

Size-resolved modeling of diatom populations: Old findings and new insights

Jonas Ziebarth¹, Werner M. Seiler², and Thomas Fuhrmann-Lieker¹

¹Institute of Chemistry and Center for Interdisciplinary Nanostructure
Science and Technology (CINSaT), University of Kassel

²Institute of Mathematics, University of Kassel

April 30, 2022

Abstract

The unique size reduction and restitution cycle of diatoms requires special attention in modeling their population dynamics. Matrix models based on distinct size classes are generally necessary to fully address periodic changes in cell size distribution and understand infra-annual algae blooms. Here, different approaches for such size-resolved models are compared, with their conclusions about oscillatory behavior in diatom populations. Particularly, the coupling of model parameters with environmental signals that may act as zeitgebers are discussed.

1 Introduction

The simulation of ecosystems by population dynamics of contained species has a long tradition[11]. Normally, coupled differential equations are used that describe the interdependence of species or the reaction to changing environmental conditions. Mostly, a species is considered as a single compartment characterized by the number of individuals, but in some models age classes such as juvenile, adult, and senescent are distinguished[44]. This idea can be transferred for treating different size classes such as in diatoms. Among unicellular organisms, diatoms are unique due to their size reduction-restitution life cycle that defines an intrinsic long-period rhythm, described by William Lewis as “sex clock”[46]. For modeling this means that the size variance has to be taken into account if real scenarios have to be described properly. Only a few contributions in literature have been devoted to treating this issue in some mathematical depth, despite the ecological significance of diatoms for carbon fixation and therefore the world climate[7] [71]. This review will examine the very different approaches for size-resolved population dynamics described in literature. After an introduction of the size reduction-restitution cycle and the important parameters

that need to be considered in models, we discuss experimentally detected deviations of the simple scheme. For instance, we will meet asynchronization of the two halves that changes the growth pattern and the timing of the proliferation process. Having set the biological background behind modeling, we introduce the basic mathematical concepts that are needed for understanding the various models. The main section is devoted to a detailed description of the different modeling approaches. Finally, some conclusions are made concerning the dependence of model parameters and thus the internal timing on external variables reflecting fluctuations in the environment.

2 The MacDonald–Pfitzer rule and the need for matrix descriptions

The size-diminution and restitution life cycle was first described independently by John MacDonald and Ernst Pfitzer in 1869[47] [60]. The mechanism is a consequence of the solid silica frustule of diatoms which prevents a growth in diameter if once formed. The two cell halves of the frustule overlap like the halves of a Petri dish. The larger half is denoted as epitheca, the smaller one as hypotheca. When a cell divides vegetatively, the missing half is replenished inside, thus forming a new hypotheca in each of the parental thecae. This leads to a decreasing mean cell size of the diatom population. On the other hand, the cell size distribution increases since the individuals inheriting the larger epitheca also continue to divide. If the generation time for all individuals is equal and no cell is removed, a distribution of size classes according to Pascal’s triangle is formed. If cells become too small for further vegetative division they change from mitosis to meiosis and switch to sexual reproduction. There are variants of sexual reproduction involving gametes and auxospores (see below) but at the end of the sexual process new initial cells of large size are generated, starting the vegetative process from the beginning (Fig. 1). In field studies, different cohorts derived from initial cells that were formed at different times (in different years) can be clearly distinguished which demonstrates the cyclic character of the process in natural populations. In order to keep the sizes within each cohort together, however, there seem to be mechanisms counteracting the broadening of the size distributions. We will discuss this important point further below.

It should have become evident that the state and further evolution of a diatom population strongly depends on the size distribution. For each individual cell, the probability between further vegetative division or sexual reproduction is decoded in its size. Conveniently, distinct size classes can be defined for mathematical modeling, so the state of the population is expressed as a vector. The transformation over time can then be implemented as a matrix acting on this vector and changing the cell size distribution. Different transition or retention parameters in this matrix represent by asymmetric division, replenishment of initial cells, and cell death. This bears some resemblance to age classes in other population models but the number of size classes needed here is typically much larger. So also in mathematical modeling diatom populations turn out to be unique and justify a separate treatment as in this review.

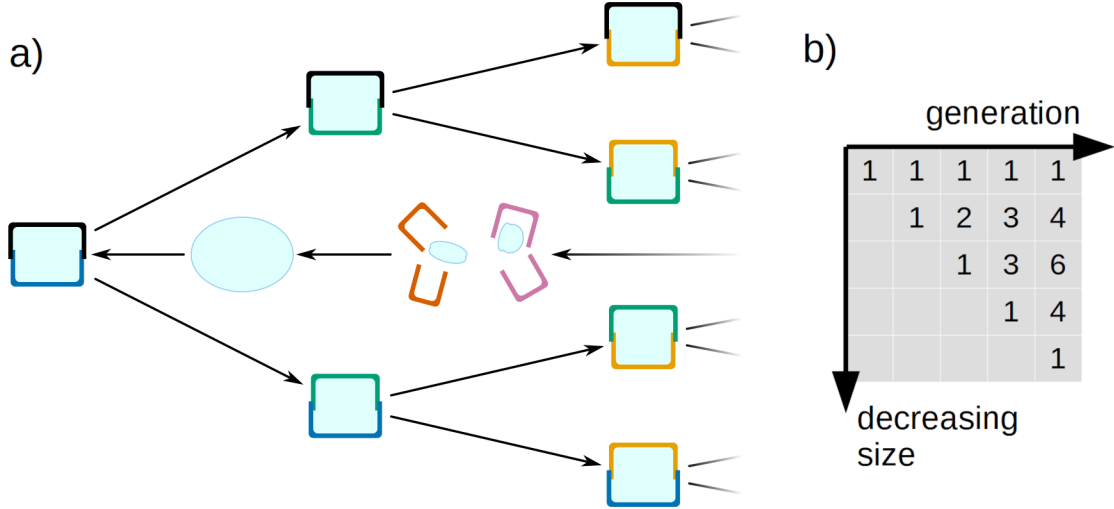


Figure 1: a) The diatom division scheme according to the MacDonald–Pfitzer rule. From left to right: vegetative process, from right to left: closing the life cycle by sexual reproduction. b) Change of size distribution from generation to generation according to Pascal's triangle.

3 Cardinal points and cycle lengths

3.1 Considered cardinal parameters

For the different sizes diatom cells may attain in the course of their life cycle, so-called cardinal points were first defined by L. Geitler of the University of Vienna. The cardinal points constitute the most important parameters needed for any size-resolved modeling of the diatom life cycle. These four cardinal points are:

- (i) the size of the initial cells, f
- (ii) the upper sexual size threshold, i.e. the upper size limit below which cells are able to produce gametes for sexual reproduction,
- (iii) a lower sexual size threshold, i.e. a lower size limit below which cells are not able any more to produce gametes,
- (iv) a final minimum size, at which cells are not able to further divide and eventually die.

That means that the cells of a culture can be divided into three subpopulations, limited by the respective cardinal points: Large cells that proliferate only by vegetative division between (i) and (ii), medium-sized cells that are capable of both vegetative and sexual reproduction between (ii) and (iii), and small cells between (iii) and (iv) that divide vegetatively until they are not viable any more. This last category is not relevant for closing the life cycle by the sexual phase, and therefore may

be neglected in studies aiming at cycle lengths. On the other hand, they contribute to the size distribution of a population as well and have to be taken into account when analyzing experimental distributions.

Since there are many cell divisions within each of these categories, the incremental size diminution in one division step, or alternatively the total number of steps between two cardinal points is a further parameter that has to be specified. It should be mentioned that size could be defined by diameter of centric diatoms, respectively the long axis in pennate diatoms, or by volume. Typically, lengths are easier detectable by microscope, whereas other methods like counting by electrical conductivity (Coulter counter) address rather the volume[58], or even the surface (fluorescence staining of cell walls)[61]. The scaling between length and volume is not always clear, since it depends on the species[5]. In long pennate diatoms like *Synedra tabulata*, length and volume are directly proportional[16].

These parameters mentioned so far affect only the sizes. For describing temporal events, the time between two mitotic cell divisions τ is an important parameter. It surely depends on environmental conditions and may be subjected to a certain distribution function (see also section 4 for a discussion on asymmetry)[1] [52]. Typically, this time has to be derived from the growth rate from counting of cells or turbidimetric measurements. The cell division rate, based on binary logarithm, can be calculated from two population numbers X_1 and X_2 with a time leap Δt by

$$\mu_D = \tau^{-1} = \frac{\log(X_2/X_1)}{\Delta t \log 2} \quad (1)$$

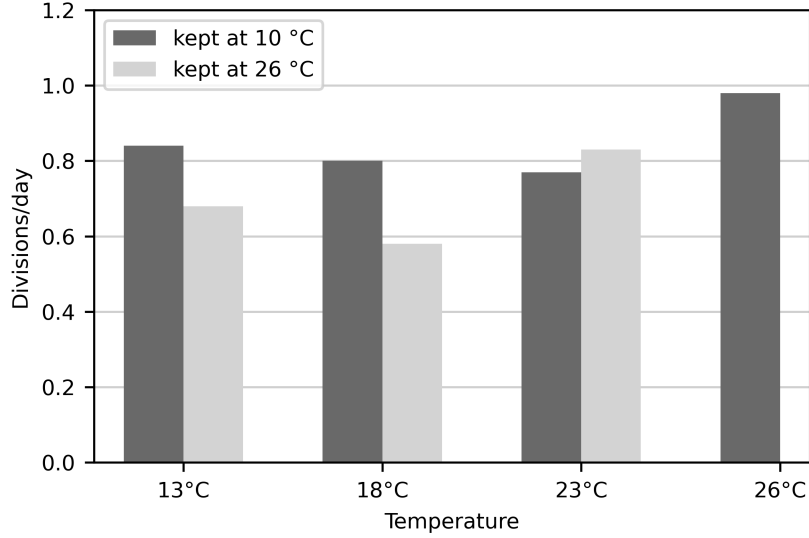
and is obtained with higher reliability by fitting $\log X$ versus time. Attention should be paid to the fact that sometimes the rate is given as growth rate $\mu = \mu_D \ln(2)$ on base of natural logarithms.

The sexual process involving gametogenesis and auxosporulation takes a certain amount of time that may be modeled as a multiple of τ .

3.2 Factors determining cardinal points

The formation of initial cells shall be considered as first step in the cycle. For simplicity in modeling, it may be reasonable to assume a certain fixed size for the initial cells. However, it can be expected that the size of initial cells is subjected to some distribution. For example, Bethge reported a ratio of 2-4 for initial cell size relative to the parental cell size for species in the *Melosira* genus[9]. For string-forming species of the genus *Aulacoseira*, a plot of initial cell diameter versus maternal filament diameter revealed an almost linear dependence, with a ratio of sizes between initial cells and their parental cells in the range 2-3[23]. More recently, Davidovich added further data on linear dependence for *Nitzschia lanceolata*, *Synedra tabulata* and *Licmophora ehrenbergii*, finding a ratio between maximum and minimum size of initial cells of 1.5 to 2[16].

In the vegetative phase, the timing for the sequence of generations follows the cell cycle. The cell cycle is controlled by some checkpoints that ensure that the right conditions for proliferation are met[34]. For instance, nutrients and silica required for



ostrearia.png

Figure 2: Temperature dependence on the doubling rate for *Haslea ostrearia*. Strains were adapted to 10°C and 26°C and then exposed to different temperatures. Constructed with data from [21].

the formation of a new theca have to be available. Notably, there is also a checkpoint for blue light that is sensed by an aureochrome photoreceptor [36]. This enables the synchronization of the cell cycle with the diurnal rhythm and can also be applied to synchronize cell cultures, as demonstrated for *Seminavis robusta* [28] and *Phaeodactylum tricornutum* [35]. Davidovich *et al.* investigated the influence of temperature on the rate of cell divisions for *Haslea ostrearia* and *Haslea karadagensis*. Data are given in figure 2 and show division times between two days at low and one day at higher temperature [21]. Williams characterized the influence of salinity for 14 species of pennate diatoms and measured maximum rates between 0.6 and 3.2 divisions/day that were found to correlate with cell volume and area/volume ratio [74]. Much slower rates are reported for *Stephanodiscus neoastraea* in an northern Irish lake with only 3-7 divisions in the spring growing period [37].

The checkpoint for the switch to sexual reproduction is one of the mysteries that is not fully understood. By artificially altering the cell size using nutritional as well as operative methods, Von Stosch at the University of Marburg was able to show in the 1960's that this threshold is indeed determined by the cell size and not by the age of a population [68] [49]. For *Stephanodiscus neoastraea*, nitrate was found to be a prerequisite to start sexual reproduction in fall [37]. For *Seminavis robusta*, it was recently found that the light quality (color) also influences sexual reproduction [10].

