}\,dx\,. \end{aligned} \tag{2.184}$$

[19] Note that it is actually not *that* simple but I like to omit some mathematical details for the sake of simplification.

The integral can be solved by substitution (c.f. the exercise in Section 2.1.7). The result is

$$2\int_{-R}^{R}\sqrt{R^2-x^2}\,dx = x\sqrt{R^2-x^2} + R^2\,\arctan\left(\frac{x}{\sqrt{R^2-x^2}}\right)\Bigg|_{-R}^{R}. \tag{2.185}$$

For both limits $x_1 = -R$ and $x_2 = R$ the square root disappears. The $\arctan$ term becomes $\arctan(\infty) = \frac{\pi}{2}$ for $x_2 = R$ and $\arctan(-\infty) = -\frac{\pi}{2}$ for $x_1 = -R$, so that we finally obtain the well known result

$$A = \pi R^2\,. \tag{2.186}$$

In polar coordinates, we can use Equation 2.183 and we have to integrate the radius from $r_1 = 0$ to $r_2 = R$ and around the full angle from $\phi_1 = 0$ to $\phi_2 = 2\pi$, so that we obtain the result within one line

$$A = \int_0^{2\pi}\int_0^R r dr\,d\phi = \int_0^{2\pi}\frac{R^2}{2}\,d\phi = \pi R^2\,. \tag{2.187}$$

Obviously, this calculation is much shorter and simpler (we avoided the complicated integral and the need to find a proper substitution).

2.5.2 Spherical Coordinates

When we switch from polar coordinates to spherical coordinates, we have to introduce a new angle ν that describes the height of a point $P(\vec{x},\vec{y},\vec{z})$ along the z axis. The projection of a point with the radial coordinate r on the z axis is then proportional to $\cos\nu$ and a projection on the x - y plane is proportional to $\sin\nu$ (Figure 27, upper right), so that Equation 2.178 turn into

$$x = r\,\cos\phi\,\sin\nu \tag{2.188a}$$

$$y = r\,\sin\phi\,\sin\nu \tag{2.188b}$$

$$z = r\,\cos\nu\,. \tag{2.188c}$$

We start again with deriving the line element $ds^2 = dx^2 + dy^2 + dz^2$ and apply the chain rule to Equation 2.188. This leads to

$$dx = \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial\phi}d\phi + \frac{\partial x}{\partial\nu}d\nu = \cos\phi\sin\nu\,dr - r\,\sin\phi\sin\nu\,d\phi + r\,\cos\phi\cos\nu\,d\nu \tag{2.189a}$$

$$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial\phi}d\phi + \frac{\partial y}{\partial\nu}\,d\nu = \sin\phi\sin\nu\,dr + r\,\cos\phi\sin\nu\,d\phi + r\,\sin\phi\cos\nu\,d\nu \tag{2.189b}$$

$$dz = \frac{\partial z}{\partial r}dr + \frac{\partial z}{\partial\phi}d\phi + \frac{\partial z}{\partial\nu}d\nu = \cos\nu\,dr + \;0\; - r\,\sin\nu\,d\nu \tag{2.189c}$$

and finally, we find the line element expressed in spherical coordinates

$$ds^2 = dr^2 + r^2\,d\nu^2 + r^2\sin^2\nu\,d\phi^2\,. \tag{2.190}$$

Comparing the line element written in Cartesian coordinates to the line element written in spherical coordinates, we find the relations

$$dx \to dr$$
$$dy \to r\,\sin\nu\,d\phi$$

$$dz \to r\, d\nu\,. \tag{2.191}$$

Therefore, the volume element turns into

$$dV = dxdydz = r^2 \sin\nu\, drd\phi d\nu\,. \tag{2.192}$$

This enables us to calculate the volume of a sphere of radius R almost within one single line. We have to integrate the radius from $r_1 = 0$ to $r_2 = R$, the angle ϕ from $\phi_1 = 0$ to $\phi_2 = 2\pi$, and $\nu_1 = 0$ to $\nu_2 = \pi$ (c.f. Figure 27, upper right) in order to cover the entire volume. Thus, the volume equals (note, that the order of integration does not matter, since R, ϕ and ν are independent from each other)

$$\begin{aligned} V = \int dV &= \int_0^{\pi}\int_0^{2\pi}\int_0^{R} r^2 \sin\nu\, drd\phi d\nu = \int_0^{\pi}\int_0^{2\pi} \frac{R^3}{3}\sin\nu\, d\phi d\nu \\ &= \int_0^{\pi} 2\pi\frac{R^3}{3}\sin\nu\, d\nu = -2\pi\frac{R^3}{3}\cos\nu\Big|_0^{\pi} = \frac{4}{3}\pi\, R^3\,. \end{aligned} \tag{2.193}$$

One can easily imagine that the same task is much more difficult to perform in Cartesian coordinates.
In the same manner, we can derive the surface area of a sphere. All points on the surface have in common that their distance to the center equals the radius R. Thus, R is constant and we do not need to integrate over this variable. Hence, the surface area can be calculated by

$$\begin{aligned} A = \int dA &= R^2 \int_0^{\pi}\int_0^{2\pi} \sin\nu\, d\phi d\nu = R^2 \int_0^{\pi} 2\pi\sin\nu\, d\nu \\ &= 4\pi\, R^2\,. \end{aligned} \tag{2.194}$$

2.5.3 Cylindrical Coordinates

A further coordinate transformation we will encounter frequently leads to cylindrical coordinates. A cylinder has a circular bottom, so that we can use the polar coordinates for x and y, and the height h, so that the z coordinate corresponds to h (Figure 27, lower panel):

$$x = r\,\cos\phi \tag{2.195a}$$

$$y = r\,\sin\phi \tag{2.195b}$$

$$z = h\,. \tag{2.195c}$$

In the same manner as before, we can calculate the line element and obtain $ds^2 = dr^2 + r^2\, d\phi^2 + dh^2$. The increments transform via

$$\begin{aligned} dx &\to dr \\ dy &\to r\, d\phi \\ dz &\to dh \end{aligned} \tag{2.196}$$

and therefore the volume of a cylinder of height H and bottom radius R equals $V = \pi R^2\, H$.

2.5.4 Gradient and Divergence Again

An important question is how the operators like gradient and divergence will transform in the new coordinates. From the above equations we can already infer, that it is probably not possible to substitute the increments dx, dy and dz just by dr, $d\phi$ and $d\nu$ in spherical coordinates and therefore, we certainly **cannot** simply write $\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial \nu} \end{pmatrix}$ for the gradient and/or the divergence. However, what we know for sure is, that therefore the chain rule must hold in any coordinate system and that the derivative of a scalar Ψ (Section 2.1.8) can be written as

$$d\Psi = \frac{\partial}{\partial x}\Psi\, dx + \frac{\partial}{\partial y}\Psi\, dy + \frac{\partial}{\partial z}\Psi\, dz = \frac{\partial}{\partial r}\Psi\, dr + \frac{\partial}{\partial \phi}\Psi\, d\phi + \frac{\partial}{\partial \nu}\Psi\, d\nu\,. \tag{2.197}$$

On the lhs we have the gradient of Ψ, $\nabla_c\Psi$ in Cartesian coordinates and on the rhs we have the gradient of Ψ, $\nabla_s\Psi$ in spherical coordinates. Thus, the above equation can be written as

$$\nabla_c\Psi \cdot \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \nabla_s\Psi \cdot \begin{pmatrix} dr \\ d\phi \\ d\nu \end{pmatrix}\,. \tag{2.198}$$

The structure of $\nabla_s\Psi$ is yet unknown, but we know from Section 2.5.2 how the increments dx, dy and dz transform into spherical coordinates and therefore can write

$$\nabla_c\Psi \cdot \begin{pmatrix} dr \\ rsin\nu\, d\phi \\ rd\,\nu \end{pmatrix} = \nabla_s\Psi \cdot \begin{pmatrix} dr \\ d\phi \\ d\nu \end{pmatrix}\,. \tag{2.199}$$

Since $\nabla_c = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$, we see, that the above equation is true, if

$$\boxed{\nabla_s = \left(\frac{\partial}{\partial r}, \frac{1}{r\sin\nu}\frac{\partial}{\partial \phi}, \frac{1}{r}\frac{\partial}{\partial \nu}\right)}, \tag{2.200}$$

that is the gradient in spherical coordinates.
In the same manner, we can derive the gradient in polar coordinates and cylindrical coordinates.

A further important operator is the divergence (Section 2.1.6) of a vector field $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ that is defined as $\nabla \cdot \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$. While the gradient generates a vector from a scalar Ψ, we have to be aware that the divergence is the dot product of the derivatives with the vector filed and therefore we have to take the derivative of the vector component v_i **and** the unit vector $\vec{e}_i$ into account, since the unit vector will also change in the new coordinate system. For example the vector $\vec{P} = \begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix}$ is written as

$$\vec{P} = 1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 5\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1\vec{e}_x + 4\vec{e}_y - 5\vec{e}_z = \sum_i v_i\vec{e}_i \tag{2.201}$$

and the derivative of one component i will have the from $d(v_i\vec{e}_i) = v_i\, d\vec{e}_i + \vec{e}_i\, dv_i$.
If we want the vector field to be written in spherical coordinates, it must have the form

$$\vec{v} = v_r\,\vec{e}_r + v_\phi\,\vec{e}_\phi + v_\nu\,\vec{e}_\nu \tag{2.202}$$

with yet unknown unit vectors $\vec{e}_r$, $\vec{e}_\phi$ and $\vec{e}_\nu$ in spherical coordinates. The divergence operator (that is a vector too, Section 2.1.6) must have the from (c.f. Equation 2.200)

$$\nabla = \vec{e}_r\frac{\partial}{\partial r} + \vec{e}_\phi\frac{1}{r\sin\nu}\frac{\partial}{\partial\phi} + \vec{e}_\nu\frac{1}{r}\frac{\partial}{\partial\nu}\,, \tag{2.203}$$

so that the dot product, hence the divergence of a vector field $\nabla\cdot\vec{v}$ equals

$$\left(\vec{e}_r\frac{\partial}{\partial r} + \vec{e}_\phi\frac{1}{r\sin\nu}\frac{\partial}{\partial\phi} + \vec{e}_\nu\frac{1}{r}\frac{\partial}{\partial\nu}\right)\cdot\left(v_r\,\vec{e}_r + v_\phi\,\vec{e}_\phi + v_\nu\,\vec{e}_\nu\right)\,. \tag{2.204}$$

This dot product yields nine addends, where the first one is for example $\vec{e}_r\frac{\partial}{\partial r}\left(v_r\,\vec{e}_r\right)$. The unit vectors in spherical coordinates are orthogonal like the Cartesian unit vectors, so that $\vec{e}_i\cdot\vec{e}_j = 0$ for $i\neq j$ and 1 else. The remaining work is a matter of diligence, so that I only like to give the result without the lengthy algebra:

$$\boxed{\nabla\cdot\vec{v} = \frac{1}{r^2}\frac{\partial\left(r^2\,v_r\right)}{\partial r} + \frac{1}{r\sin\nu}\frac{\partial\left(\sin\nu\,v_\nu\right)}{\partial\nu} + \frac{1}{r\sin\nu}\frac{\partial v_\phi}{\partial\phi}}\,. \tag{2.205}$$

Especially in Section 6 we will need another operator that is defined as $\Delta := div\,grad = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$. Having the gradient (Equation 2.200) and the divergence (Equation 2.205) in spherical coordinates, we can also write this operator in spherical coordinates too by using its definition and executing the derivatives. We obtain

$$\boxed{\Delta = \frac{\partial}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\nu^2} + \frac{1}{r^2\tan\nu}\frac{\partial}{\partial\nu} + \frac{1}{r^2\sin^2\nu}\frac{\partial^2}{\partial\phi^2}}\,. \tag{2.206}$$

2.6 Stochastic

2.6.1 Probability Theory

In mathematics *probability theory* describes the behavior of random phenomena. The central objects are random variables, stochastic processes, and events. It is in the nature of random events, that they cannot be predicted precisely. However, we can describe patterns by using the result of the *law of large numbers* and the *central limit theorem*. In this section we try to understand the basic concepts of probability theory by a simple experiment of throwing a die.

Let us assume we have a fair die with six faces like illustrated in Figure 28.

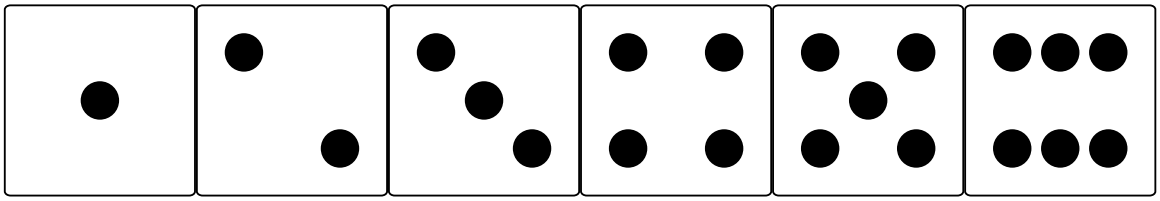

Figure 28: Illustration of a die with six faces.

Rolling the die can be understand as a *random experiment* with possible *outcomes* $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$. The whole set, that contains all realizable outcomes is called *sample space*, usually denoted as Ω. A subset $A \subseteq \Omega$ is called *event*. The symbol $\subseteq$ means "subset", including the sub case where A entirely constitutes Ω (it is somewhat the equivalent to the symbol $\leq$).
$P(A)$ is called *probability* of the event A and measures how likely it is that we observe an *outcome* being a member of A. In the example of throwing a die, events can be rolling a 3 $(A = \{3\})$, a number smaller than 4 $(B = \{\{1\}, \{2\}, \{3\}\})$, or rolling only odd numbers $(C = \{\{1\}, \{3\}, \{5\}\})$. The probability assigns every event a real number between 0 and 1.
How can the probability of rolling certain events be accessed? For the estimation of the probabilities for our experiment we make use of the *law of large numbers*, that states that for $n \to \infty$ experiments the relative frequency converges towards the probability,

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}\,, \tag{2.207}$$

where n_A denotes the absolute frequency of observing event A. In other words, it means for our example that rolling a die ten times we will observe certain fluctuations in the number of appearance, and hence of the relative frequency, of a number like "5" (or any other number on a die). But for 10^{100} rolls, the relative frequency of each number is expected to be very close to $1/6$, i. e. to its probability. In Figure 29 a die experiment is executed and it can be observed that the relative frequency converges towards the probability of 1/6. This result agrees to the probability being uniformly distributed for throwing any number in this case. Generally, there are experiments, where the probabilities are not uniformly distributed, but still the relative frequency converges towards the probability for sufficiently large numbers.
If the random variable is uniformly distributed, then the probability of observing the event A is given by,

$$P(A) = \frac{|A|}{|\Omega|}\,, \tag{2.208}$$

where $|\cdot|$ measures the number of elements in the corresponding set (i. e. the "size" of the set).

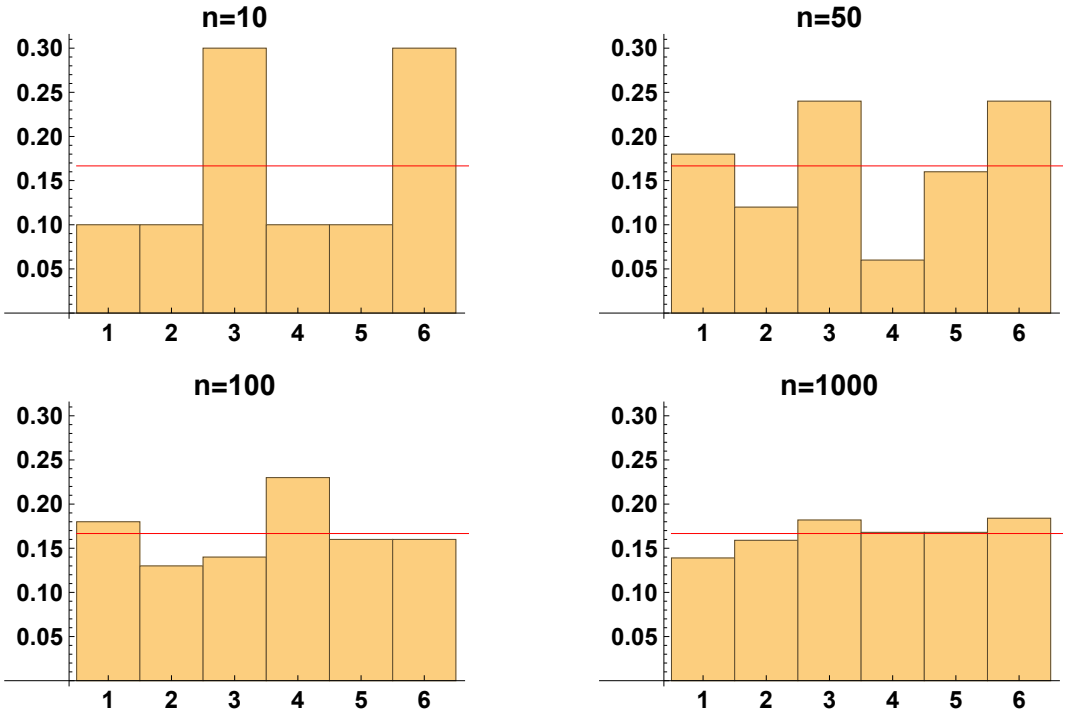

Figure 29: Experiment of rolling a die from ten to up to 1000 times. It can be seen that the relative frequency converges to the probability $1/6 \approx 0.167$.

Example I:

Let us calculate explicit probabilities for the events mentioned above by using Equation 2.208. Since $\Omega = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$*, we conclude that* $|\Omega| = 6$*. Thus, we can answer the following questions:*

1. *"What's the probability of rolling a 3"?*
   $A = \{1\} \rightarrow P(A) = \frac{1}{6}$
2. *"What's the probability of rolling a number smaller than 4"?*
   $B = \{\{1\}, \{2\}, \{3\}\} \rightarrow P(B) = \frac{3}{6} = \frac{1}{2}$
3. *"What's the probability of rolling only odd numbers"?*
   $C = \{\{1\}, \{3\}, \{5\}\} \rightarrow P(B) = \frac{3}{6} = \frac{1}{2}$

Basic operations with sets are the *union* $\cup$, *intersection* $\cap$ and *complement* $\setminus$. Suppose one set A are females and the second set B are black haired humans. The union $A \cup B$ contains all females and all black haired, including black haired males. The intersection $A \cap B$ is a set containing the elements that have both properties in common, hence black haired females. The complement $A \setminus B$ are those females that are not black haired. If an intersection yields an empty set, it is written as $A \cap B = \varnothing$. The operations are illustrated in Figure 30.

From this, we can derive some first rules, called *axioms*:

$$P(A) \geq 0, \quad A \subseteq \Omega, \tag{2.209a}$$

$$P(\Omega) = 1, \tag{2.209b}$$

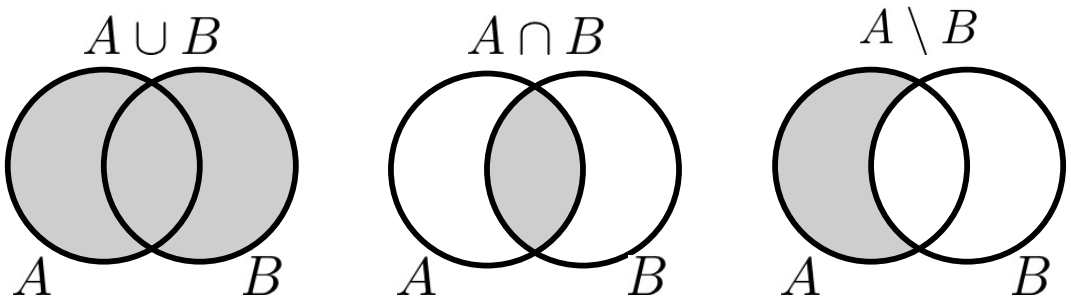

Figure 30: Basic set operations: 1. Union (left): The *union* of A and is the set off all elements of either A or B.
2. Intersection (middle): The *intersection* of A and B is the set of all elements that are both in A and B. If $A \cap B = \varnothing$ then A and B are said to be *disjoint*.
3. Complement (right): The *complement* of B in A can be seen as the "subtraction" $A - B$. In the set all elements are in A but not in B at the same time.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,. \tag{2.209c}$$

In a way, axioms follow common sense and are, by definition, not provable (in contrast to a postulate that is an ad-hoc assumption, that in principle can be proven to be incorrect). The first axiom states that the probability for an event A equals zero or lager. The second axiom states that if the set of events constitutes the entire sample space, the probability equals one. Hence, P can only have values between zero and one. The third axiom states how to calculate the probability of an event of the joint set $A \cup B$. One has to be aware, that we have to subtract $P(A \cap B)$ since we otherwise would count this subset twice.
We also state that

$$P(\varnothing) = 0 \tag{2.210a}$$
$$P(A) = 1 - P(\Omega \setminus A) = 1 - P(\bar{A})\,, \tag{2.210b}$$

hence that probability to obtain no event equals zero (first statement) and the probability for observing the complement event ("not A", $\bar{A}$) to the event A.
If for two events A and B holds $A \cap B = \varnothing$ then these events are called *mutually exclusive (ME)* or *disjoint*. For example it is not possible to get a 3 and 5 at one roll. Such events are **mutually exclusive if the occurrence of one event precludes all others**. This statement can be generalized for $i \in \mathbb{N}$ events such that it follows from Equation 2.209:

$$P(A_1 \cup A_2 \cup \ldots A_i) = P(A_1) + P(A_2) + \cdots + P(A_i) = \sum_i P(A_i) \tag{2.211}$$

If the events entirely constitute the whole sample space Ω, $A_1 \cup A_2 \cup \ldots A_i \ldots A_I = \Omega$, the sample is called *collectively exhaustive*. A consequence for ME and CE is,

$$P(A_1 \cup A_2 \cup \ldots A_i \ldots A_I) = P(A_1) + P(A_2) + \cdots + P(A_i) = \sum_{i=1}^{I} P(A_i) = 1\,. \tag{2.212}$$

It is important to keep in mind, that Equation 2.211 and Equation 2.212 are **not** valid, if the set is **not** ME!
If the events A_i are independent (IN) meaning that an event does not influence the probability to observe a subsequent event, the probability to observe A_1 **and** A_2 **and** A_3 ... (that is written $A_1 \cap A_2 \cap \ldots A_i \ldots A_I$) equals

$$P(A_1 \cap A_2 \cap \ldots A_i \ldots A_I) = P(A_1) \times P(A_2) \times \ldots P(A_i) \ldots P(A_I) = \prod_{i=1}^{I} P(A_i)\,. \tag{2.213}$$

This rule is the *Multiplication Rule*.

Example II:

A fair coin with two faces (H=head, T=tail) is flipped $N = 5$ times. What is the probability P to observe the sequence $HHTHH$?
According to Equation 2.213, $P = P(H)\,P(H)\,P(T)\,P(H)\,P(H)$. Generally, one can write for N trials and n_H times showing heads ($n_T = N - n_H$ tails since the set is ME): $P = P(H)^{n_H} P(T)^{n_T} = P(H)^{n_H} P(T)^{N-n_H}$.
For a fair coin $P(H) = P(T) = 0.5$, hence $P = (1/2)^5$.

Exercise:

Calculate P for the same situation, but for a loaded coin having the bias $P(H) = 0.2$.

Example III:

What is the probability that the events A **and** *B do* **not** *occur?*
First recall the probabilities that both events A and B occur. The probability for both events occurring is $P(AB) = P(A \cap B) = P(A)P(B)$. Consequently, the probability that A **and** *B do* **not** *occur is $P(\bar{A}\bar{B}) = 1 - P(A)P(B)$. If the events A and B are CE (then $P(B) = 1 - P(A)$), one can write $P(\bar{A}\bar{B}) = [1 - P(A)]\,[1 - P(B)]$.*
If $P(\bar{A}\bar{B}) = [1 - P(A)]\,[1 - P(B)]$ is the probability that A **and** *B do* **not** *occur, then $1 - P(\bar{A}\bar{B}) = 1 - [1 - P(A)]\,[1 - P(B)] = P(A) + P(B) - P(A)P(B)$ is the probability that either A or B occurs.*

Let me now return to the axioms (Equation 2.209): the third equation is known as *inclusion/exclusion* principle. For three sets A, B and C we write this equation as

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C) \tag{2.214}$$

and we can generalize this equation[20] even further to an infinite number of sets $A_1 \ldots A_n$, and obtain

$$P(\cup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \le i_1 < \ldots i_k \le n} P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) \right) . \tag{2.215}$$

2.6.2 Conditional Probabilities and Bayes Rules

Quite often, probabilities are not independent, but conditional and therefore require a special treatment. The *conditional probability* $P(A|B)$ measures the probability of an event A given that another event B has already occurred. $P(A|B)$ is spoken "the probability that event A occurs, given that B has occurred".
For example, the probability that *any* given person in the class likes to drink beer might

[20] Please be sure that you understand the need of the expression $P(A \cap B \cap C)$

be 25%. However, if we know that a particular person comes from Bavaria, it is much more likely that (s)he likes to drink beer, say 75%. Hence, we would write something like $P(likes\ beer) = 0.25$ and $P(likes\ beer|bavarian) = 0.75$.
The concept of conditional probabilities is very important and fundamental in probability theory. Since A and B are both subsets of Ω, it follows directly that $P(A|\Omega) = P(A)$ and $P(B|\Omega) = P(B)$. If A and B are mutually exclusive then $P(A|B) = 0$, because A and B can never occur at the same time.
If we asked for the probabilities for an event in $A \cap B$ compared to the events in B, we have to normalize by the probability for B, such that,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0\,. \tag{2.216}$$

Furthermore, $P(B)$ is called the **a priori probability** (or frequently denoted as "**prior**") and $P(B|A)$ is called **a posteriori probability**.
From the symmetry for A and B we can follow *Bayes'*[21] *Rule*:

$$P(A \cap B) = \boxed{P(A|B)P(B) = P(B|A)P(A)}. \tag{2.217}$$

According to the Bayesian notation, we define two events A and B as *independent*, if $P(A|B) = P(A)$ and $P(B|A) = P(B)$. We know already some examples for independent events such as the numbers faced up when rolling a fair die: the outcome and therefore the probability for a certain event in the second roll does not depend on the result of the first roll. The generalization for the probability of I subsequent independent events to occur leads to Equation 2.213.
Thanks to the concept of conditional probabilities, we can define the *correlation* g by

$$g = \frac{P(A|B)}{P(A)} = \frac{P(A \cap B)}{P(A)\,P(B)}\,, \tag{2.218}$$

where we find three classes of g:

- $g = 1$: no correlation, since the two events are independent and $P(A|B) = P(A)$
- $g > 1$: positive correlation,
- $g < 1$: negative correlation.

Generally, in real life one has to deal rather with conditional and correlated probabilities than with independent probabilities, for example like cooperative binding. But there are also much more trivial cases where we encounter conditional probabilities: Suppose a barrel filled with three balls, one red ball (R) and two green balls (G). In the first draw the probability of catching G (i. e. a green ball) is 2/3 and catching R is 1/3. Once a ball is caught, it is not put back into the barrel. Hence, the result of the second draw depends on the result of the first one and thus is not independent. If we would have drawn a green ball in the first trial the probability of getting a green ball in the second draw equals $P(G|G) = 1/2$. However, if we would have drawn a **red** ball in the first trial the probability of getting a green ball in the second draw is $P(G|R) = 1$ since only green balls are left after the first draw. The probability for a given result in the second (and third) draw is conditional, since it depends on the results of the preceding draws.
Thus, when expressing the probability to obtain a green ball in the second draw *under the*

[21] Thomas Bayes, 1701 - 1761

condition that we had a green ball in the first draw in Bayesian notation, we have to write $P(G|G)P(G) = 1/2 \times 2/3 = 1/3$, since the prior (the probability to obtain a ball of a particular kind in the first draw) is $P(G) = 2/3$ in this case. A *graphical model* with a *decision tree* of all possibilities and their corresponding conditional probabilities of drawing green and red balls is shown in Figure 31.

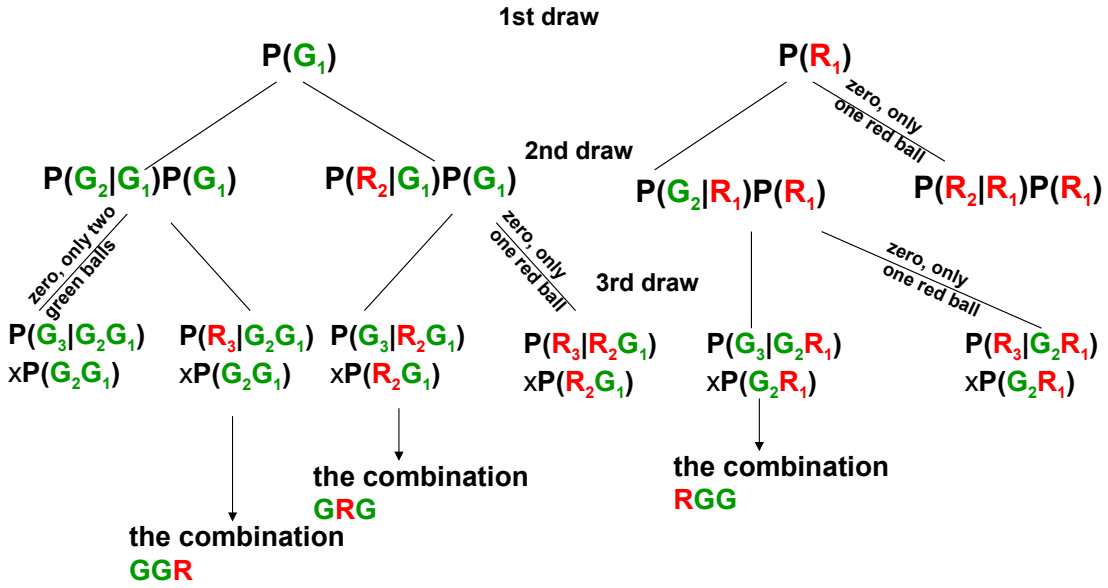

Figure 31: Decision tree for a barrel filled with one red and two green balls.

Example I:

What is the probability $P(R_3G_2G_1)$ to draw the balls in the order GGR after three draws? The probability of this combination depends on the probabilities to catch two G in the first two draws and one R after having caught two G, i.e. $P(R_3G_2G_1) = P(R_3|G_2G_1)P(G_2G_1)$. But to catch two G in the first two draws, one has to catch first one G and then a second one after the first G, i.e.

$$P(R_3G_2G_1) = P(R_3|G_2G_1)P(G_2G_1) = P(R_3|G_2G_1)P(G_2|G_1)P(G_1)\,.$$

Thus, we find that $P(R_3G_2G_1) = 1 \times 1/3 = 1 \times 1/2 \times 2/3 = 1/3$ (see also Figure 31).

Example II:

Consider a receptor with three (or more) binding sides. The first ligand has a certain prior probability $P(1) = P_*$ (c.f the example in Section 3.2.2) to bind at one of the free binding sides. The second ligand finds one binding side already being occupied by the first ligand and the next ligand finds two binding sides being occupied. In addition, the binding of the previous ligands might lead to a conformational change of the receptor that further influences the binding probabilities of the next ligands.

Example III:

Let us consider two statements:

1) *30% of all Germans will die from cancer (that is indeed the case!).*
2) *Only 10% of those having a healthy lifestyle (denoted with H) will die from cancer.*

Obviously, the chance of dying from cancer (D) depends on the life style and therefore, we can write conditional probabilities $P(D|H) = 0.1$ versus $P(D_{all}) = 0.3$. The correlation g (Equation 2.218) is $g = 0.1/0.3 = 1/3$. It is a negative correlation ($g < 1$), hence people with a healthy lifestyle die less likely from cancer.

If event B occurs depending on A there might be also a probability that B happens if A does **not** occur. Hence,

$$\begin{aligned} P(B) = P(BA) + P(B\bar{A}) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) \\ = P(B|A)P(A) + P(B|\bar{A})[1 - P(A)]\,. \end{aligned} \tag{2.219}$$

Also, the probability that event B occurs might depend on I different events, so that we generalize the above equation to

$$P(B) = \sum_{i=1}^{I} P(B|A_i)P(A_i)\,, \tag{2.220}$$

where one A_i might correspond to $\bar{A}$. This concept is called *marginalization*, since the set of events A_i that influences event B disappears on the lhs of the above equations.

Example IV:

Assume that 50% of all people have a healthy lifestyle. On the other hand, for those that do not live a healthy lifestyle, the probability to die from cancer is 80%. What is the ratio of people that will die from cancer?
From Equation 2.219 we know that $P(D) = P(D|H)P(H) + P(D|\bar{H})P(\bar{H}) = 0.1 \times 0.5 + 0.8 \times 0.5 = 45\%$.
If all (100%) people would live a healthy lifestyle, then $P(D) = 0.1 \times 1 + 0.8 \times 0.0 = 0.1$, i. e. we end up correctly with the second statement from example III.
If 75% of all people would live healthy, then $P(D) = 0.1 \times 0.75 + 0.8 \times 0.25 = 0.275$, i. e. although only a quarter of the population does not live healthy, the mortality raises almost by a factor of three (note the detailed modeling in Section 5.3)! Indeed many people in Germany and the other western countries easily reach ages around 90 years if they live slightly healthier than the average population, whereas the total average of the life expectancy is around 77 yrs.

We can turn the question around and ask how likely it is that one person who died from cancer had a healthy lifestyle ($P(H|D)$) if (s)he belonged to the population, where 75% of all people have a healthy lifestyle? Expressed as an equation, we

can write $P(H|D) = P(D|H)P(H)/P(D) = 0.1 \times 0.75/0.275 = 0.273$. *Furthermore,* $P(\bar{H}|D) = P(D|\bar{H})P(\bar{H})/P(D) = 0.8 \times 0.25/0.275 = 0.727$, *and hence* $P(H|D) + P(\bar{H}|D) = 1$.

Example V:

Imagine a horse race of n *horses (see [1]), where we would like to predict the probability of a particular result of this race. The set of* n *horses is mutually exclusive (ME), since each horse can only achieve one particular position in the race (hence,* $P(A \cap B) = 0$*). Each horse is assigned to a prior probability* $P(i)$, $P(j)$, $P(k)$, ...$P(n)$ *to win the race. The prior probabilities could have been estimated from numerous earlier races and/or from knowledge of their physical properties, their average speed etc.*
How likely is it that first horse i, *then horse* j *and third, horse* k *cross the finish line in this particular order?*
Using Bayes' notation, we write that

$$P(kji) = P(k|ji)P(ji) = P(k|ji)P(j|i)P(i)\,.$$

$P(i)$ *is the prior probability that horse* i *wins the race. After horse* i *crosses the finish line as first, horse* j *has to cross the finish line as second. Therefore, the conditional probability that event* j *occurs after* i *has occurred is* $P(j|i) = P(j)/[1 - P(i)]$. *The factor* $1 - P(i)$ *results from the fact that horse* i *is not part of the race anymore, when it has crossed the finish line. The sum of all priors equals one, since the set is CE. Once horse* i *is not part of the set anymore, the sum of all remaining priors equals* $1 - P(i)$. *In the same manner we obtain for the third horse* $P(k|ji) = P(k)/[1 - P(i) - P(j)]$. *Hence, the probability that horse* i, *horse* j *and horse* k *cross the finish line in this particular order equals*

$$P(kji) = \frac{P(i)P(j)P(k)}{[1 - P(i)]\,[1 - P(i) - P(j)]}\,.$$

Exercise I:

80% of bacteria having genotype (A) die after being exposed to a certain level of radiation, whereas the modified genotype (B) has a mortality of 5%. Consider an unknown mixture of bacteria of genotype A and B in your petri dish. After exposing this mixture to radiation of the same level, 27.5% of all bacteria have survived. What is the ratio of genotype A and B in your petri dish? Use Bayesian notation!
What is the correlation factor g for the mortality of genotype B with respect to the entire mixture?

2.6.3 Mean, Median and Variance

Most of you probably know what mean, median and variance are, but since these properties are widely used, I would like to repeat them in this section. First, I provide the definitions

and after that we will discuss some examples.
The mean μ of a quantity x is defined as

$$\boxed{\mu = \int x\, p\,(x)\, dx}\,, \tag{2.221}$$

where $p(x)$ is a *probability density function* (pdf) that gives the weight of each individual x. The pdf equals the prior probability for a given event x to occur, hence a probability per event (therefore, the term *density* is used). Thus, the integral over all possible values of x, $\int p(x)dx$ equals one by definition if x is CE.
Frequently, the mean is also called *expectation value*, often denoted as $\mu = \langle x \rangle$.
If x is discrete (i. e. one can count x), Equation 2.221 turns into

$$\boxed{\mu = \sum_i x_i\, p\,(x_i)}\,. \tag{2.222}$$

Example I:

What is the mean of numbers on a die? First, we apply Equation 2.222 because we can count the numbers (x is discrete). What is $p(x)$? The probability for each x (for each number on the die) to appear after rolling the die is identical (called uniform), thus $p(x)$ is constant. Since we have six numbers, $p(x) = 1/6$ (c.f. Equation 2.208). Therefore, Equation 2.222 equals $\frac{1}{6}(1+2+3+4+5+6) = 3.5$. The expectation value (or mean) is 3.5! **Thus, the mean can even be a number that is not necessarily an element of x!**

Example II:

Let us now consider a biased die, where the probability for each number is not uniform, but say $1/10$, $1/10$, $1/10$, $1/10$, $1/10$, $1/2$ for the numbers 1, 2, 3, 4, 5 and 6, respectively. What is the expectation value now? According to Equation 2.222: $\mu = 1/10 \times 1 + 1/10 \times 2 + 1/10 \times 3 + 1/10 \times 4 + 1/10 \times 5 + 1/2 \times 6 = 4.5$. As expected, 6 will be more frequent and the expectation value is biased towards larger numbers.

Example III:

What is the expectation value of uniformly distributed random numbers between 0 and 6, if these numbers are continuous? Again, $p(x)$ is constant since we have uniformly distributed numbers, hence $p(x) = const$, but now we need Equation 2.221 and therefore $\mu = const \int_0^6 x\, dx = 18 \times const$. But what is the value of "const"?
By definition $\int p(x)\, dx = 1$ and in our case only values between $a = 0$ and $b = 6$ are possible. Hence, $\int_a^b const\ dx = 1$ and therefore $const = 1/(b-a)$. Finally, we obtain $\mu = 18/(b-a) = 3$.

There are different definitions of the mean. Equation 2.221 and Equation 2.222 are called *arithmetic mean* that is often called *average*, although average and mean are generally **not** identical in a strict sense.
Other definitions of the mean are the *geometric mean* (often used in statistical physics)

$$\boxed{\mu_g = \left(\prod_i^n x_i\right)^{(1/n)}}, \tag{2.223}$$

the *harmonic mean* (for example for calculating the resistance in a circuit or the period of a beat frequency)

$$\boxed{\mu_h = n\left(\sum_i^n \frac{1}{x_i}\right)^{-1}}, \tag{2.224}$$

and the *weighted mean*

$$\boxed{\mu_w = \frac{\sum_i^n x_i w_i}{\sum_i^n w_i}}, \tag{2.225}$$

for example if x_i are measured values having the errors $\epsilon_i = 1/w_i$ and w_i weights the

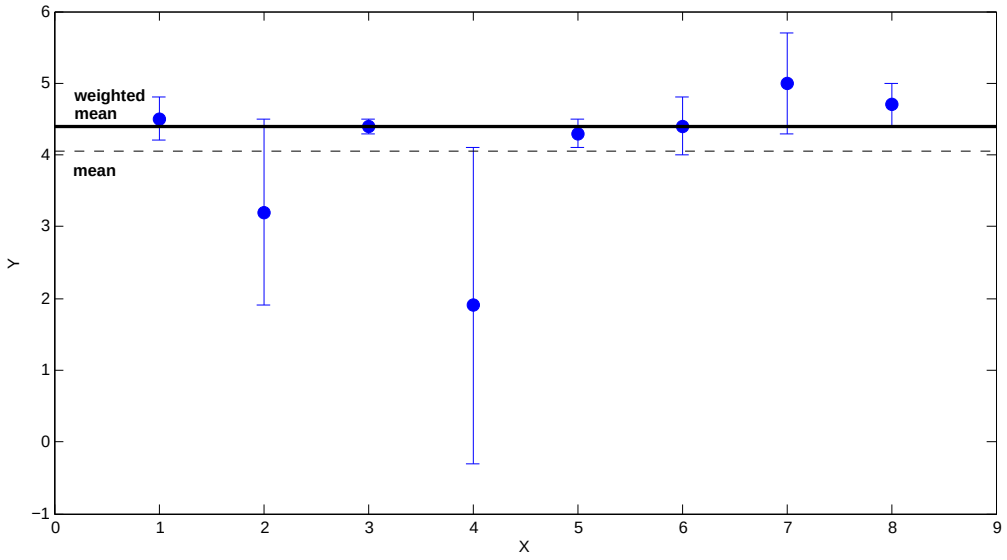

Figure 32: The weighted mean is more appropriate then the non-weighted arithmetic mean if the individual data points have error intervals of different sizes.

values according to their accuracy (so that values with large errors contribute less to the mean than accurately measured values, see Figure 32).

Since we derive the mean from statistical values, there is an interest on a measurement of the natural spread of the mean. This spread is called *variance* and is defined via

$$\boxed{var\,(x) = \int (x-\mu)^2\, p(x) dx}. \tag{2.226}$$

There are some reasons (see statistics primer) why the variance is defined as in Equation 2.226. The deviation from the mean (called *standard deviation*) is defined by

$\sigma = \sqrt{var\,(x)}$.
A further property is the *median* m that is defined by the value that divides the pdf into two parts of equal area:

$$\boxed{\int_{-\infty}^{m} p(x)\,dx =: \frac{1}{2}}\,. \tag{2.227}$$

The maximum of $p(x)$ does not necessarily coincide with the mean and/or the median. Also the median does not necessarily equal the mean, as shown in Figure 33. Note, that

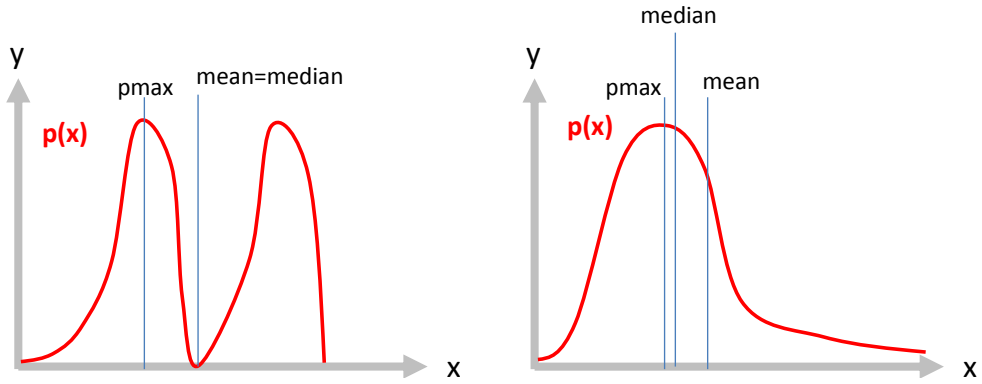

Figure 33: The mean is sometimes an inappropriate quantity and is not necessarily an element of the free variable, for example if the probability density function is bimodal (left). Also mean, median and the maximum of the probability density function are not identical in general (right). Thus, one has to take these quantities with care when reading statistics.

the median is less affected by outliers than the mean!

2.6.4 n Factorial and "n choose k"

Suppose a class room of n empty chairs and k students that wait outside to enter the room. How many different seating arrangements are possible, if the students enter the room and occupy their seats randomly?
The first student entering the room has the full choice of n seats and thus, there are n possibilities for the first student to take a seat. When the second student enters the room, one seat is already occupied by the first student and (s)he has only $n-1$ possibilities. In total, after two students entered the class room, there are $n\,(n-1)$ different seating arrangements. Then the third student enters the room finding $n-2$ free seats and hence, there are $n\,(n-1)\,(n-2)$ different seating arrangements possible. For k students, we therefore find $n\,(n-1)\,(n-2)\ldots(n-(k-1))$ different seating arrangements. If $n=k$, there are $n\,(n-1)\,(n-2)\ldots(n-n+2)\,1 = 1\times 2\times 3\times\cdots\times n$ seating arrangements. The product

$$1\times 2\times 3\times\cdots\times n =: n! \tag{2.228}$$

is defined as n *factorial*, $n!$.
With this definition, we can express the general case $k \leq n$ as

$$n\,(n-1)\,(n-2)\ldots(n-(k-1)) = \frac{n!}{(n-k)!}\,. \tag{2.229}$$

For $n=k$ we can recover the result $n!$ since $0! = 1$.
Often it is stated that $0! = 1$ by definition, but that is not true. One can actually proof that $0! = 1$. According to Equation 2.228 $\frac{n!}{(n-1)!} = n$, that also must hold for $n = 1$. Since

$n! = 1$, we obtain $\frac{1!}{0!} = 1$. Thus, we conclude that $0! = 1$.
Equation 2.229 holds if the k students are regarded as individuals, i. e. if we really care which particular student sits on which particular chair. If we however do not care which student occupies a seat as long as *a* student is sitting on a particular chair (hence, the students are "indistinguishable"), we have less possibilities to obtain different seating arrangements. For example, if student A changes the seat with student B we would not count this as a different arrangement now. For k students, we therefore would obtain $k!$ indistinguishable arrangements within each of the previous $\frac{n!}{(n-k)!}$ arrangements and thus, would obtain only

$$\frac{n!}{k!\,(n-k)!} \tag{2.230}$$

different arrangements.
The above structure appears quite frequently in stochastic and therefore is defined as

$$\boxed{\frac{n!}{k!\,(n-k)!} =: \binom{n}{k}} \tag{2.231}$$

spoken as "n choose k". It is the number of arrangements (or *states*) for k **indistinguishable** objects (or particles etc; here students) within $n \geq k$ sub states (here seats in a class room). Also in both cases (distinguishable or indistinguishable) **the order** in which the students entered the class room **does not matter**.

Another example is the following: In the German lottery one has to guess six correct integer numbers running from 1 to 49. One can imagine this as a barrel containing 49 balls numbered from one to 49. Now, we draw blindly a ball from the barrel and record the number n_1, say $n_1 = 25$. The ball is not put back into the barrel, but kept outside. We repeat this procedure five times and therefore obtain six numbers. Those who guessed these numbers right (the order of appearance of these six numbers doesn't matter) win the jackpot. What is the probability to win the jackpot?
The probability to get the first number right is $1/49 \times 6$, since it does not matter which of the six numbers one draws first. Then, in the second draw, one has to catch one ball out of 48, but it has to be one of the five remaining numbers that one guessed: $1/48 \times 5$. Hence, the probability to guess two numbers right is $\frac{6}{49} \times \frac{5}{48}$. Now we draw the third number and so on so that finally, the probability to get the six correct numbers is

$$P = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{49 \times 48 \times 47 \times 46 \times 45 \times 44} = \frac{1}{13983816}\,, \tag{2.232}$$

or in other words, we have to guess the right set of numbers out of $W = 13.983.816$ valid combinations. Generally, the probability to draw the k correct numbers out of n (if the order is **not** important) is

$$P = \frac{k!}{n!/(n-k)!} = \frac{k!(n-k)!}{n!} = 1/\binom{n}{k}\,. \tag{2.233}$$

Thus, again we have $\binom{n}{k}$ possibilities to draw k numbers out of n.

2.6.5 The Binomial Distribution

Often, one is confronted with either/or or yes/no questions, e. g. consider calculating the probability to find a certain number of purines (adenine or guanine) in a given sequence

(of a certain length) of micro RNA. This is a typical yes/no problem: no purine, or purine (yes). Another problem of the same category is predicting the number of diseased offspring (diseased vs non diseased) if the parental generation has a known genetic defect etc.
Let us assume we have a couple hosting a genetic defect and the probability p to pass it to an offspring is known. Consequently, the probability that this defect is not passed equals $1-p$. For example, p can be 25% and we now can ask for the probability P that, say $k=3$ out of $n=5$ offspring show this defect. How can we calculate P?
One possible sequence of events could be that the first two offspring are not diseased and that the last three offspring show the defect. The probability of obtaining this particular sequence of events equals

$$(1-p)\,(1-p)\,p\,p\,p = (1-p)^2\,p^3\,, \tag{2.234}$$

or for k cases among n offspring

$$(1-p)^{n-k}\,p^k\,. \tag{2.235}$$

However, this is the probability for *one* particular sequence, but we are interested in every case, where $k=3$ and $n=5$. Thus, we have to consider the number W of sequences, that give the result $k=3$ and $n=5$. Fortunately, we know from Section 2.6.4 that

$$W = \frac{n!}{k!(n-k)!} = \binom{n}{k}\,. \tag{2.236}$$

Thus, we have W sequences of interest, each of it with the probability $(1-p)^{n-k}\,p^k$. Therefore, the probability P that k out of n offspring show this defect, given p equals

$$\boxed{P(k|p,n) = \binom{n}{k}\,p^k\,(1-p)^{n-k}}\,, \tag{2.237}$$

that is the so called *binomial distribution*.
There are implemented functions in $Matlab$ ("binpdf") and R ("dbinom") that calculate P based on Equation 2.237. Thanks to these functions, it is a matter of seconds to obtain $P \approx 8.8\%$ for our example of $k=3$, $n=5$ and $p=0.25$. We can also ask for the probability to have *less* than three diseased offspring if $n=5$ and $p=0.25$, that is just the sum of $P(k=0|p,n)+P(k=1|p,n)+P(k=2|p,n) \approx 90\%$. Also this can be calculated directly by the $Matlab$ function ("bincdf"), where "cdf" stands for *cumulative density function*.
The binomial distribution is shown in Figure 34 for different k for $n=10$ and $p=0.25$.
For different p, the binomial function shifts towards higher or lower P for given k that is illustrated in Figure 35. Since the binomial function is a pdf, the sum over all k equals one:

$$\sum_{k=0}^{n} P(k|p,n) = \sum_{k=0}^{n} \binom{n}{k}\,p^k\,(1-p)^{n-k} = 1\,. \tag{2.238}$$

Example I:

Both, a male and a female carry the gene for albinism. The probability for an

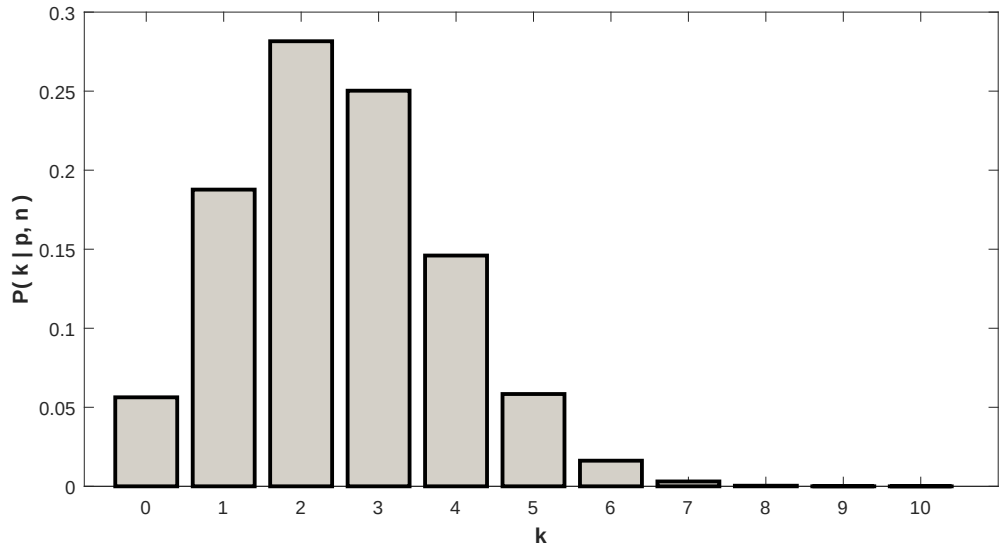

Figure 34: Binomial probability density distribution according to Equation 2.237 for $p = 0.25$ and $n = 10$.

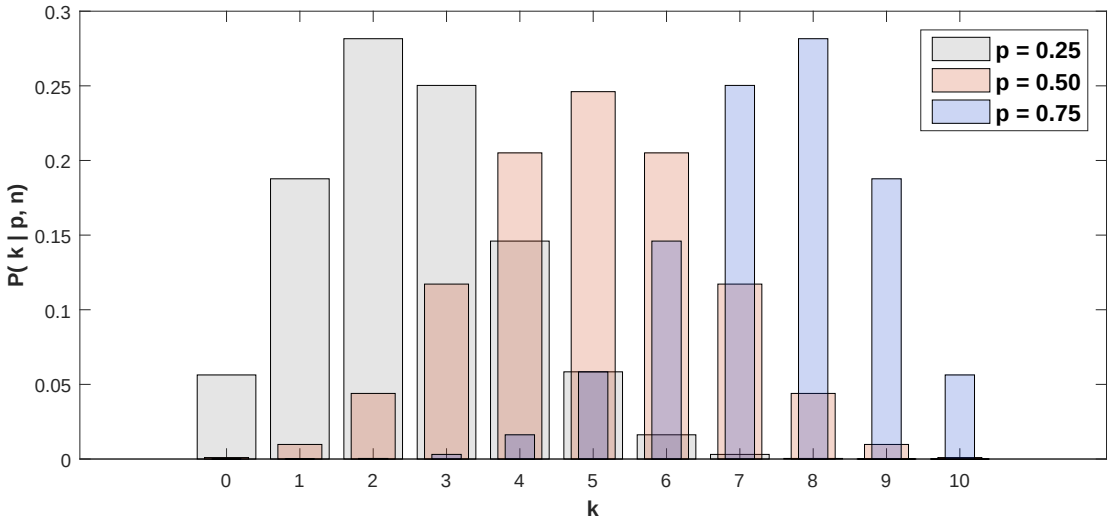

Figure 35: Binomial probability density distribution for $n = 10$ and different p. The thickness of the bars is varied for the sake of visibility.

offspring to become an albino is 0.25%. How large is the probability that exactly one offspring among five children has albinism? According to Equation 2.237

$$P(k = 1|p = 0.25, n = 5) = \frac{5!}{1!(5-1)!} \times (1/4)^1 \times (1 - 1/4)^{5-1} \tag{2.239}$$

that equals 40%. For two children being albinos $P = 26\%$ and for one child among three $P = 42\%$.

Example II *(advanced):*

The following question arose when estimating the amount of money required for a BMBF grant:
A sequencing process generates k clones that are randomly distributed over n colonies. Each colony contains only one particular, but unknown, kind of clone. How many colonies does an experimenter have to pick randomly in order to get 90% confidence for having **each clone** *at least once in the set if the prior probability for the clones is uniform?*
With the first pick, we have a probability of $1/k$ to get a particular clone, or a probability of $\left(1-\frac{1}{k}\right)$ to get not this clone. For n picks, we have a probability of $\left(1-\frac{1}{k}\right)^n$ to do not get this particular clone. In total, we have $\binom{k}{1}\left(1-\frac{1}{k}\right)^n$ ways to always miss this particular clone in n trials (we do not care whether other clones are missed as well or appear more than once). Thus, the probability to get **always** *this particular clone equals $1-\binom{k}{1}\left(1-\frac{1}{k}\right)^n$.*
In the same way we find the probability $1-\binom{k}{n_k}\left(1-\frac{n_k}{k}\right)^n$ for **always getting n_k particular clones** *in n trials. However, if we find, say, always $n_k(2)=2$ particular clones and $n_k(3)=3$ particular clones, then two clones of $n_k(3)$ could be exactly those from $n_k(2)$. For larger n_k, the overlap will be even larger and it will be more likely to get a larger overlap for larger n_k. For example, suppose $k=500$ and we will easily catch up the first 100 or 200 different clones, but then we will obtain more and more repetitions and it will be really hard to get the very last kind of clone. Thus, we have to subtract all the overlaps we do not need.*
Equation 2.214 shows how to join two sets A and B without getting the overlap and Equation 2.215 shows the inclusion/exclusion principle for an infinite number of sets. Joining the inclusion/exclusion principle with our previous considerations we finally obtain

$$P(k,n)=\sum_{n_k=0}^{k}(-1)^{n_k}\binom{k}{n_k}\left(1-\frac{n_k}{k}\right)^n. \tag{2.240}$$

The first two parts ($n_k=0$ and $n_k=1$) of this equation yield the probability $1-k\left(1-\frac{1}{k}\right)^n$ for not getting a particular clone, what we have derived already.
For $k=25$ and $P(k,n)=90\%$, we derive $n=135$ according to Equation 2.240. Note, that it was $k=5,000$ clones in the proposal for the BMBF grant.

The problem in this example is also known under the term "Coupon Collector Problem" and was first recognized by a broader audience when the company $Panini$ planned to sell their stickers of football players during the world championship in 1970.

Exercise:

Plot the binomial distribution using an internal function of $Matlab$ or R for large numbers n. How does it change with respect to the binomial distribution for small n?

According to the last example, what would we expect if the couple is very productive and produces, say twelve, children? Common sense tells us: in average $0.25\times 12=3$ (or

$n\ p = k_{expect}$) children would be albinos. The exact formulation is the mean of k that is by definition (Equation 2.222)

$$\sum_{k=0}^{n} P(k|p,n)\, k = \sum_{k=0}^{n} \binom{n}{k} p^k\, (1-p)^{n-k}\, k\,, \tag{2.241}$$

leading indeed to the result $n\, p = k_{expect}$.
Since we discuss a random process there is a natural spread around k_{expect}. This spread is introduced as variance (Equation 2.226) and we therefore find

$$\sum_{k=0}^{n} P(k|p,n)\, (k - k_{expect})^2 = n\, p\, (1-p) = var[k] = \sigma^2\,. \tag{2.242}$$

Often people use the standard deviation $\sigma = \sqrt{var}$ rather than the variance var.

2.6.6 The Poissonian Distribution

The Poissonian[22] distribution is widely used to model stochastic processes like mutations (Section 5.3), the decay of atoms or chemical reactions if only a few molecules participate at the same time and some macroscopic approximations do not apply any more (Section 5). Also single gene expression or the motion of molecular motors obey a Poissonian process (see in particular Figure 72 and also Section 6).
If we consider a binomial problem (Equation 2.237) with low probability $p << 1$ then we need a large test sample to obtain a result. Thus, one could apply a more appropriate function, **a function of rare events**, that is the *Poissonian distribution*.
If an event is rare, then $p << 1$ and $k << n$ and we are allowed to perform some approximations. If we look at Equation 2.237 we see that it contains two parts: the "n choose k" part and the $(1-p)^{n-k}$ part. Let's start with the latter.
Since $p << 1$ we can apply a Taylor expansion (Equation 2.115) around $p = 0$. The zeroth order equals $(1-p)^{n-k}$ that yields 1 for $p = 0$. The first order in the Taylor expansion originates from the first derivative

$$\frac{d\,(1-p)^{n-k}}{dp} = -\,(n-k)\,(1-p)^{n-k-1} \tag{2.243}$$

and the second derivative is

$$\frac{d^2\,(1-p)^{n-k}}{dp^2} = (n-k)\,(n-k-1)\,(1-p)^{n-k-2} \tag{2.244}$$

and so on.
Thus, the Taylor expansion around $p = 0$ equals

$$(1-p)^{n-k} \approx 1 \;\; -(n-k)\,p \;\; +(n-k)\,(n-k-1)\,\frac{p^2}{2} \tag{2.245}$$

$$-\,(n-k)\,(n-k-1)\,(n-k-2)\,\frac{p^3}{6}\,\ldots\,. \tag{2.246}$$

[22] Siméon Denis Poisson, 1781 - 1840

Since $k << n$ and $n >> 1$, only n matters and Equation 2.246 simplifies to

$$(1-p)^{n-k} \approx 1 - np + \frac{(np)^2}{2} - \frac{(np)^3}{6} \dots . \tag{2.247}$$

Comparing Equation 2.247 to Equation 2.125 we see that

$$1 - np + \frac{(np)^2}{2} - \frac{(np)^3}{6} \cdots = e^{-np} . \tag{2.248}$$

Hence, $(1-p)^{n-k} \approx e^{-np}$ for small p.
In the same manner we use Stirlings (Equation 2.131) approximation for

$$\frac{n!}{(n-k)!} \approx \sqrt{\frac{n}{n-k}} \frac{n^n e^{n-k}}{e^n (n-k)^{n-k}} \approx n^k \tag{2.249}$$

since n is large. Hence, altogether we find that

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{(np)^k e^{-np}}{k!} . \tag{2.250}$$

Often the product $n\,p$ is renamed λ. Therefore we write

$$\boxed{P(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}}, \tag{2.251}$$

that is the Poissonian distribution.

The expectation value of the Poissonian distribution equals

$$\sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda = k_{expect} \tag{2.252}$$

and the variance is

$$\sum_{k=0}^{\infty} (k - k_{expect})^2 \frac{\lambda^k e^{-\lambda}}{k!} = \lambda = k_{expect} \tag{2.253}$$

again! Since $\lambda = n\,p$, the relative error $\sqrt{var[k]}/k$ goes with $1/\sqrt{k}$. The property that variance and mean are identical is characteristic for the Poisson process. In fact, a Poisson process it identified via this property in practice. Comparing the variance of the Poisson distribution to the variance of the binomial (Equation 2.242), we see that we can recover the result for the approximation $p << 1$.
Usually, we will need the Poisson distribution when we have rates such as one event per time so that $\lambda = \nu\,t$ where ν is an average number of events per time span (e.g. reaction per second) and t is a time interval.

Example:

A given process folds two proteins per millisecond in average. How likely is it to observe two folding events within one millisecond?

First, $\nu = 2/ms$ and $t = 1ms$, thus $\lambda = 2$. Furthermore, $k = 2$ (two folding events). Hence, the probability to observe two folding events within one millisecond is $\frac{2^2}{2!}\, e^{-2} \approx 0.27$
How likely is it to observe no folding event within three milliseconds?
$t = 3ms$ and $k = 0$ yielding $\frac{6^0}{0!}\, e^{-6} \approx 0.0025$

2.6.7 The Gaussian Distribution

The binomial function becomes symmetric and bell-shaped if it is applied for large n. Since the Poisson distribution is derived from the binomial distribution for small p it shows qualitatively the same behaviour for large n. It seems that there is a common limit for large n and hence, there must be a common function that represents this limit.
Applying Equation 2.131 (large n) for Equation 2.237 we obtain

$$P(k|p,n) \approx \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k\,(n-k)}} n^n k^{-k} \left(\frac{1}{n-k}\right)^{n-k} p^k\,(1-p)^{n-k}\ . \tag{2.254}$$

The term $n\,p$ is the mean of the function (since it is the mean of the binomial distribution) and any value apart from $n\,p$ can be reached by a shift $n\,p + x$. It turns out that Equation 2.254 can be simplified if it is expressed in terms of x:

$$P(x|p,n) \approx \frac{1}{\sqrt{2\pi n\; p\,(1-p)}} \left(1 + \frac{x}{np}\right)^{-k-1/2} \left(1 + \frac{x}{n\,(1-p)}\right)^{-n+k-1/2}\ . \tag{2.255}$$

Now we apply a trick and write

$$e^{\ln\left[\left(1+\frac{x}{np}\right)^{-k-1/2}\left(1+\frac{x}{n(1-p)}\right)^{-n+k-1/2}\right]} \tag{2.256}$$

because this enables us to use the Taylor approximation for $\ln(1 + x/n)$ around $x/n = 0$ since n is large. The algebra that leads to the result is simple, but lengthy so that I just give the result

$$P(x|p,n) = \frac{1}{\sqrt{2\pi\, n\, p\,(1-p)}} exp\left[-\frac{x^2}{2\, np\,(1-p)}\right]. \tag{2.257}$$

One can show that the variance of Equation 2.257 is $\sigma^2 = np\,(1-p)$ (c.f. Equation 2.242) so that we write

$$\boxed{P(x|\sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{x^2}{2\sigma^2}\right]} \tag{2.258}$$

that is the famous **Gaussian[23] distribution**. More general, we write

$$\boxed{P(x|\sigma,\mu) = \frac{1}{\sqrt{2\pi}\sigma} C\; exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]} \tag{2.259}$$

with a constant C and the mean μ.
The Gaussian distribution is the most important probability density distribution since it is a threshold for many other functions and an overwhelming variety of very different natural

[23] Carl Friedrich Gauss, 1777 - 1855

phenomena (e.g. diffusion of particles, body size of humans, many quantum effects, error estimations etc.) can be modelled with it. Quantities obeying Equation 2.259 are called Gaussian or normally distributed. Since we derived the Gaussian distribution for large n and Stirlings approximation becomes sufficiently accurate for $n \gtrsim 10^2$, **the Gaussian approximation is only applicable for** $n \gtrsim 10^2$!. Many statistical tests (t–test, z–test, ANOVA, χ^2–test etc.) require a Gaussian distribution of the data and therefore are only applicable if $n \gtrsim 10^2$. I am aware that this fact is commonly ignored and such tests a widely used, although even if not applicable in a strict sense if n is too small. The good news is that there is a large set of alternative, mathematically applicable and more accurate tests, such as Bayesian parameter estimation or variational Bayesian density estimation.
The implemented $Matlab$ and R functions "normpdf" and "dnorm", respectively, generate random variables based on Equation 2.259. The Gaussian distribution is shown in Figure 36.
The exponent in Equation 2.259 scales the difference between a given value x and μ to

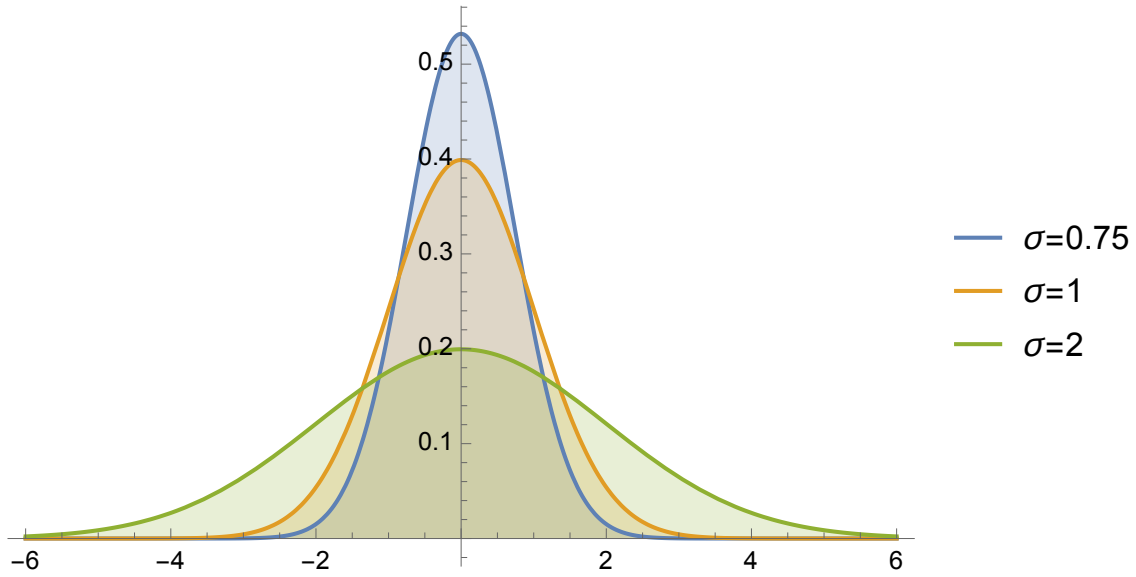

Figure 36: The normal probability distribution according to Equation 2.259 for different values of σ and $\mu = 0$. Different values for μ would shift the function along the x axis accordingly..

lengths of σ, hence it makes sense to give deviations from the mean μ in terms of 1σ, 2σ and so on. The area described by the function and vertical lines of $\pm 1\sigma$ covers ≈68% of the entire area described by the Gaussian and the $y = 0$ axes. Such an interval is called *confidence interval*. Figure 37 shows some frequently used confidence intervals.

2.6.8 Central Limit Theorem

Of course it is not a coincidence that the binomial and the Poisson distribution turn into a Gaussian distribution for large sample sizes. Generally, the arithmetic mean of independent random variables, each with a well-defined expectation value and well-defined variance, will reach a Gaussian distribution as limit for a large number n of measurements. This statement is called the **Central Limit Theorem**.
Suppose we have any probability density function $p(x)$ and we want to estimate it around its peak. Since we deal with probabilities, it is more convenient to use $L(x) = \ln p(x)$. Investigating L around its maximum (i. e. around the value x, that corresponds to the

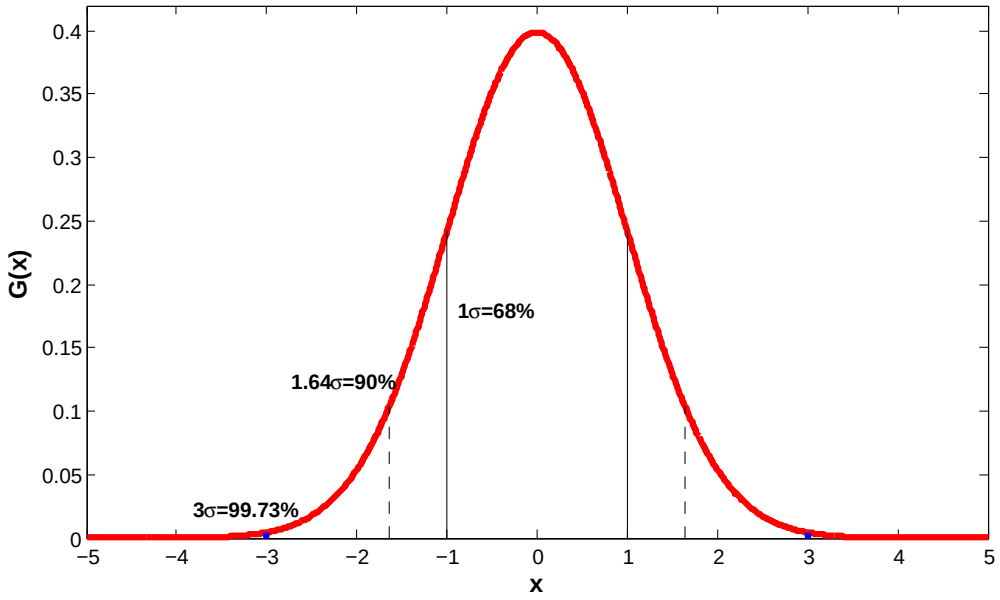

Figure 37: Confidence intervals for $\pm 1\sigma$ (solid vertical lines), $\pm 1.64\sigma$ (dashed vertical lines) and $\pm 3\sigma$ (bold blue vertical lines) corresponding to 68%, 90% and 99.73% of the entire area covered by the Gaussian probability density distribution.

probability density function (pdf)	variables	mean	variance	when to use
binomial $P(k\|p,n) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$	**n**: total number of trials/ sample size **k**: those n of a special kind **p**: probability to observe k in one trial	np	np (1-p)	If one has two choices: positive/negative; infected/ not infected "What is the probability to observe k events after n trials for a given p?"
Poisson $P(k\|\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$	**λ**: average events **λ = νt** with rate **ν** and time span **t** **k**: number of events	λ	λ	From binomial for p<<1 (rare events). "What is the probability to observe k events within t for a given λ?"
Gaussian $P(x\|\sigma,\mu) = \frac{C}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	**C**: normalization constant **σ**: standard deviation (covers ≈ 68% of the pdf) **μ**: mean (equals also the peak of the function)	μ	σ^2	Limit of the binomial and the Poisson distribution for large ($n > 10^2$) sample size

Figure 38: A small list of common probability density functions and their usage. Almost all common statistical tools are based on them.

maximum of L, hence the value $x = x_{max}$) equals investigating $p(x)$ around its maximum, it only has to be a global maximum.
Applying a Taylor series (Equation 2.115) to L yields

$$L(x) \approx L(x_{max}) + \frac{dL}{dx}\,|_{x=x_{max}}\,(x - x_{max}) + \frac{1}{2}\frac{d^2L}{dx^2}\,|_{x=x_{max}}\,(x - x_{max})^2 \qquad (2.260)$$

The first derivative of L, $\frac{dL}{dx}$, equals zero at the maximum so that

$$L(x) \approx L(x_{max}) + 0 + \frac{1}{2}\frac{d^2L}{dx^2}\,|_{x=x_{max}}\,(x - x_{max})^2 \qquad (2.261)$$

and

$$p\left(x\right) \approx e^{L(x_{max})} exp\left[\frac{1}{2}\frac{d^2L}{dx^2}\mid_{x=x_{max}}\left(x-x_{max}\right)^2\right] \tag{2.262}$$

The expression $e^{L(x_{max})}$ is a constant (because $L\left(x_{max}\right)$ is a constant) I'd like to denote as $\hat{C}$. Equation 2.262 has the structure of a Gaussian function with $\sigma^2 = \left[-\frac{d^2L}{dx^2}\mid_{x=x_{max}}\right]^{-1}$. Note, that the variance (second derivative) has a negative sign that is demanded when we search for a maximum. Hence, we can write

$$\boxed{p\left(x\right) \approx \hat{C}\ exp\left[-\frac{1}{2}\frac{\left(x-x_{max}\right)^2}{\sigma^2}\right]} \tag{2.263}$$

Thus, the probability density function $p(x)$ can be approximated by a Gaussian!
In the statistics primer I show, that the mixed second derivatives are the covariances if p depends on more than one variable. Note, that there is a mathematically more rigorous proof of the central limit theorem. One can show that an unknown, but normalized, function where only mean and variance are defined must equal a Gaussian pdf if maximum entropy is assumed. The proof argues along the line of the proofs in Section 3.1.

Exercise:

Generate random, normally distributed numbers with $Matlab$ or R and plot the corresponding histogram; first for small n, then for larger n. For which n roughly does the typical bell shape emerge and random fluctuations become negligible? Compare these values to the accuracy of Stirlings approximation (Equation 2.133). What is the conclusion for related statistical tests (t–test, z–test, ANOVA, χ^2–test etc.)?
Repeat the same procedure for the binomial function. What happens and how is this connected to the central limit theorem.

References

[1] Ken Dill, Sarina Bromberg *"Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience"*, Taylor & Francis Inc; 2nd Rev ed. (13. December 2010)

3 Thermodynamics

Biological systems are driven by chemical reactions and these chemical reactions obey the laws of thermodynamics. Thus, to some extend, we have to understand the physical meaning of these laws in order to describe biological systems correctly. One property of life is that it always works against the increase of entropy since it needs to lower entropy locally and at least temporary. However, one law of thermodynamics states, that entropy is either constant, or steadily increasing in a closed system. If so, the work against entropy is a hopeless struggle that inevitably will result in the death of the organism. Then, the biological system is in equilibrium with its environment that equals maximum entropy. The only ways to delay this process is either lowering the temperature (slowing down particle motions) or exchanging energy and/or matter with the environment (i. e. "exporting" entropy).
Often, the actual meaning of thermodynamic properties like entropy, free energy or enthalpy are not apparent and there is a lot of misunderstanding (especially concerning the interpretation of entropy, we will come to that later). Thermodynamics is usually taught from phenomenological approaches in high school and in university. This has historical reasons, since these quantities were in fact discovered phenomenological e. g. by producing steel or constructing steam engines in the nineteenth century. However, this approach is partly responsible for some of these misinterpretations.
The physical meaning of all these thermodynamic properties was not understood until physicists began to explain them by the behaviour of atoms and molecules, their motion and their states. Each individual particle has a set of particular (microscopic) properties (velocity, mass, momentum) following statistical laws (Section 2.6). Macroscopic properties like entropy, pressure and temperature are a result of these microscopic properties. For example internal energy U of a gas is related to the mean quadratic speed of the particles, hence the mean kinetic energy. Since this statistical approach unveils the true nature of these thermodynamic properties and therefore helps to understand them correctly, we like to derive them in a statistical manner in this section.

In order to derive macroscopic properties of a system, such as temperature or total energy, it is advantageous to consider the number of possible "states" of the contained particles, atoms or molecules. The term "state" has to be regarded in an abstract sense, as its definition is dependent on the system in question. For example, a state can be described by a particular energy, location or configuration (bound/unbound, folded/unfolded) of either the system as a whole, or of its individual components. The concept of states is well illustrated using the following example:
In microscopic systems, energy is quantized when the system is in a bound state (i. e. an electron in a potential trap). Let us compare these quanta of energy to the number of 'pips' (or dots) on the faces of a die. They are quantized in integer values up to "six", where each number has the same probability to appear when the die is rolled. We will call these numbers *micro states*. In our example, all micro states occur with equal probability (1/6), hence, by definition, they are uniformly distributed (i. e. they have a flat probability distribution).

Now suppose that we have N dice to represent a system of N quantized particles. It follows that the sum of all the pips corresponds to the total (internal) energy of this system. The total energy is something that we will call a *macro state*. By examining the macro states of many identical systems, it quickly becomes apparent that they are not uniformly distributed. Already for $N = 2$ we see that it is less likely to observe the system in the macro states "2" or "12" than in the macro state "7", since there is only one possible way to

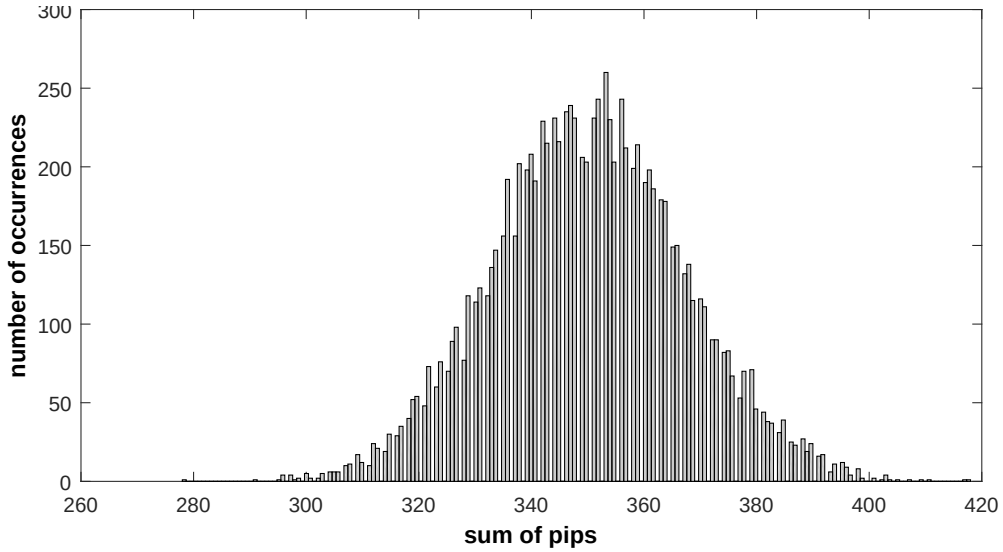

Figure 39: Sum of all pips (numbers on a die) after 100 dice rolled 10,000 times. Since the pips have a defined mean and variance, their sum follows the central limit theorem (Section 2.6.8)

generate "2" (1+1), but several possibilities to obtain "7" (1+6, 3+4, 2+5). If we consider $N = 100$, then it becomes more evident that the lowest energy state (100 times "one", i. e. all dice show the number "1"), and the highest energy state (100 times "six") are both extremely unlikely. The probability to observe these energies is $(1/6)^{100} = 1.53 \times 10^{-78}$. It is much more likely for the system to be in a state with total energy close to the mean of each roll times the number of rolls N (that is, 3.5×100). The probability that the system exhibits a state where the total energy is between 300 and 400 is 99.67%. This is nicely depicted in Figure 39.
From this example we have deduced that the most likely state of a macroscopic system is a product of the mean of its micro states. The influence of fluctuations in the micro states becomes less prominent as the number of particles is increased. This results in a more sharply peaked distribution about the mean of the macro states. The histogram in Figure 39 looks suspiciously like a Gaussian - and indeed it is, as shown in Section 2.6.8. If N is very low, then the histogram in Figure 39 smears out and finally becomes flat ($N = 1$). With this concept of micro states and macro states, we are now well prepared to dedicate ourselves to the first thermodynamic property: entropy.

3.1 The Concept of Entropy

Let us now generalize the example from the last section and say that we have N particles with I possible (micro)states each. Hence, n_i particles can be in state i (for integers $i \in [1, I]$). According to Section 2.6.4, this yields

$$W = \frac{N!}{n_1! n_2! \dots n_i! \dots n_I!} \tag{3.1}$$

different possible (macro)states of the system. Since we will deal with large N, we use Stirling's approximation (Section 2.2.4), so that Equation 3.1 reads

$$\begin{aligned} W &\approx \left(\frac{N}{e}\right)^N \frac{1}{(n_1/e)^{n_1}\,(n_2/e)^{n_2}\ldots(n_I/e)^{n_I}} \\ &= \frac{N^N}{n_1^{n_1} n_2^{n_2}\ldots n_I^{n_I}}\,\frac{e^{n_1}e^{n_2}\ldots e^{n_I}}{e^N}\,. \end{aligned} \tag{3.2}$$

The total number of particles is $N = \sum_i n_i$ and we therefore can use

$$\frac{e^{n_1}e^{n_2}\ldots e^{n_I}}{e^N} = \frac{e^{\sum n_i}}{e^N} = 1\,, \tag{3.3}$$

which simplifies Equation 3.2. Given the probability $p_i = n_i/N$ to observe a particle in state i, we are able to further simplify Equation 3.2:

$$\begin{aligned} W \approx& \frac{N^N}{(Np_1)^{n_1}\,(Np_2)^{n_2}\ldots(Np_I)^{n_I}} \\ =& \frac{N^N}{N^{\sum n_i}}\,\frac{1}{p_1^{n_1}p_2^{n_2}\ldots p_I^{n_I}} \\ =& \prod_{i=1}^{I}\frac{1}{p_i^{n_i}}\,. \end{aligned} \tag{3.4}$$

I like to loose some words about p_i here, since there are sometimes some conceptual misunderstandings. There exists the p_i of the system that generates n_i particles in state i, for example $p_i = 1/6$ for a fair die. This p_i is inherent and is the *prior probability* of the system generating a particular particle having a particular state. It is usually not possible to measure the prior p_i. However, if we observe the system sufficiently long and gain a sufficiently[24] large set of data we can divide the number of observed particles n_i found to be in state i by the total number of particles N to approximate $p_i \approx n_i/N$ (see also the law of large numbers discussed in Section 2.6.1). For $N \to \infty$ this *relative frequency* should be equal to the prior probability. For example, when rolling a die ten times we could **measure** the relative frequency for "one" $p_1 = 3/10 \neq 1/6$, in another case $p_1 = 1/10 \neq 1/6$ etc. since this process is random. For $N \to \infty$ such experiments, however, we should recover exactly the prior probability of $1/6$.

Since the number of macro states W is very large even for moderate N, it is convenient for many purposes to instead make use of $\ln W$:

$$\ln W \approx -\sum_i n_i\ \ln p_i\,. \tag{3.5}$$

This is defined to be the **total entropy** S_{tot} of the system. Since $n_i \geq 0$ and $p_i \leq 1$ it is apparent that this quantity must always be greater than or equal to zero, hence, we write:

$$\boxed{S_{tot} := lnW}\,. \tag{3.6}$$

We can now divide Equation 3.5 by the number of particles N and obtain

$$\frac{\ln W}{N} \approx -\frac{\sum_i n_i\ \ln p_i}{N} = -\sum_i p_i\,\ln p_i\,. \tag{3.7}$$

[24] What "sufficiently" really means depends on the situation.

the entropy per particle, often just called *entropy* S:

$$\boxed{S := -\sum_{i=1}^{I} p_i \ \ln p_i} \,. \tag{3.8}$$

Following Equation 3.5 we find that $S \sim \ln W$. If one joins two systems with the number of states W_1 and W_2, then the total number of states is $W_{12} = W_1 W_2$ (Section 2.6), hence $S_{12} \sim \ln(W_1 W_2) = \ln W_1 + \ln W_2 \sim S_1 + S_2$. Entropy scales with the size of the system, since one can add up the different entropies of different systems. This property is called *extensive*.
Some thermodynamic properties like temperature or pressure do not scale with the size of the system. It does not make sense to add the body temperature of two persons for example. Such properties are called *intensive*.

Before we make any further connections to thermodynamics, we will investigate the

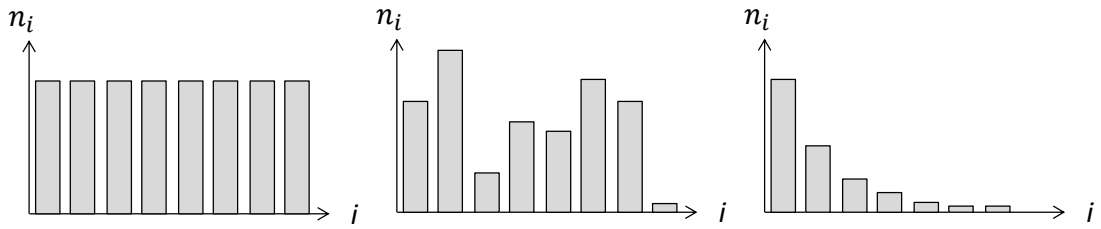

Figure 40: An uniform distribution (left) has the larges entropy possible (see Section 3.1.1), although it is the most ordered arrangement. The histogram in the middle looks least ordered, but it has the **lowest** entropy of all three. The histogram on the right has second highest entropy and shows the Boltzmann distribution (Section 3.2). This illustrates that the interpretation of entropy as measure of disorder is in fact misleading.

definition of entropy more thoroughly as it is a very important concept with many applications. In many older textbooks, entropy is described as the amount of 'disorder' in a system. For example, a messy office is disorderly and thus corresponds to high entropy, while a tidy office is ordered and therefore has low entropy. Examples of this kind are inaccurate and can be extremely misleading, as one can define a disordered office as ordered and vice versa [25]. This puzzling analogy actually arose from a statement made by Ludwig Boltzmann[26] (whom, as we shall see in the next section, made brilliant mathematical contributions to thermodynamics and entropy) in 1898 as an attempt to offer a simple interpretation of entropy. Unfortunately, it was an entirely incorrect explanation, mainly due to the fact that there was no real understanding of molecular behaviour until after his death in 1906. Treating the entropy as a function of the probability of different micro states (c. f. Equation 3.8) is unique, and much more specific and helpful than simply thinking of a system as ordered or disordered. To really understand what is behind Equation 3.8 we will go back to our example of rolling dice.

Let us assume we have $I = 6$ states that are uniformly distributed, implying that $p_i = n_i/N$ all have the same value of $p = 1/6$. Applying this to Equation 3.8 (the entropy per roll) we obtain $S = -I \ p \ \ln p = -\ln(1/6) \approx 1.8$. A configuration where all states are approximately uniformly distributed **has high entropy**, and would be regarded as ordered

[25] A counterexample is presented in Section 3.4 showing that the office example is really methodical wrong and not a matter of taste (see also Figure 40).

[26] Ludwig Eduard Boltzmann, 1844 - 1906

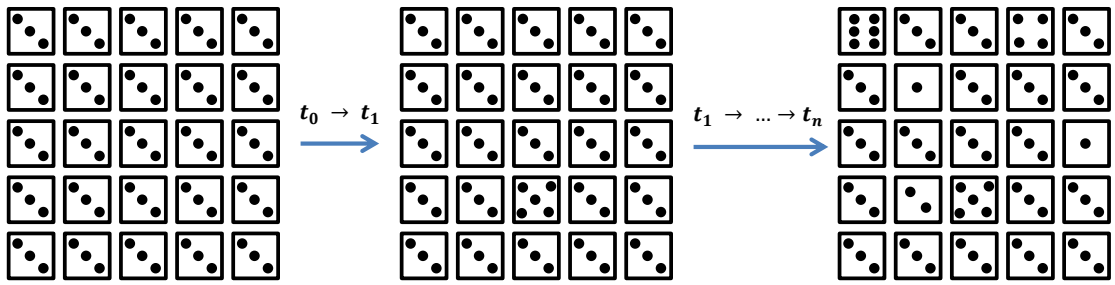

Figure 41: At every time step one die is chosen randomly to change its state (including changing back to the previous state) after a randomly chosen time step.

by certain sources (c. f. Figure 40). In Section 3.1.1 we show that a **flat distribution equals maximum entropy**, in this case. This distribution is unique in that it is only subject to a single constraint: conservation of probability ($\sum_i p_i = 1$). Many systems will have additional restrictions which alter the distribution that optimizes entropy. The most common is the law of conservation of energy, which requires that the sum of all particle energies adds to a constant value. This example will be addressed shortly in Section 3.2.
In contrast, let as now assume all p_i are zero except one (that is then $p = 1$ consequently). Hence, all particles are in the same state (all dice show the same number). Then Equation 3.8 sums up only zeros (either $p_i = 0$ or $\ln 1 = 0$) and $S = 0$. The entropy is minimal (it cannot be negative since $p_i \leq 1$). A system with this distribution of states has **minimal** S.
It is more appropriate to see entropy as a measure of information or actually a **measure of lag of information** (as it is done very successfully e. g. in bioinformatics in case of sequence alignment or motif finding or feature detection in image analysis, see Figure 121). If, for example, someone is rolling a die behind a screen and we had to guess the number without any further information, we had to assume that the probability of all numbers to appear ($p_i = 1/6$) is uniform and the entropy would be maximal ($S_{max} \approx 1.8$). However, if we had an additional information like e. g. that the rolled number is odd, then $p_i = 0$ for all the even numbers and $p_i = 1/3$ for the remaining odd numbers, hence $S \approx 1.1$. The loss of entropy equals the gain of information.

Let me return to the set of N dice and the distribution of their total sum like shown in Figure 39: usually we are talking about processes and that a closed system will approach maximum entropy after as sufficiently large time span. What does this actually mean and does the concept of entropy still makes sense for microscopic systems (i. e. small number of particles)?
If we start at $t = t_0$ with a set of dice all showing the same number faced up, i. e. all having the same state (e. g. energy), the entire system has zero entropy (c. f. Equation 3.8). We now pick one die randomly after a randomly chosen time step[27] and roll it. The die now might show another (or even the same) number faced up and we put it back. We now repeat the procedure by choosing a die randomly (that could be the same die by chance) after a randomly set time step and so on. The situation is illustrated in Figure 41.
If we measure the entropy after each time step and the mean number shown faced up, we will see prominent fluctuations if N is small. There is some chance that the histogram of numbers shown faced up after M time steps is not necessarily flat for small N, even if M is large. In a physical system we might identify the pip of each die with a micro state, say

[27] I will prove in Section 5.1 that the distribution of these randomly chosen time steps is poissonian.

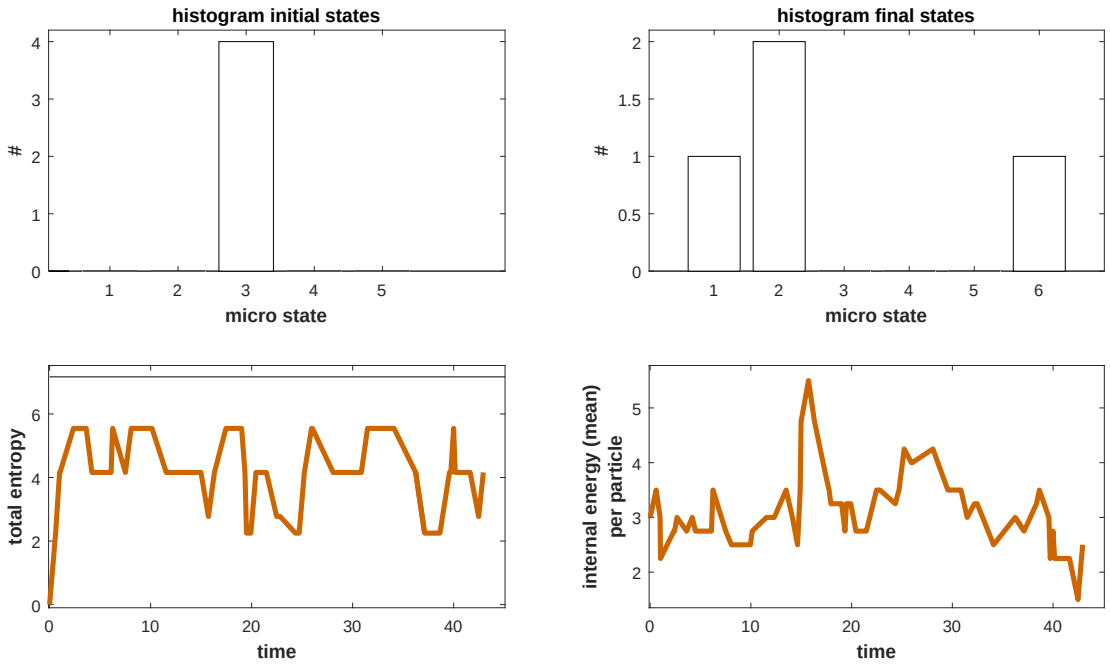

Figure 42: Course of entropy and "internal energy" of a system of $N = 4$ dice after $M = 45$ time steps.

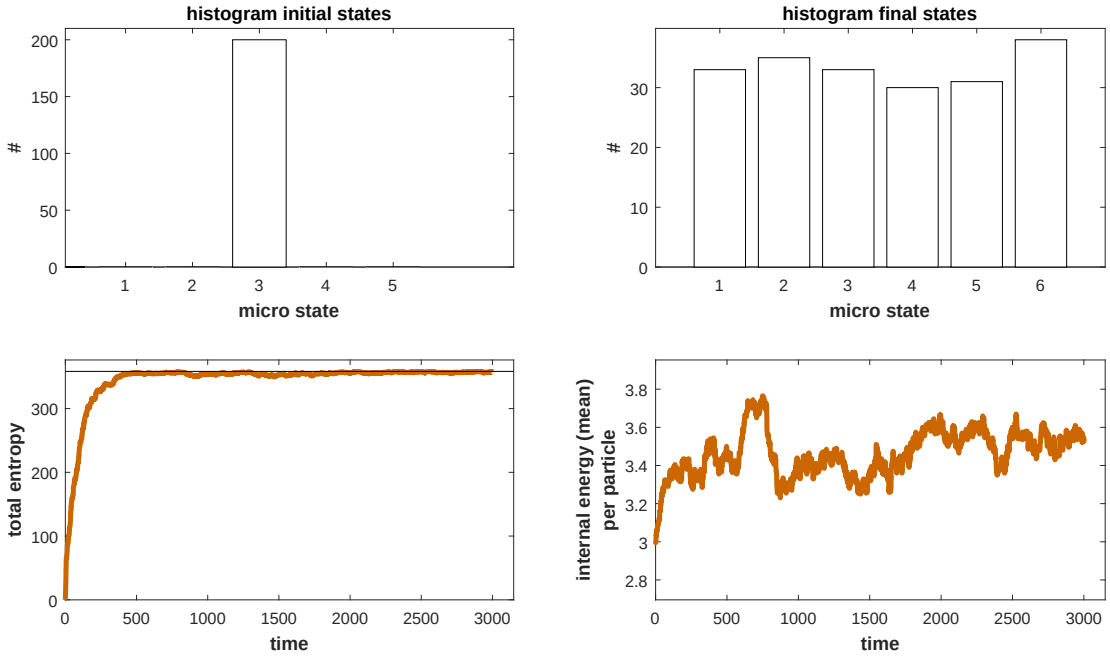

Figure 43: Same as in Figure 42 but for $N = 200$ dice and $M = 3000$ runs. Note, that the entropy of the system is now converging towards its maximum.

the energy of a particle. We will see in the next sections that the mean energy, here 3.5 per particle (die) can be identified with the internal energy. One simulated course of entropy of the system and mean energy per particle is shown in Figure 42 for $N = 4$ and $M = 45$. Indeed, the entropy does not converge to a maximum but fluctuates. Although there is a chance that we would observe an uniform distribution of the pips at the end, there is also a chance that the distribution can be anything else. Thus, for small N the total entropy of a system (not to confuse with the entropy of a roll) does not necessarily converge. This

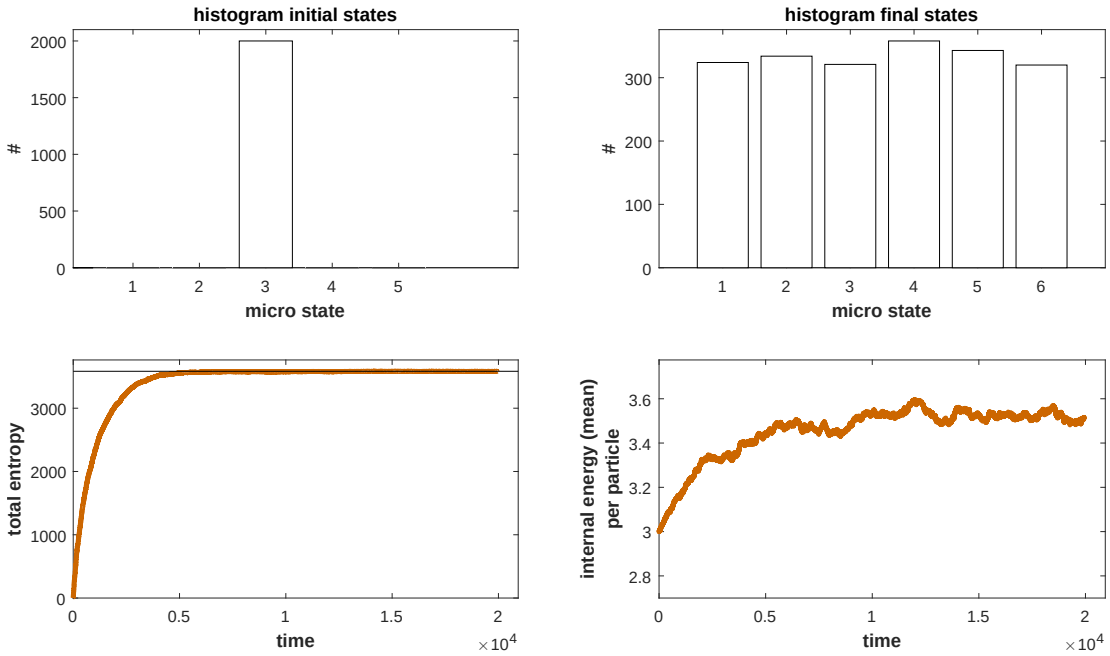

Figure 44: Same as in Figure 42 but for $N = 2000$ dice and $M = 20\,000$ runs.

becomes clear if we remember that Equation 3.8 was derived by using Stirling's approximation (Section 2.2.4) when approximating Equation 3.1 by Equation 3.2. Even more exciting we could deduce that it is impossible to sense a direction of time since entropy does not evolve towards a particular value. In fact, time does not have a direction on a microscopic scale. This is because the processes, here rolling a die, are symmetric in time. There is no direction of time when a single particle changes its state. We could also rewind the process and would not see any difference.
However, the situation changes dramatically if we repeat our experiment for N being sufficiently large in order to fit Stirling's approximation. For $N = 200$ dice (Figure 43) entropy converges to its maximum after roughly 500 time steps and the final histogram of the pips/numbers shown faced up is flat. Also the mean fluctuates much less (3.5 ± 0.2). Since N is large enough so that Stirling's approximation is valid, we have a macroscopic system. For even larger N (Figure 44), entropy and the mean (3.5 ± 0.1) fluctuate even less and the final histogram of the pips/numbers shown faced up is extremely flat. Note, that the fluctuations around the mean are normally distributed and we would just recover Figure 39 (see also Section 2.6.8)
Hence, we conclude that a **(macroscopic) closed system reaches maximum entropy, because it is the most likely macro state we could observe**. If N is increasing, the probability to observe the system at maximum entropy after some time increases even further. Once a system has reached maximum entropy, it does not change any further, **it has reached an equilibrium**, unless an external event influences the system. How *fast* the system will reach its equilibrium (i. e. maximum entropy) is not inherent to the equations for entropy and can differ dramatically for different systems. The speed of these processes is mainly linked to the temperature as we will see in the next sections. If $T = 0\,K$, there is no process, hence no sense of time.

3.1.1 An Uniform Distribution Has Maximum Entropy

I stated many times in the last section that the uniform (or flat) distribution has highest entropy. I now like to prove this statement: Consider a system of I micro states where p_i is the probability to obtain a particular micro state i. We know from before that these probabilities will be positive, and must satisfy the normalization constraint

$$P = \sum_i p_i = 1\,. \tag{3.9}$$

We now ask for the probability density distribution $p_i(i)$ that corresponds to maximum entropy. The motivation is that systems tend to evolve towards equilibrium that equals maximum entropy. The maximum entropy occurs when $dS = 0$ (necessary condition), so we will take the derivative of

$$S = S(p_1, \cdots, p_I) = -\sum_i p_i \ln p_i \tag{3.10}$$

with respect to the probabilities p_i,

$$dS = 0 = -\sum_i (1 + \ln p_i)\, dp_i\,. \tag{3.11}$$

To determine the probability distribution that results in maximum entropy, we will make use of the method of Lagrange multipliers (Section 2.1.5), where the constrain is given by Equation 3.9. According to our findings in Section 2.1.5, we can write down the condition

$$\underbrace{\frac{\partial S_i}{\partial p_i}}_{=-\ln p_i - 1} = \lambda \underbrace{\frac{\partial P}{\partial p_i}}_{=1}\,, \tag{3.12}$$

with the proportionality constant λ, which is the Lagrange multiplier.
This can then be rearranged to obtain an equation for the probability, which remains true for all i:

$$p_i = e^{(\lambda - 1)} \qquad \forall\, i\,. \tag{3.13}$$

Since λ is a constant, we see that p_i must also be constant. By simply summing our equation over all micro states (that must yield one, Equation 3.9), we find that

$$P = 1 = \sum_{i=1}^{I} p_i = \sum_{i=1}^{I} e^{(\lambda - 1)} = I e^{(\lambda - 1)}\,, \tag{3.14}$$

which implies that the distribution must be uniform! Hence, according to the above equation

$$p_i = e^{(\lambda - 1)} = \frac{1}{I}\,. \tag{3.15}$$

All p_i have the same value, one divided by the number of states that is an uniform distribution (c. f. example III in Section 2.6.3) and therefore proofs that the well ordered histogram in the left panel in Figure 40 has indeed highest entropy.

3.2 The Boltzmann Distribution

We have yet to make a connection between Equation 3.8 and thermodynamics. To show this relationship, we will derive a distribution for the micro states i under the physically relevant and very general constrain of conserved energy.
Suppose that in each state i, there are n_i particles with the energy ϵ_i, so that the total energy of the system is $E = \sum_i n_i \epsilon_i$ and the total number of particles is $N = \sum_i n_i$. Suppose further that the system is isolated and therefore does not exchange particles with its environment ($dN = 0$, but the particles are allowed to change their state so that $dn_i \neq 0$) and does not lose or gain energy ($dE = 0$). Under these conditions, the system will reach equilibrium. The equilibrium state is achieved when the entropy of the system reaches its maximum (see the dice example), i. e. when the multiplicity of the states W has reached its maximum. This implies $dW = 0$ at equilibrium.
We now have a new constraint (conservation of energy) that we must take into account when determining the distribution! This leads to the following question: What is the distribution of the energy quanta ϵ_i, if we only know the total (internal) energy ($dE = 0$) of the system?
Again, we will use $\ln W$, so Equation 3.1 reads

$$\ln W = \ln (N!) - \ln (n_1! n_2! \dots n_i! \dots n_I!) \; . \tag{3.16}$$

In order to simplify Equation 3.16 for large N we will follow our previous logic, and apply Stirling's approximation (Figure 2.2.3). Thus, we obtain

$$\begin{aligned} \ln W &\approx N \ln N - N - \sum_i \left[n_i \ln (n_i) - n_i \right] \\ &= N \ln N - N - \sum_i n_i \ln (n_i) + \sum_i n_i \\ &= N \ln N - N - \sum_i n_i \ln (n_i) + N \, . \end{aligned} \tag{3.17}$$

Here we made use of the fact that the sum of all n_i is just N, allowing us to cancel two of the original terms. This leads to

$$\ln W \approx N \ln N - \sum_i n_i \ln (n_i) \, . \tag{3.18}$$

The next step is to apply the equilibrium condition $dW = 0$ to Equation 3.18. Since W is not changing, its logarithm $\ln W$ will not change either, therefore we are allowed to use $d(\ln W) = 0$ for simplicity. Furthermore, we know that $d(\ln W) = 0$ corresponds to $dS = 0$, thus we are searching for a function of p_i yielding an extreme of S. Both terms in Equation 3.18 are of the form $x \ln x$, so we will make use of the product rule for derivatives yielding

$$\begin{aligned} d(x \cdot \ln x) &= \ln x \cdot dx + x \cdot d \ln x = \ln x \cdot dx + x \frac{dx}{x} \\ &= \ln x \cdot dx + dx. \end{aligned} \tag{3.19}$$

Applying this to Equation 3.18 we obtain

$$d (\ln W) = 0 \approx \ln (N) \, dN + dN - \sum_i \left[\ln (n_i) \, dn_i + dn_i \right] . \tag{3.20}$$

Since $\sum_i dn_i = dN = 0$, and change of the particle number in each state $dn_i \neq 0$, we arrive at

$$d(\ln W) \approx -\sum_i \ln(n_i)\, dn_i\,. \tag{3.21}$$

Since we want to find the distribution of $n_i(i)$ where $d(\ln W) = 0$ subject to the constrains $dN = 0 = \sum dn_i$ (conservation of mass) and $dE = 0 = \sum \epsilon_i dn_i$ (conservation of energy) we can again use the method of Lagrange multipliers, as we saw in the previous section for the proof of maximum entropy. Having two constrains this time, we need two multipliers λ_1 and λ_2, that are usually denoted as α and β, respectively, so that the relation

$$\underbrace{\frac{\partial}{\partial n_i}(\ln W)}_{=-\ln(n_i)} - \alpha \underbrace{\frac{\partial N}{\partial n_i}}_{=1} - \beta \underbrace{\frac{\partial E}{\partial n_i}}_{=\epsilon_i} = 0 \tag{3.22}$$

must hold. This leads to

$$\alpha + \beta\epsilon_i + \ln(n_i) = 0 \tag{3.23}$$

and solving for n_i, we find

$$n_i = e^{-\alpha} e^{-\beta\epsilon_i}\,. \tag{3.24}$$

where $e^{-\alpha}$ is a constant (since α was introduced as constant).
By summing over n_i, we can obtain an equation for N:

$$N = e^{-\alpha} \sum_i e^{-\beta\epsilon_i} = e^{-\alpha} Z\,. \tag{3.25}$$

We can now express $e^{-\alpha}$ as a much more meaningful constant N/Z. Here, Z is called the *partition function*, and is used to describe the statistical properties of a system in thermodynamic equilibrium. Its importance will become apparent in the following sections. Finally, we have

$$\boxed{n_i = \frac{N}{Z} e^{-\beta\epsilon_i}}\,, \tag{3.26}$$

the famous *Boltzmann distribution*.
Now we can easily obtain a value for the probability p_i from just by dividing Equation 3.26 by N:

$$p_i = \frac{n_i}{N} = \frac{1}{Z} e^{-\beta\epsilon_i}\,. \tag{3.27}$$

Furthermore, one can show that $\beta = 1/kT$ where T is the temperature and k is the *Boltzmann constant* $k = 1.38\cdots \times 10^{-23}$ J/K (which can be taken at face value, as the derivation is beyond the scope of this lecture). Hence, Equation 3.26 reads

$$\boxed{n_i = \frac{N}{Z} e^{-\epsilon_i/kT}}\,. \tag{3.28}$$

The Boltzmann distribution is shown in Figure 45. This figure makes it clear that one is less likely to find particles in a higher energy state (ϵ_i), and more likely to find them in a lower energy state. If a particle turns into a state of very high energy, it has to "take" it from the other particles (energy is conserved) and less energy is left for the other particles. Therefore one would expect to find a distribution like the exponential relation in Figure 45. If the temperature is increased, the probability of being in a given state i of higher energy ϵ_i increases. This means that the distribution flattens considerably, and higher energy states

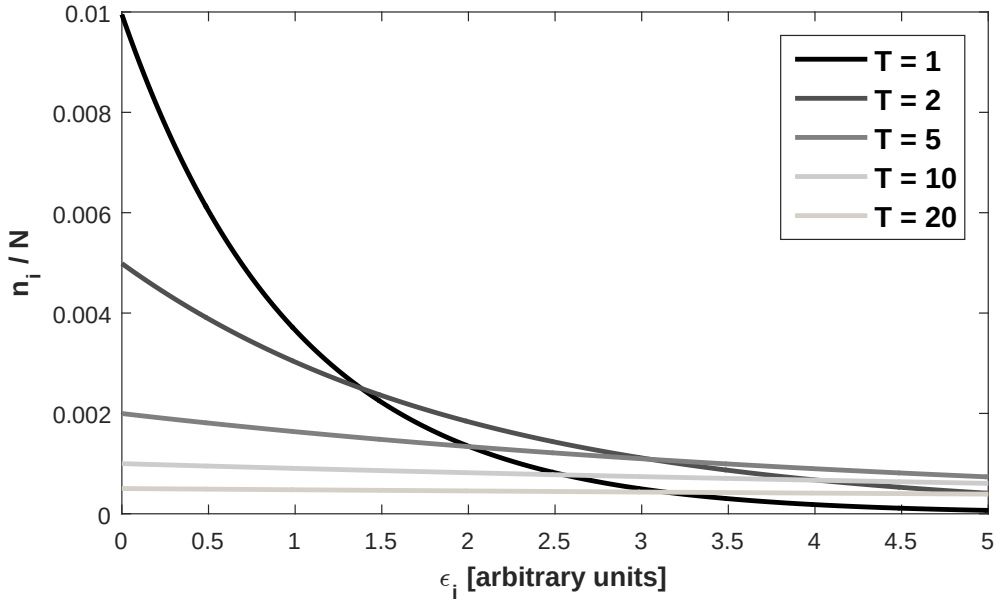

Figure 45: The Boltzmann distribution for different temperatures T (in arbitrary units). Note, that the Boltzmann distribution becomes flat for larger T.

become accessible.
Since we obtained Equation 3.26 via $d(\ln W) \sim dS = 0$, **the Boltzmann distribution appears if the system exhibits maximum entropy** [28].
Thanks to the Boltzmann distribution we can now connect the microscopic and statistical view of thermodynamics to macroscopic properties, like free energy or temperature, that we observe in our daily life. This will first be demonstrated by examining the properties of an ideal gas in Section 3.2.3, and again in Section 3.3 when we derive the thermodynamic potentials.
Generally, the number of particles per state obey a Boltzmann distribution in equilibrium if the only constrains are the conservation of energy and total number of particles. The exponential always contains the ratio of the particle energy ϵ_i in state i and the thermal energy kT. Depending on the system, the particle energy can be electrical energy qV in a electrical field of voltage V, kinetic energy $\frac{1}{2}m\,v^2$, chemical energy $\mu\,n_i$ and other forms of energy (e. g. $\sqrt{p^2c^2 + m^2\,c^4}$) or even a sum of the different kinds of energy. The scheme is always the same and some examples will follow now.

3.2.1 The Boltzmann Distribution Explains Potential Difference Across the Membrane of Nerve Cells

There is a steady state potential difference across the membrane of nerve cells, which is caused by different concentrations of ions. How can we use the Boltzmann distribution to quantify this difference?
The energy of a particle with charge q in an electrical potential of voltage V is qV (Section 2.1.8). Particle charge is represented by integer multiples of the elementary charge $e \approx 1.602 \times 10^{-19}$ coulombs. For example, potassium ions have the charge $q = +e$, magnesium ions have $q = +2e$, and an electron has $q = -e$ (note that the unit electron volt

[28] Actually $d(\ln W) = 0$ only points to an extremum and therefore could also be a minimum or an inflection point. However, Equation 3.22 implies a maximum since it leads to a negative second derivative (α and β are always positive).

eV originates from this definition).
From the Boltzmann distribution, the probability p to find an ion in the energy state qV is $p = \frac{1}{Z}e^{-qV/kT}$. Since the probability of finding an ion in a particular state is proportional to its concentration c we can write $c \sim p$. Thus, the ratio of concentrations across a membrane or between different regions in a cell is

$$\frac{c_1}{c_2} = \frac{e^{-qV_1/kT}}{e^{-qV_2/kT}}\,. \tag{3.29}$$

This yields the potential difference

$$\boxed{V_2 - V_1 = \frac{kT}{q}\ln\frac{c_1}{c_2}}. \tag{3.30}$$

Equation 3.30 is called the Nernst[29] equation.
The Nernst equation illustrates directly how a biological process (the communication between cells) is affected (and limited) by natural constants. Since the logarithm of the concentration ratio will not have a big affect on the order of magnitude of the potential difference, we look to the ratio kT/q to get an idea of the scale of the voltage. The ratio kT/q is in the mV scale, thus the voltage across a cell membrane is in the $10 - 100\ mV$ range, where thermal processes (kT) still play a role.

3.2.2 The Boltzmann Distribution Explains Ligand Binding on Receptors

Another process that can be described statistically by a Boltzmann distribution is the binding of ligands to receptors (i. e. oxygen to haemoglobin, transcription factors to DNA, etc.). Suppose that we have L ligands and one receptor with a binding site. This system has two macro states: bound and unbound. In order to determine their corresponding probabilities, we must first define and examine the possible microstates. These microstates will be characterized by the location and momentum of each unbound ligand with respect to the receptor. The set of all the possible values for position and momentum is known as the *phase space* (which is another concept developed in part by Boltzmann).
The momentum as a state is not important here since it is independent from the locations and only the right location is important (in first approximation) for the ligand to find the binding site. Let us begin with the unbound state, and say that there are N possible micro states to be occupied by our L ligands. This means that we have (from Equation 2.236)

$$\frac{N!}{L!\,(N-L)!} \tag{3.31}$$

possible ways to achieve an unbound macro state.
If each unbound ligand has the energy ϵ_u, then the total energy of the unbound system is $L\epsilon_u$. According to Equation 3.28, these states are then distributed as $e^{-L\epsilon_u/kT}$. Now if **one** ligand attaches itself to the binding site on the receptor, we have one less ligand $(L-1)$ occupying our phase space! This means there are

$$\frac{N!}{(L-1)!\,(N-L+1)!} \tag{3.32}$$

different configurations of a macro state with one ligand bound. If the bound ligand has energy ϵ_b, then the total energy of the system is $(L-1)\epsilon_u+\epsilon_b$, and results in the distribution

[29] Walther Nernst, 1864 - 1941

$e^{-[(L-1)\epsilon_u+\epsilon_b]/kT}$. A simplistic visual representation of both of our macro states can be seen in Figure 46.

Now in order to calculate the probability to observe the system in a particular state, we

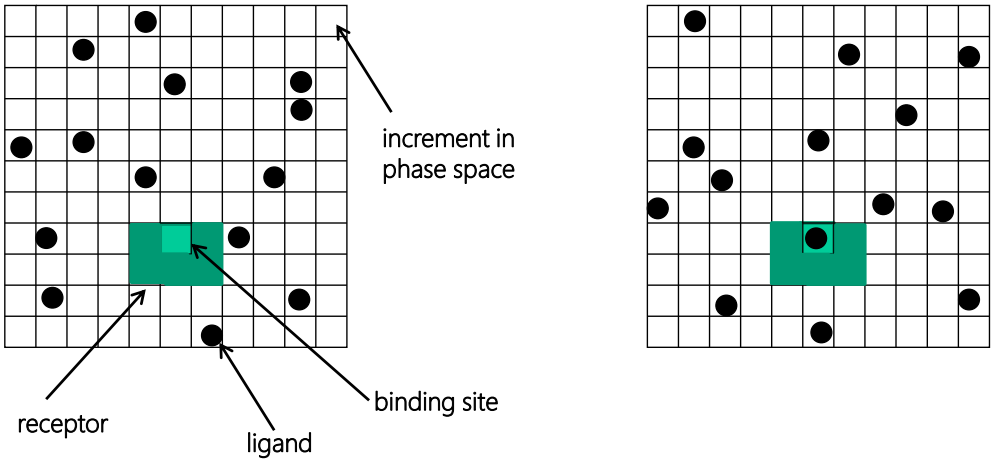

Figure 46: Left: One possible configuration of an unbound state. Right: One configuration of the bound states (one ligand bound).

need the partition function (Equation 3.25). In other words, we need the sum over all possible states weighted by their Boltzmann factor. As a result, we obtain

$$Z(N,L) = \frac{N!}{L!\,(N-L)!}e^{-L\epsilon_u/kT} + \frac{N!}{(L-1)!\,(N-L+1)!}e^{-[(L-1)\epsilon_u+\epsilon_b]/kT}. \tag{3.33}$$

We can use the approximation $\frac{N!}{(N-L)!} \approx N^L$ when $N \gg L$ (which is almost always valid for biological systems) to simplify Equation 3.33:

$$Z(N,L) = e^{-L\epsilon_u/kT}\left[\frac{N^L}{L!} + \frac{N^{L-1}}{(L-1)!}e^{-[\epsilon_b-\epsilon_u]/kT}\right]. \tag{3.34}$$

The weighted ratio of the number of all states with one ligand bound to the total number of states (bound and unbound) is then

$$p_{bound} = \frac{\frac{N^{L-1}}{(L-1)!}e^{-L\epsilon_u/kT}e^{-[\epsilon_b-\epsilon_u]/kT}}{e^{-L\epsilon_u/kT}\left[\frac{N^L}{L!} + \frac{N^{L-1}}{(L-1)!}e^{-[\epsilon_b-\epsilon_u]/kT}\right]} \tag{3.35}$$

the probability to find the system in a bound state, i. e. the probability for a ligand to bind to the receptor.

We simplify Equation 3.35 even further and introduce $\Delta\epsilon = \epsilon_b - \epsilon_u$:

$$p_{bound} = \frac{(L/N)\,e^{-\Delta\epsilon/kT}}{1+(L/N)\,e^{-\Delta\epsilon/kT}}. \tag{3.36}$$

Since L is the number of unbound ligands, and N is the number of available states that they can occupy, the ratio L/N is proportional to the ligand concentration c. We introduce a proportionality constant κ, and end up with

$$\boxed{p_{bound} = \frac{\kappa c\; e^{-\Delta\epsilon/kT}}{1+\kappa c\; e^{-\Delta\epsilon/kT}}}, \tag{3.37}$$

the *Hill*[30] equation for one ligand.
Equation 3.37 shows that the probability for a ligand to bind to the receptor depends (predictably) on the ligand concentration and also on the energy gain ($\Delta\epsilon$) that the ligand experiences while binding. If $\Delta\epsilon$ is negative, but with a large absolute value, the ligand loses a lot of energy while binding (hence, the bound state is preferred) and as expected $p_{bound} \longrightarrow 1$.

3.2.3 The Maxwell Distribution Derived from the Boltzmann Distribution

In an ideal gas (non-interacting particles with zero volume) the energy states ϵ_i are simply the kinetic energy of the individual molecules/atoms. So $\epsilon_i = \frac{1}{2}\,m v_i^2 = \frac{p_i^2}{2m}$ for a particle with velocity v_i, mass m and momentum $p_i = m\,v_i$ (Note, that the momentum p_i has nothing to do with the relative frequency p_i we used in previous sections. Unfortunately, both quantities are usually denoted as p_i). Momentum is a vector quantity, therefore the total momentum of one particle comes from the magnitude of its momenta in each direction $p_i^2 = p_{ix}^2 + p_{iy}^2 + p_{iz}^2$. For such a system, Equation 3.26 reads

$$\frac{n_i}{N} = \frac{1}{Z} exp\left(-\frac{p_{ix}^2 + p_{iy}^2 + p_{iz}^2}{2mkT}\right) . \tag{3.38}$$

This is a probability density function $\Psi_p(p_x, p_y, p_z)$. It tells us the probability to find a molecule/atom in the gas with a certain momentum as Equation 3.27 tells us the probability to find a molecule/atom in the gas with a certain energy. Probability density functions are invariably normalized, therefore the probability to find a molecule/atom in the gas with *any* momentum is one. We must perform an integral here and not a sum because momentum is continuous, meaning there are infinitely many states that the particle could occupy

$$\int_{p=0}^{p=\infty} \Psi_p\left(p_x, p_y, p_z\right) dp = \int_{p=0}^{p=\infty} \frac{1}{Z} exp\left(-\frac{p_{ix}^2 + p_{iy}^2 + p_{iz}^2}{2mkT}\right) dp = 1\,. \tag{3.39}$$

The problem is that we do not know the partition function yet. Solving the partition function would require somehow quantizing the system's momentum and space. Since by definition an ideal gas consists of molecules with zero volume, this is difficult. Fortunately, we do not need to know the partition function explicitly, and can make use of the normalization constraint. We simply have to find a constant C that normalizes the integral $\int_{p=0}^{p=\infty} \Psi_p\left(p_x, p_y, p_z\right) dp = 1$. Performing the integral in Equation 3.39 yields $C = (2\pi mkT)^{-3/2}$, so that the probability density function $\Psi_p(p_x, p_y, p_z)$ reads

$$\Psi_p(p_x, p_y, p_z) = \frac{1}{(2\pi mkT)^{3/2}} exp\left(-\frac{p_{ix}^2 + p_{iy}^2 + p_{iz}^2}{2mkT}\right) . \tag{3.40}$$

As stated already, once we know the energy of the system, we can derive all of its properties. We used the energy distribution described by Equation 3.26 to derive a distribution for the momentum of the molecules/atoms in an ideal gas, and since $p = mv$ it is straight forward to derive a speed distribution $\Psi_v(v)$. The relationship between momentum and velocity can be used to convert the distribution: $\frac{dp}{dv} = m$ and $\Psi_v(v) d^3v = \Psi_p(p_x, p_y, p_z)(dp/dv)^3 d^3v$. We need d^3v, because the speed contains the velocities in

[30] Archibald Vivian Hill, 1886 - 1977

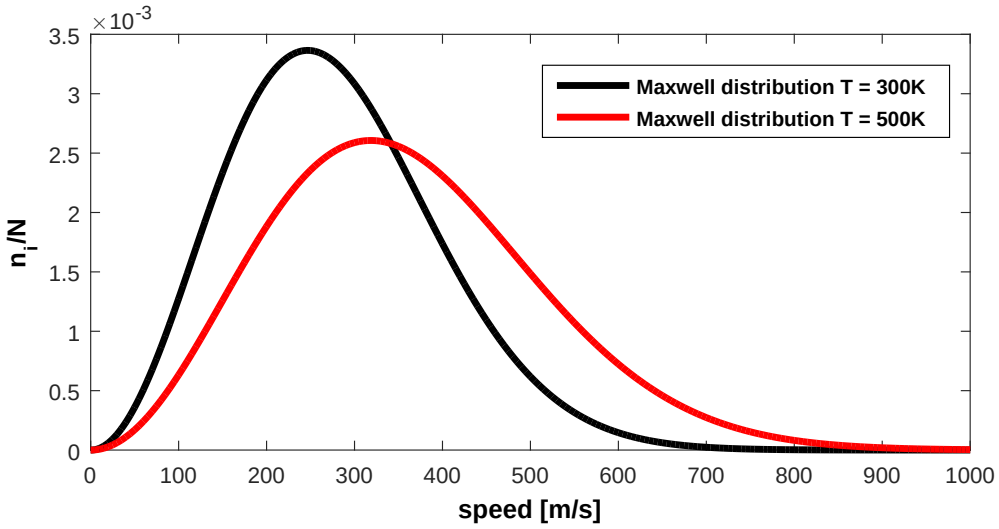

Figure 47: The Maxwell distribution for different temperatures T. The Maxwell distribution is derived from the Boltzmann distribution and shifts its maximum for higher T.

three directions and $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ (Section 2.1.1). Therefore $d^3v = 4\pi v^2\, dv$ (Section 2.5.2). Using this, Equation 3.39 turns into

$$\int_{v=0}^{v=\infty} \Psi_v\,(v)\, dv = \int_{v=0}^{v=\infty} \bar{C}\; exp\left(-\frac{mv^2}{2kT}\right) 4\pi v^2 dv = 1 \tag{3.41}$$

yielding

$$\boxed{\Psi_v\,(v) = 4\pi v^2 \left[\frac{m}{2\pi kT}\right]^{3/2} exp\left(-\frac{mv^2}{2kT}\right)}. \tag{3.42}$$

where $\bar{C}$ was another normailzation constant.
This is the *Maxwell distribution*[31], which is shown in Figure 47. It reaches its maximum $\frac{d\Psi_v}{dv} = 0$ at $v_{max} = \sqrt{\frac{2kT}{m}}$ and the mean (expectation value) $\langle v\rangle = \int_0^\infty \Psi_v v dv = 2\frac{v_{max}}{\sqrt{\pi}}$.
The mean velocity $\langle v\rangle$ implies that what we feel as temperature is caused by the speed of the particles. If $v = 0$, then $T = 0$ and vice versa.
Equation 3.42 yields also the mean quadratic value of the speed $\langle v^2\rangle = \int_0^\infty \Psi_v v^2 dv = 3\frac{kT}{m}$.
We know that the kinetic energy of the particle is a function of the speed squared, so:

$$\langle\epsilon\rangle = \frac{m}{2}\left\langle v^2\right\rangle = \frac{3}{2}kT\,. \tag{3.43}$$

Thus, the mean quadratic value of the speed yields the mean kinetic energy of a particle and therefore we obtain

$$\langle E\rangle = \langle\epsilon\rangle\, N \tag{3.44}$$

for the entire system containing N gas particles. Because $\langle E\rangle = U$ (see Section 3.3.1) the internal energy of the ideal gas is

$$U = \frac{3}{2}NkT\,. \tag{3.45}$$

This may be more familiar written terms of the gas constant $R = N_A\, k$ (where N_A is Avogadro's number[32]), and the number of moles in the system (n): $U = nRT$.

[31] James Clerk Maxwell, 1831 - 1879
[32] Lorenzo Romano Amedeo Carlo Avogadro, Conte di Quaregna e Cerreto, 1776 - 1856

3.3 Thermodynamic Potentials

3.3.1 Internal Energy, Entropy and Helmholtz Free Energy

We will now take a step further and obtain thermodynamic potentials through purely statistical derivations. If we put Equation 3.27 (Boltzmann distribution) into Equation 3.8 (entropy) we find

$$S = -\sum_i \frac{1}{Z} e^{-\beta\epsilon_i} \ln\left(\frac{1}{Z} e^{-\beta\epsilon_i}\right) , \tag{3.46}$$

that is

$$S = -\sum_i \left[-\frac{\beta\epsilon_i}{Z} e^{-\beta\epsilon_i} - \frac{1}{Z} e^{-\beta\epsilon_i} \ln Z\right] . \tag{3.47}$$

Using Equation 3.27 and $\langle\epsilon\rangle = \sum_i p_i\epsilon_i$, Equation 3.47 reduces to

$$S = \beta \langle\epsilon\rangle + \sum_i p_i \ln Z . \tag{3.48}$$

To arrive at an equation for total entropy (recall $S_{tot} = NS$) we introduce the *internal energy* U which is the average energy of the system: $\langle E\rangle = N\langle\epsilon\rangle$. Since $\sum_i p_i = 1$, and $\ln Z$ is a constant, we end up with

$$S_{tot} = \beta U + N \ln Z . \tag{3.49}$$

By convention, entropy in classical thermodynamics is multiplied by the Boltzmann constant, so that total entropy is $S_{tot} = k \ln W$ ($S \to kS$). From now on, we will use this thermodynamic entropy entropy of the entire system, so I will drop the index *tot* for simplicity. Equation 3.49 now reads

$$S = k\beta U + Nk \ln Z . \tag{3.50}$$

Using $\beta = 1/kT$ we find

$$S = U/T + Nk \ln Z . \tag{3.51}$$

This equation contains three parts: the entropy S that is a function of the distribution p_i of the states, the internal energy of the system and an additional term $Nk \ln Z$. Rearranging this equation

$$U - ST = -NkT \ln Z , \tag{3.52}$$

we see that the new term $-NkT \ln Z$ is an energy. This energy term is the difference between the internal energy and the entropy at given temperature T (ST is an energy). This means that in any thermodynamic process, a part of the energy in the system is consumed by entropy (re-arranging of the states) and the difference $-NkT \ln Z$ is the energy that can be used as work, not U. Since **temperature and volume is held constant** so far (no dT and no dV term was introduced), no work is performed on the environment by the system. Therefore, the energy difference $-NkT \ln Z$ equals the **maximum amount of useful work that can be extracted from the system**.
Following this interpretation, this energy difference is called *free energy* F

$$\boxed{F := -NkT \ln Z} . \tag{3.53}$$

Rearranging Equation 3.51, we find that

$$\boxed{F = U - TS} . \tag{3.54}$$

F is often known as the Helmholtz[33] free energy, and **is not** the same thing as the free enthalpy G (also known as Gibbs free energy). For some reason many biochemistry textbooks do not distinguish between F and G and call them both "free energy", which is simply incorrect. This would be like saying that pressure and mass density are the same since both are a kind of density, or that a magnetic field and an electric field are the same since both emerge from the electromagnetic tensor. To avoid any misunderstanding **I'll call F Helmholtz free energy and G Gibbs free energy from now on**.
From Equation 3.54 we see that F reaches its minimum when S reaches its maximum. Hence, F is small if the system has reached its equilibrium. For example, consider a system at temperature T_s that is placed in a thermal bath at temperature T_b (a thermal bath is a system large enough that it's temperature remains effectively constant when it comes into thermal contact with another system). Equilibrium occurs when $T_s \to T_b$. At this point, the entropy in the system has reached a maximum, and the free energy has reached its minimum.
As stated already, part of the internal energy U will always be "lost" to entropy. We know this loss as *heat*, defined as $Q = T\Delta S$, where ΔS is the change in entropy during a process that occurs at temperature T.**The energy that is available is the free energy F.**

We already have three thermodynamic potentials (Section 2.1.8): F, U and S (we will show that Q is not a potential in Section 3.3.4). A change of F is then

$$dF = dU - dTS - TdS\,. \tag{3.55}$$

You may have noticed that Equation 3.53 does not in any way account for changes in volume or pressure (no work performed on the environment), both of which are important to consider when chemical reactions occur. This will be addressed in the following section.

3.3.2 Enthalpy & Gibbs Free Energy and again Internal Energy and Entropy

The internal energy U is the mean of all portions of energy quanta ϵ_i, i. e. the sum (or integral) over all ϵ_i weighted by the probability density function p_i: $U := \sum_i p_i\epsilon_i$ (Section 3.3.1). In a certain sense, the internal energy tells us the energy to "create" a system out of nothing in empty space. However, the system needs volume that requires to apply the work $W = pV$. The sum of work and internal energy is called *enthalpy H*

$$H := U + pV \tag{3.56}$$

and

$$dH = dU + pdV + Vdp\,. \tag{3.57}$$

The energy of a system changes if one applies pressure to it, or if the volume is changed (Equation 3.57). If we compress a gas (i. e. decreasing the volume), we apply some work (W) and heat (Q) up the gas. On the other hand, a dense gas would freely expand into a void and cool down (that is the reason why a spray feels cold). We know already that all these processes are not completely reversible. There will be always a part of the energy (TdS) that is lost by entropy changes due to heat (Q) release. Thus, neither work, nor heat can be thermodynamic potentials (see Section 3.3.4 for a mathematical justification) - but both change the internal energy U. Another thing to keep in mind is that neither heat

[33] Hermann von Helmholtz, 1821 - 1894

nor work are a property of the system, while internal energy is. Putting these statements into equations we obtain the first law of thermodynamics:

$$dU = \delta Q - \delta W = TdS - pdV. \tag{3.58}$$

One can see that this is an adaptation of the law of conservation of energy for a thermodynamic system. A commonly used notation is δ instead of d to emphasize that Q and W are not potentials (Section 3.3.4). It is $-p\,dV$ in Equation 3.58, because it requires energy to decrease the volume.
In the case of chemical reactions the number of a particular species N_i changes. The energy that is required to insert or to remove a particle of species i is called *chemical potential* μ_i. Thus, Equation 3.58 reads

$$\boxed{dU = TdS - pdV + \sum_i^K \mu_i dN_i}, \tag{3.59}$$

where we sum over all K different species. The internal energy $U = U(S, V, N)$ in Equation 3.59 is now complete. Since U depends on S, V and N, these variables are called *natural variables* of U.
If we put Equation 3.59 into Equation 3.57 we find

$$\boxed{dH = TdS + Vdp + \sum_i^K \mu_i dN_i}. \tag{3.60}$$

Also the enthalpy $H = H(S, p, N)$ is now complete. The natural variables of H are S, p and N. In the same manner we can revisit F by using Equation 3.59 for Equation 3.54 and get

$$\boxed{dF = -SdT - pdV + \sum_i^K \mu_i dN_i}, \tag{3.61}$$

$F = F(T, V, N)$ with its natural variables T, V and N.
Using Equation 3.59 we can now pin down the entropy (Equation 3.8) on the macroscopic quantities like V, T etc:

$$\boxed{dS = \frac{1}{T}dU - \frac{p}{T}dV + \sum_i^K \frac{\mu_i}{T} dN_i}, \tag{3.62}$$

$S = S(U, V, N)$ with its natural variables U, V and N.
Having all these potentials (Equation 3.59 – Equation 3.62) we can now derive any macroscopic quantity from the partition function and/or the micro states. For example if Z is known and $dT = dN = 0$ for a given system we use $F = -kT \ln Z$ (Equation 3.53) and $dF/dV = -p$ (Equation 3.61) to derive the pressure.
We can now ask whether there is an equivalent to the Helmholtz free energy F if we start from enthalpy H, i. e. a *free enthalpy*? Indeed, we can write

$$G := H - TS \tag{3.63}$$

and

$$dG = dH - TdS - SdT. \tag{3.64}$$

With Equation 3.60 we find

$$\boxed{dG = Vdp - SdT + \sum_{i}^{K} \mu_i dN_i} \tag{3.65}$$

the free enthalpy or *Gibbs* [34] *free energy*.
G depends on the natural variables p, T and N. Under laboratory conditions or in a cell it is very easy to keep T or p constant in contrast to e. g. keeping S constant using H as a potential. Thus, G is the most suitable thermodynamic potential for us. Similar to F, G reaches its minimum in equilibrium.

The thermodynamic potentials seem to be something abstract, but the following example should clarify some things: Suppose you are the board engineer on the space ship enterprise and you want to beam Mr Spock on the surface of a planet. How much energy do you need? First, you have to conserve the energy of Mr Spock himself, that is his internal energy $E_{spock} = U$. Then, when Mr Spock appears on the planet, his volume replaces the volume of the surrounding gas on the planet. This requires work, and costs the energy pV. The energy to create Mr Spock out of nothing equals $H = U + pV$.
When Mr Spock replaces the volume by his own volume (Figure 48), the locations and momenta of the gas particles change, i. e. the micro states re-arrange. According to Equation 3.8 this changes entropy and therefore gives an additional energy term TS. This part only leaves the (Helmholtz) free energy $F = U - TS$. One can imagine the reverse process, when removing Mr Spock from the planet again. The energy we would recover as work would be F (not U) plus the work done by the collapsing atmosphere. Thus, you have to compensate the part of the Helmholtz free energy **and** the work that is required to replace the volume (pV). Hence, the **total** energy E you need is $E = U + pV - TS$, that is just G! Note the sign of TS – it is negative, meaning that you **gain** this part of energy, because you can only get $F = U - TS$ in the reverse process (and then $-pV$).

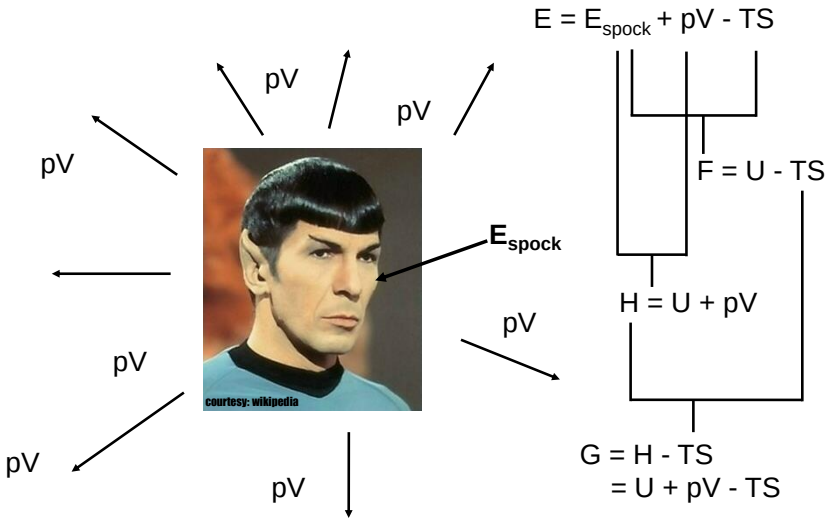

Figure 48: The energy that is required to beam Mr Spock depends on his internal energy E_{spock} as part of the free energy F that takes reconfiguration of states into account and the energy that is required to replace his volume against the pressure. The sum of all components is Gibbs free energy G.

[34] Josiah Willard Gibbs, 1839 - 1903

3.3.3 Legrende Transformations

Occasionally, you will probably hear something about *Legrende*[35] *transformations* in connection with thermodynamics. The strict mathematical formulation is not trivial, but we can boil it down to something that is actually very simple.
Suppose you use a thermodynamic potential like U, but it is hardly feasible to keep entropy constant in an experiment (c. f. Equation 3.59). Thus, U is inappropriate and you look for a potential that is more suitable. Taking the entropy part in Equation 3.59 we see that $TdS = d(TS) - SdT$ and we write Equation 3.59 as

$$dU = d\left(TS\right) - SdT - pdV + \sum_{i}^{K} \mu_i dN_i \tag{3.66}$$

and join $d(TS)$ to the lhs of Equation 3.66:

$$d\left(U - TS\right) = -SdT - pdV + \sum_{i}^{K} \mu_i dN_i. \tag{3.67}$$

Since $F = U - TS$, Equation 3.66 is identical to Equation 3.61. We changed from U to F and $dT = 0$ is by far easier to accomplish than $dS = 0$. This transition is called *Legrende transformation*.
If we go back to Equation 3.59 and we don't want to keep the volume constant we just set $pdV = d(pV) - Vdp$ and derive

$$d\left(U + pV\right) = TdS + Vdp + \sum_{i}^{K} \mu_i dN_i \tag{3.68}$$

the enthalpy H (Equation 3.60).
Using Legrende transformations, we can infer any thermodynamic potential, which allows us to conveniently tailor the equations to our needs! Legendre transformations can also be used with quantities that are not thermodynamic potentials (which can be proven using so called Maxwell relations).
Thanks to Legrende transformations we can summarize the thermodynamic potentials as $d(pV) = dH - dU = dG - dF$ and $d(TS) = dU - dF = dH - dG$. Figure 49 shows what the most suitable thermodynamic potential is, depending on the conditions.

	V=const	p=const
S=const	dU	dH
T=const	dF	dG

Figure 49: What the thermodynamic potential of the choice is depends on the conditions.

[35] Adrien-Marie Legendre, 1752 - 1833

3.3.4 Why isn't Heat a Thermodynamic Potential?

In the previous sections I stated that heat (Q) is not a potential, meaning that δQ is not an exact differential (Section 2.1.8). Since it is not immediately clear why this is true, I would like to prove it in this section.
We know that $\delta Q = TdS$ (Equation 3.58) and can use this to write Equation 3.59 in terms of Q so that we have $Q = Q(S, V, N)$

$$\delta Q = dU + p\,dV - \mu\,dN\,. \tag{3.69}$$

We will begin by assuming that heat is in fact a potential, and it will become apparent that this is a false assumption. So, we start with

$$\begin{aligned} dQ = \frac{\partial Q}{\partial V}dV + \frac{\partial Q}{\partial S}dS + \frac{\partial Q}{\partial N}dN =& dU + p\,dV - \mu\,dN \\ =& \frac{\partial U}{\partial S}dS + \frac{\partial U}{\partial V}dV + \frac{\partial U}{\partial N}dN + p\,dV - \mu\,dN \\ =& \frac{\partial U}{\partial S}dS + \left(\frac{\partial U}{\partial V} + p\right)dV + \left(\frac{\partial U}{\partial N} - \mu\right)dN. \end{aligned} \tag{3.70}$$

Comparing the dV parts in Equation 3.70 leads to

$$\frac{\partial Q}{\partial V} = \frac{\partial U}{\partial V} + p\,. \tag{3.71}$$

If Q was a potential, then the mixed second derivatives must be identical (Equation 2.97). Let's try $\frac{\partial^2 Q}{\partial S \partial V}$; hence, write down the derivative of Equation 3.71 wrt entropy S:

$$\frac{\partial^2 Q}{\partial S \partial V} = \frac{\partial^2 U}{\partial S \partial V} + \frac{\partial p}{\partial S} \tag{3.72}$$

If we compare the dS parts in Equation 3.70 you can see that

$$\frac{\partial Q}{\partial S} = \frac{\partial U}{\partial S} \tag{3.73}$$

and therefore

$$\frac{\partial^2 Q}{\partial V \partial S} = \frac{\partial^2 U}{\partial V \partial S}\,. \tag{3.74}$$

Now, the mixed second derivatives of Q must be identical (combining Equation 3.72 with Equation 3.74)

$$\frac{\partial^2 U}{\partial S \partial V} + \frac{\partial p}{\partial S} = \frac{\partial^2 U}{\partial V \partial S}. \tag{3.75}$$

Since U is a potential $\frac{\partial^2 U}{\partial S \partial V} = \frac{\partial^2 U}{\partial V \partial S}$, which we can just rename to L for convenience. Equation 3.75 now reads

$$L + \frac{\partial p}{\partial S} = L. \tag{3.76}$$

We see that Equation 3.76 (or Equation 3.75) does state a contradiction - the lhs does **not** equal the rhs! Since all the mathematical operations that we performed are permitted, **the initial statement that Q is a potential must be wrong!**

Let's try the next example with enthalpy. According to Equation 3.60, enthalpy depends on S, p and N. On the other hand, according to the definition of enthalpy

$$dH = d\left(U + pV\right) = dU + pdV + Vdp. \tag{3.77}$$

Inserting $U = U(S, V, N)$ (Equation 3.59) into Equation 3.77 yields

$$dH = \frac{\partial U}{\partial S}dS + \left(\frac{\partial U}{\partial V} + p\right)dV + \frac{\partial U}{\partial N}dN + Vdp\,. \tag{3.78}$$

Comparing Equation 3.60 to Equation 3.78 we see that the dp part is identical - this fits anyway. But there is no dV part in Equation 3.60, whereas it appears in Equation 3.78. Is this a contradiction? No, because $\frac{\partial U}{\partial V} = -p$ (Equation 3.59) and therefore the dV part in Equation 3.78 is identical to zero. Thus, only the dS part and the dN part are interesting. Again, we check whether the mixed second derivatives of H are identical: $\frac{\partial H}{\partial S} = T$ (Equation 3.60), that equals $\frac{\partial U}{\partial S}$ (Equation 3.78). According to Equation 3.60, $\frac{\partial U}{\partial S} = T$ too. Hence, already the first derivatives are identical - thus the second derivatives must be identical too.
Comparing the dN part of Equation 3.78 in the same way we find that $\frac{\partial H}{\partial N} = \mu = \frac{\partial U}{\partial N}$. Again, already the first derivatives are identical. If both first derivatives of H are identical, then the mixed second derivatives are identical too. Hence, **H is a potential.**
In the same manner one can show that S, U, G and F are potentials too, whereas $\delta W = p\,dV$ is not. T is a potential too, but it turns out, that it is not very practical.

3.4 Entropic Forces

Since entropy is a thermodynamic potential like internal energy or Gibbs free energy, it implies a force (Section 2.1.8) $f \sim grad\, S$. These forces are called *entropic forces*. Entropic forces drive ions to form a crystal, appear as hydrophobic force, osmotic pressure or cause different quantum effects. Recently, there has even been debate about whether or not gravity can be considered an entropic force! It is quite natural to guess entropy as the underlying principle of many fundamental forces, because it is the only property that justifies a certain direction of time (since it describes irreversible processes, Section 3.1).
An entropic force of biological relevance is the **Asakura-Oosawa**[36] force that leads to conglomeration of macromolecules in a solution (like proteins in a cell). In such a scenario, neither temperature, pressure, nor the internal energy of the particles changes. Then, the Helmholtz free energy reads

$$dF = \underbrace{dU - SdT}_{=0} - T\, dS\,. \tag{3.79}$$

Using Equation 3.53, we find the relation

$$\frac{k}{Z}\, dZ = dS\,, \tag{3.80}$$

expressing the change of entropy per particle depending on the change of the partition function Z.
The interior of a cell is densely packed with different kinds of molecules. Let us consider a simplified picture with larger, squared, macro molecules and smaller, suspended particles of radius r (Figure 50). The centers of the smaller particles cannot access a volume close to the larger molecules (dashed boxes in Figure 50) because then the smaller molecules and the larger molecules would touch and they cannot get closer. If the dashed boxes overlap, the smaller molecules cannot access this volume, they "feel" a reduced volume (Figure 50, left). However, if the larger molecules get even closer, so that the dashed boxes overlap completely, the volume that cannot be accessed by the smaller molecules decreases, so that their **accessible volume increases** (Figure 50, right).
Suppose we have a given total volume V_{tot} and the volume ν of the small particles that are the solvent for the larger molecules. Thus, we have $m = V_{tot}/\nu$ different volume cells that can be occupied by one of the smaller molecules. In contrast to ligand - receptor binding (Section 3.2.2) the energy of the molecules does not change. Therefore, V_{tot} describes the real physical volume and ν the volume of the smaller particles and not an abstract phase space or the phase space increment, respectively.
With N particles and m volume cells, we have (Equation 2.236)

$$\frac{m!}{N!\,(m-N)!} \approx \frac{m^N}{N!} \tag{3.81}$$

different states. Hence, the partition function reads $Z \approx m^N/N!$. What changes from the left picture in Figure 50 to the right one is that the larger molecules in the picture on the right are arranged in that way, that the accessible volume for the smaller particles increases by dV_{tot}. The individual volume ν of each particle doesn't change and the number N of

[36] Sho Asakura, 1927 - today and Fumio Oosawa, 1922 - today. Note that different transcriptions of the names exist.

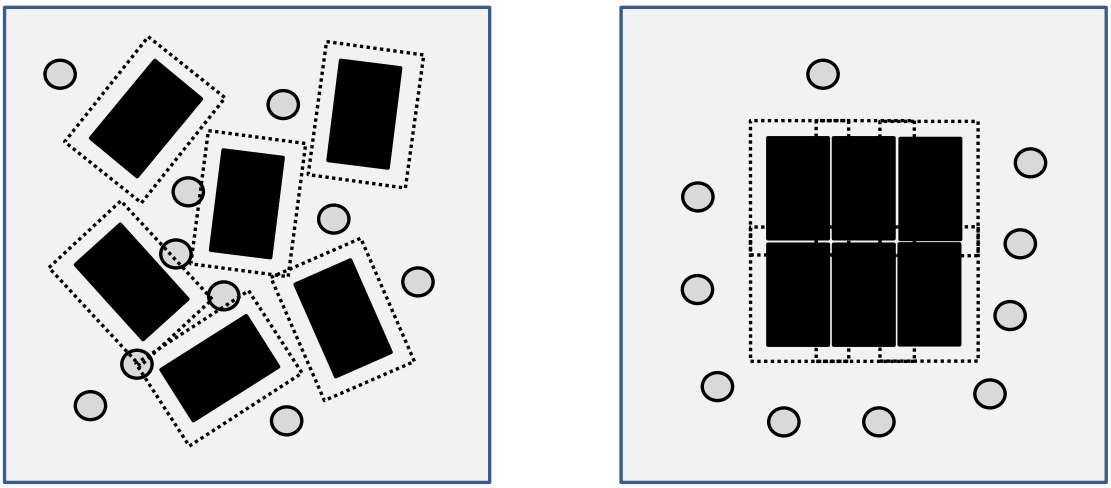

Figure 50: Macromolecules tend to conglomerate in a solution of smaller molecules even if each molecule is electrical neutral. This effect is caused by the gain of volume accessible for the smaller molecules that increases entropy.

particles does not change too, so that only **the total volume that is accessible to smaller particles does change**. Thus,

$$dZ = d\left[\frac{m^N}{N!}\right] = d\left[\frac{\left(\frac{V_{tot}}{\nu}\right)^N}{N!}\right] = \frac{N}{N!}\frac{V_{tot}^{N-1}}{\nu^N}dV_{tot} \tag{3.82}$$

and therefore

$$dS = \frac{kN}{V_{tot}}\,dV_{tot}\,. \tag{3.83}$$

dV_{tot} is the part of the volume that is gained for the smaller molecules.
We see that dS is positive, when dV_{tot} is positive. Increasing the volume accessible for the smaller molecules (that equals decreasing the *effective* volume of the larger molecules) leads to increasing entropy. Thus, the left picture in Figure 50 has a **lower entropy** than the right picture, although it looks less ordered. Let us assume that the smaller molecules are of spherical shape with the radius r, then they have the volume $\nu = 4\pi r^3/3$ and if they cannot access a given region between the macro molecules, they lose the volume $dV_{tot} = 4\pi r^2 dr$ and

$$dS = \frac{kN}{V_{tot}}4\pi r^2 dr\,. \tag{3.84}$$

The corresponding entropic force $f \sim grad\, S$ equals

$$\frac{dS}{dr} \sim f \sim \frac{kN}{V_{tot}}4\pi r^2\,. \tag{3.85}$$

According to Equation 3.51, the proportionality constant is $-T$ (just set dU to zero in Equation 3.51) and finally we obtain

$$\boxed{f = -4\pi\frac{kNr^2}{V_{tot}}T} \tag{3.86}$$

the Asakura-Oosawa force.
Note that the force is attractive (negative sign), so that the macro molecules will conglomerate, even if they are electrically neutral, and that the force increases with higher

temperature because the particles can move faster. It seems a bit contra intuitive that the ordered state (Figure 50, right) has higher entropy than the disordered state (Figure 50, left), but this is because describing entropy with order and disorder is just wrong and has nothing to do with the true meaning of entropy. The Asakura-Oosawa force is even frequently called "structure ordering force".

By definition pressure p is force per area. The surface area of a sphere is $4\pi r^2$ and the pressure is

$$\boxed{p = -\frac{n}{V_{tot}} kT}. \tag{3.87}$$

Equation 3.87 describes the *osmotic pressure*. This is the pressure needed to keep the n macro molecules away from conglomerating depending on temperature and solute (not solvent) concentration n/V_{tot}. Separating the macro molecules or keep them in one part of the volume, e. g. by a semi-permeable membrane, requires the same pressure, but changes its sign.
Often, Equation 3.87 (with positive sign) is called *van't Hoff*[37] equation.

[37] Jacobus Henricus van 't Hoff, Jr., 1852 - 1911

3.5 Chemical Reactions

We derived thermodynamic laws from pure statistical considerations. We will now treat chemical reactions in the same manner. This way of deriving chemical laws is necessary in order to understand the treatment in Section 5.2.
We introduced the chemical potential in Equation 3.59 more or less out of the blue. Although it is clear that the consumption or formation of different species of particles (such as in a chemical reaction) will cost or gain some energy, introducing a new, unknown variable like μ seems a bit provisional. Thus, I would like to give a statistic justification for the expression $\mu\, dN$.
We know from Section 3.2 the Boltzmann distribution as the distribution that tells us the probabilities to find particles in a certain energy state. The sum of the Boltzmann factors of each state is the partition function or partition sum Z. For a system of N **indistinguishable** particles, each of it may have access to the different energy states, we can therefore find the partition function

$$Q = \frac{Z^N}{N!}\,. \tag{3.88}$$

of the whole system if the **states are not overlapping** meaning that there are no interactions (or only very weak interactions) between the particles (e. g. if the particles are not too dense or within a certain temperature range). The direct connection from the microscopic partition sum to the macroscopic thermodynamic properties is given by Equation 3.53. We put Equation 3.88 into Equation 3.53 and apply Stirlings approximation (Equation 2.133). This results in

$$F = -NkT\ \ln\left(\frac{eZ}{N}\right)\ . \tag{3.89}$$

How does (Helmholtz) free energy change, if the number of particles changes (that is what happens during a chemical reaction)? This is just the derivative of F with respect to N and we find

$$\frac{\partial F}{\partial N} = -kT\ \ln\left(\frac{Z}{N}\right)\ . \tag{3.90}$$

We call this derivative μ, the chemical potential, and we see from Equation 3.61 that indeed $\frac{\partial F}{\partial N} = \mu$! Hence,

$$\boxed{\mu = -kT\ \ln\left(\frac{Z}{N}\right)}\,. \tag{3.91}$$

We discussed at the end of Section 3.3 that the free enthalpy G (or Gibbs free energy) is the most suitable potential to describe chemical reactions in biological systems.
Let us consider the simple reaction

$$A \underset{k_-}{\overset{k_+}{\rightleftharpoons}} B \tag{3.92}$$

where we have two kinds of particles N_A and N_B and the forward and back reaction rate constants k_+ and k_-, respectively (Section 4.3). Equation 3.65 reads then

$$dG = Vdp - SdT + \mu_A dN_A + \mu_B dN_B. \tag{3.93}$$

All processes in living biological systems occur in aqueous solutions leading relatively fast to temperature and pressure equilibrium – even for non-equilibrium processes. Usually, non-equilibrium processes are driven by $d\mu$ in biological systems. Therefore, we set $dT = dp = 0$. In equilibrium $dG = 0$, so that Equation 3.93 simplifies to

$$-\mu_A dN_A = \mu_B dN_B \tag{3.94}$$

in our specific case[38].
The total number of particles $N_{tot} = N_A + N_B$ is conserved so that $dN_{tot} = 0$ and $dN_B = -dN_A$ that leads to

$$\boxed{\mu_A = \mu_B} \tag{3.95}$$

under equilibrium conditions.
With Equation 3.91 we find that

$$\frac{N_B}{N_A} = \frac{Z_B}{Z_A}. \tag{3.96}$$

The structure of the two partition sums is known from Section 3.2, but we have to be

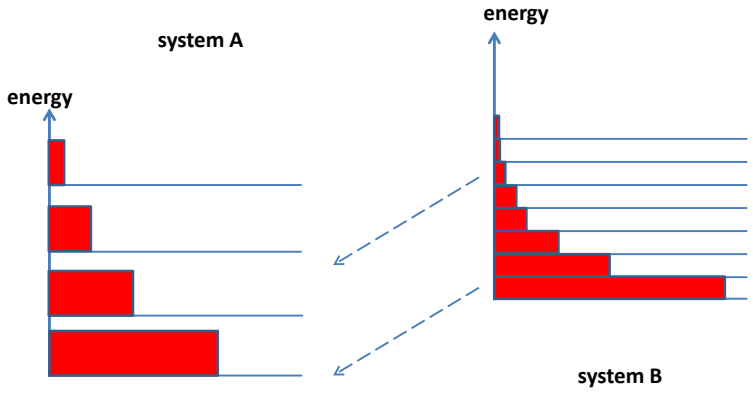

Figure 51: A chemical reaction means joining two (or more) systems with different energy states and different state densities. The joint system attributes as many particles as possible to the lowest energies.

aware that the energy levels and the density of the energy states in A and B are generally different, as illustrated in Figure 51. We denote $Z_A = \sum_i e^{-\epsilon_{Ai}/kT}$ and $Z_B = \sum_j e^{-\epsilon_{Bj}/kT}$ with different indices i and j to account for that. We can see in Figure 51 that both systems have a ground state - the energy level with the lowest energy ϵ_{A0} and ϵ_{B0}, respectively. Therefore, we write Equation 3.96 as

$$\frac{N_B}{N_A} = e^{-(\epsilon_{B0}-\epsilon_{A0})/kT} \frac{1 + e^{-(\epsilon_{B1}-\epsilon_{B0})/kT} + e^{-(\epsilon_{B2}-\epsilon_{B0})/kT} \ldots}{1 + e^{-(\epsilon_{A1}-\epsilon_{A0})/kT} + e^{-(\epsilon_{A2}-\epsilon_{A0})/kT} \ldots}. \tag{3.97}$$

The sums on the rhs of Equation 3.97 are a kind of reduced partition functions (I denote as $\bar{Z}_A$ and $\bar{Z}_B$, respectively). They show the individual energy levels with respect to the ground states. This makes sense, since we can only measure energy differences (i. e. it is a relative quantity).
If we use the definition of the *equilibrium constant* K we finally obtain

$$\boxed{K := \frac{N_B}{N_A} = e^{-(\epsilon_{B0}-\epsilon_{A0})/kT} \frac{\bar{Z}_B}{\bar{Z}_A}}. \tag{3.98}$$

A closer look at Equation 3.98 reveals that for example if $\epsilon_{B0} > \epsilon_{A0}$ the exponential is negative and therefore the ratio K is shifted in favour of A (ignoring the ratio of the partition functions, see later) and the reaction will go in the direction to A if $K < 1$. Hence, lower

[38] Having $\frac{\partial F}{\partial N} = \mu$ in mind shows that we could have started with F too. This maybe explains why life scientists often just use the term *free energy* for G, although they are not identical in general.

energy states are generally favoured.
Note, that K is actually not a real constant. It depends on pressure and temperature, but both are held constant in biological systems.

Let us now consider the reaction

$$aA \; + \; bB \underset{k_-}{\overset{k_+}{\rightleftharpoons}} cC\,. \tag{3.99}$$

In the same manner as before we obtain

$$\mu_A dN_A + \mu_B dN_B + \mu_C dN_C = 0 \tag{3.100}$$

as equilibrium condition.
It is useful to introduce a new variable ξ, where $dN_C = cd\xi$, $dN_A = -ad\xi$ and $dN_B = -bd\xi$. This is always possible, since a, b and c are just constants. Therefore, we find that

$$(c\mu_C - a\mu_A - b\mu_B)\, d\xi = 0 \tag{3.101}$$

is the equilibrium condition and we find (similar to Equation 3.98)

$$\boxed{K = \frac{N_C^c}{N_B^b N_A^a} = \frac{\bar{Z_C}^c}{\bar{Z_B}^b \bar{Z_A}^a} e^{-(c\epsilon_{C0} - a\epsilon_{A0} - b\epsilon_{B0})/kT}}. \tag{3.102}$$

Let us discuss the meaning behind Equation 3.102: If K is small, the reaction does not shift towards C. This occurs when the partition functions of A and B are large. In other words, if the number of possible configurations (states) of A and B is large. This can happen if the solution is too dilute and each particle has many volume increments to occupy, meaning that the likelihood of A and B coming together is low. This can also happen if the geometry of the particles is complicated such that A and B can only react if they are oriented in a particular direction to each other (think about ligand $-$ receptor binding where the ligand has to find the binding side). Also, the reaction has to be energetic favourable for the reactants, which is expressed in the exponential.
Equation 3.102 also states that a chemical reaction can occur even if it is **not energetically favourable from its internal energy** if the ratio of the partition sums overcomes the exponential. **The reaction is driven by entropy in this case.** Also, if the reaction is energetically favourable, the ratio of the partition sums can inhibit the reaction - the cost of entropy would be too large. That is a again a justification for choosing Gibbs free energy ($\equiv$ free enthalpy) as the favourable potential.

3.6 Summary and Further Reading

We saw in this short section that thermodynamics is nothing but doing statistics on an atomic level by counting states, deriving probability density distributions (that is mostly the Boltzmann distribution) and calculating means that equal macroscopic thermodynamic properties. In that sense, thermodynamics is not much more than flipping coins and rolling dice (Section 2.6). The intention of this section was to demystify some thermodynamic concepts and to show that everything can be straightforwardly derived from few principle assumptions (energy is conserved, maximum entropy approach).
To underline the biological implications, we discussed some examples like ligand binding, the potential in nerve cells and osmotic pressure. For further reading I strongly recommend [1] that is filled with biological examples including the complete, but brilliantly explained, mathematical background. The level of this book equals pretty much that what is expected from you in quantitative bioscience and illustrates how tightly mathematical modeling and the understanding of biological systems is bound together.
Often, it is very helpful to read about the same topic in another textbook to complement the acquired knowledge. I therefore suggest [2] and [3] for further reading.
Also, thermodynamics is required to understand the dynamics of chemical reactions, the driving processes of life. These are now investigated in more detail in the next main chapter.

References

[1] Ken Dill, Sarina Bromberg *"Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience"*, Taylor & Francis Inc; 2nd Rev ed. (13. December 2010)

[2] Philip Nelson *"Biological Physics: Energy, Information, Life"*, W H Freeman & Co (July 2003)

[3] Daniel V. Schroeder *"An Introduction to Thermal Physics"*, Addison Wesley Pub Co Inc; ed. New. (8. December 1999)

4 Reaction Kinetics

Chemical reactions are driven by the laws of thermodynamics as showed at the end of Section 3. A chemical reaction occurs if the system is in a non-equilibrium state and the reaction will last until the equilibrium is reached. The dynamics of these processes are subject to reaction kinetics. It turns out that the dynamics of a chemical reaction are described by differential equations. Classical reaction kinetics deal with large numbers of molecules so that small statistical fluctuations are negligible. Such a system is fully deterministic and no probabilistic properties are included. This works for amounts of more than one hundred molecules/atoms and on time scales in which a single reaction cannot be resolved. However, when investigating single molecule dynamics and reactions with a small number of molecules, we have to take statistical fluctuations into account that is subject to Section 5. This approach is actually more general since classical reaction kinetics, that will be discussed here, is a special case of stochastic kinetics for large numbers and "low" temporal resolution.
The idea of this chapter is to explain how the properties of chemical reactions can be explored without solving the equations analytically by an qualitative analysis of their dynamics. The scope and the outline of this chapter follows the approach in [3] that I strongly recommend for further reading.

4.1 Basic Nomenclature

Chemical reactions are described by ordinary differential equations (ODEs, see Section 4.3). Therefore, I first like to give an overview about the nomenclature:

Ordinary Differential Equation (ODE) denotes an equation that contains total derivatives of the form $\frac{d^k f(x)}{dx^k}$, where $k \in \mathbb{N}$. An ODE is called "ODE of n^{th}" order if the highest order derivative is of order n.
A generic example for an ODE of first order is the equation for exponential growth:

$$\frac{dx}{dt} = k\,x \tag{4.1}$$

Partial Differential Equation (PDE) denotes an equation that contains partial derivatives of form $\frac{\partial^k f(x)}{\partial x^k}$, where $k \in \mathbb{N}$.

A PDE is called "PDE of n^{th}" order if the highest order partial derivative is of order n. PDEs get important once the system we want to describe is multidimensional.
An example of a PDE of second order is the one dimensional[39] wave equation:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \tag{4.2}$$

Linear Differential Equation denotes a differential equation which does only contain x in linear terms, but no higher powers of x. Linear differential equations can be easily[40] solved analytically, even in higher dimensions.
An example is again exponential growth:

$$\frac{dx}{dt} = k\,x \tag{4.3}$$

[39] These are actually two dimensions: one in time and one in space.
[40] There exists a recipe that can be used to solve any linear differential equation.

Nonlinear Differential Equation denotes a differential equation that is not linear and thus contains higher powers of x. Nonlinear differential equations are usually hard to solve analytically. However there are still ways to derive interesting and useful properties from such systems. Moreover many interesting phenomena in biology[41] can only be explained by nonlinear equations.
As an example the exponential growth equation can be easily altered to:

$$\frac{dx}{dt} = k\,x^2\,. \tag{4.4}$$

[41] For example the Hodgkin-Huxley model for describing action potentials in neurons is basically a nonlinear PDE.

4.2 One-dimensional First Order Differential Equations

In the case of reaction kinetics usually the change in the system is fully determined by its current state. Thus reaction kinetics can be represented in one dimension as

$$\frac{dx}{dt} = f(x) \tag{4.5}$$

For example, $\frac{dx}{dt} = k\,x^2$. The structure of any chemical reaction is always that of Equation 4.5, where the term $f(x)$ can be any function of x and maybe some other additional variables.
Frequently, people use Newton's notation ($\dot{x} := \frac{dx}{dt}$) for the absolute temporal derivative of x so that we write Equation 4.5 in the more compact form $\dot{x} = f(x)$. This equation might look pretty simple but if $f(x)$ is non-linear there might be no explicit solution for the trajectory $x(t)$. For such cases we need to restrict ourselves to a more qualitative analysis of the system that will be introduced now.

4.2.1 Qualitative Analysis of a Phase Portrait

To gain some intuition of how systems of differential equations can be analysed, even if there is no explicit solution for the trajectories, we will start with a basic example. Instead of plotting the solution $x(t)$, which in unknown, of the ODE we can visualize the function $\dot{x} = f(x)$ itself. Such a plot is called *phase portrait*. An example of a phase portrait is shown in Figure 52. One has to be aware that the phase portrait does not show $x(t)$, but the temporal derivative of x. One can treat this system as a particle moving on a line (here the x-axis). Since the flow $\dot{x}$ is given by $\dot{x} = f(x)$ the value $f(x)$ tells us how the particle moves at position x. For example if $f(x)$ is *positive*, x is *increasing* with time since $\dot{x} = f(x)$. If $f(x)$ is *negative*, x is *decreasing* with time. In the case of $f(x) = 0$, it means that there is no change of the quantity x since $\dot{x} = 0$. Thus, we can distinguish three qualitatively different cases:

$f(x) > 0$	the system moves to the right	$x(t)$ is increasing	
$f(x) = 0$	the system does not move at all	$x(t)$ is constant	(4.6)
$f(x) < 0$	the system moves to the left	$x(t)$ is decreasing	

The x-values for which $f(x) = 0$, i. e. the zeros, are called *fixed points* and are usually denoted by x^*. There are two different kinds of of fixed points, unstable fixed points, also called *repellers*, and stable fixed points, also called *attractors*. Attractors are stable in the sense that small perturbations will vanish after a certain time and the system will return to its initial state. At a repeller, small perturbations would increase with time and the system will not return to its initial state. Note, that a system that is **exactly** at the fixed point does not change regardless of the kind of the fixed point.
By looking at the phase portrait in Figure 52 we can deduce which fixed points are stable and which are unstable. The fixed point x_1^* is not stable. If a small perturbation would move x to the left, $f(x)$ (the derivative of x) would be negative and x would decrease even further. In case a perturbation would move x to the right, $f(x)$ would be positive and x would increase even further. The same applies for the fixed point x_3^*.
In contrast, a small shift to the left at fixed point x_2^* would increase x ($f(x)$ is positive) until it has reached the initial value ($\dot{x} = 0$) and a small perturbation to the right would decrease x ($f(x)$ is negative) until it has reached x_2^*. Hence, this fixed point is stable.

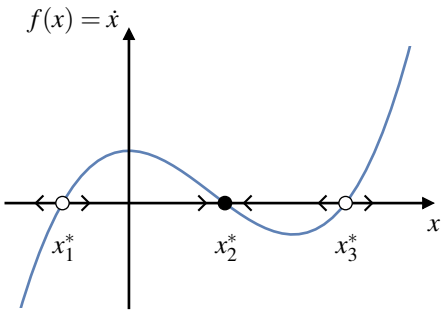

Figure 52: Exemplary phase portrait of an one-dimensional ODE of form $\dot{x} = f(x)$. The fixed points x_1^*, x_2^*, x_3^* are denoted by the three dots on the x-axis. By general convention filled dots denote stable fixed points and empty dots represent unstable fixed points. The arrows between the fixed points indicate the direction of the flow of the system. It is essential to keep in mind that the system only exists on the x-axis.

Thus, the stability of a fixed point can be inferred by the slope of $f(x)$:

$$\begin{aligned} \left.\frac{df(x)}{dx}\right|_{x=x^*} &> 0 \qquad \text{repeller/unstable} \\ \left.\frac{df(x)}{dx}\right|_{x=x^*} &< 0 \qquad \text{attractor/stable} \end{aligned} \tag{4.7}$$

Hence, only applying the rules Equation 4.6 and Equation 4.7 to Figure 52 gives already many insights to the qualitative behaviour of the system, without actually solving the equations. Sometimes it might be of interest at which x the system accelerates or decelerates. These points are called *turning points* of $x(t)$ and correspond to the minima (acceleration) and maxima (deceleration) of $f(x)$.