Typically, new initial cells are formed within auxospores. The kinetics of auxospore formation depends on the detailed mechanism of sexual reproduction of which a large variety is known [39] [13]. Some bypasses exist that allow single cells to form new enlarged initial cells without the need of combining genetic material stemming from another cell. One example is vegetative enlargement [68], another is automixis in which

gametes produced in one parental cell immediately fuse to form the auxospore[13]. In these uniparental cases there is no need to consider encounter dynamics of gametes by diffusion. In biparental auxosporulation we distinguish between homothallic strains in which one clonal strain produces two types of gametes, and heterothallic strains, in which different mating types exist in separate strains. Since gametes from different parents have to combine it is no wonder that the success rate of auxosporulation depends on the cell density[64].

The lower limit for sexual reproduction is not described for all diatom species. If such a lower limit exists, the life cycle is called a “closed cycle”, if not, it is referred to as an “open cycle”[39].

Finally, aging and natural cell death shall be discussed. Obviously, there is some difficulty to maintain the silica pattern with possibly self-organized pores in very small frustules. Also, there would be a crowding of organelles, so it is understandable that there is some size limit for viability, and very small cells will eventually die. Jewson addressed the important question whether some aging exists in the vegetative phase of larger cells i.e. whether there is a maximum number of divisions an epitheca can undergo or whether the cells will leave the population after a certain number of smaller offspring[37] [38]. For *Stephanodiscus neoastraea* he deduced a lifespan of six to eight generations and extrapolated a similar value for *Aulacoseira subarctica*. By the removal of older cells from the scheme, an asymmetry is induced that changes the size-resolved population dynamics. The occurrence of such asymmetries and consequences for modeling will be further treated in section 4.

3.3 Experimental data

Experimental data normally stem from investigations of the size distribution of natural diatom populations in some time intervals. Such data may have some gaps due to the organization of the field study and are influenced by strong weather fluctuations and singular natural events. Nevertheless, they give important insight into the cycles of real ecosystems. Notable studies include the investigation of populations in lakes, such as *Tabellaria fenestrata* and *Aulacoseira (Melosira) islandica* in the Zürichsee[56], *Nitzschia sigmaidea* in Blackford Pond[48], and *Aulacoseira subarctica* and *Stephanodiscus* sp. in Lough Neagh[37] [38]. Investigations of sediments that made the assignment of distinct years possible might open a window in the past[56]. Only in some cases in which cultures can be held successfully in the laboratory, parameter data under controlled conditions are available. The most conspicuous finding of these investigations are multimodal distributions of cell sizes which are used for calculating life cycle length, or more precisely the periodic occurrence of initialization events by large cells. For plotting and detection of deviations from Gaussian distributions, Rankit plots[59] have been favorably used as visualization tool[48]. In table 1, some proposals for life cycle lengths are compiled. Interpretation of data might be ambiguous, so some data have been reanalyzed later. For example, Jewson suggested a shorter life cycle for *Aulacoseira islandica* than the one originally reported by Nipkow, whereas for *Tabellaria fenestrata* Mann suggested a longer life cycle. For *Aulacoseira subarctica* he proposed a theoretical maximum cycle of 15 years

corresponding to over 100 divisions, while the upper sexual size threshold is already reached after 4-6 years.

Species	Cycle length	Size range	Ref.
<i>Seminavis robusta</i>	<1 year	25(50)–73 μm	[12] [14]
<i>Eucampia zodiacus</i>	$\sim$ 1 year	10(15)–80 μm	[57]
<i>Cocconeis scutellum</i>	1 year	15–49 μm	[50]
<i>Fragilariopsis kerguelensis</i>	$\sim$ 1 year	10(31)–76 μm	[42][4]
<i>Pseudo-nitzschia delicatissima</i>	<1 year	8(80)–94 μm	[2]
<i>Pseudo-nitzschia multistriata</i>	2 years	30(55)–82 μm	[15]
<i>Tabularia fasciculata</i>	1–1.5 years	10(113)–248 μm	[19]
<i>Haslea karadagensis</i>	<1–2 years	22(52)–97 μm	[20]
<i>Stephanodiscus</i> sp.	1–2 years	25–66 μm	[62]
<i>Haslea ostrearia</i>	1–2.5 years*	17(68)–140 μm	[18]
<i>Melosira helvetica</i>	3–4 years	4–31 μm	[9]
<i>Pseudo-nitzschia australis</i>	2 years	75(121)–170 μm	[33]
<i>Pseudo-nitzschia pungens</i>	3 years	75(115)–185 μm	[33]
<i>Pseudo-nitzschia multiseriata</i>	3 years	39–146 μm	[17]
<i>Chaetoceros dichæta</i>	>3 years*	6(15)–44 μm	[5]
<i>Stephanodiscus</i> sp.	3–4 years	20(34)–74 μm	[37]
<i>Coscinodiscus wailesii</i>	<4 years*	90(200)–350 μm	[55]
<i>Nitzschia sigmoidea</i>	>5 years	200(300)–424 μm	[48]
<i>Tabellaria fenestrata</i>	2–8 years	40(40)–90 μm	[56] [48]
<i>Aulacoseira subarctica</i>	4–6 years	3(8)–14 μm	[38]
<i>Aulacoseira islandica</i>	40 years	4–22 μm	[56] [48]

Table 1: Reported life cycles and cardinal points of selected species. Cycle lengths marked with a star are rough estimates based on the reported size diminution speeds and the size range. The size in brackets marks the size at which cells were observed to become sexually inducible.

4 Asymmetry, delay and Fibonacci growth

4.1 The Müller model

Since the introduction of the MacDonald–Pfitzer rule in 1869, it has been evident that the theory cannot represent full reality. The rule is commonly used as a simple description of the diatoms life cycles but over time diatomists have repeatedly disputed its simplicity. Some have specifically rejected the idea that diatoms divide uniformly and would thus generate a binomial distribution of cell sizes.

In 1883, Otto Müller proposed an approach to the cell division scheme which involves a delay [53]. He defined a process of simultaneous division as the vegetative cell division in which cells divide into two daughter cells which then in turn continue

to divide simultaneously again. Müller, however, found that this was not in agreement with his observations of some diatom species. Instead, if the larger or older daughter cells would divide first and the younger daughter would delay division by one generation, his findings could be explained. Furthermore, a simultaneous division would lead to a large number of auxospores at one point. As this cannot be found in nature, a delay of some kind should slow down the formation of auxospores and the progression of the size reduction.

The diatom *Ellerbeckia (Melosira) arenaria* possesses particular modifications of the girdle band that allow the deduction of the age of the cells and thecae. Müller first found evidence of a delay when investigating diatoms of the genus *Terpsinoë*. These chain forming diatoms often form groups of three (triplets) hinting at a faster division of one cell. Lacking the proper tools to assign the age of the thecae, he could not prove his assumption. However, in *Ellerbeckia arenaria* the girdle bands can indicate the age of the theca. Because of its division mechanism, this species has girdle bands only on one theca opposite to a thickening [54] or a step [38] on the next cell (inset B Fig. 3). During division, a new girdle band is formed and smaller thecae are enclosed. Again, the girdle band ends at a step which is formed on the adjacent new valve of the next cell (inset A Fig. 3). Hence, two parameters can be assigned to each theca in the chain, thereby creating four types: the existence or absence of a step and whether the thecae are enclosed or free. Every cell is then found to be part of a twin or triplet group which has free thecae as its boundary. If the knowledge about the parameter sequence in the chain and the grouping of the cells is combined, the age of each theca can be deduced and the formation of the chain could be traced back until the initial cell. Applying this method, Müller concluded that the younger cell which contains the smaller half of the mother cell and a fresh even smaller theca, must delay its division for one generation. He hypothesised that the older daughter would inherit richer material, by quality or quantity, from their mother.

In 1884, following up on his first publication, Müller presented a deeper analysis of the composition of the *Ellerbeckia arenaria* chains[54]. As the motivation of the analysis was to find the influence of the delay on auxosporulation, it is important to understand the number of cells and their corresponding sizes first. The delayed division scheme generates a chain whose total number of individuals follows the Fibonacci sequence (Fig. 3). The Fibonacci sequence dominates the division process in all aspects of single theca generation as well as frustule grouping. When transitioning into the next generation, a twin of an older and a younger cell produces one additional cell which turns it into a triplet with a fresh triplet-twin and an undivided triplet cell. A triplet reproduces another triplet from the triplet-twin and a twin from the triplet cell. After the first few divisions, this leads to Fibonacci growth of twin and triplet numbers. Furthermore, Fibonacci growth also occurs with the numbers of the four different theca types. Although, in this case there is one exception for the initial cell which can cause a variation by two thecae depending on the initial cells composition. After establishing the total number of cells, the number of cells in a certain size after a random number of generations are essential since the cells of the smallest size can auxosporulate. These numbers are obtained with arithmetic sequences of the corresponding order. Thus, the onset of auxosporulation in the n th

generation in a simultaneous division scheme is delayed to the $2n + 1$ th generation. However, the numbers of auxospores that are produced from this generation on are the same as with simultaneous division.

Further investigation of the chain, provides insights into the symmetry of the chains. Here, symmetry is defined as the symmetry of the arrangement of the four types of theca from each end of the chain. The individuals can be grouped in different levels. The first level would be the twins and triplets which in turn can be regrouped on multiple levels. The analysis and breakdown of this grouping leads to three key findings. Firstly, the previous elements of the chain are always reproduced until the point where the chain ended in the previous generation. Secondly, the entire chain is symmetrical with the exception of up to four cells in the center and the outmost thecae. Lastly, the chain can be broken down into smaller groups which are in turn symmetrical except for their outmost thecae. The length of these groups is also derived from Fibonacci growth.

4.2 The Laney model

The previous section discussed asymmetric division and favoring of the older daughters. On the contrary, Laney *et al.* noting that much recent research has been focusing on phenotypic differences of daughter cells in diatoms, but not on the timing of frustule inheritance, just found the opposite, namely favoring of the younger daughters[43].

Laney *et al.* investigated the division rates and cell volume of *Ditylum brightwellii* using a modern time-lapse imaging approach[43]. The pictures were taken at low density using infrared microscopy. Because of higher density after only few divisions and close proximity of the cells the analysis was mostly performed manually and therefore had to be limited. Nevertheless, daughters inheriting the hypotheca were found to need about 4 % less total light dosage to divide. The smaller cells also received more of the mother’s cytoplasm. Interestingly, the division speed did not seem to accumulate over multiple generations. Nonetheless, results were comparable to observations in *E. coli* which suggested a common behavior across species. A linear regression of the received additional cytoplasm and the light dosage supported a relationship between the volume of cytoplasm and division rate as well. Moreover, the analysis indicated that the younger daughters have an advantage independent of cytoplasm volume. This would be in agreement with the previously introduced theory that the quality of inherited material is better. Specifically, it is assumed that the smaller daughter receives material that is fresh and received less oxidative stress. Thus, the chance for success in sexual reproduction at the end of a size-diminution cycle would be higher.

The authors propose that the asymmetrical division of diatoms could also play a role in tuning the diatom sex clock. This would be an important modeling parameter for the clock besides the cardinal points that is depending on the diatom life cycle. Ecologically, a faster division of smaller cells would lead to more copies of a mother cells genome and permit more of its progeny to reach sexual reproduction. Müller had already concluded that delays in cell division fundamentally alter the growth speed

and size distribution of a population. Laney *et al.* remark that it would be essential to further investigate the influence of different division speeds and subsequently improve modeling approaches. As the delays in both cases are assigned to different halves, a wide spectrum of systematic asymmetry in diatom division is suggested.

5 Continuous vs. discrete modeling

Population dynamics represents a typical *evolution problem*: we are interested in the evolution of certain quantities over time. A fundamental question in the design of mathematical models is then whether time is considered as a continuous or as a discrete quantity, as these options lead to rather different type of equations. A general overview over various mathematical approaches to population dynamics can be found in the textbook[11].

In everyday life, we perceive time as a continuous quantity, i.e. mathematically it is identified with the real numbers $\mathbb{R}$. In typical models, we then also treat quantities like the number X of diatoms in a given population as real numbers, although they are actually integers. Since X is usually a rather large number, this does not lead to relevant errors and allows us to consider X as a function $\mathbb{R} \rightarrow \mathbb{R}$ for which we can derive *differential equations*. Some classical introductions into their basic theory are the books[32] [70] [73].

In population dynamics, it is often also natural to think in “generations”: one combines all births within a certain time span into one generation and enumerates the different generations, i.e. mathematically time is now identified with the integers $\mathbb{Z}$. In principle, this makes it also natural to describe X by integers, i.e. as a function $\mathbb{Z} \rightarrow \mathbb{Z}$. However, most models contain real parameters like birth rates or survival probabilities and thus one must consider X as a real-valued function $\mathbb{Z} \rightarrow \mathbb{R}$ satisfying *difference equations*. Here, some typical introductory textbooks are[24] [40] [63].

In practice, one usually deals with systems and not with scalar equations. Thus, one deals with vector-valued functions $\mathbf{X}$ from either $\mathbb{R}$ or $\mathbb{Z}$ to $\mathbb{R}^n$. The meaning of the components of $\mathbf{X}$ can be rather different. They may represent different species or compartments, but vectors also naturally appear in age structured (Leslie[44]) models or in the case of diatoms in size structured models. If one uses a model with both age and size structure, then it may be useful to consider $\mathbf{X}$ actually as a matrix-valued function from $\mathbb{R}$ or $\mathbb{Z}$ to $\mathbb{R}^{m \times n}$. For notational simplicity, we will in this section consider such models still as vector-valued with vectors in $\mathbb{R}^{mn}$.