4.2.2 Population Growth

Before investigating chemical reactions, let us first discuss a simple, but very intuitive model: the model of unbound growth. The most common model for growth is simply exponential given by

$$\frac{dN}{dt} = r\,N\ . \tag{4.8}$$

Equation 4.8 states, that the change of number of individuals N at given time t is a product of growth rate r and the number of individuals (higher number of individuals can generate more offspring). This model, however, is not applicable in many cases since it describes a population that grows infinitely for $r > 0$. Most biological systems are limited by resources and space so that a decelerating part has to be included into Equation 4.8. The straight forward solution of this problem is to remove the constant growing rate r and replace it with a rate $R(N)$ that decreases if N is too large so that Equation 4.8 reads

$$\frac{dN}{dt} = R(N)\,N\ . \tag{4.9}$$

To get bound growth $R(N)$ should be decreasing with increasing population size N. Theoretically all kinds of decreasing functions could be possible for $R(N)$. Usually, people start with a simple model that would be a linearly decreasing growth rate like

$$R(N) = r\,(1 - \frac{N}{\kappa})\ , \tag{4.10}$$

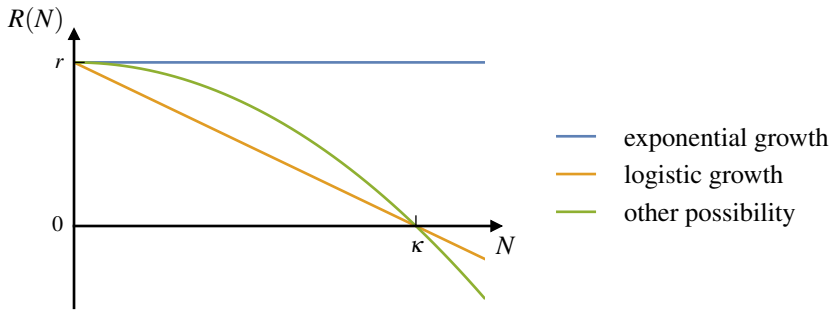

Figure 53: Three possible growth rates $R(N)$ as functions of the population size N. A constant growth rate (blue) yields exponential growth (Equation 4.8). A linearly decreasing growth rate (yellow. Equation 4.10) gives rise to a so called logistic growth. These are however not the only possibilities of how the growth changes as function of the population size. The green line illustrates another possibility.

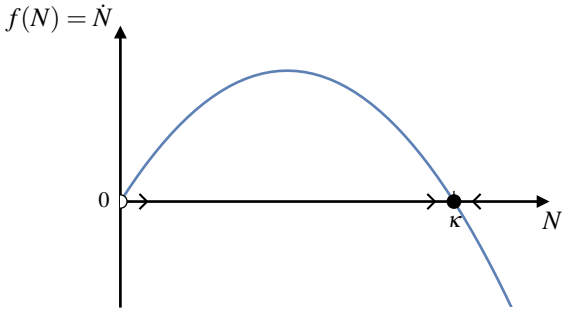

Figure 54: The phase portrait of the Verhulst equation with its two fixed points $N_1^* = 0$ (unstable) and $N_2^* = \kappa$ (stable). $\tilde{N} = \kappa/2$ is a turning point which represents a deceleration of the growth of the population.

where κ is the so called *carrying capacity*. According to Equation 4.10, the unit of κ is a number and it sets the limit upto which size the population is sustained by its environmental conditions. When N approaches κ, the growth decelerates continuously until it would reach $N = \kappa$[42]. At the point $N = \kappa$, $R(N) = 0$ and the size of the population does not change. If $N > \kappa$, $R(N)$ would be negative and the number of individuals decreases. The change of the growth rate $R(N)$ as function of population size is shown in Figure 53. By assuming $R(N)$ to be linearly decreasing we get the following ODE,

$$\frac{dN}{dt} = r\,N\,(1 - \frac{N}{\kappa})\,, \tag{4.11}$$

which is known as the *Verhulst equation* of bound growth.[43].

4.2.3 Qualitative Analysis of the Verhulst Equation

The Verhulst equation can now be analyzed using the phase portrait as explained in Section 4.2.1. The Verhulst ODE (Equation 4.11) is just a parabola and the fixed points of the dynamical system are straightforward to find via requiring,

$$\frac{dN}{dt} \stackrel{!}{=} 0 = r\,N\,(1 - \frac{N}{\kappa})\,. \tag{4.12}$$

[42] Note, that it would reach $N = \kappa$ only for $t \to \infty$

[43] Pierre-Francois Verhulst, 1804 - 1849

Therefore, the two fixed points are $N_1^* = 0$ and $N_2^* = \kappa$. The existence of the fixed point $N_1^* = 0$ is obvious because there cannot be any reproduction if there is no population. The stability of the fixed points can be deduced by investigating the phase portrait (Figure 54). Another way to derive the stability is to calculate df/dN, evaluated at the respective fixed-point, as done in Equation 4.7.

$$\left.\frac{df(N)}{dN}\right|_{N=N_1^*} = r\left(1 - 2\frac{0}{\kappa}\right) = r > 0 \qquad \Rightarrow \text{ unstable} \tag{4.13}$$

$$\left.\frac{df(N)}{dN}\right|_{N=N_2^*} = r\left(1 - 2\frac{\kappa}{\kappa}\right) = -r < 0 \qquad \Rightarrow \text{ stable} \tag{4.14}$$

Since this system has only one stable fixed point, it is called *globally stable*. This means that wherever the system starts it will always end up at N_2^*, with the only exception of starting exactly at the other fixed point $N_1^* = 0$. The stability of the second fixed point $N_2^* = \kappa$ is also the reason why κ is referred to as the *carrying capacity*, since $N = \kappa$ is the population size at which the system always returns to.
Another interesting feature is the turning point $\tilde{N} = \kappa/2$. As can be seen in Figure 54, the turning point is a maximum and therefore the system decelerates when it has passed it. Thus, we can summarize by this rather simple analysis:

- the population will always evolve towards $N_2^* = \kappa$
- the population growth accelerates till $\tilde{N} = \kappa/2$ and then decelerates
- the population grows fastest at $\tilde{N} = \kappa/2$ (maximum of $\dot{N}$)
- above $N_2^* = \kappa$ the population decreases towards $N_2^* = \kappa$

Although we did not solve the Verhulst ODE yet, we can now plot the solution $N(t)$ qualitatively for three different representative initial conditions ($t = 0$):

- $N(t = 0) > 0$, and $N(t = 0) < \kappa/2$ (blue curve in Figure 55)
- $N(t = 0) > \kappa/2$, and $N(t = 0) < \kappa$ (orange curve in Figure 55), and
- $N(t = 0) > \kappa$ green curve in Figure 55).

4.2.4 The Verhulst Equation and its Exact Solutions

We now derive the exact solution of the Verhulst equation and compare it to the findings from the previous section.
Rearranging Equation 4.11 in order to separate the variables dt and dN[44] leads to

$$\frac{dN}{N\left(1 - \frac{N}{\kappa}\right)} = r dt \; . \tag{4.15}$$

The rhs of Equation 4.15 can be easily integrated, but the lhs appears to be more complicated. Therefore, we use the trick of substitution to simplify the expression on the lhs of Equation 4.15. We rename $N = 1/z$ and obtain $dz/z^2 = -dN$. Inserting this substitution for Equation 4.15 leads to

$$-\frac{dz}{z - \frac{1}{\kappa}} = r dt \; , \tag{4.16}$$

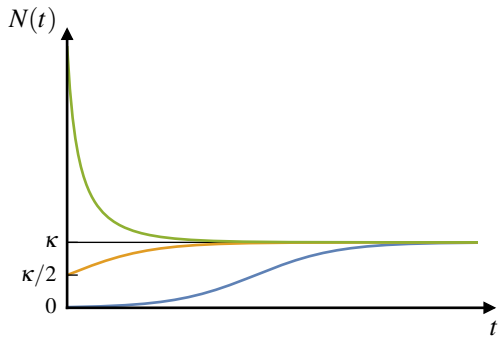

Figure 55: Trajectories of the solution of the Verhulst equation for three different initial population sizes N_0 (also compare to the discussion in Section 4.2.3). $N = \kappa$ is a global fixed point since all trajectories evolve asymptotically towards it.

that has the same structure as Equation 2.82e and we can therefore easily integrate both sides of the equation and find

$$\ln\left(z - \frac{1}{\kappa}\right) = -rt + C. \tag{4.17}$$

Substituting $N = 1/z$ back and denoting the initial condition $N(t = 0) = N_0$ leads to the exact solution of the Verhulst equation:

$$\boxed{N(t) = \frac{\kappa N_0 e^{rt}}{\kappa + N_0\left(e^{rt} - 1\right)}}. \tag{4.18}$$

The structure of Equation 4.18 is called *logistic equation*. The parameter r appears in the exponential and has the unit "per time" which underlines that r is a *growth rate*.
By taking the limit of $t \to \infty$ we can verify that the exact solution also evolves towards $N_2^* = \kappa$, independent of the initial conditions ($N_0 > 0$).

$$\lim_{t\to\infty} N(t) = \lim_{t\to\infty} \frac{\overbrace{\kappa\, N_0\, e^{rt}}^{\to\infty}}{\kappa + \underbrace{N_0\,\left(e^{rt} - 1\right)}_{\to\infty}} = \lim_{t\to\infty} \frac{\kappa\, N_0\, e^{rt}}{N_0\, e^{rt}} = \kappa$$

The trajectories for three different initial population sizes are plotted in Figure 55. All the properties we deduced from the qualitative analysis in Section 4.2.3 can be confirmed, especially the fixed points and the turning point. This illustrates that the relatively simple qualitative analysis using a phase portrait unveils already most of the properties of such a dynamical system. Moreover, in general exact solutions for trajectories of dynamical systems obeying nonlinear differential equations do not necessarily exist. In this regard, the Verhulst equation is an exception and even here finding the exact solution is more cumbersome than the qualitative analysis. Thus, the concept of the qualitative analysis is a powerful tool for understanding dynamical systems, in particular before setting up as simulation or developing the model even further.

4.2.5 Linear Stability Analysis

In Section 4.2.1 we derived rules for the stability of a fixed point by interpreting the phase portrait, followed by a qualitative approach in Section 4.2.4. However, for more complicated

[44] We can treat dN and dt as variables since they originate from the limit of finite differences.

nonlinear and multidimensional models, the stability of a fixed point cannot always be depicted from the phase portrait that simple. Therefore, we will discuss a more rigorous stability analysis here.
The stability of a fixed point can be inferred by the response of the system after a small perturbation $\epsilon(t)$. Adding the perturbation to the fixed point, we can write

$$x(t) = x^* + \epsilon(t) \ , \tag{4.19}$$

where x^* is the fixed point to be analysed. Rewriting Equation 4.19 as $x(t) - x^* = \epsilon(t)$ we see that the perturbation simply describes the distance[45] of the system from the fixed point. If $\epsilon(t)$ increases with time the fixed point is unstable and if $\epsilon(t)$ decreases with time it is stable. Thus, it requires to know the temporal evolution of $\epsilon(t)$ in order to infer the stability, i. e.

$$\frac{d\epsilon(t)}{dt} = \frac{d}{dt}\left[x(t) - x^*\right] = \frac{dx(t)}{dt} = f(x) = f(x^* + \epsilon) \ ,$$

where $\dot{x}^* = 0$ (by definition of the fixed point) was used. Since there is no general solution to such an ODE, but $\epsilon(t)$ is small, we can use the Taylor expansion (Section 2.2.2) of $f(x^* + \epsilon)$ up to the linear order:

$$\frac{d\epsilon(t)}{dt} = f(x^* + \epsilon) \overset{\text{Taylor}}{\approx} f(x^*) + \left.\frac{df(x)}{dx}\right|_{x=x^*} (x - x^*) \tag{4.20}$$

The expression $f(x^*) = 0$ by definition since x^* is a fixed point. Furthermore, $x - x^* = \epsilon$ (Equation 4.19). The first derivative evaluated at the particular value x^* is a constant, so that we define

$$\rho := \left.\frac{df(x)}{dx}\right|_{x=x^*} \tag{4.21}$$

Therefore, Equation 4.20 reads

$$\frac{d\epsilon(t)}{dt} \approx \rho\epsilon \ . \tag{4.22}$$

This approach is called *linearisation*, since we use the Taylor expansion up to the linear term. Equation 4.22 is just the ODE for exponential growth and is therefore solved by

$$\epsilon(t) = \epsilon_0 e^{\rho t} \ . \tag{4.23}$$

Equation 4.23 tells us the temporal behaviour of the small perturbation $\epsilon(t)$. If ρ is **negative**, then the small perturbation **decreases** exponentially with time and therefore x^* is stable. If ρ is **positive**, the small perturbation **increases** exponentially with time and therefore x^* is unstable. By using the definition of ρ (Equation 4.21) we get the same rule as in Section 4.2.1.

$$\begin{aligned} \left.\frac{df(x)}{dx}\right|_{x=x^*} &= \rho > 0 \qquad \text{repellent/unstable} \\ \left.\frac{df(x)}{dx}\right|_{x=x^*} &= \rho < 0 \qquad \text{attractor/stable} \end{aligned} \tag{4.24}$$

But in addition to that the more rigorous analysis also gives us a way to judge how stable or unstable a fixed point is. The larger $|\rho|$ is the more unstable or stable the fixed point is since ρ determines how fast the perturbation from the fixed point increases or decreases. The parameter $1/\rho$ can be seen as a characteristic timescale, describing how fast or slow the system is attracted or repelled from the fixed point.

[45] To be more rigorous the distance would actually be $|\epsilon(t)|$. The sign of ϵ however would drop out in the following derivation anyway. Therefore, the absolute value is omitted to avoid confusion and keep the notation concise.

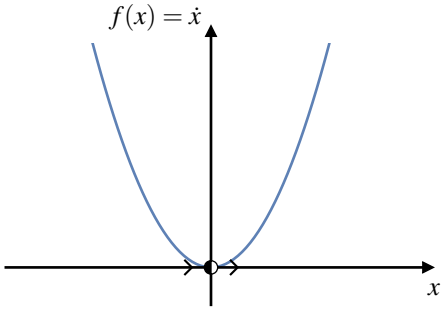

Figure 56: Phase portrait of a half-stable fixed point/saddle point. The fixed point is stable when the system evolves from negaitve x and unstable when evolving from positive x.

4.2.6 Saddlepoints

So far we ignored the case

$$\left.\frac{df(x)}{dx}\right|_{x=x^*} = 0 \tag{4.25}$$

which is called a *half-stable fixed point* or *saddle point*. An example for a saddle point is given in Figure 56. A saddle point is stable if the system approaches it from one side, but unstable when approached from the other side. Thus the flow does not change its direction at the saddle point (in contrast to an "actual" fixed point). The saddle point is an extreme bottleneck for the systems since it can only evolve towards a particular direction. Note, that for a saddle point the linear stability analysis as explained in Section 4.2.5 does not work, since we cannot ignore higher order terms in the Taylor expansion if the first two terms are zero. In order to analyze saddle points one has to consider higher order derivatives of $f(x)$. This shows that the linear stability analysis has its limits, though in general it is a powerful tool.

4.3 Chemical Reactions

4.3.1 From Chemical Reactions to Rate Equations

Consider a chemical reaction $A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} P$ in non equilibrium. The goal of this section is to find the corresponding equation describing the time evolution of such a non equilibrium system. Assuming we are interested in the time evolution of $[P]$, it is useful to look separately at the process through which we gain P and those through which we loose P. The only way P is produced is via the reaction $A + B \overset{k_+}{\rightarrow} P$. We will consider this part of the reaction first. The probability of a particle A to meet a particle B to form the product P is proportional to the product of their concentrations $[A]$ and $[B]$, respectively, and the reaction rate k_+. Thus, although the reaction rate k_+ is constant, the reactants meet more frequently if they are more abundant and less frequently if they are rare. Hence, the concentration of the product $[P]$ changes with time according to (for a more rigorous derivation see Section 5.2.1)

$$A + B \overset{k_+}{\rightarrow} P : \qquad \frac{d[P(t)]}{dt} = k_+ [A] \cdot [B] \; . \tag{4.26}$$

This kind of ODE is often called a *rate equation*. Equation 4.26 is a rate equation describing the gain of P, but there is also loss of P through $P \overset{k_-}{\rightarrow} A + B$. Using the same argument as above yields:

$$P \overset{k_-}{\rightarrow} A + B : \qquad \frac{d[P(t)]}{dt} = -k_- [P] \tag{4.27}$$

The sign in Equation 4.27 is negative, because P is depleted.
Combining the gain and loss term yields:

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} P : \qquad \frac{d[P(t)]}{dt} = k_+ [A] \cdot [B] - k_- [P] \tag{4.28}$$

Equation 4.28 has now the familiar form $\dot{x}_1 = f(x_1, x_2, x_3)$ that was introduced in Section 4.2 (Equation 4.5). This approach can be generalized to a reaction of the type $n\,X + m\,Y \overset{k_+}{\rightarrow} Z$, when n particles of X and m particles of Y have to meet at the same time to form the product Z. Thus, we obtain

$$\frac{d[Z(t)]}{dt} = k_+ [X]^n \cdot [Y]^m \; . \tag{4.29}$$

In the same manner, rate equations can be derived for any case of reaction or set of N reactions that always can be written in the structure $\dot{x}_i = f(x_1, x_2, \ldots, x_i, \ldots, x_N)$.

4.3.2 Enzymatic Reaction

An important type of chemical reactions is the basic enzymatic reaction

$$E + S \underset{k_-}{\overset{k_+}{\rightleftharpoons}} ES \overset{r}{\rightarrow} E + P \; , \tag{4.30}$$

where E denotes the enzyme, S the substrate, ES the enzyme-substrate complex and P the product. We allow a back reaction from the ES-complex to its components, but no back reaction from the enzyme that once has released the product.

We want to derive a set of equations that describes the temporal behaviour of the concentration of each reactant. Let us start with the enzyme itself: The enzyme is "created" via the r reaction and the k_- reaction that yields the two gain terms $r\,[ES] + k_-\,[ES]$. On the other hand the enzyme is consumed via the k_+ reaction, which yields the loss term $-k_+\,[E]\cdot[S]$. Hence, in total we obtain the following rate equation for $[E]$:

$$\frac{d[E]}{dt} = -k_+\,[E]\cdot[S] + r\,[ES] + k_-\,[ES] \tag{4.31}$$

The substrate S is depleted via the k_+ reaction and generated via the k_- reaction. Therefore, we find the rate equation for the substrate analogously:

$$\frac{d[S]}{dt} = -k_+\,[E]\cdot[S] + k_-\,[ES] \tag{4.32}$$

In the same manner we obtain for the temporal behaviour of the concentration of the ES-complex:

$$\frac{d[ES]}{dt} = k_+\,[E]\cdot[S] - r\,[ES] - k_-\,[ES] \tag{4.33}$$

The rate equation of the product is simply:

$$\frac{d[P]}{dt} = r\,[ES] \tag{4.34}$$

These four equations are the *coupled* ODEs, that are necessary to describe the complete time evolution of this enzymatic reaction. The attribute *coupled* in this case means that the time dependent variables ($[E]$, $[S]$, $[ES]$, and $[P]$) do not only appear in the ODE describing their own time evolution, but also in the ODEs for the time evolution of the other variables. This set of ODEs has again the structure $\dot{x}_i = f(x_1, x_2, \ldots, x_i, \ldots, x_N)$, where $N = 4$. Thus, the ODEs influence each other and cannot be solved separately. This particular model of an enzymatic reaction is named after its discoverers, Michaelis-Menten-model[46]. Often, the model is used to investigate the steady state at the equilibrium where the ES-complex concentration does not change, $\frac{d[ES]}{dt} = 0$. Inserting this constraint into Equation 4.33 yields:

$$\frac{[E]\,[S]}{[ES]} = \underbrace{\frac{r + k_-}{k_+}}_{\text{const.}} =: K_m \tag{4.35}$$

where K_m is the well known *Michaelis constant*.

[46] Leonor Michaelis, 1875 - 1949; Maud Leonora Menten, 1879 - 1960

4.4 Two Dimensional Systems

Coupled differential equations, as the one described in Section 4.3.2, become increasingly difficult the more dimensions [47] the system has. A system with only three dimensions can already exhibit chaotic behaviour, in the sense that small changes in initial conditions result in completely different outcomes. This makes the analysis of such systems very complicated, since even the rounding errors of a numeric simulation can change the qualitative behaviour of the system dramatically. The enzymatic reaction in Section 4.3.2 is already four dimensional and hard to visualize, but an investigation of a N-dimensional system concerning fixed points and stability is mathematically not very different from the one dimensional case as discussed in Section 4.2.2. For $N \geq 2$, the approaches are identical and we therefore will restrict ourselves to the analysis of a two dimensional system in the following.
A two dimensional dynamical system can be written as follows.

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y)\,. \end{aligned} \tag{4.36}$$

4.4.1 Linearization and Classification of Fixed Points

In one-dimension there exist only three kinds of fixed points: stable fixed points, unstable fixed points, and saddle points. The classification of a fixed point was derived from the slope in the phase portrait (Section 4.2.5 and Section 4.2.2). In a two dimensional case, we have derivatives in two different directions. Both derivatives at the fixed point can be positive or negative, or one derivative can be positive and the derivative wrt the other variable can be negative. This illustrates that the diversity of fixed points is more complex for higher dimensional systems. How many qualitatively different classes of fixed points do exist in the two-dimensional case? As in the one dimensional case, we can linearize around the fixed point by considering a small perturbation $(\epsilon_x, \epsilon_y)^T$ from the fixed point $(x^*, y^*)^T$

$$\begin{aligned} x(t) &= x^* + \epsilon_x(t) \\ y(t) &= y^* + \epsilon_y(t)\,. \end{aligned}$$

Analogous to the one dimensional case, we can investigate the time evolution of the perturbation that is given by:

$$\begin{array}{rcccccc} \frac{d\epsilon_x(t)}{dt} = & \frac{d}{dt}\left[x(t) - x^*\right] = & \dot{x} = & f(x,y) = & f(x^* + \epsilon_x, y^* + \epsilon_y) \\ \frac{d\epsilon_y(t)}{dt} = & \frac{d}{dt}\left[y(t) - y^*\right] = & \dot{y} = & g(x,y) = & g(x^* + \epsilon_x, y^* + \epsilon_y)\,. \end{array}$$

Using the 2D-Taylor expansion the last term can be approximated up to linear order:

$$f(x^* + \epsilon_x, y^* + \epsilon_y) \overset{\text{Taylor}}{\approx} f(x^*, y^*) + \underbrace{\left.\frac{\partial f}{\partial x}\right|_{x^*,y^*}}_{=:a} \cdot \epsilon_x + \underbrace{\left.\frac{\partial f}{\partial y}\right|_{x^*,y^*}}_{=:b} \cdot \epsilon_y$$

$$g(x^* + \epsilon_x, y^* + \epsilon_y) \overset{\text{Taylor}}{\approx} g(x^*, y^*) + \underbrace{\left.\frac{\partial g}{\partial x}\right|_{x^*,y^*}}_{=:c} \cdot \epsilon_x + \underbrace{\left.\frac{\partial g}{\partial y}\right|_{x^*,y^*}}_{=:d} \cdot \epsilon_y$$

[47] In this context the dimension of a system is the number of time dependent variables describing the state of the system.

As in the one-dimensional case we can use that $f(x^*, y^*)$ and $g(x^*, y^*)$ are both zero. The first derivatives evaluated at the fixed point give constant numbers, which are defined as a, b, c, d (analogous to the definition of ρ in Section 4.2.5). Thus, we obtain the following equation which exhibits the same structure as Equation 4.22 in Section 4.2.5:

$$\begin{aligned} \frac{d\epsilon_x(t)}{dt} &= a\,\epsilon_x + b\,\epsilon_y \\ \frac{d\epsilon_y(t)}{dt} &= c\,\epsilon_x + d\,\epsilon_y\,. \end{aligned} \tag{4.37}$$

Equation 4.37 can be rewritten in a more compact form using matrix notation (Equation 2.11), that leads to

$$\dot{\vec{\epsilon}} = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{=:A} \vec{\epsilon} \qquad \text{with } \vec{\epsilon} = \begin{pmatrix} \epsilon_x \\ \epsilon_y \end{pmatrix} \tag{4.38}$$

The matrix A that contains the derivatives is called *Jacobian*[48] *matrix*. This notation is not only beneficial because it is very concise, it also allows us to use the whole mathematical machinery of linear algebra, especially eigenvalues and eigenvectors. But what is the meaning of eigenvalues and eigenvectors in this context. From Section 2.1.2 we know that for an eigenvector $\vec{v}$ with eigenvalue λ the following holds:

$$\dot{\vec{v}} = A \cdot \vec{v} = \lambda\,\vec{v} \tag{4.39}$$

that has the same structure as Equation 4.38 and we can just exchange ϵ with v and write

$$\dot{\vec{\epsilon}} = A \cdot \vec{\epsilon} = \lambda\,\vec{\epsilon} \tag{4.40}$$

The vector $\vec{\epsilon} = \begin{pmatrix} \epsilon_x \\ \epsilon_y \end{pmatrix}$ points to the direction of the small perturbation and the temporal evolution of $\vec{\epsilon}$, $\dot{\vec{\epsilon}}$, is just $\lambda\,\vec{\epsilon}$, hence the vectors $\vec{\epsilon}$ and $\dot{\vec{\epsilon}}$ are parallel to each other. Since any vector $\vec{w} = \alpha\,\vec{v}$ is also an eigenvector this results in the fact that a system located in one of the eigendirections will always stay on that eigendirection. This is visualized in Figure 57.

Another way to realize this property is by noting that Equation 4.39 is simply the equation for exponential growth since the derivative of the exponential ($\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}$) is again an eigenvalue problem: it yields the function itself times a constant. Hence,

$$\vec{\epsilon}(t) = e^{\lambda\,t}\,\vec{\epsilon} \tag{4.41}$$

that has the same structure as Equation 4.22. These similarities illustrate even further that a system lying on one eigendirection will always stay on that particular eigendirection while accelerating or decelerating depending on the sign of the eigenvalue λ. If λ is positive the system is repelled from the fixed point on the corresponding eigendirection, if λ is negative the system is attracted towards the fixed point on that eigendirection. Hence, the eigenvalues determine the stability of the system along the eigendirections.
In order to derive the eigenvalues we need to solve the eigenvalue problem (c. f. Section 2.1.2). The characteristic equation is given by:

$$0 \stackrel{!}{=} \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)\,(d-\lambda) - c\,b$$

[48] Carl Gustav Jacob Jacobi (actually Jacques Simon), 1804 - 1851

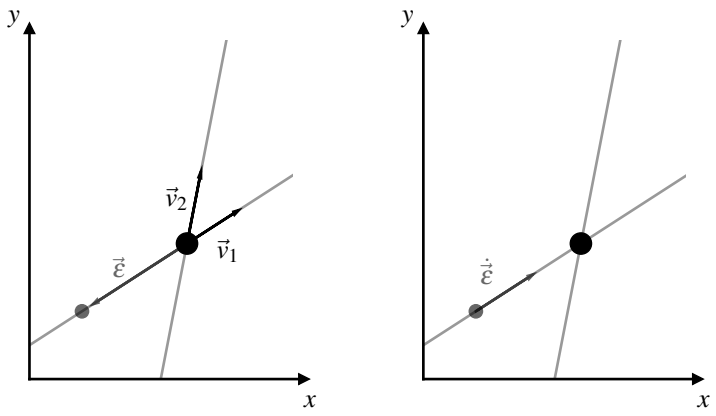

Figure 57: The eigendirections of a fixed point (large black dot). The left plot shows the eigendirections as grey lines, which are determined by the the direction of the eigenvectors $\vec{v}_1$ and $\vec{v}_2$. The system (small black dot) is perturbed from the fixed point by the vector $\vec{\epsilon}$. The right plot shows the flow $\dot{\vec{\epsilon}}$ of the perturbed system at that point. The change is parallel to the eigendirection, and since the system already is on the eigendirection, this will result in the system staying on that eigendirection. In this example the fixed point is stable in $\vec{v}_1$ direction since the system is pulled back towards the fixed point along this eigendirection.

$$\begin{aligned} 0 &\stackrel{!}{=} \lambda^2 - \underbrace{(a+d)}_{=:\tau}\,\lambda + \underbrace{(a\,d - c\,b)}_{=:\Delta} \\ 0 &\stackrel{!}{=} \lambda^2 - \tau\,\lambda + \Delta \end{aligned} \tag{4.42}$$

In the last line we use the definition of the *trace* τ, as the sum of the diagonal elements, and the definition of the two-dimensional *determinant* Δ (Equation 2.25). The eigenvalues are obtained by solving the quadratic equation

$$\lambda_1 = \frac{\tau + \sqrt{\tau^2 - 4\,\Delta}}{2} \qquad \lambda_2 = \frac{\tau - \sqrt{\tau^2 - 4\,\Delta}}{2}\,. \tag{4.43}$$

Note, that it is sufficient here to take the eigenvalue problem and its solution as a kind of recipe. The deeper meaning and the geometrical interpretation of eigenvalues and eigenvectors are throughoutly explained in terms of principle component analysis in the statistics primer.

It can be inferred analogous to Equation 4.22 by Equation 4.41 that the sign of λ determines whether the fixed point is stable or unstable in the respective eigendirection. In contrast to the one dimensional cases, the eigenvalues can also have an imaginary part, precisely whenever $\tau^2 - 4\,\Delta < 0$. Note that both eigenvalues either simultaneously have an imaginary part or not, since the $\tau^2 - 4\,\Delta$ term is identical in both equations. Therefore, the diversity of fixed points including attractor, repeller and saddle points are extended by those containing an imaginary part. In summary the *trace* τ and the *determinant* Δ are sufficient to classify most fixed points. The different kinds of fixed points are shown in Figure 58, and an overview for the classification scheme based on the values of τ and Δ is shown in Figure 59.

The different fixed points are classified as follows:

saddle-nodes $\Delta < 0$

One of the eigenvalues is negative and one is positive, without an imaginary part. Therefore one eigendirection is stable and the other one is unstable.

non-isolated fixed points $\Delta = 0$

At least one of the eigenvalues equals zero. Thus there is no attraction/repulsion in

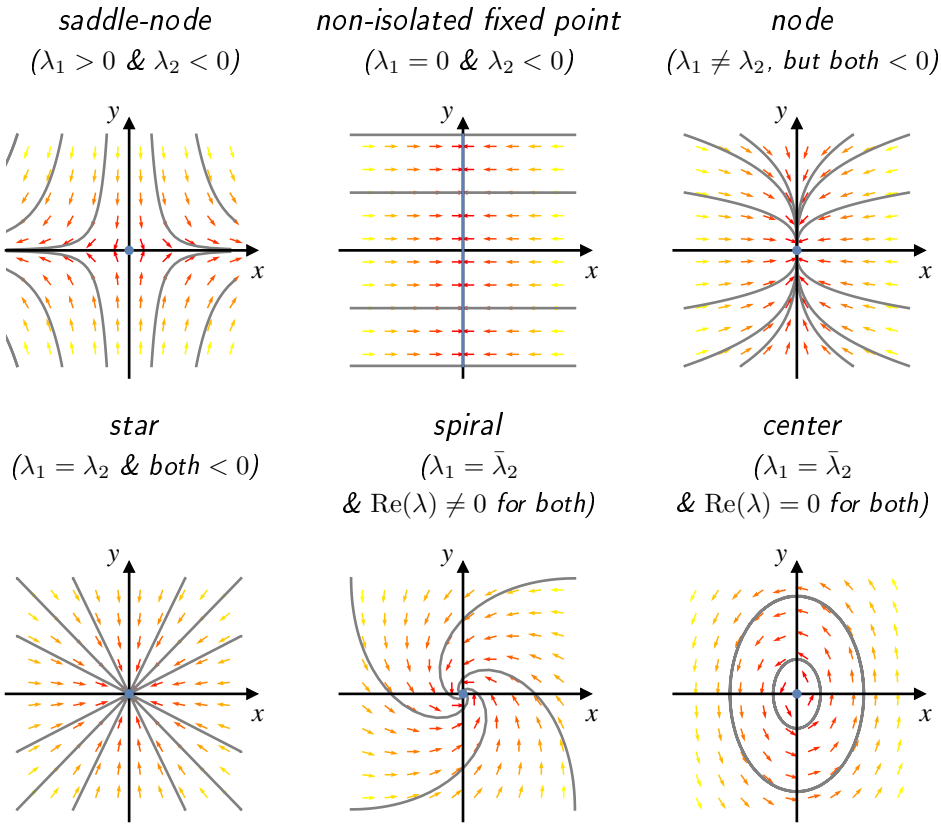

Figure 58: Different possible fixed points of a two-dimensional dynamical systems. The green/gray lines show representative trajectories for the vectorfield (small coloured arrows) determined by given pairs x and y. The non-isolated fixed point, the node, the star, and the spiral (see Section 4.4.4 for a biological example) are all stable. The unstable counterparts of these fixed points (switching the sign of the real part of the eigenvalues) would give rise to the same plots with the vectors showing in the opposite direction.

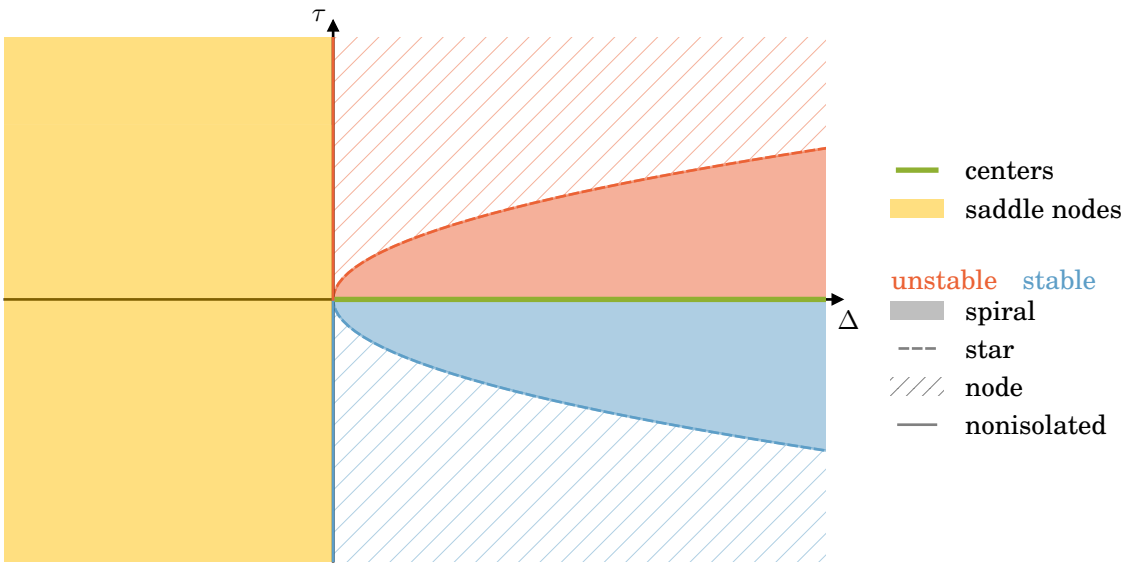

Figure 59: The properties of the fixed points are characterized by the trace τ and the determinant Δ of the linearized system (Equation 4.38, c. f. [3, chapter 5]). The red area denotes the unstable regime and the blue area the stable regime. The non-isolated fixed points coincide with the τ-axis since $\Delta = 0$, and the centers coincide with the Δ-axis. The parabola is given by threshold $\tau^2 = 4\Delta$ (Equation 4.43).

this direction, which results in a fixed "point" actually being a straight line (in the eigendirection with $\lambda = 0$). Therefore, the system does not change once it reaches the "fixed-line". If both eigenvalues are zero the whole plane is fixed, in the sense that the system will stay at whatever initial position it started from.

node $\Delta > 0$ and $\tau^2 > 4\,\Delta$
Both eigenvalues have the same sign (both attracting or both repelling) and their imaginary part equals zero. If $\tau < 0$ the node is stable, and if $\tau > 0$ it is unstable.

stars $\Delta > 0$ and $\tau^2 = 4\,\Delta$
The two eigenvalues are identical because the square-root becomes zero. Therefore, both eigendirections repell/attract equally strong. If $\tau < 0$ the star is stable, and if $\tau > 0$ it is unstable.

spirals $\Delta > 0$ and $\tau^2 < 4\,\Delta$
Both eigenvalues are complex due to the negative sign of the term under the square root. The real part is the same for both, whereas the imaginary part has opposite signs (they are complex conjugates). This results in a circular motion around the fixed point in addition to repulsion/attraction in both eigendirections. Thus, the resulting trajectories are spirals. If $\tau < 0$ the spiral is stable, and if $\tau > 0$ it is unstable.

center $\Delta > 0$ and $\tau = 0$ Both eigenvalues become purely imaginary, with opposite signs. The resulting trajectories are concentric ellipses around the fixed point[49].

As one can see in the list above, except for saddle nodes the sign of the trace τ defines the stability of the fixed point. Note, that one can show that a system is stable if **all** the eigenvalues of the Jacobian matrix have negative real parts.

4.4.2 Visualization

The visualization of the different fixed points in Figure 58 implicitly requires a mathematical basis. The trajectories are determined by the derivatives at the location (x, y). For the one-dimensional system we plotted the position x and the corresponding flow given by $\dot{x} = f(x)$ yielding a two dimensional graph. In a two-dimensional system as described in the previous subsection, we need to know the location by the (x, y)-position, the flow in the x-direction given by $f(x, y)$, and the flow in the y-direction given by $g(x, y)$. In total this means four variables requiring a four-dimensional plot for an analogous visualization.
In order to circumnavigate this problem, it is reasonable to start by using the x-y-plane, called *phase plane*, as the coordinate system for our plot, since this is the physical space occupied by the system. In addition, we have to include the information given by $f(x, y)$ and $g(x, y)$ (the flow). Usually, three different kinds of plots are used for visualizing $f(x, y)$ and $g(x, y)$, which can be seen in Figure 60 (the plots in Figure 58 are one of these kinds).

Vector plot: The plots in Figure 58 show $f(x, y)$ and $g(x, y)$ as a vectorfield where each vector $\vec{v}$ is defined by:

$$\vec{v} := \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$$

[49] This is similar to the non-isolated fixed points in the sense, that now there are non-isolated *trajectories* rather than non-isolated fixed points.

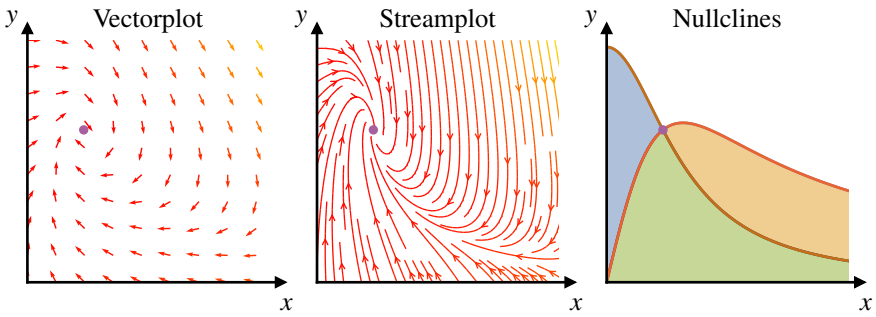

Figure 60: Three different plots for the same dynamical system (the fixed point is indicated by the lilac dot.)

Usually, the length of the vector indicates directly the magnitude of the vectorfield at a given location. However, it might be more appropriate to indicate the magnitude $|\vec{v}|$ by a colour code and to normalize all vectors to the same length. Hence, the direction is given by the normalized vector $\vec{w}$ such that

$$\begin{array}{lcll} \frac{1}{|\vec{v}|}\,\vec{v} & = & \vec{w} & \Rightarrow \text{ direction} \\ |\vec{v}| & \xrightarrow{colormap} & \text{color} & \Rightarrow \text{ speed} \end{array}$$

and the color indicades the speed of the flow.

Stream plot: A stream plot shows various representative trajectories with their direction. The colour of the stream lines again indicates the speed (magnitude) of the flow. Since the solution of a set of ODEs is unique and therefore the trajectories cannot split at any given point (there is only one possible direction of flow), the *different trajectories can never intersect* (c. f Equation 4.36). The stream lines can, however, get infinitely close to each other which may lead to the impression that they merge or intersect.

Nullcline plot: In particular in Section 4.4.4 we will use the visualization via *nullclines*. Nullclines are defined by the lines where either $\dot{x} = 0$ or $\dot{y} = 0$, that separates the phase plane into several regions (called *phases*) in which the system exhibits a different behaviour. The point where the two nullclines intersect is a fixed point since here $\dot{x} = 0 = \dot{y}$.

Note that the first two plots (vector plot and stream plot) can only be realized by using a computer. Plotting nullclines is best suited for getting a first (and quantitative) overview of the behaviour of the system (see Section 4.4.4).

4.4.3 Exact Solutions of Linear ODEs

In Section 4.2, we discussed the properties of one dimensional ODEs based on the phase portrait and the exact solution of the Verhulst equation of bound growth. In the same manner, we derive now the exact solution of a linear ODE in two dimensions. This subsection is written for dedicated students that are interested in the detailed mathematical background and it will not be subject of the lecture. Section 4.4.4 can be understood without this subsection.
A two dimensional set of **linear** ODEs can be written as

$$\dot{\vec{\xi}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \vec{\xi} \qquad \text{with } a, b, c, d \in \mathbb{R}\,. \tag{4.44}$$

To get an idea of the mathematical structure, I like to discuss three different sub cases in the following:

Diagonal Matrix A diagonal matrix like

$$\dot{\vec{\xi}} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \vec{\xi} \tag{4.45}$$

represents two uncoupled ODEs, because it equals

$$\dot{x} = a\,x$$
$$\dot{y} = d\,y$$

when written in its components. The equations above can be solved easily by the exponential equation

$$x(t) = c_1\, e^{a\,t}$$
$$y(t) = c_2\, e^{d\,t}$$

and using the initial conditions $x(0) = x_0$ and $y(0) = y_0$ we obtain

$$x(0) = c_1\, e^{a\,0} = c_1 = x_0$$
$$y(0) = c_2\, e^{d\,0} = c_2 = y_0$$

the two constants c_1 and c_2. The overall result written in vector notation is.

$$\vec{\xi}(t) = \begin{pmatrix} x_0\, e^{a\,t} \\ y_0\, e^{d\,t} \end{pmatrix} \ . \tag{4.46}$$

This toy example can be used to create stars ($a = d$), nodes (a and b having the same sign), saddle nodes (a and b having opposite sign), and non-isolated fixed points (either a or b equally zero). Such equations were actually used, to plot the trajectories for the respective graphs in Figure 58.
Apart from deriving the exact solution for this particular matrix, the case of a diagonal matrix is also of conceptual interest, since the eigenvalues and eigenvectors are already given by

$$\lambda_1 = a \qquad \lambda_2 = d \qquad \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

Any point in a vector space can be described by a linear combination of eigenvalues and eigenvectors. For example the point $\vec{P} = (2; 1; -4)$ is actually a linear combination of

$$\begin{pmatrix} 2 \\ 1 \\ -4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 4 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -4 \end{pmatrix} \ .$$

The matrix above is diagonal, because the coordinate axis which span the vector space are the eigenvectors. It might be for some reason that the coordinate system is not well chosen (for example it has been turned or is non-orthogonal) so that the matrix is not diagonal. Then, the matrix can be diagonalized by solving the characteristic equation.
If the λ are the eigenvalues and $\vec{v}$ are the eigenvectors (c. f. Equation 4.39), we can use the linear combination as ansatz for the solution of ξ as

$$\vec{\xi}(t) = c_1\, e^{\lambda_1\, t}\, \vec{v}_1 + c_2\, e^{\lambda_2\, t}\, \vec{v}_2 \quad . \tag{4.47}$$

The constants c_1 and c_2 can be determined via the initial conditions (starting point). How this is done will become clear in the next case.

Circles and Ellipses The matrix describing the linear ODEs can also exhibit the form

$$\dot{\vec{\xi}} = \begin{pmatrix} 0 & \mp\beta \\ \pm\gamma & 0 \end{pmatrix} \vec{\xi} \qquad \text{with } \beta, \gamma > 0 \tag{4.48}$$

and we can solve for the eigenvalues and eigenvectors. We obtain

$$\lambda_1 = \pm i\sqrt{\beta\,\gamma} \qquad \lambda_1 = \mp i\sqrt{\beta\,\gamma} \qquad \vec{v}_1 = \begin{pmatrix} i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} \qquad \vec{v}_2 = \begin{pmatrix} -i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix}$$

and therefore can use the general ansatz (Equation 4.47) to get:

$$\vec{\xi}(t) = c_1\, e^{\pm i\sqrt{\beta\gamma}\,t} \begin{pmatrix} i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} + c_2\, e^{\mp i\sqrt{\beta\gamma}\,t} \begin{pmatrix} -i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} \tag{4.49}$$

$$= \begin{pmatrix} \pm i\sqrt{\beta/\gamma}\left(c_1\, e^{i\sqrt{\beta\gamma}\,t} - c_2\, e^{-i\sqrt{\beta\gamma}\,t}\right) \\ c_1\, e^{i\sqrt{\beta\gamma}\,t} + c_2\, e^{-i\sqrt{\beta\gamma}\,t} \end{pmatrix} \tag{4.50}$$

We determine the constants from the initial condition ($t = 0$) and obtain the starting point

$$\vec{\xi}(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \pm i\sqrt{\beta/\gamma}(c_1 - c_2) \\ c_1 + c_2 \end{pmatrix}.$$

This leads after some basic algebra to the constants c:

$$c_1 = \frac{1}{2}\left(y_0 \mp i\sqrt{\gamma/\beta}\,x_0\right) \qquad c_2 = \frac{1}{2}\left(y_0 \pm i\sqrt{\gamma/\beta}\,x_0\right) \tag{4.51}$$

Inserting the constants into Equation 4.50 yields:

$$\begin{aligned} \vec{\xi}(t) &= \begin{pmatrix} \pm i\sqrt{\beta/\gamma}\,\frac{y_0}{2}\left(e^{i\sqrt{\beta\gamma}\,t} - e^{-i\sqrt{\beta\gamma}\,t}\right) & - & i\sqrt{\beta/\gamma}\,i\sqrt{\gamma/\beta}\,\frac{x_0}{2}\left(e^{i\sqrt{\beta\gamma}\,t} + e^{-i\sqrt{\beta\gamma}\,t}\right) \\ \frac{y_0}{2}\left(e^{i\sqrt{\beta\gamma}\,t} + e^{-i\sqrt{\beta\gamma}\,t}\right) & \mp & i\sqrt{\gamma/\beta}\,\frac{x_0}{2}\left(e^{i\sqrt{\beta\gamma}\,t} - e^{-i\sqrt{\beta\gamma}\,t}\right) \end{pmatrix} \\ &= \begin{pmatrix} \mp y_0\sqrt{\beta/\gamma}\,\sin(\sqrt{\beta\,\gamma}\,t) & + & x_0\cos(\sqrt{\beta\,\gamma}\,t) \\ y_0\cos(\sqrt{\beta\,\gamma}\,t) & \pm & x_0\sqrt{\gamma/\beta}\,\sin(\sqrt{\beta\,\gamma}\,t) \end{pmatrix}. \end{aligned}$$

For $\beta \neq \gamma$ his trajectory represents an ellipse with its principle components corresponding to the x and y axes (since we started with a linear combination of the eigenvectors and eigenvalues in the ansatz). The elliptical shape of the trajectory is not that obvious, but by choosing either $x_0 = 0$ or y_0, we find that

$$\vec{\xi}(t) = \begin{pmatrix} \mp y_0\sqrt{\beta/\gamma}\,\sin(\sqrt{\beta\,\gamma}\,t) \\ y_0\cos(\sqrt{\beta\,\gamma}\,t) \end{pmatrix} \quad \text{or} \quad \vec{\xi}(t) = \begin{pmatrix} x_0\cos(\sqrt{\beta\,\gamma}\,t) \\ \pm\, x_0\sqrt{\gamma/\beta}\,\sin(\sqrt{\beta\,\gamma}\,t) \end{pmatrix}. \tag{4.52}$$

Moreover, if $\beta = \gamma$, the trajectory turns into a circle, since all pre-factors cancel out. For both cases the system moves clockwise around the fixed point[50]. When switching the signs of β and γ (red signs) the solution is almost identical except that now the direction of rotation is counter-clockwise.

[50] This can be seen by setting $x_0 = 0$ in Equation 4.52. At $t = 0$, the x-coordinate equals zero and gets negative if t increases. The y-coordinate has its maximum value at $t = 0$ and decreases with increasing t, that equals a clockwise rotation when being plotted in the xy-plane.

Spirals A spiral trajectory is obtained if the ODEs have the form

$$\dot{\vec{\xi}} = \begin{pmatrix} \alpha & \mp\beta \\ \pm\gamma & \alpha \end{pmatrix} \vec{\xi} \qquad \text{with } \beta, \gamma > 0 \text{ and } \alpha \in \mathbb{R}\,. \tag{4.53}$$

The calculation of the eigenvalues and eigenvectors leads to

$$\lambda_1 = \alpha \pm i\sqrt{\beta\,\gamma} \qquad \lambda_1 = \alpha \mp i\sqrt{\beta\,\gamma} \qquad \vec{v}_1 = \begin{pmatrix} i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} \qquad \vec{v}_2 = \begin{pmatrix} -i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix}$$

and again using the general ansatz like in Equation 4.47, we find

$$\vec{\xi}(t) = c_1\, e^{(\alpha \pm i\sqrt{\beta\gamma})\,t} \begin{pmatrix} i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} + c_2\, e^{(\alpha \mp i\sqrt{\beta\gamma})\,t} \begin{pmatrix} -i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} \tag{4.54}$$

$$= e^{\alpha\,t} \underbrace{\left[c_1\, e^{\pm i\sqrt{\beta\,\gamma}\,t} \begin{pmatrix} i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} + c_2\, e^{\mp i\sqrt{\beta\,\gamma}\,t} \begin{pmatrix} -i\sqrt{\beta/\gamma} \\ 1 \end{pmatrix} \right]}_{\text{=ellipse equation (Equation 4.49)}}. \tag{4.55}$$

Equation 4.55 is almost identical to the equation we derived for ellipse/circle, except for the prefactor $e^{\alpha\,t}$. Since this prefactor equals one for $t = 0$, the solution for the coefficients c_1 and c_2 is exactly the same as in the previous case (Equation 4.51). Thus, we can simply use the previously derived result and multiply it with the new prefactor to obtain the final solution:

$$\vec{\xi}(t) = e^{\alpha\,t} \begin{pmatrix} \mp y_0\sqrt{\beta/\gamma}\,\sin(\sqrt{\beta\,\gamma}\,t) & + & x_0\,\cos(\sqrt{\beta\,\gamma}\,t) \\ y_0\,\cos(\sqrt{\beta\,\gamma}\,t) & \pm & x_0\sqrt{\gamma/\beta}\,\sin(\sqrt{\beta\,\gamma}\,t) \end{pmatrix}.$$

This trajectory has the same properties as the corresponding circle/ellipse, with the key difference that the distance from the fixed point changes according to the prefactor $e^{\alpha\,t}$. If $\alpha < 0$ the distance decreases exponentially and therefore the fixed point is stable. For $\alpha > 0$ the distance increases exponentially and the fixed point is unstable.

4.4.4 Glycolysis and Limit Cycles

One of the most suitable examples for discussing a set of ODEs using a phase portrait is glycolysis. Glycolysis is one of the most important reactions in the body. The glycolysis reaction chain actually contains dozens of reactions that can be fortunately simplified to two reaction equations (Sel'kov[51], 1968) that are

$$\frac{d[X]}{dt} = -[X] + a\,[Y] + [X]^2\,[Y] \tag{4.56}$$

$$\frac{d[Y]}{dt} = b - a\,[Y] - [X]^2\,[Y]\,. \tag{4.57}$$

where X represents ADP and Y represents Fructose 6-phosphate (F6P), a and b are positive constants. This system is a non linear (since the term $[X]^2$ appears) set of coupled ODEs and has exactly the structure of Equation 4.36. Note, that mass (total concentration) is not conserved in this system, since $\frac{d[X]}{dt} + \frac{d[Y]}{dt} = b - [X] \neq 0$.
In order to get an idea of for the dynamics of the system it is useful to consider the nullclines. As described in Section 4.4.2, nullclines are the lines where either $[\dot{X}]$ or $[\dot{Y}]$ is zero. Setting

[51] Evgeni (Zhenya) Sel'kov, 1937 - today

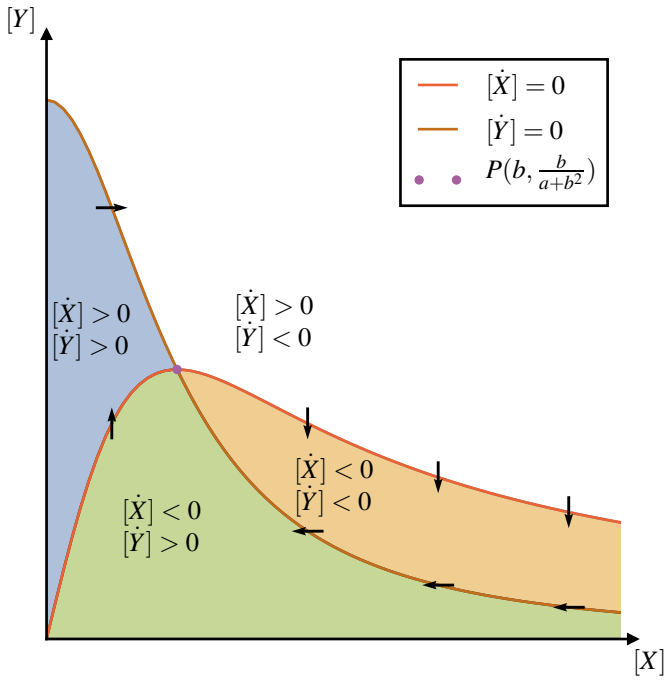

Figure 61: Phase diagram of nullclines of Sel'kovs model for glycolysis. The nullclines divide the phase diagram into four different (colour coded) regions. Depending on the nullcline either $[\dot{X}]$ or $[\dot{Y}]$ is zero which results in a purely vertical or horizontal flow, respectively, on the nullclines (arrows).

$[\dot{X}] = 0$ in Equation 4.56 and rearranging the equation in order to express it in terms of Y we obtain a nullcline I like to denote as $[Y]_1$ and setting $[\dot{Y}] = 0$ in Equation 4.57 and solving for $[Y]$ we derive a second nullcline I like to denote as $[Y]_2$. These equations yield the nullclines in the $x - y$ plane:

$$[Y]_1 = \frac{[X]}{a + [X]^2} \qquad\qquad [Y]_2 = \frac{b}{a + [X]^2}$$

The phase space with the nullclines is shown in Figure 61.
The nullclines intersect at $[Y]_1 = [Y]_2$ and divide the phase space into four different phases. The point of intersection $P = ([X]^*, [Y]^*)$ is a fixed point since both, $[\dot{X}]$ and $[\dot{Y}]$, are zero. Therefore, we obtain the solution for P:

$$P = \left(b, \frac{b}{a + b^2}\right) . \tag{4.58}$$

On the nullcline $[Y]_1$, the flow points only in vertical direction since (by definition) $[\dot{X}] = 0$, whereas the flow vector points into a horizontal direction ($[\dot{Y}] = 0$) on the nullcline $[Y]_2$. But how does the flow behave in the different sections of the phase space?
If we set $[Y]$ in Equation 4.56 to $[\bar{Y}] = [Y] + \epsilon$, where ϵ denotes a small shift, we obtain the equation

$$\frac{d[\bar{X}]}{dt} = [\dot{X}] + \epsilon \left([X]^2 + a\right) . \tag{4.59}$$

Since $[\dot{X}] = 0$ equals the nullcline $[Y]_1$, the derivative of $[X]$ above/below the nullcline equals $\epsilon\left([X]^2 + a\right)$. The constant a and the concentration $[X]$ is always positive, so that above ($\epsilon > 0$) the $[Y]_1$-nullcline $[\dot{X}] > 0$ and below ($\epsilon < 0$) the nullcline $[\dot{X}] < 0$. Similarly, we obtain, that above the $[Y]_2$-nullcline $[\dot{Y}] < 0$ and below the nullcline $[\dot{Y}] > 0$. The sign

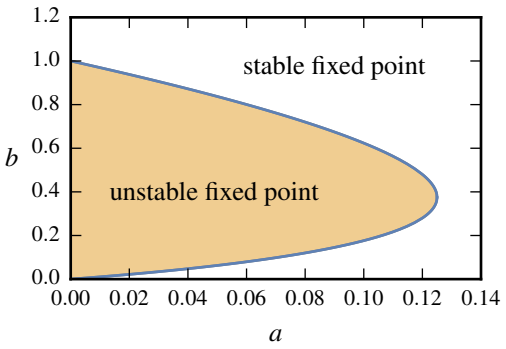

Figure 62: The two states (stable/unstable) for glycolysis depending on the parameter a and b.

of $[\dot{X}]$ and $[\dot{Y}]$ can only change at the nullclines because in order to change the sign the field has to pass through zero somewhere in between. The different signs of the derivatives of $[X]$ and $[Y]$ are shown in Figure 61.
Let us now return to the fixed point and derive whether P is an attractor or a repeller. We use the techniques derived in section Section 4.4.1. The first step is to find the Jacobian Matrix A of the dynamical system evaluated at the fixed point P (c. f. Equation 4.38).

$$A = \begin{pmatrix} \frac{d[\dot{X}]}{d[X]}\Big|_P & \frac{d[\dot{X}]}{d[Y]}\Big|_P \\ \frac{d[\dot{Y}]}{d[X]}\Big|_P & \frac{d[\dot{Y}]}{d[Y]}\Big|_P \end{pmatrix} = \begin{pmatrix} -1 + 2[X][Y]|_P & a + [X]^2|_P \\ -2[X][Y]|_P & -a - [X]^2|_P \end{pmatrix}$$

$$= \begin{pmatrix} -1 + 2\,\frac{b^2}{a+b^2} & a + b^2 \\ -2\,\frac{b^2}{a+b^2} & -a - b^2 \end{pmatrix} . \tag{4.60}$$

Therefore the determinant Δ and the trace τ are

$$\Delta = a + b^2 \qquad \text{and} \qquad \tau = -\frac{b^4 + (2a-1)\,b^2 + a + a^2}{a + b^2} . \tag{4.61}$$

From Section 4.4.1 we know that if $\tau > 0$ the fixed point is unstable and if $\tau < 0$ the fixed point is stable. In order to judge for which values of the parameters a and b the fixed point is either stable or unstable we calculate the dividing line, $\tau = 0$, between these two states. Setting $\tau = 0$ in Equation 4.61 and solving for b yields the condition

$$b = \sqrt{\frac{1}{2}\left(1 - 2a \pm \sqrt{1 - 8a}\right)} .$$

The resulting line dividing the system into the two states is shown in Figure 62 together with the stable and unstable area in the parameter space (a, b). For a throughout classification of the fixed point we need to check whether it is a spiral/center, $\tau^2 \geq 4\,\Delta$, or a normal node, $\tau^2 < 4$ (c. f. Section 4.4.1 and the figures therein). In this case, however, it is very laborious to find the two states analytically and therefore the analysis is disregarded in this section.
A single fixed point in a biological system might be problematic, since such a system would diverge (either $[X]$ or $[Y]$ getting infinitely large for $t \to \infty$) if the fixed point is unstable. If the fixed point would be unstable, a small perturbation would lead to a flow pointing away from it. If the system would reach the phase indicated with white colour in Figure 63, $[Y]$

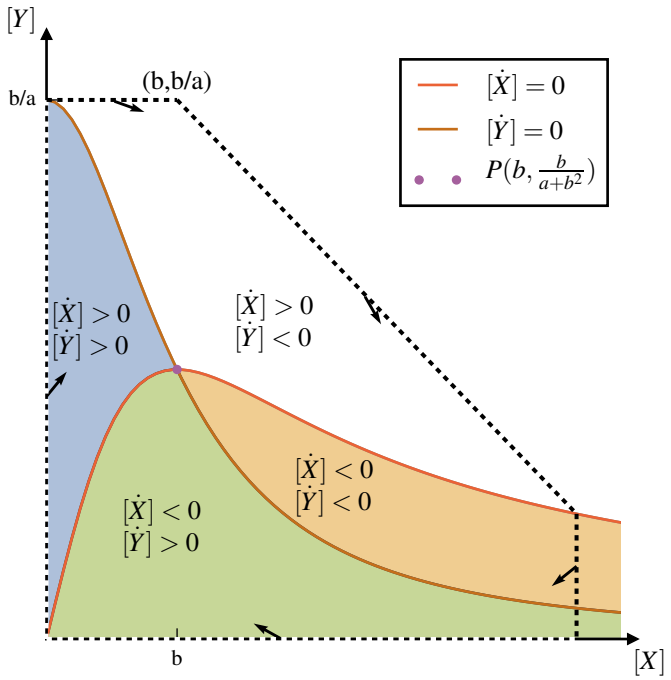

Figure 63: Trapping region for glycolysis according to Sel'kovs model. The diagonal has the slope -1, and starts at $(b, b/a)$ from the top-left and continues until it intersects with $[Y]_1$ at the bottom-right. The black arrows represent the flow at the respective boundary line.

would decrease and $[X]$ increase, so that the flow would point to the lower right (indicated by the black arrow in Figure 63). Depending on the angle, the system would sooner or later cross the nullcline $[Y]_1$ that would result in a purely vertical motion ($[\dot{X}] = 0$). When crossing $[Y]_1$, the system enters the orange phase in Figure 63 resulting into a motion to the lower right, until crossing the nullcline $[Y]_2$ and entering the green shaded phase. According to $[\dot{X}]$ and $[\dot{Y}]$ in the green phase, the system moves to the upper left until crossing $[Y]_1$ and entering the blue shaded area in Figure 63. In total, this results in a clockwise circular trajectory around the unstable fixed point.
Even if the system is far away from the fixed point, hence above the diagonal $[Y] = -[X]$, it would perform such a circular motion close to the fixed point. This can be seen when dividing Equation 4.57 by Equation 4.56, so that we obtain the expression $\frac{d[Y]}{d[X]}$, and applying $[Y] \to \infty$ and $[X] \to \infty$ (far away from the fixed point). The result is $\frac{d[Y]}{d[X]} = -1$, i. e. the diagonal itself. Thus, the system would move along the diagonal $[Y] = -[X]$ until it reaches the nullcline $[Y]_1$ and then performing a clockwise spiral motion as described above.
For the flow to point inwards if the system is *on* the diagonal, the condition $[\dot{X}] > -[\dot{Y}]$ has to be satisfied. This leads to

$$-[X] + a\,[Y] + [X]^2\,[Y] < -(b - a\,[Y] - [X]^2\,[Y])$$
$$[X] > b$$

when inserted into the glycolysis Equation 4.56 and Equation 4.57. Since by construction each point on the diagonal has larger $[X]$-values than b this condition is satisfied.
Hence, independent from the initial conditions, the system would circulate around a particular region if the fixed point is unstable. Since the system cannot escape this region, it is called *trapping region* (dashed line in Figure 63). The system cannot escape the trapping

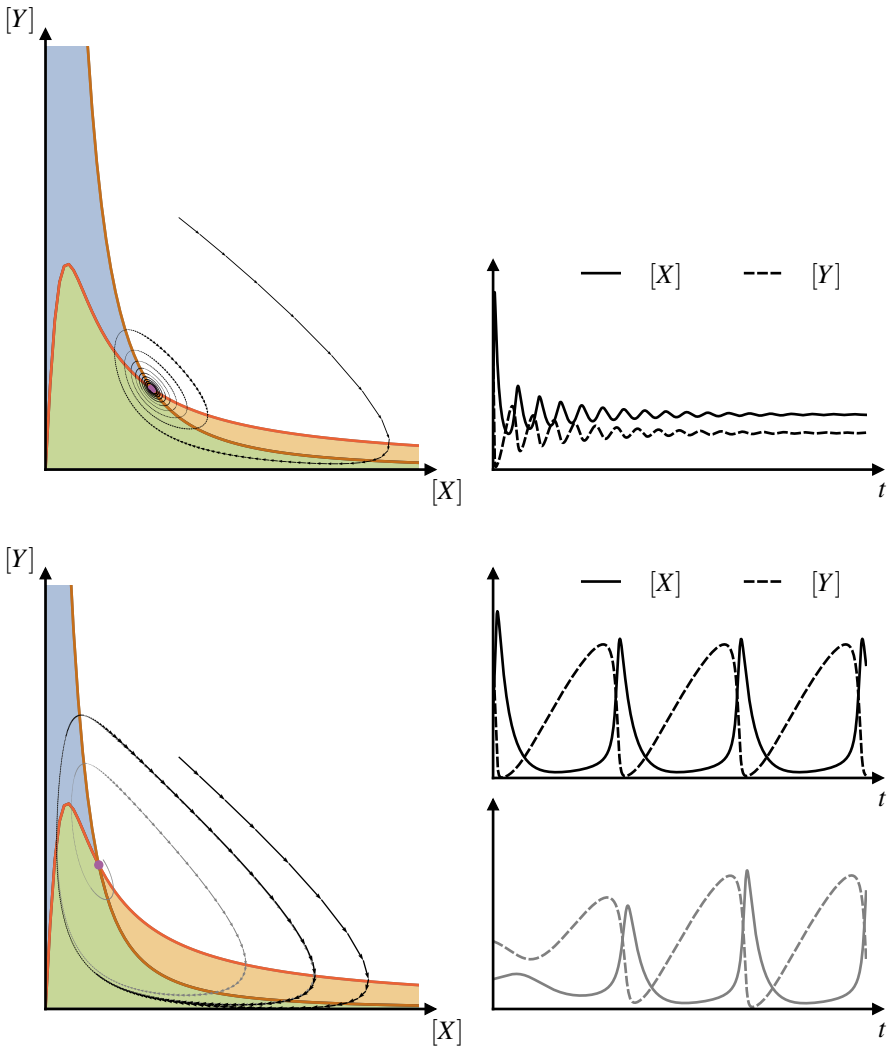

Figure 64: Trajectories of Sel'kovs glycolysis model for the two cases of a stable fixed point (top) and an unstable fixed point with resulting limit cycle (bottom). The temporal evolution of $[Y]$ and $[Y]$ to the corresponding scenarios is derived by solving the glycolysis ODEs numerically and illustrated on the right column.

region while continuously being repelled by the fixed point. The consequence is that the system moves along a *limit cycle*, which is a closed trajectory (loop) around the fixed point that the system approaches asymptotically for $t \to \infty$. This might seem like a different version of a non-isolated fixed point (Figure 58) but the key difference is that on a fixed point the system does not move in any direction whereas on a limit cycle the system stays in motion while being constrained to the limit cycle. Such limit cycles are important since they provide an inner clock of the system (e. g. circadian rhythm, menstrual cycle etc.).
If the fixed point is stable, the system would move along a spiral clockwise trajectory approaching the fixed point and asymptotically reaching steady state. The limit cycle and the steady state scenario are shown in Figure 64.
Depending on the parameters a and b, the system would either perform a limit cycle or turn into steady state - even if it is perturbed. Such a behaviour is called *self regulation* and is an important feature in biological systems. Self regulating circuits are relevant especially in living systems because constant conditions (temperature, PH value, etc.) are preferred. This example illustrates the self regulation of the glucose level in blood via self regulating glycolysis reactions.

4.5 n - Dimensional ODEs: The Goodwin Oscillator

We now extend the considerations about stability of a system to a n-dimensional feedback loop. Such a system was notably applied to a biological background by Goodwin[52] in 1963 [2]. The Goodwin model was mainly derived for modeling gene regulatory networks and it exists in many different versions. For the original work I like to refer to [2]. The notation of the variables and the structure of this subsection was adopted from the helpful summary of this topic in [1]. The principle of such a feedback loop is illustrated in Figure 65. In contrast to the previous sections, we do not investigate the entire mathematical concept, but deduce the principle properties, based on the knowledge about simpler systems.
In a feedback cascade, the $i+1^{th}$ reaction is directly influenced by the preceding i^{th} reaction. We know already (c. f. Equation 4.21) that if

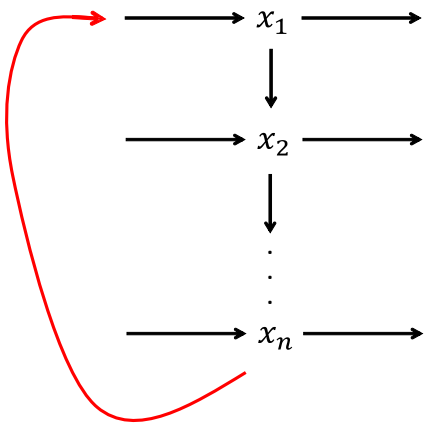

Figure 65: Simplified scheme of a feedback cascade illustrating the principle of the Goodwin oscillator.

$$\frac{\partial}{\partial x_i}\left(\frac{dx_{i+1}}{dt}\right) > 0 \tag{4.62}$$

the system tends to diverge. Hence, the synthesis of x_i destabilizes the system. The quantity x_i is therefore called *activator*.
If x_i and x_{i+1} are coupled such that

$$\frac{\partial}{\partial x_i}\left(\frac{dx_{i+1}}{dt}\right) < 0 \tag{4.63}$$

the system is stabilized and x_i is called an *inhibitor*.
We also know that the inverse of Equation 4.62 and Equation 4.63 equals a time scale within which the system becomes unstable or relaxes, respectively. An entire feedback loop is called positive if

$$\prod_i^n \frac{\partial}{\partial x_i}\left(\frac{dx_{i+1}}{dt}\right) > 0\ , \tag{4.64}$$

where $x_{n+1} \equiv x_1$ and the loop is called negative if

$$\prod_i^n \frac{\partial}{\partial x_i}\left(\frac{dx_{i+1}}{dt}\right) < 0\ . \tag{4.65}$$

[52] Brian Carey Goodwin, 1931 - 2009

The Goodwin oscillator has the form (see page 30 in the original paper [2])

$$\dot{x}_1 = \frac{a}{1+x_n^\rho} - \alpha_1 x_1 \tag{4.66a}$$
$$\dot{x}_2 = c_1 x_1 - \alpha_2 x_2 \tag{4.66b}$$
$$\vdots$$
$$\dot{x}_n = c_{n-1} x_{n-1} - \alpha_n x_n \ . \tag{4.66c}$$

At steady state ($\dot{x}_i = 0$) these equations can be written as

$$\alpha_1 x_1 = \frac{a}{1+x_n^\rho} \tag{4.67a}$$
$$\alpha_2 x_2 = c_1 x_1 \tag{4.67b}$$
$$\vdots$$
$$\alpha_n x_n = c_{n-1} x_{n-1} \tag{4.67c}$$

and we therefore can combine the first two equations to

$$\frac{\alpha_1 \alpha_2}{c_1} x_2 = \frac{a}{1+x_n^\rho} \ , \tag{4.68}$$

the first three equations to

$$\frac{\alpha_1 \alpha_2 \alpha_3}{c_1 c_2} x_3 = \frac{a}{1+x_n^\rho} \tag{4.69}$$

and so on. Thus, in steady state, the Goodwin equations can be written in the compact form

$$\boxed{x_n \frac{\prod_i^n \alpha_i}{\prod_i^{n-1} c_i} = \frac{a}{1+x_n^\rho}} \ . \tag{4.70}$$

The products of the constants α_i and c_i are also constants as well as the pre-factor a yielding altogether the constant C, so that Equation 4.70 can be written in the even more compact form

$$C\, x_n = \frac{1}{1+x_n^\rho} \ . \tag{4.71}$$

Investigating the stability of a n-dimensional system is not different from investigating a two or three dimensional set of ODEs. In order to predict the behaviour of the system after a small perturbation, we have to calculate the Jacobian matrix $\frac{\partial \dot{x}_i}{\partial x_j} = a_{ij}$ (c. f. Equation 4.21 and Section 4.4.1). From Equation 4.66, we find that for example $\frac{\partial \dot{x}_1}{\partial x_1} = -\alpha_1$ and $\frac{\partial \dot{x}_i}{\partial x_i} = -\alpha_i$ that are the diagonals of the matrix. Furthermore, $\frac{\partial \dot{x}_2}{\partial x_1} = c_1$ or in general $\frac{\partial \dot{x}_{i+1}}{\partial x_i} = c_i$ that gives the entries below the diagonal. We also find that the first row ($i = 1$) has the entry $\frac{\partial \dot{x}_1}{\partial x_n} = -\frac{a\rho x_n^{\rho-1}}{\left(1+x_n^\rho\right)^2}$ in the last column ($i = n$). All other entries in the Jacobian matrix equal zero. Thus, the entire matrix reads

$$A = \begin{pmatrix} -\alpha_1 & 0 & & \dots & -\frac{a\rho x_n^{\rho-1}}{\left(1+x_n^\rho\right)^2} \\ c_1 & -\alpha_2 & 0 & \dots & 0 \\ 0 & c_2 & -\alpha_3 & \dots & 0 \\ & & \vdots & & \\ 0 & 0 & \dots & c_{n-1} & -\alpha_n \end{pmatrix} \ . \tag{4.72}$$

As discussed in Section 4.4.1) and in particular in Section 4.4.3, the real part of **all** eigenvalues λ of A have to be negative in order to obtain a stable system. Solving the characteristic equation $det(A - \lambda I)$ leads after some extensive but simple algebra to the condition

$$-\prod_i^n (\alpha_i + \lambda) = a \frac{\rho x_n^{\rho-1}}{(1 + x_n^\rho)^2} \prod_i^{n-1} c_i \ . \tag{4.73}$$

Inserting the condition for stability again leads to extensive algebra and I therefore only give the result

$$\rho > \frac{1}{(\cos[\pi/n])^n} \frac{1}{1 - C x_n} \ , \tag{4.74}$$

that is the condition for an oscillation (repeller).
The quantity ρ equals the cooperativity or Hill[53] coefficient, i. e. the number of particles required for a reaction that have to be present at the same time. In the minimal model proposed by Goodwin $n = 3$ so that $C = \frac{\alpha_1 \alpha_2 \alpha_3}{a\, c_1 c_2}$ and $\cos \pi/n = 0.5$. This leads to $\rho > 8$ in steady state (Equation 4.71) that is considered as an unlikely high cooperativity for a biological system. Such a conceptional problem can be avoided for example by including a biological reasonable Michaelis - Menten term for $\dot{x}_n$. In such a case, the required inequality relation is a function of the speed of the enzyme and large Hill coefficients are not necessary. For a deeper discussion, however, I refer to further reading.
A simulation of the three dimensional Goodwin oscillator is shown in Figure 66.

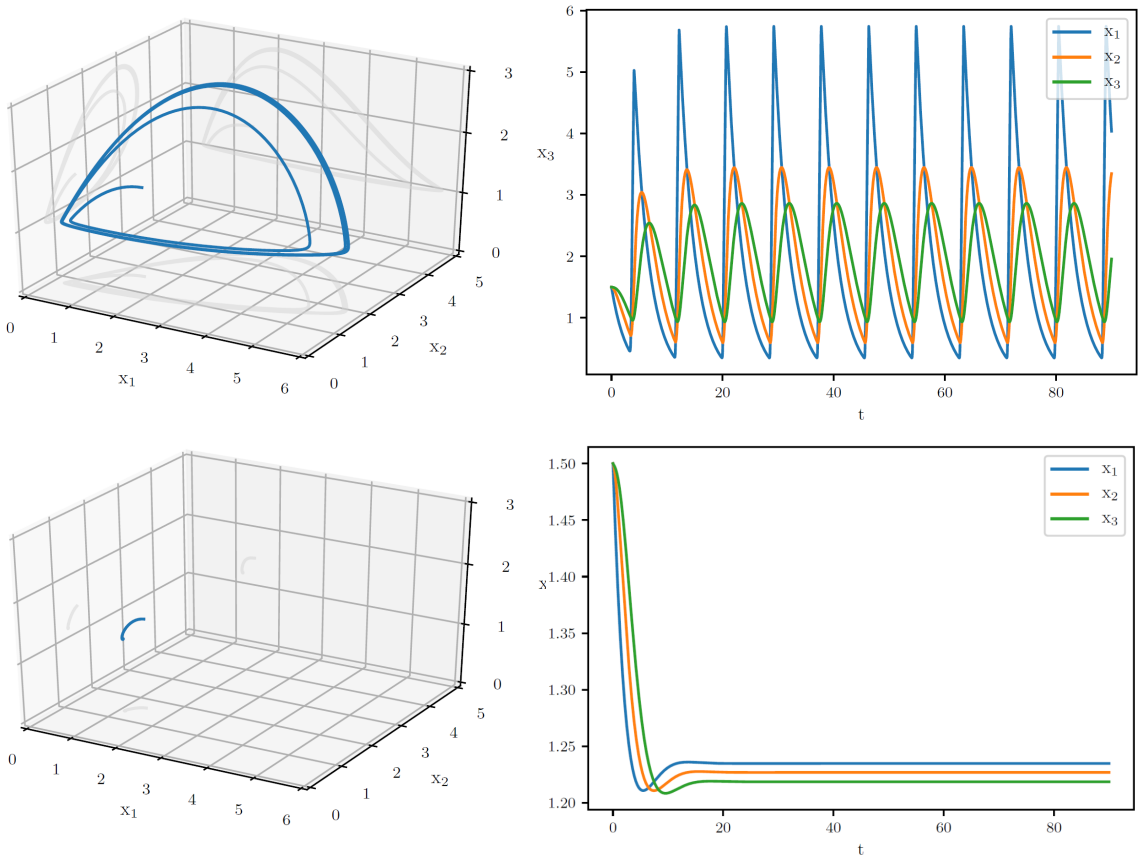

Figure 66: Simulation of the Goodwin oscillator ($n = 3$) once when obeying Equation 4.74 (oscillation, upper panel) and when violating Equation 4.74 (no oscillations, lower panel). All units are arbitrary.

[53] Archibald Vivian Hill, CH OBE FRS, 1886 - 1977

References

[1] Emery D. Conrad, *"Mathematical Models of Biochemical Oscillations"*, Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia 1999.

[2] Brian C. Goodwin, *"Temporal Organization in Cells - A Dynamic Theory of Cellular Central Processes"*, Academic Press, London and New York 1963.

[3] Steven H. Strogatz, *"Nonlinear Dynamics and Chaos"*, Perseus Book Publishing, 1994.

5 Stochastic Ordinary Differential Equations

In Section 4 we described reaction kinetics with ordinary differential equations (ODEs) and discussed stability and self regulation. ODEs are fully deterministic and well describe a system with a large number of particles being involved in the reaction process. However, when describing single molecule processes like protein folding or gene expression, individual stochastic processes become important. These processes have to be taken into account and in contrast to deterministic ODEs, the trajectories of a system can now be very different, even for exactly identical initial conditions and boundary conditions. The equations describing such a system should converge to deterministic ODEs, when applied to a large number of molecules. Such a behaviour is called *correspondence principle*.
Before studying processes like gene expression, we need some conceptual mathematical preparation in order to set up a proper model. The mathematical tools we derive now are the basics of quantum mechanics in physics and therefore most scientists that are doing their research in these biological fields are theoretical physicists. Also I as a theoretical physicist was surprised that one can apply these methods in an almost one - to - one manner. However, it is time that biologists and biochemists should learn these methods in order to be able to describe their own systems, to learn the "language" of theoreticians and to get more independent from external researchers.
In contrast to a lecture in theoretical physics where everything is derived from an abstract algebraic (and thus, more general) approach, I like to introduce this topic in a more intuitive way without extending the mathematical machinery too much. It turned out in the lectures that a good introduction into this topic is the poissonian stepper, that will be explained now.

5.1 From the Poissonian Stepper to the Master Equation

Suppose a kinesin molecule moving along a micro tubule. This molecule moves only in one particular direction on the micro tubule, it never performs a step backwards but only moves forward, say in positive x direction along the micro tubule. The moment *when* the molecule performs a step is random. If many different molecules would start their walk at the same position at the same time and if we would stop the time, say after t seconds have elapsed, we would find that not all kinesin molecules are located at the same position. We would actually find a kind of distribution of molecules versus location (Figure 67). We can now ask for this distribution and therefore can ask for the probability P to find a molecule at a particular position n at given time t. If we do our experiment with these kinesin molecules we can also ask for a distribution of times t that the molecules have arrived at position n. In a deterministic process, all molecules would have arrived at the same position n at the same time t.
Generally, n must not necessarily be a position. It can be the state "folded" or "unfolded" in case of protein folding or "mutated" and "not mutated" when modeling the onset of cancer by point mutations (Section 5.3). Therefore, we rather denote n as a *state*. The important thing here is, that we only allow the system to jump from state n to state $n+1$, but do not permit a jump back (i. e. from $n+1$ to n) as a first simplification (that will be relaxed in Section 6).
The probability for jumping from one state to the next state within a given time step dt is called *hopping rate*, that I like to denote as ν. The hopping rate is constant over all n and has the unit probability per time. The situation is illustrated in the upper panel of Figure 68.

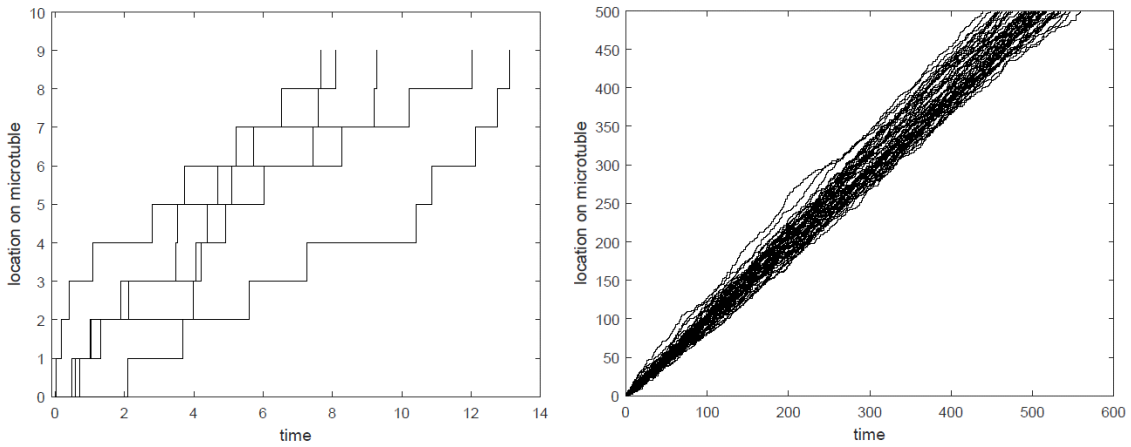

Figure 67: Random motion of kinesin molecules along a microtuble. Due to the stochasticity, the molecules arrive at a particular position at different times. We might ask for the distribution of these arrival times.
Left: Five different runs for ten steps.
Right: Fifty different runs for 500 steps.

We now assume, that we can resolve the motion of the poissonian stepper within a time

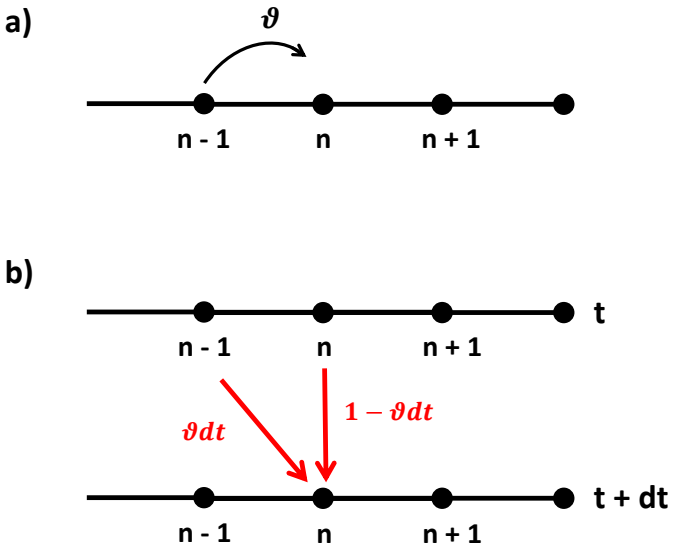

Figure 68: The poissonian stepper with a constant hopping rate ν moves only towards increasing states n (upper panel). The same situation including the time coordinate is shown below.

step dt in such a way, that dt is small enough so that the stepper can either perform *one* step only, or *none*. The time increment dt is always small enough that never two or three (or more) events occur. Since the unit of ν is probability per time and dt **is small**, the product $\nu\, dt$ equals the probability that the system jumps from one state to the next state. What is now the probability $P(n, t + dt)$ to find the system in state n at time $t + dt$? Either the system was in state $n - 1$ at time t and jumped with the probability $\nu\, dt$ to state n, or it was already in state n and stayed in this state within dt. The probability for staying in n within dt is just the complementary probability for a jump, hence $1 - \nu\, dt$. States higher than n are disregarded, since the system cannot jump back and states lower

than $n-1$ are ignored too since we assumed that dt is small enough in order to observe only one (or none) step (see also Figure 68, lower panel).
Thus, the probability $P(n, t+dt)$ to observe the system in state n at time dt equals the probability that the system was in state $n-1$ at t and performed a jump ($\nu\, dt$) plus the probability that the system *is already* in state n at t and *did not* performed a jump ($1-\nu\, dt$) or expressed in equations:

$$\begin{aligned} P(n, t+dt) &= P(n-1, t)\, \nu dt + P(n, t)\, (1 - \nu dt) && \text{(5.1a)} \\ &= P(n-1, t)\, \nu dt + P(n, t) - P(n, t)\, \nu dt\,. && \text{(5.1b)} \end{aligned}$$

The probability distribution $P(n, t+dt)$ we are aiming on is yet unknown, but since dt is small, we can perform a Taylor approximation (Section 2.2.2) up to the linear order and obtain

$$P(n, t+dt) \approx P(n, t) + \frac{dP(n, t)}{dt}\, dt. \tag{5.2}$$

Combining Equation 5.1b and Equation 5.2 leads to

$$\boxed{\frac{dP(n, t)}{dt} = \nu\, P(n-1, t) - \nu\, P(n, t)}, \tag{5.3}$$

the so-called *master equation*.
The master equation expresses the temporal behaviour of the probability of a system being in state n at time t as the sum of terms that lead to the state (gain term, the first addend on the rhs) and terms that lead away from the states (loss terms, the second addend on the rhs in Equation 5.3). In general, a system could switch from state n to *any* other state $m \neq n$ or back from m to n, each case with its own hopping rate $\nu_{mn} \neq \nu_{nm}$, so that we have to sum all gain and loss terms for the general mater equation

$$\boxed{\frac{dP(n, t)}{dt} = \sum_m \nu_{mn}\, P(m, t) - \sum_m \nu_{nm}\, P(n, t)}. \tag{5.4}$$

5.1.1 The Generating Function

But let us return to our simplified model described by Equation 5.3. This equation is a stochastic ODE because it contains the probabilities $P(n, t)$ and $P(n-1, t)$ instead of deterministic quantities like e. g. a concentration. Obtaining an analytical solution $P(n, t)$ to the master equation is rarely straight forward. Fortunately, it is not always necessary to know the entire solution but only its moments (like the mean or the variance) or the behaviour of the solution for a limited parameter range (e. g. around its maximum). In the simple case of Equation 5.3 an analytical solution exists, but I like to introduce the trick of using a so-called *generating function* G (since it also helps to treat the more complicated cases e. g. Figure 72) here although a straight forward way would be feasible as well.
We remember the fact that a function can be written as a polynomial using the Taylor expansion (Equation 2.115), e. g. we found that for example $\sum_{n=0}^{\infty} \frac{z^n}{n!} = e^z$ (Equation 2.125) or that (Section 2.6.6) $\sum_{n=0}^{\infty} \binom{c}{n} z^n = (1+z)^c$. Hence, there is always the structure of a pre-factor and a polynomial z^n, so that we can define the generating function

$$\boxed{G(z, t) = \sum_{n=0}^{\infty} P(n, t) z^n}, \tag{5.5}$$

where $P(n,t)$ is the pre-factor in the Taylor series. This structure seems to appear quite sudden, but the purpose of such an ansatz becomes clear, when we perform the fist temporal derivative

$$\frac{d}{dt}G(z,t) = \sum_{n=0}^{\infty} \frac{dP(n,t)}{dt}\, z^n\,, \tag{5.6}$$

because we then can insert the master equation (Equation 5.3) on the rhs and obtain

$$\begin{aligned} \frac{d}{dt}G(z,t) &= \sum_{n=0}^{\infty} \left[\nu\, P(n-1,t) - \nu\, P(n,t)\right]\, z^n & \text{(5.7a)} \\ &= \sum_{n=0}^{\infty} \nu P(n-1,t) z^n \,-\, \sum_{n=0}^{\infty} \nu P(n,t) z^n\,. & \text{(5.7b)} \end{aligned}$$

The rhs of Equation 5.7b are two independent sums that reach from $n = 0$ to infinity. We can therefore rename the index $n - 1 = n'$ of the first sum and write

$$\frac{d}{dt}G(z,t) = \sum_{n=0}^{\infty} \nu P(n',t) z^{n'+1} \,-\, \sum_{n=0}^{\infty} \nu P(n,t) z^n \tag{5.8}$$

and rename back again $n' = n$. This might seem dubious, but the naming of an index is just arbitrary and it can be shifted in any direction. Also note, that $P(n \neq 0, t = 0) = 0$ and $P(n = 0, t = 0) = 1$ so that $G(z, t = 0) = 1$, since the system starts at $n = 0$ (that is again an arbitrary setting that simplifies the math).
Including the index shift, the derivative of the generating function can be written as

$$\begin{aligned} \frac{d}{dt}G(z,t) &= \sum_{n=0}^{\infty} \nu P(n,t) z^{n+1} \,-\, \sum_{n=0}^{\infty} \nu P(n,t) z^n & \text{(5.9a)} \\ &= \sum_{n=0}^{\infty} \left[z^{n+1} - z^n\right]\, \nu P(n,t) & \text{(5.9b)} \\ &= \underbrace{\sum_{n=0}^{\infty} z^n P(n,t)}_{G(z,t)}\, \nu\, [z-1] & \text{(5.9c)} \\ &= \nu\, [z-1]\, G(z,t)\,. & \text{(5.9d)} \end{aligned}$$

The last step in Equation 5.9d is important, because the solution of this equation is rather simple. Taking the initial conditions discussed above into account, we can now solve Equation 5.9d and obtain

$$G(z,t) = e^{\nu(z-1)t}\,. \tag{5.10}$$

In the first glimpse, Equation 5.10 and Equation 5.5 do not exhibit any similarities. In particular, we aimed on solving for $P(n,t)$ that, however, is not present in Equation 5.10. We have to express Equation 5.10 as a sum of polynomials containing $P(n,t)$ as a pre-factor. The solution can be written as $G(z,t) = e^{-\nu t} e^{\nu z t}$. If we now use the Taylor series of $e^{\nu z t}$ (Equation 2.125), we find that

$$G(z,t) = e^{-\nu t} \sum_{n=0}^{\infty} \frac{(\nu t)^n}{n!} z^n \tag{5.11}$$

and obtain the same structure as Equation 5.5. Comparing Equation 5.11 to Equation 5.5 leads to the solution of the master equation (Equation 5.3):

$$\boxed{P(n,t) = \frac{(\nu t)^n}{n!}\, e^{-\nu t}} \tag{5.12}$$

that is the Poisson distribution (Equation 2.251). Hence, the random process with a constant hopping rate leads to a Poisson distribution (therefore, the name poissonian stepper originates from) of the probability to find the system in state n at time t. That Equation 5.12 is indeed the solution can be verified by inserting it into Equation 5.3.
I like to discuss the generating function a bit further before returning to the poissonian stepper. As mentioned before, we sometimes need only the moments of $P(n,t)$ that can be calculated from $G(z,t)$. By construction of $G(z,t)$ (Equation 5.5), the k^{th} moment can be calculated via

$$z^k \frac{\partial^k G(z,t)}{\partial z^k}\bigg|_{z=1} = \frac{\partial^k G(z,t)}{\partial \left(\ln z\right)^k}\bigg|_{z=1} \tag{5.13}$$

where I used $\frac{\partial}{\partial(\ln z)} = \frac{\partial z}{\partial(\ln z)}\frac{\partial}{\partial z} = z\frac{\partial}{\partial z}$ in the last step.
Let us for example calculate the first two moments. The first moment is the mean ($k=1$) and we find that

$$z\frac{\partial G(z,t)}{\partial z}\bigg|_{z=1} = \sum_{n=0}^{\infty} P(n,t)\, n z^n|_{z=1} = \sum_{n=0}^{\infty} P(n,t)\, n \tag{5.14}$$

yields the mean of $\langle n(t)\rangle = \sum_{n=0}^{\infty} P(n,t)\, n$ weighted by its probability density distribution $P(n,t)$.
The second moment ($k=2$) can be derived by

$$\begin{aligned}
z^2\frac{\partial^2 G(z,t)}{\partial z^2}\bigg|_{z=1} &= z^2 \sum_{n=0}^{\infty} \frac{\partial}{\partial z}\left[P(n,t)\, n z^{n-1}\right]\bigg|_{z=1} && \text{(5.15a)}\\
&= z^2 \sum_{n=0}^{\infty} \left[P(n,t)\, n(n-1) z^{n-2}\right]\big|_{z=1} && \text{(5.15b)}\\
&= \sum_{n=0}^{\infty} \left[P(n,t)\, n(n-1)\right] && \text{(5.15c)}\\
&= \langle n^2(t)\rangle - \langle n(t)\rangle\,, && \text{(5.15d)}
\end{aligned}$$

that leads to the variance, $\sigma_n = \sqrt{\langle n^2\rangle - \langle n\rangle^2}$. In this way, all higher moments of $P(n,t)$ can be derived without actually knowing the function.

5.1.2 The Waiting Time

Although the poissonian stepper has a constant hopping rate, the time that elapses between two hopping events varies due to the random nature of this process (see also Figure 67). Thus, one might ask for the distribution of times between two hopping events, the *waiting time distribution* $\omega(t)$ and its mean and variance. We introduced already the probability $\nu\, dt$ that the system performs one step from n to $n+1$ within the time slot dt and the probability $1-\nu\, dt$ that the system does not change and stays at the current position n.
Let us put the poissonian stepper on the initial position ($n=0$). The probability, often

denoted as *survival probability*, that **no** step occurred after the time t has elapsed equals (according to Equation 5.12) $P(0,t) = e^{-\nu t}$. Since the subsequent steps are independent, this relation holds for any step from state n to the next state $n+1$. Hence, a typical time scale $t = \tau$ would be

$$\tau = \frac{1}{\nu} \ln \left[\frac{1}{P(0,t)} \right] . \tag{5.16}$$

The probability that the stepper **has** moved (no matter if it was one step, two steps or any other number) equals just the complementary probability $\bar{P} = 1 - e^{-\nu t}$. $\bar{P}$ is a cumulative density function (cdf), not a probability density function (pdf) since it gives the probability for an event *until* the time t and not *at* time t. For $t \to \infty$ the probability $\bar{P}$ approaches one, that makes sense, because as longer one waits, it becomes more likely, that the system has left the previous position.
The corresponding pdf, the change of $\bar{P}$ wrt time, hence $\frac{d}{dt}\bar{P} = \nu e^{-\nu t}$, is the waiting time distribution

$$\omega(t) = \nu e^{-\nu t} \tag{5.17}$$

we aimed on.
We can therefore calculate the mean waiting time

$$\int_0^\infty t\, \omega(t)\, dt = \int_0^\infty t\, \nu e^{-\nu t}\, dt = \frac{1}{\nu} . \tag{5.18}$$

Hence, typically (but not always!) after the time $t = 1/\nu$ the system has changed its state. This time scale corresponds to a probability of $P(0,t) = 1/e$ (comparing Equation 5.18 to Equation 5.16).
The poissonian stepper is a very helpful concept, since it can be used not only for a kinesin molecule, but also for modeling cancer incidence by point mutations (Section 5.3), gene expression (Figure 72) or single molecule reactions (Section 5.2.1). Therefore, let us briefly summarize the properties of a poissonian stepper:

$P(n,t) = \frac{(\nu t)^n}{n!} e^{-\nu t}$	
$P(0,t) = e^{-\nu t} \to \tau = \frac{1}{\nu} \ln \left[\frac{1}{P(0,t)} \right]$	waiting time between two events
$\omega(t) = \nu\, e^{-\nu t}$	waiting time distribution
$\frac{dP(n,t)}{dt} = \nu\, P(n-1,t) - \nu\, P(n,t)$	master equation

5.2 Single Molecule Reactions and the Gillespie Algorithm

We now have the tools to investigate single molecule reactions by taking random fluctuations into account. In Section 4, we worked with concentrations $[A]$ and $[B]$ and modeled the reaction kinetics with deterministic ODEs. Since defining a concentration does only make sense for large numbers of molecules, I like to call this the *macroscopic* or *classic* approach. In the *microscopic* case, that we will discuss now, we count single molecules. For example considering the simple reaction $A \xrightarrow{k} B$: $A(t)$ and $B(t)$ are *not* concentrations, but number of molecules. If we start with say ten ($A(t=0)=10$) molecules, the transformation from one A molecule to state B is poissonian. Hence, each time we would run this reaction a particular state, say $A=5$, would have been reached at a different time. The probability for reaching state n at time t obeys the distribution in Equation 5.12.

5.2.1 Solving the Reaction $A \xrightarrow{k} B$

The reaction $A \xrightarrow{k} B$ equals the poissonian stepper in Section 5.1. Since there is only a forward reaction (the decay of A), the system cannot move back to any previous state. The reaction rate k corresponds to the hopping rate ν, i. e. $k \widehat{=} \nu$, and the number of molecules A corresponds to the index of states n. The only difference now is that the probability that one event occurs (i. e. a transformation from $A \xrightarrow{k} B$) is also a function of $A(t)$. If $A(t)$ is large, it is more likely to observe an event in a given time span. According to Equation 5.16, the time elapsing between two successive events is

$$\boxed{\tau = \frac{1}{A(t)k} \ln\left[\frac{1}{P(0,t)}\right]}. \tag{5.20}$$

In order to understand the stochastic nature of this process, let us assume we would like to simulate this reaction numerically. The probability P that *one* reaction occurs within the time interval $[t, t+dt]$ is proportional to $A(t)k\,dt$. One naive approach would be to choose a random number r between zero and one and to compare it to the product $A(t)k\,dt$. If $r < A(t)k\,dt$, the probability for a reaction to occur is large and one A molecule turns into state B, hence $A(t+dt) = A(t) - 1$. If however $r > A(t)k\,dt$, that will be the case towards the end of the reaction since $A(t)$ decreases, no reaction occurs in the given time interval and $A(t+dt) = A(t)$.
However, numerically we have to divide the time into discrete intervals $dt \to \Delta t$ and there will be many intervals where $A(t)kdt << 1$, especially at the end of the reaction when $A(t)$ is low. Thus, we would have many steps in our simulation, where no reaction occurs, that is very inefficient. On the other hand, for large $A(t)$ or at the beginning of the reaction, the product $A(t)k\,dt$ can be larger then one and setting it into relation with r does not make sense.

The solution to the problem is that instead of performing a time step and checking whether a reaction has occurred, we only perform a step in the simulation *if* a reaction occurs. The time increment is then not fixed, but determined by Equation 5.20 that yields subsequent time intervals τ_1, τ_2, ... The value of τ also depends on $P(0,t)$. Equation 5.20 maps this probability to τ according to a poissonian process, i. e. for an uniformly distributed $P(0,t)$, the time intervals τ_1, τ_2, ... are poisson distributed. Hence, a proper and effective algorithm for such a simulation would be the following procedure:

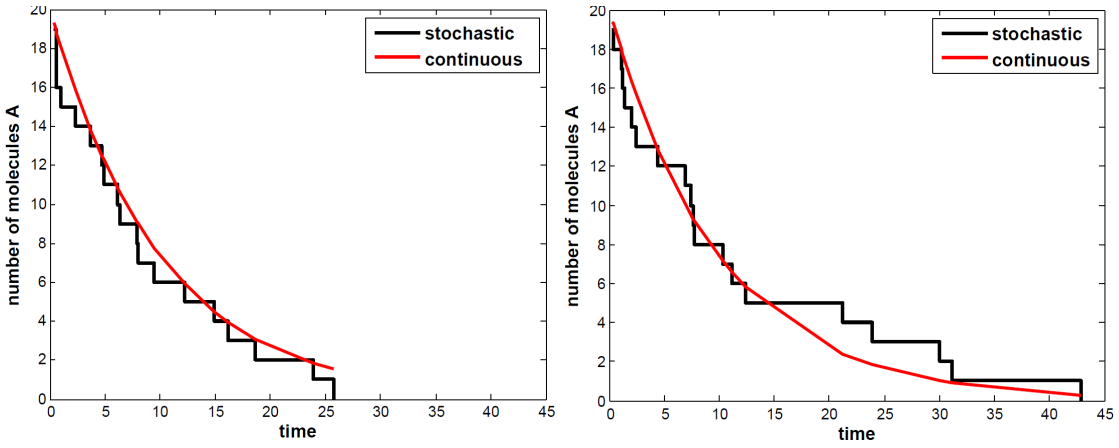

Figure 69: Two different runs of the stochastic reaction $A \xrightarrow{k} B$ (black) using the Gillespie algorithm. Both reactions started with $n_0 = 20$ molecules and a constant rate k of 0.1 molecules per time unit. The difference between the two runs is caused by the stochastic character of the process (Equation 5.20) - in contrast to the analytical solution (red) that can be approximated from the mean of many stochastic runs (or large A) that produces a continuous and always identical exponential decay.

1. generate a random number $r =$ rand (uniformly distributed between 0 and 1)
2. calculate the time elapsing to the next event (Equation 5.20): $\tau = \frac{1}{A(t)k} \ln\left[\frac{1}{r}\right]$
3. change the state; here: $A(t+\tau) = A(t) - 1$
4. go to the first step.

Since r is random, the process (τ) is random and it follows a poissonian distribution (Equation 5.20). The number of molecules $A(t)$ is decreasing by one unit in every time step τ, where τ tends to increase (within the statistical fluctuations) because $A(t)$ appears in the denominator of Equation 5.20. This procedure is called *Gillespie algorithm*[54].
Two runs of the simulation of this reaction is shown in Figure 69.

Exercise I:

Write a code performing the reaction $A \xrightarrow{k} B$ using the Gillespie algorithm (avoid loops). Verify the correctness of your code by generating Figure 69.

What is the corresponding master equation of this reaction? As already mentioned, the reaction $A \xrightarrow{k} B$ equals the poissonian stepper and we therefore define $P(n,t)$ as the probability of having n molecules of A at a given time t (c. f. Equation 5.3). We sum up all the possible transitions that lead to this state, that is only $P(n+1,t)$ (no back reaction) and all transitions that lead away from $P(n,t)$, i. e.

$$\frac{dP(n,t)}{dt} = -\text{away from } P(n,t) + \text{to } P(n,t) \quad n \text{ molecules of } A\,. \tag{5.21}$$

[54] Joseph L. Doob, 1910 - 2004 and Daniel Thomas Gillespie, 1938 - today

The difference to the poissonian stepper is that the probability of a transition is proportional to the amount of molecules n. If the system contains $n+1$ molecules and each molecule can decay with the rate k, then the probability that one decay event occurs is $k\,(n+1)$. Recall, that we define our time steps always in such a way that we either observe one event or none. Thus, the master equation reads

$$\boxed{\frac{dP(n,t)}{dt} = -knP(n,t) + k(n+1)P(n+1,t)}\,. \tag{5.22}$$

Let us denote the number of molecules for $t=0$ as n_0. The probability for having n_0+1 molecules at $t=0$ equals zero so that the second addend on the rhs of Equation 5.22 vanishes for the initial conditions, whereas $P(n_0, t=0) = 1$ by definition. Thus, the master equation under the initial condition reads

$$\begin{aligned}\frac{dP(n_0,t)}{dt} &= -kn_0\,P(n_0,t)\\ P(n_0,t) &= \underbrace{P(n_0,t=0)}_{=1}\,e^{-kn_0\,t} = e^{-kn_0\,t}\,.\end{aligned}$$

We can solve the master equation e. g. using the generating function (Section 5.1.1) that, however, requires extensive algebra [2] and I therefore like to give the final result:

$$\boxed{P(n,t) = e^{-knt}\binom{n_0}{n}\left[1-e^{-kt}\right]^{n_0-n}}\,. \tag{5.23}$$

Equation 5.23 is the exact solution of the master equation and fully describes the single molecule reaction $A \xrightarrow{k} B$.

Exercise II:

Proof that Equation 5.23 is the correct solution by inserting $P(n,t)$ to Equation 5.22.

When running a simulation (e. g. using the Gillespie algorithm) of this process with a large number ($n_0 >> 100$) of molecules, we obtain a smooth exponential decay (c. f. red curve in Figure 69) in contrast to a step function we find when running the simulation for a small number of molecules (black curve in Figure 69). The reason is that the statistical fluctuations in this process (e. g. the length of the time steps τ_i, Equation 5.20) even out for large n. The same effect can be observed when running the simulation for small n, but with a large number of repetitions. The mean curve $\langle n(t)\rangle$ of all these runs equals again a smooth exponential decay. This phenomenon is called *correspondence principle*. Thus, the classical, macroscopic reaction is a sum of the contributions from all the single molecule reactions.
This effect can also be seen in the model. According to the definition of the mean (integral or sum of a quantity, here n, weighted by the probability density function), the mean curve $\langle n(t)\rangle$ equals

$$\langle n(t)\rangle = \sum_{n=0}^{n_0} n(t)P(n,t)\,. \tag{5.24}$$

Again, solving this equation requires some simple, but extensive, algebra. The solution to $\langle n(t) \rangle$ is

$$\langle n(t) \rangle = n_0 \, e^{-k\,t} \,, \tag{5.25}$$

the exponential decay.
Equation 5.25 is the solution of the deterministic and macroscopic ODE $\frac{d[A]}{dt} = -k\,[A]$, that is the equivalent to the stochastic (microscopic) ODE Equation 5.22. In this mathematical context, we therefore actually justified the modeling of (macroscopic) chemical reactions with ODEs as done in Section 4. Deterministic ODEs work only for the sub-case of large n ($\gtrsim 100$), whereas stochastic ODEs like Equation 5.22 work for small n *and* large n.

5.2.2 Solving the Reaction $A \underset{k_2}{\overset{k_1}{\rightleftarrows}} B$

Let us now go a step further and include the back reaction. The forward and backward rate k_1 and k_2, respectively, are constant so that the transformation from one state to the other is poissonian. The forward rate times the number of molecules $A(t)$ equals the probability ν_+ to observe one event in forward direction within a small time interval dt, hence $\nu_+ = A(t)\,k_1$. the same applies for the backward direction so that $\nu_- = B(t)\,k_2$. Thus, the probability that *any* reaction occurs within a time interval equals $\nu_0 = A(t)k_1 + B(t)k_2$ and therefore, the waiting time for *any* reaction equals

$$\tau = \frac{1}{\nu_0} \ln \left[\frac{1}{P(0,t)} \right] . \tag{5.26}$$

In order to set up the Gillespie algorithm, one has to generate a random number r_1 for deriving τ in Equation 5.26 and generate a *second* random number r_2 to decide *which* of the two reactions occurs. Both random numbers are uniformly distributed in the interval between zero and one. Since $\frac{A(t)\,k_1}{\nu_0} + \frac{B(t)\,k_2}{\nu_0} = 1$, the decision threshold whether the forward reaction or the backward occurs is the comparison of r_2 to these two probability ratios. Hence, the Gillespie algorithm for this reaction would include the following recursive steps:

1. generate a random number $r_1 =$ rand (uniformly distributed between 0 and 1)
2. calculate the time elapsing to the next event (Equation 5.26): $\tau = \frac{1}{\nu_0} \ln \left[\frac{1}{r_1} \right]$
3. generate a second random number $r_2 =$ rand (uniformly distributed between 0 and 1)
4. compute *which* reaction occurs:

$$A(t+\tau) = \begin{cases} A(t) + 1 & \text{if } r_2 < B(t)k_2/a_0 \rightarrow B(t+\tau) = B(t) - 1 \\ A(t) - 1 & \text{if } r_2 > B(t)k_2/a_0 \rightarrow B(t+\tau) = B(t) + 1 \end{cases} \tag{5.27}$$

5. go to the first step.

The simulation of the reaction $A \underset{k_2}{\overset{k_1}{\rightleftarrows}} B$ is illustrated in Figure 70.
In the same manner as in Section 5.2.1, we find the master equation by summarizing all

loss and gain terms. The system is characterized by n molecules of A and m molecules of B. Thus, we would need the probability to observe the system at the state with n and m molecules at time t, i. e. $P(n,m,t)$. Since here the number of molecules is conserved ($m+n = const$) it is sufficient to use the notation $P(n,t)$ only. Hence, the master equation of this system is

$$\begin{aligned}\frac{dP(n,t)}{d} =& -k_1 n\, P(n,t) + k_1(n+1)\, P(n+1,t) \\ & - k_2 m\, P(n,t) + k_2(m+1)\, P(m+1,t)\,.\end{aligned}$$

We again solve the master equation and calculate the mean $\langle n(t)\rangle$

$$\langle n(t)\rangle = \sum_{n=0}^{n_0} n(t) P(n,t) \tag{5.28}$$

and also the mean for B, $\langle m(t)\rangle$. One can imagine, that the calculations are very extensive and I therefore like to give the final result only, that reads

$$\boxed{\frac{d}{dt}\langle n(t)\rangle = -k_1 \langle n(t)\rangle + k_2 \langle m(t)\rangle}, \tag{5.29}$$

that is the equivalent of the macroscopic case $\frac{d}{dt}[A] = -k_1 [A] + k_2 [B]$.

5.2.3 The Lotka-Volterra Model (Predator-Prey)

The method of single molecule dynamics is also very useful for population studies. In Section 4 we discussed some properties of bound growth and found the Verhulst equation. This equation is a deterministic ODE, but we can also treat such a model stochasticly. Thereby, I like to introduce the predator-prey model or *Lotka-Volterra*[55] model, that, as the name implies, models the dynamics between two species. One specie acts as predator and consumes the preys. However, if a predator does not encounter a prey, it starves and gets "depleted". Usually, in textbooks, the predators are wolfs W and the preys are lambs L (or foxes vs. rabbits). Often, also an empty region E, where neither a lamb, nor a wolf is present, is included.
If a lamb encounters an empty region, it starts to reproduce immediately and our first equation is

$$L + E \xrightarrow{k_1} L + L\,. \tag{5.30}$$

As mentioned, if a wolf does not meet a lamb after some time, it starves and finally dies and leaves an empty region, so that the second equation is

$$W \xrightarrow{k_2} E\,. \tag{5.31}$$

There are two options of what happens when a wolf meets a lamb: either the lamb gets eaten and the wolf reproduces, so that we have

$$W + L \xrightarrow{k_3} W + W \tag{5.32}$$

[55] Alfred James Lotka, 1880 - 1949; Vito Volterra, 1860 - 1940

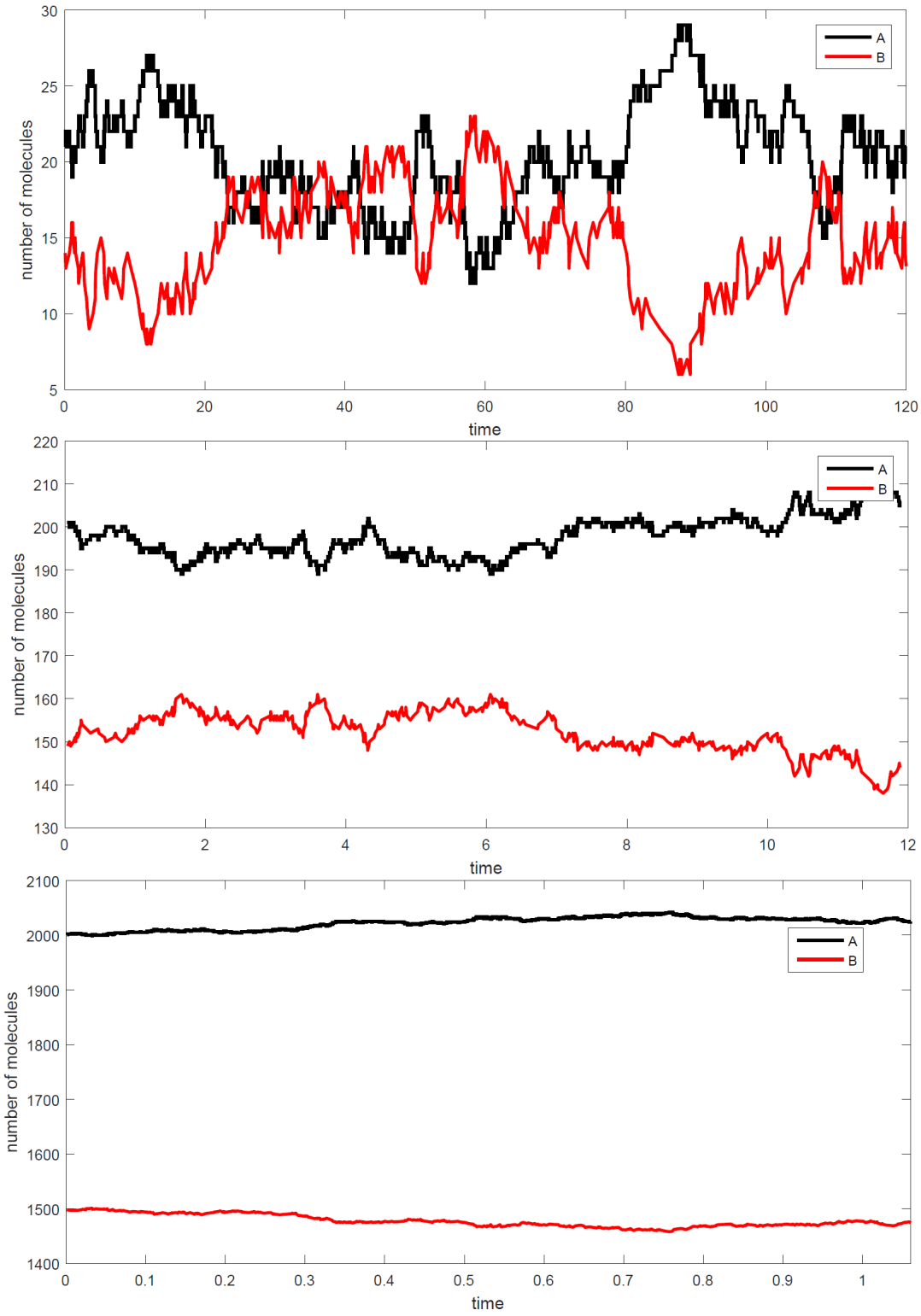

Figure 70: Simulation of the reaction $A \underset{k_2}{\overset{k_1}{\rightleftarrows}} B$ for $k_1 = 0.1$ and $k_2 = 0.15$ using the Gillespie algorithm. The statistical fluctuations become negligible for larger number of molecules (from upper to lower panel).

or, the wolf does not reproduce itself and leaves an empty region where the lamb has been:

$$W + L \overset{k_4}{\rightarrow} W + E\,. \tag{5.33}$$

Of course, such a model is very simple and therefore has its limits in predicting the population dynamics (for example, we need actually two wolfs/lambs of different sex to get a reproduction term).
We can see from these equations, that the total number N of individuals (treating E as an "individual" too) is conserved so that $N = n_L + n_W + n_E$ if n_i is the number of the corresponding species i. The Gillespie algorithm works as in the previous case (Section 5.2.2): we first have to calculate the time elapsing until *a* reaction occurs and then have to decide *which* reaction occurs.
From the first reaction (Equation 5.30), we see that the probability P_L to get one lamb out of N individuals equals $P_L = n_L/N$. Then, we have one lamb less and $N \to N-1$, so that the probability to obtain one E in order to perform the first reaction is $P_E = n_E/(N-1)$. Hence the probability that the first reaction occurs (i. e., L and E have to meet) within a time interval dt is

$$\nu_1 = k_1 P_L P_E \tag{5.34}$$
$$= k_1 \frac{n_L}{N} \frac{n_E}{N-1} \tag{5.35}$$
$$= k_1 \frac{n_L (N - n_L - n_W)}{N (N-1)} \tag{5.36}$$

The term $\frac{n_L(N-n_L-n_W)}{N(N-1)}$ corresponds to $A(t)$ from the example in Section 5.2.2. The only difference is that $A(t)$ is a number and $\frac{n_L(N-n_L-n_W)}{N(N-1)}$ is a probability in order to keep the units of the different k constant. If we would use the numbers n_E, n_L and n_W directly, the unit of k_1 would be "per time per number squared" (like time per square-mol), whereas k_2 would have the unit "per time per number". This simplification does not change the overall model.
In the same manner we obtain

$$\nu_2 = k_2 \frac{n_W}{N} \tag{5.37}$$

for the second reaction,

$$\nu_3 = k_3 \frac{n_L}{N} \frac{n_W}{N-1} \tag{5.38}$$

for the third reaction and finally,

$$\nu_4 = k_4 \frac{n_L}{N} \frac{n_W}{N-1} \tag{5.39}$$

for the fourth reaction.
As in Section 5.2.2, $\nu_0 = \nu_1 + \nu_2 + \nu_3 + \nu_4$ gives the rate that any of those four reactions occur, so that

$$\tau = \frac{1}{\nu_0} \ln \left[\frac{1}{r_1} \right] \tag{5.40}$$

is the time that elapses between two subsequent reactions.
Like before we use a second uniformly distributed random number r_2 between one and zero to choose between the four cases. The Gillespie algorithm can now be set up in the same manner as before:

1. generate a random number $r_1 =$ rand (uniformly distributed between 0 and 1)

2. calculate the time elapsing to the next event (Equation 5.26): $\tau = \frac{1}{\nu_0} \ln\left[\frac{1}{r_1}\right]$

3. generate a second random number $r_2 =$ rand (uniformly distributed between 0 and 1)

4. compute *which* reaction occurs:

   - if $r_2 < \frac{\nu_1}{\nu_0}$, reaction one (Equation 5.36) occurs $\Rightarrow \; L(t+\tau) = L(t) + 1$
   - if $\frac{\nu_1}{\nu_0} < r_2 < \frac{\nu_1+\nu_2}{\nu_0}$ reaction two (Equation 5.37) occurs $\Rightarrow \; W(t+\tau) = W(t) - 1$
   - if $\frac{\nu_1+\nu_2}{\nu_0} < r_2 < \frac{\nu_1+\nu_2+\nu_3}{\nu_0}$ reaction three (Equation 5.38) occurs $\Rightarrow \; W(t+\tau) = W(t) + 1$ and $L(t+\tau) = L(t) - 1$
   - and finally if $\frac{\nu_1+\nu_2+\nu_3}{\nu_0} < r_2 < 1$ reaction four (Equation 5.39) occurs $\Rightarrow \; L(t+\tau) = L(t) - 1$

5. go to the first step.

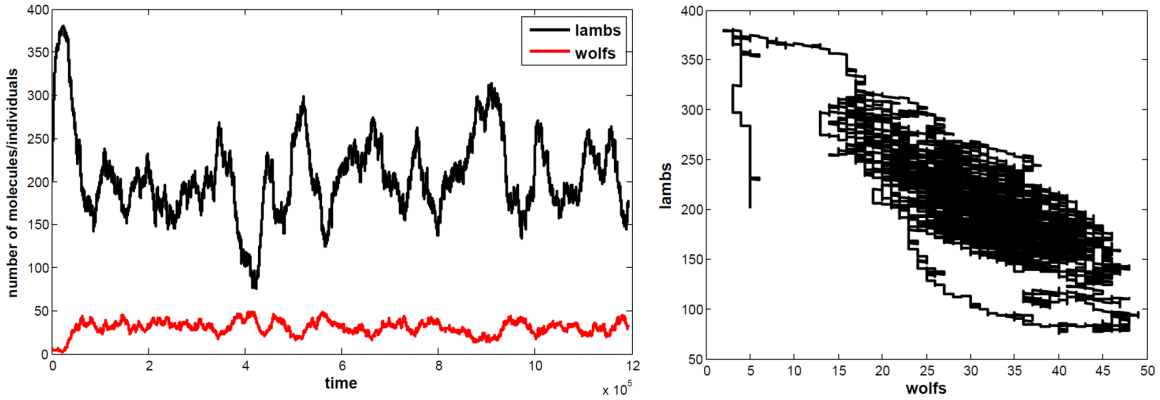

Figure 71: One run of the Lotka-Volterra model (predator-prey) using the Gillespie algorithm (left: number of individuals vs time; right the phase portrait). In contrast to the treatment of a fully deterministic model in Section 4, the populations do not approach a steady state or a limit cycle, since stochastic fluctuations are too prominent.

Mathematically, the Lotka-Volterra model is very similar to Sel'kov's model of glycolysis (comparing Equation 5.37 and Equation 5.38 to Equation 4.56 and Equation 4.57) and therefore exhibits the same behaviour. The evolution and the phase plot of the population of lambs and wolfs is illustrated in Figure 71.

Exercise:

Write a code simulating the Lotka-Volterra model using the Gillespie algorithm. Verify the correctness of your code by generating Figure 71. Modify the code in order

to model glycolysis and compare the results to the findings in Section 4. What happens close to the limit cycle condition?

5.2.4 A Simple Gene Expression Model

Let us now consider a simple model for gene expression. The minimal model contains at least the interaction between DNA and mRNA and the subsequent synthesis of a protein. Such a model is mathematically explained for example in [3] and I like to give some background information that helps to understand the first pages and the first equations therein.
I denote the transcription rate as k_R and the number of $mRNA$ molecules at a given time t as r and the depletion rate for the $mRNA$ molecules as γ_R. The transcription part is described by

$$\frac{dr(t)}{dt} = k_R - \gamma_R\, r(t)\,. \tag{5.41}$$

On the other hand, $mRNA$ is used to build the proteins p with a rate k_P which are depleted with rate γ_P. The equation for translation then reads

$$\frac{dp(t)}{dt} = k_p\, r(t) - \gamma_P\, p(t)\,. \tag{5.42}$$

The situation is illustrated in Figure 72. Note, that there is no $-k_p\, r(t)$ part in Equation 5.41, since $mRNA$ is not consumed while generating a protein.

Furthermore, I denote $w(r,p,t)$ as the probability density function to find r $mRNA$

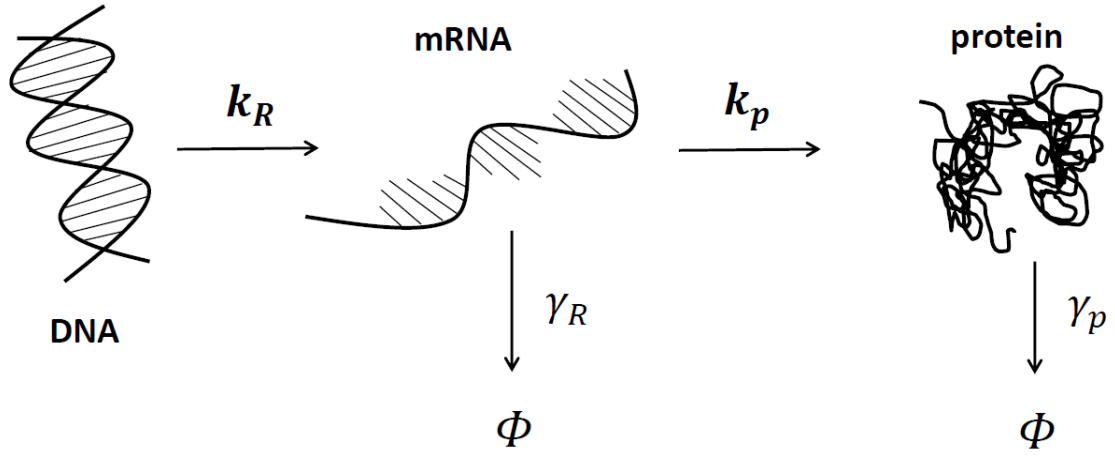

Figure 72: Sketch of a simple gene expression model.

molecules and p proteins at the time t. The master equation is obtained by summing all gain and loss terms that lead to the particular state $w(r,p,t)$ and away from it. There are four neighbouring states of $w(r,p,t)$, hence eight transitions. The contributions to the master equation are illustrated in Figure 73.

Hence, the master equation reads

$$\begin{aligned}\frac{dw(r,p,t)}{dt} = {} & k_R\, w(r-1,p,t) + \gamma_R\,(r+1)\, w(r+1,p,t) + k_P\, r\, w(r,p-1,t) \\ & +\gamma_P\,(p+1)\, w(r,p+1,t) - \left[k_r + \gamma_R\, r + k_P\, r + \gamma_P\, p\right]\, w(r,p,t)\,.\end{aligned} \tag{5.43}$$

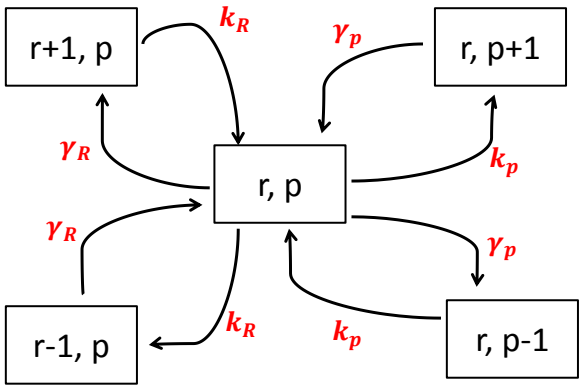

Figure 73: On the construction of the master equation of a simple gene expression model in Figure 72.

The relative complicated structure of Equation 5.43 (although it is the minimal model) implies that deriving its solution is not that trivial. Indeed, there exists no straight forward way to derive the analytic solution $w(r,p,t)$. Therefore, we now apply the trick with the generating function (Section 5.1.1) in order to derive the mean and the variance of r and p (since these quantities are measured in an experiment). In contrast to Equation 5.43 where G is a function of n and t only, the generating function is now a function of r, p and t, hence the polynomials in G must be of power of r and p. Applying the same structure as Equation 5.5, the generating function is defined as

$$\boxed{G(y,z,t) = \sum_{r,p=0}^{\infty} w(r,p,t)\, y^r\, z^p}\,. \tag{5.44}$$

We know, that the sum of $w(r,p,t)$ over all possible configurations of p and r must be one, since the system must have any r and p (conservation of probability), so that

$$G(y,z,t)|_{y=z=1} = \sum_{r,p=0}^{\infty} w(r,p,t) = 1\,. \tag{5.45}$$

The first moments (mean) of r and p are then generated with (Equation 5.13)

$$\frac{\partial}{\partial y} G(y,z,t)\bigg|_{y=z=1} = \langle r \rangle \tag{5.46a}$$

$$\frac{\partial}{\partial z} G(y,z,t)\bigg|_{y=z=1} = \langle p \rangle\ . \tag{5.46b}$$

The second moment of p is derived via

$$\frac{\partial^2}{\partial z^2} G(y,z,t)\bigg|_{y=z=1} = \sum_{r,p=0}^{\infty} \frac{\partial}{\partial z}\left[p\, w(r,p,t)\, y^r\, z^{p-1}\right]\bigg|_{y=z=1} \tag{5.47}$$

$$= \sum_{r,p=0}^{\infty} \left(p^2 - p\right)\, w(r,p,t) \tag{5.48}$$

$$= \langle p^2 \rangle - \langle p \rangle \tag{5.49}$$

and the second moment of r is derived in the same manner by

$$\frac{\partial^2}{\partial y^2} G(y,z,t)\bigg|_{y=z=1} = \langle r^2 \rangle - \langle r \rangle \ . \tag{5.50}$$

A further useful quantity is the mean of the product of r and p that can be obtained by

$$\frac{\partial^2}{\partial y \partial z} G(y,z,t)\bigg|_{y=z=1} = \langle rp \rangle \ . \tag{5.51}$$

Like in Section 5.1.1 we now calculate the temporal derivative of the generating function (Equation 5.44) in order to insert the master equation. We can therefore express the master equation in terms of the generating function and find that

$$\begin{aligned}\frac{\partial}{\partial t} G(y,z,t) = k_R\, y\, G(y,z,t) + \gamma_R \frac{\partial}{\partial y} G(y,z,t) + k_P\, y\, z \frac{\partial}{\partial y} G(y,z,t) + \gamma_P \frac{\partial}{\partial z} G(y,z,t) \\ -k_R G(y,z,t) - \gamma_R\, y \frac{\partial}{\partial y} G(y,z,t) - k_P\, y \frac{\partial}{\partial y} G(y,z,t) - \gamma_P\, z \frac{\partial}{\partial z} G(y,z,t)\,.\end{aligned} \tag{5.52}$$

The gene expression process will reach steady after a certain time, so that $\frac{dr(t)}{dt} = \frac{dp(t)}{dt} = 0$, yielding (Equation 5.41 and Equation 5.42) $r = k_R/\gamma_R$ and $p = k_P k_r/(\gamma_R \gamma_P)$. Steady state also requires that $\frac{\partial}{\partial t} G(y,z,t) = 0$, and therefore Equation 5.52 gives

$$(1-y)\left[\gamma_R \frac{\partial}{\partial y} G(y,z,t) - k_R\, G(y,z,t)\right] = (1-z)\left[k_P\, y \frac{\partial}{\partial y} G(y,z,t) - \gamma_P \frac{\partial}{\partial z} G(y,z,t)\right] . \tag{5.53}$$

The derivatives of G in Equation 5.53 are the mean of r (Equation 5.46a) and p (Equation 5.46b), respectively. We also know from Equation 5.45 that $G(y=1,z=1,t)=1$ due to conservation of probability. Hence, from all different combinations of y and z that solve Equation 5.53, we take $z=1$ and $y=1$ in order to find a solution in the most convenient way.
We first set $z=1$ so that the rhs of Equation 5.53 equals zero and therefore obtain the lhs

$$(1-y)\left[\gamma_R \frac{\partial}{\partial y} G(y,1,t) - k_R\, G(y,1,t)\right] = 0\,. \tag{5.54}$$

This equation is valid if either $y=1$, or if the expression in the square brackets equals zero, or both. For the second case we find

$$\gamma_R\ \langle r \rangle = k_R\, G(y,1,t)\,. \tag{5.55}$$

where we used Equation 5.46a in the last step.
We now set $y=1$ that is a solution of Equation 5.54 and obtain (Equation 5.45) finally the mean of r

$$\boxed{\langle r \rangle = \frac{k_R}{\gamma_R}}\,. \tag{5.56}$$

Returning to Equation 5.53 and now setting first $y=1$ we obtain the rhs of Equation 5.53. In the same way as before, this leads to (using Equation 5.46a and Equation 5.46b)

$$\boxed{\langle p \rangle = \frac{k_P\, k_R}{\gamma_P\, \gamma_R}}\,. \tag{5.57}$$

Similarly, we use Equation 5.53 and the derivatives of $G(y,z,t)$ to find the higher moments of r and p that leads to

$$\boxed{\sigma_r^2 = \langle r \rangle} \tag{5.58}$$

and

$$\boxed{\langle pr \rangle = \frac{k_R \langle p \rangle + k_P \langle r \rangle \left(\langle r \rangle + 1 \right)}{\gamma_R + \gamma_P}} \tag{5.59}$$

and finally

$$\boxed{\sigma_p^2 = \frac{k_P}{\gamma_P} \langle pr \rangle + \langle p \rangle - \langle p \rangle^2}\,. \tag{5.60}$$

The expressions from Equation 5.56 to Equation 5.60 are all observables that can be measured in an experiment and compared to a theoretical model. Thanks to the generating function, we could derive all these quantities without actually solving the master equation. Although it is the minimal model, the underlying algebra is relatively extensive.
Typical values for γ_P/γ_R are a few minutes versus a few hours. The ratio k_P/γ_R is called "burst size" giving the number of proteins produced per $mRNA$ which is in the order of a few dozens.
Often, a stochastic system is characterized by its randomness ρ (sometimes called **Fano**[56] **factor**), that is defined by the ratio of the variance to the mean. If randomness is large (large variance), the system is ruled by stochastic behaviour, if ρ is small, then the system is rather deterministic. In our case

$$\rho_R = \frac{\sigma_r^2}{\langle r \rangle} = 1 \tag{5.61}$$

for the amount of $mRNA$. This ratio equals exactly that of a poissonian process (c. f. Equation 2.252 and Equation 2.253).
The Fano factor of the protein production process is

$$\rho_P = \frac{\sigma_p^2}{\langle p \rangle} = 1 + \frac{k_P/\gamma_R}{1 + \gamma_P/\gamma_R} \approx 1 + \frac{k_P}{\gamma_R} \approx 10 - 50\,. \tag{5.62}$$

[56] Ugo Fano ForMemRS, 1912 - 2001

5.3 Poissonian Cancer Models

An important poissonian process is the development of cancer. Cancer arises through the accumulation of mutations in oncogenes and tumor suppressor genes. Not all mutations lead to cancer, only some particular mutations, so-called *driver mutations*, are required to develop cancer. The question is how many such mutations are required and how are they related to the overall risk of getting cancer, depending on life style and environmental influence (chemicals, radiation, etc).
A first investigation in such a direction was performed in the 1950's by [1]. They observed from cancer mortality statistics data, that the log age of the patients and the log death rate yields a linear function (Figure 74).
A simplified assumption (that will be improved in the next lines) of the underlying process

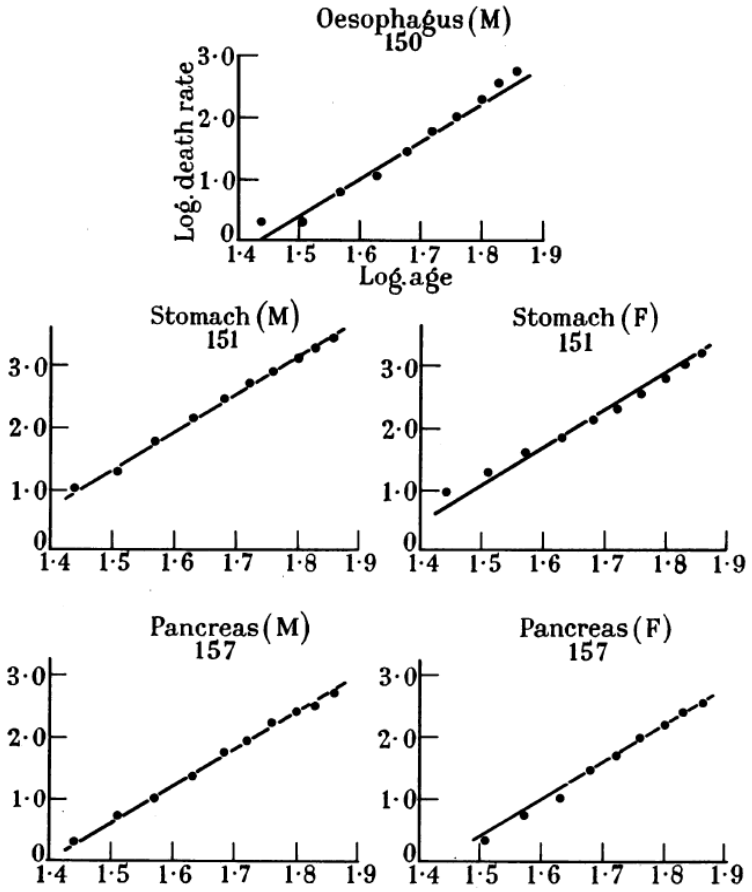

Figure 74: Connection between age (in years) and death rate (per million) of different cancer types (M=male, F=female). The figure was taken from [1].

might be that only point mutations are responsible for cancer. According to the findings in Section 5.1.2 the probability P that a particular gene has been mutated equals $1 - e^{-\nu t}$. A reasonable simplification is that the mutation rate ν is small, so that we can expand P with a Taylor series to the linear order and obtain $P \approx \nu t$. This approximation should be valid throughout the life time (the incidence is in the order of some cases per thousand, see Figure 74) of a patient although νt reaches one after a sufficiently large t (see also Figure 75).
Let us now consider a series of n subsequent point mutations, each with its own mutation rate ν_i. In order to get n mutations, the system must have performed $n - 1$ steps (c.

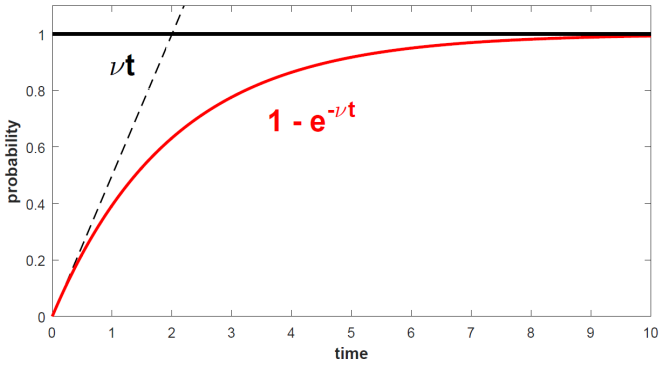

Figure 75: The probability that a particular gene has been mutated (red solid line) and its approximation for a small mutation rate ν (black dashed line).

f. Section 5.1 and Figure 76). The probability $P(t)$ to observe *any* combination of these n mutations equals $P(t) = t^{n-1} \prod_{i=1}^{n} \nu_i$. However, the order of these mutations might matter, so that the probability that *one particular* sequence of mutations occurred equals

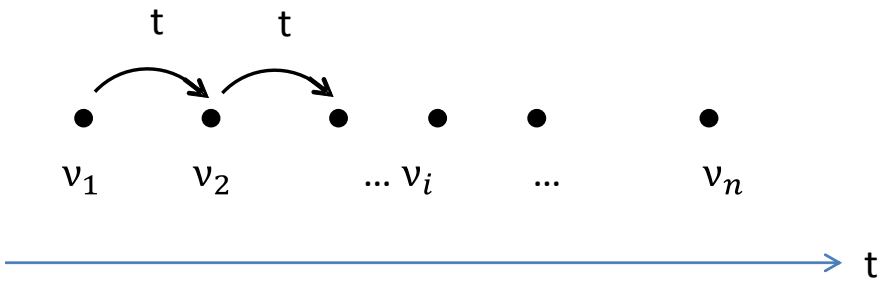

Figure 76: Illustration of the process of n subsequent point mutations.

$$P(t) = \frac{t^{n-1}}{(n-1)!} \prod_{i=1}^{n} \nu_i \,. \tag{5.63}$$

The death rate in Figure 74 is proportional to the probability of cancer incidence $P(t)$. Taking the log of $P(t)$

$$log\,[P(t)] = log\left[\frac{\prod_{i=1}^{n} \nu_i}{(n-1)!}\right] + (n-1)\,log(t) \tag{5.64}$$

leads to the linear relation observed in Figure 74. Thus, according to Equation 5.64, the slope of the plot in Figure 74 equals the number of point mutations minus one $(n-1)$. According to [5] these point mutations (or single base substitutions, SBS) explain most of the alterations in the genome, in particular for colorectal cancer and breast cancer, but only two third in case of medulloblastoma.
A better and therefore more realistic model was proposed by [4]. The probability for a mutation is proportional to the base pair length I_i of the affected gene i so that we can introduce an average mutation rate ν and substitute the individual mutation rate $\nu_i = \nu\, I_i$ and obtain

$$P(t) = \nu^n \frac{t^{n-1}}{(n-1)!} \prod_{i=1}^{n} I_i \,. \tag{5.65}$$

A further step towards a more realistic model is the assumption that not always a fixed number n of driver mutations is required to generate cancer. There is some evidence for a minimal number $min(n) = m$ (≈ 2) and a maximal number $max(n) = M$ (≈ 7). Also not only one, but different combinations of exactly n driver mutations lead to a detectable cancer type. Therefore, we define (according to [4]) j_n as the index of all mutually exclusive combinations of exact n driver mutations yielding cancer and $I_i(j_n)$ as the length of the corresponding gene i in the n gene combination j_n. In order to derive $P(t)$, we now have to sum over all combinations and obtain

$$P(t) = \sum_{n=m}^{M} \underbrace{\sum_{j_n} \nu^n \frac{t^{n-1}}{(n-1)!} \prod_{i=1}^{n} I_i(j_n)}_{I_n} . \tag{5.66}$$

The underbraced part I_n equals the incidence for one particular cancer type at age t through the occurrence of a fixed number of n mutations, taking all mutually exclusive combinations into account.
If now the mutation rate ν is increased by a factor $x > 1$ due to cancerogenes, we obtain a factor x^n

$$P(t) = \sum_{n=m}^{M} x^n \, I_n \, . \tag{5.67}$$

Hence, an increase of the mutation rate by the factor x leads to an increase of incidence by a factor x^n.

Exercise:

Smokers have a 16 to 25 fold higher risk (I_s) to get lung cancer compared to non-smokers (I_{no}). The mutation rate of smokers is 3.23 times higher compared to non-smokers. Estimate, how many driver mutations are required to lead to lung cancer?

The correct answer to the exercise is $n \approx 3$. Hence, we can infer the number of driver mutations just from a fit of the data. The problem is however, that not all n necessarily have the same frequency, but follow a certain (unknown) distribution. We therefore measure only a weighted average

$$x^{\bar{n}} = \frac{\sum_{n=m}^{M} x^n \, I_n}{\sum_{n=m}^{M} I_n} \tag{5.68}$$

from the fold increase x^n and the corresponding $\bar{n}$. What we *want* to measure is however

$$\tilde{n} = \frac{\sum_{n=m}^{M} n \, I_n}{\sum_{n=m}^{M} I_n} \, . \tag{5.69}$$

Fortunately, one can show [4] that there is at least an estimate $m \leqslant \tilde{n} \leqslant \bar{n}$.
We so far accounted only for point mutations, but epigenetic mutations like gene fusion, translocation, amplification and methylation changes are disregarded yet. Let therefore i_p be the index of all mutually exclusive combinations of p (epi)genetic mutations and $\nu_k(i_p)$ the mutation rate of gene i in the p gene combination i_p, so that Equation 5.66 turns to

$$P(t) = \sum_{p,n} \sum_{j_n, i_p} \prod_{k=1}^{p} \nu_k(i_p) \nu^n \frac{t^{p+n-1}}{(p+n-1)!} \prod_{i=1}^{n} I_i(j_n) \, . \tag{5.70}$$

A further refinement is to account for the conditional nature of the probabilities $\nu_k(i_p)$, that however goes beyond the scope of the script. I therefore refer to [4].

References

[1] P. Armitage and R. Doll, *"The Age Distribution of Cancer and a Multi-stage Theory of Carcinogenesis"*, Br J Cancer. 1954 Mar; 8(1): 1-12.

[2] Radek Erban, Jonathan Chapman, Philip Maini, *"A practical guide to stochastic simulations of reaction-diffusion processes"*, arXiv:0704.1908 [q-bio.SC], 2007.

[3] Johan Paulsson, *"Models of stochastic gene expression"*, Physics of Life Reviews 2 (2005) 175-157.

[4] Cristian Tomasetti, Luigi Marchionni, Martin A. Nowak, Giovanni Parmigiani and Bert Vogelstein, *"Only three driver gene mutations are required for the development of lung and colorectal cancers"*, Proc Natl Acad Sci U S A. 2015 Jan 6; 112(1): 118-123.

[5] Bert Vogelstein, Nickolas Papadopoulos, Victor E. Velculescu, Shibin Zhou, Luis A. Diaz Jr., Kenneth W. Kinzler, *"Cancer Genome Landscapes"*, Science, vol 339, 2007.

6 Diffusion

Diffusion is one of the most interesting and fundamental processes for pattern formation in biological systems, but also plays a mayor role in other disciplines like physics, chemistry or economics. This phenomenon is understood as motion of particles along a concentration gradient, i. e. from a high concentration of molecules or atoms to lower concentrations. This coincides with our daily experience (milk clouds in coffee) and seems to be a natural behaviour. However the origin of the process is actually not immediately clear. How could the particles know in which direction they have to move in order to follow the gradient and how can pattern emerge if diffusion tends to even out gradients?
In 1827 Robert Brown[57] observed a random and complicated motion of particles in fluids and later Albert Einstein[58] described the trajectory of such a particle using the model of a random walk, caused by the collisions of these particles with the constituent molecules of the fluid. These investigations lead to a deeper understanding of diffusion as a random walk and connected the microscopic fluctuations to macroscopic properties like the diffusion constant. Along this line, the aim of this section is to describe the diffusion process as consequence of the random behaviour of the Poissonian stepper that directly guides to Ficks[59] laws and their generalized versions the Smoluchowski[60] equation and the Fokker - Planck[61] equation.

6.1 Fick's Second Law in 1D

In Section 5.1 I introduced the concept of the Poissonian stepper that only moves towards higher states, hence from state n to $n+1$ with the hopping rate ν. The hopping rate has the unit of probability per time. This model leads directly to the corresponding master equation. We now expand the concept and also allow a motion of the stepper along decreasing states, hence from n to $n-1$. Therefore, I like to denote the hopping rate in positive direction as ν_+ and in negative direction as ν_- (see Figure 77).

After taking all loss and gain terms for $\frac{dP(n,t)}{dt}$ into account, the corresponding master equation reads

$$\frac{dP(n,t)}{dt} = \nu_+ P(n-1,t) + \nu_- P(n+1,t) - (\nu_+ + \nu_-) P(n,t) \,. \tag{6.1}$$

Like in Section 5.1, we always choose a time interval dt that is sufficiently small to observe either one step or none. Therefore, the spatial increment Δn equals always one. Also the hopping rates are no functions of location, i. e. they are constant for every n.
Let us now consider the sub case $\nu_+ = \nu_- = \nu$ so that Equation 6.1 simplifies to

$$\frac{dP(n,t)}{dt} = \nu \left[P(n+1,t) + P(n-1,t) - 2\,P(n,t) \right] . \tag{6.2}$$

The structure on the rhs of this equation looks remarkably like the definition of the second derivative

$$\frac{d^2 f(x)}{dx^2} := \lim_{\Delta x \to 0} \frac{f(x+\Delta x) + f(x-\Delta x) - 2\,f(x)}{\Delta x^2}$$

[57] Robert Brown, 1773 - 1858
[58] Albert Einstein, 1879 - 1955
[59] Adolf Fick, 1829 - 1901
[60] Marian Smoluchowski, 1872 - 1917
[61] Adriaan Daniël Fokker, 1887 - 1972 and Max Karl Ernst Ludwig Planck, 1858 - 1947

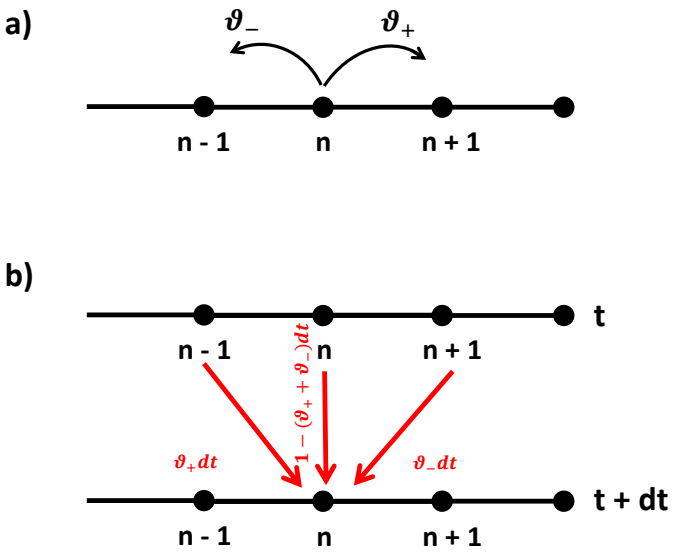

Figure 77: The Poisson stepper in one spatial direction with a forward hopping rate ν_+ and the backward hopping rate ν_- (c. f. Figure 68) along a sequence of states (upper panel) and with the time coordinate (lower panel).

from Section 2.1.3, where $P(n,t)$ equals $f(x)$. One difference is that the spatial increment Δx corresponds to the increment of the states $\Delta n = 1$ and therefore seems to be missed in the denominator of Equation 6.2, that is actually 1^2.
Another difference it that P is a function of n *and* t so that we have to use the partial derivative. Hence, for small Δn (that equals one) or for many possible states n, Equation 6.2 approaches the limit

$$\frac{\partial P(n,t)}{\partial t} = \nu \frac{\partial^2 P(n,t)}{\partial n^2} \,. \tag{6.3}$$

Equation 6.3 represents the macroscopic (or classical) limit of the master equation (Equation 6.2) and is therefore fully deterministic. But what is a possible application of Equation 6.3? One can identify the states n with locations x (like the kinesin molecule on the micro tubule in Section 5.1) of a particle, so that a motion in the states corresponds to a motion along a physical distances. Therefore, the location x is proportional to the state number n, hence $n \sim x$. The probability P equals the probability to find a particle at location x at time t that is then proportional to the concentration c, i. e. $c(x,t) \sim P(n \sim x, t)$. Then, ν is also a different constant, that is commonly denoted as D so that Equation 6.3 finally reads

$$\boxed{\frac{\partial c(x,t)}{\partial t} = D \frac{\partial^2 c(x,t)}{\partial x^2}} \,. \tag{6.4}$$

This equation is known as *Fick's second law of diffusion*. The constant D is called diffusion constant and has the unit length squared divided by time (e.g. m^2/s).
Equation 6.4 explains diffusion in one spatial direction in an isotropic homogeneous (D is not a function of time or location) environment. The solution of this equation can be derived by solving it in the inverse space after a Fourier transformation, that is subject to Section 7.1. I therefore give here the solution for a point source ($c(t=0) = c_0$ for $x = 0$ and $c(t=0) = 0$ elsewhere) without derivation:

$$c(x,t) = \frac{c_0}{\sqrt{4\pi D t}} \; exp\left[-\frac{1}{2} \cdot \frac{x^2}{2\,Dt}\right] \,. \tag{6.5}$$

The solution of Equation 6.4 is a Gaussian distribution (Section 2.6.7) with variance of $\sigma^2 = 2Dt$ and the mean at $\mu = 0$. When time elapses, the variance grows with t so that the Gaussian peak gets broader while the height of the peak decreases (division by the factor $\sqrt{t}$ in Equation 6.5), hence the particles migrate from higher concentration to lower concentration. According to Equation 6.5, a typical time scale τ of this process is

$$\boxed{\tau = \frac{x^2}{2D}} \tag{6.6}$$

with a diffusion length λ

$$\boxed{\lambda = \sqrt{2D\,\tau}}\,. \tag{6.7}$$

For $t \to \infty$, the peak is infinitely broad and flat; the concentration gradient has been evened out. The situation is illustrated in Figure 78.

For $t \to 0$, the peak becomes infinitely high and thin, that is an artifact of our simple

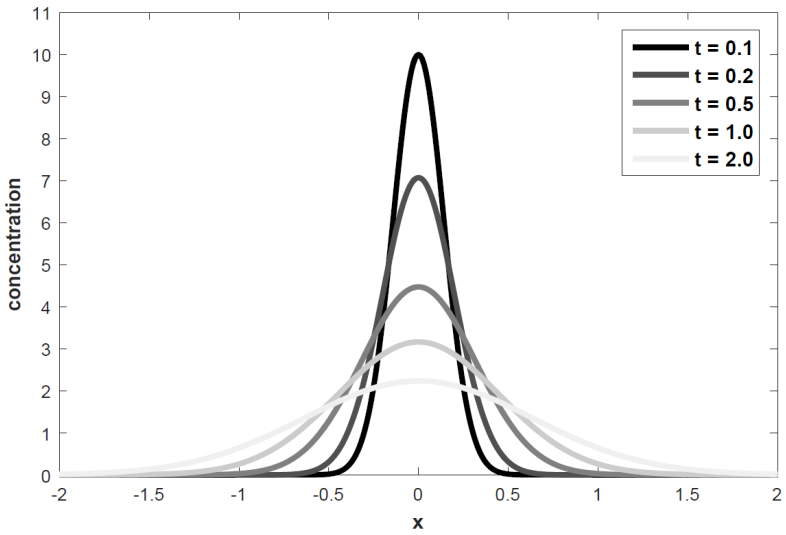

Figure 78: Visualization of Ficks second law in one dimension for different time steps (arbitrary units), see also Equation 6.4.

model (usually, the distribution $c(x, t = 0)$ has a finite limit).

6.2 Diffusion is a Random Walk

The behaviour of the solution of Equation 6.4 shown in Figure 78 seems to be intuitive, but how do the molecules in a diffusive process know in which direction they have to move in order to decrease the gradient? The answer is actually already given in the model of the Poissonian stepper we used to derive Fick's second law.
Suppose the Poissonian stepper performed N steps in total after a certain time span ΔT and that it has performed N_+ steps in positive direction and N_- steps in negative direction. Each time before the system performs step, the choice of the direction is random (left or right) and it is independent from the previous step. The probability to perform a step into positive direction after the time increment dt is $\nu_+ dt$ and $\nu_- dt$ for negative direction. Such an undirected behaviour is called a *random walk*.
We can now ask for the probability density distribution $P(N_+, N_-|N, \nu_+, \nu_-)$ that the system has performed N_+ steps in positive direction and $N_- = N - N_+$ in negative direction after it has performed N steps in total, given the hopping rates. Since the system can either jump to the left or a jump to the right[62], it is a binomial situation and therefore, the probability density distribution is binomial

$$P(N_+, N_-|N, \nu_+, \nu_-) = \frac{N!}{N_+!N_-!}\, \nu_+^{N_+}\, \nu_-^{N_-} \tag{6.8}$$

$$= \frac{N!}{(N-N_-)!N_-!}\, (1-\nu_-)^{(N-N_-)}\, \nu_-^{N_-}. \tag{6.9}$$

We are actually done now, but we want to explore the classical limit, hence the behaviour for $N \to \infty$. We learned in Section 2.6.7 that the limit of the binomial probability density distribution for large N is the Gaussian probability density distribution. To prove this we used Stirlings approximation (Equation 2.131) in Section 2.6.7. Applying this to Equation 6.9 leads to the approximation

$$\begin{aligned}\ln\left[P(N_+, N_-|N, \nu_+, \nu_-)\right] \approx N\ln N - N_-\ln N_- - (N-N_-)\ln(N-N_-) \\ +N_-\ln\nu_- + (N-N_-)\ln\nu_+.\end{aligned} \tag{6.10}$$

Like in Section 2.6.7, we like to approximate $P(N_+, N_-|N, \nu_+, \nu_-)$ around its maximum. For example, we might ask for the most likely value of N_- that I like to denote as $\bar{N}_-$. Thus, we calculate the derivative of $\ln P(N_+, N_-|N, \nu_+, \nu_-)$ wrt N_- and set it to zero

$$\frac{d\ln\left[P(N_+, N_-|N, \nu_+, \nu_-)\right]}{dN_-} = -\ln N_- - 1 + \ln(N-N_-) + 1 + \ln\nu_- - \ln\nu_+ = 0\,, \tag{6.11}$$

that leads to the condition

$$\bar{N}_- = \frac{\nu_- N}{\nu_- + \nu_+}\,. \tag{6.12}$$

Since the system has to perform a step in any direction, we know that $\nu_- + \nu_+ = 1$ and the maximum is located at $\bar{N}_- = \nu_- N$.
This is an intuitive result. If $\nu_- = \nu_+ = 0.5$ we expect that the system has performed more or less 50 steps in each direction if $N = 100$. For another run it might be $N_- = 40$ and $N_+ = 60$, but a result like $N_- = 1$ and $N_+ = 99$ should be rare (like flipping a coin). This means that the system usually does not change its total position $N_{tot} = N_+ - N_-$ a lot, if $\nu_- = \nu_+$ and that an inequality $\nu_- \neq \nu_+$ implies a drift or a net motion into the direction

[62] The system can also perform no step within dt, but this is not counted as "step". This is a slight, but important, difference to ν in Section 5.1.

of larger ν.
In the next step we approximate $P(N_+, N_-|N, \nu_+, \nu_-)$ around its maximum $P(\bar{N}) = P_{max}$ by a second order Taylor (Section 2.2.2) series as done for a related case in Section 2.6.8, that gives

$$\ln P(N_+, N_-|N, \nu_+, \nu_-) \approx \ln P_{max} + \left.\frac{\partial \ln P(N_+, N_-|N, \nu_+, \nu_-)}{\partial N_-}\right|_{N_-=\bar{N}_-} (N_- - \bar{N}_-) + \left.\frac{\partial^2 \ln P(N_+, N_-|N, \nu_+, \nu_-)}{\partial N_-^2}\right|_{N_-=\bar{N}_-} (N_- - \bar{N}_-)^2 . \tag{6.13}$$

The first derivative vanishes at $N_- = \bar{N}_-$ and P_{max} is a constant. The second derivative can be obtained from the derivative of Equation 6.11. Altogether, this leads to the expression

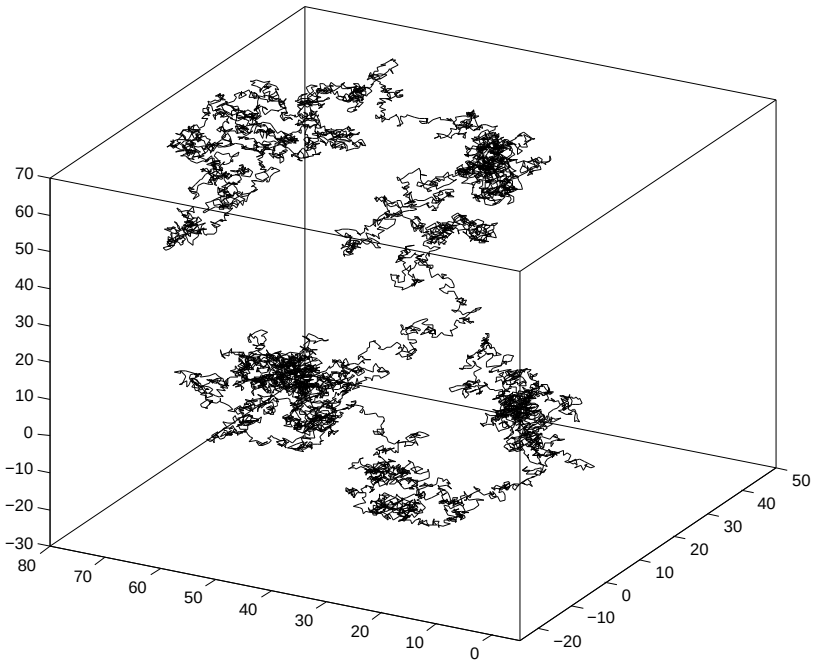

Figure 79: Simulated 3D random walk of 10 000 steps for a stained particle.

$$P(N_+, N_-|N, \nu_+, \nu_-) \approx \frac{1}{\sqrt{2\pi\nu_-(1-\nu_-)N}}\, exp\left[-\frac{1}{2}\cdot\frac{(N_- - \bar{N}_-)^2}{\nu_-(1-\nu_-)N}\right] . \tag{6.14}$$

The resemblance of Equation 6.14 to Equation 6.5 is remarkable. The variance now corresponds to $\sigma^2 = \nu_-(1-\nu_-)N$, that grows with increasing N. Since N, the number of steps, is proportional to time t, $\sigma \sim t$ like in Equation 6.5. The mean equals $\bar{N}_- = \nu_- N$ so that there is no net motion for $\nu_- = \nu_+$. Comparing Equation 6.14 to Equation 6.5 (that we derived for $\nu_- = \nu_+$), we find that

$$8\, Dt = N\,, \tag{6.15}$$

yielding a direct connection from the total number of steps of the Poissonian stepper to the measurable macroscopic quantity D. This connection and the identical structure of Equation 6.14 and Equation 6.5 illustrates that diffusion is caused by a random walk of the

particles. Therefore, diffusion from higher concentration to lower concentration appears just because of the random, undirected motion of the particles (the variance, hence the blur, is proportional to t or N). Due to the undirected motion, the random walk does not lead to a net motion of the Poissonian stepper if $\nu_- = \nu_+$. Therefore, a suitable distance measurement is the blur, the root mean square (RMS) $\sqrt{\nu_-(1-\nu_-)N}$.
A possible trajectory of a 3D random walk of a particle is illustrated in Figure 79.

6.2.1 The Biased Random Walk of E. Coli

An interesting application of the three dimensional random walk is the chemotaxis of E. Coli. When E. Coli searches for food (sugars, amino acids), it is interested in finding the location with the highest concentration. But how does it work?
E. Coli is too small to measure a concentration gradient directly. Therefore, it performs a random walk (called tumble) and measures the concentration every four seconds (that is its memory). After four seconds it compares the actual concentration to the one measured four seconds ago and therefore obtains a concentration gradient. Now, E. Coli moves in the direction of the gradient (called run) and repeats with the tumble phase. In this way, E. Coli approaches the source of food.

Since E. Coli has these two phases of locomotion, random walk and directed motion, it

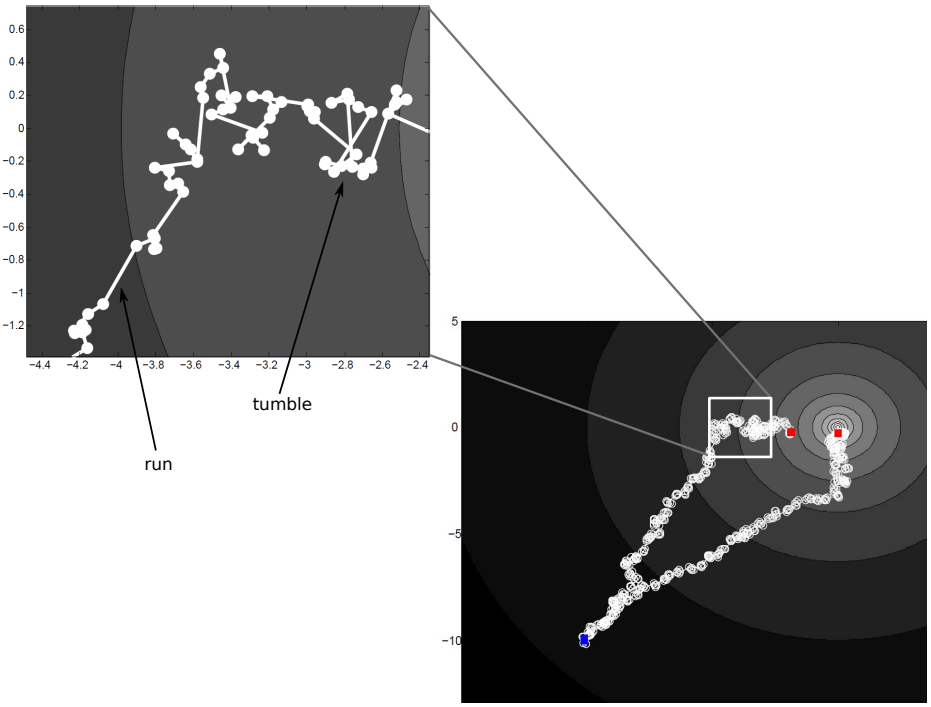

Figure 80: The chemotaxis of E. Coli can be simulated by a biased random walk. E. Coli measures the gradient of particular chemicals (here color coded; white corresponds to high concentration) after some random steps (tumbling motion), since it is too small to measure the gradient directly, and then moves towards the gradient (run phase) before entering a new tumbling phase.
The plot shows two random trajectories originating at identical starting point (blue).

is called a *biased random walk*, because it is biased by the concentration gradient (hence $\nu_- = \nu_+$ is **not** valid and the random walk **has** a direction). A simulation of E. Coli's chemotaxis is illustrated in Figure 80.

If simulating the chemotaxis of several thousand E. Coli (with identical starting point) and plotting a histogram of their locations for each time step I, we find, that the histogram

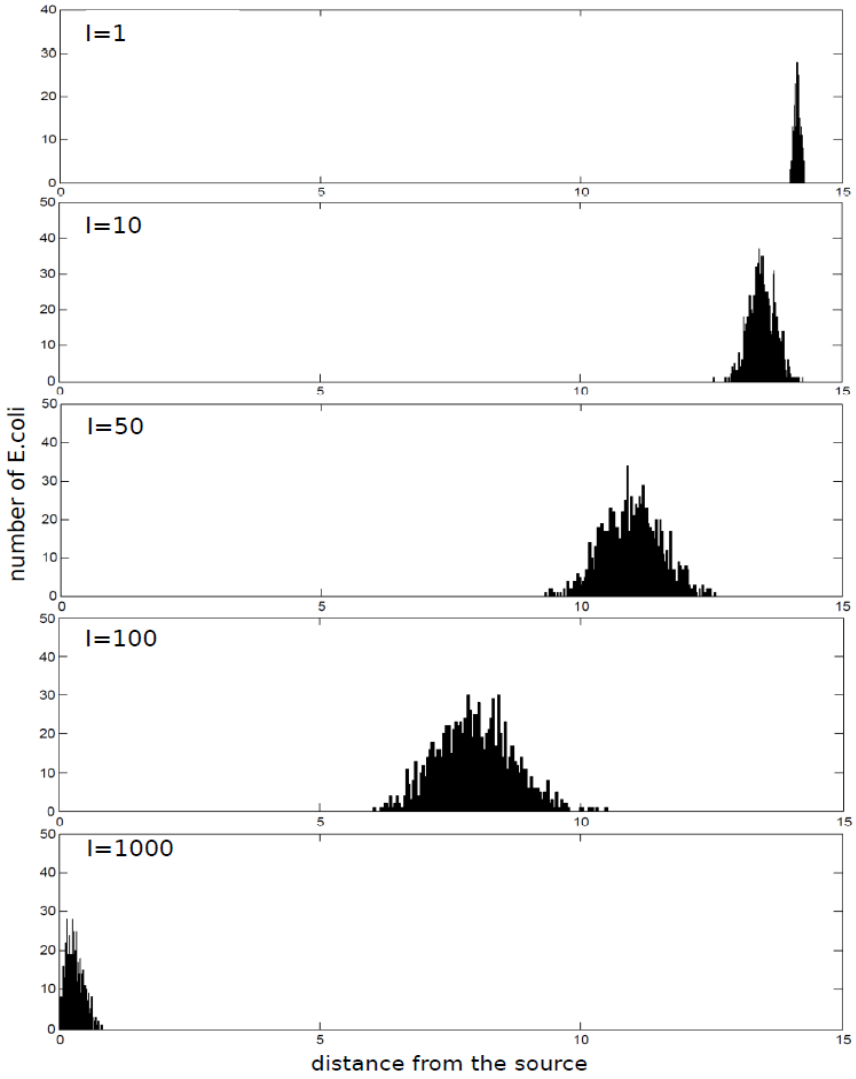

Figure 81: The simulated motion of 1000 E. Coli towards the gradient after I =1, 10, 50, 100 and 1000 time steps. The histogram smears out while the E. Coli approach their target. This is a common diffusive effect discussed in Section 6.4.

looks like a Gaussian (c. f. Equation 6.14 and Equation 6.5) that approaches the food source. While approaching, the Gaussian smears out (variance grows with time), that is a common diffusive phenomenon and will be discussed in Section 6.4. Such a simulation is illustrated in Figure 81.
Note, that the gradient is not always recognized if it is too flat. Some E. Coli will not be able to find the source and will perform a real random walk, while some still find it. This phenomenon is called *bifurcation* and occurs because not all E. Coli have a four second memory and not all have the same sensitivity for sugars and/or amino acids.

6.2.2 The Orientation of Macromolecules is a Random Walk

Here is a nice example of application of the concept of a random walk I found in [1].
Macromolecules are composed out of many identical segments. Due to their segments, the same kind of macromolecules can appear in different configurations. The effective length of a macromolecule (with the same number of segments) or its end-to-end distance is influenced by statistical fluctuations. Usually, a macromolecule does not appear to be

stretched out to its total length, but it is most likely somehow curled and bended, even if it is electrically neutral. Such a behaviour can be explained by a random walk model. Suppose we have a macromolecule of N segments where we denote the number of segments that are oriented to the right as n_R and those oriented to the left as n_L. The probability for a segment to be orientated to the left should be the same as for the right way orientation, hence 0.5 in both cases. This model is of course very simplified, because the segments can have any random orientation. But even then, the idea works, but is a bit more complicated. Therefore, I like to stick to the conceptual identical, but simplified version with only two directions. Also, this kind of model is identical to the Poissonian stepper in Section 6.1. The end-to-end distance R of the segments would be the difference between n_R and n_L, hence $R = (n_R - n_L)a = \Delta n a$, where a is the length of a segment. Note, that R is not the effective length, that would be the extension of the curled molecule. I show an example for a small macromolecule with three segments in Figure 82.