In this section, we will briefly survey some basic mathematical facts for both types of equations with an emphasis on *linear* systems where much information can be gleaned with the help of elementary linear algebra. This concerns in particular the stability of equilibria and the existence of periodic solutions (i.e. oscillations). In many situations, it is relevant for the understanding of a linear system to have information about the dominant eigenvalue of the system matrix. In the context of population dynamics, such information can often be obtained rather easily with the help of the Perron–Frobenius Theorem. Therefore, we will also briefly discuss this theorem and its application.

5.1 Discrete dynamical systems

Let $\mathbf{X}$ be a function $\mathbb{Z} \rightarrow \mathbb{R}^n$ whose values are n -dimensional vectors decoding different species or size classes. Alternatively, we may consider $\mathbf{X}$ as an infinite sequence of vectors

$$(\dots, \mathbf{X}(-2), \mathbf{X}(-1), \mathbf{X}(0), \mathbf{X}(1), \mathbf{X}(2), \dots) .$$

As in evolutionary problems one is typically only interested in the future, one considers $\mathbf{X}$ only from some starting index n_0 on, which is usually chosen to be 0. Hence we will consider from now $\mathbf{X}$ as a function $\mathbb{N}_0 \rightarrow \mathbb{R}^n$ or as a sequence $(\mathbf{X}(0), \mathbf{X}(1), \mathbf{X}(2), \dots)$.

An (*autonomous*) *first-order system of difference equations* for $\mathbf{X}$ is of the form

$$\mathbf{X}(t+1) = \mathbf{f}(\mathbf{X}(t)) \quad (2)$$

with some function $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Equation (2) may be considered as a rule describing the population in generation $t+1$, if the population in generation t is known. Consequently, it is common to consider *initial value problems*: prescribing the values $\mathbf{X}(0) = \mathbf{X}_0 \in \mathbb{R}^n$, the rule (2) determines uniquely the population $\mathbf{X}(t)$ in all generations $t > 0$.

An *equilibrium* (or *steady state*) *solution* of an autonomous system is defined by a vector $\mathbf{V} \in \mathbb{R}^n$ such that $\mathbf{f}(\mathbf{V}) = \mathbf{V}$ (such vectors are also called *fixed points* of the function $\mathbf{f}$). Indeed, in this case the constant function $\mathbf{X}(t) = \mathbf{V}$ for all $t \in \mathbb{N}_0$ obviously solves (2). A *periodic solution with period ℓ* (for some $\ell > 1$) is defined by a vector $\mathbf{V} \in \mathbb{R}^n$ such that $\mathbf{f}^\ell(\mathbf{V}) = \mathbf{f}(\mathbf{f}(\dots(\mathbf{f}(\mathbf{V})))) = \mathbf{V}$ and $\mathbf{f}^j(\mathbf{V}) \neq \mathbf{V}$ for any exponent $1 \leq j < \ell$. Indeed, setting $\mathbf{X}(0) = \mathbf{V}$, $\mathbf{X}(1) = \mathbf{f}(\mathbf{V})$, $\mathbf{X}(2) = \mathbf{f}^2(\mathbf{V})$, $\dots$, $\mathbf{X}(\ell-1) = \mathbf{f}^{\ell-1}(\mathbf{V})$, we can construct a solution by continuing with $\mathbf{X}(t) = \mathbf{X}(t-\ell)$ for all indices $t \geq \ell$.

A *second-order system of difference equations* is of the form

$$\mathbf{X}(t+2) = \mathbf{f}(\mathbf{X}(t), \mathbf{X}(t+1)) , \quad (3)$$

i. e. the population in generation $t+2$ depends not only on the population in the last generation $t+1$ but also on the last but first generation t . This may be interpreted as a model for delay effects. In an *initial value problem*, it is now necessary to prescribe both the values $\mathbf{X}(0) = \mathbf{X}_0$ and $\mathbf{X}(1) = \mathbf{X}_1$, as now the rule (3) is able to determine the population $\mathbf{X}(t)$ only in all generations $t > 1$. One can extend in an obvious manner to higher-order systems.

The simplest difference equations are the linear ones and we will concentrate here on these. A *linear, first-order system of difference equations with constant coefficients* has the form

$$\mathbf{X}(t+1) = A\mathbf{X}(t) \quad (4)$$

where $A \in \mathbb{R}^{n \times n}$ is a real matrix with constant entries. In the context of population dynamics, A is often called the *propagator matrix*. Obviously, the zero vector (or more generally any vector $\mathbf{V}$ with $A\mathbf{V} = \mathbf{V}$) defines an equilibrium for such a system. Furthermore, such systems can be closed in solved form via elementary linear algebra.

We restrict here to the simplest case that A is a diagonalizable matrix which is also the most relevant case for applications. For such a matrix, we can find n linearly

independent eigenvectors $\mathbf{V}_1, \dots, \mathbf{V}_n$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ (i. e. solving the matrix equation $A\mathbf{V}_k = \lambda_k \mathbf{V}_k$; note that we do not require that $\lambda_i \neq \lambda_j$ for $i \neq j$). If we furthermore assume that all these eigenvalues and -vectors are real, then a fundamental system of real solutions is given by the functions

$$\mathbf{X}_k(t) = \lambda_k^t \mathbf{V}_k, \quad 1 \leq k \leq n \quad (5)$$

and the general solution is the linear combination $\mathbf{X}(t) = \sum_{i=1}^n c_i \mathbf{X}_i(t)$ with arbitrary coefficients $c_i \in \mathbb{R}$. The unique solution of an initial value problem with $\mathbf{X}(0) = \mathbf{X}_0$ is then obtained by solving the linear system of equations $\sum_{i=1}^n c_i \mathbf{V}_i = \mathbf{X}_0$ for the coefficients c_i (i. e. one expresses the vector $\mathbf{X}_0$ in the eigenbasis $(\mathbf{V}_1, \dots, \mathbf{V}_n)$).

If the matrix A also possesses complex eigenvalues, then (5) still represents a fundamental system of solutions, but at least some of them are complex. It is, however, easily possible to derive an alternative fundamental system consisting entirely of real solutions. If $\lambda = \rho e^{i\omega} \in \mathbb{C}$ is a complex eigenvalue of A , then its complex conjugate $\bar{\lambda} = \rho e^{-i\omega}$ is also an eigenvalue. Furthermore, if $\mathbf{V} \in \mathbb{C}^n$ is an eigenvector for λ , then $\bar{\mathbf{V}}$ is an eigenvector for $\bar{\lambda}$. Decomposing the eigenvector in its real and imaginary part, $\mathbf{V} = \mathbf{W}_{\text{re}} + i\mathbf{W}_{\text{im}}$, we can replace in our fundamental system the two complex solutions corresponding to λ and $\bar{\lambda}$ by the following two real ones:

$$\begin{aligned} \mathbf{X}_{\text{re}}(t) &= \rho^t [\cos(\omega t) \mathbf{W}_{\text{re}} - \sin(\omega t) \mathbf{W}_{\text{im}}], \\ \mathbf{X}_{\text{im}}(t) &= \rho^t [\sin(\omega t) \mathbf{W}_{\text{re}} + \cos(\omega t) \mathbf{W}_{\text{im}}]. \end{aligned} \quad (6)$$

The expressions (5) and (6) describing a real fundamental system of (4) easily allow us to make certain statements about the asymptotic behavior of solutions. If there exists an eigenvalue λ_k with $|\lambda_k| > 1$, then for any initial condition with $c_k \neq 0$ the solution will be unbounded for $t \rightarrow \infty$, i. e. the population of at least one species will become arbitrary large; one says that the system is *unstable*. Conversely, if all eigenvalues satisfy $|\lambda_k| < 1$, then any solution will converge for $t \rightarrow \infty$ towards $\mathbf{0}$, i. e. all species get extinct; mathematically, one speaks of an *asymptotically stable* system. If the absolute value is greater than 1 for some eigenvalues and less than 1 for some others, then one has a *saddle point*. (6) indicates that complex eigenvalues are related to *oscillatory behavior*: if $\rho < 1$, the oscillations are damped, whereas for $\rho > 1$, the amplitude grows towards infinity.

The situation is more complicated, if eigenvalues λ_k exist with $|\lambda_k| = 1$. If λ_k is real, this means that $\lambda_k = \pm 1$. In the case $\lambda_k = +1$, any eigenvector $\mathbf{V}_k$ defines an equilibrium; in the case $\lambda_k = -1$, a periodic solution with period 2 is given by $\mathbf{X}(t) = (-1)^t \mathbf{V}_k$ for any eigenvector $\mathbf{V}_k$. If λ_k is complex, then one can observe for appropriate initial conditions sustained oscillations. One should note that such oscillations do not necessarily lead to periodic solutions. For $\lambda_k = e^{i\omega_k}$, periodic solutions will only arise if ω_k is a rational multiple of 2π . Otherwise, one finds *almost periodic* solutions which never exactly repeat themselves, but come arbitrarily close to doing so.

Figure 4a shows in green the region in the complex plane where all eigenvalues must be located for an asymptotically stable equilibrium, namely inside the unit circle. For sustained oscillations, one must also have at least some eigenvalues on the

unit circle itself (shown in blue). The figure also shows that complex eigenvalues always appear in complex conjugated pairs, i.e. the eigenvalues lie symmetric with respect to the real axis. The *Jury test*[45] allows for effectively determining the number of eigenvalues outside of the unit circle (indicating instability) without explicitly computing the eigenvalues which is quite useful for higher-dimensional systems.

We remark that the case that the propagator matrix A is not diagonalizable can also be handled in closed form. The situation is more complicated then and one needs the so-called *Jordan normal form* of A . In this case, further possibilities for unstable behavior exist, namely multiple eigenvalues of absolute value 1 without the corresponding number of linearly independent eigenvectors. In such situations, one no longer has an exponential instability but only a polynomial one.

We thus conclude that using linear algebra allows us to solve any system of the form (4) with a fundamental system composed of the solutions (5). This explicit form furthermore permits statements about the asymptotic behavior of the solution. Assume for simplicity that one eigenvalue, say λ_1 , has the greatest absolute value, i.e. $|\lambda_1| > |\lambda_k|$ for all $k > 1$. We know from the discussion above that for the initial data $\mathbf{X}_0 = \sum_{i=1}^n c_i \mathbf{V}_i$ the solution is given by $\mathbf{X}(t) = \sum_{i=1}^n c_i \lambda_i^t \mathbf{V}_i$. It is now easy to see that we obtain for large times t

$$\lim_{t \rightarrow \infty} \frac{\mathbf{X}(t)}{|\lambda_1|^t} = c_1 \mathbf{V}_1. \quad (7)$$

Thus in the long run the solution mode corresponding to the eigenvalue λ_1 dominates all other modes which consequently effectively disappear, as they become negligibly small in comparison. If several eigenvalues have the same absolute value, then the long time behavior is controlled exclusively by the corresponding modes.

For nonlinear first-order systems of the form (2), some of the above sketched ideas about linear systems can be reused. Assume that $\mathbf{V}$ is an equilibrium of (2) and that the right hand side $\mathbf{f}$ is partially differentiable with respect to all of its arguments. Then we can form the *Jacobian* of $\mathbf{f}$ at $\mathbf{V}$

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{V}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{V}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{V}) & \dots & \frac{\partial f_n}{\partial x_n}(\mathbf{V}) \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (8)$$

and in many cases the local behavior of (2) near $\mathbf{V}$ can be analyzed by studying its *linearization* $\mathbf{Y}(t+1) = J\mathbf{Y}(t)$. If for example all eigenvalues of the Jacobian J have an absolute value less than 1, then the equilibrium $\mathbf{V}$ of (2) is asymptotically stable. Such a simple relationship between linear and nonlinear stability exists, however, only for *hyperbolic* equilibria, i.e. if none of the eigenvalues of J has absolute value 1. Otherwise, more sophisticated methods must be used.