Having three segments, where each segment has two possibilities for its orientation, we

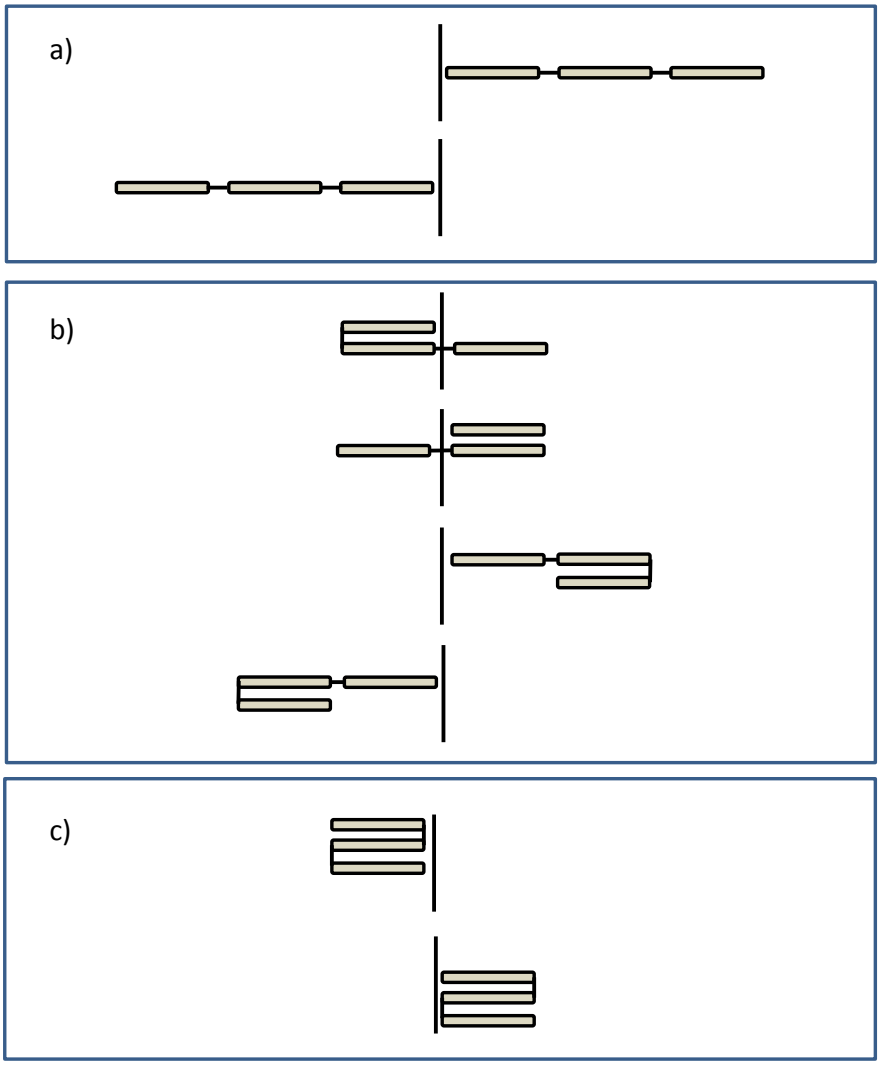

Figure 82: Different configurations of a "macromolecule" having three segments. See also the related figure in [1].

obtain $2^3 = 8$ total configurations for the molecule. For example all segments can be orientated to the right (panel a) in Figure 82) so that $R = 3a$ and also the effective length equals three segments. We can have the same situation for all segments being orientated

to the left. Thus, we have $\Omega = 2$ possibilities to find the molecule in a completely stretched state.
Another possibility is that two segments are orientated in one direction and one segment is orientate in the opposite direction, so that $R = a$. There are $\Omega = 6$ such configurations (panel b) and c) in Figure 82), where four of them have an effective length of two segments (panel b)) and two configurations have an effective length of one unit (panel c)). Although each orientation of each segment has the same prior probability (0.5), they produce outputs with different probabilities (end-to-end distance and effective length). Thus, we expect to find the macromolecule most likely in a somehow folded or curled state. Of course, a real macromolecule has $N = 100 \ldots 1000$ segments and it is not appropriate to count the possibilities by hand.
The probability to find a macromolecule of N segments with a certain n_R is (c. f. Equation 2.237 and Equation 6.9)

$$P\left(n_R, N\right) = \frac{N!}{n_R!\left(N - n_R\right)!}\left(\frac{1}{2}\right)^N. \tag{6.16}$$

Inserting $R = (n_R - n_L)a = \Delta n a$ into Equation 6.16 and using $N = n_R + n_L$, we obtain

$$P\left(n_R, N\right) = \frac{N!}{\left(\frac{N}{2} + \frac{R}{2a}\right)!\left(\frac{N}{2} - \frac{R}{2a}\right)!}\left(\frac{1}{2}\right)^N. \tag{6.17}$$

We now perform exactly the steps like in Section 6.2.1: we take $\ln P$, apply Stirling's approximation (that requires $N \gtrsim 100$) and derive the Taylor series of second order from this expression. The result is

$$\boxed{P\left(R, N\right) = \frac{2}{\sqrt{2\pi N}} exp\left(-\frac{R^2}{2Na^2}\right)}, \tag{6.18}$$

again a Gaussian probability density distribution.
Thus, most likely we will find the molecule with an end-to-end distance close to zero (Equation 6.18 peaks at $R = 0$), i. e. somehow curled or folded. The probability to find the molecule stretched completely ($R = Na$) is $P(Na, N) = 2/\sqrt{2\pi e N}$. This probability would be less than 5% for a macromolecule with 100 segments.

6.2.3 Fick's First Law

Fick's first law essentially states, that particles always move from a region of higher concentration to regions of lower concentration. What seems to be caused by a directed process can also be understood by an undirected random walk of particles.
Suppose, we have two boxes: one with a higher concentration (on the left) and one with a lower concentration (on the right) of particles and if we join them and remove the wall between the boxes we have a situation like illustrated in Figure 83. Each particle moves in 3D space, hence, performs motions in six different directions. All these directions are independent from each other. Let us denote the number of particles that can cross the dashed line between the boxes from left to right in the next motion with k_L and those that cross the line from right to left with k_R. In total, the left box hosts N_L particles and the right box hosts N_R particles. All particles move completely random and uncorrelated, so that in both boxes the probability p for a particular particle to cross the dashed line and the probability $1 - p$ to not cross the dashed line is equal.

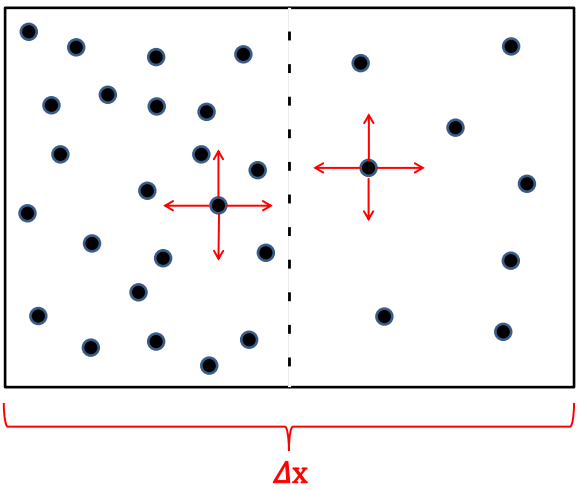

Figure 83: Diffusion between two regions of different concentration.

According to Equation 2.237, the probability that k_L particles from N_L particles in the left have crossed the dashed line within a given time t equals

$$P_L(k_L|N_L) = \frac{N_L!}{k_L!\,(N_L - k_L)!} p^{k_L}(1-p)^{N_L - k_L} \tag{6.19}$$

and the probability that k_R particles cross the dashed line from the opposite direction is

$$P_R(k_R|N_R) = \frac{N_R!}{k_R!\,(N_R - k_R)!} p^{k_R}(1-p)^{N_R - k_R}\,. \tag{6.20}$$

The means of the particle flows (Equation 2.241) are $\langle k_L \rangle = N_L\,p$ and $\langle k_R \rangle = N_R\,p$. We would interpret the difference of the two flows as a net flux $\vec{\Xi} = \langle k_L \rangle - \langle k_R \rangle = p\,(N_L - N_R) = p\,\Delta N$. Hence, the net flow from higher concentrations to lower concentrations occurs only because it is more likely for a particle to be located in the high concentration part and thus more likely to cross the dashed line from this side.
The volume of both boxes together is $V = A\Delta x$ and the concentration of all particles is $c = \frac{N}{V} = \frac{N}{A\Delta x}$. This concentration changes according to $\Delta c = \frac{\Delta N}{V} = \frac{\Delta N}{A\Delta x}$. Expressing the net flux in terms of the concentration we find

$$\vec{\Xi} = \langle k_L \rangle - \langle k_R \rangle = p\,A\,\Delta x\,(-\Delta c)\,. \tag{6.21}$$

It is minus Δc because the net flux is **positive** from left to right, if the gradient of the concentration ($\Delta c = c_R - c_L$) is **negative** (it points from right to left).
Now we define a flux density $\vec{\phi}$ that is flux per area per time (actually a current density, see Section 2.1.6), hence $\vec{\phi} = \frac{\Delta\vec{\Xi}}{A\Delta t}$. In this way we eliminate A and find that the flux density is

$$\vec{\phi} = \frac{\langle k_L \rangle - \langle k_R \rangle}{A\Delta t} = -p\frac{\Delta c \Delta x}{\Delta t} \tag{6.22}$$

We see that $\frac{\Delta x}{\Delta t}$ is a velocity, the flow velocity. We could also interpret the term $\frac{\Delta c}{\Delta t}$ as change of the concentration wrt time and dividing Equation 6.22 by Δx would lead to the structure of the continuity equation (Equation 2.65). Hence, apart from a proportionality constant, we derived the same equation from different approaches: conservation of mass in the case of the continuity equation and the statistical motion of particles in this approach.

However, Equation 6.22 is better known in a different arrangement. If we multiply the rhs of Equation 6.22 with $\frac{\Delta x}{\Delta x}$ we obtain

$$\vec{\phi} = -p\frac{\Delta x^2 \Delta c}{\Delta t \Delta x}. \tag{6.23}$$

We rename $p\frac{\Delta x^2}{\Delta t} = D$ and for infinitely small spatial steps Δx, Equation 6.23 turns into (generalizing for 3D)

$$\boxed{\vec{\phi} = -D\, grad\, c}, \tag{6.24}$$

that is *Fick's first law*. Fick's first law states that there is a flux per area and time if there is a concentration gradient. Since the sign is negative, the flux works against the gradient in order to equalize it. The proportionality constant D is the diffusion constant we know already that has the unit length squared per time. This makes sense because it states the penetration of a given area per time due to particles. That D is the actual diffusion constant we know from Section 6.1 can be shown be inserting Fick's first law into the continuity equation (Section 2.1.6)

$$\frac{\partial c}{\partial t} = -div\, \vec{\phi} = -div\, (-D\, grad\, c) = D\, \Delta c. \tag{6.25}$$

for constant D. Equation 6.25 is just Fick's second law (Section 6.1). Note, that D is actually not a constant, since it depends on temperature and may also depend on the spatial coordinates in an inhomogeneous environment (see Section 6.5).
Fick's first law is not complete, since the gradient will be depleted if the flux goes on. But if the gradient becomes flat, also the flux will decrease. Thus, Fick's first law is only applicable for a steady state situation and therefore Fick's second law is required to describe the process.

References

[1] Rob Phillips, Jane Kondev, Julie Theriot *"Physical Biology of the Cell"*, Garland Science, New York, USA, 2009.

6.3 Fick's Second Law in Higher Dimensions

We found in the previous sections that the concept of the Poissonian stepper and the random walk is very fruitful and that it can be used to understand different phenomena. I therefore like to discuss Fick's second law further by expanding it for two and three spatial dimensions. The Poissonian stepper from Section 5.1 is now able to move also in a further direction m, that is perpendicular to n. Every state at given time t is now characterized by the set (n, m) and we ask for the probability $P(n, m, t)$ to find the system in the particular state at time t. As before, the hopping rates in n direction are denoted as ν_- and ν_+. In the same manner, I denote the corresponding hopping rates in m direction as μ_- and μ_+. Again, we choose dt it that way, that the system can only change one state parameter or none, e. g. it can never jump from (n, m) to $(n+1, m+1)$ directly within one time step, but would have to move from (n, m) to $(n, m+1)$ first and then to $(n+1, m+1)$. This two dimensional grid of states is shown in Figure 84.

We write down the master equation by considering all loss and gain terms and obtain

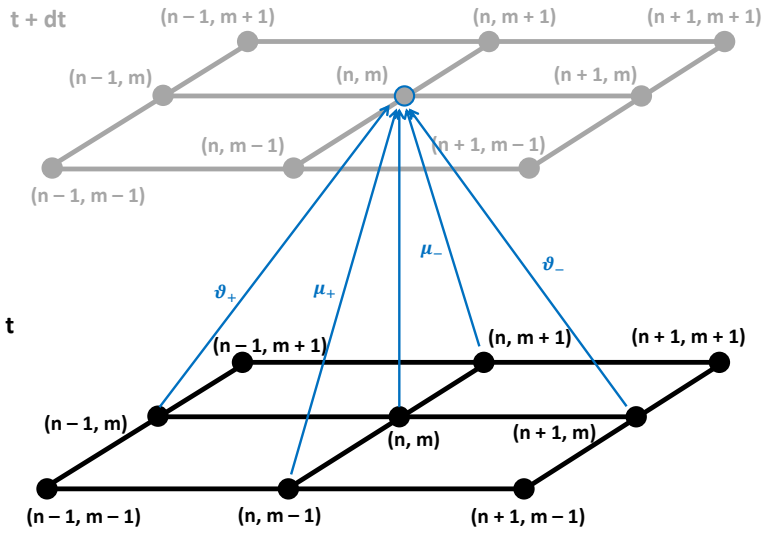

Figure 84: Visualization of the concept of the Poissonian stepper in 2D (see also Section 5.1).

$$\begin{aligned}\frac{dP(n,m,t)}{dt} = \quad & \mu_+ P(n,m-1,t) - \mu_+ P(n,m,t) \\ & + \mu_- P(n,m+1,t) - \mu_- P(n,m,t) \\ & + \nu_+ P(n-1,m,t) - \nu_+ P(n,m,t) \\ & + \nu_- P(n+1,m,t) - \nu_- P(n,m,t)\,.\end{aligned} \tag{6.26}$$

In the case of an undirected random walk, $\mu_+ = \mu_- = \mu$ and $\nu_+ = \nu_- = \nu$, but not necessarily $\nu = \mu$, since the diffusion along the different states might be different. Hence, diffusion is homogeneous, but not isotropic. In such a case we can simplify Equation 6.26 and derive

$$\begin{aligned}\frac{dP(n,m,t)}{dt} = \quad & \mu\,[P(n,m-1,t) + P(n,m+1,t) - 2\,P(n,m,t)] \\ & + \nu\,[P(n-1,m,t) + P(n+1,m,t) - 2\,P(n,m,t)]\,.\end{aligned} \tag{6.27}$$

Analogous to Section 5.1 we can identify the n direction with the physical spatial direction x and the m direction with y with the increments $\Delta n = 1$ and $\Delta m = 1$, respectively. The two addends in Equation 6.27 equal the definition of the second derivative (Equation 2.49) of a function $f(x, y)$ and we therefore can approximate Equation 6.27 to

$$\frac{\partial}{\partial t}P(x,y,t) = \mu\frac{\partial^2}{\partial y^2}P(x,y,t) + \nu\frac{\partial^2}{\partial x^2}P(x,y,t)\,, \tag{6.28}$$

which, due to proportionality of the concentration c to probability P leads to

$$\frac{\partial}{\partial t}c(x,y,t) = D_y\frac{\partial^2}{\partial y^2}c(x,y,t) + D_x\frac{\partial^2}{\partial x^2}c(x,y,t)\,. \tag{6.29}$$

In a similar way, we would find the equation for $P(x, y, z, t)$ in three dimensions and an additional term $\frac{\partial^2}{\partial z^2}P(x, y, z, t)$ would appear. The structure of Equation 6.28 would always be the same for any number of dimensions. For our purposes, the 3D case is fully sufficient. If diffusion is isotropic, the diffusion constants in all directions are identical and we can write $D_x = D_y = D_z = D$. Using the common definition of the sum $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} =: \Delta$ we can write for the 3D case of homogeneous, isotropic diffusion

$$\boxed{\frac{\partial}{\partial t}c(x,y,z,t) = D\;\Delta c(x,y,z,t)}\,. \tag{6.30}$$

The operator Δ (not to change with the finite difference Δ) is called *Laplace operator*[63]. It just is a shortcut for the recipe to add all the second spatial derivatives. There is the connection $\Delta = div\,grad$ (Section 2.1.6) that can be proofed by executing the operations explicitly. Equation 6.30 is Fick's second law in three dimensions.

The solution of Equation 6.30 for a point source can be derived by solving the master

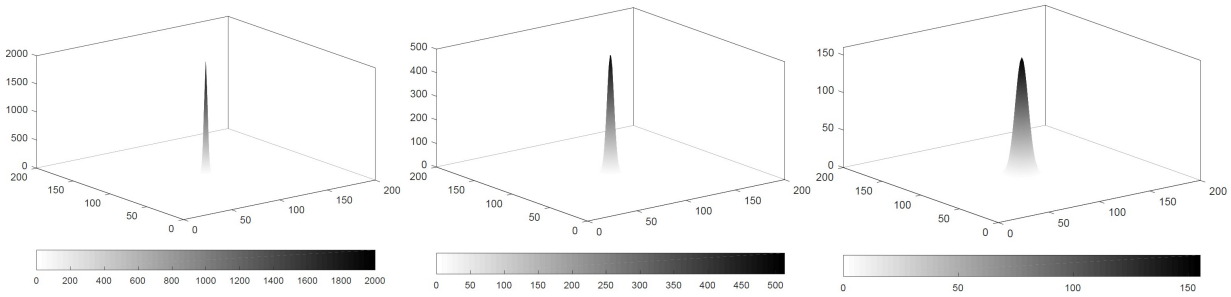

Figure 85: Temporal evolution of Equation 6.31 (in two dimensions) for three different time steps. Note the increasing broadness of the peak and the scale of the z-axis.

equation (Equation 6.26) in three dimensions and calculating the means (as done in Section 5.2.1) that leads to very extensive algebra, or by using a Fourier transformation (that is shown in Section 7.1). Therefore, I like to give the solution, that is similar to Equation 6.5 and reads

$$c(x,y,z,t) = \frac{c_0}{(4\pi Dt)^{3/2}}\;exp\left[-\frac{1}{2}\cdot\frac{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}{2\,Dt}\right]\,. \tag{6.31}$$

The concentration has its largest value at $(x_0,\, y_0,\, z_0)$ and the variance, hence the broadness, of the peak increases with time t while the height of the peak decreases. The temporal behaviour of Equation 6.31 is shown in Figure 85 (in two dimensions only for the sake of better visibility).

[63] Pierre-Simon, marquis de Laplace, 1749 - 1827

6.3.1 Diffusion Into a Cell

Let us discuss an example of a simple diffusion process (see also [1]). Imagine a spherical cell with radius a that passes a chemical through its membrane by diffusion. This chemical is depleted in the interior of the cell so that the concentration $c = 0$ for $r \leq a$. Thus, there is a constant concentration gradient from the inner part of the cell to its environment. The equilibrium concentration far away from the cell is denoted at c_0. We now ask for the concentration profile around the cell and the entire flux into the cell.
After a certain time, the system would have reached steady state, assuming, that the cell continuously consumes the chemical. The diffusion process is described by Equation 6.30, where the rhs equals zero in steady state and we obtain

$$D\,\Delta c = 0\,. \tag{6.32}$$

We could solve Equation 6.32 directly, but we can also use the spherical geometry of the cell and therefore use the Laplace operator Δ in spherical coordinates (Section 2.5). If the cell is a perfect sphere in first approximation the diffusion occurs radially symmetric and all derivatives in the spherical Laplace operator disappear, except those wrt r. This would be not the case for Cartesian coordinates and the derivation of the solution would be complicated. Thus, for spherical coordinates Equation 6.32 can be expressed as

$$D\,\Delta c = D\left[\frac{\partial^2 c}{\partial r^2} + \frac{2}{r}\frac{\partial c}{\partial r}\right] = D\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c}{\partial r}\right)\right] = 0\,. \tag{6.33}$$

Since c now only depends on t, the partial derivatives can be treated as total derivatives and we can integrate both sides of the equation. First integration gives

$$A_1 = r^2\frac{\partial c}{\partial r} \tag{6.34}$$

and integrating again leads to

$$A_2 - \frac{A_1}{r} = c(r) \tag{6.35}$$

with two yet unknown constants A_1 and A_2.
If $c = c_0$ far from the interior of the cell then $c(r = \infty) = c_0$ that gives $A_2 = c_0$. If the chemical is depleted inside the cell, then $c(r \leq a) = 0$ that yields $A_1 = a\,c_0$. With these constants, we obtain the concentration profile

$$\boxed{c(r) = c_0\left(1 - \frac{a}{r}\right)}\,. \tag{6.36}$$

Note, that solving this problem became fairly easy since we used the Laplace operator in spherical coordinates. Guess how inconvenient it would have been if we would have taken its Cartesian version. The concentration profile $c(r)$ is shown in Figure 86.

Having, c, we now can calculate the total flux into the cell. The flux density (flux per area and time) is given by Fick's first law (Section 6.2.3). Since we know c as function of r (Equation 6.36) we can now calculate its gradient (again in spherical coordinates, see Section 2.5):

$$grad\,c(r) = a\frac{c_0}{r^2}\,. \tag{6.37}$$

Thus, according to Section 2.1.6, the flux density is

$$\vec{\phi} = Da\frac{c_0}{r^2} \tag{6.38}$$

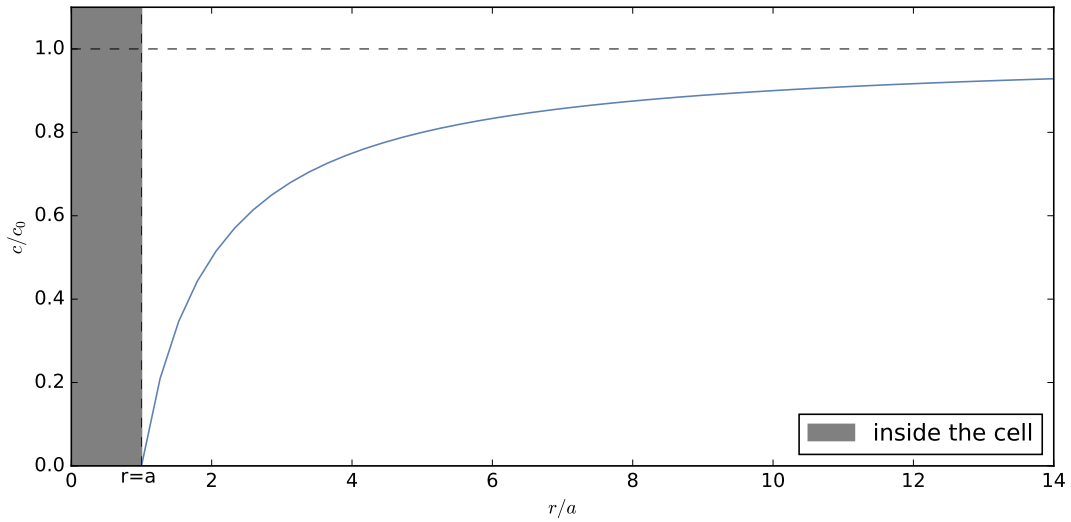

Figure 86: Concentration profile (blue solid line) at a spherically symmetric cell of radius a caused by diffusion. The equilibrium concentration far from the cell is denoted as c_0 (horizontal dashed line).

and the total flux (integrated over the entire surface $dA = r^2 \sin\theta d\phi d\theta$ of the cell) is (see end of Section 2.5)

$$\int \vec{\phi}\, d\vec{A} = \int_{-\pi}^{\pi} \int_{0}^{2\pi} Da \frac{c_0}{r^2} r^2 \sin\theta \, d\phi \, d\theta = \boxed{4\pi a D c_0} \tag{6.39}$$

Note, that the entire flux is independent of r, since mass (concentration times volume) is conserved. The factor $k = 4\pi a D$ is called *rate coefficient* and is usually measured in $\frac{l}{mol \cdot s}$.

References

[1] Rob Phillips, Jane Kondev, Julie Theriot *"Physical Biology of the Cell"*, Garland Science, New York, USA, 2009.

6.4 Diffusion with Drift: The Smoluchowski Equation

When introducing the 1D Poissonian stepper for deriving Fick's second law in 1D (Section 6.1), we derived the master equation

$$\frac{dP(n,t)}{dt} = \nu_+ P(n-1,t) + \nu_- P(n+1,t) - [\nu_+ + \nu_-]\ P(n,t)\,,$$

and set $\nu_+ = \nu_-$ to derive the actual diffusion equation. It turned out that there is no net motion of the stepper in such case, that became more clear when interpreting the process as random walk. We now relax this condition and keep $\nu_+ \neq \nu_-$. By applying a trick by including the zeros $\frac{1}{2}\nu_+ P(n-1,t) - \frac{1}{2}\nu_+ P(n-1,t) = 0$ and $\frac{1}{2}\nu_- P(n-1,t) - \frac{1}{2}\nu_- P(n-1,t) = 0$ we can write the master equation after some rearranging as

$$\begin{aligned}\frac{dP(n,t)}{dt} = &-\frac{\nu_+ - \nu_-}{2}\left[P(n+1,t) - P(n-1,t)\right] \\ &+ \frac{\nu_+ + \nu_-}{2}\left[P(n+1,t) + P(n-1,t) - 2P(n,t)\right].\end{aligned} \tag{6.40}$$

The purpose of this exercise is to make it easier to infer the structure of the equation. The second part on the rhs looks familiar since it is again the second derivative (Equation 2.40), but now the prefactor equals the average of both hopping rates. The first addend on the rhs equals the definition of the first derivative (Equation 2.33). The prefactor here is a drift term. If ν_+ is larger than ν_- (recall that both values are probabilities and therefore always positive), the system will in average most of the time move to larger n and if ν_- is larger, it will move to smaller n. Considering the structure of the derivatives, the limit of the master equation is

$$\frac{\partial P(n,t)}{\partial t} = -(\nu_+ - \nu_-)\frac{\partial}{\partial n}P(n,t) + \frac{(\nu_+ + \nu_-)}{2}\frac{\partial^2}{\partial n^2}P(n,t). \tag{6.41}$$

Like in the previous sections, we now link these probabilities to particle concentrations c and introduce some proportionality factors and obtain

$$\frac{\partial}{\partial t}c(x,t) = -v_x\frac{\partial}{\partial x}c(x,t) + D\frac{\partial^2}{\partial x^2}c(x,t)\,. \tag{6.42}$$

Let us discuss the units of these prefactors in order to understand what they really are. On the lhs of Equation 6.42 we have the units concentration per time and on the rhs in the first addend we find the spatial derivative of the concentration that has the units concentration per length. Hence, in order to keep the units consistent, the unit of the prefactor v_x must be $\frac{\text{length}}{\text{time}}$, i. e. indeed a (drift) velocity, meaning that an object performing this random walk attains a (macroscopic) net flow velocity in one direction.
The second addend on the rhs of Equation 6.42 contains the second spatial derivative of c so that the unit equals concentration per length squared. Thus, the unit of the prefactor D must be $\frac{\text{length}^2}{\text{time}}$. Hence, D is indeed a diffusion constant.
For a 3D system, we would set up the master equation in the same manner and the equivalent of Equation 6.42 is

$$\boxed{\frac{\partial}{\partial t}c(x,y,z,t) = -\vec{v}\cdot grad\ c(x,y,z,t) + D\,\Delta c(x,y,z,t)}\,, \tag{6.43}$$

the *Smoluchowski equation*.
Like in Section 6.1, I like to give the solution of this equation for a point source with

concentration c_0 that equals

$$c(x,y,z,t) = \frac{c_0}{(4\pi Dt)^{3/2}}\ exp\left[-\frac{1}{2}\cdot\frac{(x-v_x\ t)^2+(y-v_y\ t)^2+(z-v_z\ t)^2}{2Dt}\right]. \quad (6.44)$$

This equation is almost identical to Equation 6.31 with the difference, that the coordinates of the maximum $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ change with time by $\begin{pmatrix} v_x t \\ v_y t \\ v_z t \end{pmatrix}$. Hence, while the peak becomes broader by $2\,Dt$ and lower by $(4\pi\ Dt)^{3/2}$ while time elapses, it also moves along the velocity vector $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$. The dynamics of the solution is illustrated in Figure 87 in two dimensions for the sake of visibility. Since we have a drift term, the behaviour of the solution is a result

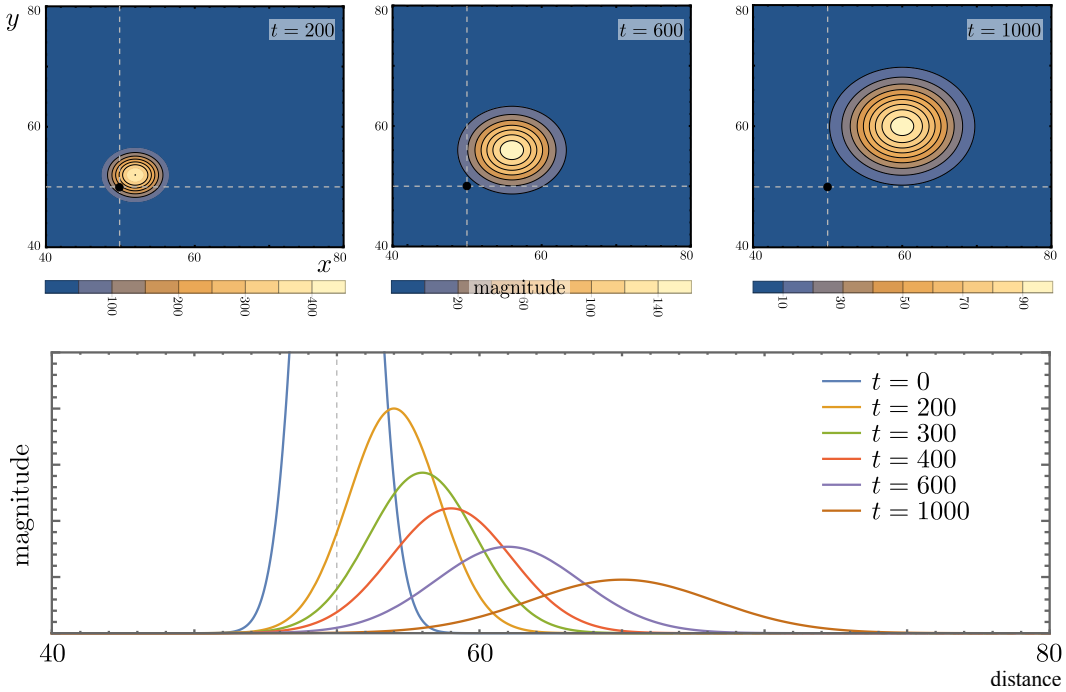

Figure 87: Behaviour of the (2D) solution of the Smoluchowski equation after 200, 600 and 1000 time steps (upper panels). The velocity points to the upper right and the black dot indicates the initial position of the distribution. One can see, that the source gets blurred while moving along the diagonal ($v_x = v_y$ here).
The concentration profile from the solution of the Smoluchowski equation (1D) for different times t is shown in the lower panel. Units are arbitrary.

of the superposition of the diffusion itself and the drift velocity. Thus, while the particles are performing a net motion into a particular direction, they constantly undergo diffusion.

6.4.1 The Gradient in a Test Tube and the Smoluchowski - Einstein Relation

Imagine you put a well mixed solution of molecules into a test tube. After some time this solution separates into different layers with increasing concentration towards the bottom of the test tube until the mixture reaches an equilibrium. At equilibrium, the formed concentration gradient is stable so that the system reaches steady state. Still, the particles move and undergo diffusion, but there is no net flux between the layers so that we can set

$\frac{\partial c}{\partial t} = 0$ (steady state) in Equation 6.42. Diffusion occurs only along the test tube that is a 1D scenario in first approximation, so that Equation 6.42 turns into

$$D\frac{\partial c(x)}{\partial x} = v\, c(x)\,, \tag{6.45}$$

with x being the height of a particle in the test tube.
The significant force which leads to the formation of the concentration gradient is gravity F_G that works against the drag force F_D in the solution. It can be shown, that the drag force is $F_D = \xi\, v$ (where v is the velocity of the particle and ξ is a proportionality constant, the drag coefficient) if there is no turbulence in the medium. Hence, the process has reached equilibrium if $F_G = F_D$ and therefore we can express the flow velocity for each of the particles in a test tube as $v = \frac{F_G}{\xi}$ and Equation 6.45 turns into

$$D\frac{\partial c(x)}{\partial x} = c(x)\frac{F_G}{\xi}\,. \tag{6.46}$$

Since Equation 6.46 depends only on one variable, x, we can write the total derivative and therefore use the separation of the variables and find that

$$\frac{dc(x)}{c(x)} = \frac{F_G}{D\,\xi}dx. \tag{6.47}$$

While integrating both sides of Equation 6.47, we have to perform the integral $\int F_G\, dx$, which represents (by definition) the work $W(x)$ done against the friction during the motion of the particles. Therefore, the solution of Equation 6.47 is

$$c\,(x) = c_0\; exp\left[-\frac{W(x)}{D\xi}\right] \tag{6.48}$$

that is the concentration profile in the test tube in equilibrium.
The concentration profile looks suspiciously like the Boltzmann distribution (Equation 3.28) and indeed: the exponent in Equation 6.48 contains the ratio between two energies, the potential energy of the particle represented by work $W(x)$ and $D\xi$. The motion of the particles is caused by thermal energy, so that according to Section 3.2 the Boltzmann equation ($\sim exp\left[\frac{\epsilon}{kT}\right]$) applies for equilibrium and we therefore find the relation

$$D\xi = kT \quad \Rightarrow \quad \boxed{D = \frac{kT}{\xi}}, \tag{6.49}$$

that is the famous *Einstein-Smoluchowski relation*.
The proportionality factor ξ depends on the size and shape of the molecules. For example one can show that $\xi = 6\pi\eta r$ for a spherical body of radius r suspended in a liquid with viscosity η (*Stokes relation*, Section 8). The importance of the Einstein - Smoluchowski relation is that it connects the size r of a particle (microscopic) to the (macroscopic) diffusion constant that is easy to measure. Hence, for the first time in history it was possible to measure the sizes of molecules and atoms from their measurable diffusive behaviour in a solution. It thereby also proved that Brownian motion is indeed a random walk (c. f. Section 6.2).
Combining the Einstein-Smoluchowski relation to the findings in Section 6.1 we can link the diffusion time τ (Equation 6.6) to particle size r and temperature T and obtain

$$\boxed{\tau = 3\frac{\pi\eta\, r}{kT}x^2} \tag{6.50}$$

and the diffusion length λ (Equation 6.7) equals

$$\boxed{\lambda = \sqrt{\frac{1}{3}\frac{kT}{\pi\eta\, r}\,\tau}}\,. \tag{6.51}$$

6.5 Fokker-Planck Equation

The last subsection in this section is devoted to the *Fokker - Planck equation*. In the previous section, we relaxed the condition, that the hopping rates a equal for any direction by setting $\nu_+ \neq \nu_-$. If μ is the hopping rate in y direction and λ the hopping rate in z direction, we can always set $\mu_- \neq \mu_+$ and $\lambda_- \neq \lambda_+$ that leads to the drift vector $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ and the three dimensional Smoluchowski equation (Equation 6.44).
If we also allow $\nu_\pm \neq \mu_\pm \neq \lambda_\pm$, the diffusion constant D will be different for any spatial direction, but still constant **along** the particular direction. If we now allow the hopping rates to be a function of location $\nu_\pm(x,y,z) \neq \mu_\pm(x,y,z) \neq \lambda_\pm(x,y,z)$, both the diffusion "constant" and the drift velocity will not only be different for any spatial direction, but also different **along/at** any direction/location. This is the most general case of diffusion.
If the location of the particle is described by the vector $\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and if the drift velocity $\vec{v}$ is a function of location, we can write $\vec{v}(\vec{x})$. If diffusion depends on the direction (three coordinates x, y and z) **and** location (again three coordinates), it is not a vector, but a 3×3 matrix (often denoted as σ).
The derivation of the master equation in such a case works in principle as for ordinary diffusion (Fick's second law, Section 6.1) and diffusion with drift (Smoluchowski equation, Section 6.4), but leads to very intensive algebra. I therefore recommend further literature for the devoted students. The resulting equation in 3D is

$$\frac{\partial P(\vec{x},t)}{\partial t} = -\, div\left[\vec{v}(\vec{x},t)P(\vec{x},t)\right] + \frac{1}{2}\Delta\left[\sigma^2(\vec{x},t)P(\vec{x},t)\right] , \tag{6.52}$$

where the factor $1/2$ in front of the diffusion term (c. f. the second addend on the rhs of Equation 6.41) is usually not included in the diffusion expression.
Equation 6.52 can be generalized even further for the N dimensional case that leads to the actual Fokker - Planck equation

$$\boxed{\frac{\partial P(\vec{x},t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left[v_i(\vec{x},t)P(\vec{x},t)\right] + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j}\left[D_{ij}(\vec{x},t)P(\vec{x},t)\right]}, \tag{6.53}$$

where $\vec{x}$ and $\vec{v}(\vec{x})$ are $N-$dimensional vectors for location and drift velocity, respectively, and $D_{ij}(\vec{x},t)$ is a so-called diffusion tensor given by $D_{ij} = \frac{1}{2}\sum_{k=1}^{N} \sigma_{ik}(\vec{x},t)\sigma_{jk}(\vec{x},t)$, and $\sigma(\vec{x},t)$ is a $N \times N$ matrix. Equation 6.53 describes any diffusion process one can think about and appears very often in physics. For example the Fokker-Planck equation finds its application in quantum physics as the *Schrödinger*[64] *equation*

$$i\hbar\frac{\partial \Psi(\vec{x},t)}{\partial t} = \left[\frac{-\hbar^2}{2m}\Delta + V(\vec{x},t)\right]\Psi(\vec{x},t)\,.$$

The probability $P(\vec{x},t)$ is substituted by a (complex) wave function $\Psi(\vec{x},t)$, where $|\Psi(\vec{x},t)|^2$ gives indeed the probability density function for a particle of mass m at location $\vec{x}$ at time t. The expression $V(\vec{x},t)$ is a potential energy, for example the Coulomb[65] energy an electron

[64] Erwin Rudolf Josef Alexander Schödinger, 1887 - 1961
[65] Charles Augustin de Coulomb, 1736 - 1806

feels in an atom. Hence, up to some constants, the Fokker - Planck and the Schrödinger equation give an analogous description of the evolution of the probability density in time. The only difference is the diffusion constant $D = \frac{i\hbar}{2m}$.

For our purposes the 3D case is fully sufficient. For example in three dimensions, if $\sigma^2(\vec{x},t) \equiv 2D(\vec{x},t) = const$ we obtain

$$\begin{aligned} \frac{\partial P(\vec{x},t)}{\partial t} &= -\, div\,[\vec{v}(\vec{x},t)P(\vec{x},t)] + D\Delta P(\vec{x},t) = \\ &= -\, P(\vec{x},t) div\,[\vec{v}(\vec{x},t)] - \vec{v}(\vec{x},t)\, grad\,[P(\vec{x},t)] + D\,\Delta P(\vec{x},t)\,. \end{aligned} \tag{6.54}$$

One can show that $div\,[\vec{v}(\vec{x},t)] = 0$ for an incompressible fluid (Section 8), that is a good approximation for biological systems, in which the relevant range of pressure is such that the compressibility of water is negligible. Thus, the Fokker - Planck equation turns into its sub-case

$$\frac{\partial P(\vec{x},t)}{\partial t} = -\vec{v}(\vec{x},t)\, grad\,[P(\vec{x},t)] + D\,\Delta P(\vec{x},t)$$

the Smoluchowski equation. For $\vec{v}(\vec{x},t) = 0$ (no drift) we obtain just Fick's second law $\frac{\partial P(\vec{x},t)}{\partial t} = D\,\Delta P(\vec{x},t)$.

7 Diffusion Reaction and Pattern Formation

In Section 5 and Section 6 we acquired many theoretical tools that are all actually a preparation for answering some important questions we will address in this section now. These questions are for example: How does an undifferentiated embryo evolve to a highly organized organism with different cell types of completely different functions and how do these cells "know" into what they have to evolve? The process of differentiation is called *morphogenesis* and biological systems face the problem of forming stable concentration gradients against the omnipresent process of diffusion for a given time span in order to form stable patterns. These gradients can be spatial concentration gradients of certain chemicals that inhibit or activate the expression of some genes, which in turn, may guide the process of differentiation. The underlying mechanisms in this regard were completely unknown until mid of the 20^{th} century and only in the last few decades some light was shed on the physical laws of morphogenesis.
It turned out that all these processes are guided by diffusion reactions, i. e. diffusion of particles in a certain balance with producing and/or consuming certain chemical compounds. The mathematical structure of the underlying equations always equals the scheme

$$\frac{\partial c(x,y,z,t)}{\partial t} = D\Delta c(x,y,z,t) + \text{reaction term.} \tag{7.1}$$

In order to understand these processes that lead to pattern formation and cell differentiation, we have to solve the diffusion equations now explicitly (that we have circumnavigated in Section 6). Therefore, we have to acquire a further mathematical tool that are the tricks to solve *partial differential equations*.

7.1 Solving the Diffusion Equations

Partial differential equations (PDEs) like Equation 7.1 are usually not straight forward to solve and mathematicians usually apply some tricks. One way of solving PDEs is to transform them into the reciprocal space by Fourier transformation (FT), in which they are easier to solve. According to Section 2.4, a function $f(\vec{x})$ can be transformed from spatial coordinates (represented by $\vec{x}$) into inverse coordinates (spatial frequency $\vec{k}$) by

$$\int_{-\infty}^{\infty} f(\vec{x},t)\, e^{-i\vec{k}\vec{x}} dV = F(\vec{k})\,. \tag{7.2}$$

Even if one does not understand where this recipe really comes from it is worth to apply it to a PDE in order to see what happens. Applying the FT on both sides of the PDE does not change its mathematical content, but the new structure we obtain might be easier to solve compared to the original PDE. Let us for example look at the lhs of Fick's second law (Equation 6.30) and apply the FT (here in 1D in order to keep it simple, that, fortunately, is not so much different from the 3D case). We find that

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{\partial c(x,t)}{\partial t}\, e^{-ikx} dx = \frac{\partial}{\partial t}\int_{-\infty}^{\infty} c(x,t)\, e^{-ikx} dx = \frac{\partial}{\partial t} C(k,t) \\ &\text{hence,}\quad \frac{\partial c(x,t)}{\partial t} \rightarrow \frac{\partial C(k,t)}{\partial t}\,. \end{aligned} \tag{7.3}$$

It seems that we did not gain much, since we obtained exactly the same structure, but let us investigate the rhs of Fick's second law, that contains the second spatial derivative of c

and we find that

$$\int_{-\infty}^{\infty} \frac{\partial^2 c(x,t)}{\partial x^2}\, e^{-ikx} dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(\frac{\partial c(x,t)}{\partial x}\right)\, e^{-ikx} dx\,. \tag{7.4}$$

The integral seems to be complicated, but it is nothing but the product of two functions: the exponential and the derivative of c, so that we can use partial integration (Section 2.1.7)

$$\int_{-\infty}^{\infty} u'v dx = uv|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} uv' dx$$

and define

$$v = e^{-ikx} \tag{7.5a}$$

$$u' = \frac{\partial}{\partial x}\left(\frac{\partial c(x,t)}{\partial x}\right)\,. \tag{7.5b}$$

From that, we can perform the partial integration of Equation 7.4 and obtain

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(\frac{\partial c(x,t)}{\partial x}\right)\, e^{-ikx} dx = \frac{\partial c(x,t)}{\partial x}\, e^{-ikx}\Big|_{-\infty}^{\infty} - (-ik)\int_{-\infty}^{\infty} \frac{\partial c(x,t)}{\partial x}\, e^{-ikx} dx\,. \tag{7.6}$$

What is the meaning of the first expression on the rhs of Equation 7.6? We know that, by conservation of mass, the concentration of molecules $c(x,t)$ must be always finite, so that for $x \to \pm\infty$, $c(\pm\infty, t) \to 0$. If so, the curve of $c(x,t)$ must get very flat towards $x \to \pm\infty$, so that the derivative $\frac{\partial c(x,t)}{\partial x}$ also approaches zero at infinity, hence there is no flux out. This argument holds especially for positive x since the exponential function pulls down the term to lower values even faster. If not, we would always sum up a finite contribution of c over an **infinite** length, leading to an infinitely large amount of material. Since this is not the case for any reasonable biological system, we can set this part to zero. However, note that the first addend on the rhs of Equation 7.6 does not equal zero in general, just for our special conditions.
Thus, only the second part remains, so that we obtain

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(\frac{\partial c(x,t)}{\partial x}\right)\, e^{-ikx} dx = -(-ik)\int_{-\infty}^{\infty} \frac{\partial c(x,t)}{\partial x}\, e^{-ikx}\, dx\,, \tag{7.7}$$

that is nothing but an iterative procedure showing how to come from the n^{th} derivative under the integral on the lhs to the $(n-1)^{th}$ derivative under the integral on the rhs. Therefore, we just repeat the procedure and find that

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(\frac{\partial c(x,t)}{\partial x}\right)\, e^{-ikx} dx &= -(-ik)\int_{-\infty}^{\infty} \frac{\partial c(x,t)}{\partial x}\, e^{-ikx}\, dx \\
&= \underbrace{c(x,t)\, e^{-ikx}\big|_{-\infty}^{\infty}}_{=0} -(-)(-ik)(-ik)\underbrace{\int_{-\infty}^{\infty} c(x,t)\, e^{-ikx}\, dx}_{C(k,t)} \\
&= -k^2\, C(k,t)
\end{aligned} \tag{7.8}$$

Hence, under the FT, the expression $\frac{\partial^2 c(x,t)}{\partial x^2}$ turns into $-k^2\, C(k,t)$, so that we can summarize:

$$\frac{\partial c(x,t)}{\partial t} \longrightarrow \frac{\partial C(k,t)}{\partial t} \tag{7.9a}$$

$$\frac{\partial^2 c(x,t)}{\partial x^2} \longrightarrow -k^2\, C(k,t)\,. \tag{7.9b}$$

Therefore, Fick's second law reads

$$\boxed{\frac{\partial C(k,t)}{\partial t} = -Dk^2\, C(k,t)}\,, \tag{7.10}$$

and the PDE turned into an ODE after we applied the FT since the derivative applies only wrt t, but not wrt t *and* x anymore. An ODE is easy to solve and we therefore can write down the result

$$C(k,t) = C(k,0)\, e^{-Dk^2\, t} \tag{7.11}$$

immediately. Hence, the idea of applying a FT is that under some conditions PDEs can turn into ODEs.
The expression $C(k,0)$ equals the initial concentration profile in inverse spatial coordinates (*the momentum space*). The function $C(k,0)$ can be anything, not necessarily only a point source. To obtain the solution in spatial coordinates, we have to perform the inverse FT, that is by definition (Section 2.4)

$$\begin{aligned} c(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} C(k,t)\, e^{ikx}\, dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} C(k,0)\, e^{-Dk^2\, t}\, e^{ikx}\, dk \\ &= \frac{1}{2\pi} e^{-\frac{x^2}{4Dt}} \int_{-\infty}^{\infty} C(k,0)\, e^{-\left(\sqrt{Dt}k - \frac{i}{2\sqrt{Dt}}x\right)^2}\, dk\,, \end{aligned} \tag{7.12}$$

where the identity

$$-\left(Dk^2t - ikx\right) = -(\sqrt{Dt}k - \frac{i}{2\sqrt{Dt}}x)^2 - \frac{1}{4Dt}x^2$$

was used in the last step.
Equation 7.12 can now be solved by partial integration, similar to the approach for deriving Equation 7.8. Therefore, I only like to give the result

$$\boxed{c(x,t) = \int_{-\infty}^{\infty} c(\bar{x},0)\frac{1}{\sqrt{4\pi Dt}} e^{-\frac{1}{4Dt}(x-\bar{x})^2} d\bar{x}}\,, \tag{7.13}$$

that is the general solution of Fick's second law in 1D.
If we compare this solution to Equation 6.5 we see that now we have to perform an additional integration in order to derive an explicit expression for $c(x,t)$. The reason is that for the general solution, the initial condition $c(\bar{x},0)$ can be any concentration profile. This profile is multiplied with the Gaussian diffusion term we know from Equation 6.5 with a subsequent integration (such a mathematical structure is called *folding*), meaning, that any pattern of the concentration profile $c(\bar{x},0)$ undergoes diffusion and will be evened out while time elapses. In order to distinguish the variable over which we integrate from the coordinate x, the notation $\bar{x}$ is used here.
For a point source that has the value c_0 at $\bar{x} = x_0$ and zero elsewhere, we use the definition of the delta function δ that is

$$\delta(x - x_0) = \begin{cases} 1, & \text{for } x = x_0 \\ 0, & \text{for } x \neq x_0 \end{cases}\,, \tag{7.14}$$

so that $c(\bar{x},0) := c_0\,\delta(x-x_0)$. Then, the product of $c(\bar{x},0)$ with the Gaussian in Equation 7.13 always yields zero, except for $x = x_0$ so that the integral (that is just the limit of a sum, Section 2.1.7) is only $\neq 0$ at $x = x_0$, where $c(\bar{x},0) = c_0$, so that the solution is

$$c(x,t) = c_0 \frac{1}{\sqrt{4\pi D t}} e^{-\frac{1}{4Dt}(x-x_0)^2}\,, \tag{7.15}$$

that is identical to Equation 6.5 for $x_0 = 0$.

7.2 The Bicoid Profile

The most famous and one of the best studied examples of morphogenesis is the development of the fruit fly *Drosophila melanogaster*. In 1995 Nüsslein-Volhard[66] was awarded with the Nobel Prize in Physiology and Medicine, for her distinguished work on this model organism. She identified the protein *Bicoid* being the first substance which was shown to act as a morphogen. The concentration gradient along the anterior-posterior axis of this protein in the Drosophila embryo is responsible for the separation of the head from the tail. This separation first becomes visible by the appearance of a so-called *cephalic furrow* (Figure 88), that preferentially divides the embryo always with the same proportions.

The source generating Bicoid is maternal mRNA located at the head of the embryo. This

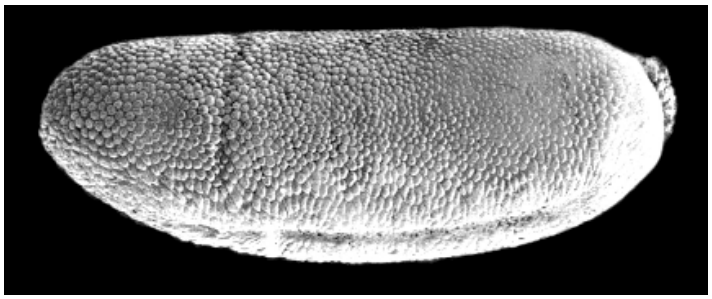

Figure 88: Electron microscopy image of a Drosophila embryo exhibiting the cephalic furrow, roughly one-third of the distance from the left. Image was taken from [1], referring to E. F. Wieschaus.

source is small compared to the size of the embryo and a good approximation might be a point source like the delta function in Equation 7.14. This source continuously produces Bicoid molecules diffusing through the embryo. If diffusion and production would be the only processes, the embryo would be filled completely with Bicoid after some time and no gradient would emerge. Therefore, Bicoid has to be depleted by another process, leading to the formation of a stable gradient at steady state.

We now set up a theoretical model step by step and compare the predictions from this model to the actual measured quantities. This section follows the discussion of the Bicoid gradient in [1] to which I refer to for further reading.

From experiments, the measured Bicoid profile (Figure 89) can be fitted with an exponential

$$c(x) = c_0 \, e^{-x/\lambda} \, , \tag{7.16}$$

where x denotes the position on the anterior-posterior axis, c_0 the concentration at $x = 0$ and λ equals a characteristic length scale. The location of the cephalic furrow x_{cf} corresponds to a concentration c_{cf}, so that we can write

$$x_{cf} = \lambda \ln \left[\frac{c_0}{c_{cf}} \right] \, . \tag{7.17}$$

The measured characteristic length in the embryo is $\lambda \approx 100 \; \mu m$ and the value of x_{cf} is accurate by 1%. In order to keep the proportions of the embryos constant, the location of the cephalic furrow and other features need to have a certain accuracy. The question is how the system keeps these proportions constant for different initial conditions. For example the amount of maternal mRNA in the embryo, i. e. the strength of the source, cannot be controlled by the female and is subject to strong variations.

[66] Christiane Nüsslein-Volhard, 1942 -

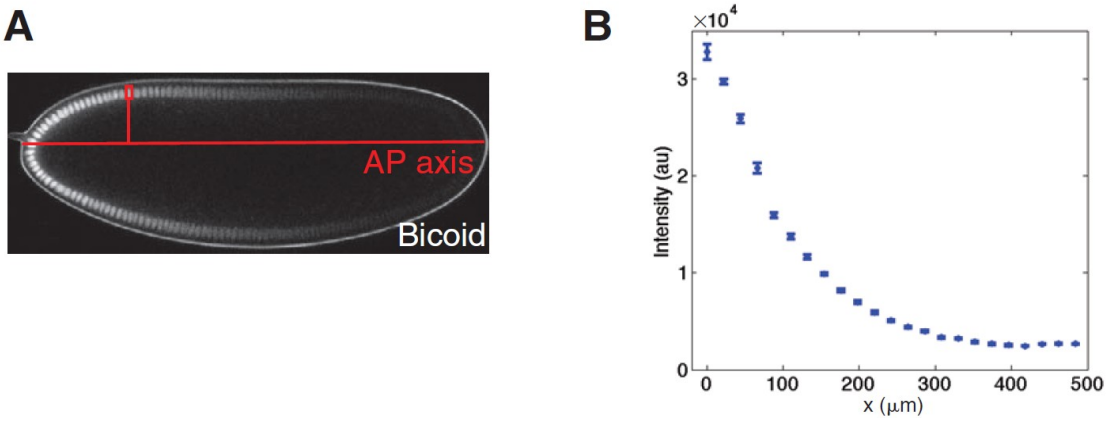

Figure 89: The Bicoid gradient in a Drosophila embryo (left, A) indicated by the fluorescence of the stained cells at the rim and the corresponding concentration profile (right B) measured along the anterior - posterior (AP) axis (red horizontal line in A). Figures are taken from [2].

A minimal model would contain at least a diffusion part and the depletion of Bicoid in order to obtain a gradient in steady state. Thus, we can start with the most basic diffusion equation

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} - \frac{1}{\tau}c(x,t)\,, \tag{7.18}$$

where τ is the lifetime of the Bicoid protein, i. e. its characteristic survival time against its degradation. Equation 7.18 has the structure of Equation 7.1.
In any case, the condition $\frac{\partial c(x,t)}{\partial x}\Big|_{x=L} = 0$ must be valid if L denotes the length of the embryo, since there is no flux out. In steady state Equation 7.18 simplifies to

$$\frac{\partial c_s(x,t)}{\partial t} = 0 \implies D\frac{\partial^2 c_s(x,t)}{\partial x^2} = \frac{1}{\tau}c_s(x,t) \tag{7.19}$$

that is an ODE since the derivative is performed wrt one variable only. Therefore, the solution reads

$$c_s(x) = c_0\, exp\left[-\frac{x}{\sqrt{D\tau}}\right]. \tag{7.20}$$

Comparing the equation above to the findings from the experiments (Equation 7.16), we can conclude that $\frac{1}{\sqrt{D\tau}} \equiv \frac{1}{\lambda}$, from which follows that $\lambda = \sqrt{D\tau}$. Thus, the characteristic length is a function of the diffusion constant and the survival time of the Bicoid protein.
Now, taking the derivative of c_s with respect to x to check for the no flux condition, we find the relation

$$\frac{\partial c(x)}{\partial x}\Big|_{x=L} = -\frac{c_0}{\sqrt{D\tau}}\, exp\left[-\frac{x}{\sqrt{D\tau}}\right]\Big|_{x=L}, \tag{7.21}$$

which must be 0. Since x is always positive this condition can only be fulfilled for $x \to \infty$ or at least approximately if $L >> \lambda$, i. e. if the embryo is large compared to the characteristic length. Typical sizes of Drosophila embryos at this state are $L \approx 500\ \mu m$ and $\lambda \approx 100\ \mu m$, so that $c(x) \sim e^{-\frac{L}{\lambda}} = e^{-5} \approx 0.0067$. This value is in the same order of magnitude (1%) as the required accuracy for the location of the cephalic furrow; or in other words, the embryo is indeed sufficiently large compared to the characteristic length.
The strength R of the source (amount of mRNA) located at the head can be defined by

the gradient at $x = 0$

$$\left.\frac{\partial c(x)}{\partial x}\right|_{x=0} = -\frac{c_0}{\sqrt{D\tau}}\,, \tag{7.22}$$

where $R := -D\left.\frac{\partial c(x)}{\partial t}\right|_{x=0}$, and using the relation above (see also the notation in [1]) it follows that $R = \sqrt{\frac{D}{\tau}}\, c_0$.

The model described by Equation 7.18 well explaines the observed quantities at steady state but it is not complete, since it does not contain the source term and a complete model would also yield the dynamics of the system, where the steady state scenario is included as a sub-case for $t \to \infty$. Therefore, we now include the point source term $\delta(x)$ (c. f. Equation 7.14) for modeling the maternal mRNA in Equation 7.18 and obtain the model

$$\frac{\partial c(x,t)}{\partial t} = \underbrace{D\frac{\partial^2 c(x,t)}{\partial x^2}}_{\text{diffusion}} - \underbrace{\frac{1}{\tau}c(x,t)}_{\text{degradation}} + \underbrace{\kappa\delta(x)}_{\text{source term}}\,. \tag{7.23}$$

Note, that modeling the source with a delta function is a simplification. A further refinement of the model would be a source function $j(x,t)$ that has a certain extension along the x coordinate. In order to characterize the source properties, we have multiplied the source term with a yet unknown constant κ, which has probably some relations to R.

Equation 7.23 is a PDE that is difficult to solve in spatial coordinates. Fortunately, we know from Section 7.1 that this equation can be turned into a much simpler ODE by performing the transformations

$$\begin{aligned}
\frac{\partial c(x,t)}{\partial t} &\to \frac{\partial C(k,t)}{\partial t}\,,\\
\frac{\partial^2 c(x,t)}{\partial x^2} &\to -k^2\, C(k,t)\,,\\
-\frac{1}{\tau}\int_{-\infty}^{\infty} c(x,t)\, e^{-ikx}\, dx &= -\frac{1}{\tau}\, C(k,t)\,,\\
\kappa\int_{-\infty}^{\infty} \delta(x) e^{-ikx}\, dx &= \kappa\,,
\end{aligned} \tag{7.24}$$

where the helpful property $\int_{-\infty}^{\infty}\delta(x)f(x) = f(0)$ of the delta function was used for the source term (for an extended source, we would have to integrate over $j(x,t)$, that might be more difficult). The above transformations turn the PDE into the much simple equation

$$\frac{\partial C(k,t)}{\partial t} = -C(k,t)\left[Dk^2 + \frac{1}{\tau}\right] + \kappa\,. \tag{7.25}$$

Integrating both sides of this equation wrt time t, we find the solution for $C(k,t)$ being

$$C(k,t) = \frac{\kappa}{Dk^2 + 1/\tau}\left[1 - e^{-\left(Dk^2+\frac{1}{\tau}\right)t}\right]\,. \tag{7.26}$$

Since this solution exists only in momentum space, we have to transform it back into spatial coordinates by the inverse FT

$$c(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C(k,t) e^{ikx}\, dk\,. \tag{7.27}$$

This is a certain weakness of the approach now, because such integrals are often hard to solve. One trick is to solve the integral in the space of complex numbers. The underlying

math is not trivial and takes some pages of algebra. I therefore give the final solution without the extensive mathematical derivation, that is

$$\boxed{c(x,t) = \frac{\kappa\lambda}{2D}\left[e^{-\frac{x}{\lambda}} - \frac{1}{2}e^{-\frac{x}{\lambda}}\,\mathrm{erfc}\left(\frac{\frac{2Dt}{\lambda} - x}{\sqrt{4Dt}}\right) - \frac{1}{2}e^{+\frac{x}{\lambda}}\,\mathrm{erfc}\left(\frac{\frac{2Dt}{\lambda} + x}{\sqrt{4Dt}}\right)\right]}. \qquad (7.28)$$

The expression *erfc* is a short cut for $\mathrm{erfc}\,(q) := \frac{2}{\sqrt{\pi}}\int_q^\infty e^{-\bar{q}^2}d\bar{q}$, where $\mathrm{erfc}\,(\infty) = 0$ that is called *complementary error function*.
The Bicoid gradient is one of the modest pattern emerging during the morphogenesis process in Drosophila and already requires quite a portion of algebra that leads to a reltively complicated solution for $c(x,t)$. The model is also still very simplified, since we included no extended mRNA source (c. f. [2]), assumed an infinite length of the empryo and solved the equations in 1D. A Drosophila egg has an approximately elliptical shape so that the Laplacian operator Δ must be solved in elliptical coordinates, that makes the math even more challenging. Some pattern are far more complicated than those of Bicoid, for example the stripy structure caused by the Gt-protein ([3]). One might wonder about the mathematical challenges in such a case. Thus, even simple biological problems turn out be complicated in modeling.
Usually, the mathematical preparation required to find the solution $c(x,t)$ is not discussed in the related papers, since it is assumed to be known by the devoted reader. Therefore, Equation 7.28 is just the third equation in [2]. Hence, what took us tens of hours of lectures is just the intro part on the first pages of a typical biophysics paper.
Equation 7.28 gives now the full dynamics of the Bicoid gradient formation (see Figure 90).

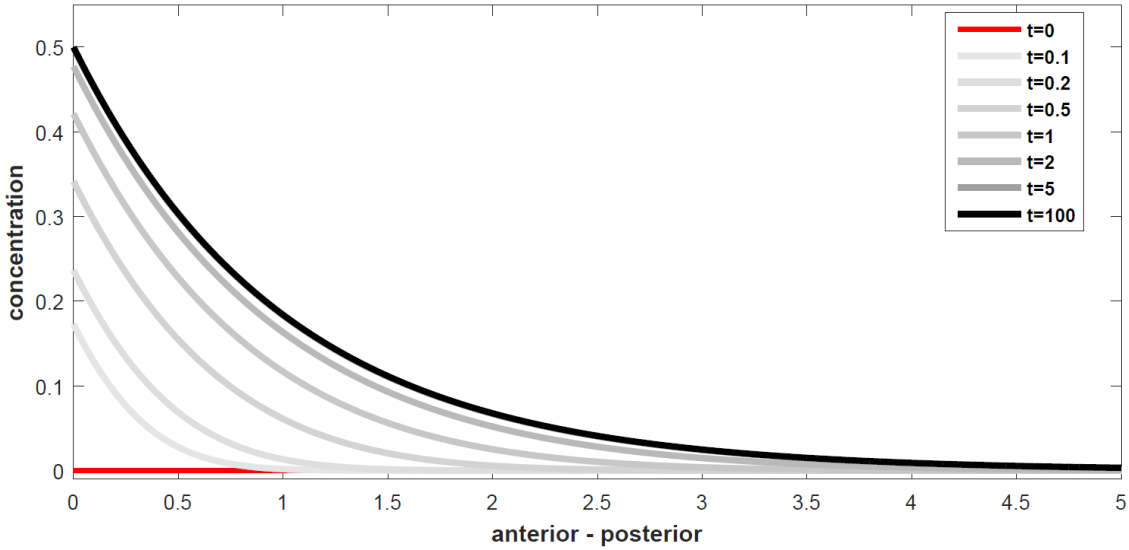

Figure 90: Formation of the Bicoid gradient in a Drosophila embryo for different time steps (in arbitrary units) along the anterior-posterior axis as described by Equation 7.28. Note, that the concentration profile reaches steady state for large t (black line).

Especially, we can recover the steady state solution by applying the limit $t \to \infty$ leading to

$$c(x, t\to\infty) \equiv c_s(x) = \frac{\kappa\lambda}{2D}\left[e^{-\frac{x}{\lambda}} - 0 - 0\right] = \frac{\kappa\lambda}{2D}e^{-\frac{x}{\lambda}}\,, \qquad (7.29)$$

that we can compare to Equation 7.20 unveiling that $\frac{\kappa\lambda}{2D}$ corresponds to c_0, and thus to $c_0 = \sqrt{\frac{\tau}{D}}R$ which further implies that $\kappa = 2R$.

Although the current model completely explains the formation of the Bicoid gradient and links the model parameters κ and τ to the observed quantities c_0, D and λ, it has the weakness that the location of the cephalic furrow x_{cf} (Equation 7.17) depends on the source strength R via c_0. A situation that we actually wanted to avoid.
There are many ways to modify the model now in order to solve this problem. For example, one cloud change the source itself and include an extended source, or, as proposed in [1] hypothesize that the Bicoid molecule is only depleted in a (rare) dimer state. The concentration of the dimer scales with $c^2(x,t)$ since dimerisation occurs with the probability proportional to $c^2(x,t)$ and we need another constant μ scaling for the dimer concentration. Then, the original model Equation 7.18 would read

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} - \frac{1}{\tau\mu}c^2(x,t)\,, \tag{7.30}$$

with the corresponding steady state solution

$$c_s(x) = \frac{6D\tau\mu}{(x+x_0)^2}, \qquad x_0 := \frac{12D^2\tau\mu}{R}^{1/3}. \tag{7.31}$$

If $R >> 1$, the constant x_0 becomes negligibly small, and the concentration $c(x)$ is much less dependent from the strength of the source. The only constrain is that the source strength R has to be large, but the actual amount of the maternal mRNA is not important. Hence, the proportions of the features in the Drosophila embryo are now less independent from the initial conditions.
Although Equation 7.31 is different to Equation 7.20, it might be that a concentration profile proportional to $\frac{1}{x^2}$ mimics an exponential profile in the experiments, especially taking measurement errors into account.

References

[1] William Bialek *"Biophysics: Searching for Principles"*, Princeton University Press, Princeton NJ, 2012

[2] Oliver Grimm, Mathieu Coppey, Eric Wieschaus *"Modelling the Bicoid gradient"*, Development 2010 137: 2253-2264; doi: 10.1242/dev.032409

[3] Johannes Jaeger, Svetlana Surkova, Maxim Blagov et al., *"Dynamic control of positional information in the early Drosophila embryo"*, Nature 430, 368-371: (15 July 2004); doi:10.1038/nature02678

7.3 Activator vs Inhibitor and the Formation of Fur Pattern

The Biocoid gradient was formed by the balance between diffusion and depletion of the substance. Depletion was caused by a second constituent that was included as a loss term in the diffusion equation. However, no possible feedback may caused by the Bicoid molecules acting on the depleting substance was taken into account. Therefore, it was sufficient to model the system with only one equation. We can imagine a system, where the amount of the constituent like that depleting Bicoid (let us denote this substance as u_2) depends on the amount of Bicoid (that we denote as u_1). If there is too much u_1, it activates u_2 that in turn depletes u_1. The substance u_2 will deplete by a loss term until enough u_1 is formed to generate or activate u_2 again. If we want to include a feedback term for such a system, we have to include a second equation. Since these equations are diffusion equations, they must obey the form (c. f. Equation 7.1)

$$\frac{\partial u_1}{\partial t} = D_{u_1}\Delta u_1 + f_{u_1}(u_1, u_2) \tag{7.32a}$$

$$\frac{\partial u_2}{\partial t} = D_{u_2}\Delta u_2 + f_{u_2}(u_1, u_2)\,. \tag{7.32b}$$

Since u_1 activates u_2 it is called *activator* and since u_2 depletes, i. e. inhibiting u_1, it is called *inhibitor*. The expressions $f_{u_1}(u_1, u_2)$ and $f_{u_2}(u_1, u_2)$ can be any reaction terms containing the connection (feedback) between the two constituents u_1 and u_2. Not all functions $f_{u_1}(u_1, u_2)$ and $f_{u_2}(u_1, u_2)$ lead to pattern formation, but there are some functions that cause remarkable pattern of biological relevance that will be introduced (without derivation) now. For a deeper discussion, I like to refer to [1].
One system that leads to a stable gradient formation can be modeled ([1]) by

$$\frac{\partial u_1}{\partial t} = D_{u_1}\Delta u_1 + \rho_{u_1}\left(\frac{u_1^2}{u_2} - u_1\right) \tag{7.33a}$$

$$\frac{\partial u_2}{\partial t} = D_{u_2}\Delta u_2 + \rho_{u_2}\left(u_1^2 - u_2\right) \tag{7.33b}$$

where the ρ_{u_i} are called *removal rates* (the required loss term) and $f_{u_1}(u_1, u_2) = \rho_{u_1}\left(\frac{u_1^2}{u_2} - u_1\right)$ and $f_{u_2}(u_1, u_2) = \rho_{u_2}\left(u_1^2 - u_2\right)$. The idea is that such a system leads to gradient formation, even if the initial conditions, i. e. the concentrations $u_1(x, y, z, t = 0)$ and $u_2(x, y, z, t = 0)$ are random and noisy so that the system starts completely featureless. The gradients emerging in steady state can work as "seeds" for the actual pattern formation.
A particular model discussed in [1] is

$$\frac{\partial a}{\partial t} = D_a\Delta a + \rho_a\frac{a^2}{(1 + \kappa_a a^2)h} - \mu_a\, a + \sigma_a \tag{7.34a}$$

$$\frac{\partial h}{\partial t} = D_h\Delta h + \rho_h\, a^2 - \mu_h\, h + \sigma_h\,. \tag{7.34b}$$

Here, κ_a is a saturation constant for the activator a and μ_a and μ_h are the removal rates for the activator and the inhibitor h, respectively. The parameter σ_i are basic production terms for the activator and the inhibitor. It is always possible to scale $\mu_a = \rho_a$ and $\mu_h = \rho_h$ for arbitrary units (for example if the time increment Δt is scaled to one, see also the appendix in [1]), so that Equation 7.34a and Equation 7.34b can be transformed into Equation 7.33a and Equation 7.33b, respectively.
Pattern emerge ([1] and references therein), if $D_h \gg D_a$ (typically 10^{-6} cm^2/s vs 10^{-8} cm^2/s) and if $\mu_h \gg \mu_a$. A particular simulation is explained in the next section.

7.3.1 Numerical Treatment

Finding an analytical solution for one particular diffusion equation was difficult (Section 7.2), finding a solution for a system of diffusion equations is often literally impossible. Therefore, a system described by the type of Equation 7.33a and Equation 7.33b can only be solved numerically. The lhs of the PDE determines the temporal evolution of the system, so that for a given state at time t, the properties of the system at time $t + \Delta t$ can be calculated. Starting from the initial conditions $t = 0$, the values for the inhibitor and the activator can be derived for any time slice until the last time step T. As the procedure implies, time, and also the spatial coordinates, have to be discretizised into small increments.

For a 2D simulation, u_1 and u_2 are functions of x, y and t so that we define the

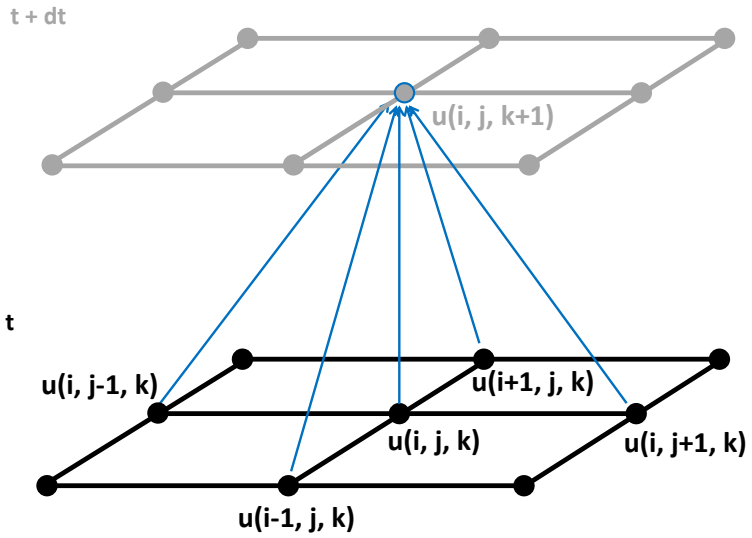

Figure 91: Geometrical illustration of the Laplace operator of the diffusion term in a diffusion reaction. The future point $u(i, j, k + 1)$ is not only influenced by its previous state $u(i, j, k)$, but also depends on the previous states of its neighbouring points $u(i - 1, j, k)$, $u(i + 1, j, k)$, $u(i, j - 1, k)$ and $u(i, j + 1, k)$. See also Figure 84 for comparison.

corresponding increments as $x_{n+1} - x_n = \Delta x$, $y_{n+1} - y_n = \Delta y$ and $t_{n+1} - t_n = \Delta t$, respectively, for the n^{th} step in the simulation. The results are stored in a matrix u for each reactant during the simulation. In a 2D system, each matrix has *three* indices i, j and k, that stand for the coordinates x, y and t, respectively. Then, the lhs of the diffusion equation can be approximated by (Equation 2.33)

$$\frac{\partial u}{\partial t} \approx \frac{u\,(i, j, k + \Delta t) - u\,(i, j, k - \Delta t)}{2\,\Delta t}\,. \tag{7.35}$$

The rhs of the diffusion equations contains the Laplace operator that equals $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ for two dimensions in Cartesian coordinates, so that we need the second derivatives. According to Equation 2.40, Equation 2.48a, Equation 2.48b, Equation 2.49a and Equation 2.49b, the Laplace operator can be approximated by

$$\begin{aligned}\Delta u \approx &\frac{u\,(i + \Delta x, j, k) - 2\,u\,(i, j, k) + u\,(i - \Delta x, j, k)}{\Delta x^2} \\ &+ \frac{u\,(i, j + \Delta y, k) - 2\,u\,(i, j, k) + u\,(i, j - \Delta y, k)}{\Delta y^2}\,.\end{aligned} \tag{7.36}$$

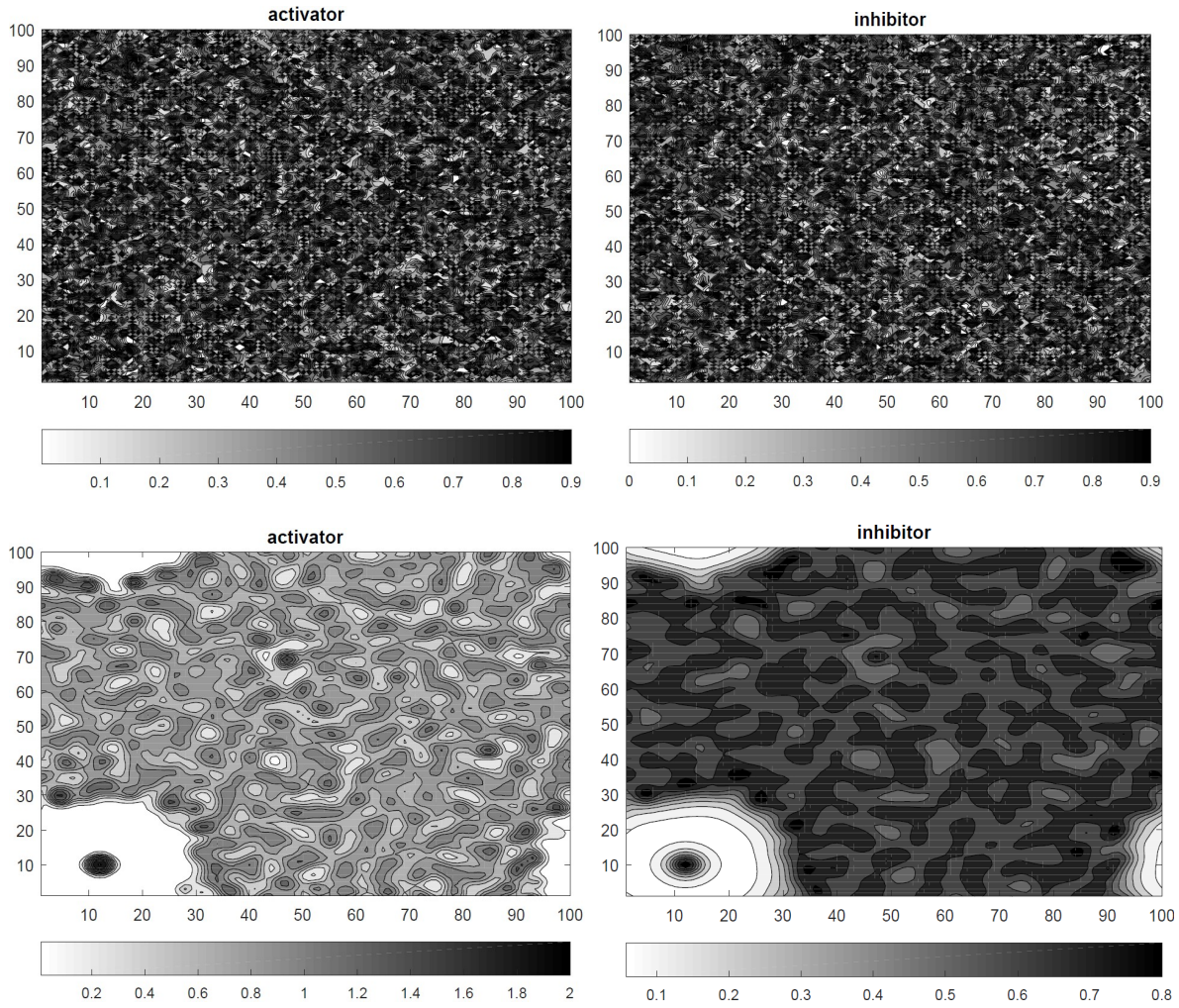

Figure 92: Simulation of the formation of a stable gradient (bottom left, lower panel) according to Equation 7.34. The boundary conditions are cyclic and the plots are derived for random initial conditions (upper panel) and after 500 time steps (bottom).

The indices of the cells in a matrix are integer numbers starting at 1 and increasing in steps of 1 so that it is most suitable (by not necessary) to scale the system in order to set $\Delta t = \Delta x = \Delta y = 1$. In this scale, every change of the coordinates in the system by the increments correspond to a change in the index of the matrix. Using this scale and combining Equation 7.35 and Equation 7.36, the entries of a matrix u are filled iteratively according to

$$\begin{aligned} u(i,j,k+1) \approx 2\left[u\left(i+1,j,k\right) - 2\,u\left(i,j,k\right) + u\left(i-1,j,k\right)\right] \\ +2\left[u\left(i,j+1,k\right) - 2\,u\left(i,j,k\right) + u\left(i,j-1,k\right)\right] \\ +u(i,j,k-1) + 2\,f(u_1,u_2)\,, \end{aligned} \tag{7.37}$$

once the initial conditions are chosen.
One can see in Equation 7.37 that a point at $u(i,j,k+1)$ in the future is influenced not only by its previous state $u(i,j,k)$, but also by the previous states of its neighbouring points $u(i-1,j,k)$, $u(i+1,j,k)$, $u(i,j-1,k)$ and $u(i,j+1,k)$. This connection is caused by the second derivatives from the Laplace operator. The geometrical illustration is shown in Figure 91.
As an example, the results of the simulation of the diffusion reaction model given by Equa-

tion 7.34 in [1] for $D_a = 0.005$, $D_b = 0.2$, $\rho_a = 0.01$, $\rho_b = 0.02$ and $\kappa_a = 0.25$ for a grid size of $L_x \times L_y = 100 \times 100$ points are shown in Figure 92. The system starts with random noise (upper panel in Figure 92) and no structure is visible. A stiff gradient emerges on the lower left edge after 150 time steps, becomes flatter soon afterwards but remains as stable structure (lower panel, after 500 time steps) for $t \longrightarrow \infty$. The boundary conditions are cyclic ($u(0, j, k) = u(L_x, j, k)$ and $u(i, 0, k) = u(i, L_y, k)$, where L_x and L_y are the size of the matrix in the respective directions) in order to avoid numerical artifacts on the edges of the grid. Note, that the important issue is that even if there was not any structure present in the initial conditions, a stable gradient formed almost inevitable. Repeated simulations exhibit that the gradient will appear at random positions, but always with the same properties (maximum of the peak, shape and slope). This gradient then may act as the initialization for the next pattern formation,

7.3.2 Fur Pattern

Such a next step can be the formation of pattern that are more complicated than just a gradient. Many common fur pattern can be modeled with a relative small effort using the diffusion equations of the type (see [1] for explanations)

$$\frac{\partial a}{\partial t} = D_a \Delta a + \rho_a \left[\frac{a^2 s}{1 + \kappa_a a^2} - a \right] , \tag{7.38a}$$

$$\frac{\partial s}{\partial t} = D_s \Delta s + \frac{\sigma_s}{1 + \kappa_s y} - \frac{\rho_s a^2 s}{1 + \kappa_a a^2} - \mu_s s \, , \tag{7.38b}$$

$$\frac{\partial y}{\partial t} = \rho_y \frac{y^2}{1 + \kappa_y y^2} - \mu_y y - \sigma_y a \, . \tag{7.38c}$$

Once a gradient is formed by the reactions modeled by Equation 7.34, the reactions Equation 7.38 take place. For example one obtains the pattern of **giraffe fur or crocodile skin** if we set

$$\begin{array}{lllll} D_a = 0.015 & \rho_a = 0.025 & \mu_s = 0.00075 & \sigma_s = 0.00225 & \kappa_a = 0.1 \\ D_s = 0.03 & \rho_s = 0.0025 & \mu_y = 0.003 & \sigma_y = 0.00015 & \kappa_s = 20 \\ & \rho_y = 0.03 & & & \kappa_y = 22 \end{array}$$

in that unit system where a time step corresponds to the shift of one index in the matrices for a, s and y.
We obtain the pattern of **cheetah fur** with the identical equations, but slightly different parameter

$$\begin{array}{lllll} D_a = 0.01 & \rho_a = 0.05 & \mu_s = 0.003 & \sigma_s = 0.0075 & \kappa_a = 0.5 \\ D_s = 0.1 & \rho_s = 0.0035 & \mu_y = 0.003 & \sigma_y = 7 \times 10^{-5} & \kappa_s = 0.3 \\ & \rho_y = 0.03 & & & \kappa_y = 22 \, . \end{array}$$

The simulations are visualized in Figure 93.
It is remarkable, that the equations are identical, but only a change in the diffusion constants and the reaction rates leads to completely different pattern. Most of the common fur and skin pattern and also for example the morphogenesis of a fly eye ([1]) can be modeled with these equations. This of course does not explain the underlying processes, but it reveals a fundamental connection between phenomena, that seem to be completely different in the first glimpse.

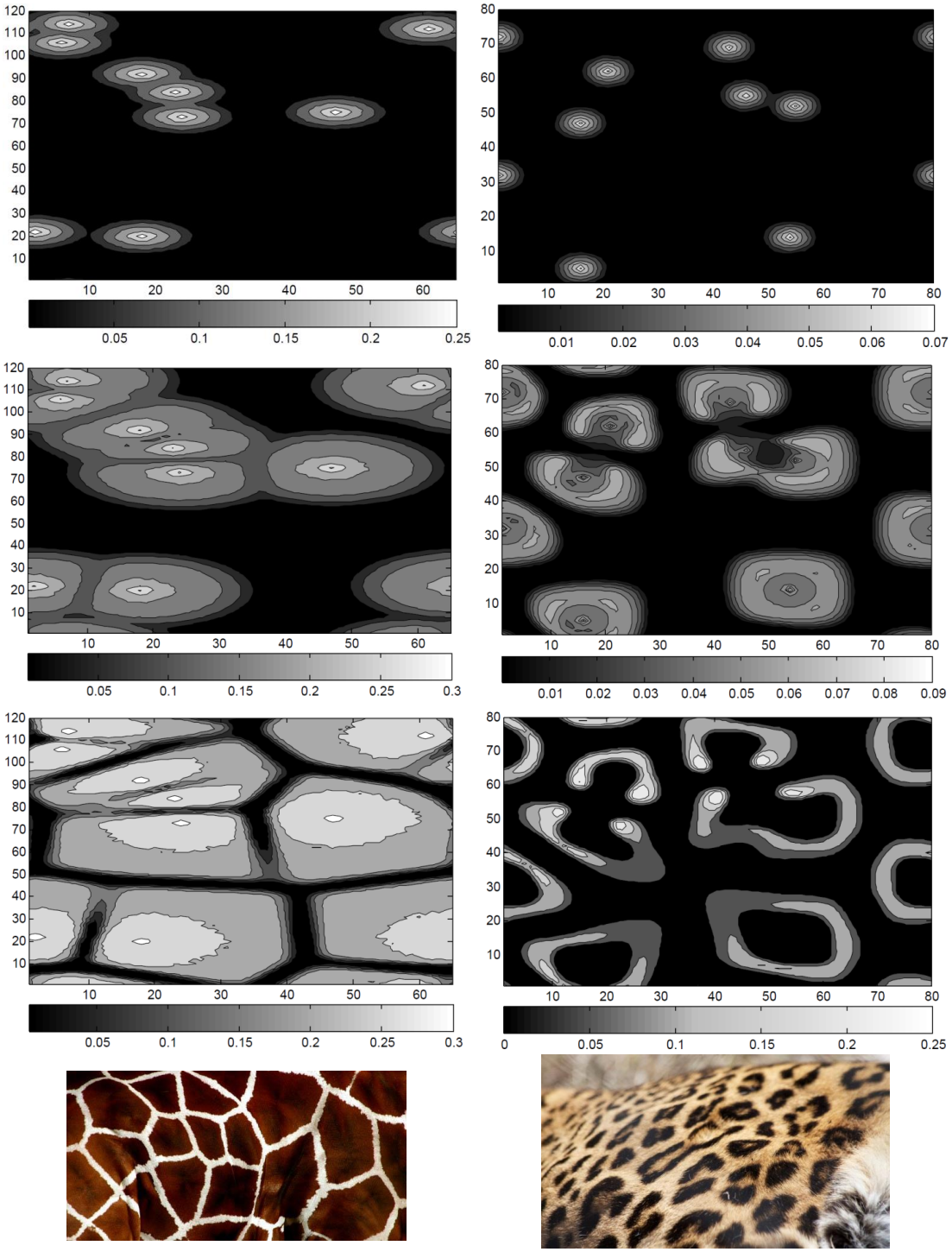

Figure 93: Simulation of the formation of fur pattern based on Equation 7.38.
Left: Formation of giraffe fur pattern starting from gradients (Section 7.3.1) as "seeds" (upper panel) and shown after 500 (middle) and $2,500$ (bottom) time steps. The used grid size was 120×65 points with cyclic boundary conditions.
Right: Formation of cheetah fur pattern (initial condition in the top panel) shown after $1,500$ (middle) and $3,000$ (bottom) time steps. The used grid size was 80×80 points, again with cyclic boundary conditions. The figures correspond to figures $9a$ to $9c$ in [1].