5.2 The Perron–Frobenius Theorem

It has become apparent from the discussions in the previous section that information about the size of the eigenvalues of a matrix are crucial for assessing the stability

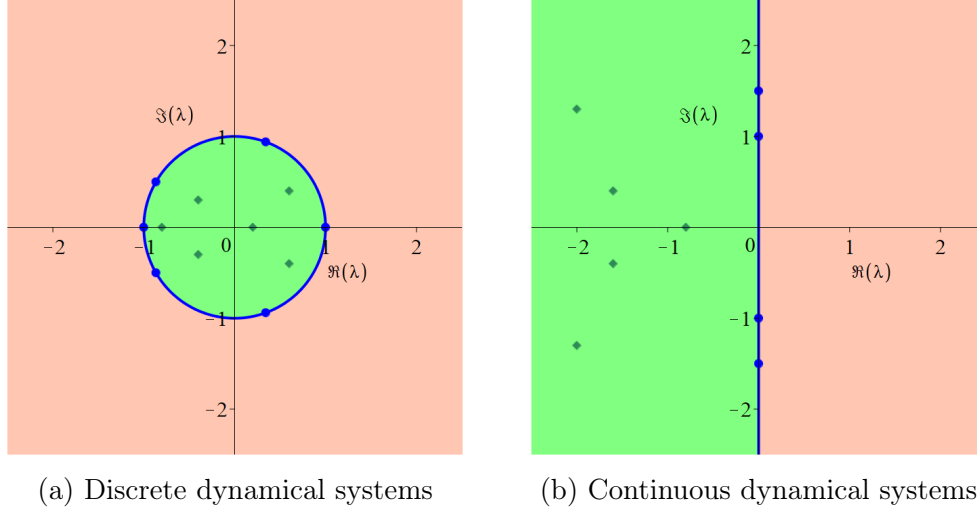


Figure 4: Stability and oscillatory regions (green and blue area, resp.) in the complex plane for (a) discrete and (b) continuous dynamical systems. The green and blue marks show possible locations of the eigenvalues for an asymptotically stable system and for a system with sustained oscillations, respectively.

of solutions or equilibria on the one hand and for the search for periodic solutions on the other hand. For small systems, it is often possible to determine explicitly all eigenvalues by hand or with the help of a computer algebra system. For larger systems, one usually fails here. The *Perron–Frobenius Theorem* provides information for certain matrices about the dominant eigenvalue which determines the long time behavior of a dynamical system as discussed above. We refer to the monograph[8] for an in-depth treatment of everything discussed in this section.

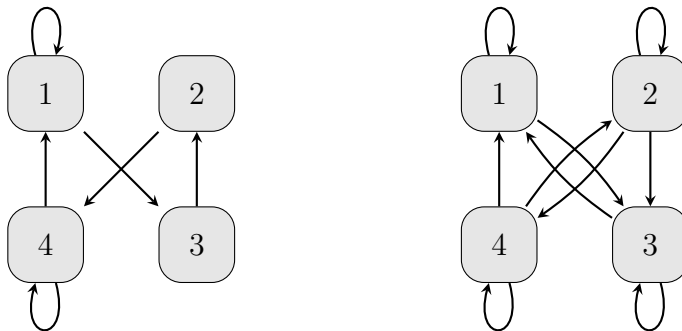
We first introduce some necessary terminology. Let $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ be a square matrix. The maximum of the absolute values $|\lambda|$ of the eigenvalues is the *spectral radius* $\rho(A)$ of A (note that it is always a real number, even if all the eigenvalues are complex). We say that A is *reducible*, if there exists a permutation matrix P such that PAP^{-1} is in block triangular form; otherwise A is *irreducible*. A *positive* matrix A has only positive entries, i. e. if $a_{ij} > 0$ for all indices i, j , and in a *non-negative* matrix also zero entries are allowed, i. e. $a_{ij} \geq 0$ now. Such matrices appear naturally in stochastic problems, but also in many biological applications like population dynamics.

A property like irreducibility is not easy to check directly from its definition. A practically useful approach is based on a directed graph $G(A)$ associated with any non-negative matrix $A \in \mathbb{R}^{n \times n}$. It has n vertices $1, \dots, n$ and a directed edge goes from i to j , if $a_{ij} > 0$. Thus, this graph encodes where the positive entries are located in A . A *path* of length ℓ connecting two vertices is a sequence of edges $i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_\ell$; it is called a *loop*, if $i_\ell = i_0$. The graph $G(A)$ is *strongly connected*, if for any pair (i, j) there exists a path from i to j . One can now prove that the matrix A is irreducible, if and only if $G(A)$ is strongly connected.

We consider as a simple example the following two non-negative matrices

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \quad (9)$$

which possess the following associated directed graphs



It is easy to see that in the graph $G(A)$ on the left we have a loop $1 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ which connects all vertices. Hence, $G(A)$ is strongly connected and the matrix A is irreducible. In the graph $G(B)$ on the right, there exists no possibility to reach 2 or 4 from 1. Hence, the graph is not strongly connected and the matrix B is reducible.

The Perron–Frobenius Theorem makes a number of statements about matrices with specific properties. The first assertion is that the spectral radius $\rho(A)$ of a non-negative matrix A is always an eigenvalue and that it always possesses an eigenvector with non-negative entries. Thus, we always find a non-negative *real* eigenvalue of maximal absolute value and in applications like population dynamics where negative values make no sense in solutions it is very important to have in addition an eigenvector without negative entries.

If the matrix A is in addition irreducible, then $\rho(A)$ is a simple eigenvalue and its one-dimensional eigenspace can be generated by a positive eigenvector (and $\rho(A)$ is the only eigenvalue with such an eigenvector). In the context of population dynamics, this means that no species is extinct. Furthermore, all other eigenvalues λ with $|\lambda| = \rho(A)$ are simple, too. The total number of such eigenvalues is called the *index of cyclicity* of A . If $\rho(A)$ is the only eigenvalue with this property, then A is a *primitive* matrix. The index of cyclicity can again be read off the graph $G(A)$; it is the greatest common divisor of the lengths of the different loops existing in $G(A)$. For the concrete irreducible matrix A defined in (9), the situation is very easy: we noted already above that there exists a loop of length 4 passing through all vertices; in addition, we obviously have loops of length 1 like $4 \rightarrow 4$ and hence the greatest common divisor of all lengths can only be one (there is no need to check whether further loops exist), so that A is a primitive matrix.

5.3 Continuous dynamical systems

The theories of differential and difference equations, respectively, share many similarities, but also differ in crucial aspects. We consider now functions $\mathbf{X}: \mathbb{R} \rightarrow \mathbb{R}^n$ whose values are n -dimensional vectors and always tacitly assume that $\mathbf{X}$ is sufficiently often differentiable. An (*autonomous*) *first-order system of differential equations* for $\mathbf{X}$ is of the form

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) \quad (10)$$

with $\dot{\mathbf{X}}$ denoting the first derivative of $\mathbf{X}$ (taken componentwise) and with some function $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ (the *vector field* corresponding to the equation (10)). Thus at each time t , the system allows us to compute the rate of change $\dot{\mathbf{X}}(t)$ from the current state $\mathbf{X}(t)$. Again it is common to consider *initial value problems* where one prescribes the initial state $\mathbf{X}(0) = \mathbf{X}_0 \in \mathbb{R}^n$, as normally any differential equation possesses infinitely many solutions and such an initial condition selects then a particular one.

An *equilibrium* (or *steady state*) *solution* of an autonomous system is defined by a vector $\mathbf{V} \in \mathbb{R}^n$ such that $\mathbf{f}(\mathbf{V}) = \mathbf{0}$, i. e. by the zeros of the vector field $\mathbf{f}$. Indeed, in this case the constant function $\mathbf{X}(t) = \mathbf{V}$ for all $t \in \mathbb{R}$ obviously solves (10). A *periodic solution with period* $T > 0$ is a solution $\mathbf{X}(t)$ satisfying $\mathbf{X}(t + T) = \mathbf{X}(t)$ for all $t \in \mathbb{R}$ (with T the smallest value with this property). It is generally a hard task to decide whether or not a system of the form (10) possesses any periodic solution.

While it was obvious for systems of difference equations in the form (2) that any initial value problem has a unique solution, this represents for differential equations a theorem: the *Picard–Lindelöf Theorem*[70] asserts that, given $\mathbf{X}(0)$, the system (10) has a unique solution $\mathbf{X}(t)$ for t in some open interval I containing 0 – under mild conditions on $\mathbf{f}$ which are in applications almost always satisfied. In general, it is a non-trivial task to decide whether I is infinite or whether the solution $\mathbf{X}(t)$ becomes infinite already for some finite time $t_s > 0$. Only for linear systems, it is guaranteed that one may choose $I = \mathbb{R}$.

A *linear, first-order system of differential equations with constant coefficients* has the form

$$\dot{\mathbf{X}}(t) = A\mathbf{X}(t) \quad (11)$$

where $A \in \mathbb{R}^{n \times n}$ is a real matrix with constant entries. Any vector $\mathbf{V}$ with $A\mathbf{V} = \mathbf{0}$ (and thus in particular the zero vector) defines an equilibrium. As in the discrete case, any such system can be explicitly solved with elementary linear algebra and we restrict our discussion to diagonalizable matrices. Thus, we assume again that $\mathbf{V}_1, \dots, \mathbf{V}_n$ are n linearly independent eigenvectors with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. A fundamental system of solutions is then given by

$$\mathbf{X}_k(t) = e^{\lambda_k t} \mathbf{V}_k, \quad 1 \leq k \leq n. \quad (12)$$

For solving the initial value problem with $\mathbf{X}(0) = \mathbf{X}_0$, we must again express the initial vector $\mathbf{X}_0 = \sum_{k=1}^n c_k \mathbf{V}_k$ as a linear combination of the eigenvectors with coefficients c_k and obtain the unique solution $\mathbf{X}(t) = \sum_{k=1}^n c_k \mathbf{X}_k(t)$.

If some eigenvalues are complex, then the same is true for the corresponding eigenvectors. But as in the discrete case, we can rewrite complex members of the

fundamental system (12) as real ones. Again this is based on the fact that with $\lambda = \rho + i\omega$ also its complex conjugate $\bar{\lambda} = \rho - i\omega$ is an eigenvalue of A and if $\mathbf{V} \in \mathbb{C}^n$ is an eigenvector for λ , then $\bar{\mathbf{V}}$ is an eigenvector for $\bar{\lambda}$. Decomposing the eigenvector in its real and imaginary part, $\mathbf{V} = \mathbf{W}_{\text{re}} + i\mathbf{W}_{\text{im}}$, we can replace in our fundamental system the two complex solutions corresponding to λ and $\bar{\lambda}$ by two real ones:

$$\begin{aligned}\mathbf{X}_{\text{re}}(t) &= e^{\rho t} [\cos(\omega t) \mathbf{W}_{\text{re}} - \sin(\omega t) \mathbf{W}_{\text{im}}], \\ \mathbf{X}_{\text{im}}(t) &= e^{\rho t} [\sin(\omega t) \mathbf{W}_{\text{re}} + \cos(\omega t) \mathbf{W}_{\text{im}}].\end{aligned}\tag{13}$$

One key difference to the discrete case is that here ρ and ω are the real and the imaginary part of λ , respectively, whereas in the discrete case ρ was the modulus and ω the argument of λ .

Based on the explicit expressions (12) and (13), we can again make statements about the asymptotic behavior of solutions. If the real part $\Re(\lambda_k)$ of some eigenvalue λ_k (for $\lambda_k \in \mathbb{R}$ obviously $\Re(\lambda_k) = \lambda_k$) is positive, then the solution will be unbounded for $t \rightarrow \infty$ for any initial condition with $c_k \neq 0$. Thus, the population of at least one species/size class will become arbitrarily large and the system is *unstable*. On the other hand, if all eigenvalues have a negative real part, then any solution will converge for $t \rightarrow \infty$ to the origin, so that all species/size classes get extinct. This represents an *asymptotically stable* system. If positive and negative real parts exist among the eigenvalues, then the origin is a *saddle point*. As in the discrete case, complex eigenvalues indicate an *oscillatory behavior* which is damped for $\rho < 0$ and unbounded for $\rho > 0$.

The situation is more complicated, if eigenvalues λ_k exist with a vanishing real part. If λ_k is real, this means $\lambda_k = 0$ and the matrix A is singular. In this case, we have not only one equilibrium at the origin, but at least a whole line of equilibria. If $\lambda_k \neq 0$, it is purely imaginary, $\lambda_k = i\omega_k$, and periodic solutions with frequency ω_k exist for appropriate initial conditions. They correspond to sustained oscillations.

Using the theory of the Jordan normal form, these results can be extended also to non-diagonalizable matrices. While the fundamental system above was composed of pure exponential functions, now also products of polynomials with exponential functions appear as components. Assume now for simplicity that one eigenvalue, say λ_1 , has the real part with the largest absolute value, i.e. $|\Re(\lambda_1)| > |\Re(\lambda_k)|$ for all $k > 1$. Now we find that for large times the solution $\mathbf{X}(t) = \sum_{k=1}^n c_k e^{\lambda_k t} \mathbf{V}_k$ for the initial data $\mathbf{X}_0 = \sum_{k=1}^n c_k \mathbf{V}_k$ satisfies

$$\lim_{t \rightarrow \infty} e^{-\lambda_1 t} \mathbf{X}(t) = c_1 \mathbf{V}_1.\tag{14}$$

Hence, the fundamental mode corresponding to the eigenvalue λ_1 dominates the long term behavior of solutions.

These results can to some extent be applied to study equilibria of nonlinear systems of the form (10), too. Assume that $\mathbf{V}$ is an equilibrium of (10). As in the discrete case, we use the Jacobian J of the right hand side $\mathbf{f}$ at $\mathbf{V}$ defined by (8). We call $\mathbf{V}$ a *hyperbolic* equilibrium, if no eigenvalue of J has a vanishing real part. For such equilibria, the local solution behavior of (10) near $\mathbf{V}$ is the same as for the linear

system $\dot{\mathbf{Y}} = J\mathbf{Y}$. Hence, the equilibrium is asymptotically stable if all eigenvalues have a negative real part.