References

[1] A. J. Koch and H. Meinhardt *"Biological pattern formation: from basic mechanisms to complex structures"*, Rev. Mod. Phys. 66, 1481, 1994

8 Fluid Dynamics and Microfluidics

An important field in synthetic biology, evolutionary biology and medical biochemistry is microfluidics. This technique is required to generate supramolecular assemblies like liposomes that may act for example as a carrier for drugs; or to investigate the interactions between cellular compounds and sub units in synthetic biology [1; 2]. Such processes like mixing and separation of fluids require high precision and a certain level of control over the fluid dynamics in volumes from micro liter down to femto liter scale. Therefore, some background knowledge about fluid dynamics is required in order to understand the technique itself and of course to a certain extend the content of related specialist literature.
The good news is that we do not need the complete math of fluid dynamics, since our medium is always liquid water in a low pressure regime at constant temperature and (mostly) without turbulence. These approximations simplify the underlying math tremendously. The key part is the understanding of the meaning of the *Navier - Stokes - equation* since everything we need is derived by it. Fortunately, the Navier - Stokes - equation can be derived from very simple assumptions (conservation of mass and energy).
A typical experiment to produce supramolecular assemblies is illustrated in Figure 94. Many

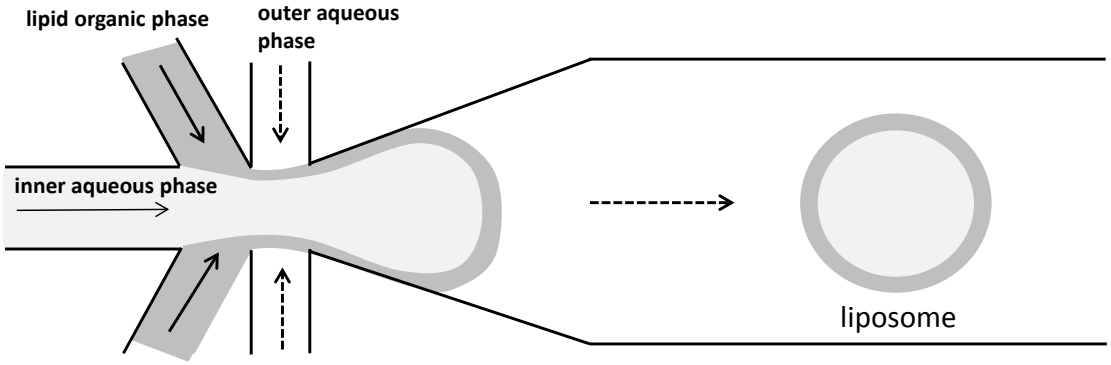

Figure 94: Assembly of a liposome in a micro channel using the double emulsion (two aqueous phases) technique.

physical processes occur in an experiment like the one shown in this image. For example when the diameter of the micro channel changes, the flow velocity changes accordingly. Due to the conservation of mass and the incompressibility of water, the flow velocity decreases if the diameter increases and vice verse. The formation of the lipid bilayer obviously has to do something with surface tension and the growth of such a droplet is determined by the flow through the micro channel. All these quantities have to be calculated in order to set up a working experiment.

8.1 The Navier - Stokes - Equation

All the phenomena mentioned above can be quantified with one equation, the Navier - Stokes - equation (NSE)[67]. Essentially, the NSE is just Newton's law $\vec{F} = m\vec{a}$. Newton's law is fundamental and it that sense always applicable - in particular for our purposes. It turned out, that it is more convenient to use densities instead of the actual quantities. For example we know already the mass density ρ that is mass per volume. In the same manner

[67] Claude Louis Marie Henri Navier, 1785 - 1836 and Sir George Gabriel Stokes, 1. Baronet PRS, 1819 - 1903

we can define the force density $\vec{f}$ that is force per volume and the energy density ϵ that is energy per volume. The advantage is that the volume term cancels out in equations and connections to other quantities are easier to show.
For example pressure p has the same unit (force per area equals $\frac{kg}{m\,s^2}$) as energy density (energy per volume equals $\frac{J}{m^3} = \frac{kg}{m\,s^2}$) and therefore the force density is proportional to the pressure gradient $\vec{f} \sim \nabla p$ since the gradient of energy is proportional to a force (Section 2.1.8). Thus, we write Newton's law with densities

$$\vec{f} = \rho\,\vec{a}\,. \tag{8.1}$$

The rhs of Newton's law contains acceleration, hence the temporal derivative of velocity. This quantity might be useful for us since a change of diameter in a micro channel causes a change of flow velocity (hence an acceleration/deceleration). The flow velocity is a vector $\vec{v}$ with its components

$$\vec{v}(\vec{x},\vec{y},\vec{z}) = \begin{pmatrix} v_x(\vec{x},\vec{y},\vec{z}) \\ v_y(\vec{x},\vec{y},\vec{z}) \\ v_z(\vec{x},\vec{y},\vec{z}) \end{pmatrix} \tag{8.2}$$

that themselves can depend on location $\vec{r} = (\vec{x},\vec{y},\vec{z})$.
According to our findings in Section 2.1.4, the temporal derivative of the velocity field is

$$\begin{aligned}
\vec{a}(\vec{x},\vec{y},\vec{z}) =& \frac{d}{dt}\vec{v}(\vec{x},\vec{y},\vec{z}) \\
=& \frac{\partial}{\partial\vec{x}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\frac{d\vec{x}}{dt} + \frac{\partial}{\partial\vec{y}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\frac{d\vec{y}}{dt} + \frac{\partial}{\partial\vec{z}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\frac{d\vec{z}}{dt} + \frac{\partial}{\partial t}\vec{v}(\vec{x},\vec{y},\vec{z}) \\
=& \frac{\partial}{\partial\vec{x}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\vec{v}_x + \frac{\partial}{\partial\vec{y}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\vec{v}_y + \frac{\partial}{\partial\vec{z}}\vec{v}(\vec{x},\vec{y},\vec{z})\,\vec{v}_z + \frac{\partial}{\partial t}\vec{v}(\vec{x},\vec{y},\vec{z}) \\
=& \vec{v}(\vec{x},\vec{y},\vec{z})\;(\nabla\vec{v}(\vec{x},\vec{y},\vec{z})) + \frac{\partial}{\partial t}\vec{v}(\vec{x},\vec{y},\vec{z})\,,
\end{aligned} \tag{8.3}$$

where the expression

$$\nabla\vec{v}(\vec{x},\vec{y},\vec{z}) = \begin{pmatrix} \partial_x v_x(\vec{x},\vec{y},\vec{z}) & \partial_y v_x(\vec{x},\vec{y},\vec{z}) & \partial_z v_x(\vec{x},\vec{y},\vec{z}) \\ \partial_x v_y(\vec{x},\vec{y},\vec{z}) & \partial_y v_y(\vec{x},\vec{y},\vec{z}) & \partial_z v_y(\vec{x},\vec{y},\vec{z}) \\ \partial_x v_z(\vec{x},\vec{y},\vec{z}) & \partial_y v_z(\vec{x},\vec{y},\vec{z}) & \partial_z v_z(\vec{x},\vec{y},\vec{z}) \end{pmatrix} \tag{8.4}$$

was used in the last step[68]. The meaning of ∇ is not just a gradient that turns a scalar into a vector by performing the derivatives like we know from a potential ϕ where $\vec{f} = grad\,\phi \equiv \nabla\phi$ (Section 2.1.4). Now, we need the derivative of a vector field $\vec{v}(\vec{x},\vec{y},\vec{z})$ with its components $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$. Such a structure is called *tensor*, since it indeed was first found by describing surface *tension*. For our purposes we actually do not need to know what a tensor really is, but I like to mention it since it occurs frequently in microfluidics and fluid dynamics textbooks.

Exercise I:

Ensure that you really understood Equation 8.4 by showing that $(\nabla\vec{v})\,\phi \neq (\vec{v}\,\nabla)\,\phi$ *where* ϕ *is a scalar field; and that in contrast* $\vec{v}\cdot\nabla\vec{v} = (\nabla\vec{v})\,\vec{v}$.

[68] Also with the common short cut $\partial_i \equiv \frac{\partial}{\partial i}$.

With these notations, Newton's law can be written as

$$\rho\left(\frac{\partial\vec{v}}{\partial t}+\vec{v}\,\nabla\vec{v}\right)=\vec{f}\,. \tag{8.5}$$

Often the absolute derivative of the velocity field is written as $\frac{D}{Dt}\vec{v}:=\frac{\partial\vec{v}}{\partial t}+\vec{v}\,\nabla\vec{v}$.
What are possible forces acting on our liquid? As discussed already, a pressure gradient implies a force density. Imagine a vertical water column where we measure the pressure along the depth z from the surface to the bottom. Pressure is caused by the weight of the fluid so that it increases with depth. Hence, if the pressure equals p at given z, it equals $p+dp$ at $z+dz$ (Figure 95).

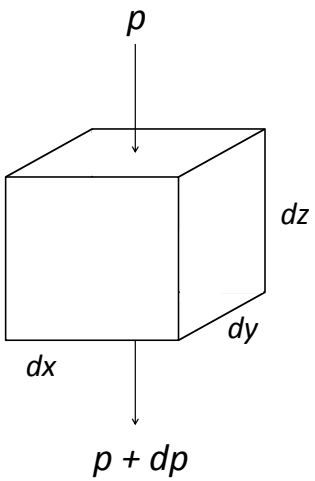

Figure 95: Pressure as a function of depth z in a fluid column implies a force under gravity by the weight of the liquid.

Pressure is defined as force per area, hence $p:=\frac{dF}{dA}$. Thus, the force in z direction implied by the pressure gradient in Figure 95 is $dF_z=p\,dA-(p+dp)\,dA$, where $dA=dx\,dy$. Hence, $dF_z=-\frac{dp}{dz}\,dV$ or for the force density $df_z=-\frac{dp}{dz}$. Since force is a vector, we can formulate this relation for any direction and therefore obtain the result

$$\vec{f}=-\nabla\,p\,. \tag{8.6}$$

It is important to note, that the gradient points in the opposite direction of the force.
Another force is friction in the liquid caused by viscosity. It is intuitive to add such a term because it is a difference whether one presses water (small viscosity) or honey (high viscosity) through the same pipe. Thus, some of the kinetic energy gets lost by the dissipative drag force that contributes to Equation 8.5. In a laminar flow (no turbulence), the liquid can be separated into infinitely thin layers of height Δy where the drag force is caused by friction between these layers (Figure 96). Friction between the layers causes a velocity gradient along the vertical (y) axis. Thus, the force required to move these layers must be proportional to $\frac{\Delta v}{\Delta y}$. Also, friction increases for larger area A between the layers, where it can act on. Therefore, the drag force is also proportional to A. Introducing a (yet unknown) proportionality constant η, we can write

$$F_y=\eta A\,\frac{\Delta v}{\Delta y}\,. \tag{8.7}$$

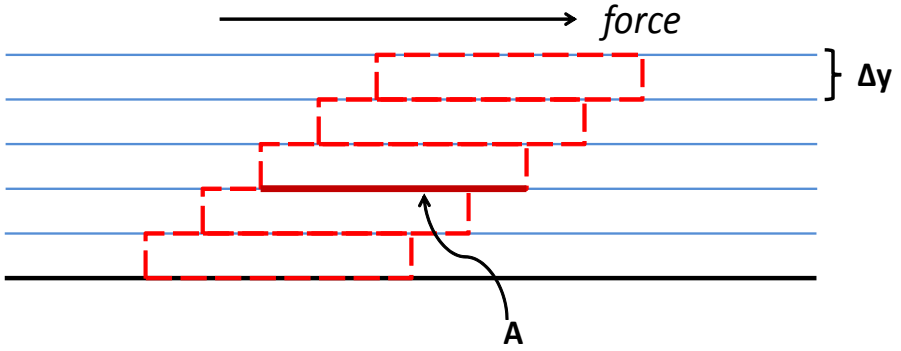

Figure 96: Model for explaining friction caused by viscosity in a liquid.

The proportionality constant is called *dynamic viscosity*.
We recall that pressure is force per area and that the gradient of pressure equals force density that we have on the rhs in Equation 8.5. Therefore, $f_y = \eta\, grad_y \left(\frac{\Delta v}{\Delta y}\right)$ or for infinitely small increments $f_y = \eta \frac{d^2v}{dy^2}$. This construction is valid for any spatial component, so that the force density required to overcome friction caused by viscosity η equals

$$\vec{f} = \eta\, \Delta\vec{v}\,, \tag{8.8}$$

where $\Delta = \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}$ is the Laplace operator here.
Combining all these forces with Equation 8.1, we finally obtain the NSE

$$\boxed{\rho\left(\frac{\partial\vec{v}}{\partial t} + \vec{v}\,\nabla\vec{v}\right) = -\nabla p + \eta\,\Delta\vec{v} + \vec{f}_{ext}}\,, \tag{8.9}$$

where the force density at the end of the equation obtained the index "ext" to emphasize its origin as further external force (electrical forces, gravity etc).
The NSE express the balance between inertia on the lhs and external and internal forces on the rhs. The expression $\frac{\partial\vec{v}}{\partial t}$ is called *local acceleration*, whereas the expression $\vec{v}\,\nabla\vec{v}$ is denoted as *advective acceleration*. The velocity profile can be any function with the sub-case of a linear relation. If the velocity profile is linear, the liquid is called *Newtonian fluid*.
The NSE as derived here is valid for laminar flow (otherwise the friction term must be different) only. For the density ρ, no constrain was made so far so that it can be any function $\rho = \rho(\vec{x}, \vec{y}, \vec{z})$. However, it is often stated that water is incompressible. The question is now what this means concerning ρ and for the NSE in general and how it is expressed mathematically, and, whether incompressibility is a good approximation or not.

8.1.1 Water is (Almost) Incompressible

If mass in a volume element is conserved, we can use the Fokker - Planck equation (Equation 6.52) to express the density as

$$\begin{aligned}\frac{d\rho}{dt} &= D\,\Delta\rho - div\,(\rho\vec{v}) \\ &= D\,\Delta\rho - \vec{v}\,grad\rho - \rho\,div\vec{v}\,.\end{aligned} \tag{8.10}$$

If a fluid is incompressible, the density is constant, hence neither a function of time or of any spatial coordinates. Therefore, $\Delta\rho$, $grad\,\rho$ and $\frac{d\rho}{dt}$ equal zero. The only expression that remains in Equation 8.10 is

$$\boxed{div\,\vec{v} = 0}\,. \tag{8.11}$$

In other words, the velocity vector field has to be free of sinks and sources. This constrain is a condition needed to solve the NSE and to obtain an unique solution.
However, strictly speaking the statement that a substance is incompressible is almost certainly wrong. Also water is actually compressible. Compressibility κ is defined as the relative change of a volume $\frac{\Delta V}{V}$ depending on the change of pressure Δp. If a gas is compressed, it releases heat. If the entire work required for compression is transferred into heat, then the entropy of the compressed system stayed constant (c. f. Section 3). Since entropy is constant, the corresponding compressibility κ_S is called *isentropic compressibility*.
If the compression occurs isothermal, all the work is put into the reconfiguration of states (within the system) and the entropy changes. This process of compression is called *isothermal* and the corresponding *isothermal compressibility* is usually denoted as κ_S. According to the relative change of volume, the compressibilities are defined as

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \quad \text{and} \tag{8.12a}$$

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_S\,, \tag{8.12b}$$

respectively.
Usually, compression is neither completely isothermal nor completely isentropic, but a mixture of both. However, as a good approximation most of the compression processes are almost isentropic, since thermal heat diffusion is much slower than the propagation of audio acoustic distortions. For example the speed of sound (sound waves are propagating periodic distortions) in water c is in the order of $1,500\ m/s$ - much faster than any heat transport in this medium. Thus, if water is compressible, κ_S would be the relevant quantity.
One can show, that

$$\kappa_S = \frac{1}{\rho\, c^2} \tag{8.13}$$

and taking $\rho \approx 1,000\ kg/m^3$, the isentropic compressibility of water in a biological relevant temperature range is in the order of $10^{-9}\ /Pa$ (note, that κ_T is in the same order of magnitude). This means (Equation 8.12b) that a change of pressure by $\Delta p = 10^7\ Pa$ causes a relative change in the volume of only 1%. The atmospheric pressure on the terrestrial surface is in the order of $p = 10^5\ Pa$ (corresponding to an air column of $10\ km$ height and $1\ m^2$ area with constant density of $1\ g/l$), or in other words, a pressure difference of $\Delta p = 10^7\ Pa$ equals the weight of $10^3\ tons$ per square meter on the terrestrial surface. Such pressure differences do not occur in any biological system. Thus, **for our purposes, treating water as an incompressible liquid is a very good approximation.**

8.2 The Reynolds Number

The NSE together with the divergence of the velocity field fully describe the dynamics of a fluid. One may ask now whether the dynamics in a system do not change with the scale of the system. For example one can build a micro channel and an experimental set up like in Figure 94 that is in μm size. Does the same system behave differently if the set up gets blown up to meter scale, i. e. if it gets 10^6 times large? One might first think that such a scale difference does not matter, but as larger (and therefore as more massive) the system, as more important inertia becomes.
Every quantity has a value and a corresponding unit. The length of one meter is written "$1\ m$". If we change the scale, e. g. from meter to centimeter, the same quantity is written as $100\ cm$. The value increased while the scale decreased. This might sound trivial, but it has some consequences. For example the density $[\rho] = kg/m^3$ would change by $\frac{kg}{m^3} = \frac{kg}{(10^2)^3 cm^3}$ to $10^{-6}\ kg/cm^3$. Thus, the density decreases while the scale also decreases.
If we would change the scale l to a new scale $\bar{l}$ (like from meter to centimeter), for example with a factor λ, the old length scale is substituted by $l = \bar{l}\,\lambda$. Since velocity is length per time, also the velocity would change by $v = \bar{v}\,\nu$ (all vector arrows are left away further on for the sake of clarity) that consequently leads to a new time scale $t = \bar{t}\,\tau$. Also related quantities like the gradient and divergence (both with unit $1/length$) and the Laplace operator (unit $1/length^2$) would change too.
Such a re-scaling would turn the lhs of the NSE into

$$\frac{\rho\nu^2}{\lambda}\left(\frac{\partial\bar{v}}{\partial\bar{t}} + \bar{v}\,\bar{\nabla}\bar{v}\right). \tag{8.14}$$

If we only scale the size of the system, quantities like density or viscosity remain untouched. Imagine the micro channel filled with water scaled up by a factor of 10^6. It still contains water with the identical viscosity[69] as before. Pressure equals force per area, so that there is no natural scaling of pressure anyway.
Scaling the viscosity part in the NSE leads to

$$\eta\,\frac{\nu}{\lambda^2}\,\bar{\Delta}\bar{v}\,. \tag{8.15}$$

Hence, both, the inertia part and the viscosity part of the NSE, have their pre-factors when scaled to a different size of the system. **If the dynamics of the system has to be identical for different scales, the *ratio* of the two pre-factors has to be constant.** This ratio

$$\frac{\rho\nu^2}{\lambda} \,/\, \eta\,\frac{\nu}{\lambda^2} \tag{8.16}$$

is a dimensionless factor

$$\boxed{\Re e = \frac{\rho\lambda\nu}{\eta}} \tag{8.17}$$

called *Reynolds number*[70].
The quantities ρ and η are the properties of the medium, where λ equals a length scale of an object in the medium and ν its velocity wrt the flow velocity. The dynamic viscosity of water around 20^oC is in the order of $10^{-3}\ Pa\,s$. A bacteria of μm length and with a typical velocity of some hundred $\mu m/s$ swimming in water has a Reynolds number of 10^{-4}.

[69] Note, that often textbooks use the *kinematic viscosity* η/ρ instead.
[70] Osborne Reynolds FRS, 1842 - 1912

A human of $1\ m$ scale has a Reynolds number of 10^{+4} in the same medium. This is a difference of eight orders of magnitude. Thus, the dynamics are not comparable at all. If a human would like to feel how bacteria feel when they swim in water, one has to change a parameter in Equation 8.17 in order to obtain the same Reynolds number. Hence, to compensate a factor ratio of 10^8 between the different Reynolds numbers, the viscosity has to be increased by this ratio. However, a viscosity of $\eta_{water} \times 10^8$ "feels" like concrete just before it sets. Hence, bacteria swimming in water feel like us if we would swim in concrete! Generally, smaller objects have a smaller Reynolds number and therefore "feel" a higher viscosity. The reason is that the smaller objects, like bacteria, have an inertia that is almost negligible compered to friction in the particular liquid. That is why larger objects, like humans in water or gigantic supertanker, can coast whereas bacteria not (c. f. Figure 97). Laminar flow exists for $\Re e \lesssim 0.1$. Then the drag force $\vec{F}$ can be approximated by $\vec{F} = -\gamma\,\vec{v}$ with the drag coefficient γ. On the other hand, force is just $\vec{F} = m\,\frac{d\vec{v}}{dt}$, so that the velocity

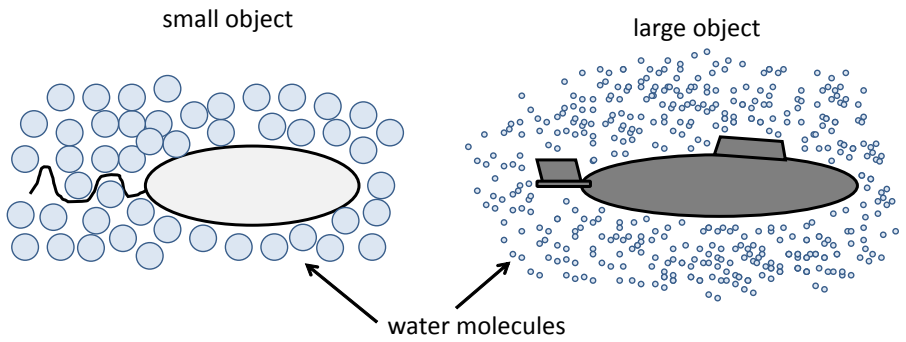

Figure 97: Due to the "grain size" of the medium, smaller objects (left) have a lower Reynolds number than larger objects (right). For larger objects, hence larger Reynolds numbers, the NSE are dominated by inertia, whereas the dynamics of smaller objects are dominated by viscous forces.

v_0 of an object if it is not moved actively decreases according to

$$v(t) = v_0\, e^{-\frac{\gamma}{m}\,t}\,, \tag{8.18}$$

where $\tau = \frac{m}{\gamma}$ is a typical time scale.
It can be shown that $\gamma = 6\pi\eta\, r$ for a spherical body of radius r (*Stokes law*). For a homogeneous body $m = \frac{4}{3}\pi\, r^3\rho$ and therefore, the characteristic time can be expressed as

$$\tau = \frac{2\rho\, r^2}{9\eta} = \frac{2\, r}{9\, v_0}\,\Re e\,. \tag{8.19}$$

This time scale corresponds to a characteristic length $l = \tau\, v_0$.

Exercise I:

What is the characteristic length of bacteria swimming in water? Compare this value to the size of an atom. What is the characteristic length and time scale of a human swimming in water? Compare the dynamic viscosity of water, honey, cytoplasm and concrete. What does it tell you?

8.3 Flow Through a Pipe: The Law of Hagen - Poiseuille

For a micro channel, the scale is sufficiently small so that $\Re e \lesssim 0.1$. Therefore, inertia can be neglected and the lhs of the NSE (8.9) disappears, leading to

$$0 = -\nabla p + \eta \, \Delta \vec{v} \, . \tag{8.20}$$

The micro channel is of cylindrical shape with radius R and length z (Figure 98), so that we express the Laplace operator in cylindrical coordinates and the equation can be written as

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v_z(r)}{\partial r} \right) = \frac{1}{\eta} \frac{\partial p}{\partial z} \, . \tag{8.21}$$

We can integrate this equation two times wrt r in order to remove the derivatives and to obtain the velocity profile in z direction along the radial coordinate r in the micro channel. While integrating two times, we obtain two integration constants c_1 and c_2, respectively and find that

$$v_z(r) = \frac{1}{4\eta} \frac{\partial p}{\partial z} r^2 + c_1 \, lnr + c_2 \, . \tag{8.22}$$

For $r = 0$, the coordinates are located in the center of the channel (Figure 98). If $r = 0$,

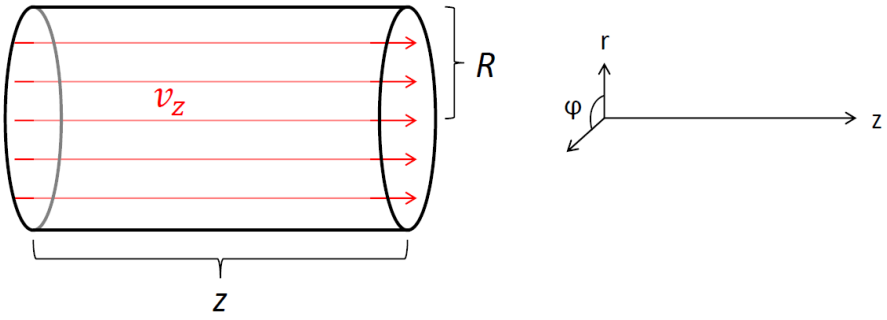

Figure 98: A cylindrical pipe, like a micro channel and the corresponding cylindrical coordinates (right).

the second addend in Equation 8.22 would run to infinity. However, we expect to have a finite value of $v_z(r = 0)$. Therefore, c_1 must equal zero. If we expect the flow velocity in z direction to be zero at the edge of the pipe, hence $v_z(R) = 0$, we can determine the second constant and finally obtain for both

$$c_1 = 0 \quad \text{and} \tag{8.23}$$

$$c_2 = -\frac{1}{4\eta} \frac{\partial p}{\partial z} R^2 \, . \tag{8.24}$$

The condition $v_z(R) = 0$ is called *no-slip condition* and is usually a good approximation. For $v_z(R) \neq 0$, one can approximate to which length $R + l_s$ one would have to extend the radius given the velocity profile so that $v_z(R + l_z) = 0$. The quantity l_s is called *slip length*. Combining Equation 8.22 with Equation 8.24, the velocity profile in z direction along the r axis equals

$$v_z(r) = \frac{1}{4\eta} \frac{\partial p}{\partial z} \left(R^2 - r^2 \right) \, , \tag{8.25}$$

yielding a parabolic profile.
We can calculate the mean velocity $\langle v_z(r)\rangle$ over the cross section $\pi\,R^2$ of the channel

$$\begin{aligned}\langle v_z(r)\rangle =& \frac{1}{\pi R^2}\int v_z(r)\,dA \\ =& \frac{1}{\pi R^2}\int v_z(r)\,2\pi r\,dr \\ =& \frac{1}{2}\,\frac{1}{4\eta}\,\frac{\partial p}{\partial z}\,R^2 = \frac{1}{2}\,v_z(r=0)\end{aligned} \tag{8.26}$$

and find that it is half the velocity at the center of the micro channel.
In order to calculate the time required to fill a liposome like illustrated in Figure 94, we need the flow rate or volume flow

$$\frac{dV}{dt} = \int_A \vec{v}\,d\vec{A} \tag{8.27}$$

through the cross section of the channel. Since the flow moves only in z direction (Figure 98), it is parallel to the normal vector on the cross sectional area and $\vec{v}\,d\vec{A} = v\,dA$. Thus, it is sufficient to multiply the mean velocity $\langle v_z(r)\rangle$ with the cross sectional area $\pi\,R^2$ and we obtain the volume flow

$$\boxed{\frac{dV}{dt} = \frac{\pi}{8\eta}R^4\,\frac{\partial p}{\partial z}}\,. \tag{8.28}$$

This relation is called *the law of Hagen - Poiseuille*[71] and it is worth to mention, that the volume flow increases with R^4.
With this relation, we can now calculate the time τ required to fill a spherical liposome of radius $\bar{r}$. If the liposome is filled with the fluid, it changes its volume by

$$\frac{dV}{dt} = \frac{d}{dt}\left(\frac{4}{3}\pi\,\bar{r}^3\right) = 4\pi\,\bar{r}^2\frac{d\bar{r}}{dt}\,. \tag{8.29}$$

Combining Equation 8.28 with Equation 8.29 and integrating from $t=0$ to $t=\tau$ leads to the result

$$\tau = \frac{32}{3}\,\frac{\bar{r}^3}{R^4}\,\frac{\eta}{\partial P/\partial z}\,. \tag{8.30}$$

In order to form a droplet like shown in Figure 94, the flow has to be decelerated by increasing the radius of the micro channel. Since mass (and therefore volume, if $\rho = const$) is conserved, the same volume $\pi\,R_1^2\Delta z_1$ at the part of the channel of radius R_1 equals the volume $\pi\,R_2^2\Delta z_2$ when the flow passes the channel at radius R_2 (Figure 99) in z direction at the time span Δt. Thus, the relation $\pi\,R_1^2\Delta z_1 = \pi\,R_2^2\Delta z_2$ must hold.
Since $\Delta z = v\,\Delta t$ the only variable that changes is the flow velocity. Therefore, the above relation equals $\pi\,R_1^2\,v_1\,\Delta t = \pi\,R_2^2\,v_2\,\Delta t$, hence it follows that

$$v_2 = \frac{R_1^2\,v_1}{R_2^2}\,. \tag{8.31}$$

If the radius increases linearly (Figure 94 and Figure 99), say with the function $R(z) = R_1 + a\,z$ (for a positive constant a), the flow velocity decreases by

$$v_2 = \frac{R_1^2\,v_1}{(R_1 + a\,z)^2}\,. \tag{8.32}$$

[71] Gotthilf Heinrich Ludwig Hagen, 1797 - 1884 and Jean Léonard Marie Poiseuille, 1797 - 1869

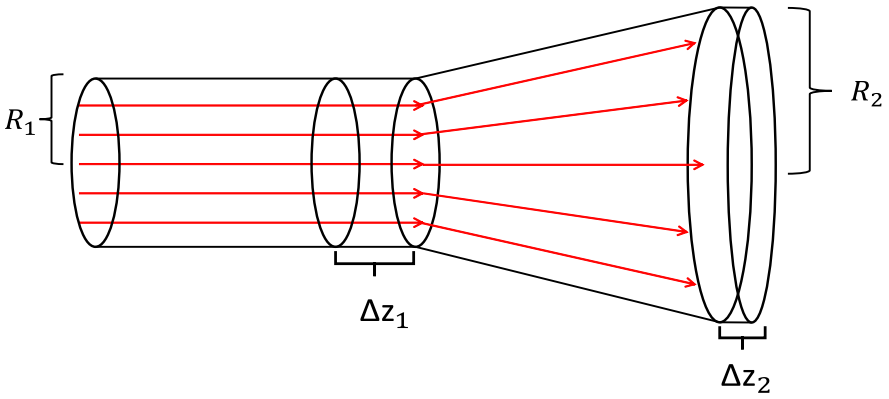

Figure 99: Due to the conservation of mass, the volume flow is decelerated if the radius of the pipe is increasing.

Depending on the shape of the channel, any function $R_2(z)$ can be inserted into Equation 8.31 in order to calculate the flow velocity.
If η is not negligible and for the no-slip condition, the velocity profile as function of r is not constant and follows Equation 8.25 leading to the law of Hagen - Poiseuille (Equation 8.28). However, if η **is** negligible, inertia dominates (large $\Re e$) and if there are no external forces, the NSE can be written as

$$\rho \frac{d\vec{v}}{dt} = -grad\, p\,. \tag{8.33}$$

If the flow velocity only changes by the geometry of the channel (i. e. by $R \neq const$) and not actively, $\frac{\partial \vec{v}}{\partial t} = 0$. With these constrains and for a flow in z direction only, the NSE can be written as

$$\rho \frac{\partial \vec{v}}{\partial \vec{z}} \frac{d\vec{z}}{dt} = -\frac{\partial p}{\partial z}\,, \tag{8.34}$$

where $\frac{d\vec{z}}{dt} = \vec{v}_z$. Note, that we can rewrite $\frac{d}{d\vec{z}}\left(\frac{\vec{v}^2}{2}\right) = \vec{v}_z \frac{d\vec{v}}{d\vec{z}}$, so that Equation 8.34 turns into

$$\frac{d}{d\vec{z}}\left(\frac{\rho\, \vec{v}^2}{2} + p\right) = 0\,. \tag{8.35}$$

We were allowed to change the partial derivative wrt z into the absolute derivative since all the quantities are only z - dependent. Integrating this equation leads to

$$\boxed{\frac{\rho\, \vec{v}^2}{2} + p = const}\,. \tag{8.36}$$

that is called *Bernoulli Equation*[72].
This equation is important since it states, that the sum of kinetic energy density of the fluid and pressure is constant. If the kinetic energy increases, pressure must decrease and vice verse. A fast flow lowers the pressure, whereas a slowly moving liquid corresponds to higher pressure.
Including gravity, the similar relation

$$\boxed{\frac{\rho\, \vec{v}^2}{2} + p + \rho g h = const} \tag{8.37}$$

[72] Daniel Bernoulli FRS, 1700 - 1782

can be found, where g equals the local gravitational acceleration and h the height of the fluid column. Therefore, the speed of flow can be measured by the height of the fluid column for constant outer pressure, like illustrated in Figure 100.

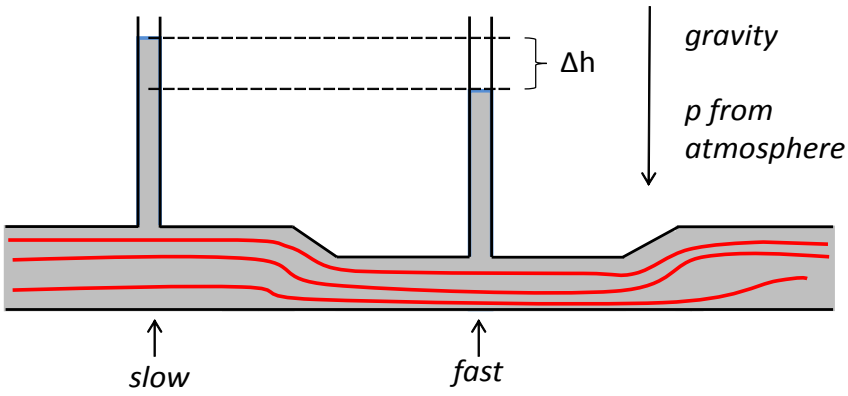

Figure 100: Bernoulli equation at work: if the radius of the pipe decreases, the flow velocity increases (Equation 8.31), leading to a drop of pressure in the pipe (Equation 8.36). If the outer pressure is constant, like in our atmosphere, this decrease of pressure can be measured by the reduced column height (Equation 8.37). The difference Δh can be increased even further for higher flow velocities.

8.3.1 Hydrodynamic Capacity and Resistance

Some quantities in fluid dynamics have some similarities to quantities in electronics and electrodynamics and therefore find their way into corresponding papers.
In electronics, the current I, measured in ampere A can be understood as a flow of charges, similar to the volume flow $\frac{dV}{dt}$ in fluid dynamics. The current I is driven by the voltage U, actually the voltage difference ΔU, measured in volts V. The analogue quantity in fluid dynamics is the pressure difference Δp. The strength of the current in electronics also depends on the resistance R (measured in ohms Ω) in the conducting medium. One can show in electrodynamics that the relation between R, U and I is

$$\Delta U \frac{1}{R} = I \,, \tag{8.38}$$

that is *Ohms law*[73].
In a similar way, we can suspect that a related law holds in fluid dynamics

$$\Delta p \frac{1}{R_H} = \frac{dV}{dt} \,, \tag{8.39}$$

where R_H denotes the *hydrodynamic resistance*. The full analogy in electronics is actually the resistance per length L, since this quantity really refers to the properties of the conducting medium and not to its absolute length. Rewriting the law of Hagen - Poiseuille (Equation 8.28) where the expression $\frac{\partial P}{\partial z}$ equals the pressure difference over the length L of the pipe, hence $\frac{\Delta p}{L}$, we find that

$$\Delta p \frac{1}{\frac{8\eta}{\pi R^4} L} = \frac{dV}{dt} \,, \tag{8.40}$$

[73] Georg Simon Ohm, 1789 - 1854

so that the hydrodynamic resistance

$$R_H = \frac{8\eta}{\pi R^4} L \tag{8.41}$$

and the resistance per length $R_h = R_H/L$

$$R_h = \frac{8\eta}{\pi R^4} \,. \tag{8.42}$$

A further analogy is the electronic capacity C and the relation

$$I = C \, \frac{d\Delta U}{dt} \tag{8.43}$$

that states how fast the voltage changes if it is evened out by the flux of charges, for example in a capacitor.
Since pressure changes if there is a flow in or flow out of particles, the analogous relation in fluid dynamics must be

$$\frac{dV}{dt} = C_H \, \frac{d\Delta p}{dt} \,, \tag{8.44}$$

with the *hydrodynamic capacity* C_H.
Combining Ohms law with Equation 8.43 leads to an exponential decay of the voltage difference in a capacitor

$$\Delta U(t) = \Delta U_0 \, e^{\frac{-1}{RC} t} \tag{8.45}$$

with the characteristic decay time $\tau = RC$. Note, that the flow rate $\frac{dV}{dt}$ works against the pressure gradient like the current I works against the voltage gradient so that it is an exponential **decay** if the gradients are positive and vice verse.
In the same way, we find the corresponding relation in fluid dynamics by combining Equation 8.40 and Equation 8.44 leading to

$$\Delta p(t) = \Delta p_0 \, e^{\frac{-1}{R_H C_H} t} \tag{8.46}$$

an exponential law for the pressure difference.
The meaning of the hydrodynamic capacity becomes clear when we combine the definition of compressibility (Equation 8.12b and Equation 8.12a) with Equation 8.44 leading to the relation

$$\kappa = \frac{1}{V} \, C_H \,. \tag{8.47}$$

This states, that the capacity equals a compressibility of the liquid or that it acts as a blind volume. If we release high pressure in a liquid, it partly expands (since κ is not exactly zero) causing an additional flow $\frac{dV}{dt}$, from that part of the volume that was compressed before. For example sea level would be roughly 30 meters higher if water would be completely incompressible. If we would release the ocean water somehow from the pressure caused by its own weight, sea level would raise by these 30 meters yielding an expanding volume, hence $\frac{dV}{dt} > 0$.
Finally, the last common analogy to electronics and electrodynamics is the *hydrodynamic permeability* σ defined via (c. f. Equation 8.28)

$$\frac{dV}{dt} = \frac{A}{\eta} \, \sigma \nabla p = -\frac{A}{\eta} \, \sigma \frac{p_{low} - p_{high}}{L} \,, \tag{8.48}$$

where A denotes the cross sectional area penetrated by the flow. For the law of Hagen - Poiseuille (Equation 8.28), the permeability equals $\frac{R^2}{8}$. If the radius is large, more material passes through the pipe and therefore σ is large. But permeability is not only a geometric factor. It also covers properties of the material *in* the pipe. For example, the pipe could be filled with soil, or a sponge or any porous medium, so that σ is reduced. Equation 8.48 was first found empirically by Dracy[74] but could be derived much later, after the NSE was found.

[74] Henry Philibert Gaspard Darcy, 1803 - 1858, therefore known as "Darcy's law"

8.4 Surface Tension

Expanding the liposome in Figure 94 costs energy that is provided by the kinetic energy of the volume flow. The work W required to expand a certain volume V equals $dW = p\,dV$ (recall that pressure p is an energy density). While the liposome is expanding, it increases its surface. Thus, the work must also be proportional to the surface area $dW \sim dA$ or with a yet unknown proportionality constant $dW = \alpha\,dA$. The liposome has almost a spherical shape of radius r so that the balance between volume work and surface work

$$p\,4\pi r^2\,dr = \alpha\,8\pi r\,dr \tag{8.49}$$

yields the relation

$$\boxed{p = \frac{2\alpha}{r}}. \tag{8.50}$$

The pressure p denotes actually a pressure difference between the interior of the liposome and the pressure outside of the liposome. The liposome is stable if it reaches a radius where this relation holds.
But why is there actually a resistance against the volume work while expanding the surface? Atoms and/or molecules in the liposome have six nearest neighbours. Each neighbour binds the molecule with a certain binding energy so that the total binding energy sums up to a certain value ϵ. A particle at the surface of the liposome however, has only five neighbours (see Figure 101) and the total binding energy sums up to only $5/6\,\epsilon$. This discrepancy causes an energy deficit of $1/6\,\epsilon$. If we denote the mean binding length as L we can define

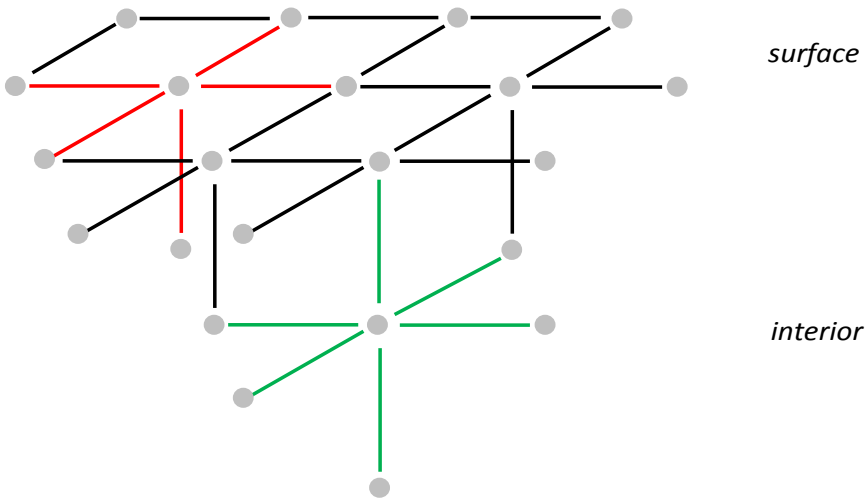

Figure 101: A particle in a solid or in a fluid has six neighbours (green) if it is located in the interior, whereas surface particles (red) have only five neighbours causing a binding energy deficit of $1/6$ of the total binding energy ϵ. This difference of binding energy causes surface tension.

the energy deficit per surface area $1/6\,\epsilon\,/L^2$. This deficit causes a tension on the surface particles so that our proportionality constant α is related to $1/6\,\epsilon\,/L^2$.
Of course, there are not always only six neighbours and there are also second order effects (e. g. contributions from particles in the extended neighbourhood) and the structure of liquids or solids is not that simple, but the key message is that there is always this energy deficit between surface particles and inner particles.
The surface tension α depends on the particular material and its environment. For example

$\alpha = 0.073\ N/m$ for water in air under biological conditions and the surface tension of a lipid bilayer varies between 0.08 and $0.3\ N/m$, depending on the environment (again for biologically relevant conditions). For generating droplets like in Figure 94 pressure differences around $10^3\ Pa$ are needed, i. e. 1% of the atmospheric pressure, leading to droplets of $100\ \mu m$ size. Thus, cells of $1\ \mu m$ size can withstand a pressure difference of roughly $10\ Pa$.
At a certain droplet size, gravity becomes important and the droplet looses its spherical shape. Surface tension and gravity are in balance if $\vec{f}_{grav} = -grad\, p$, hence

$$\rho\, g = -\frac{\partial}{\partial r}\left(\frac{2\alpha}{r}\right), \tag{8.51}$$

that leads to

$$r = \sqrt{\frac{\alpha}{\rho\, g}}. \tag{8.52}$$

Inserting typical values for α and ρ (water), the droplet looses its spherical shape if it exceeds radii of several millimeter, that is well above the size of a droplet ($0.1\ mm$) in a micro channel. Thus, ignoring gravity in the micro channel as external force as done in Section 8.3 is fully justified.

References

[1] L. R. Arriaga, E. Amstad and D. A. Weitz, *"Scalable single-step microfluidic production of single-core double emulsions with ultra-thin shells"*, Lab Chip, 2015, 15, 3335

[2] Siddharth Deshpande, Yaron Caspi, Anna E.C. Meijering, Cees Dekker1 *"Octanol-assisted liposome assembly on chip"*, nature communications, 2016, DOI: 10.1038/n-comms10447

9 Experimental Methods

I present a few experimental methods in this section that are widely used in biology, biochemistry and biophysics. In particular, I focus on how to beat the spatial resolution limit of visible light (super resolution) and the physics required to perform the different steps of image processing in electron microscopy. Also atomic force microscopy and X-ray crystallography are important experimental methods in life science and thus are briefly discussed here. This should give you a better grasp of how the outputs are generated, as well as how to interpret them. Furthermore, this section should demonstrate the importance of some of the physical methods you have learned to life sciences research.

9.1 Optics and Super Resolution

Before we begin with our discussion on light microscopy, we must first review several important properties of light in order to understand how "optical tricks" like super resolution work.

Light is electromagnetic radiation that propagates with speed of $c = 299\,792\,458\ \mathrm{m/s}$ in a vacuum. Amazingly, no matter what happens, the speed of light in a vacuum does not change. Even if you are moving parallel to a beam of light, its relative speed is still c! The speed of light, and its independence of any reference frame, are not only predicted theoretically, but are also well confirmed experimentally. **The speed of light in a vacuum is a fundamental natural constant.** However, light *is* affected by the medium through which it propagates. When traveling through air, water or some other medium, light propagates at a speed $\tilde{c} < c$. As we will see later, the amount at which c is decreased plays a significant role in how light refracts as it moves between media.

There has long been discussion on whether light is a wave or a particle. We now know that neither is true, and that by treating light as either a particle or as a wave eventually leads to contradictions. Light is a phenomenon that can only be described fully and coherently by a theory called "Quantum Electrodynamics". Some of the abstract equations yielded by this theory can be greatly simplified in certain situations. Under these conditions, the behaviour of light may appear to be analogous to more accessible concepts such as waves (like water or sound waves) or to particles (like small, fast moving balls). Thus, the so-called wave-particle duality of light is really just a way to describe quantum behaviour using more simplistic classical models.

As any chemist will tell you, molecules are not "balls", and they certainly aren't connected by any "sticks". It is often instructive, however, to treat chemical bonds as such in order to explain how many chemicals (statistically) interact, without having to account for why Xenon isn't actually inert when in a specific gaseous solution at $77K$. The point here is that light is neither a wave nor a particle, but we can be opportunistic in our experiments. If it is advantageous to describe light as a wave for a particular situation, then we do so. If, in another situation, the model of a particle is more appropriate, then we use this model.

Let us now introduce a few important equations for the two frameworks which we will primarily be considering. When treating light as a wave, we must be able to relate its wavelength λ and frequency f to its speed c. This simple and fundamental equation is given by

$$\boxed{c = \lambda f}\,. \tag{9.1}$$

When treating light as a "particle" (called a photon), we must be able to assign an energy E to it, despite it having no (rest)mass. The equation has been determined to be

$$\boxed{E = h\,f}\,, \tag{9.2}$$

where $h = 6,626\,069\,57 \cdot 10^{-34}$ Js is the Planck[75] constant.
Visible light makes up only a tiny portion of the full electromagnetic spectrum (Figure 102). The different colours are described by wavelengths between $\sim 400\,nm$ (blue) and $\sim 800\,nm$ (red), corresponding with energies (Equation 9.2) on the order of $1\ eV$ (two orders of magnitude greater than the thermal energy $k_B\,T$ at $300\ K$).

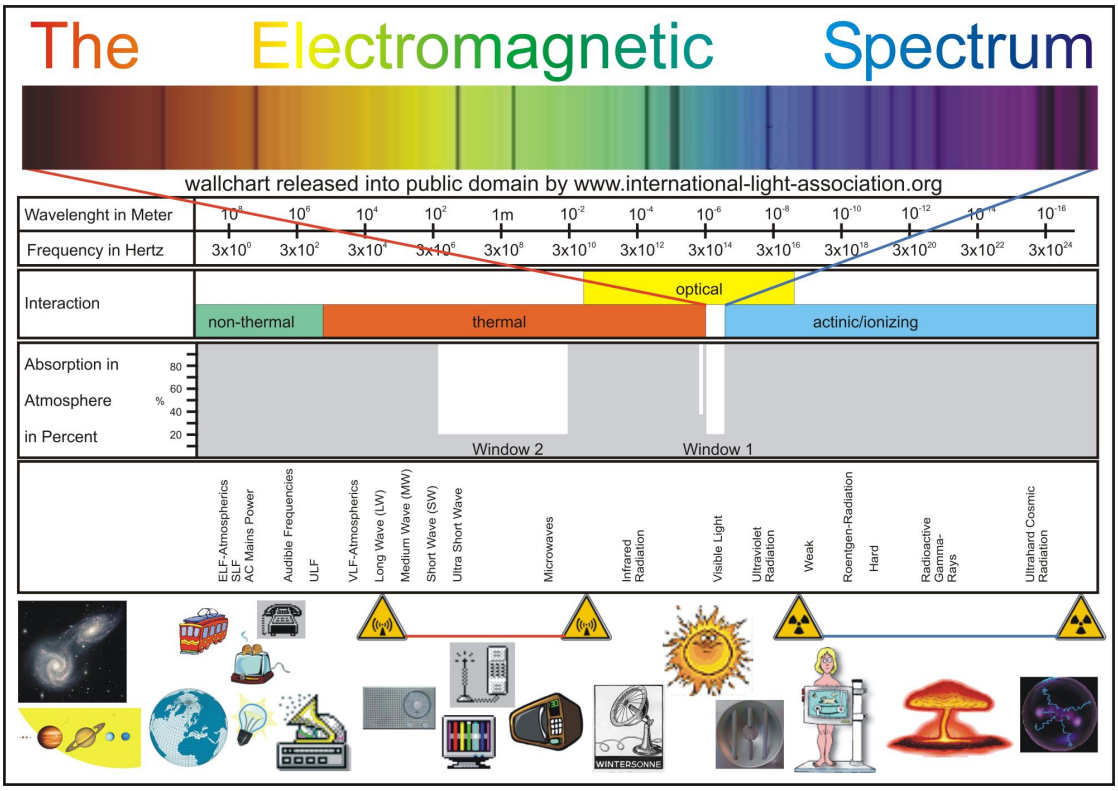

Figure 102: The electromagnetic spectrum from radio waves (left) to gamma rays (right). Image courtesy: `www.international-light-association.org`

9.1.1 About Waves

Usually, treating light as a wave is fully sufficient for our purposes, but we should always keep in mind that the "wave" is just a model. A wave is associated with a sine or cosine curve like

$$y = A_0\,\sin(\phi) + C\,, \tag{9.3}$$

where A_0 denotes the amplitude and C a possible offset. If the sine wave is a function of time t, the angle ϕ changes according to $\phi = \omega\,t$ with the angular frequency ω. Hence, while time elapses, the value of y changes periodically with the period T. If the wave is moving along the x direction, the argument in Equation 9.3 changes according to $\phi = x\,k$ where k is the wave number, the inverse of the spatial domain x (such as angular frequency ω is the inverse of time t; see also Section 2.4). Thus, the value of y can change periodically for different times t at given x or, while the wave is propagating in space, y can change periodically for different x a given t (like a "frozen" wave). Therefore, the angle ϕ is a

[75] Max Karl Ernst Ludwig Planck, FRS, 1858 - 1947

function of both and we obtain

$$y = A_0 \sin(x\,k - \omega\,t) + C\,. \tag{9.4}$$

For $x\,k - \omega\,t = const$ we obtain values of constant phase (and thus equal y) in the wave. Since the wave is periodic with the period T (that corresponds to 2π), the y values repeat after $t = T$ for fixed x and therefore one cycle is complete for $2\pi = \omega\,T$. Hence, we obtain the relation

$$\omega = \frac{2\pi}{T} = 2\pi\,f \tag{9.5}$$

with the definition of the frequency $f = \frac{1}{T}$.
If we fix time t and vary location x (like moving along a frozen wave), the y values repeat after each full wavelength λ, so that we find $\lambda\,k = 2\pi$ leading to the relation

$$k = \frac{2\pi}{\lambda}\,. \tag{9.6}$$

Combining this equation with Equation 9.5 and Equation 9.1, we find the important relation between speed, wave number and angular frequency

$$\boxed{c\,k = \omega}\,. \tag{9.7}$$

We can also introduce a phase shift $\Delta\phi = \nu$, that leads to the most general version of a sine wave:

$$y = A_0 \sin(x\,k - \omega\,t + \nu) + C\,. \tag{9.8}$$

The parameters defining a sine wave are illustrated in Figure 103.

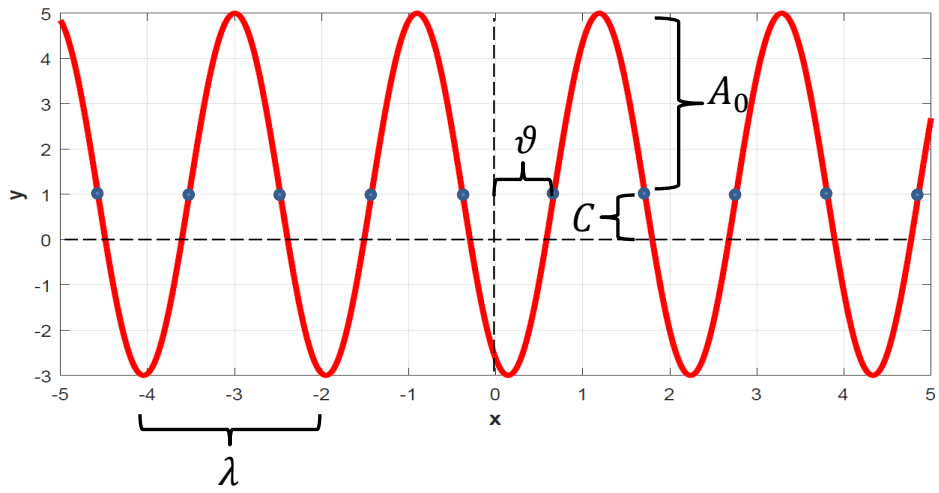

Figure 103: A sine wave and its parameters amplitude A_0, phase shift ν, offset C and wavelength λ.

We learned in Section 2.3.1 that sine and cosine are two different parts of the complex exponential and therefore, a wave function can be written as

$$A(x,t) = A_0\,e^{i\,(x\,k - \omega\,t + \nu)}\,, \tag{9.9}$$

where we set the offset C to zero for convenience.
Such a wave is called a *plane wave* since the regions of constant phase $x\,k - \omega\,t = const$ (for example the wave crests) are parallel lines, i. e. parallel *wave fronts*. A plane wave is

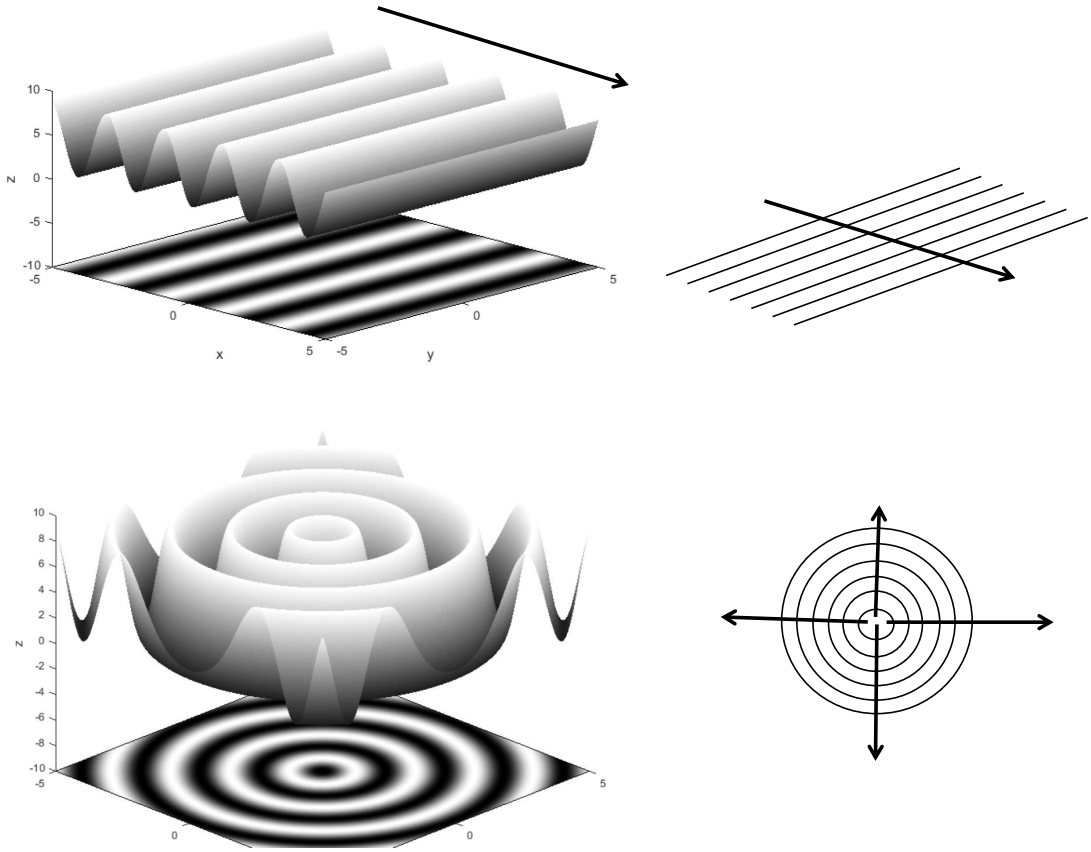

Figure 104: A plane wave (upper panels) and a spherical wave (in 2D) shown in the lower panels. Instead of visualizing the entire wave, usually only the regions of constant phase, for example the wave crests, are illustrated (here by black lines). The black arrows indicate the direction of propagation. Note that the distance between the black lines, i. e. the wavelength stays constant.

illustrated in the upper panel of Figure 104. When visualizing a plane wave in optics not the entire sinusoidal wave is illustrated, but the lines of constant phase.

While for a plane wave the regions of constant phase correspond to parallel wave fronts, the wave fronts for a *spherical* wave correspond to concentric rings (in 2D, Figure 104, lower panels) or concentric spheres in 3D. Since a sphere is defined by the radius $\vec{r}$ with $|r| = \sqrt{x^2 + y^2 + z^2}$, the corresponding wave function reads

$$A(\vec{r}, t) = A_0 \, e^{i\,(\vec{r}\vec{k} - \omega\, t + \nu)} \,. \tag{9.10}$$

If energy is conserved, the amplitude A_0 of a spherical wave must decrease with increasing distance r. In fact, the underlying physical theories lead to the more correct equation

$$A(\vec{r}, t) = \frac{A_0}{r} \, e^{i\,(\vec{r}\vec{k} - \omega\, t + \nu)} \,, \tag{9.11}$$

where $A_0 \rightarrow \frac{A_0}{r}$. Furthermore, what we sense and what is detected in experiments as *intensity* I is not A itself, but the absolute value of A squared, i. e.

$$I = |A(\vec{r}, t)|^2 = \frac{|A_0|}{r^2} \,. \tag{9.12}$$

Hence, we do not measure the amplitude A or A_0 itself, but the intensity I, where the information of phase $\phi = \vec{r}\vec{k} - \omega\, t + \nu$ gets lost. This lost of information becomes important for X-ray crystallography.

9.1.2 Diffraction, Refraction, Reflection and Scattering

Treating light as a wave should help us to understand phenomena like diffraction, refraction, reflection and scattering. If we consider light as wave, we could observe for example water waves and study their behaviour in order to draw conclusions for light waves. What happens for example if water waves pass an obstacle (e. g. rocks near a coast) The waves at the edges of the rocks generate new, circular waves that intersect behind the rocks and form a characteristic pattern. Such pattern are called *diffraction pattern*. The new circular waves are called *secondary wavelets* or *elementary waves* and the entire phenomenon is known as *Huygens*[76] *principle*. Thus, a definition for diffraction is

Definition: "According to the Huygens principle, diffraction is a result of wave interference".

We will see in the next pages that the Huygens principle is a very strong tool.
Just as water will not pass through a cliff, we see shadows when light is shone on a macroscopic object. However, if the size of this object is on the same order as the wavelength of the incident light (as the size of the water waves are on the order of the size of the rocks), then we see a diffraction pattern. A schematic visualization of the Huygens principle explaining diffraction patterns is shown in Figure 105. The distance between the wave crests is constant, indicating that the wavelength is constant, i. e. that we have monochromatic light (literally "having one colour") in Figure 105.
Reflection is a much more familiar occurrence than diffraction, yet it may be explained in

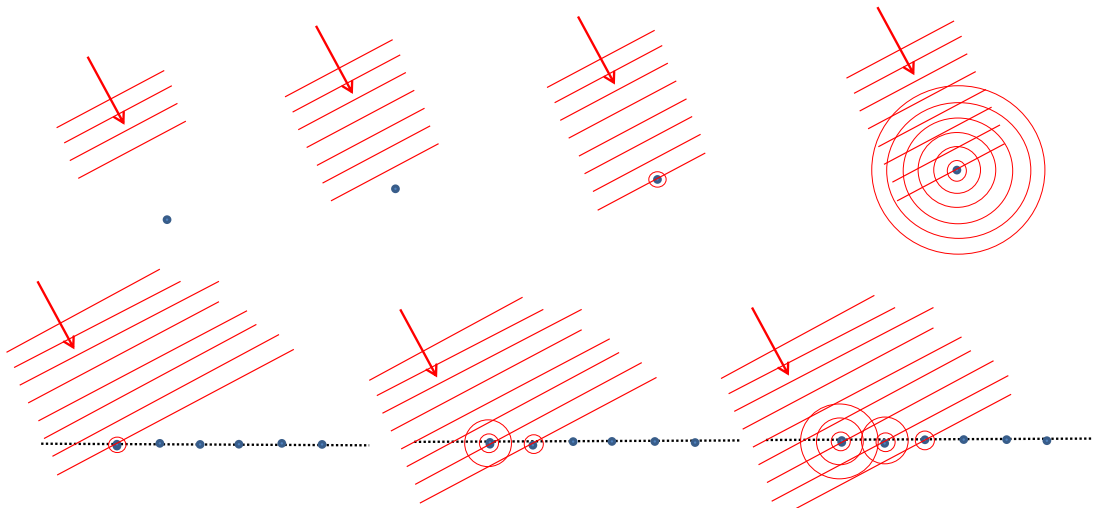

Figure 105: Diffraction explained by the Huygens principle: when a wave front (indicated by the red lines, with the red arrow pointing in the direction of wave propagation) hits an obstacle, e. g. light waves hitting an atom or water waves hitting a rock, it causes circular waves.

the same way: with the Huygens principle (Figure 105). Let us consider the waves emerging upwards from the obstacles in Figure 105. If we join the wave fronts of equal phase, we find a new direction of propagation (black arrow in Figure 106, left). Parts of the incident light thus changed direction and are now moving back away from the obstacles. The outgoing angle with respect to the normal of the surface plane of the obstacles is always equal to the incoming angle. This is shown in Figure 106, where the incoming angle is α and the

[76] Christiaan Huygens, 1629 - 1695

outgoing angle α'. This phenomenon is called reflection.

Definition: "Reflection is the change in direction of a wave front at an interface between two different media such that the wave front returns into the medium from which it originated."

When light penetrates an *optically dense*[77] medium of refraction index n, its speed is decreased to $\tilde{c}(n) = c/n$. This leads to a small change in the direction of propagation because a decreased speed results in a decreased wavelength (c. f. Equation 9.1). While perhaps not intuitive, this effect can again be easily explained by the Huygens principle (Figure 107, left). The effect of changing the direction of propagation when penetrating a medium is called refraction.

Definition: "Refraction is the change in direction of propagation of a wave due to a change in its transmission medium."

Let α be the angle between the direction of motion of an incoming beam of light and a line perpendicular to the border between two media of refractive indices n_1 and n_2, respectively. Now let β be the angle between the direction of motion of light and the line perpendicular to the border between the two media after the light has entered the second medium (Figure 107), one can derive the relation

$$n_1 \sin\alpha = n_2 \sin\beta\,. \tag{9.13}$$

The refractive indices are defined as the ratio of the speed of light in vacuum to the speed

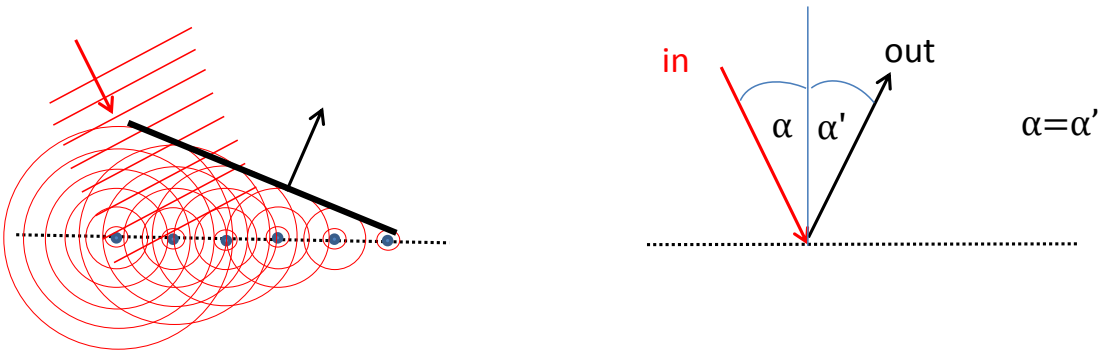

Figure 106: Reflection explained by the Huygens principle as part of diffraction. The surface plane is indicated by the black dotted line with its normal (blue vertical line).

of light in the medium ($n_1 = c/c_1$ and $n_2 = c/c_2$, respectively). Consequently, the refractive index of vacuum equals exactly $n = 1$ by definition. Note, that the refractive index is also a function of frequency. Blue and red light (and also radio waves and gamma rays) have different n for a given medium (except in vacuum, where n is always one). Refraction is the key mechanism involved in the focusing of light with lenses. A perfect lens would focus light with parallel wave fronts onto a single focal point.

So far we have assumed that an elementary wave is emitted immediately after the incident wave front hits an obstacle (hence no phase shift) and that there is a perfect energy transfer

[77] An optical dense medium is one with a relatively large refractive index n, like e. g. diamond. Note that optical density is not necessarily connected to mass density.

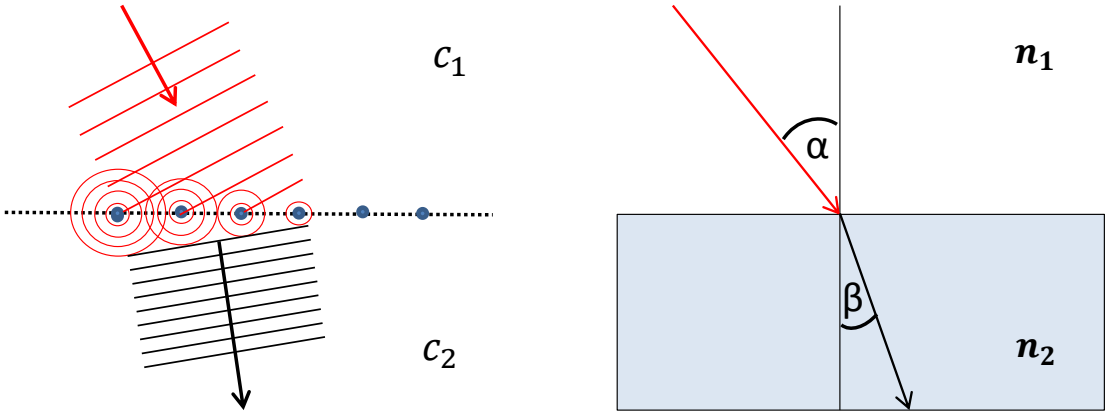

Figure 107: When light penetrates an optically dense medium its wavelength and speed c are decreased ($c_1 > c_2$) while its frequency remains constant (note the different distances between the red wave fronts and the black wave fronts on the left). This leads to a tilted wave front (c. f. Figure 106) within the medium and therefore an altered direction of propagation, i. e., refraction.
For the opposite situation, i. e. $c_1 < c_2$, the angle β is *larger* than α, since the wavelength in the second medium is larger now. In this case light moves from a optical dense into a less dense medium.

to the emitted wave (called *elastic scattering*). However, when a photon hits a particle it may lost some of its energy to its surroundings. Such an interaction is known as *inelastic scattering*. Depending on the initial energy of the photon and the particle it hits, the lost energy may excite atoms, or cause vibrational or rotational motion.
If a photon is scattered inelastically, the wavelength of the resultant light will be larger than it was before (compare Equation 9.1 and Equation 9.2). Scattering in terms of the Huygens principle is shown in Figure 108. Since the wavelength is changing and there is generally a phase shift, a new common wave front cannot emerge.

Definition: "Scattering is the deflection of the direction of propagation of the wave front from a straight path due to irregularities in the propagation medium, particles, or in the interface between two media."

If the sun is setting (or rising), the light rays must travel a greater distance through the atmosphere than during the day, thus increasing the amount of light that scatters inelastically off of debris (e. g., dust). The debris is heated by this interaction, taking energy from the light and effectively increasing its wavelength. As more light loses energy, we see more orange and red in the sky.
Consider the special case of *coherent* scattering, where the phase is conserved. This is exactly what we have been referring to in the previous subsections. In fact, X-ray diffraction in crystals can be treated as elastic, coherent scattering at the atoms of the crystal or as reflection at the crystal plane (see Section 9.2). Reflection, refraction and diffraction are simply different types of scattering - they emerge from the same physical principle.

9.1.3 Diffraction limits Resolution

Imagine a point source, i. e. an infinitely small black spot. To which extend can we resolve this object under a microscope? According to the Huygens principle, every light

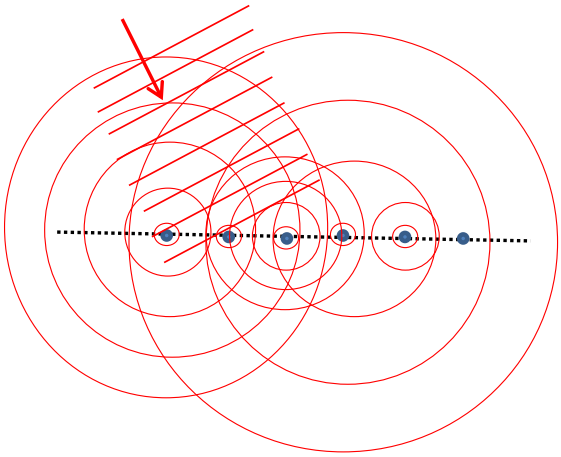

Figure 108: When a wave undergoes scattering, the emitted elementary waves are not in phase wrt the incident wave and they have a different wavelength. A common new wave front does not exist.

wave causes secondary waves when interacting with an obstacle. For macroscopic objects and the wavelength of visible light, these secondary waves form a common wave front right above the surface of the object so that diffraction processes are negligible, since the wavelength is small compared to the size of the object - we see the actual image of the object.
If we now shrink the object to a size comparable to the wavelength of visible light, diffraction effects become more prominent and the light waves do not just reflect at the object, they will cause a complex diffraction pattern due to the superposition of numerous elementary waves. One can calculate that the image of a point source, i. e. its diffraction pattern, would look like what is shown in Figure 109.
A point source like object, would not appear as a point with no extension, but as blurred circular object with weak concentric rings that get less visible towards the rim. The intensity profile of this object is called *point spread function (PSF)* (Figure 109), since diffraction "spreads" an actually dimensionless object to a certain size.

The PSF of a point source consists mainly of its central peak characterized by the full width at half maximum (FWHM) defining its broadness. Since the central peak hosts more than 99% of the intensity, the concentric rings are barely visible and only the main peak is apparent. Therefore, the complex structure of a PSF (see Figure 109, zoom in the upper right) is often disregarded and the central peak is usually approximated by a simple Gaussian. Often, the 2D projection of a point source PSF onto the $x-y$ plane (Figure 109, lower right) is called *Airy disc*[78].

If the wavelength λ of light is larger than the object, it can be treated as point source and we obtain a PSF. If the object is much larger than the wavelength, diffraction is negligible and we see the actual object. Hence, the PSF becomes broader with increasing λ. The PSF of a point source for different wavelengths and the overlapping PSFs of two point sources are shown in Figure 110.
Two point sources are considered "resolved" if their PSFs overlap at the FWHM (Figure 110, lower right). Since the FWHM of the PSF is a function of λ, the resolution is a function of λ too. Based on these considerations, one can show that the minimal resolvable distance

[78]Sir George Biddell Airy, 1801 - 1892

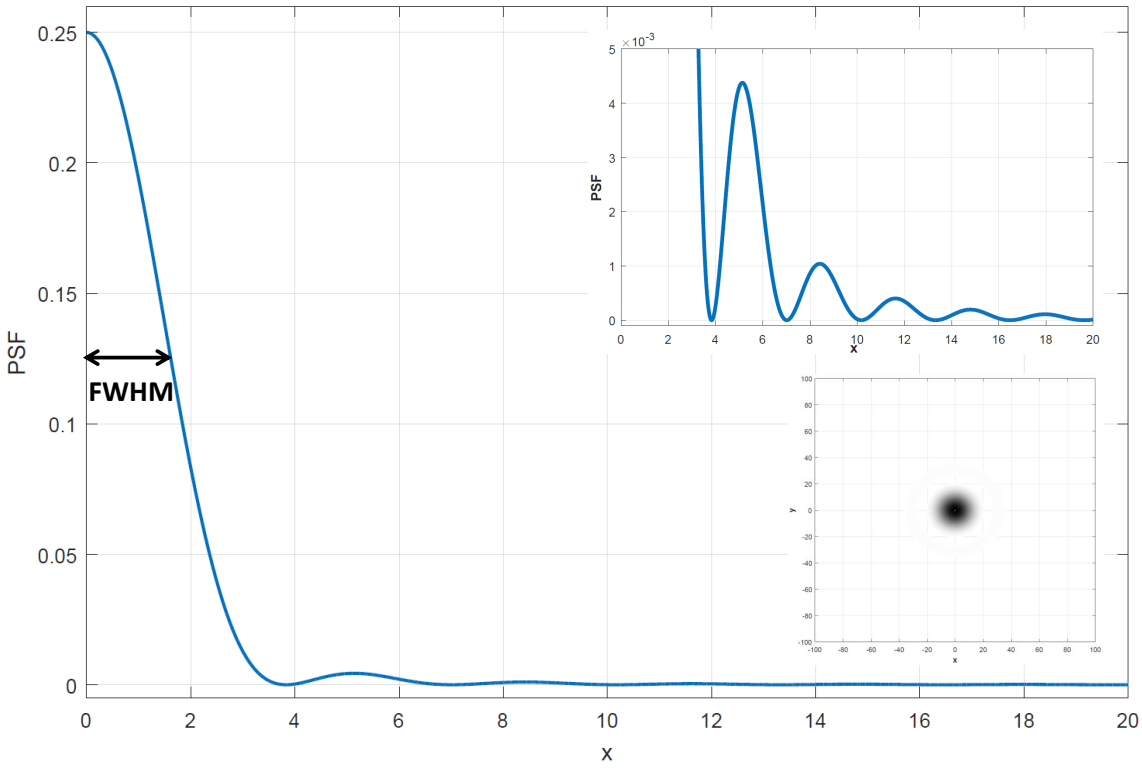

Figure 109: Due to diffraction effects a point source would appear as blurred circular object with concentric rings (lower right, rings are barely visible in linear scale). The intensity profile (PSF, blue graph) along a radial coordinate exhibits a strong maximum at the center and smaller peaks for larger distances from the center (see also the zoom in the upper right). The broadness of the central peak is characterized by the full width at half maximum (FWHM).

D between two point sources is

$$\boxed{D \approx \frac{\lambda}{2}}, \tag{9.14}$$

that was first found by Ernst Abbe[79] and is therefore called *Abbe limit*.
Hence, if we want to increase resolution, we have to decrease the wavelength. Thus, with visible light, D is in the order of $200\ nm$.

A resolution of $200\ nm$ is well sufficient for larger cells, but too less for organelles or viruses, not to mention even the largest proteins. An idea might be to construct an X-ray microscope ($\lambda \approx 10^{-9}\ m$). X-ray detectors are well sophisticated, but it is not possible to construct X-ray lenses of the required quality. The reason is that either absorption (e. g. in class for soft X-rays) is to high or the refractive index of almost any common material is very close to $n = 1$ at X-ray energies. Thus, an X-ray microscope would have a focal length of some kilometers. Some new materials have a higher refractive index at X-ray energies, but the lenses have such a bad quality, that the gain of resolution is lost due to optical distortions of the lenses.
There are many ways to circumnavigate this problem: X-ray crystallography (Section 9.2) uses the diffraction pattern directly to construct the actual image from a Fourier synthesis (Section 2.4) so that an X-ray microscope is not necessary.

[79] Ernst Karl Abbe, 1840 - 1905, who was the first one constructing microscopes from mathematical calculations and quantitative considerations. Before that, people manufactured microscopes in a more or less "trial and error" like fashion.

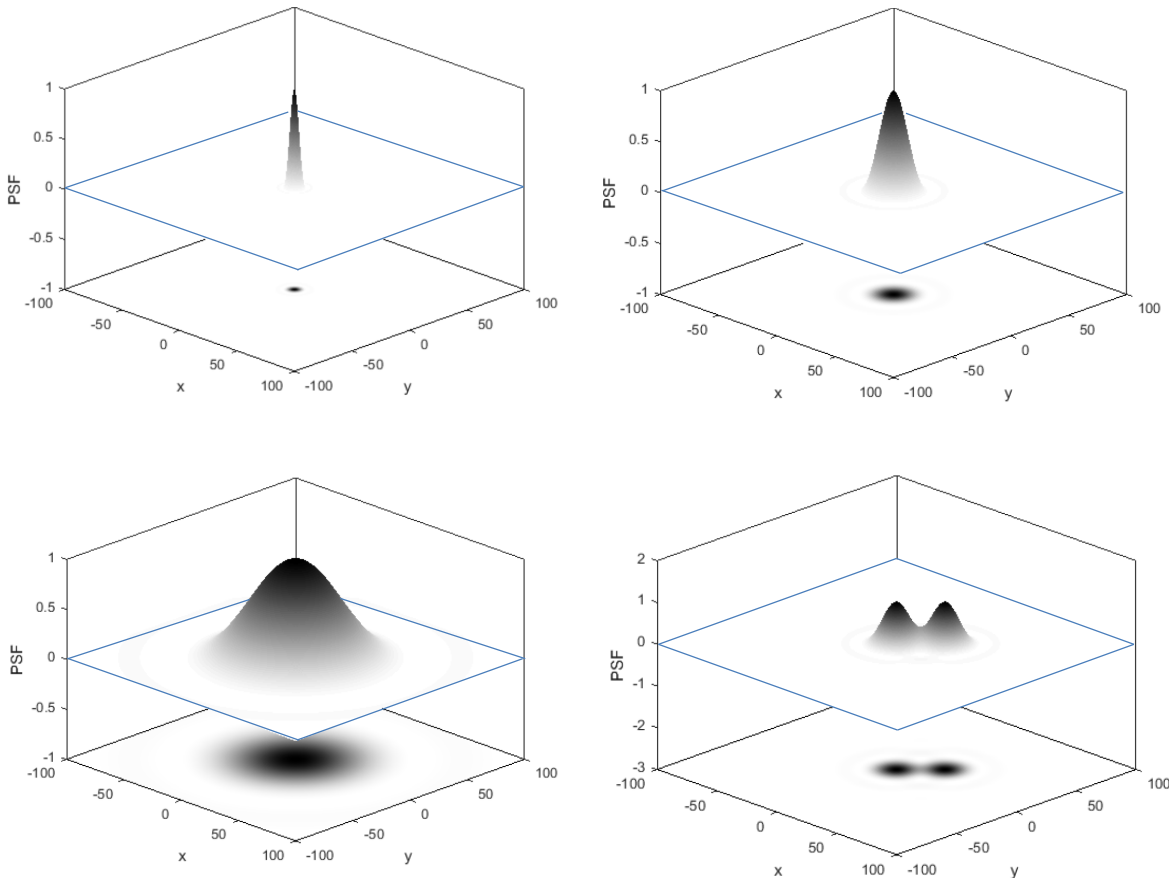

Figure 110: The PSF of a point source for different wavelengths λ and the PSF of two close point sources ($\lambda = 2,\ 6$ and 20 in arbitrary units, clockwise from upper left). Note, that the FWHM increases with increasing λ.

For electron microscopes (Section 9.3), the wavelength is easily adjustable, down to $\lambda \approx 10^{-10}\ m$ or even below and magnetic lenses are far easier to manufacture than X-ray lenses.
Another set of methods is summarized as super resolution, that is subject to the next sections. These are optical methods that reach below the diffraction limit (Equation 9.14) by using some optical tricks. They do not *violate* the diffraction limit (by definition, a physical law cannot be violated, otherwise it is not a physical law), but they use additional information (either from the PSF or from diffraction pattern) to construct the actual image.

9.1.4 Structured Illumination Microscopy (SIM)

The first super resolution trick is brilliant in its simplicity, utilizing the phenomenon of moiré[80] fringes. Many people will recognize moiré fringes from old TV screens. When there is a periodic pattern to be displayed, such as a fence in the background, the overlaying of the colour raster results in fringes being exhibited. This is known as the moiré effect. More generally, we see moiré patterns whenever two similar and repetitive semitransparent patterns are superimposed. Furthermore, when the patterns are rotated with respect to each other, the fringes change dramatically.
The idea is now to superimpose the image of an object that is below the resolution limit

[80] This time not named after its discoverer. Moiré is French for "marbled," like the appearance of rippled silk.

with a known pattern. This results in fringes that are larger than the object itself. From this moiré fringes and the known pattern that was superimposed, it is possible to derive the image of the original object (Figure 111).