Figure 4b shows on the right hand side in green the region in the complex plane where all eigenvalues must be located for an asymptotically stable equilibrium, namely in the left half plane. The *Routh–Hurwitz criterion*[70] allows for detecting this without explicitly computing of the eigenvalues and requires only calculation of certain determinants. For sustained oscillations, one must also have at least some eigenvalues on the imaginary axis (shown in blue).

5.4 Extensions and generalizations

We mentioned above that higher-order difference equations may be considered as a model for delay effects. This is not the case for continuous models: a higher-order differential equation has nothing to do with a delay; one has to generalize to *delay differential equations*. In the simplest case an autonomous first-order system of delay differential equations takes the form

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t), \mathbf{X}(t - \tau)) \quad (15)$$

where $\tau > 0$ denotes the delay. One may also consider several delays $\tau_1, \dots, \tau_r$ by allowing that $\mathbf{f}$ depends $\mathbf{X}(t - \tau_1), \dots, \mathbf{X}(t - \tau_r)$. In addition, delay models frequently incorporate integrals (one then speaks of a *distributed delay*).

The theory of delay differential equations is considerably more involved than that of ordinary differential equations. This starts already with the formulation of initial value problems. Whereas for an ordinary differential equation like (10) it suffices to prescribe a single value $\mathbf{X}(0)$ (implying that the solution space is finite-dimensional), an initial condition for (15) has typically the form $\mathbf{X}(t) = \mathbf{X}_0(t)$ for $-\tau \leq t \leq 0$, i. e. $\mathbf{X}_0$ is no longer a constant vector but a function defined on the interval $[-\tau, 0]$ (implying that the solution space is now infinite-dimensional like for most partial differential equations). Using the method of steps, i. e. looking at intervals of the form $[k\tau, (k+1)\tau]$ for $k = 0, 1, 2, \dots$, one can use the Picard–Lindelöf Theorem to prove the (local) existence of a unique solution for the described initial value problem.

The simplest form of a linear delay system with constant coefficients is

$$\dot{\mathbf{X}}(t) = A\mathbf{X}(t) + B\mathbf{X}(t - \tau) \quad (16)$$

with two matrices $A, B \in \mathbb{R}^{n \times n}$. If one searches for exponential solutions as in the case of linear systems of ordinary differential equations, then the equation for the “eigenvalues” λ takes the form $h(\lambda) = \det(\lambda E_n - A - e^{-\tau\lambda}B) = 0$ and thus is no longer simply a polynomial in λ . One can show that h is an entire function entailing many special properties of the equation $h(\lambda) = 0$, i. e. that there are at most countably many solutions and that these do not possess a finite accumulation point. Most remarkably, for any real number $r \in \mathbb{R}$ there exist only finitely many solutions with $\Re(\lambda) > r$ and again the dominating solution modes are those corresponding to the solutions λ with maximal real part.

Delays can lead to oscillations even in very simple linear models, if the delay τ is greater than some threshold. For example, in the simplest scalar delay equation $\dot{X}(t) = -aX(t - \tau)$, the solutions will start to oscillate whenever $a\tau \geq e^{-1}$. The oscillations will be damped, as long $a\tau \geq \pi/2$; for even greater values of $a\tau$, the amplitude of the oscillations will grow with t .

We refrain from delving deeper into the theory of delay differential equations and refer instead to textbooks like [66] for an introduction or to [22] [29] for in-depth treatments of different aspects.

So far, we have always assumed that $\mathbf{X}$ depends only on the time t . Of course, in reality populations will also show a spatial distribution or in the case of marine species like diatom a dependence on the water depth. This leads to models with *partial differential equations* for functions $\mathbf{X}(t, \mathbf{r})$ where the spatial variable $\mathbf{r}$ may again be vector-valued. The dimension of $\mathbf{r}$ depends on the modeling. If one considers simply a water column to include the effect of the water depth, then $\mathbf{r}$ would be a scalar quantity, namely the depth. If one studies the distribution of a population over some region on the surface of the earth, $\mathbf{r}$ would be two-dimensional (cartesian coordinates or longitude and latitude). If the region is part of an ocean and water depth is taken into account, $\mathbf{r}$ would be even three-dimensional. As a general rule of thumb, both the theoretical analysis and the numerical simulation of a system of partial differential equations become rapidly more complicated and expensive with each additional dimension. As it suffices for our purposes here, we will restrict to the case that $\mathbf{r}$ is a scalar quantity x .

In the context of population dynamics (and many other biological applications), a frequent approach to formulating distributed models consists of simply adding a *diffusion term* to model the spatial spreading. This leads then to *reaction-diffusion equations* of the form

$$\frac{\partial \mathbf{X}}{\partial t}(t, x) = \frac{\partial}{\partial x} \left(D(t, x) \frac{\partial \mathbf{X}}{\partial x}(t, x) \right) + \mathbf{f}(\mathbf{X}(t, x)) \quad (17)$$

where D is a diagonal matrix containing the diffusion coefficients for the different species (which may depend on t and x , although in many models they are assumed to be constant) and $\mathbf{f}$ represents the reaction part of the equation. Thus (17) may be considered as a distributed refinement of the simpler model $\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t))$. Such equations appear in many biological, ecological and chemical applications, but also in geology, physics and many other fields of science. Some classical textbooks describing their mathematical theory and some applications are [26] [30] [67].

From a mathematical point of view, (17) represents a parabolic system. As the only nonlinear part is the reaction term $\mathbf{f}$, many results from the theory of linear partial differential equations can be exploited in its analysis. Nevertheless, this nonlinearity leads to many interesting phenomena like traveling waves solutions or pattern formation as discovered by Turing[72]. One usually studies *initial boundary value problems* for a system like (17). Here, the spatial variable x is constrained to some closed interval $[a, b]$ and one prescribes an initial condition $\mathbf{X}(t_0, x) = \mathbf{X}_0(x)$ for all $x \in [a, b]$ with a given function $\mathbf{X}_0(x)$ and two boundary conditions $\mathbf{X}(t, a) = \mathbf{X}_a(t)$, $\mathbf{X}(t, b) = \mathbf{X}_b(t)$ for all $t > t_0$ with two given functions $\mathbf{X}_a(t)$, $\mathbf{X}_b(t)$. Using methods

from functional analysis, one can show that under modest conditions on the reaction term $\mathbf{f}$ and the prescribed functions $\mathbf{X}_0(x)$, $\mathbf{X}_a(t)$ and $\mathbf{X}_b(t)$ a unique solution exists.

5.5 Stochastic branching models

In the approaches discussed so far, individuals are grouped into classes in order to define a state vector on which propagating matrices are applied. Stochastic branching processes constitute a different approach of following the fate of individual cells during proliferation[41]. Interestingly, we encounter in principle the same matrices with the same eigenvalues and eigenvectors if the concept is extended to multiple species or size classes. The general idea is as follows: At a certain time, a cell or generally a particle, is generated. This particle has a lifetime that underlies a certain distribution. After the end of the lifetime, a number of new cells are generated, each starting a new process. There are also variants in which new particles are generated during the lifetime, but these are mainly applicable for higher organisms, so we will not discuss them further in this context. Important, however, is that the newly generated progeny may occur in different type classes. This type space can be attributed to different size classes, so the treatment leads to matrix equations just as in the case of differential or difference equation systems. An important point is that also non-Markovian processes can be easily written down due to the introduction of finite lifetimes. In a Markov process, the future development of a system only depends on the present state, in non-Markovian processes also states in the past contribute to the development. Actually, a system in which the previous generation contributes may be classified as a second-order Markov chain, and so on, so a true non-Markov process has to take the complete history into account.

In a multitype branching process, the state is characterized by a vector $X(t)$ with the number of particles (cells) of the different types (for diatoms: size classes). To make it simple, we first discuss a one-dimensional system (i.e. one type) with one initial particle at $t = 0$ with lifetime realization $\tau(\omega)$ and progeny number $Y(\omega)$. Then, after time t , the number of particles is given by

$$X(t) = \begin{cases} 1 & t < \tau \\ \sum_{i=1}^Y X^{(i)}(t - \tau), & t \geq \tau, \end{cases} \quad (18)$$

which is a recurrence relation because each individual starts a new process.

Now, different models can be considered for the lifetime distribution. The simplest case is that each particle has the same fixed lifetime of one generation (Galton–Watson process). In this case the process is Markovian and a stochastic version of a discrete system is reproduced, with all the properties given above. In the second special case of a lifetime distribution that falls as a single exponential, also a Markov system results, which can be proved analytically[41]. The general case is the Bellman–Harris process in which τ is an arbitrary non-negative random variable from another distribution. Here, a true non-Markovian, age-dependent process results.

The key for treating such problems is to rewrite it with probability generating functions. Such probability generating functions apply to non-negative random vari-

ables such as the number of particles or the number of progeny and are polynomials on a symbolic new variable s with the possible values as exponents and the probabilities p_i as coefficients. For instance, the probability generation function for the number of offspring is defined as

$$f(s) = \sum_i p_i s^i \quad (19)$$

in which i are the possible values for Y . In the case of exact two daughters, $f(s)$ is simply s^2 since the probability for the value $Y = 2$ is exactly 1. For the total number of particles we write $F(s, t)$ with the same definition and take i as possible values for X . The mean (expectation value $\mathbb{E}$) of the particle number after time t is given then conveniently by the first derivative by s at $s = 1$

$$\mathbb{E}[Z(t)] = F'(1, t) = \left. \frac{\partial F(s, t)}{\partial s} \right|_{s=1}. \quad (20)$$

Higher moments of $X(t)$ are represented by higher derivatives, so the variance is given by $\mathbb{V}(X) = F''(1) + F'(1) - (F'(1))^2$.

With these two probability functions we can rewrite the above expression and bring it into a single line by introducing a cumulative distribution function for the lifetimes $G(\tau)$, resulting in

$$F(s, t) = \begin{cases} s, & t < \tau \\ f[F(s, t - \tau)], & t \geq \tau. \end{cases} \quad (21)$$

$$= s[1 - G(t)] + \int_0^t f[F(s, t - u)] dG(u). \quad (22)$$

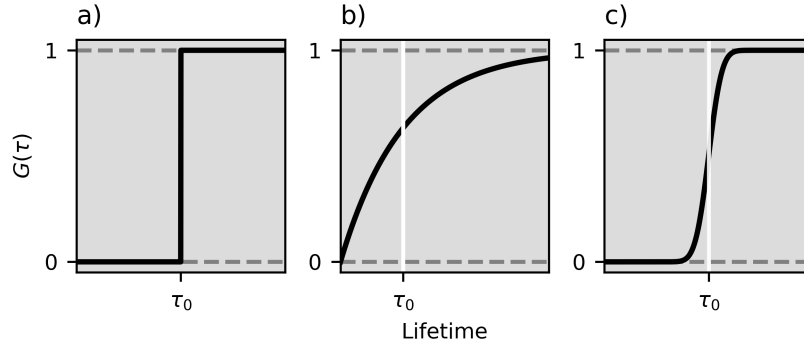


Figure 5: Three cumulative distributions functions for generation times (“lifetimes” of a Bellman–Harris particle): a) fixed lifetime, b) exponentially decreasing lifetime distribution c) lifetime with Gaussian distribution.

Examples for cumulative distribution functions are a step function at the fixed lifetime τ_0 for the Galton–Watson process, the function $G(\tau) = 1 - \exp(-\tau/\tau_0)$ for an

exponentially decreasing lifetime distribution, or the integral of the Gaussian function for a normal distribution with τ_0 as mean value .5 The computational task in solving Bellman–Harris processes is to treat the resulting integral equation to obtain self-consistent solutions, representing steady-states, or follow the temporal evolution.

In multitype branching processes, the offspring may occur in different types (classes), so $X(t)$ and $F(s, t)$ become vectors and a progeny matrix M is needed that sorts the offspring according to some rules to these classes. This matrix is also non-negative and similar to the propagation matrices of difference equations; thus the evolution is also subjected to the Perron–Frobenius theorem[6].

6 Simulation models

6.1 The Schwarz *et al.* model

Schwarz *et al.* developed a stochastic model based on experimental data to find optimal values for Markov-type propagation matrices[65]. They reduced the number of size classes from around 40 as expected for the modeled species *Pseudo-nitzschia delicatissima* to seven. The size distribution was described as a superposition of seven normal distributions with respective means and widths (Gaussian mixture model). The number of seven size classes has been found as optimal number according to the Bayesian information criterion preventing overfitting with insignificant parameters added by further Gaussian functions.

$$\frac{dX(d)}{dd} = \sum_{i=1}^7 A_i \exp\left(\frac{(d - \mu_i)^2}{2\sigma_i^2}\right) \quad (23)$$

	1	2	3	4	5	6	7
μ_i in μm	80.8	75.4	64.5	43.8	51.6	28.7	17.9
σ_i in μm	0.5	0.5	7.1	4.0	4.2	3.6	2.0
limits (μm)	>80.0	>74.5	>57.0	>48.0	>37.5	>21.0	>14.0
range in μm	5.5	5.5	17.5	9.0	10.5	16.5	7.0
τ (days)	5.1	3.0	15.6	15.0	31.9	168.7	∞

Table 2: Size clusters for *P. delicatissima* according to Schwarz *et al.*[65]. Given is mean μ_i , standard deviation σ_i (from graph), lower boundary, cluster range and holding time τ for the Gaussian functions and their derived clusters, respectively.