Any image can be resolved up to a certain point and the resolution (Equation 9.14)

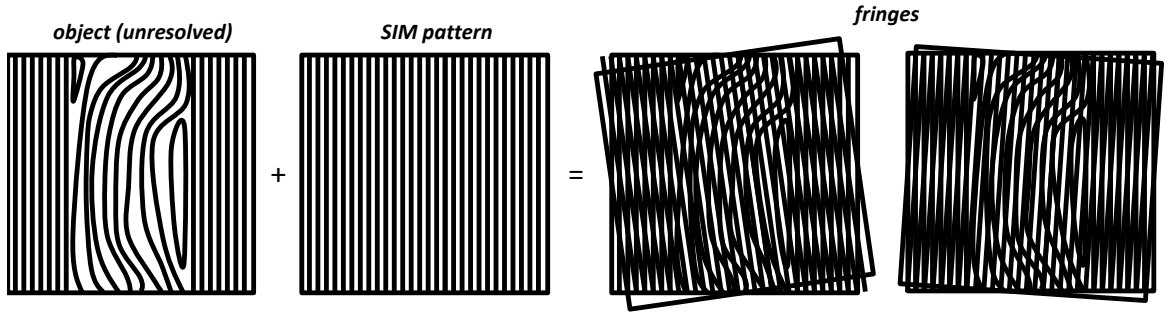

Figure 111: An unresolved object superimposed by a regular pattern causes fringes (so called moiré fringes) that are larger than the object itself. Since the fringes are characteristic for every individual object and the SIM pattern is known, the actual object can be reconstructed.

corresponds to a particular frequency k_0 in the inverse domain (Section 2.4) such that

$$k_0 \sim \frac{2}{\lambda}\,. \tag{9.15}$$

Every given spatial frequency k_i that is less than k_0 is resolved so that accessible information in an image is located within the circle of radius $r \leq k_0$ in the inverse domain, hence the Fourier image (Section 2.4, see also Figure 112, upper left). Superimposing a pattern with a fixed periodic structure (i. e. with the single spatial frequency k_1) causes fringes with a frequency of $|k_i - k_1| < k_0$ (see also [2] for details). Hence, information with a spatial frequency at $k_i \leq 2\,k_0$ has been made accessible so that the highest observable spatial frequency is now at $k_0 + k_1$ (Figure 112, upper right). Since k_1 is also limited by diffraction, the maximal spatial frequency that is accessible equals $2\,k_0$ so that the resolution (Equation 9.14) is improved by a factor of two. While we will see methods that far surpass the resolution of SIM, it is important to note that it still provides significantly better results than conventional light microscopes.

Turning the pattern, i. e. changing the angle of superposition, is equivalent to changing the angle in Fourier space without changing the spatial frequency (Section 2.4). In order to cover the entire inverse domain one must take images of the moiré pattern at multiple angles of superposition, usually in $60°$ steps (Figure 112, lower left), i. e. six to eight exposures per image.

To enhance the resolution even further, one can use a SIM pattern including higher harmonic frequencies of k_1 (Figure 112, lower right), so called *non linear SIM*. In principle, resolution can be increased to infinity but is bound by the signal to noise ratio in practice, since the intensity goes reciprocal with the order of the harmonics (c. f. Figure 20 in Section 2.4). In practice, the superimposed patterns are generated by a $1 \times 1\ cm$, 1700×1700 pixel screen. On such a screen, a pixel is below the μm range (recall that the wavelength of visible light is in the $0.1\ \mu m$ range). The advantage of a screen is that one can use different patterns. The light must be monochromatic and is therefore generated by a laser. The experimental set up (Figure 113) is still moderate.

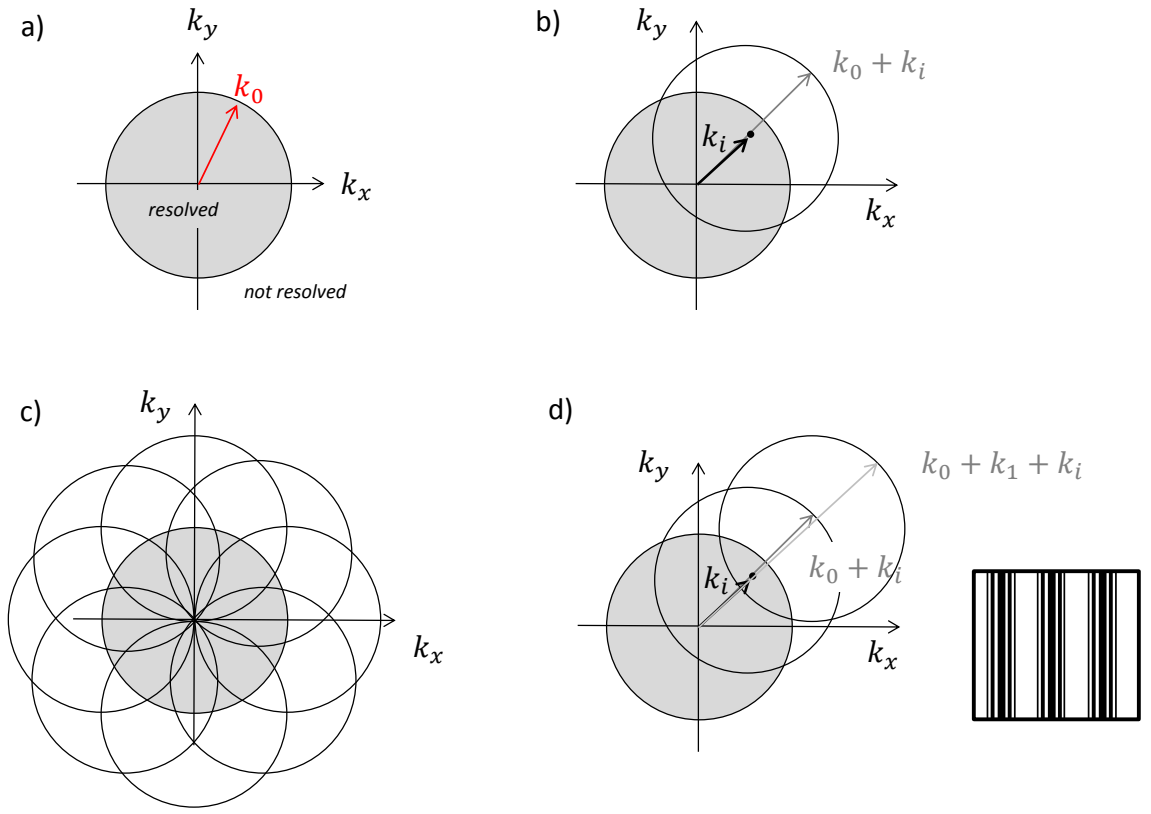

Figure 112: Accessible information in an image corresponds to a ring of radius k_0 in the inverse domain (Fourier space). Any information within this ring is resolved (a)). Superimposing a SIM pattern of spatial frequency k_0 shifts the k vector towards higher resolution (b)). In order to obtain the entire information, the SIM pattern is turned that is equivalent to changing the angle in Fourier space (c)). Higher harmonics in the SIM pattern can increase resolution even further (d)) without a hard theoretical limit.

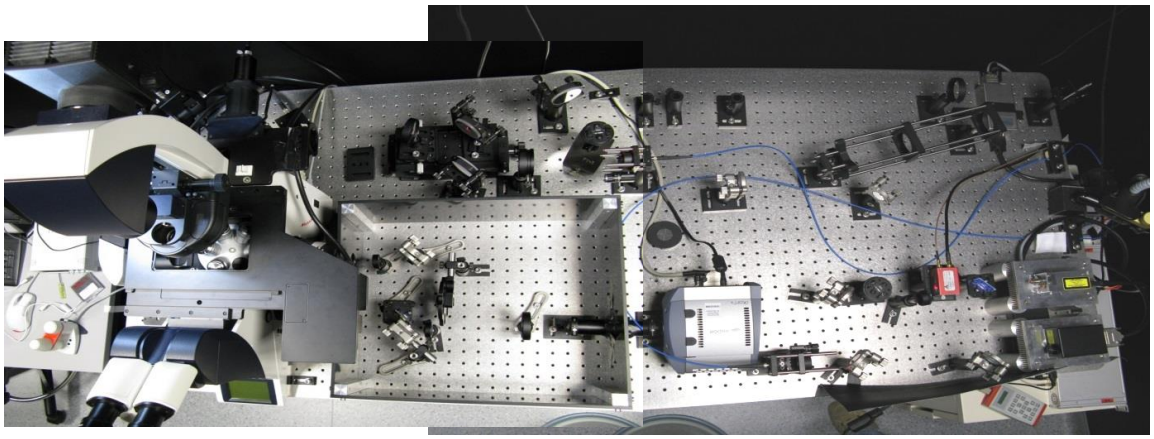

Figure 113: An experimental SIM set-up as used here in the genecenter (Image courtesy: Christophe Jung).

9.1.5 Stimulation Emission Depletion (STED)

In common microscopy fluorescent dyes may be excited by a laser with a particular wavelength. As the atoms relax to their ground state, they release photons with wavelengths longer than that of the incoming beam. For example, we may excite a dye with blue light, then observe green light being emitted. It is quite straightforward to filter the resultant

beam in order to obtain an image exhibiting only the emitted light. However, as we know, the image of the specimen has a spatial resolution restricted by diffraction. Even a perfect optical facility would yield an Airy disc as an image from a point source.
The idea of stimulated emission depletion (STED, see [3]) is that if the excited dye is illuminated with a separate beam (called the STED beam) of a specific wavelength, then the atom will instantaneously emit a photon and return to its ground state before spontaneous emission can occur. The wavelength required for stimulated emission is such that the energy of the STED photons matches the energy difference between the excited and ground states of the dye. If the excitation (EXC) beam is superimposed with a beam leading to stimulated emission, then the molecules are excited by the EXC beam before the STED beam causes them to immediately re-emit. However, if we tailor the stimulating beam such that its center (the STED spot) has zero intensity (a hole), then all of the molecules in this region will not be affected by stimulated emission but will instead undergo spontaneous emission later on.
The edges of the EXC Airy disc are thus cut out, leaving us with a resultant beam that has a better resolution than that dictated by the diffraction limit, see Figure 114. Of course, the diameter of the inner hole of the STED spot is also limited by diffraction, but since only a few photons are required to induce stimulated emission, no principle limit exists for the spatial (lateral) resolution of the resultant beam.
Such a STED beam can be generated by a phase plate, where the phase of the incoming light beam is continuously shifted from zero to 360 degrees in a clockwise (or counter clockwise manner), depending on the location of penetration. One can show that in such an set-up the phase shift in the center of the plate is undefined, corresponding to zero intensity. In theory, only the intensity of the the STED spot I_{STED} has to be increased in order to decrease the size of the resulting spot. A detailed calculation would show that Equation 9.14 would turn into

$$d \approx \frac{\lambda}{2} \frac{1}{\sqrt{1 + \frac{I_{STED}}{I_{Sat}}}}, \tag{9.16}$$

where the variable I_{Sat} denotes the saturation intensity of the particular dye.
With the facilities that are currently available, it is possible to reach down to a lateral spatial resolution of only $2.5\,nm$. This is more than a hundred times better than the ordinary diffraction limit (Equation 9.14)! STED is so powerful that even for moderate set-ups one could reach down to $\approx 20\,nm$ and therefore its inventor, the physicist Stefan Hell[81], was awarded with the Nobel Price in 2014.
However, there are some drawbacks to using STED. For instance, the method only works for special dyes (e. g. Rhodamine B, Atto 532 and Atto 590), as they must have specific energy levels and the life time of a stimulated excitation state must be much shorter than that of a spontaneous excitation state (c. f. Figure 114, right). An incorrect choice of dye could lead to the stimulating beam lifting the electron to a higher excitation state with a much longer lifetime. These conditions greatly limit the number of suitable dyes.

9.1.6 Photo Activation Localization Microscopy (PALM)

The decay of an electron from an excited state to its previous state in a fluorescent dye follows a Poissonian process (Section 2.6.6). If the dye is very dilute and the decay time is relatively large, one can monitor the blinking of single molecules as they emit a photon

[81] Stefan Walter Hell, 1962 – today

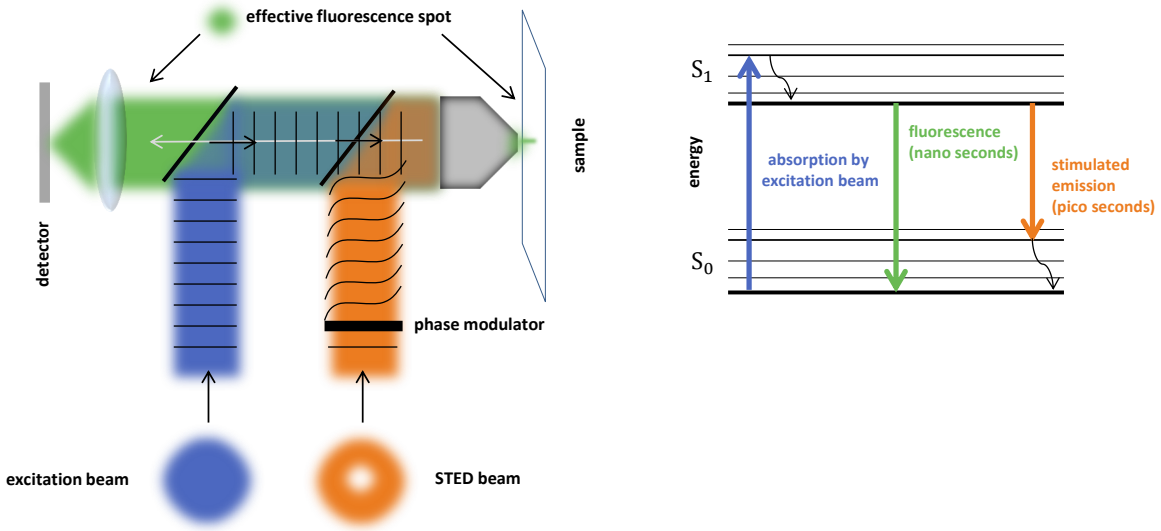

Figure 114: **Left:** Schematic view of a STED setup. The excitation beam would lead to the common fluorescence in the sample. Superimposing the excitation beam with a doughnut shaped beam that leads to simulated emission (hence, STED beam) reduces the extension of the fluorescence spot such that the final spot (green) has a much smaller FWHM than the initial excitation beam. Note, that still all beams are diffraction limited.
Right: Common fluorescence and forced (stimulated) emission have to be at different wavelengths in order to filter out the STED part. S_0 and S_1 denote atom orbitals of lower and higher energy, respectively. Energy levels are indicated by the black horizontal lines.

during relaxation.

Each blinking molecule is a point source and therefore results in an Airy disc diffraction

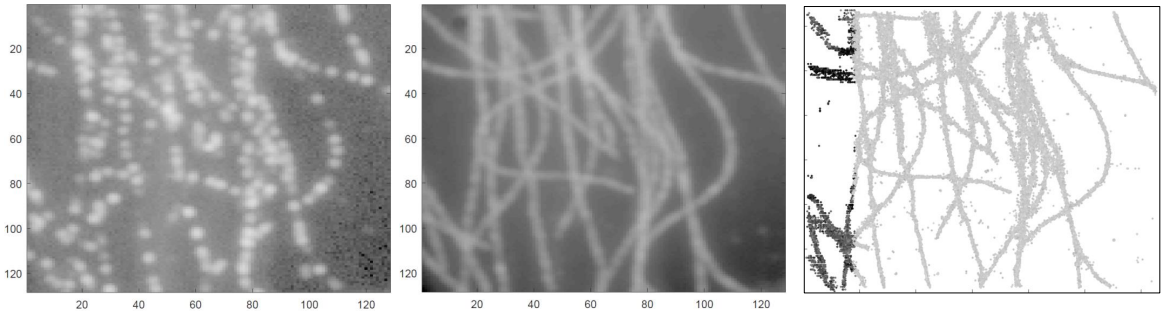

Figure 115: Fluorescence raw image of filaments (left), the co-added and noise filtered image (middle, from 500 individual images) and the final PALM image after fitting the PSF of each signal. Note the improvement of the resolution for the intersecting filaments in the center of the images. Although the images were processed with a simple, relatively short Matlab code, the resolution improved down to $70\ nm$ (right) compared to $200\ nm$ (left). Data Courtesy: Christophe Jung

pattern. The center of this disc can be fitted with a Gaussian emission profile, where the center of the Gaussian corresponds to the location of the molecule and the fitting error is much lower than the diffraction limited spatial resolution (see [1]). This procedure is then repeated for each blinking molecule.

In this way, one can monitor the locations of single molecules. The experimental set-up has to be adjusted in order to ensure that the probability of two subsequent emissions occurring within a short time is very low. To achieve this, one must either repeat the illumination of the specimen several times (after bleaching) in order to map the locations of the molecules

properly, or illuminate the specimen with very dim light in order to provide a continuous measurement. A PALM image of filaments is shown in Figure 115.
The major benefits of using PALM are that one can achieve resolutions comparable to those of STED with a relatively cheap experimental set-up. However, numerous images are required to monitor the locations of single molecules.

9.1.7 Comparison of Super Resolution Techniques

All super resolution methods have advantages as well as drawbacks. For a given study, the method that is chosen depends on the experimental requirements and on the (computational and financial) feasibility.
A small overview of the advantages and disadvantages of the three major super resolution techniques is provided in Table 1.

	+	−
SIM	• only nine images are needed • $\approx 100\ ms$ time resolution • no special preparation required • works with any existing fluorophore	• only $\approx 100\ nm$ resolution • computationally intensive
STED	• $20 - 30\ nm$ spatial resolution • no image post-processing	• temporal resolution $\approx$ minutes • special fluorophores required
PALM	• $20 - 30\ nm$ spatial resolution • relatively cheap	• requires large number of raw images • temporal resolution $\approx$ minutes • computationally intensive • requires photo activated fluorophores

Table 1: Advantages and disadvantages of the different super resolution methods.

9.1.8 Image Processing with Matlab

Whenever an image is generated, an image analysis or first an image processing has to be done before the actual scientific work follows. Even if the experimenter does not recognize it, often modern sophisticated cameras perform an internal image processing. The term image processing includes steps like normalization, enhancing edges, improving contrast or removing noise in images as well as actual image analysis like segmentation or pattern recognition. Be it if we want to remove noise in a cryo-EM image as a preparation for particle finding (c. f. Section 9.3) or if we want to identify cancerous cell structures in a stack of images showing tissue sections or if we just want to know number and size of cells in a given image. All these things can be automated with a much lower error rate and of course orders of magnitudes faster as if done manually.
For all these tasks, different labs use different software, which are however, written for more or less specific work flows. In principle all of these tasks can be performed with the *Image Processing Toolbox* in *Matlab*, especially if combined with the *Machine Learning*

Toolbox[82]. I therefore assume for this section that the reader has some basic experience with the programming language *Matlab*. For simplicity, I will usually use the default settings of the internal *Matlab* functions but will explicitly address changed settings if they become necessary.
Note, that image processing is not removing, adding or altering information – it is not "photoshopping" the data to something that we *might* want to find; it is more like extracting information in the sense of data mining and obeys (of course) the rules of scientific exactness and traceability.

Reading and Displaying the Image: Any common image format can be read in by the *Matlab* function *imread*, where the input is a string (name of the image including format extension) that returns a matrix I, e. g. like

```
I = imread('MyImage.jpg');
```

The matrix I is an $m \times n \times k$ matrix, where $m \times n$ refers to the size of the image in pixel. For example a four megapixel image might have $m \times n = 2160 \times 2160$. The variable k refers to the levels in the matrix. A gray scaled image (commonly denoted as black and white) has $k = 1$ since every pixel has just one value, often scaled between one (black) and 255 (white), whereas a color image has three levels ($k = 3$). The pixel values in the respective levels refer to red, green and blue, thus called *RGB image*. Every color can be composed out of a mixture of red, green and blue, such as the color vector in *Matlab* can point to any color.
For example the complete color image can be displayed by

```
figure;
imshow(I);
```

whereas

```
figure;
imshow( I(:,:,1) );
```

displays only red and so on. Figure 116 was generated by the commands

```
I = imread('imageColour.jpg');


figure;
subplot(2,2,1); imshow(I);
subplot(2,2,2); imshow(I(:,:,1));
subplot(2,2,3); imshow(I(:,:,2));
subplot(2,2,4); imshow(I(:,:,3));
```

[82]For example I created Figure 115 with a code I wrote in *Matlab* performing PALM.

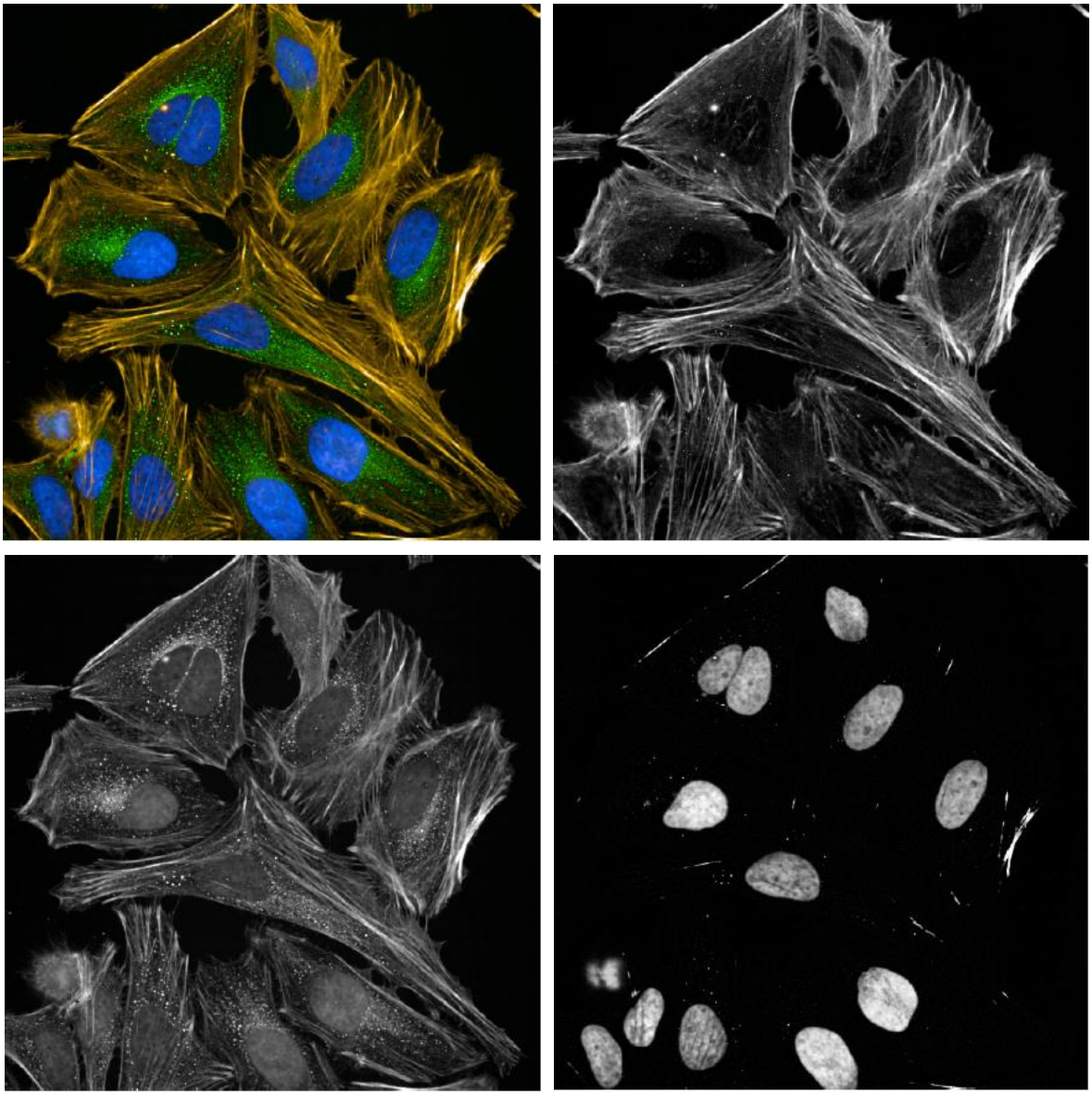

Figure 116: RGB color image and the same picture, but only in red, green and blue color channel (from upper left to lower right), respectively (Image courtesy: Christophe Jung).

which displays the whole color image (first figure) and the same figure, but red, green and blue only. This is in particular useful if we want to investigate objects of different staining, like the nuclei in this image. The color filtered images do of course not show the true colors, because $I(:,:,k)$ is also just a matrix containing some numbers and *Matlab* does not know whether the values refer to colors in an image. In principle, any matrix can be displayed as an image using *imshow*.
For some purposes we might want to turn a color image into a gray scaled image, which can be done by using the *rgb2gray* function:

```
1 I = imread('imageColour.jpg'); % reading image

3 Igray = rgb2gray(I);           % get corresponding gray scaled image
```

where the matrix $Igray$ is now a $m \times n \times 1$ image.

Noise Filtering: *Matlab* provides a variety of noise filtering functions. Every noise filtering method has its own advantages and drawbacks and the methods have to be chosen according to the properties of the images. What these methods have in common is that they take advantage over the fact that noise is random and (usually) uncorrelated, whereas an actual feature is not. A simple but very effective method to reduce noise is just co-adding the images if a number of exposures of the same object are available. Assuming that the different exposures are properly adjusted, the images can just be added like ordinary matrices. After dividing the final co-added image by the total exposure time, the noise should have been canceled out, since it undergoes random fluctuations around a mean value, whereas a feature would have summed up.
A code-block loading and adding N noisy exposures of one image, that are enumerated with "1", "2" etc in their name extension, might look like this

```
for i=1:1:N

      % loading images ''MyImage_noise_1.jpg'', 'MyImage_noise_2.jpg''
      % etc by using num2str in order to append index

      I_noise = imread(['MyImage_noise_' num2str(i) '.jpg']);

        if i==1

           I_noise_toadd =  I_noise;

        else

           % adding current image to added images
           I_noise_toadd = I_noise_toadd + I_noise;

        end

end

% dividing the final image by the total exposure time totExp
% (either known or derived from header readout)

I_added_final = I_noise_toadd / totExp;
```

I added up ten noisy images of the image shown in Figure 116. While a single image is extremely noisy, the situation has been considerably improved after co-adding (Figure 117).
A closer look on the co-added image in Figure 117 reveals that not all noise features are removed. Also in many cases only one exposure is available and an alternative method for noise reduction is needed.
One important property of noise is that the features have one or two pixel extension whereas actual features in the images are larger. Hence, in inverse space, noise is most prominent at higher spatial frequencies. A Fourier noise filter takes advantage over this property by transforming the image into inverse space and by removing spatial frequencies k that are larger than a given threshold k_0 (c. f. Section 2.4). The back transformed image then does not contain any information at these spatial frequencies anymore. A drawback is that if k_0

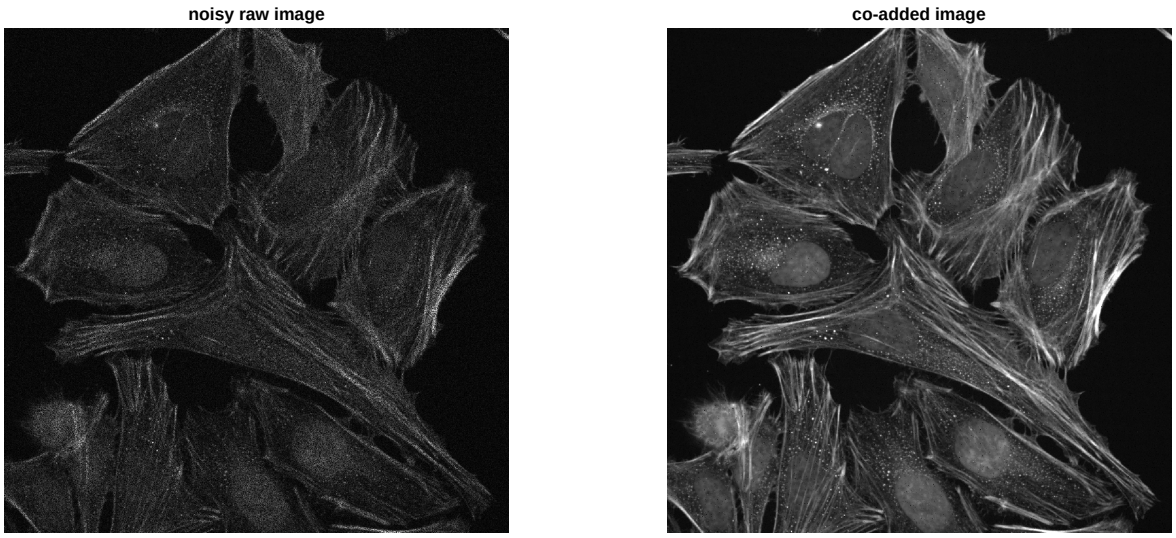

Figure 117: Noisy single exposure (left) and the final co-added image (right; ten single exposures).

is too small, we remove information from larger features, hence blur the images whereas if k_0 is too large, not enough noise might have been removed. There is always this kind of trade-off, but we can perform some useful estimations: one pixel has the spatial frequency of $k = 1^{-1} = 1$ (inverse space, c. f. Section 2.4), an object with two pixel extension has an spatial frequency of $k = 2^{-1} = 0.5$ etc. Thus, the image does not contain any information above $k > 1$, but if noise features have a size of one or two (or more) pixel extension, an appropriate threshold would be $k_0 = 0.5$, $k_0 = 0.4$ or even a bit less.
Fortunately, *Matlab* has an internal function, called *fft2*, **f**ast **f**ourier **t**ransformation, that performs the Fourier transformation of an image (the "2" in *fft2* stands for two dimensions) and we therefore do not need to perform the transformation explicitly. Let us transform the gray scaled image from Figure 117 (right) into the inverse space and display it:

```
1 [M, N] = size(Igray);          % determining the size of the image
      F = fft2(Igray,M,N);       % Fourier transformation into inverse
3                                % space
  figure; imshow(F);             % displaying the Fourier image
```

In line 2 we determined the size of the image in order to tell *fft2* when it has to stop integration (c. f. Section 2.4.2), since there is no spatial frequency below $k_{min} = (M^2 + N^2)^{-0.5}$. The Fourier transformed image of Figure 117 (right) is shown in Figure 118, left.

The pixel coordinates of the images start in the upper left at $(1, 1)$ according to the rows and columns in the matrix representing the image. The same is true for the Fourier transformed image, but it is more appropriate for the next steps to have the coordinate origin in the center of the image. Therefore, we apply the function *fftshift* by

```
1     F_shifted = fftshift(F);
```

and center the coordinate origin (Figure 118, right). Applying *fftshift* again would lead back to the original Fourier image F.
The center of the image has the coordinate $(N/2, M/2)$, if the matrix representing the

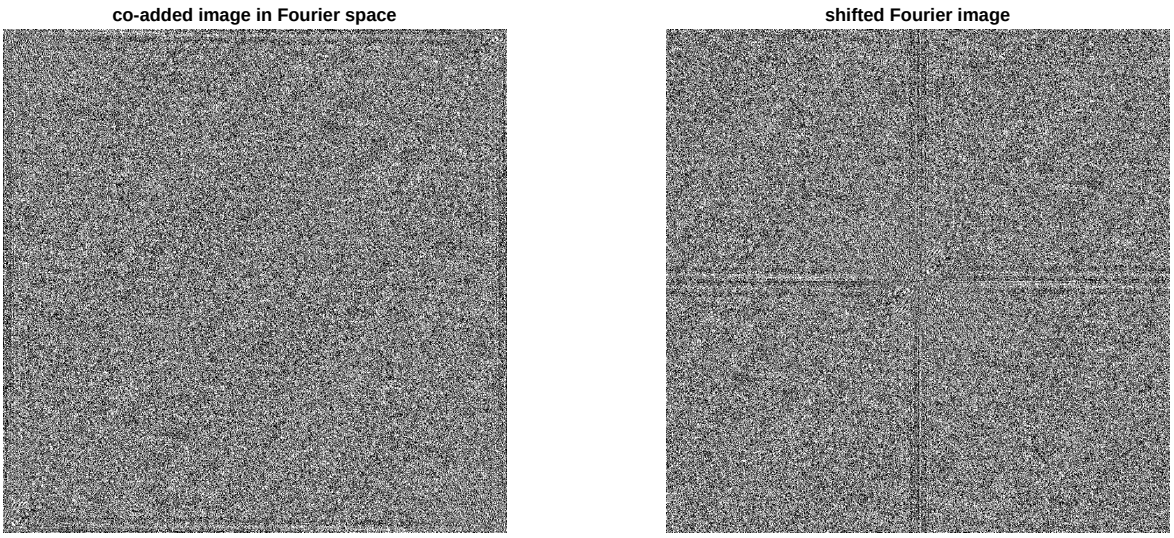

Figure 118: Fourier transformed image of the co-added image in Figure 117 (left) and the shifted image with the coordinate origin at the center of the image (right).

image has M rows and N columns and due to applying *fftshift* the image center corresponds to the coordinate origin of the spatial frequencies (Figure 118, right). Any point of the coordinate (i, j) therefore has the radius

$$r^2 = \left(i - \frac{N}{2}\right)^2 + \left(j - \frac{M}{2}\right)^2 \tag{9.17}$$

to the coordinate origin and hence $|r|$ corresponds to a particular spatial frequency in the Fourier image. We now have to measure r for every pixel (i, j) (that can be done by two nested *for* loops) and compare it to the threshold k_0 (Figure 119). If r is larger than k_0 (measured in pixel coordinates), we have to set the pixel value (i, j) in the shifted Fourier image to zero. We remove any information, that is supposed to be dominated by noise, for spatial frequencies above k_0 in the image. Finally, we have to shift the image back (by applying *fftshift* again) and we have to transform the Fourier image back into spatial coordinates by using the inverse Fourier transformation using the function *ifft2* (i stands for inverse). A corresponding code-block is shown below:

```
k_0 = 0.3; % threshold for frequency cut-off

% loop runs along rows (M) and columns (N)

    for i = 1:N
        for j = 1:M

            % measure the distance between image center and
            % current position
            r2 = (i - N/2)^2 + (j - M/2)^2; % radius r^2

            if  r2 > ( ((N+M)/2) * k_0)^2

                %set power of these frequencies to zero
                F_shifted(i,j) = 0;

            end
```

```
        end
    end

% transform filtered image back
I_fft = real(ifft2(fftshift(F_shifted)));
```

Note, that the pixel values in the image equal the amplitudes in the complex plane (c. f. Section 2.4.2) and we therefore only need the real part of the vector (last line in the code-block). The result of the noise filtering is shown in Figure 120. While the noisy pixel disappeared, the image got, however, slightly blurred.

Since the higher spatial frequencies are removed and only lower frequencies pass, the

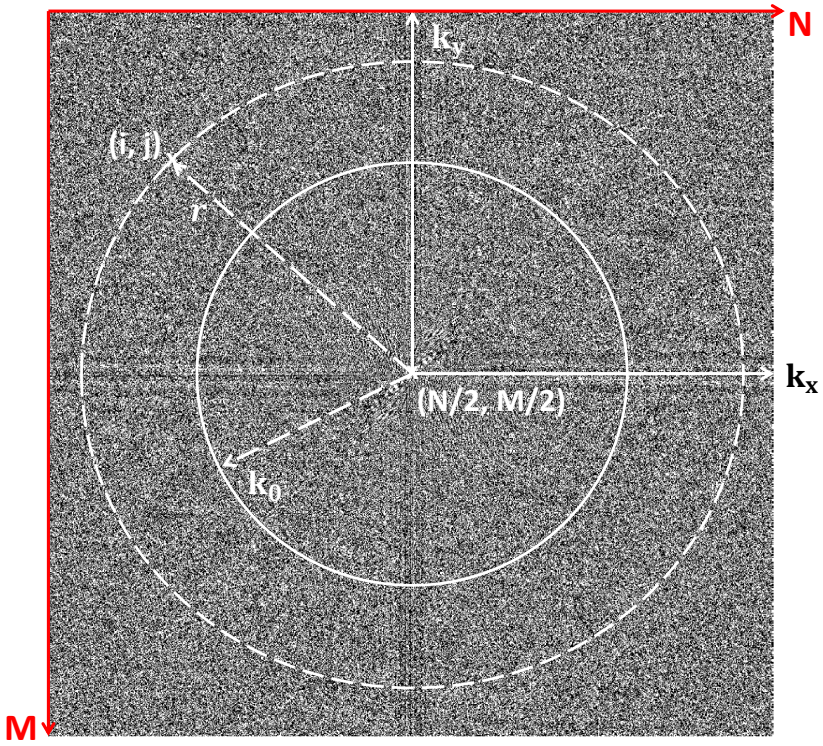

Figure 119: The location of a pixel in a $N \times M$ shifted Fourier image can be described by the two coordinates (i, j). The mode of the vector r pointing from the coordinate origin (having the pixel coordinates $(N/2, M/2)$) of the spatial frequencies k_x and k_y, respectively, to the location of the pixel corresponds to a spatial frequency that can be compared to a given threshold frequency k_0.

Fourier filter is a *low pass* filter.

There are other low pass filter like the median filter (*medfilt2*) and the Gaussian filter (*imgaussfilt*): The median filter calculates the median of the environment (3×3 pixel by default) of each pixel which is then the value output of the pixel. The idea behind this smoothing is, that the value of a noise pixel sticks out from its environment (that can be also seen in Figure 120, upper left) and therefore can be reset to the median of the values of its neighboring pixel. If the environment is chosen to be too large, the image appears blurred.

Pixel that are more distant to the reference pixel in a given environment can be weighted less in order to adapt the smoothing. If the weight follows a Gaussian distribution (Section 2.6.7), the filter is called a *Gaussian* filter. The weighting of the pixel can be controlled by the standard deviation σ. Also here, there is the off-trade between leaving too much

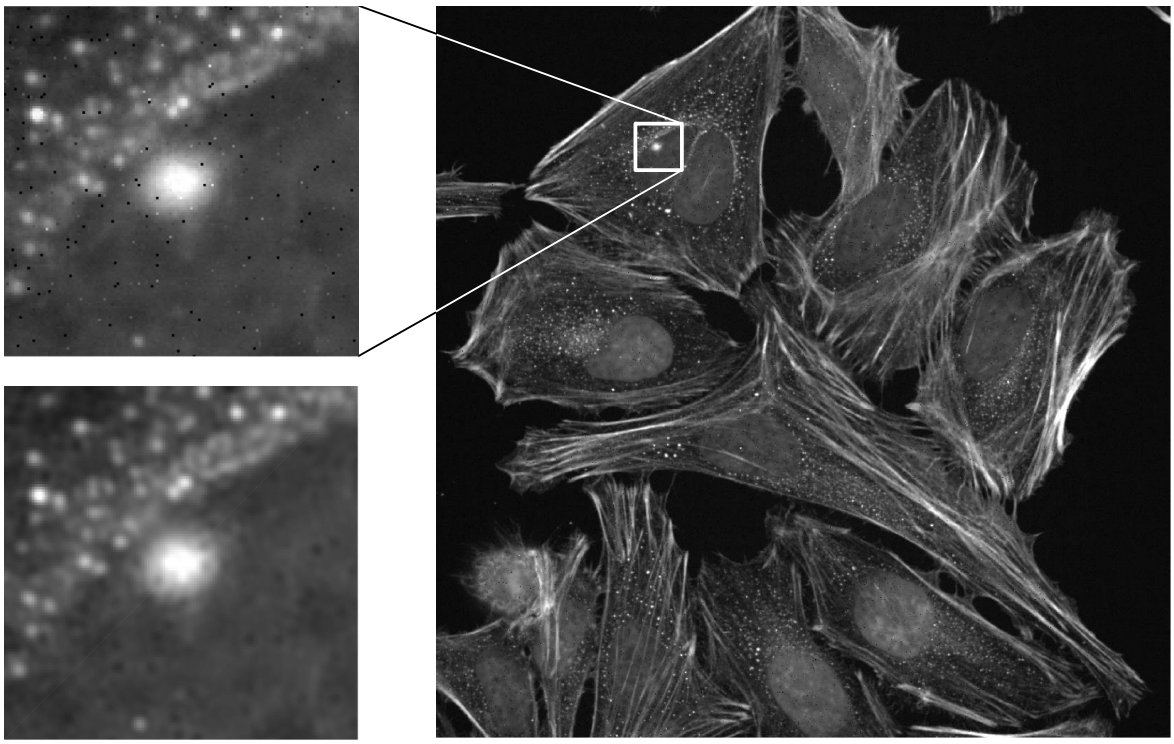

Figure 120: Zoomed region (white square) of the co-added image with noise features (upper left) and after Fourier noise filtering (lower left).

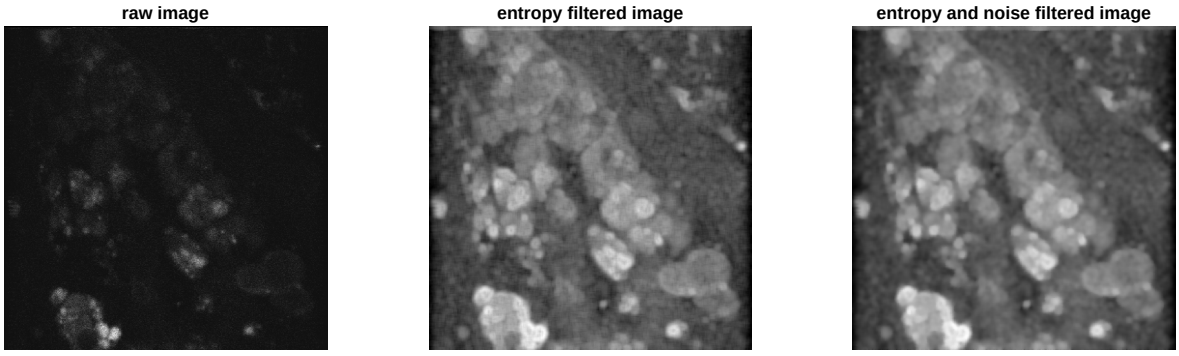

Figure 121: Raw image of a dendrite cross section (left) and after applying an entropy filter (middle) and a subsequent median filter (right). All filter were applied with the *Matlab* default settings. Image courtesy: Baccara-Jale Hizli.

noise (σ too small) or removing too much information (σ too large).
The problem of blurring can be partly avoided by using the entropy filter *entropyfilt*. A featureless region in a gray scale image that is superimposed by noise has high entropy since the pixel values are more or less uniformly distributed (Section 3.1). If there is an edge or any other feature in the region, the pixel values are not uniformly distributed and therefore entropy decreases (that is consistent with treading entropy as a measure of information and **not** as a measure of order or disorder, see Section 3.1). Thus, we can therefore measure the entropy within the environment (default setting is 9×9 pixel) of every pixel and assign it either to noise (high entropy) or to an actual feature (low entropy). This method is in particular useful when edges or particles have to be detected. At which entropy threshold a feature is regarded as noise or as actual feature and the optimal size of the pixel environment can be learned by the software when the *Image Processing Toolbox* is combined with

tools of the *Machine Learning Toolbox*.
An example where the entropy filter (default settings) has been applied to an image showing dendrite cross sections is shown in Figure 121.
The *Matlab* code-block for the entropy filter reads:

```
I_entr = entropyfilt(I);    % entropy filter

figure; imshow(I_entr,[]); % displaying
                           % the "[]" helps scaling the gray values
                           % for better visibility (but
                           % doesn't change the data)
```

Feature Detection and Segmentation: Pattern recognition or feature detection in images is called segmentation. Suppose we want to investigate a particular feature in more detail, like e. g. determining number and size of the nuclei shown in Figure 116. In this case, we of course could count the number of nuclei manually, but suppose you have a stack of numerous images, each containing a larger number of nuclei. Counting them manually would be time consuming and erroneous, not mention size determination.
In this example, the nuclei are blue stained and thus most prominent in the blue channel $I(:,:,3)$, see Figure 116 (lower right). Before we start with the actual segmentation, we remove the background features in the image by using the function *imopen*. But what is considered background? The nuclei appear as larger disc like objects, hence we remove everything that does not have this structure. This can be performed with the structure element function *strel*, that scans the image for objects having a certain structure (that can be a disc, diamond like shape, square, line, arbitrary etc) of a certain maximum size given in pixel. The nuclei are much larger than 30, 40 or 50 pixel radius, whereas the background features are smaller, scattered and not connected. Thus, removing the background with *imopen* using *strel* will work with

```
I = imread('imageColour.jpg');
Blue = I(:,:,3);                % nuclei are stained blue

R = 35;                         % radius in pixel
Str = strel('disk', R);         % disc shaped objects with less than
                                % 35 pixel radius

Blue_bkg = imopen(Blue,Str);    % removing objects with less than
                                % 35 pixel radius

figure;                         % displaying
imshow(Blue_bkg,[]);
title('blue channel (background subtracted)');
```

that generates Figure 122 (left). The nucleus of the dying cell (lower left) and the pegged nuclei in the upper left in Figure 122 (left) are clearly visible.
The actual pattern recognition part is performed by the function *bwconncomp*. The required input image must be a binary image (therefore the "bw") that is the case here since the image already contains only one color channel. The function *bwconncomp* returns a

list CC of connected components that have been recognized as features:

```
CC = bwconncomp(Blue_bkg);
```

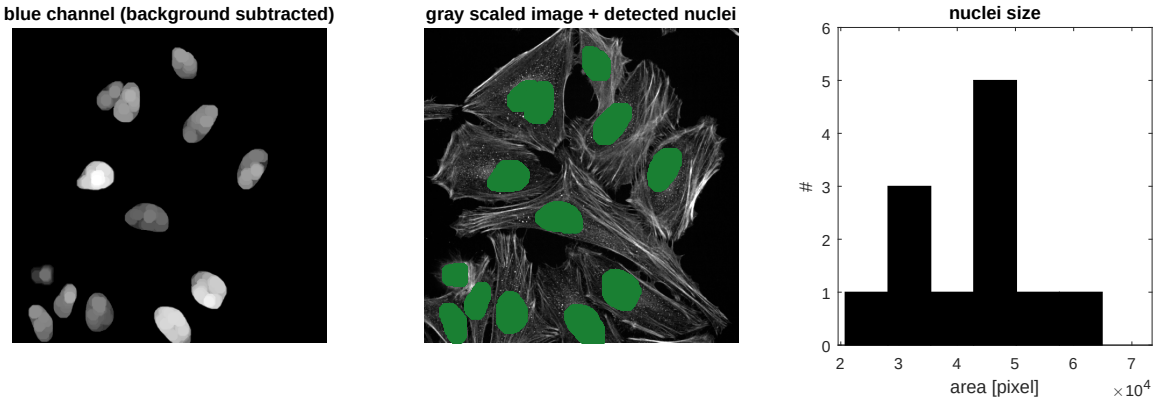

Figure 122: Nuclei from Figure 116, lower right (blue channel) after background subtraction using *imopen* (left) and superimposed on the original gray scaled image (middle). The size of the nuclei in pixel is shown on the right.

CC is a structure containing four fields among one, $CC.NumObjects$, yields the number of objects and $CC.PixelIdxList$ contains the identifiers (positive integers) of the pixel of the objects found by *bwconncomp*. By eye, we find thirteen nuclei in the image and indeed $CC.NumObjects = 13$. Hence, counting objects of a certain kind is not a difficult task!
For determining the size of the nuclei, i. e. the projected area seen in the image, one has to count the number of pixel belonging to a particular object. This can be done by generating a pixel list $PixelList$ of every feature/region that is assigned in $CC.PixelIdxList$ by extracting the region properties with *regionprops*

```
label = labelmatrix(CC);
S     = regionprops(label,'PixelList');
```

The variable $label$ is a matrix containing the identifier (here starting with 1 and ending with 13 since the function found thirteen regions) of each feature/region located at the corresponding pixel in the image. Hence, the size of $label$ equals the size of I.
The function *regionprops* returns a structure S where the number of fields equals the number of detected features/regions and each field contains the numerical matrix hosting the coordinates of the pixel belonging to the particular feature/region. Hence, each matrix contains two columns (i and j) and the number of rows equals the number of pixel making up the particular feature/region. Such a pixel list can be accessed via $S(1).PixelList$, $S(2).PixelList \ldots S(13).PixelList$.
The following code-block illustrates how the pixel coordinates can be accessed in a *for* loop and how are they used to over plot the original image with the detected features (as a cross check) and how a histogram of the area of the features found by *bwconncomp* is generated:

```
[A, ~] = size(S);       % size A of S equals number of features
```

```
M       = zeros(A,1);   % defining matrix of length A

figure;                 % displaying original image
imshow(rgb2gray(I),[])
title('gray scaled image + detected nuclei');

for i=1:1:A

    % extracting pixel list of feature i
    pxlList = S(i).PixelList;

    % determining number of pixel = area
    [a, ~]  = size(pxlList);

    % storing area of feature i
    M(i)    = a;

    % over plotting (this is one! possibility)
    hold on
    scatter(S(i).PixelList(:,1), S(i).PixelList(:,2),1,...
            'MarkerEdgeColor',[0.1 .5 .2],...
            'MarkerFaceColor',[0.1 .5 .2]);

end

% generating bar plot histogram
figure;
[where, val] = hist(M,round(A/2));
bar(val,where,'BarWidth',1,'FaceColor',[0 0 0])
ylim([ 0  max(where)*1.2 ])
xlim([ min(val)*0.8  max(val)*1.2 ])
xlabel('area [pixel]')
ylabel('#')
title('nuclei size');
```

This code-block generates Figure 122 middle and right.
Note, that these are only few examples of the large diversity of image processing tools provided by *Matlab*. An excellent overview with examples is given in [4]. The authors even provide their well commented source codes for free download.

References

[1] Eric Betzig et al., *"Imaging Intracellular Fluorescent Proteins at Nanometer Resolution"*, Science 2006 Sept; vol. 313, no. 5793, 1642-1645.

[2] Mats G. L. Gustafsson, *"Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution"*, PNAS 2005 Sept; vol. 102, no. 37, 13081-13086.

[3] Stefan W Hell, *"Toward fluorescence nanoscopy"*, Nature 2003 Oct; 21, 1347-1355 (2003).

[4] E Hodneland et al., *"CellSegm – a MATLAB toolbox for high-throughput 3D cell segmentation"*, Source Code for Biology and Medicine 2013, 8:16.

9.2 X-Ray Crystallography

Let us now return to the Huygens principle (Section 9.1.2) and discuss a plane, monochromatic wave hitting atoms that are arranged in a regular grid as shown in Figure 123. Such a symmetric arrangement is called a crystal. Let further the atoms be arranged in a lattice with spacing d between adjacent planes and let ν be the angle of incidence of the wave. Like in Section 9.1.1, the incoming wave leads to spherical secondary waves when interacting with an atom. We assume elastic scattering, so that energy is conserved and the wavelength of the incoming wave equals that of the emitted one. Furthermore, the secondary wave is emitted immediately, such that there is no phase shift between the incoming wave and the emitted wave (coherent scatter). Under such conditions, we can join wave fronts of equal phase and compare the paths of two waves with each other. The path difference between the wave interacting with the atom on the uppermost plane to the wave hitting the atom in the second plane equals $2\,\Delta s$ (Figure 123, lowest panel).

On the other hand, we can express Δs as $\Delta s = d\,\sin\nu$. We have a common wave front,

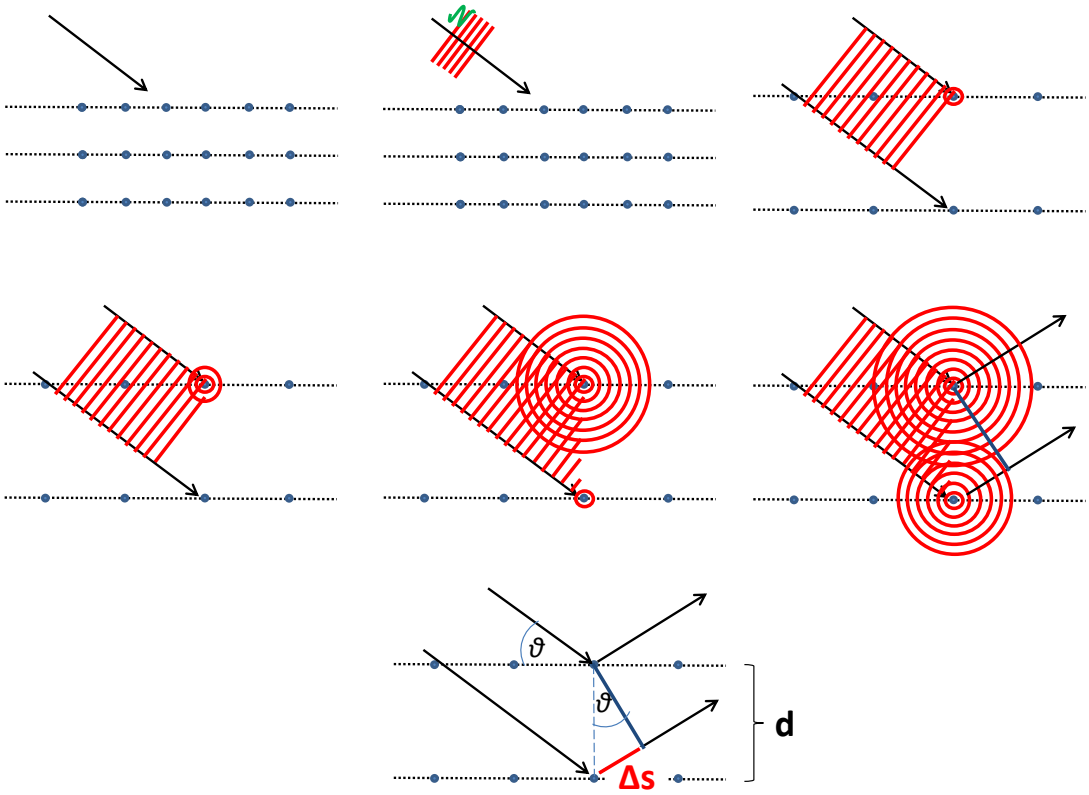

Figure 123: Diffraction in a crystal lattice explained by the Huygens principle. The elastic coherent scattering on the atoms in a crystal equals the reflection of rays on the crystal planes.

if the two rays are in phase, i. e. if the path difference $2\Delta s$ (see Figure 123) equals the wavelength λ times an integer m. If the two rays are in phase, they add up to a new common wave front of twice the amplitude. Hence, if we would place a screen behind the crystal, we would see a signal, if the condition

$$\boxed{2d\,\sin\nu = m\lambda} \tag{9.18}$$

is valid. This equation is known as **Braggs**[83] **law**.

A schematic experimental setup is illustrated in Figure 124. We can see in Figure 123 that

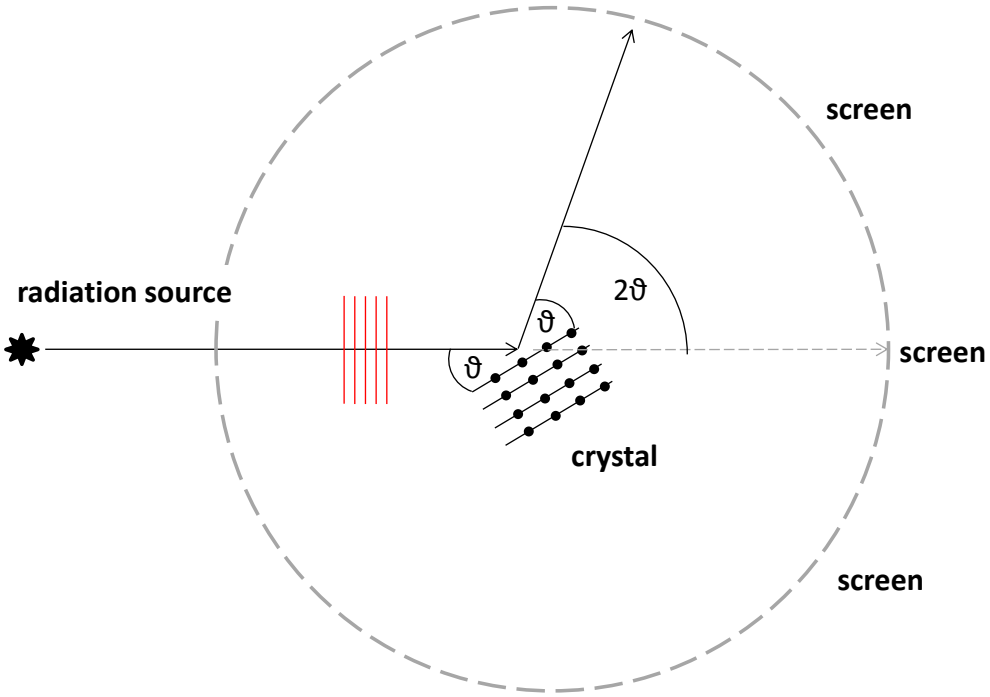

Figure 124: Schematic experimental set up for deriving the diffraction pattern of a crystal. The pattern are visible on the screen under any angle 2ν for different integers m if obeying Equation 9.18. The wave fronts are illustrated by the red parallel lines (c.f. also Figure 123).

we can join more than two different wave fonts that are in the same phase, so that we also would obtain different angles ν_m and therefore different m. Thus, there are many solutions for Bragg's law and we would expect to find a complex image of diffraction pattern on the screen.

Figure 123 also unveils, that the **elastic coherent scattering of light at the atoms of a crystal equals the reflection of light at the crystal planes**. The wavelength is known and set by the experiment, the angle ν can be chosen, hence we are able to determine the spacing d between the crystal planes. According to the construction in Figure 123, this only works, if the incident light is monochromatic, parallel (refers to the direction of propagation, i. e. also to the orientation of the wave fronts) and coherent (i. e. in phase).

We know today, that d is in the order of some $10^{-10}\ m$. Visible light has a wavelength of some $10^{-6}\ m$. Ideally, we have many reflections on the screen, where Bragg's law is fulfilled, so that we expect m in Equation 9.18 to be large. Thus, with this rough estimate, Bragg's law would read $10^{-10} \sin\nu \approx 10^{-5}$, or $\sin\nu \approx 10^{5}$. Since the sinus cannot reach values above one, Bragg's law is not fulfilled and we would not observe any diffraction pattern with visible light. Therefore, in order to obey Equation 9.18, λ has to be in the order of some $10^{-9}\ \mathrm{m}$ and thus the experiment can only be performed with X-rays.

If λ would be even smaller (e.g. gamma rays), the angle ν under which we would find diffraction pattern would be close to zero, even for larger m and thus the rays would just pass through the crystal without generating a useful pattern.

In Figure 123, we compare the phase shift between only two representative rays. However,

[83] Sir William Henry Bragg OM, KBE, PRS, 1862 - 1942 and his son Sir William Lawrence Bragg, 1890 - 1971 who, at the age of 25, became the youngest Nobel Prize laureate.

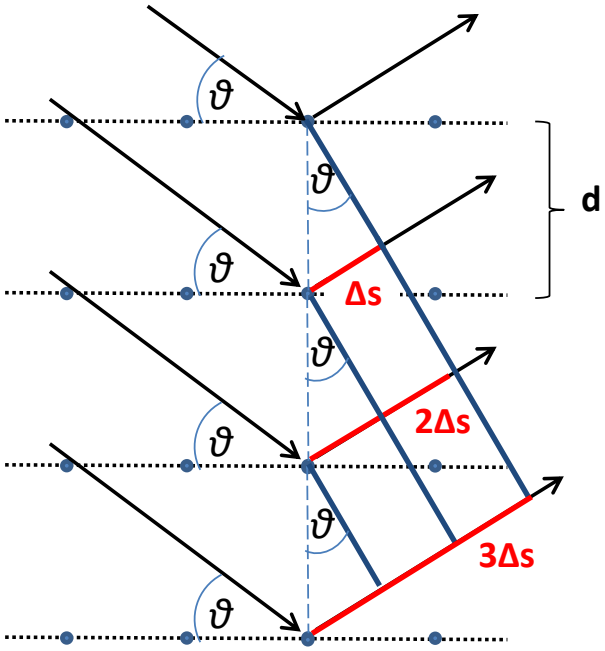

Figure 125: Diffraction in a crystal illustrated by several representative rays. The phase shift of the n^{th} ray equals $2\,n\Delta s$ wrt the upper most ray. For observing constructive interference, this phase shift has to be $m\lambda$.

in a crystal, or even in the tiniest crystallized protein, we have numerous atoms on which the rays can scatter. Since the crystal has a regular structure, also the phase shift $\Delta\phi$ caused by the different path length can be calculated in a straight forward manner (Figure 125): the first ray can be described by the wave function (Section 9.1.1) $\Psi_0 = A\,e^{i\,\omega t}$. Caused by the path length difference, the second ray has a phase shift of $\Delta\phi = \frac{4\pi\,\Delta s}{\lambda}$ (Equation 9.6) wrt the first ray and the third ray has a phase shift of $2\Delta\phi = 2\frac{4\,\pi\Delta s}{\lambda}$ wrt the first ray and so on (Figure 125). Thus, the total amplitude Ψ_{tot} measured at a particular point on the screen is proportional to

$$\begin{aligned}\Psi_{tot} &= A\,e^{i\,\omega t} + A\,e^{i(\omega t+\Delta\phi)} + A\,e^{i(\omega t+2\Delta\phi)} + A\,e^{i(\omega t+3\Delta\phi)}\ldots \\ &= A\,e^{i\,\omega t}\sum_{n=0} e^{i\,n\Delta\phi} \underbrace{=}_{Equation\ 9.18} A\,e^{i\,\omega t}\sum_{m=0} e^{i\,m\frac{4\pi\,d\sin\nu_m}{\lambda}}\,. \end{aligned} \tag{9.19}$$

Comparing the sum at the end of this equation to what we have learned in Section 2.4.2, we see that the image on the screen is **not the actual image of the crystal, but its Fourier transformation**. Thus, in order to derive the lattice spacing d and the actual structure of the crystal, we have to perform a Fourier synthesis of the diffraction pattern. We will return to this problem in Section 9.2.2.

9.2.1 The Crystal Lattice and Miller Indices

Even for the simplest crystal, the definition of a crystal plane and therefore the lattice spacing d in Equation 9.19 and Equation 9.18 is not unique. For a cubic crystal structure, there are fifteen different planes we must define to examine all of the spacings between atoms, that is illustrated in Figure 126. In order to derive the 3D structure of a crystal from its diffraction pattern, we have to take all these different planes into account. To tackle this problem, we must first be mathematically consistent in how we define the orientations

of such a plane in a given lattice.

The orientations of a given plane may be described in terms of a three-dimensional

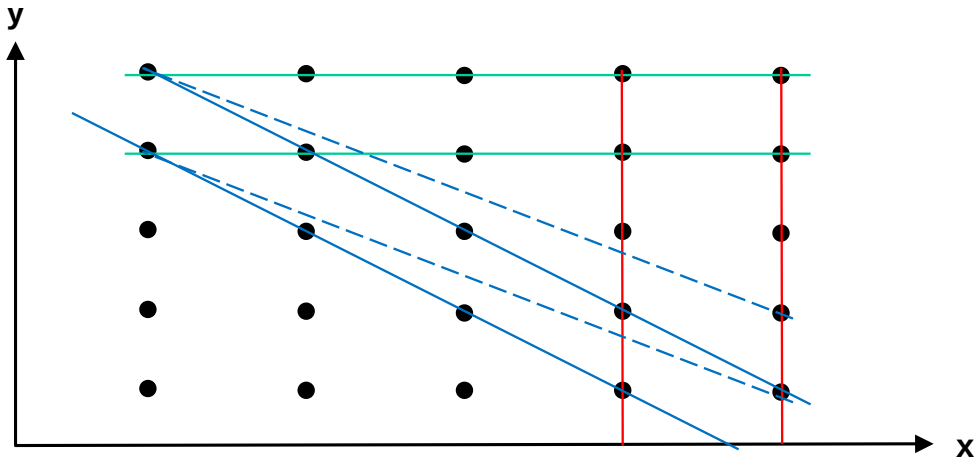

Figure 126: Possible representative crystal planes (lines) in a simple crystal. The atoms are indicated by black dots.

coordinate system. The more complex the geometry of a crystal lattice is, the more planes it contains. It is important to identify each plane, as each contributes to the refraction pattern and must be accounted for.

Let a be the spacing between the atoms in the x direction and b the spacing in the y

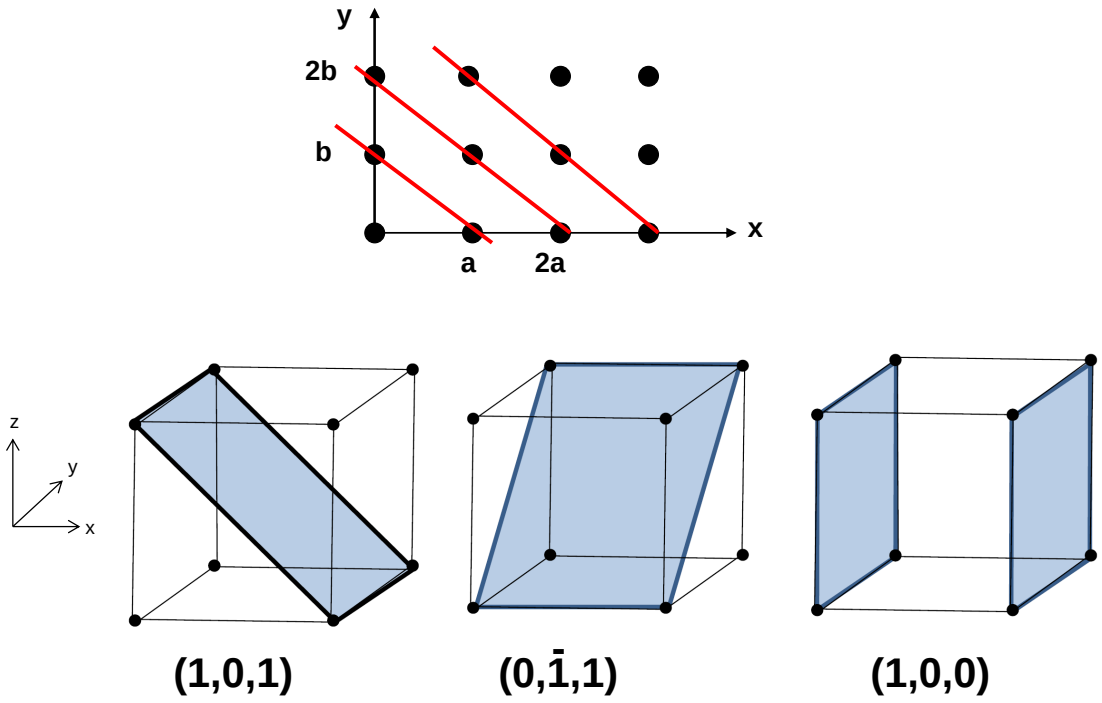

Figure 127: Top: Parallel lines (red) indicating a particular plane in a cubic crystal lattice. Bottom: Three of fifteen possible planes in a cubic crystal along with their Miller indices.

direction. The line $ay + bx = ab$ then intersects the x-axis ($y = 0$) at $x_0 = a$ and the y-axis ($x = 0$) at $y_0 = b$, the line $ay + bx = 2\,ab$ intersects at $x_0 = 2\,a$ and $y_0 = 2\,b$, and so on (Figure 127, upper panel). Dividing the equations for these lines by ab, we obtain

the form

$$\frac{x}{a} + \frac{y}{b} = n \tag{9.20}$$

for n parallel lines that intersect the coordinate axes at $x_0 = n\,a$ and $y_0 = n\,b$.
Extrapolating to three-dimensional space, we find

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = n\,. \tag{9.21}$$

Written in vector form, this is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \circ \begin{pmatrix} 1/a \\ 1/b \\ 1/c \end{pmatrix} = n\,. \tag{9.22}$$

The above equations describe a plane that intersects the coordinate axes at $(n\,a, 0, 0)$, $(0, n\,b, 0)$ and $(0, 0, n\,c)$.
Crystals are periodic, so the number n is an integer and thus the vector $\vec{v}_0 = (1/a, 1/b, 1/c)$ in Equation 9.21 tells us the orientation of the plane in the crystal. Note that $\vec{v}_0$ is orthogonal to the plane itself.

Exercise I:

Verify that Equation 9.20 is just the rearranged standard form $y = c_1\,x + c_2$ of a linear function. What are c_1 and c_2 in this regard?

The indices that make up $\vec{v}_0$ are known as **Miller[84] indices**, and it is a common convention to define them as $1/a = h$, $1/b = k$ and $1/c = l$. Using this convention, Equation 9.21 turns into

$$hx + ky + lz = n\,. \tag{9.23}$$

Miller indices are usually written as integers in lowest terms; i. e., their greatest common divisor should be one. For example, the indices $(1, 0, 1)$ represent the plane intersecting at $x = 1$, $y = \infty$ and $z = 1$, and therefore describe a plane that is parallel to the y axis (intersection at infinity) and that has a slope of -1 in the $x - z$ plane. I show a few more examples in the lower panel of Figure 127 (note that negative values are assigned with a bar, e.g. $\bar{1}$).
Utilizing Miller indices, we can now alter Equation 9.18 to account for orientation. One can show that the distance d between a point $P = (x_0, y_0, z_0)$ and a plane (Equation 9.23) is

$$d = \frac{|h\,x_0 + k\,y_0 + l\,z_0 - n|}{\sqrt{h^2 + k^2 + l^2}} \tag{9.24}$$

Exercise II:

Show that Equation 9.24 is indeed the distance between two adjacent planes. Hint: use the fact that $\vec{v}_0$ is orthogonal to the plane and that the distance between any point P and the plane is the length of the orthogonal vector pointing from P to the plane.

[84] William Hallowes Miller FRS, 1801 - 1880

In the case of a cubic lattice, all atoms along each coordinate axis are separated by the same distance a from each other. For parallel planes, then, we have $n = 1$, $n = 2$, and so on. Due to the periodicity of crystals though, the distance d between adjacent planes of the same $(h,\, k,\, l)$ will be identical (upper panel of Figure 127), and therefore it is sufficient to only calculate the distance between the coordinate origin $P_0 = (0, 0, 0)$ and the closest $(n = 1)$ plane.
Thus, the distance d between adjacent planes of any (h, k, l) reads

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}\,, \tag{9.25}$$

that is illustrated in Figure 128 four a 2D cubic lattice. One can show that there are fifteen

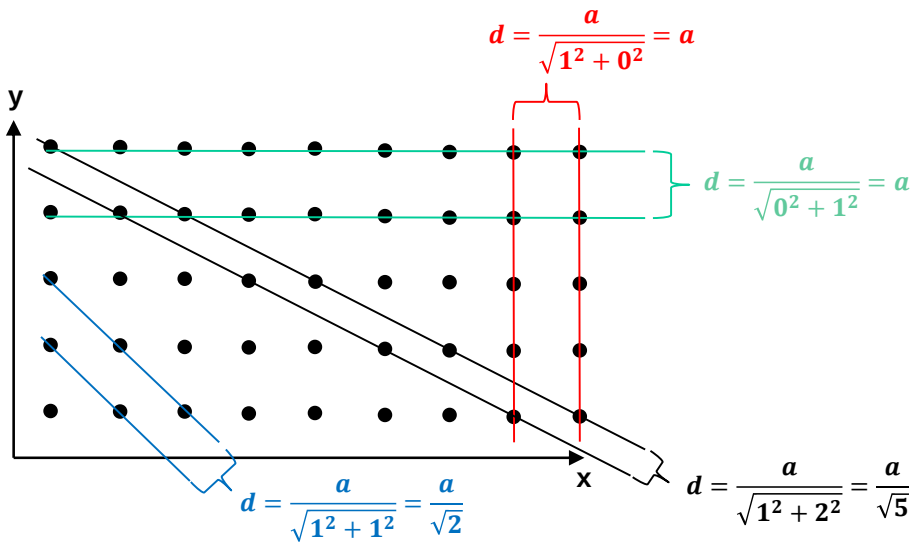

Figure 128: Four different representative lattice spacing in a (2D) cubic crystal with the lattice constant a. Note, that for a sufficiently large crystal, the number of different spacing (and thus different miller indices) is infinite.

main planes in a 3D cubic crystal (Figure 126 upper panel shows three of them) and that there are fourteen main crystal structures in 3D, so called *Bravais lattices*[85].
Using Equation 9.25, Bragg's law (Equation 9.18) becomes

$$\boxed{2\frac{a}{\sqrt{h^2 + k^2 + l^2}}\,\sin\nu = m\,\lambda} \tag{9.26}$$

for a **cubic** crystal.
For non cubic crystals or for non orthogonal lattices (e.g. hexagonal structures), Equation 9.26 looks more complicated and protein crystals are obviously not as geometrically simple as a cube, but the principle is always the same: Once a diffraction pattern is obtained, a software program derives the three dimensional structure of the crystal by performing a Fourier synthesis (Section 2.4). How this process works in detail is subject to the next section.

9.2.2 Image Formation

Let us now summarize, what we know about the diffraction process. Bragg's law (Equation 9.18) tells us under which conditions we observe diffraction pattern, but it does not

[85] Auguste Bravais, 1811 - 1863

yield any intensities. When a light ray scatters at an atom it actually scatters at the outer electron shell. Larger atoms like iron host more electrons than e.g. carbon or oxygen and therefore, the scatter will be more prominent. Hence, intensity matters if we want to solve the structure of a complex bio molecule. How to calculate intensity is given in Equation 9.19. This equation tells us that the image will be always a Fourier analyzed image of the actual object. This is also true in our daily life with visible light, but the ratio between d and λ in the last part of Equation 9.19 will be then close to zero ($\lambda_{vis} >> \lambda_{x-ray}$), so that the sum would give us

$$\Psi_{tot} = A\, e^{i\,\omega t} \sum_{n=0} 1\,, \tag{9.27}$$

that is the actual image itself. Hence, the effect of diffraction is negligible in this case. We discussed already in this regard in the last section that Bragg's law (Equation 9.18) does not yield a solution for visible light.
Since scatter actually occurs at the electrons, the final structure we obtain from a crystal equals an electron density map. If we denote the electron density $\rho(\vec{r})$ in an unit cell as a function of the coordinates $\vec{r}$ in the unit cell, we can weight the contributions to the total amplitude in Equation 9.19 by $\rho(\vec{r})$. Assuming that $\rho(\vec{r})$ is a continuous function of $\vec{r}$, the sum in Equation 9.19 becomes an integral over $\vec{r}$. Thus, the total amplitude at a particular location on a detector is proportional[86] to

$$e^{i\,\omega t} \int \rho(\vec{r})\, e^{2\pi\, i\, \vec{k}\vec{r}}\, d\vec{r}\,. \tag{9.28}$$

The volume over which we have to integrate in the above equation is actually the volume of the entire object, but since the crystal is repetitive, we only need to integrate over one unit cell and just obtain the contributions of N identical unit cells.
We know that every point in the crystal that yields a scatter amplitude that is not zero obeys Bragg's law (last part of Equation 9.19) and that any point at $\vec{r}$ located on a crystal plane must be expressed by Equation 9.23. If we furthermore introduce the vector $\vec{s} = \begin{pmatrix} h \\ k \\ l \end{pmatrix}$ and use the definition of $|\vec{k}|$ (see Equation 9.6), one can show after some algebra (that I like to omit here), that Equation 9.28 turns into

$$F(\vec{s}) := C\, e^{i\,\omega t} \int \rho(\vec{r})\, e^{2\pi\, i\, \vec{s}\vec{r}}\, d\vec{r}\,. \tag{9.29}$$

The above equation is called *structure factor* with a yet unknown proportionality constant C. For a perfect simple crystal, where all atoms have the same scatter properties, this equation turns into a sum again and we obtain

$$F(\vec{s}) := C\, e^{i\,\omega t} \sum_{j=1}^{J} \rho_i\, e^{2\pi\, i\, \vec{s}\vec{r}_j}\,, \tag{9.30}$$

where we sum over all atoms j and their location $\vec{r}_j$ in the unit cell. The quantity ρ_j is a scattering factor indicating the strength of the scatter. Usually, ρ_j and C are combined into one variable. The proportionality factor C depends on the experimental set up (strength of the x-ray source, distance to the detector etc.) and can be ignored since at the end only

[86] Note that this only works if absorption and multiple scattering is negligible so that the amplitude stays constant throughout the object.

relative intensities are important.
What we measure in the experiment is $F(\vec{s})$, what we want to obtain is the crystal structure, hence its electron density $\rho(\vec{r})$. We know from Section 2.4.2, that we just have to apply the inverse transformation that leads to

Figure 129: A negative object (a hole in a plane or a cavity in a crystal structure) causes the same diffraction pattern as a positive object (a disc or an atom in a crystal structure) if both have the same shape. They can only be distinguished by their phase.

$$\boxed{\rho(\vec{r}) = e^{i\,\omega t} \sum_h \sum_k \sum_l F(\vec{s})\, e^{-2\pi\, i\, \vec{s}\vec{r}}}\,. \tag{9.31}$$

The above equation is just what the crystallography software does. Since $\vec{s}$ contains the orientation of the plane, we can derive the crystal structure from the diffraction pattern on the screen.

However, there is a small, but delicate problem that is known as *phase problem*. Unfortunately, what we measure is not the amplitude itself, hence not $F(\vec{s})$ directly - what we consider *intensity*, i. e. the pattern on the screen, is **the absolute value of $F(\vec{s})$, squared**. The factor $e^{i\,\omega t}$, that corresponds to a phase $\Delta\phi$ in $e^{i\,\Delta\phi}$, vanishes when the absolute value squared is taken (c.f. Section 2.3.1). Thus, an intensity I can be caused by an infinite set of $F(\vec{s})$, since

$$I \sim |F(\vec{s})\, e^{i\,\Delta\phi}|^2 = |F(\vec{s})|^2\, |e^{i\,\Delta\phi}|^2 = |F(\vec{s})|^2 \tag{9.32}$$

and $\Delta\phi$ can have any real value. This means that different features in the unit cell might cause the same pattern on the screen (see Figure 129). Mathematically, the phase information is lost and there is no way to recover it. Fortunately, there is a trick called *staining* that enables us to "anchor" the phase. When including a grid of heavy metals like tungsten, gold or even uranium with a huge electron cloud, they will cause very prominent diffraction pattern on the screen. Then we know that this bright spot is caused by an atom and not e.g. by a cavity in the crystal structure. Based on this anchoring, the software is able to reconstruct the phase at least approximately.

9.3 Cryo-Electron Microscopy

In this section I will discuss cryo-electron microscopy (Cryo-EM), with an emphasis on the steps of image processing and analysis, 3D reconstruction and determining the spatial resolution as well as exploring the math and physics behind Cryo-EM that is more critical to properly interpreting the data. I therefore focus less on sample preparation and virtrification. For further reading, I recommend the review paper [1] and [2] and the references therein.

9.3.1 Why Electrons?

We know that the spatial resolution of light microscopes is limited by diffraction, and that it therefore depends on the wavelength of the light used. It then seems logical that we could improve the resolution by simply increasing the energy of the light. However, the reality is not so simple. As you may recall from Section 9.1.3, the refractive index of a material is also wavelength dependent. In most materials, the refractive indices for wavelengths in the X-ray and γ-ray ranges are too close to one to build proper lenses. It might seem that one could use materials with a higher refractive index, such as aluminum or boron. But absorption in these materials is too high, resulting in prominent optical aberrations. Consequently, we are unable to build useful X-ray or γ-ray microscopes.
Super resolution techniques (Section 9.1) allow us to reach down to the nm scale, thus beating the diffraction limit by up to two orders of magnitude. But this is not always enough, because in order to resolve atoms we need a way to reach the Ångström scale.

In the 1920's, de Broglie[87] demonstrated in his PhD thesis that massive particles, especially electrons, exhibit wave-particle-duality as well. The *de Broglie wavelength*, as the wavelength of a particle is now known, of an electron λ_e is connected to its momentum p_e by the equation

$$\lambda_e = \frac{h}{p_e}\,. \tag{9.33}$$

An important consequence of this equation is that by increasing the momentum, the wavelength of a particle decreases and is therefore adjustable. An electron is a charged particle that can be accelerated in an electric field of voltage U. The potential energy $E_{pot} = e\,U$ in the field is converted into kinetic energy $E_{kin} = \frac{1}{2} m_e\, v^2$ (c. f. Section 2.1.8) leading to the relation

$$v = \sqrt{\frac{2\,eU}{m_e}}\,, \tag{9.34}$$

where m_e is the rest mass of the electron and e its charge ($1.602176565 \times 10^{-19}\ C$).
Furthermore, since electrons are charged particles, they can be focused by magnetic fields. With modern equipment, attaining voltages on the order of $U = 10^5\ V$ is quite straightforward. At such voltages, electrons reach relativistic speeds (c. f. Equation 9.34) and we must alter our earlier equations in order to include the relativistic gain of mass. If you are unfamiliar with general relativity, that is not important, just know that at speeds close to the speed of light (in vacuum) some effects become important and we must account for them. After doing so and combining Equation 9.33 with Equation 9.34, we write the wavelength of a relativistic electron as

$$\boxed{\lambda_e = \frac{h}{\sqrt{2\,m_e\,eU\left(1 + \frac{eU}{2\,m_e\,c^2}\right)}}}\,. \tag{9.35}$$

[87] Louis-Victor-Pierre-Raymond, 7th duc de Broglie, 1892-1987

Hence, we see that the major advantage of using electrons instead of light is that **we can change the wavelength of the electron, and therefore the resolution, by changing the voltage**. With this knowledge, we are able to achieve very high resolutions without much difficulty. For example, at only $U = 10^5\ V$ the electron reaches a speed of $1.64 \times 10^8\ m/s$ and a wavelength of $\lambda_e = 3.7 \times 10^{-12}\ m$. At $U = 10^6\ V$ the electron has a wavelength of only $\lambda_e = 0.87 \times 10^{-12}\ m$, a resolution far beyond the size of even a hydrogen atom. It was immediately clear to much of the scientific community that this insight was invaluable, and by 1931 the first electron microscope had already been created by Knoll[88] and his PhD student Ruska[89]. The diffraction limit of light was then surpassed already in December 1933.
Cryo-EM is getting the method of choice in structural biology and is going to out-compete other methods like X-Ray crystallography (Section 9.2) and therefore nicely demonstrates the fruitful interaction between biology and theoretical physics (here: quantum mechanics and relativity).
Electrons have a mean free path, i. e., an average distance traveled between collisions, of a few cm in air and only a few μm in solids. Therefore, the electron beam has to be situated in a vacuum tube in order to be effective. The situation is complicated by the unfortunate denaturing of organic specimens in a vacuum. We may remedy this by freezing the specimen. In doing so, we must be careful to freeze it fast enough and to a low enough temperature ($-135°C$) that the natural features are as conserved as possible and no crystals can be formed that could destroy the inner structure of the specimen. The trade-off of achieving such low resolutions is then that it is impossible to investigate our specimens in vivo.
An EM image is strongly limited by the number of electrons, with counts usually less than 20 electrons per square Ångström, in practice less than 10 electrons per second per pixel ([1]). Consequently, EM images will be very noisy and image processing will be a crucial part in the data analysis. Here we see another advantage of cryogenic temperatures, as the ice partially protects the specimen against radiation damage (recall that β-rays emitted during radioactive decay are nothing but fast electrons).
Since the electrons in the beam are interacting with the electrons in the specimen, an EM image equals an electron density map. However, in particular biological samples mainly contain light atoms like C, N, O and H with a less prominent electron shell then e. g. gold or tungsten and therefore EM images suffer from low contrast. I will address all these issues in the next sections.

9.3.2 Image Formation

In transmission EM, the method on which we will be focusing, the specimen we are investigating is located between the source of the electron beam and the detector. As we will be studying the electrons that pass through the specimen and reach the detector, we must consider all of the possible effects that could influence our data before discussing actual image formation.

The most straightforward effect is the elastic or inelastic scattering of beam electrons as they are repelled by the electron clouds of the specimen. We also observe that the beam electrons decelerate as they move through the specimen and may then release this energy loss in the form of X-ray photons (called *bremsstrahlung*[90]). Additionally, electrons may

[88] Max Knoll, 1897 - 1969
[89] Ernst August Friedrich Ruska, 1906 - 1988
[90] Taken from German and means "braking radiation".