A univariate clustering analysis of an experimental distribution (with cell size as the single variable) led to the definition of the seven classes with respective boundaries, as given in table 2. It should be noted that in this procedure for the determination of the boundaries, the contained size ranges may differ substantially from cluster to cluster.

Having defined the distinct size classes, the propagation matrix for continuous time modeling was obtained by fitting experimental data for a laboratory strain of *Pseudo-nitzschia delicatissima*[2] with size distributions followed for 265 days. Choosing a

monoclonal strain of a single mating type, sexual reproduction was prevented, so that the vegetative phase could be modeled in isolation. Allowing only a transition into the next size class with smaller size and abandon the reentry by the sexual phase, the continuous, time-independent propagation matrix was constructed as

$$Q = \begin{pmatrix} -q_{12} & q_{12} & & 0 \\ & \ddots & \ddots & \\ & & -q_{67} & q_{67} \\ 0 & & 0 & 0 \end{pmatrix}. \quad (24)$$

The diagonal elements correspond to the negative inverse holding times that were as an initial guess extracted from the scaling law of the size distribution versus time plot (values in table 2). These holding times vary significantly for the different size classes and do not scale with the size range of an individual class. For instance, they are conspicuously long for size class 6 without an identifiable reason.

The fitted entries were tested with parallel data of two other strains by predicting the time evolution of the size distribution. In order to estimate the confidence intervals for being in the experimentally determined states after certain amounts of time, the Fisher information approach was used. The Fisher information $I(\theta)$ expresses the significance of parameter values θ that are contained in a probability distribution function $f(X; \theta)$ and is defined as the expectation value of the second derivative $-\log f$ versus the parameter.

$$I(\theta) = - \int \left(\frac{\partial^2 \log f(X; \theta)}{\partial \theta^2} \right) f(X; \theta) dX. \quad (25)$$

In this case, the parameter is time with n different distributions at certain time intervals. Schwarz *et al.* calculated 95% confidence intervals for the time estimates as

$$\bar{t} \pm 1.96 \sqrt{1/(nI(t))} \quad (26)$$

and showed that the predicted times for the observed size distributions match the experimental time within these limits, thus verifying the model.

In extension of the model, an 8th class was introduced representing the sexual phase, being accessible only from size class 6 and 7. The average holding time was set to 2 days after which class 1 is repopulated (not explicitly stated in the paper). In this case, a stationary distribution with population maximum in the 5th class was obtained.

6.2 The D'Alelio *et al.* model

D'Alelio *et al.* developed a parametrized model that aimed at explaining experimental data for *Pseudo-nitzschia multistriata* taken over a time span of 10 years in the Gulf of Naples[15]. It is formulated as a continuous model with a sampling time of one day. From the original paper it is not completely conclusive how the size classes are defined, but the largest, initial cells were set to 80 μm , the smallest size to 30 μm ,

and the upper sexual threshold to 55 μm . The resolution of the output was 2.5 μm which would correspond to 20 or 21 size classes. Even if cells were collected to size classes and hence the model is denoted as matrix model, there seems to be a certain tracing of individual cells. This can be deduced from a size-depending rate of cell size decrease, given as

$$R(d) = 10^{-7} \mu\text{m} / \text{division}(d/\mu\text{m})^{2.5} \quad (27)$$

(units added here, d represents cell size), and the statement that the total number of cells was kept below 10^4 by randomly removing cells in each time step. The size-dependent expression for $R(d)$ exhibits a larger reduction for larger cells and was justified by comparing the output of the model with the observed data.

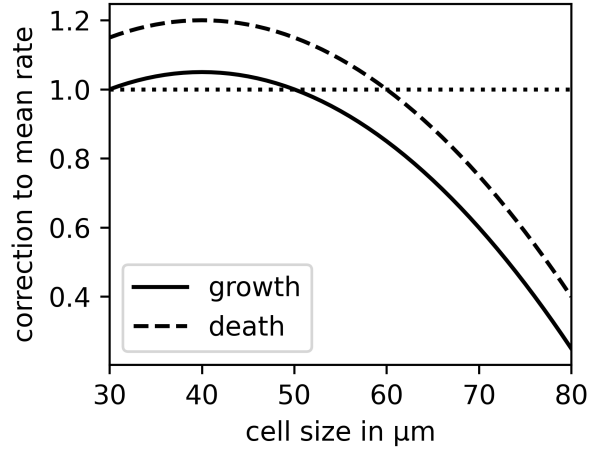


Figure 6: Size-dependent growth parameter described by a parabolic function.

Similarly, the growth and death rates were taken as cell size dependent (Figure 6):

$$g(d) = g_0[0.25 + 0.04(d/\mu\text{m}) - 5 \cdot 10^{-4}(d/\mu\text{m})^2] \quad (28)$$

$$k(d) = k_0[0.4 + 0.04(d/\mu\text{m}) - 5 \cdot 10^{-4}(d/\mu\text{m})^2] \quad (29)$$

Also for these equations, the choices of prefactors in the second order polynomial were derived from comparison with experimental data. The basic rates were modulated according to the seasons: g_0 was set tentatively from 0.69 to 1.40 per day during the bloom season and from 0.07 to 0.20 per day during the rest of the year. Coupled to g_0 , k_0 varied accordingly from 0.05 to 1.00 per day.

The probability for auxosporulation was 2%, with instantaneous reentering into the process (an experimentally determined delay of 2-3 days was neglected deliberately). The authors compared different timing scenarios for sexual reproduction with a gating function $\theta(t)$ between 0 and 1. While the assumption of a continuous sexual reproduction irrespective of the seasons and an annual sex time window during the bloom failed to reproduce the experimental observations, the assumption of biennial sexual periods displayed the best matching. If sex would occur less than every 4th year, the model predicted local extinction of a population.

The parameters used in the model are compiled in table 3 for reference.

Parameter	Code	here
Size	s	d
Size of initial cells	Init size	80 μm
Size of smallest cells	Small size	30 μm
Gametangia threshold size	Sex size	55 μm
Parameterized growth rate	Pgr	$g(d)$
Parameterized death rate	Pdr	$k(d)$
Rate of cell size decrease	Red	$R(d)$
Probability of forming gametes	Sex prob	2%
Timing of sexual reproduction	Sex time	$\theta(t)$

Table 3: Parameters used in the model of D’Alelio *et al.* Given is the original denotation in the paper [15] as well as an abbreviated version used here in the mathematical expressions.

6.3 The Hense–Beckmann model

Hense and Beckmann[31] developed a rather complex model of the population dynamics of diatoms and embedded it into the context of a complete ecosystem, which leads to a system of partial differential equations combined with a discrete system. The basis is a compartmental model PPND of a water column consisting of four compartments: a diatom species B , a bulk phytoplankton P , a limiting nutrient N and detritus D . This model takes the form of a reaction-diffusion system as in (17) where the vector $\mathbf{X}$ consists of the four component (P, B, N, D) , given as biomass in nitrogen equivalents. The equations for B and D are augmented each by a term describing a sinking along axis z , whereas P and N are assumed to be buoyant. More precisely, one obtains the following model

$$\begin{aligned}
\frac{\partial B}{\partial t} &= \mu_B B - (\gamma_B + \delta_B(B + P))B + A_v \frac{\partial^2 B}{\partial z^2} - w_B \frac{\partial B}{\partial z} , \\
\frac{\partial P}{\partial t} &= \mu_P P - (\gamma_P + \delta_P(B + P))P + A_v \frac{\partial^2 P}{\partial z^2} , \\
\frac{\partial N}{\partial t} &= -\mu_P P - \mu_B B + \tau D + A_v \frac{\partial^2 N}{\partial z^2} , \\
\frac{\partial D}{\partial t} &= (\gamma_B + \delta_B(B + P))B + (\gamma_P + \delta_P(B + P))P - \tau D + A_v \frac{\partial^2 D}{\partial z^2} - w_D \frac{\partial D}{\partial z} .
\end{aligned} \tag{30}$$

The parameters and typical values are listed in 4. Here, the growth rates $\mu_{B/P}$ of the species B and P , respectively, depend on the temperature T , the nutrient N and the light I in the water column:

$$\mu_{B/P} = \mu_{0,B/P} \mathcal{F}_{B/P}(T) \frac{N}{k_{N,B/P} + N} \frac{I}{k_{I,B/P} + I} . \tag{31}$$

Here, the temperature dependency for the species P follows an empirical law of a constant times 1.066^T [25], whereas for diatoms B an additional Gaussian term is used defining an optimal temperature[51]. The temperature itself plays the role of

an external forcing with seasonal variation. For the light, self-shading is taken into account by modeling its depth dependency as

$$I(z) = I_0 \exp \left(k_w z + k_c \int_0^z (P(z') + B(z') + D(z')) dz' \right). \quad (32)$$

Parameter	Symbol	Value
Diatom maximum growth rate	$\mu_{0,B}$	0.98 day ⁻¹
Diatom optimal temperature	T_{opt_B}	8 °C
Diatom nutrient half saturation	k_{N_B}	0.16 mmol N m ⁻³
Diatom light half saturation	k_{I_B}	33.6 Wm ⁻²
Attenuation coefficient for water	k_w	0.08 m ⁻¹
Attenuation coefficient for biomass	k_c	0.07 m ² mmol N ⁻¹
Diatom cell lysis rate	γ_B	0.01 day ⁻¹
Diatom grazing rate	δ_B	0.10 m ³ day ⁻¹ mmol N ⁻¹
Subsurface vertical mixing coefficient	A_v	10 ⁻⁵ m ² s ⁻¹
Diatom sinking rate	w_B	0.12 m day ⁻¹
Nutrient remineralization rate	τ	0.1 day ⁻¹
Detritus sinking rate	w_D	1.5 m day ⁻¹

Table 4: Selected parameters used in the Hense–Beckmann PPND model. Parameters for the second, bulk phytoplankton species (index P instead of B) are omitted here.

In this relatively classical form, the size reduction and restitution cycle of the diatoms is not yet taken into account. For this purpose, Hense and Beckmann developed a rather detailed individual based model of the diatom life cycle called DiaLCM. It aggregates a fixed number c of diatoms with identical status into one entity called agent, so that a smaller number of individual cells has to be followed in total. The biomass of the individually modeled diatoms b_i constitutes the total biomass, i.e.

$$B = c \sum_i b_i. \quad (33)$$

For each cell, five processes are taken into account: (i) Cell growth: Depending on temperature and limited by nutrient and light availability as described above but with size dependent parameters, individual cylindrical cells uptake biomass and change their shape by extending into the pervalvar direction due to a fixed diameter. (ii) Cell division: When a size dependent maximum biomass is reached, the biomass is split evenly between epivalvar and hypovalvar daughter, but the hypovalvar daughter gets a smaller diameter, calculated with a constant factor (iii) Cell mortality: Cells die due to size independent cell lysis and size dependent grazing (iv) Auxosporulation: If cells are below the upper sexual threshold (after 80% of the maximum number of generations) and meet additional conditions on available energy and cell concentration, they form new cells of maximum size after an developmental period. (v) Vertical movements in the water column: This is modeled by size dependent sinking and turbulent mixing in

the surface region. Since cells change their shape, an equivalent hydrodynamic radius is calculated for this and other processes.

The precise modeling of these processes is too complex to be presented here in detail. For example, in size dependent aspects the authors have to recalculate between biomass, volume, surface area and equivalent radius. Empirical functional relationships between the various relevant quantities are assumed and taken from the literature. On the whole, the DiaLCM model depends on almost 20 parameters for which values must be chosen. Some variable ranges and parameters are compiled in 5. The authors report from a comparison of the results of simulations of the DiaLCM model with the results of laboratory experiments conducted by Armbrust and Chisholm[3] where they obtained a good agreement. In particular, they could observe periodic changes.

Quantity	Symbol	Value
Number of cells per agent	c	10^4
Number of agents	M	$2 \cdot 10^5$ – $3.1 \cdot 10^6$
Cell biomass	b_i	0.18–1.82 pmol N cell ⁻¹
Equivalent radius	r_{ei}	2.61–6.89 μm
Pervalvar length	h_i	3.85–17.69 μm
Number of generations	G	100
Biomass ratio epivalve:hypovalve	$\delta_e : \delta_h$	50:50
Maximum age of epivalve	a_{max}	7 divisions
Generation threshold for sex	g_a^{crit}	$0.8 G$
Daily mean energy threshold for sex	E_{crit}	12 Wm^{-2}
Minimum cell concentration for sex	c_{crit}	$10^7 \text{ cells m}^{-3}$
Developmental period	τ_{dev}	1 day

Table 5: Selected variables and parameters used in the Hense–Beckmann DiaLCM model as far as mentioned in the text. For a complete list please refer to the appendix of the original article[31].