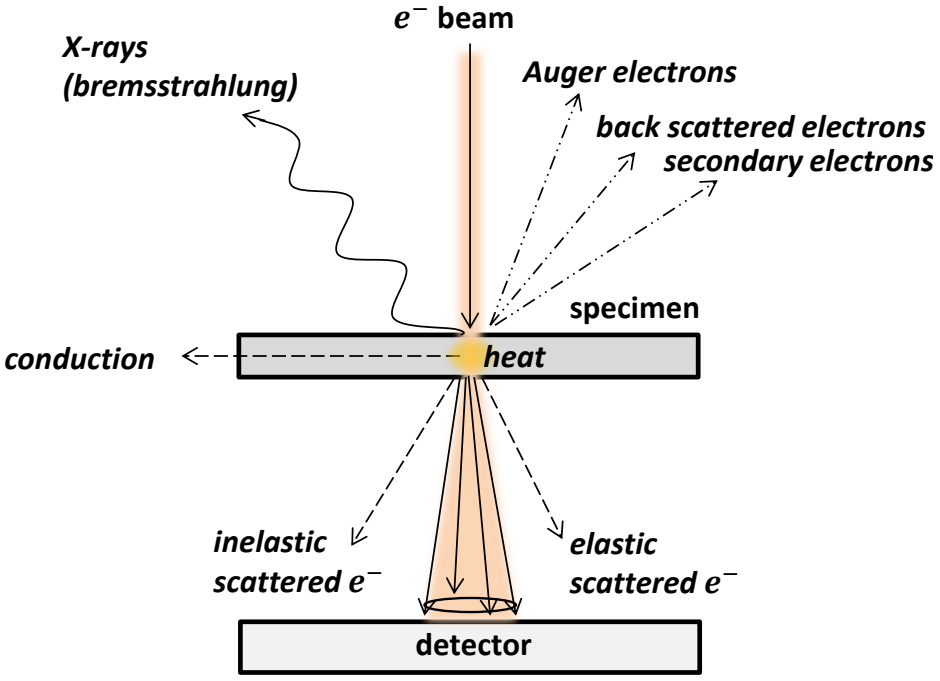

Figure 130: A sketch of possible effects when an electron beam hits a specimen.

collide with and replace electrons in the specimen's atoms, releasing them as secondary electrons. The electrons originally belonging to the beam may occupy an excited state, and if they relax to a lower energy state they will release a photon. Such a process is referred to as *cathodal luminescence*. The emitted photon may itself then collide with and release another electron. These are known as *Auger*[91] *electrons*.
Apart from all of the possible collisions, it is important to remember that electrons, unlike light, are charged particles. As such, they will charge the specimen, leading to structural damage, atomic displacement, deposition or even mass loss.
It is clear that radiation damage is the limiting factor for electron microscopy, and due to all of the effects (see Figure 130) only a small portion of the beam ends up passing through the specimen and reaching the detector. We therefore obtain raw images that are very noisy and some spatial information is lost due to scattering (Figure 130).
Furthermore, the specimens that we investigate often have a very complicated three-dimensional shape. However, each particle is randomly oriented and our raw image is made up of two-dimensional projections on the detector. The lost 3D information must be reconstructed.

9.3.3 Detectors and Sampling

When a transmitted electron impacts the charge coupled device (CCD) detector, it causes a charge cloud. This charge cloud leads to a decrease in voltage via the resistance coupled to each pixel. The measured drop of voltage, proportional only to the number of electrons,

[91]Pierre Victor Auger (1899 - 1993). The effect was actually discovered a year earlier by Lise Meitner (1878 - 1968), however her paper was rejected because of her sex. Her role as co-discoverer of nuclear fission was also ignored when the research was awarded the Nobel Prize. The scientific community finally honored her in some way nearly 30 years after her death, when element 109 officially became known as *Meitnerium* in 1997.

is then used to generate each pixel value. We are therefore left with a grey scale image. The grey scale value only varies linearly with the voltage change in a certain range but at the detection limit and for larger values (close to the saturation point), the charge cloud may not be adequately reflected in the grey scale for large values.
The continuous spatial distribution of the specimen's electron density is transformed into a discrete and digitalised CCD image, thus requiring suitable sampling. If a signal is sampled at a rate of at least double its own (here spatial) frequency f, and this sampling is done in phase with the signal, then no information is lost. Hence, in turn, a feature is detected, if its extension is at least two pixel or larger. This statement comprises the *Nyquist*[92] *theorem*, and signals that are accurately $2f$ sampled are said to be Nyquist sampled (ses Figure 131). When a signal is undersampled it cannot be properly detected, leading to artificial features (called aliases, see Figure 131, very right) in the Fourier space that will also influence the reconstructed 3D image later on. Note that the moiré pattern back in Section 9.1.4 appeared as a consequence of the undersampling of the object due to the superimposed pattern.

Modern detectors use a direct single pixel readout, in contrast to older CCD detectors.

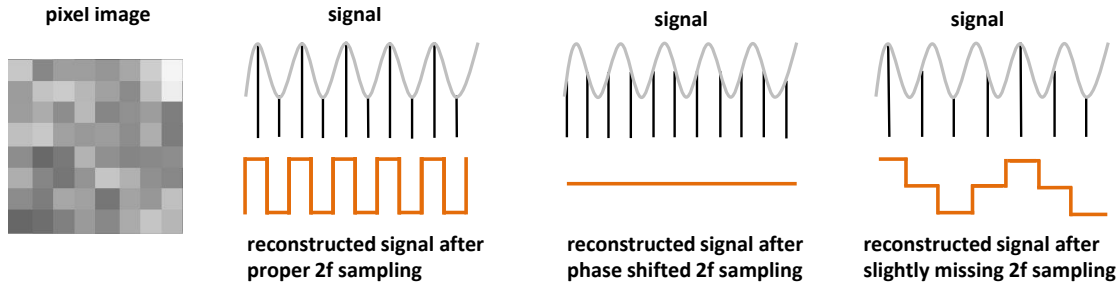

Figure 131: The CCD detector turns the analogue signal into a discrete pixel image (very left). However, the spatial frequency f of the pixels has to be at least twice the spatial frequency of the signal. A signal cannot be reconstructed if the sampling rate is equal to $2f$ but phase shifted, or if the sampling rate is less than $2f$ (very right).

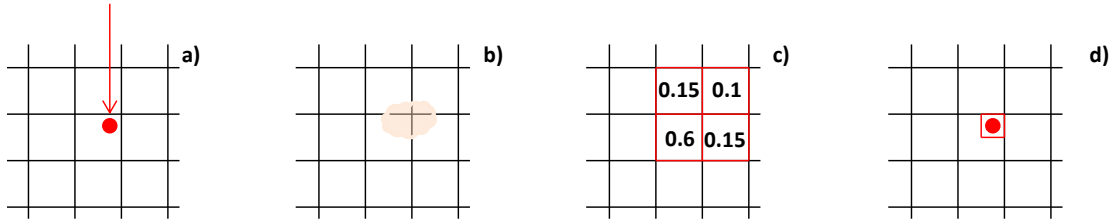

Figure 132: Super resolution for CCD detectors: An electron impacting the detector (a) causes a charge cloud (b) that affects the neighbouring pixels. The pixel values (c) can now be used to reconstruct the location of the impact more accurately (d).

However, inevitable interpolation errors caused by the image processing steps limit the attainable resolution to three pixel in practice. If the charge cloud caused by the impacting electrons also affects neighbouring pixels, we see a similarity with super resolution techniques (Section 9.1). This is because even on the sub-pixel scale we are able to reconstruct the electrons position by weighting the contributions of each pixel, see Figure 132 for a schematic view.

[92] Harry Nyquist, 1889 - 1976

9.3.4 Noise Filtering, Particle Picking and CTF Correction

Before we can process the image, we must first have it pre-processed, by which I mean we must filter out as much of the noise as possible. Remember that because of the risk of radiation damage, the raw images may be extremely noisy. In fact, signal to noise ratio can be in the order of one or two only. An example of such an image for some ribosomes is given in Figure 133. In some raw images the particles are barely visible. To remedy this, the first step that has do be done is to apply some noise cleaning techniques. These are usually median filter, Gaussian convolution, entropy filter or Fourier filter, as also discussed in detail in Section 9.1.8. These noise filters are standard equipment of the image processing software and do not need to be implemented manually. However, understanding their functionality and their drawbacks helps to interpret the data and possible artifacts.

A zoomed region of the image in Figure 133 is shown in Figure 134 after removing the

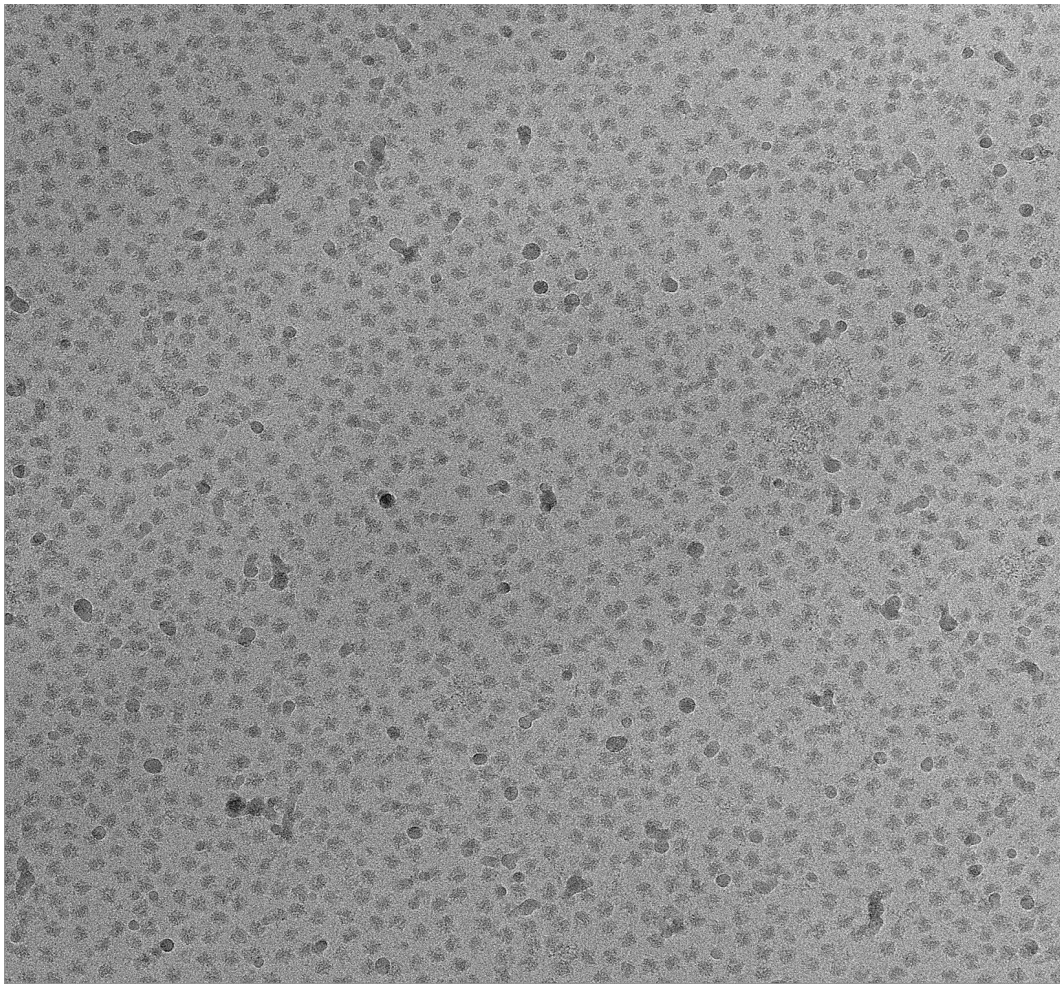

Figure 133: The 70S Ribosome of E. coli. Image courtesy: Hanna Kratzat

background noise by gradient filtering using the *Matlab* function *graydiffweight* and applying the subsequent entropy filter *entropyfilt* for edge recognition (see Section 9.1.8). The edges are outlined using amongst others the structure element function *strel* (Section 9.1.8). Averaging the noise out by adding together several images (frames) as demonstrated in Section 9.1.8 is usually not an option since exposing the same region of the specimen many times increases the probability of radiation damage.

After noise cleaning, the software proceeds with detecting the particles. Often, particles

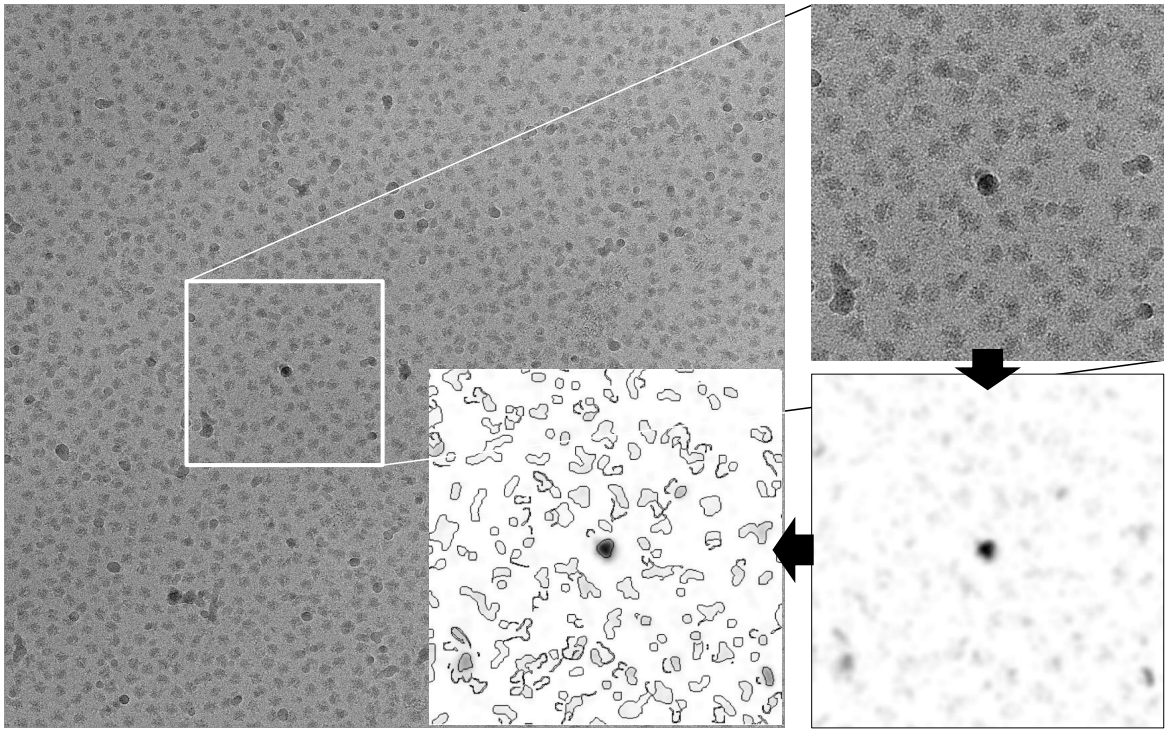

Figure 134: EM image in Figure 133 and a zoomed region of the raw image (upper right). The two further sub figures show this region after removing noise by gradient filtering with the *Matlab* function *graydiffweight* and applying the entropy filter *entropyfilt* for edge enhancing. Arrows indicate the order. Image courtesy: Hanna Kratzat

are selected via *micrographs*, small boxes/windows containing a magnified image of an item. This is a crucial step since the quality of selected particles has a major influence on the subsequent analysis like particle classification and 3D reconstruction. Including poorly detected particles precludes a successful 3D reconstruction and structure determination. Therefore, these steps are still partly done manually since fully automated detection is still error prone. Even particle selection by an experienced scientist might lead to errors since humans tend to focus on features that appear to be more familiar to them and therefore less frequent features are omitted. In semi-automated approaches, the software detects putative features in the micrograph and the user rejects poorly identified particles after visual inspection – that of course also might be biased. Often, templates are used as an initial step, but also this method introduces a bias. Indeed, this might have played a role in some structure determinations (see [1] and references therein, in particular page 443). The problem is well illustrated by the image containing the outlined edges in Figure 134: some features are clearly particles, but some features detected by the software are clearly artifacts and in some cases, the kind of object is not clear. Unfortunately, this is a very typical problem.

We are used to defining the contrast of a macroscopic object by the amplitude of the scattered light we receive. In other words, we see contrast as the relative brightness of an object to its surroundings. Since the interaction of the electrons with the light atoms in biological samples is weak, the wave front (recall that electrons have wavelike properties, Section 9.3.1) is deformed and not absorbed like light in the "macroscopic world". The objects in the Cryo-EM world are *phase objects*, i. e. the contrast in an image arises from the phase shift caused by the interaction of the electron beam with the specimen. One can

imagine this difference by picturing two transparent foils lying upon each other. If one foil were rippled, we would recognize this by the deformation of the wave front of the reflected light (hence, interference pattern) and not by absorption since the foils are transparent.
This has some consequences in the "EM world": We are not used to a contrast depending on the size of the object. However, in EM the contrast of an object depends on the spatial frequency. To quantify the dependence of the contrast on spatial frequency, we use what is called the *contrast transfer function (CTF)*. The CTF is a quasi-periodic function in reciprocal space (c. f. Section 2.4) which depends among other parameters on the optical settings of the microscope like the coherence of the electron beam and the defocus settings. The CTF may be very complex and is not restricted to positive values. When the CTF returns a negative value, it simply means that the contrast has been flipped - objects that appeared black would then appear white, and vice versa. At a zero of the CTF, there is no contrast at the corresponding spatial frequency. Here, information is lost. Low amplitudes of the CTF correspond to an attenuation of the image at the corresponding spatial frequencies. Ideally, the CTF would be constantly equal $+1$ for all spatial frequencies. An example of a typical CTF is shown in Figure 135 and some nice examples of different CTFs are shown in [2], Figure 2 therein.
The effects of the CTF can be corrected by mainly two different ways. The CTF changes

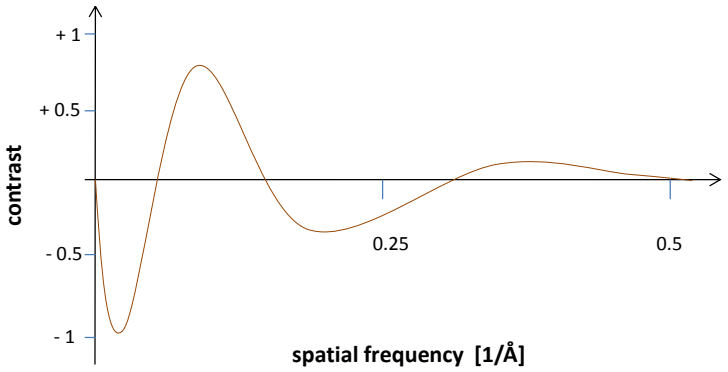

Figure 135: Example of a contrast transfer function (CTF).

for different defoci and in particular its zeros are changing and information that has not been accessible in a particular defocus (in focus, contrast is close to zero anyway) setting becomes now available (non-zero contrast). Usually, three different exposures, hence three different defoci settings are made. There is always the trade-off between higher contrast and less resolution in defocus versus less contrast and higher resolution in focus. However, taking three exposures increases the radiation load for the specimen and therefore an alternative is to use a *phase plate* for CTF correction. The CTF itself can be calculated from the (Fourier) power spectrum of each image. Finally, the images (in Fourier space) are normalized by the calculated CTF.
The contrast can be enhanced by staining the specimen with salt containing heavy metals. The electron shells of heavy metal atoms like gold, tungsten or uranium are a order of magnitude larger than those of biologically relevant atoms (C, N, O, H).
Fortunately, there are numerous software packages available that perform all the image processing steps discussed here and in the subsequent sections. Most common are packages like *SPIDER* (one of the fist that became available), *EMAN2* (most popular, many options), *FREALIGN*, *RELION* (most suitable for beginners) and *SPARX*. For details, see [1] and [2] and the references therein.

9.3.5 Classification, Averaging and 3D Reconstruction

The images of the ribosomes in Figure 133 are 2D projections of their 3D shape. Each ribosome is oriented randomly and casts its own projection (Figure 136, upper left). If we were to have a mixture of different molecules, the situation would be even more difficult. In order to reconstruct the true 3D shape, additional mathematical preparation is required. First, the objects have to properly classified according to their orientation and kind. Most of the algorithms are more or less differently modified versions of the *K-means* clustering method or employ a *principal component analysis (PCA)* (see the bioinformatics script/lectures) – or both. Due to the low contrast and the unfavorable signal to noise ratio in biological samples, it is necessary to average thousands of projections of the single particles. Noise is random and should average out, whereas the signal is not random and will instead add up (c. f. the discussion in Section 9.1.8).
The next step is aligning the different projections of the object. Suppose the ribosome is rotated by an angle ϕ in the $x-y$ plane, i. e. about the z axis. We can introduce new axes x' and y' that are turned with the same angle so that they describe the position of a point $P(x,y)$ in terms of this rotation. How can we express the new coordinate system in terms of the old coordinate system? The location of a point $P(x,y)$ was determined by the vector $\vec{x}+\vec{y}$ in the old coordinate system. Due to the rotation, x becomes x' via the equation

$$\vec{x}' = \vec{x}\cos\phi + \vec{y}\sin\phi\,, \tag{9.36}$$

where ϕ is positive for angles measured counter-clockwise (see also Section 2.5). In the same way, we find

$$\vec{y}' = -\vec{x}\sin\phi + \vec{y}\cos\phi\,. \tag{9.37}$$

The z axis is kept fixed so that $z=z'$. Thus, we can express a rotation about the z axis with a matrix R_z:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{R_z} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{9.38}$$

See the upper right diagram in Figure 136 for a visualization of this transformation.

For any random projection, the ribosome may also be turned about the y axis by the angle χ (lower left diagram in Figure 136). We obtain the rotation matrix R_y in the same manner as we derived the matrix R_z. We can then express our new coordinates in terms of a rotation in the $x-z$ plane as:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\chi & 0 & -\sin\chi \\ 0 & 1 & 0 \\ \sin\chi & 0 & \cos\chi \end{pmatrix}}_{R_y} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{9.39}$$

Finally, we derive an analogous matrix for a rotation about the x axis by an angle ψ (lower right diagram in Figure 136) and obtain:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix}}_{R_x} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{9.40}$$

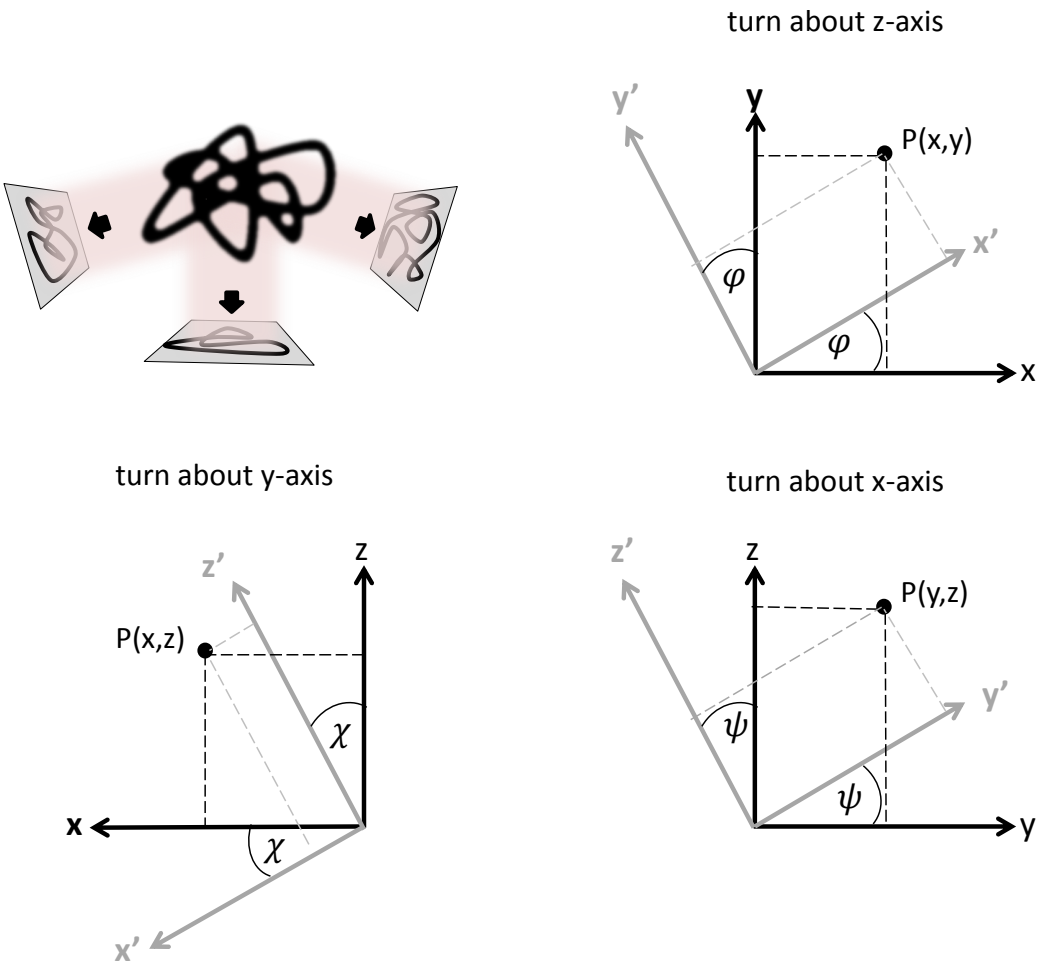

Figure 136: A 3D object produces a random 2D projection in the EM image (upper left). The 3D geometry can be inferred by fitting the 2D projections with the three rotation angles ϕ, χ and ψ (upper right to lower right).

These three angles are called *Euler angles*.
One must execute all three rotation matrices (the order doesn't matter) in order to describe *any* possible rotation. The general rotation matrix is then:

$$\boxed{\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{R_x R_y R_z}_{R_{xyz}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}} \tag{9.41}$$

The image can also be shifted in a particular direction by a vector $\vec{v}$ by a constant amount m. For example, images may be shifted with respect to the background image. The new coordinates are then given by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \vec{t} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + m\vec{v} \tag{9.42}$$

Combining Equation 9.41 and Equation 9.42, we come to the full transformation of a pixel coordinate $\vec{r}_j$:

$$\boxed{\vec{r'}_j = \begin{pmatrix} x'_j \\ y'_j \\ z'_j \end{pmatrix} = R_{xyz}\vec{r}_j + \vec{t}_j} \tag{9.43}$$

But how can we judge whether the alignments and projections are good fits?

Suppose we have two images showing the same picture, which we denote as f_1 and f_2. Each pixel j has a value at the coordinates $\vec{r}_j$ (or at the point $P(x_j, y_j)$). For simplification, we assume for a moment that the pixel values are either 1 (black) or 0 (white, i. e., no signal). If we superimpose the images in a random orientation with respect to each other, we will obtain many overlapping black and white pixels. Therefore, if we multiply the values of the superimposed pixel pairs $(P_1(x, y)\, P_2(x, y))$, we will obtain many zeroes (no matches: 0×1, 1×0, as well as some random matches: 0×0). Some black pixels will also overlap, yielding values of 1 (see Figure 137).

When the images are properly aligned, we will only have 0×0 pairs and 1×1 pairs.

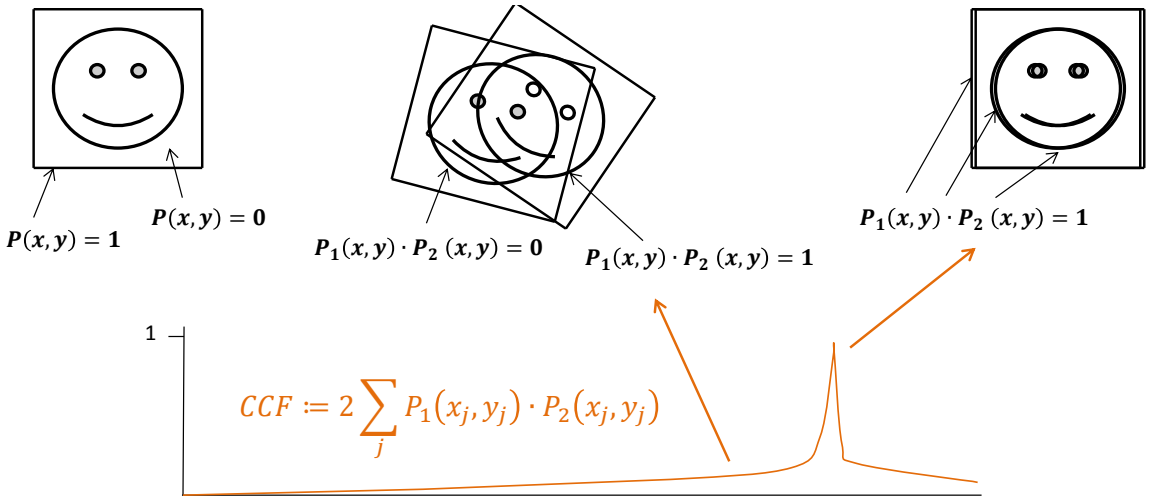

Figure 137: The principle of cross correlation.

Summing up these values would thus yield a larger value than in the previous case.
The sum of the products $P_1(x_j, y_j)\, P_2(x_j, y_j)$ for each pixel pair j is therefore a measure of the matching quality. A high value (usually normalized to one) indicates a good match, while a low value indicates no match (Figure 137). This is the principle of the *cross correlation* method, and the function $\sum_j P_1(x_j, y_j)\, P_2(x_j, y_j)$ is called the *cross correlation function (CCF)*.
Relating this concept back to our EM image with objects of arbitrary rotation and shift (Equation 9.43), we define:

$$E^2_{1;2}(R_{xyz}, \vec{t}) = \sum_j \left[f_1\left(\vec{r}_j\right) - f_2\left(R_{xyz}\vec{r}_j + \vec{t}\right)\right]^2$$

$$= \sum_j \left[f_1\left(\vec{r}_j\right)\right]^2 + \sum_j \left[f_2\left(R_{xyz}\vec{r}_j + \vec{t}\right)\right]^2 - \underbrace{2\sum_j f_1\left(\vec{r}_j\right) f_2\left(R_{xyz}\vec{r}_j + \vec{t}\right)}_{CCF}. \qquad (9.44)$$

Note the appearance of the CCF as the last addend in Equation 9.44.
Usually, the best match is considered to be about the peak of the CCF. However, there may be more than one peak, such as when the object exhibits a certain symmetry. For example, the protein complex in Figure 138 exhibits three peaks due to its rotational symmetry.

When correlating an object with images of itself at different exposures, as we do when

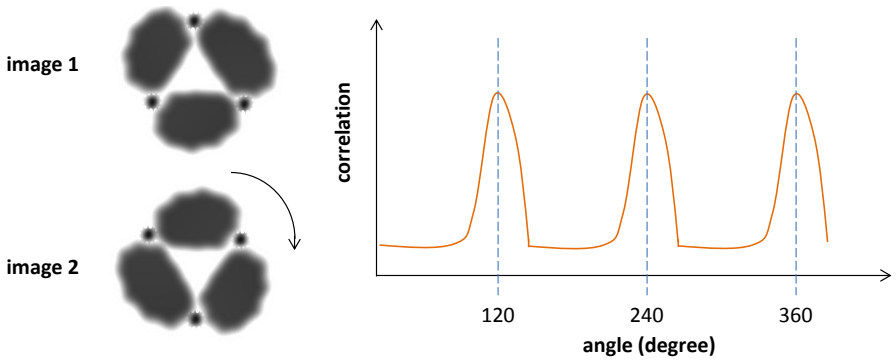

Figure 138: An object with rotational symmetry causes more than one peak in the CCF.

trying to reduce noise, a related quantity is the *correlation coefficient*:

$$\rho_{1;2} = \frac{cov(f_1; f_2)}{\sqrt{var(f_1)var(f_2)}} = \frac{\sum_j \left[f_1\left(\vec{r}_j\right) - \langle f_1\rangle\right]\left[f_2\left(\vec{r}_j\right) - \langle f_2\rangle\right]}{\sqrt{\sum_j \left[f_1\left(\vec{r}_j\right) - \langle f_1\rangle\right]^2 \sum_j \left[f_2\left(\vec{r}_j\right) - \langle f_2\rangle\right]^2}} \,. \tag{9.45}$$

where we have defined the mean value for all J pixel pairs as:

$$\langle f_i\rangle = \frac{1}{J}\sum_j^J f_i\left(\vec{r}_j\right) \,. \tag{9.46}$$

The correlation coefficient is, by its definition, normalized such that the values are between -1 (maximal linear negative/anti correlation) and $+1$ (maximal linear positive correlation), where a value of zero indicates no correlation.
The Euler angles and shift parameter are then varied in order to maximize the CCF and/or the correlation coefficient. The actual 3D reconstruction can start from an initial guess if the rough geometry of the object and/or preferred orientation is known. The Euler angels are calculated in an iterative refinement process by comparing calculated projections from the current model to the map of all recorded projections. An *ab initio* approach uses the *central section theorem* which states that the 2D Fourier transformed images of all 2D projections are central sections of the 3D Fourier transformed image of the object. Each 2D projection has three common lines with their 3D origin ([4]).

9.3.6 Determining the Resolution: The Fourier Shell Correlation

Theoretically, resolutions down to sub atomic level should be easily possible (Section 9.3.1). However, as we have seen, EM images are noisy, the resolution is limited by the pixel size of the detector (Nyquist theorem, Section 9.3.3), due to scattering, spatial information is lost and the 3D reconstruction is erroneous. These are only *some* effects that lower the actual resolution. But what is regarded resolved? It turns out that determining the resolution in Cryo-EM is not a trivial task.
Usually, the data is split into two random sets. One then performs the previously described steps in image processing, particle finding and 3D reconstructions on each set individually, but exactly in the same manner with the identical settings. The two projections of the same object are then transformed into the Fourier space and their correlation is calculated

(similar to Equation 9.44, but in the Fourier space). The resulting curve is called *Fourier shell correlation (FSC).*
The FSC is a function of the spatial frequency, due to it being done in the Fourier space. If the FSC is close to unity, then the two different 3D reconstructions lead to the same result. This should be the case for smaller spatial frequencies, because it corresponds to the correlation of larger features. Hence, the object is said to be resolved at the corresponding spatial frequency. For larger spatial frequencies, the FSC should drop and ideally reaches zero (no correlation, not resolved) at some point. But the drop is not necessarily sharp and usually occurs gradually (Figure 139). Furthermore, at higher frequencies, noise becomes more dominant and one finally only correlates noise (recall that noise features have typically pixel size), that just by chance yields values above zero.
What is now an acceptable threshold? In the earlier days, scientists argued, that it should

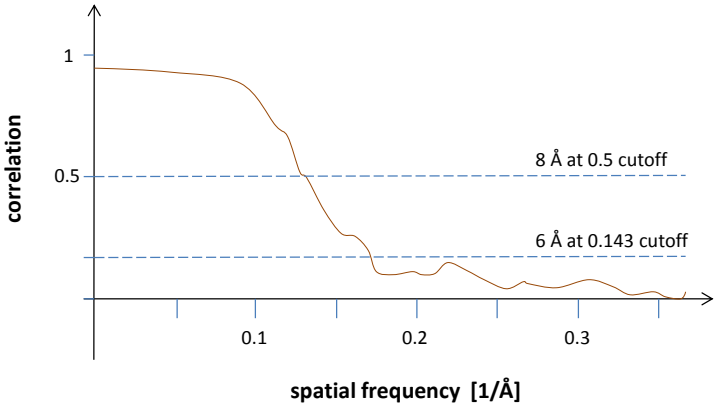

Figure 139: The Fourier shell correlation (FSC). Usually, everything up to $FSC = 0.5$ is regarded as resolved, yielding the threshold between resolved spatial frequencies and unresolved spatial frequencies. This cutoff is often regarded too conservative and one can argue that a cutoff at 0.143 (called "golden standard") is more reliable (see discussion in [3]).

be common convention to set the threshold at the value $FSC = 0.5$, although this is not a unanimous choice and no mathematical justification was given. Applying this threshold to the FSC shown in Figure 139, one would get a resolution of 8 Ångström. The $FSC = 0.5$ cutoff is a very conservative threshold. However, one can show that a fixed threshold does not take the individual symmetry of the object and the number of voxels in the Fourier shell into account. Thus, resolutions obtained by the same threshold, but for different experiments are not necessarily comparable (see also the discussion in the last sections of [3]).
Another criterion is that details are not resolved if only noise is randomly correlated. Therefore, if the noise spectrum is well known, on can set a 3σ or often 5σ threshold for the FSC being above the noise tail to avoid random correlation. A different approach is to compare well known structures obtained by other methods like X-ray crystallography to those calculated from EM images. As a kind of figure of merit it turned out that an appropriate threshold is about $FSC = 0.143$, the so-called *golden standard*, that is nowadays frequently used and which leads to better resolutions (Figure 139).
All these thresholds have to be taken with caution since they can formally reach spatial resolutions beyond the Nyquist limit. As a rule of thumb, resolutions should in any case not be below two third of the Nyquist limit ([3] and references therein). Therefore, while values below it are indeed published, any claims that it has been surpassed are dubious.

Compared to the theoretical resolution (Equation 9.35), we lost two orders of magnitude by the effects discussed above. However, modern EM reach down to some Ångström resolution, thanks to better image processing algorithms and improved detector techniques and thus Cryo-EM will probably out compete X-ray crystallography in the next years.

References

[1] Yifan Cheng, Nikolaus Grigorieff, Pawel A. Penczek and Thomas Walz, *"A Primer to Single-Particle Cryo-Electron Microscopy"*, Cell. 2015 Apr 23; 161(3): 438 – 449.

[2] Mariusz Czarnocki-Cieciura, Marcin Nowotny, *"Introduction to high-resolution cryo-electron microscopy"*
"Wprowadzenie do wysokorozdzielczej kriogenicznej mikroskopii elektronowej", Postepy Biochemii 62 (3) 2016.

[3] Marin van Heel and Michael Schatz, *"Fourier shell correlation threshold criteria"*, Journal of Structural Biology 151 (2005) 250-262

[4] A. Singer and Y. Shkolnisky, *"Three-Dimensional Structure Determination from Common Lines in Cryo-EM by Eigenvectors and Semidefinite Programming"*, SIAM J. Imaging Sci., 4(2), 543-572.

9.4 Atomic Force Microscopy (AFM)

9.4.1 A Brief Introduction of Atomic Forces

Atomic forces are the forces between, not within, atoms, as well as between different molecules (molecular forces), hence they are actually forces between the electron shells of the atoms and are therefore electrical forces[93]. Such forces are not necessarily associated with bonds, and may even be macroscopically palpable.
If two different surfaces are close together, then the electron clouds of the atoms of one surface begin to "feel" the electron clouds of the other surface. This gives rise to three possible situations. First, dipoles on both sides of the surfaces (if present), can influence each other and cause the molecule to rotate until the the poles are aligned such that the resulting force is attractive. Second, a dipole on one side could also induce the polarization of a nearby, non-polarized surface. And finally, small quantum fluctuations that lead to tiny inhomogeneities in the electron clouds of the atoms may cause a temporary polarization. If the polarization is strong enough, then it could induce a permanent dipole with the opposing surface. Such dipoles exert small forces individually, but summed over an entire surface may result in a macroscopically relevant force.
As the surfaces are brought closer together, the electron clouds are forced to overlap. The

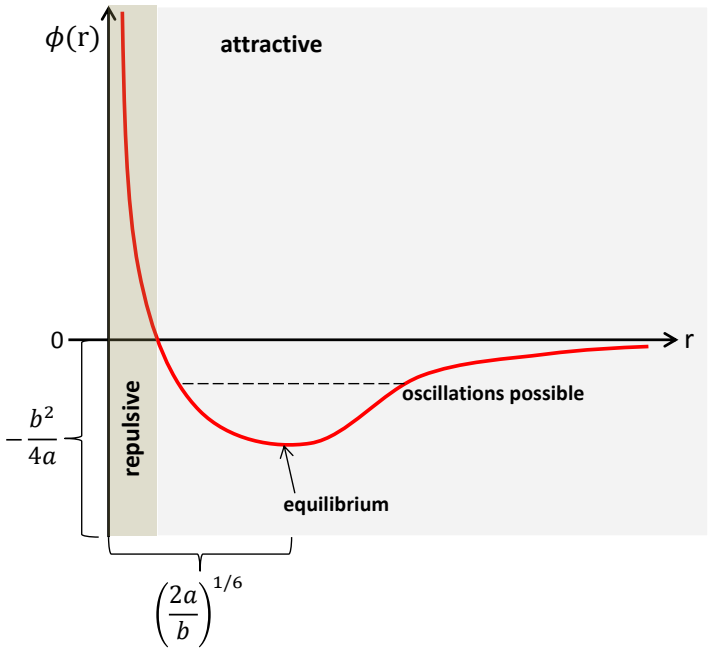

Figure 140: The Lennard-Jones potential $\phi(r)$ contains a repulsive part for small r (Pauli repulsion) and an attractive part for larger r (Van-der-Waals attraction).

high energy cost of this results in a repulsive force, known as *Pauli*[94] *repulsion*. In between the attractive region of the *van-der-Waals*[95] *forces* and the repulsion regime, there exists a stable equilibrium state. Thus, there exists a potential well that governs the distances between atoms and molecules where they are "trapped" together.
The derivation of the potential ϕ (i. e. the integral of the force over the separation

[93] in contrast to nuclear forces - the actual atomic forces
[94] Wolfgang Ernst Pauli, 1900 - 1958
[95] Johannes Diderik van der Waals, 1837 - 1923

distance r) that describes this situation is not trivial and requires some quantum mechanical knowledge. I therefore just provide the result, which is the so-called *Lennard-Jones*[96] *potential*:

$$\boxed{\phi_{LJ}(r) = \underbrace{\frac{a}{r^{12}}}_{\text{repulsive part}} - \overbrace{\frac{b}{r^6}}^{\text{attractive part}}} \,. \tag{9.47}$$

The constants a and b in Equation 9.47 depend on the material, while the attractive and repulsive parts are quantitative descriptions of the Van-der-Waals forces and of Pauli repulsion, respectively. A graphical representation of the Lennard-Jones potential is given in Figure 140.
The potential energy of a charge q in a potential ϕ is just $E_{pot} = q\phi$ so that the minimum of the Lennard-Jones potential equals minimum energy, i. e. an equilibrium state. A particle that is removed from this equilibrium by a small perturbation falls back to lower energies and then performs oscillations around the minimum (like a jumping spring without friction, c. f. Equation 2.123).

The potential energy E_{pot} between two charges q is described by the Coulomb[97] potential in electrostatics:

$$E_{pot} = \frac{1}{4\pi\epsilon_0\epsilon}\int \frac{q}{r}\,dq = \frac{1}{4\pi\epsilon_0\epsilon}\frac{q^2}{2r} \tag{9.48}$$

where r is the distance between the two charges and $\epsilon_0 = 8.8541878176 \times 10^{-12}\ F/m$ is the absolute permittivity of free space (a conversion factor like the gravitational potential factor G, since we measure the quantities in arbitrary units like seconds, meter etc.).
The permittivity, and thus the electric field, are affected by the medium being passed through. This can be caused by small dipoles in the medium that become polarized by the electric field. Due to the opposing positive and negative charges, the field attenuates. Counter-intuitively, this corresponds to an increased permittivity, as permittivity is not a measure of how much is permitted, but rather of how much a material resists an electric field. The relative permittivity ϵ weights ϵ_0 to account for polarization effects. By definition, a vacuum will give zero resistance, so its relative permittivity $\epsilon = 1$. However, in water, where dipole interactions are prominent, $\epsilon \approx 60 \ldots 90$ (depending on temperature).
We have seen in previous sections that thermal fluctuations govern the microscopic motion of molecules. For charged particles, their interactions are complicated by their electrical energy. At what point does thermal noise overcome electrical energy? Setting the right hand side of Equation 9.48 equal to the thermal energy kT (T in Kelvin, see also Section 3), and writing the charge q as an integer number Z times the elementary charge $e = 1.602\,176\,565 \times 10^{-19}\,\mathrm{C}$, we rearrange to solve for distance and come to:

$$r = \lambda_B = \frac{Z^2e^2}{8\pi\epsilon_0\epsilon kT}\,. \tag{9.49}$$

This is a famous result, where the distance λ_B, known as the *Bjerrum*[98] *length*, gives the largest separation at which electrical energy dominates thermal noise. The length λ_B is an order of magnitude estimate and may differ by a factor of two in the literature. Under physiological conditions, λ_B is usually less than a manometer.

[96] Sir John Edward Lennard-Jones KBE, FRS, 1894 - 1954
[97] Charles Augustin de Coulomb, 1736 - 1806
[98] Niels Bjerrum, 1879 - 1958

The situation is further complicated by the inclusion of charge cloud dynamics in our description. As electric charge must be conserved, so must the electrostatic flux around a charge cloud. Therefore, we apply the Laplace operator to the potential ϕ, just as we did in Section 6.1, Equation 6.30 for concentration (mass conservation) in our discussion of the law of diffusion (Fick's second law), and obtain for a cloud of charge density (instead of mass density) ρ:

$$\Delta\phi(r) = -\frac{1}{\epsilon_0\epsilon}\rho(r)\,. \tag{9.50}$$

Note that if $\rho(r)$ describes a point source, Equation 9.48 (for a single molecule) is the solution of Equation 9.50.
A radially symmetric charge cloud greatly simplifies our calculation by applying Equation 2.206 for writing the Laplace operator in spherical coordinates, yielding:

$$\frac{d^2\phi_i(r)}{dr^2} = -\frac{1}{\epsilon_0\epsilon}\rho_i(r)\,. \tag{9.51}$$

Equation 9.51 describes the potential of a radially symmetric charge cloud as "felt" by an ion i. But what is ρ_i?
The total charge Q is the sum over all charges,

$$Q = \sum_j Z_j\, e\, N_j\,, \tag{9.52}$$

where N_j is the number of ions of a particular type. The charge density is given by $\rho = \frac{Q}{V}$, so

$$\rho_i(r) = \sum_j Z_j\, e\, n_j\,, \tag{9.53}$$

where $n_j = \frac{N_j}{V}$ is the *number density* of ion type j.
The number density of the charged particles also depends on their location in the cloud. A charge feels less of the potential if it is located at the edge of the cloud and can therefore be dragged away more easily by thermal noise. Hence, the location (or the probability of a particular location) of a particular ion depends on the ratio of electrical energy to thermal energy. We can write describe this for ρ_i using the Boltzmann distribution (Section 3.2, Equation 3.26):

$$\rho_i(r) = \sum_j Z_j\, e\, n_j = \sum_j Z_j\, e\, n_{0j}\, \underbrace{e^{-\frac{Z_j e\phi_i(r)}{kT}}}_{\text{Equation 3.26}}\,, \tag{9.54}$$

where n_{0j} is the number density in the center of the cloud. Plugging this back into Equation 9.51, we obtain the *Poisson-Boltzmann* equation:

$$\frac{d^2\phi_i(r)}{dr^2} = -\frac{1}{\epsilon_0\epsilon}\sum_j Z_j\, e\, n_{0j}\, e^{-\frac{Z_j e\phi_i(r)}{kT}} \tag{9.55}$$

At this point, it seems like we could just solve for ϕ_i. In doing so, however, we would be violating the principle of superposition, because $\phi(r)$ depends on ρ_i and vice versa. Equation 9.51 would need a further non-linear term in order to account for that (c. f. Section 4.2). Instead, we search for a solution by expanding Equation 9.51 into a Taylor series (Section 2.2.2). Our result truncates terms of higher orders (the *Debye-Hückel*[99]

[99] Peter Joseph William Debye ForMemRS, 1884 - 1966 and Erich Armand Arthur Joseph Hückel ForMemRS, 1896 - 1980

approximation), thus **restricting the validity of the solution only to dilute solutions**. The approximate solution of Equation 9.55 is known as the *Debye-Hückel potential*:

$$\boxed{\phi_i(r) \approx \frac{A}{r} e^{-\kappa r}} \tag{9.56}$$

where A is a constant fixed by the boundary conditions and

$$\kappa = \sqrt{\frac{\sum_j Z_j^2 \, e^2 \, n_{0j}}{\epsilon_0 \epsilon \, kT}} \, . \tag{9.57}$$

The quantity $\frac{1}{\kappa}$ is a length called the *screening length* or the *Debye length*, and characterizes the extent to which the electrostatic effect of this charge cloud persists. It is often written in terms of the *ionic strength*

$$I = \sum_j Z_j^2 \, n_{0j} \, . \tag{9.58}$$

There is a three-dimensional equivalent to the Debye length known as the *Debye volume* $\sim \left(\frac{1}{\kappa}\right)^3$, which characterizes a sphere of electrostatic influence.
Under physiological conditions, the Debye length is $\approx 2 \; nm$.

9.4.2 The Atomic Force Microscope

Light microscopy create images more or less in 2D only. The ***a**tomic **f**orce **m**icroscopy (AFM)* offers the opportunity to get an image of the two directions on the surface of the specimen and, in addition, its profile, the z-direction.
The idea of AFM is rather simple, but it requires sophisticated techniques that work properly at nm scale. These techniques became available in last quarter of the 20^{th} century so that the first AFM was developed by three scientists[100] in 1986.
The key part of an AFM is a cantilever made out of silicon with a very thin tip (Figure 141, left) that is guided close to the surface of the specimen. The atomic forces between the surface of the specimen and the tip lead to motions of the tip in z direction. This motion is measured by a laser, reflecting its light from the tip onto a photo diode. Due to the laser reflection, the motion of the tip is magnified on the diode and an amplifier increases the signal so that it becomes visible on a screen.
The diode is separated into four quadrants to measure vertical and horizontal motions of the tip. The cantilever itself is fixed, but the specimen is moved slowly by piezo crystals[101].
The cantilever is modeled as a spring with the spring constant κ (not to change with Debye length in Equation 9.57). In first approximation, the energy E that is required to deflect the tip by a length z is

$$E = \frac{1}{2} \kappa z^2 \, , \tag{9.59}$$

see Section 2.2.2
Ideally, the tip at rest is only influenced by thermal noise ($E = \frac{1}{2} \, kT$). Thus, the spring constant can be measured by

$$\kappa = \frac{kT}{\langle z^2 \rangle} \, . \tag{9.60}$$

[100] Gerd Binnig (1947 - today), Calvin Quate (1923 - today) and Christoph Gerber (1942 - today)
[101] Piezo crystals deform under voltage. In this way one can perform motions in the nm range depending on the applied voltage.

What is the mean of z^2?
Thermal noise leads to fluctuating deflections of the cantilever. These deflections are measured over time and a subsequent Fourier analysis (Section 2.4) yields the oscillation frequency of the cantilever in the power spectrum. Because the deflection is the amplitude of the oscillation, the integral over the power spectrum equals the mean of the weighted quadratic deflection that is $\langle z^2 \rangle$.
If the spring constant is known, one can account for the bending of the tip under its own mass – the tip correction. The force needed to bend the tip is (c. f. Section 2.1.8)

$$\frac{d}{dz}\left(\frac{1}{2}\kappa z^2\right) = \kappa z\,. \tag{9.61}$$

One can measure the softness of the specimen if the spring constant of the tip is less than the spring constant of some surface features on the specimen. AFM measure forces in the $10^{-12}\ N$ scale. Among other, there are three main modes for measuring the surface profile.

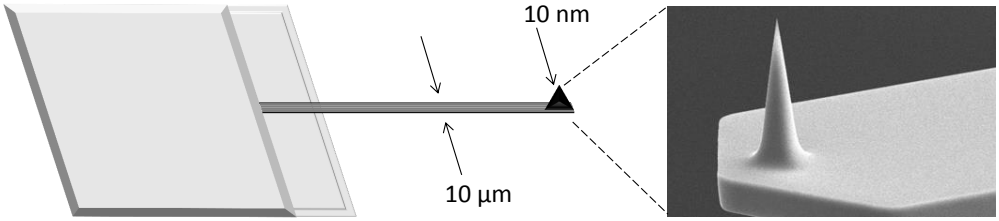

Figure 141: AFM cantilever with tip (left) and a real electron microscopy image of the tip (right; taken from `www.nanoscience.com`.

Contact Mode: If the cantilever is dragged close to the surface of the specimen, i. e. at the repulsive part of the Lennard-Jones-potential (Figure 140), it is deflected by repulsive forces. Either one measures the deflection in order to obtain the profile, or the force is measured that is required to keep the tip at constant z. This mode is called contact mode. Alternatively, at small distances z the tunnel current between tip and surface, that is a function of z, can be measured instead of the tip reflection.

Tapping Mode: The contact mode might damage the surface. Also, if the specimen contains water, the tip can be trapped by capillary forces due to a water meniscus. To avoid these drawbacks, the tip is held in a position apart form the surface but still close enough (a few nm) where it is affected by the attracting van-der-Waals forces (see Figure 140). The tip is excited to perform oscillations by a piezo crystal so that it starts to oscillate in its resonance frequency around the minimum in the Lennard-Jones-potential (Figure 140) and therefore, the tip touches the surface only for a short time. The amplitude and frequency of this oscillation depends on how close the tip is to the equilibrium point (Figure 140). This method is called *tapping mode*. The topographic image of the surface is obtained by the compensation of this changes in order to keep the average z constant all over the scan. The big advantage of the tapping mode is that it is gentle enough to scan lipid layers in water or single molecules without changing the molecular conformation.

Non Contact Mode: At larger separations, the AFM operates in *non contact mode* that is most suited for even softer surfaces or if numerous scans over the same regions of the sample are required. It could be that images of soft or liquid surfaces obtained with non contact mode differ from those obtained in contact mode.

9.4.3 Folding and Binding of Molecules

The tip can be used to investigate the conformation mode of molecules.
Imagine for example something like a cystin molecule bound to a protein and a polymer (Figure 142). The tip can stick at the polymer if in contact mode. Moving the tip upwards in z direction requires some force that can be measured. The force increases further for larger z until it jumps suddenly to much lower values. What happened is that the curled polymer is stretched until a part, that was folded previously, was released and turned upwards ((a) in Figure 142).

The same phenomenon can appear for several times, depending on the conformation of

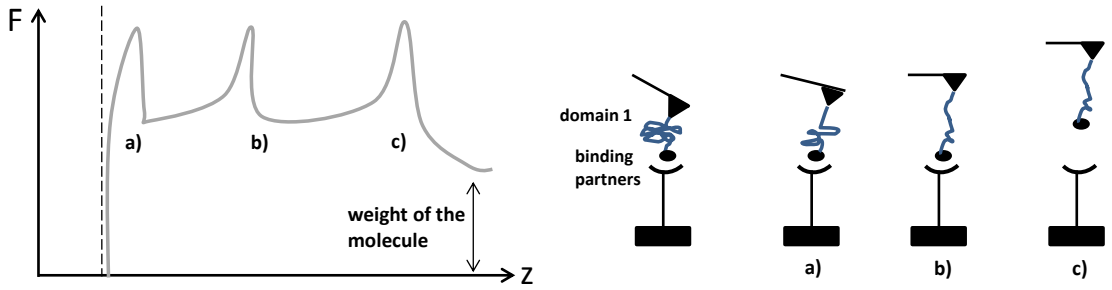

Figure 142: Required force F to move the tip in z direction (left) if the tip sticks on a macromolecule (bottom) with different domains. Stretching the molecule by increasing z first turns the freely joint parts of the polymer upwards (a and b) and may eventually lead to the destruction of the bound (c).

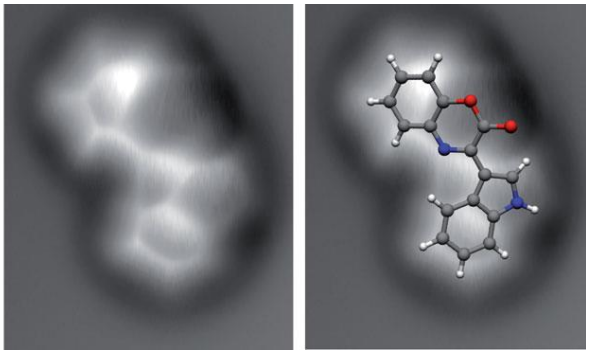

Figure 143: The AFM image of Cephalandole A (left) and the same image overlaid with the corresponding molecular model. Image taken from [1].

the polymer, if z is increased further on ((b) in Figure 142). In this way one can distinguish between a chain like polymer with freely joint parts (exhibiting these jumps in the force curve; Figure 142, left) and measure the length of the unfolded domain; and a worm like polymer leading to more gradual features in the force curve.
If z is increased even further, then the force might drop further ((c) in Figure 142). This occurs if the bound is torn by the tip and the two domains are now separated.

The limiting factor of AFM is actually the size of the tip. AFM easily reaches resolutions in the nm range or even down to some Angströms that is sufficient to resolve even smaller single molecules (Figure 143). It is even possible to reach atomic resolutions under some circumstances.
AFM can be performed under physiological conditions (in liquids; no special coatings and no vacuum is required) that is quasi in vivo and reduces the probability to measure artifacts. However, AFM is very slow and only areas of a few dozens μm radius can be scanned within an acceptable time. AFM can be affected by non linear effects like hysteresis in the tip deflection or in the piezo crystal. Generally, the measurement of the force and the deflections is affected by non-linearity since Equation 9.60, that is used to calculate the spring constant, is derived by a Taylor expansion truncating terms of higher (i. e. non linear) orders.

References

[1] Gross, L. et al., *"The chemical structure of a molecule resolved by atomic force microscopy"*, Science 325, 1110 – 1114 (2009).

9.5 Mass Spectrometry

The principle of mass spectrometry was proposed by J. J. Thomson[102] who also built the first mass spectrometer in 1912 although some attempts were already made by e. g. William Prout[103] in the early 19^{th} century. While the first mass spectrometer reached resolutions of only $\frac{m}{\Delta m} \sim 13$, modern facilities surpass a value of $\frac{m}{\Delta m} \sim 10^7$ with a sensitivity down to 10^{-15} grams. Thus, an application in bio science suggests itself.
As for Cryo−EM and other methods, mass spectrometry is a wide field and a throughout and complete description extends the scope of this script by far. Therefore, I like to focus on the mass spectrometry of proteins and in particular on Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT−ICR MS) only since it illustrates a further application of Fourier transformation which is discussed already in detail in Section 2.4, and since it is fast (μs resolution) and has the highest mass resolution among the large diversity of different methods. For more details and further reading I recommend [1] which gives also very detailed and clear examples that help the reader checking the own understanding.
The principle work flow of a mass spectrometer is to ionize the analyte and to separate the molecules and their fragments according to their mass to charge ratio by using electric and magnetic fields. While it is straight forward to derive the absolute amount of the compounds in the analyte by the mass spectrum, deriving a sequence, e. g. of a protein, is not trivial and sometimes even not unambiguous.

9.5.1 Ionization

There are two main approaches to ionize bio molecules. Electrospray ionization (ESI) is a relatively "soft" method that leaves fragile molecules intact and is best suited for complex molecules like proteins. However, ESI is very sensitive to sample contamination that has to be taken into account while preparing the analyte.
The sample is dispersed in a fine aerosol within a strong electric field of $\sim 10^6\, V/m$ at $\sim 10^3\, V$ that leads to the ionization of the molecules. The analyte contains volatile organic compounds that leads to an effective evaporation into small droplets. During the evaporation, the sample gets ionized and the charge density increases on the droplets. Eventually, the repulsive electrical forces of the ions overcome the surface tension of the droplet (beating the so-called Rayleigh[104] limit) leading to a complete fission of the droplets, leaving only the ions.
The second method is called Matrix Assisted Laser Desorption/Ionization (MALDI) for which the analyte is co-crystallized within a matrix (usually 3,5-dimethoxy-4-hydroxycinnamic acid i. e. sinapinic acid, α-cyano-4-hydroxycinnamic acid, i. e. alpha-cyano or 2,5-dihydroxybenzoic acid, also known as DHB). The ratio of analyte to matrix molecules is in the order of $1 : 10^3$ mol, but could wary by an order of magnitude. The ionization occurs via pulsed laser shots onto the sample (called desorption or ablation) where the laser wavelength corresponds to the absorption energy of the matrix (cf. Equation 9.2). Like ESI, MALDI is a relatively soft ionization method and is most suited to analyze larger bio molecules, such as DNA, sugars (and their polymers) or complex proteins.

[102] Sir Joseph John Thomson OM, 1856 – 1940
[103] William Prout, 1785 – 1850
[104] John William Strutt, 3rd Baron Rayleigh OM PC PRS, 1842 – 1919

9.5.2 The Physics: Separating according to $\frac{m}{z}$

It is impossible to understand mass spectrometry and to interpret the results correctly, if the physical background of this method is not clear. Thus, I like to give a brief overview about the basic physical principles behind mass spectrometry.
Once, the analyte is ionized, it is analyzed in a vacuum cavity that is penetrated by an electric and magnetic field - a so called *Penning trap*[105]. We know that an electrically charged particle of charge z and mass m that moves with a velocity $\vec{v}$ is deflected in a magnetic field of flux density $\vec{B}$ according to the Lorentz[106] force

$$\vec{F}_L = z\left(\vec{v}\times\vec{B}\right). \tag{9.62}$$

Due to the deflection, the particle changes its direction. This change equals a curve in space and one can assign a curve radius r (Figure 144, *a)*) at the turning point and a corresponding angular frequency ω. Due to this change of direction, the particle feels the centrifugal force

$$\vec{F}_C = m\,\vec{\omega}\times(\vec{\omega}\times\vec{r}). \tag{9.63}$$

If velocity and magnetic flux density are perpendicular (i. e. $\vec{v}\perp\vec{B}$), the vector product in Equation 9.62 reads $|\vec{v}\times\vec{B}| = v\,B$ and if the particle passes through an homogeneous magnetic field of sufficiently large spatial extension, its trajectory will eventually follow a circular motion (hence $\vec{\omega}\perp\vec{r}$). Then, the Lorentz force is balanced by the centrifugal force and we can join Equation 5.64 with Equation 9.63, leading to

$$z\,v\,B = m\,\omega^2\,r, \tag{9.64}$$

where r equals the radius of the circle (Figure 144).
The particle will circumnavigate the circumference $2\pi\,r$ within the period $T = \frac{1}{f} = \frac{2\pi}{\omega}$, i. e. with the velocity

$$v = \frac{2\pi\,r}{T} = \frac{2\pi\,r\,\omega}{2\pi} = r\,\omega. \tag{9.65}$$

Substituting the velocity in Equation 9.64 by Equation 9.65 leads to the well known ratio

$$\boxed{\frac{m}{z} = \frac{B}{\omega}} \tag{9.66}$$

where ω is usually denoted as *cyclotron frequency* ω_C. The magnetic flux density B is given by the experiment and ω_C can be measured very precisely and therefore, one can infer the mass to charge ratio $\frac{m}{z}$ of the ion.
Based on Equation 9.66, we can estimate the expected cyclotron frequency in a magnetic field of a flux density of some Tesla (note that Earth's magnetic field has $\approx 10^{-5}\ T$), a typical value in a mass spectrometer. The elementary charge is in the order of $10^{-19}\ C$ and the mass will be in the order of some u up to some hundred u ($u \approx 10^{-27}\ kg$), hence, the expected cyclotron frequency is in the order of $10^{12}\ Hz$ (plus/minus one or two magnitudes). This frequency corresponds to microwaves or far infrared (c. f. Figure 102). Thus, once the ions are trapped in a cavity, it is easy to measure the cyclotron frequency by using an external periodically changing electrical field. The measured voltage will drop when the external field is in resonance with the cyclotron frequency.
As Figure 144, *b)* suggests, the $\vec{B}$ – field forces the ion to move in a circle, but there is

[105] Frans Michel Penning, 1894 – 1953
[106] Hendrik Antoon Lorentz, 1853 – 1928

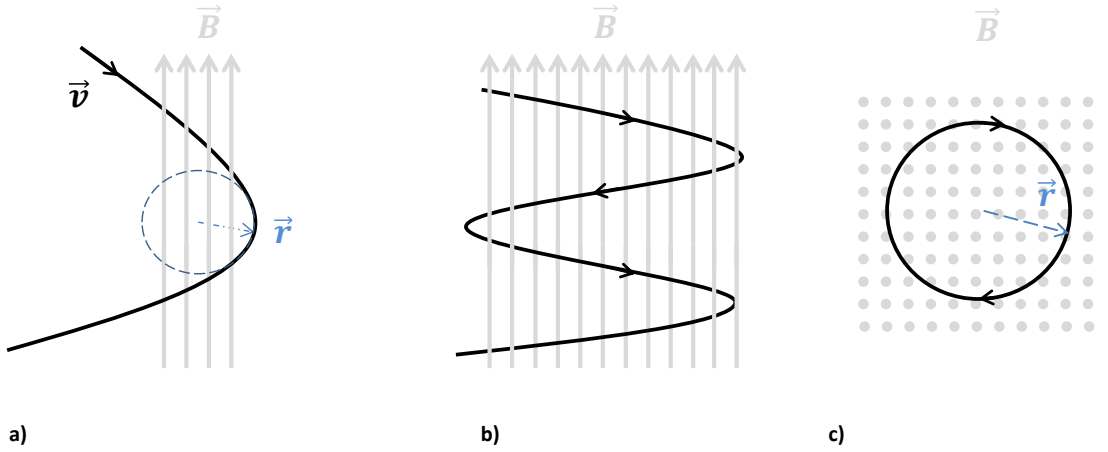

Figure 144: Trajectories (black solid lines) of a positively charged particle in a magnetic field $\vec{B}$. For negatively charged particles, the trajectory points in the opposite direction, but exhibits the same geometry. The charged particle is deflected in a magnetic field $\vec{B}$ by the Lorentz force (*a*), Equation 9.62. If the magnetic field is homogeneous and if the field density is high enough and/or the field is sufficiently spatially extended, the particle is forced to perform a circular (actually spiral) motion (*b*) with curvature radius $\vec{r}$ so that centrifugal and Lorentz force are equal (Equation 9.64). Panel *c)* shows the same as in *b)*, but viewed from above.

still some freedom in axial direction (parallel to the magnetic field lines) such that the ion actually performs a spiral motion. This motion would lead to a drift and the ion would eventually leave the cavity. The idea of a *Penning trap* is that in addition to the magnetic field, an electric field is used to remove the freedom of motion in axial direction. This is achieved by three electrodes: one electrode (negatively charged if the ion is positive and vice verse) is ring shaped and "wrapped" around the magnetic field lines and two other electrodes are two endcaps, having both the same electric potential (both positively charged if the ion is positive and vice verse) and are located along the axial directions (Figure 145). Both, the ring electrode and the endcaps have the shape of hyperboloids of revolution. This configuration creates a saddle point of the electrical potential for the ion in the center of the ring.

If an ion moves into the *Penning trap*, it will move in axial direction towards the saddle point and surpass it, since its momentum is not necessarily exactly zero at this point. Thus, the ion will approach one endcap, which acts repulsively and therefore the ion will be pulled back to the saddle point and again surpass it with non-zero momentum. The ion will therefore approach the other endcap, that also repels the ion and so on. Hence, the ion will oscillate still in axial direction around the saddle point, but does not leave the trap.
We know from Section 2.2.2, that such an oscillation can be modeled with

$$E \approx \frac{1}{2}\, k \Delta x^2 \tag{9.67}$$

for small displacements Δx (c. f. Equation 2.123) in axial direction. Here, the energy E that is required for the displacement equals the potential energy $E = z\Delta U$ in the electrical field with a voltage U, that changes wrt location while moving around the saddle point. The constant k is still something like a spring constant. Since the motion is an oscillation, we can express the displacement by

$$\Delta x(t) = \Delta x_0(t)\, \sin(\omega_E\, t) \tag{9.68}$$

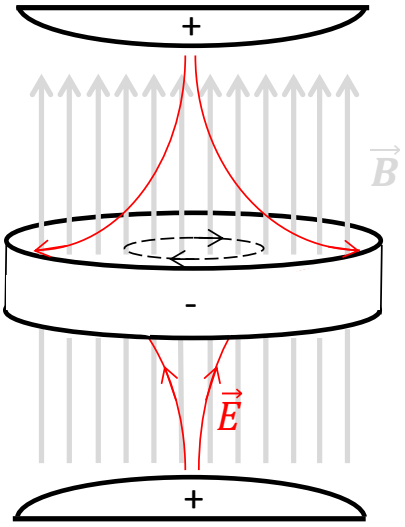

Figure 145: A *Penning trap*: An homogeneous magnetic field superimposed by an electric field $\vec{E}$ generated by two endcaps with the same electrical potential (here positive) and a ring shaped electrode (here negatively charged). Both, the ring electrode (if cut along the image plane here) and the endcaps have the shape of hyperboloids of revolution. A (here positively charged) ion would perform a circular motion (dashed circle) around the saddle point in the electrical potential created by the electrodes.

with an angular frequency ω_E.
The force F acting on the ion is (c. f. Section 2.1.8)

$$F = \frac{dE}{d\Delta x} = m\frac{d^2\Delta x}{dt^2}\,. \tag{9.69}$$

Inserting Equation 9.67 and Equation 9.68 into Equation 9.69, leads to the relation

$$\omega_E = \sqrt{\frac{k}{m}} = \sqrt{\frac{z\Delta U}{m}\frac{1}{d^2}} \tag{9.70}$$

with a factor d (here $d = \frac{\Delta x_0}{\sqrt{2}}$ in this simplified approach) that is determined by the geometry of the *Penning trap*. The frequency ω_E is called *axial frequency*.
In practice, the motion of the ions is more complicated and the actual resonance frequency $\bar{\omega}$ that is measured is a combination of ω_E and ω_C and one can show that

$$\boxed{\bar{\omega}^2 = \left(\omega_C - \frac{\Delta U}{2d^2\,B}\right)^2 + \left(\frac{\Delta U}{2d^2\,B}\right)^2 + \frac{z\,\Delta U}{m\,d^2}}\,. \tag{9.71}$$

Since $\bar{\omega}$ is measured and U and B are given by the experiment, the ratio m/z can be determined.

9.5.3 The Mass Spectrum

Depending on its mass and charge, every ion will have its own resonance frequency $\bar{\omega}$ and the measured signal is a composition of all these contributions, where the amplitude for a given frequency is proportional to the number of ions having the particular mass to charge ratio. The frequency spectrum, hence the $\frac{m}{z}$ spectrum, can be analyzed via a *Fourier transformation* (Section 2.4). As already mentioned, FT−ICRMS is the most accurate method

reaching resolutions up to $\frac{m}{\Delta m} \approx 10^{10}$ and even modest facilities easily reach $\frac{m}{\Delta m} \approx 10^7$; and it is also the fastest method having a temporal resolution of $\Delta t \approx 1\ \mu s$.
But what do we expect to see in the $\frac{m}{z}$ spectrum? We know, that every element has a distinct mass (usually measured in u) and thus also bio molecules (and their fragments) will have a mass which is a combination of these element masses. The dominant peak is called *base peak* and the amplitudes in the spectrum are normalized such that the height of the base peak equals 100%. Also the charge is quantized: ideally all ions have $z = \pm 1e$, but even if not, any charge z will be an integer times the elementary charge e. Thus, we would expect to see a *discrete* spectrum with distinctive peaks. The x−axis of the spectrum corresponds to the mass/charge ratio (m/z) and is measured in units of u/e, which is called a *Dalton*[107].
We even know, that only a limited number of different mass combinations is possible, since for example proteins are made of (only) twenty different amino acids and they themselves contain mainly H, C, N, O and two (cysteine and methionine) contain sulfur. Also the fragments we obtain after digestion and ionization (Section 9.5.1) are not arbitrary (see later in Section 9.5.4). Hence, it seems feasible to infer the components and the sequence of a given peptide (or any other molecule) by the corresponding mass spectrum. However, we have to take two things into account: The first problem is that the atom masses *are not* integers times u. For example a carbon atom does not have the mass $12\ u$, but $12.001\ u$, iron does not have the atomic mass of $56\ u$, but $55.845\ u$ and so on. Thus, we have to distinguish between the *nominal mass* and the *monoisotopic mass*:

nominal mass: atomic mass of the predominant isotope rounded to the next integer, e. g. $12\ u$ for carbon, $14\ u$ for nitrogen etc.

monoisotopic mass: actual atomic mass of the most abundant isotope, e. g. $12.011\ u$ for carbon, $14.007\ u$ for nitrogen etc.

The difference between nominal mass and monoisotopic mass is called *mass defect*. One reason is that the nucleons (protons and neutrons) in the nucleus are bound and the loss of binding energy equals a loss of mass ($E = m\,c^2$). Thus, the sum of masses of the nucleons making up a nucleus is *larger* than the mass of the nucleus[108]. The other reason is that the unit u is defined as $\frac{1}{12}^{th}$ of the mass of ^{12}C that has six neutrons and six protons. However, heavier atoms (led, gold, tungsten etc) tend to have more neutrons than protons and therefore cannot have atomic masses that are an integer product of u.
The second fact we have to take into account is already implied by the notation of *monoisotopic mass*: we have a mixture of different isotopes of each atom, hence different masses. A selection of isotopes of some biologically relevant atoms and their relative abundances is given in Table 2.

According to the abundances of its isotopes, the atoms of every single element cause several peaks in the mass spectrum. Suppose for example we would measure the mass of Cl_2, how many peaks would we detect and what would be their amplitudes?
According to Table 2, chlorine has two main isotopes, ^{35}Cl with a mass of $\approx 35\ u$ and a

[107] John Dalton, 1766 − 1844
[108] This works of course for any binding energy. Due to the loss of gravitational binding energy, Earth's mass is one million tons less than the sum of the mass of its components.

isotope	rel. abundance
$^{1}\mathrm{H}$	98.90 %
$^{2}\mathrm{H} \equiv \mathrm{D}$	0.02 %
$^{12}\mathrm{C}$	98.90 %
$^{13}\mathrm{C}$	1.10 %
$^{14}\mathrm{N}$	99.60 %
$^{15}\mathrm{N}$	0.40 %
$^{16}\mathrm{O}$	99.60 %
$^{17}\mathrm{O}$	0.04 %
$^{18}\mathrm{O}$	0.20 %
$^{32}\mathrm{S}$	94.99 %
$^{33}\mathrm{S}$	0.75 %
$^{34}\mathrm{S}$	4.25 %
$^{35}\mathrm{Cl}$	75.77 %
$^{36}\mathrm{Cl}$	trace
$^{37}\mathrm{Cl}$	24.23 %
$^{54}\mathrm{Fe}$	5.85 %
$^{55}\mathrm{Fe}$	–
$^{56}\mathrm{Fe}$	91.75 %
$^{57}\mathrm{Fe}$	0.28 %

Table 2: Relative abundances of isotopes of biologically relevant elements (selection, see also [1] and [2]). Note, that in contrast to most other elements, chlorine has *two* abundant isotopes causing several prominent peaks in the mass spectra of substances containing e. g. $\mathrm{CH_3Cl}$.

relative abundance of 75.77 % and ^{37}Cl with a mass of $\approx 37\ u$ and a relative abundance of 24.23 %. The molecule Cl_2 can be generated by *four* combinations of the isotopes: $^{35}\mathrm{Cl} - ^{35}\mathrm{Cl}$, $^{37}\mathrm{Cl} - ^{35}\mathrm{Cl}$, $^{35}\mathrm{Cl} - ^{37}\mathrm{Cl}$ and $^{37}\mathrm{Cl} - ^{37}\mathrm{Cl}$, but we will find *three* peaks in the spectrum since the second and the third combination are indistinguishable. The first combination has a relative abundance of 75.77 %×75.77 %=57.76 %, the mixed combinations will have a relative abundance of 75.77 %×24.23 %×2=36.48 % and finally the last combination has a relative abundance of 24.23 %×24.23 %=5.76 %.
The base peak corresponds to the $^{35}\mathrm{Cl} - ^{35}\mathrm{Cl}$ combination and will be normalized to 100 %. Thus, the peak of the mixed combination has the amplitude 36.48/57.76=65.5 % and the peak generated by the two heavier isotopes has an amplitude of 5.76/57.76=10.3 %. In summary we will see three peaks in the mass spectrum:

- at 70 u with 100% amplitude (base peak)
- at 72 u with 65.5% amplitude
- at 74 u with 10.3% amplitude

The *average mass* of Cl_2 reflects the mixture of the different isotope combinations according to the weighted mean and we obtain $70\ u \times 0.5776 + 72\ u \times 0.3648 + 74\ u \times 0.0576 = 70.96\ u$. One can easily imagine that molecules which are more complex than chlorine, like amino acids (the monoisotopic and the average masses of amino acids are shown in Table 3) or sugars, generate a distinctive, but predictable pattern of peaks in the mass spectrum according to the isotope ratios of their components. Since the isotope abundances are well known, the pattern in the mass spectrum are like a finger print.

1−letter code	3−letter code	Chemical formula	Monoisotopic	Average
A	Ala	C_3H_5ON	71.03711	71.0788
R	Arg	$C_6H_{12}ON_4$	156.10111	156.1875
N	Asn	$C_4H_6O_2N_2$	114.04293	114.1038
D	Asp	$C_4H_5O_3N$	115.02694	115.0886
C	Cys	C_3H_5ONS	103.00919	103.1388
E	Glu	$C_5H_7O_3N$	129.04259	129.1155
Q	Gln	$C_5H_8O_2N_2$	128.05858	128.1307
G	Gly	C_2H_3ON	57.02146	57.0519
H	His	$C_6H_7ON_3$	137.05891	137.1411
I	Ile	$C_6H_{11}ON$	113.08406	113.1594
L	Leu	$C_6H_{11}ON$	113.08406	113.1594
K	Lys	$C_6H_{12}ON_2$	128.09496	128.1741
M	Met	C_5H_9ONS	131.04049	131.1926
F	Phe	C_9H_9ON	147.06841	147.1766
P	Pro	C_5H_7ON	97.05276	97.1167
S	Ser	$C_3H_5O_2N$	87.03203	87.0782
T	Thr	$C_4H_7O_2N$	101.04768	101.1051
W	Trp	$C_{11}H_{10}ON_2$	186.07931	186.2132
Y	Tyr	$C_9H_9O_2N$	163.06333	163.1760
V	Val	C_5H_9ON	99.06841	99.1326

Table 3: Masses (in u) of amino acids [2] within the peptide chain (residue). The corresponding mass of the free amino acid is larger by one H_2O molecule, hence one would add a mass of $\approx 18\,u$.

9.5.4 Interpreting the Results

As seen in the previous section, determining the components of a bio molecule like a protein is almost trivial. However, inferring the *sequence* of a protein is a tricky task. In order to understand how the sequence of a protein can be derived from the mass spectrum I like to repeat some facts that should be well known to life scientists.
We know, that amino acids cannot be joined arbitrarily. A peptide will always end with the C− terminus $-COOH$ and will always start with the N−terminus $-NH_2$. Amino acids can only be joined by combining the C− terminus of one amino acid with the N−terminus of another one. We therefore get the junction $-CO-NH-C-$ and a water molecule H_2O with a mass of $18\,u$. Before being ionized, the protein is fragmented by a protease like e. g. trypsin, that mainly chops of the amino acids somewhere at their $-CO-NH-C-$ junction. If for example the peptide gets fragmented at the $-CO-NH-$ bond, the lhs amino acid will not end with $-COOH$, but with $-(C=O)^+$ since OH^- is missing if water is absent. But the peptide can also be fragmented at $-CHR-CO-$ (hence $-C-C-$) or at $-NH-CHR$, hence at $-N-C-$. Also the positive charge can be kept either at the C− terminal or at the N− terminal side. Thus, there is the following convention to denote the different cleavages at the amino acid in the peptide chain

- if the cleavage occurred at $-C-C-$: a_n (if positive charge is kept by N− terminal) or x_n (if positive charge is kept by C− terminal)
- if the cleavage occurred at $-C-N-$: b_n (if positive charge is kept by N− terminal) or y_n (if positive charge is kept by C− terminal)

- if the cleavage occurred at $-N-C-$: c_n (if positive charge is kept by N− terminal) or z_n (if positive charge is kept by C− terminal)

where the index n indicates the number of amino acids contained in the fragment. The notation is illustrated in Figure 146.

For example considering the spectrum shown in Figure 147 we find many peaks of different

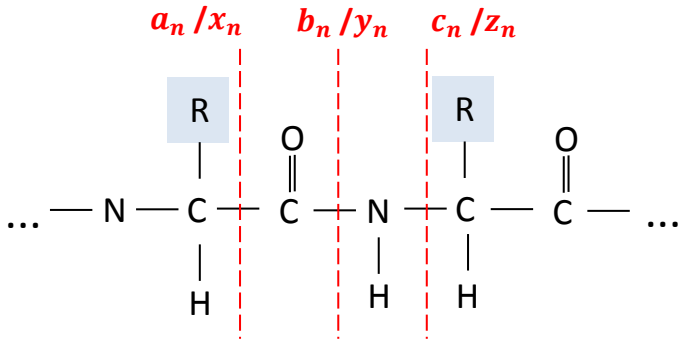

Figure 146: Notation of peptide fragmentation. If the positive charge is kept by the N− terminal, the letters a, b and c are chosen for the respective cleavage position, whereas the letters x, y and z are used if the positive charge is kept by the C− terminal.

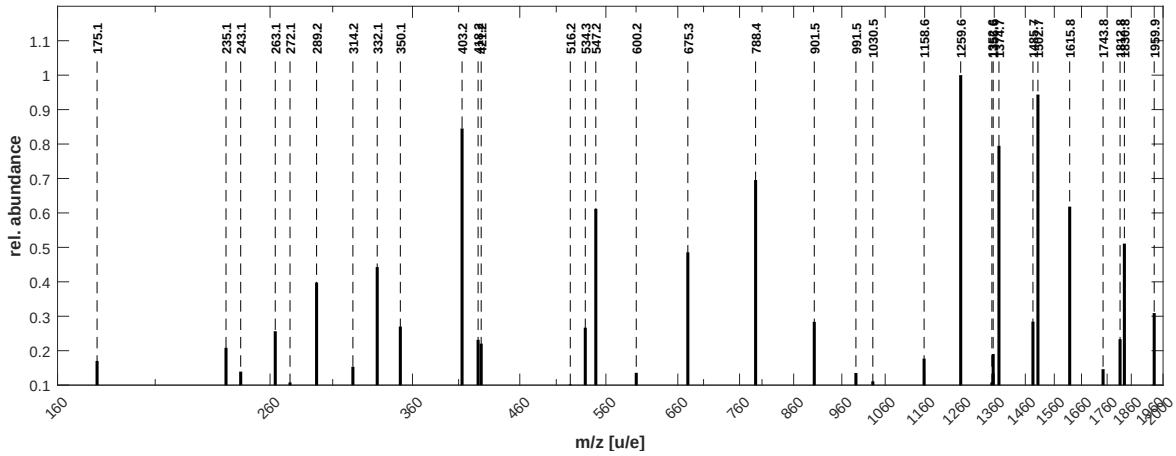

Figure 147: Mass spectrum of the peptide $DFSALESQLQDTQELLQEENR$. For clarity, peaks with a height of less than 10% are omitted. Data taken with permission from Victor Solis.

heights that might look confusing in the first glimpse. However, we can guess that peaks corresponding to masses below $\approx 220\,u$ might indicate the presence of single amino acids, although this interpretation is not unambiguous since some amino acids like glycine can form dimers in this mass range. We also know usually that the enzyme that was used to cut the peptide prefers particular sides and we thus do not need to consider the whole combinatorical space of fragment combinations.

Hence, starting with the first peak at $175.1190\,u/e$ we could guess that it is caused by only *one* amino acid. Looking up the possible candidates in Table 3 we find that for example arginine has an average mass of $156.1875\,u$ and thus the free amino acid (including H_2O with a mass of $18.015\,u$) would have a total mass of $174.2025\,u$. If arginine was cut at position b or y (Figure 146) and got protonated, i. e. the positive charge (an hydrogen atom of mass $1\,u$) was kept at the $C-$terminal, then we would see a peak at $\approx 175.20\,u$.

The next peak at $235.1077\,u/e$ can be interpreted as the fragment DF (aspartic acid and phenylalanine). The mass of the fragment equals $115.0886\,u\ +\ 147.1766\,u$ (Table 3), hence $\approx 262.3\,u$ in the peptide chain. However, when cut at position a (Figure 146) a CO of mass $\approx 28\,u$ gets lost and an hydrogen atom is added, leading to a final mass of $\approx 235.1\,u/e$. If cut on position b, the CO group is still present and when protonated (one hydrogen atom) we arrive at a mass of $\approx 263.3\,u$. In deed, this peak is also clearly detected in the spectrum shown in Figure 147 and therefore the presence of a DF fragments seems evident.
Of course when identifying the *composition* of an fragment, we do not automatically infer its *sequence* (here DF vs FD). Fortunately, the experiments usually yield many different fragments of different sizes with overlapping information. For example the sequence of the fragment $QELLQ$ (c. f. Figure 147) can hardly be identified if only the fragments QE, $QELL$ or LQ are present in the spectrum, but as soon as the fragment LL is detected the number of possible combinations gets reduced from $\frac{5!}{2!2!} = 30$ to $\frac{4!}{2!} = 12$ (see Section 2.6.4) and so on.
In practice, a software calculates the most likely sequence or the spectrum is submitted to a data base like [2] yielding the most likely spectrum (spectra) explaining the data.

Exercise:

Isoleucine and leucine have the same mass. Discuss possible experiments that help to distinguish between both.

References

[1] Edmond de Hoffmann & Vincent Stroobant *"Mass Spectrometry"*, Wiley 2007; 3rd edition.

[2] Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A. *'ExPASy: the proteomics server for in-depth protein knowledge and analysis"*, Nucleic Acids Res. 31:3784-3788(2003); PUBMED: 12824418.

END