For the coupling of the two models, the diatom compartment B is removed from the PPND model, i. e. the equation for B is removed from (30) and replaced by the DiaLCM model. This approach allows determination of all B dependent terms in the remaining equations in (30). Hense and Beckmann report that in the coupled model the strict periodicity of the DiaLCM model under “laboratory conditions” disappears. They observed that under the influence of environmental factors achieved through the coupling with the PPND model, the size reduction and restitution cycle can be significantly accelerated or decelerated. Due to the complexity of their model with its many parameters and assumptions, it is, however, difficult to see which factors are mainly responsible for these observations.

6.4 The Terzieva–Terziev model

Terzieva and Terziev followed a non-Markovian approach and applied the Bellman–Harris stochastic branching concept to the cell division scheme[69]. This allows them

to model a flexible time between cell divisions. Explicitly, the time between two divisions, constituting the lifetime of a particle in the multitype Bellman–Harris scheme, is drawn from a Gaussian normal distribution with mean τ_0 and variance σ^2 . Originally, for the size range between 30 μm and 150 μm of a generic diatom, five size classes (types) have been defined which are reduced to four by uniting the two lowest classes (i. e. 1 from 30 to 80 μm , 2 from 80 to 105 μm , 3 from 105 to 130 μm , 4 from 130 to 150 μm). For the individual cells of these four types, the following individual generation functions are constructed:

$$f_1(\mathbf{s}) = 1 - \rho + \rho s_1 , \quad (34)$$

$$f_i(\mathbf{s}) = 1 - \rho + \rho s_{i-1} s_i \quad \text{for } i = 2, 3, 4 \quad (35)$$

forming together the four-dimensional vector $\mathbf{f}$ with the four-dimensional symbolic variable $\mathbf{s}$ as argument. ρ denotes the number of offspring after dividing in each of the daughter types. ρ is not subjected to a distribution and represents the probability of survival until the next division. The lack of asymmetry between the daughter types reveals that one type essentially represents one distinct size, otherwise the number of progeny would be less for the smaller type. Also, as a probability, ρ is meant to be not larger than 1. In the multitype Bellman–Harris variant, the number of progeny is expressed as matrix (particle production mean matrix M) with elements given as

$$m_{i,j} = \left. \frac{\partial f_i}{\partial s_j} \right|_{\mathbf{s}=\mathbf{1}} \quad (36)$$

performing the differentiation at the vector of ones $\mathbf{1} = (1, 1, 1, 1)$, i.e.

$$M = \begin{pmatrix} \rho & 0 & 0 & 0 \\ \rho & \rho & 0 & 0 \\ 0 & \rho & \rho & 0 \\ 0 & 0 & \rho & \rho \end{pmatrix}. \quad (37)$$

Note that a sexual size restoring mechanism is not implemented in this matrix.

For the ensemble of particles with individual lifetimes, a vectorial generation function was then constructed, with integrals over the cumulative lifetime distribution functions $G(\tau)$ as in the single type case with components

$$F_i(\mathbf{s}, t) = s_i(1 - G_i(t)) + \int_0^t f_i(\mathbf{F}(\mathbf{s}, t - \tau)) dG_i(\tau). \quad (38)$$

Note that in the integral the generating function vector $\mathbf{F}$ is used as argument for the individual generating functions f to define its own components, hence the recursion.

A matrix of means for the ensemble $U(t)$ was then similarly constructed, in which the last row gives the expectation of the population $\mathbf{X}(t)$ for the starting condition of one initial cell

$$\mathbf{X}(t) = \begin{pmatrix} u_{11}(t) & 0 & 0 & 0 \\ u_{21}(t) & u_{22}(t) & 0 & 0 \\ u_{31}(t) & u_{32}(t) & u_{33}(t) & 0 \\ u_{41}(t) & u_{42}(t) & u_{43}(t) & u_{44}(t) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The respective elements of $U(t)$ are given by

$$u_{ij}(t) = 1 - G_i(t) + \rho \int_0^t (1 - G_i(t-u)) d \sum_{n=1}^{\infty} \rho^{n-1} (G_i(t))^{*n} \quad (39)$$

in which $*n$ denotes the n th convolution[6].

The paper of Terzieva and Terziev, which is not easily accessible at first glance, demonstrate how multitype Bellman–Harris procedures may be applicable to size-resolved population dynamics of diatoms in principle, but numerical simulations are still lacking.

6.5 The Fuhrmann-Lieker *et al.* model

Fuhrmann-Lieker *et al.* understood the diatom population as a discrete dynamical system[27]. The authors varied parameters and expanded their model to investigate whether a linear matrix model could produce the diatom clock’s oscillations in cell and population size. For that, a basic model was introduced which assumed the population to consist of discrete size classes according to the MacDonald–Pfitzer scheme. Consequently, cells would divide simultaneously at a discrete periodical rate. The number of cells in a certain size class is denoted in a population vector $\mathbf{X}(\mathbf{t})$ with elements x_i . The distribution of the next generation $t+1$ is obtained by multiplication of the vector with the propagator matrix (as introduced in (4)):

$$P = \begin{pmatrix} \alpha p & p & 0 & \cdots & 0 & 0 \\ 0 & \beta p & p & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta p & p \\ s & 0 & 0 & \cdots & 0 & \beta p \end{pmatrix}. \quad (40)$$

The propagator matrix contains the parameters of survival and auxosporulation rate as well as an optional asymmetry factor according to table 6. For simplification of P and later analytical treatment, the parameter σ was introduced.

In addition to the basic model, four different modifications were examined. Asymmetry between two daughter cells was considered according to the observations of Müller and Laney[53] [43]. In a second modification the authors applied two different methods to introduce a nonlinearity. First, to account for a biparental sexual reproduction, the auxosporulation probability was modified:

$$x_{n-1,t+1} = \frac{s x_{0,t}^2}{(1 + 2s x_{0,t})}. \quad (41)$$

Second, the natural saturation was included by addition of a Ricker function which limits the maximal size of the population by modifying the survival probability depending on the population size:

$$p = p_0 e^{-\gamma(\sum_i x_i)_t}. \quad (42)$$

Parameter	Meaning	Model
N	number of size classes	all models
i	index for size class ($0 \leq i < N$)	all models
p	survival probability of theca	all models
s	probability of auxosporulation	all models
α	survival factor of smallest cells	all models
β	asymmetry factor between daughters	all models
σ	$(s/p)^{1/n}$	all models
t	generation as time scale	all models
γ	inverse carrying capacity	saturation, aging models
m	number of age classes	aging models
j	index for age class ($0 \leq j < m$)	aging models

Table 6: Parameters used in the Fuhrmann-Lieker *et al.* model.

The third modification involves the aging of daughter cells as it has been observed in various microorganisms to control population development. The authors follow Jewson’s findings that the daughter cells containing the epitheca would age and subsequently die after a certain number of divisions[37]. To include this mechanism, another dimension is added to the population vector, thus turning it into a matrix with a size and an age dimension. Upon division, the cell containing the epitheca moves along the age axis of the matrix into the older class while the hypothecal cell moves into the smaller size class along the size axis. The last variation of the model introduced a sinusoidal oscillation of the survival and auxosporulation parameters to mimic seasonal changes.

The authors evaluated their models by generational development of total cell number, number of auxospores, mean size and the variance of the sizes. Furthermore, eigenvalues and eigenvectors of the systems were analyzed (as described in section 5). The authors conclude that a continuous oscillation as in natural environments could only be obtained by including temporal modulations into the model. With certain combinations of parameters and variations of the model a very long decay time of millions of generations for initially occurring inherent oscillations was found. Under natural conditions it could be assumed that these oscillations would appear stable since no system could stay undisturbed for such a long period. Nevertheless, an equilibrium is always reached in theory as long as no fluctuation is included.

As an example of numerical results from the model, in Figure 7 the situation for two different scenarios is compared. On the left side the population dynamics and the trajectory of size distribution is followed for the original MacDonald–Pfitzer scheme, showing an approach towards an equilibrium population and size distribution after a few life cycles. On the right side, the assumption of Jewson about a limited lifespan of individual cells is added. Here, the oscillations persist for a long time and decay only slowly.

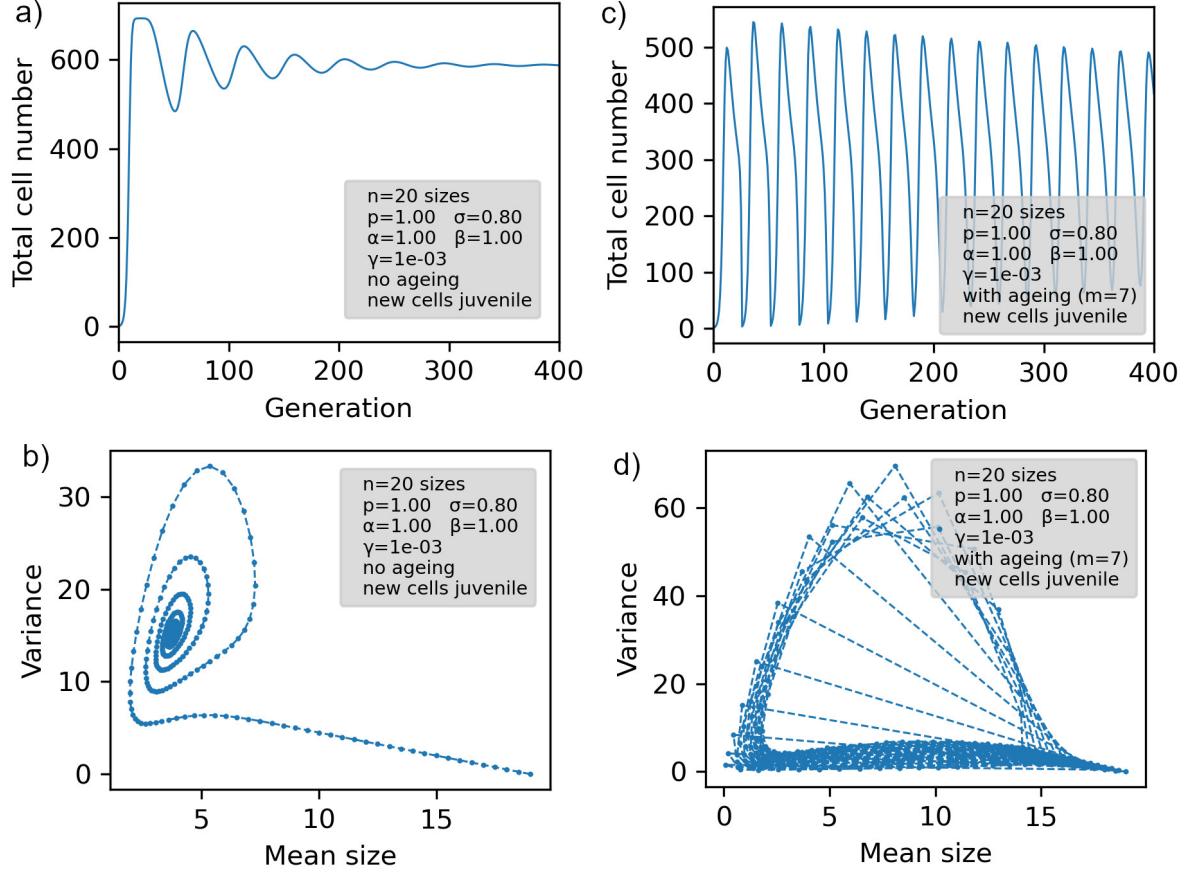


Figure 7: Direct comparison of the Fuhrmann-Lieker *et al.* model with and without aging. a) Development of the population in a small model with 20 size classes in which cells die only due to overpopulation (saturation by Ricker term) b) Trajectory of the size distribution in a), represented by mean size and variance. c) Same model as in a), but thecae live for only 7 generations as suggested by Jewson. Long-term oscillations are induced. d) Trajectory of the size distribution in c).

7 Oscillatory behavior

7.1 Reproduction of experimental data

Models have to be capable of describing present data and giving predictions for the future, also in alternative scenarios with modified parameters. In this section we show for the few existing long-term studies agreements and difficulties in the application of the models. Comparison mainly concentrates on the observation of distinct cohorts, each started by a synchronous auxosporulation event which starts the reduction-restitution cycle.

Figure 8 can be taken as a representative example for long-term studies. For 10 years, the size distribution in the population of *Pseudo-nitzschia multistriata* in the Gulf of Naples was monitored by taking samples in regular intervals[15]. During

bloom phases *P. multistriata* was abundant, so sampling was easy, but the bloom did not occur every year with the same cell concentration. In samples in the rest of the year, the diatoms occurred less frequently, so the statistic significance of the data is also less. A bimodal distribution of cell sizes was observed in many of the samples, but the assignment of a certain cell size to one of the assumed cohorts can be ambiguous if the modes overlap. From the bimodal distribution, a total of 7 cohorts was identified, each of which survive for three years and generate a new cohort in the second year, so a biannual rhythm can be seen. The first phase from the initial cell to midsize cells could only be postulated with no experimental data because of the low frequency of occurrence. The problem of observing auxosporulation and the generation of new initial cells was discussed in depth by Mann[48]. Since the total percentage of cells switching from vegetative growth to gametogenesis is very small, auxosporulation is a very rare event, when compared to all cells that are still in the vegetative phase. In reading the further size decrease from the graph, one has to take into account that model parameters change in the bloom phase (black bar at the bottom) and in the following stationary phase (white bar). This feature is further discussed in the next section. As this example shows, finding total agreement of a model and experimental data is not easy. Similar studies confirm this picture[48] [49]. Sometimes multimodal behavior can be seen, but sometimes the size classes merge, and cohorts are not easily distinguished. Therefore, the only hope is to be able to deduce some of the parameters that would represent an ideal life cycle, neglecting stochastic fluctuations of real populations. In the laboratory, chances are better, but also here unforeseen events and discontinuity in the culture may obstruct the analysis.

7.2 Coupling to external rhythms

In all models the question arises if the parameters used in the differential equations and matrices can be taken as constant or if they may be subjected to external influences. Mann and others suggested that auxosporulation occurs only within a few weeks each year[48]. So there must be a coupling of the auxosporulation probability not only to size but also to photo- or thermoperiod. Wherever doubling rates are given, they are also subjected on fluctuations in nutrient and light conditions. Typically they vary roughly with some typical period that can be defined as *zeitgeber* rhythm, in the terms of chronobiology. *Zeitgeber* can be implemented in the models as gating functions (see discussion of the D’Alelio *et al.* model) or, as all periodic signals can be described as Fourier series of the fundamental frequency and its higher harmonics, as sine functions. Then, the intrinsic frequency given by the size reduction and restitution cycle (the “sex clock”) and the annual cycle of the *zeitgeber* interact.

In Figure 9 the behavior is shown when the auxosporulation parameter σ is varied sinusoidally in the Fuhrmann-Lieker *et al.* model. Since the auxosporulation probability is given by $s = p\sigma^n$ it mimics a confined time window with high probability for sexual reproduction and a longer phase with a strongly reduced probability. In the resulting plot of the population dynamics, both periodicities are involved (a), with the result that auxosporulation does not take place significantly in every year but with a burst in the years when sexually productive cells are abundant during the

mating season (b). Fourier transformation of the cell number reveal the frequency coupling on short and long terms. At the beginning of the oscillations, the Fourier spectrum shows the signals for the size cycle at 0.018/generation and the annual cycle at 0.025/generation, but also mixed frequencies by nonlinear coupling. For example, the dominant Fourier peak at 0.07/generation is just the difference frequency of both underlining rhythms (c). In the long term however, the *zeitgeber* rhythm controls the dynamic behavior (d) and becomes the only surviving mode. This again is a result of the Perron–Frobenius theorem that forbids a dominant oscillating eigenvector in the matrices describing the size class structure to which the external signal is not subjected.

8 Conclusion

We have seen that there are many different approaches possible in modeling the size resolved population dynamics of diatoms. They all have in common that the population is sorted into classes of distinct cell sizes in order to describe the vegetative size reduction process. Then the similarity ends. Under the proposed models, there are systems of differential equations, difference equations, as well as individual-based stochastic processes. The models are either constructed to describe an experimentally obtained data set for a distinct species, or are developed as rather generic models to understand the principles of size-dependent population dynamics. In any case, a good model for population dynamics of diatoms takes cell size and environmental parameters into account. Such models will be important for simulating whole ecosystems that contain diatoms as one of the most influential primary production sources and carbon dioxide sequestration agents.

Acknowledgment

Funding by the Research Training Group “Biological Clocks on Multiple Time Scales” (GRK 2749) of the German Research Foundation (DFG), the Graduate School “Biological Clocks” (P/979) of the University of Kassel and the bilateral project ANR-17-CE400036/DFG-391322026 SYMBIONT is gratefully acknowledged.

We thank Nico Kubetschek and Giagos-Ion Chlomoudis for proofreading and help in formatting formulae and figures.

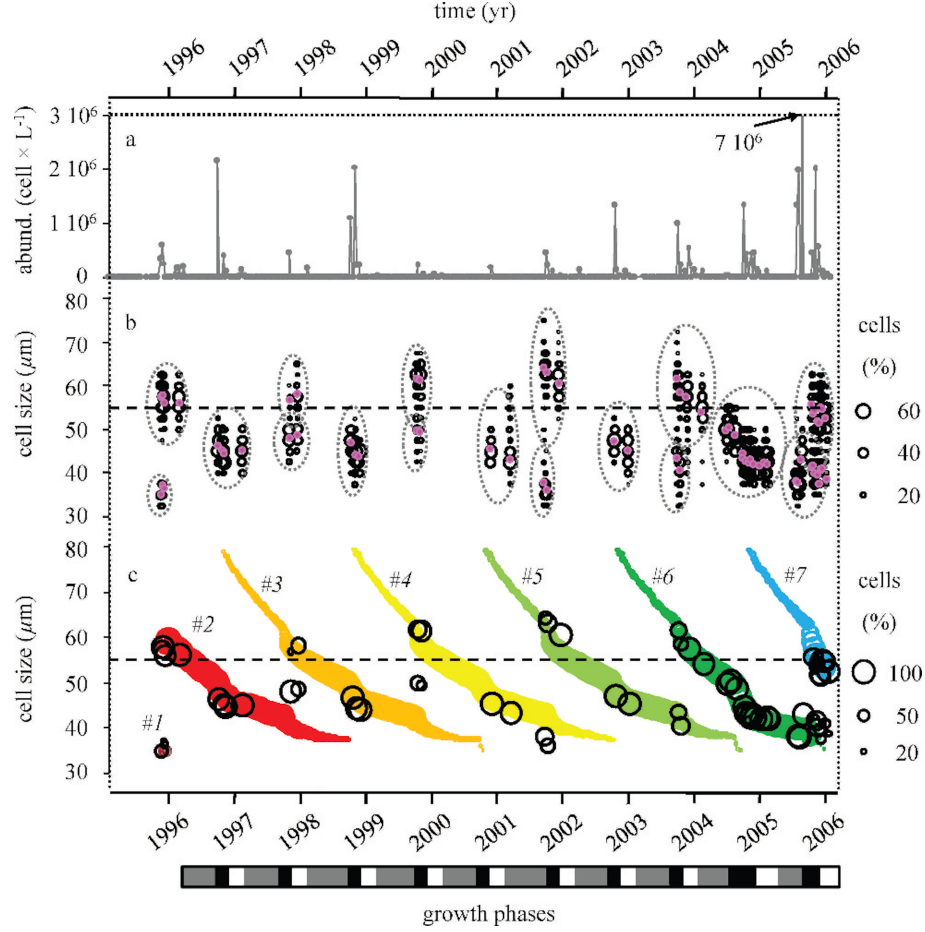


Figure 8: Cell size dynamics of *Pseudo-nitzschia multistriata* according to D'Alelio *et al.*[15]. Upper graph: cell abundance at the sampling period. Center: Division of the population according to cell size into subpopulations (cohorts, dotted ellipses) and size classes (solid circles). The mean cell size of the cohorts is indicated in pink. The dashed line indicates the postulated upper size threshold for sex. Lower graph: Comparison with the model. Different model cohorts are emphasized by different colors. Changes of cell size reduction speed are attributed to different growth phases: normal growth (gray in the bar below), bloom phases (black), and stationary phases (white). Reproduced with permission.

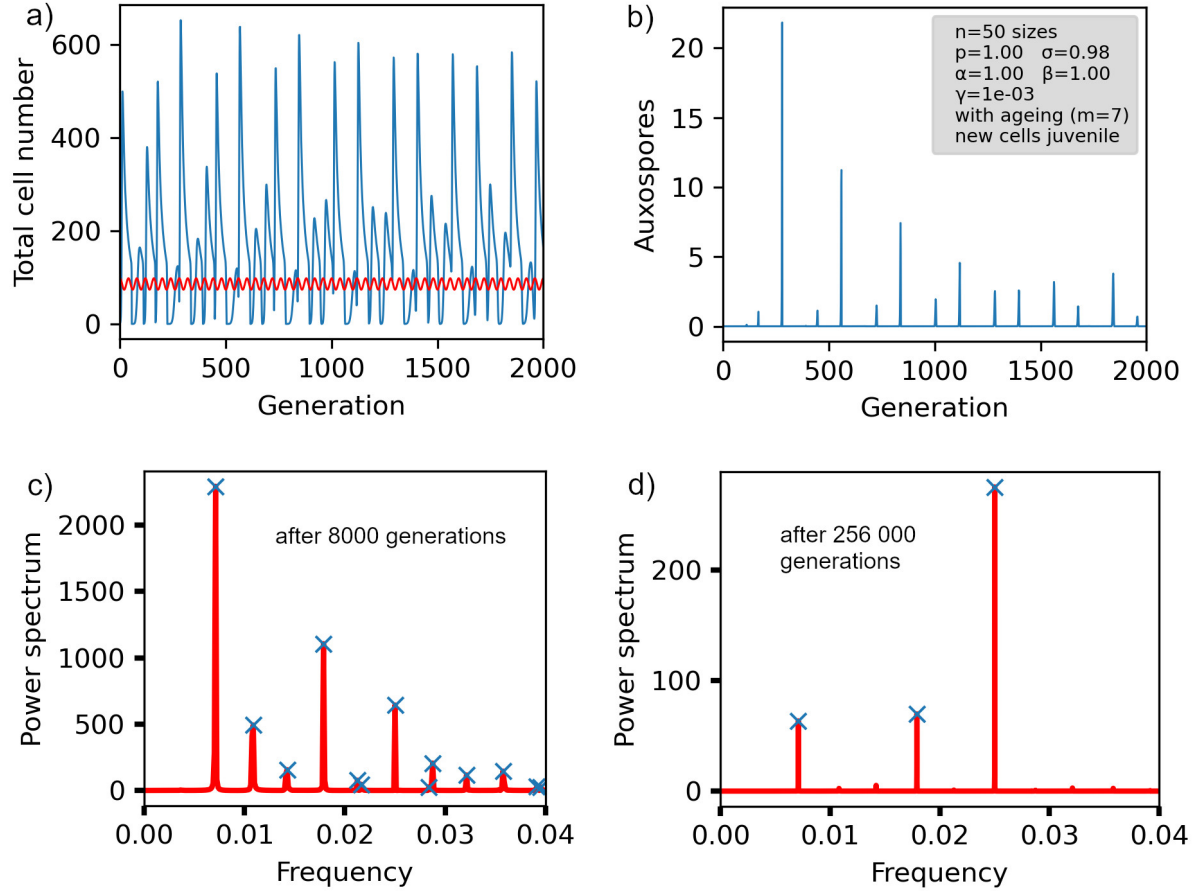


Figure 9: Temporal development of population parameters and dynamics in the Fuhrmann-Lieker *et al.* model with *zeitgeber* rhythm. a) Population size and *zeitgeber* modulation of the parameter σ with a period of 40 generations, multiplied by 100 for better scaling (red), b) Number of auxospores, showing years of low and high auxosporulation c) Fourier transform of the total cell number for the first 8000 generations. Temporal frequency is given as 1/generation d) Fourier transform of the total cell number for the first 256000 generations.

References

- [1] W. Admiraal. Influence of light and temperature on the growth rate of estuarine benthic diatoms in culture. *Marine Biology*, 39(1):1–9, 1977. doi: 10.1007/BF00395586.
- [2] A. Amato, L. Orsini, D. D’Alelio, and M. Montresor. Life cycle, size reduction patterns, and ultrastructure of the pennate planktonic diatom *Pseudo-nitzschia delicatissima* (Bacillariophyceae). *Journal of Phycology*, 41:542–556, 2005. doi: 10.1111/j.1529-8817.2005.00080.x.
- [3] E. Armbrust and S. Chisholm. Patterns of cell size change in a marine centric diatom: Variability evolving from clonal isolates. *Journal of Phycology*, 28:146–156, 1992.
- [4] P. Assmy, J. Henjes, V. Smetacek, and M. Montresor. Auxospore formation by the silica-sinking, oceanic diatom *Fragilariopsis kerguelensis* (Bacillariophyceae). *Journal of Phycology*, 42(5):1002–1006, 2006. doi: 10.1111/j.1529-8817.2006.00260.x.
- [5] P. Assmy, D. U. Hernández-Becerril, and M. Montresor. Morphological variability and life cycle traits of the type species of the diatom genus *Chaetoceros*, *C. dictyota*. *Journal of Phycology*, 44(1):152–163, 2008. doi: 10.1111/j.1529-8817.2007.00430.x.
- [6] K. Athreya and P. Ney. *Branching processes*. Die Grundlagen der mathematischen Wissenschaften in Einzeldarstellungen 196. Springer, New York, 1972.
- [7] A.-S. Benoiston, F. M. Ibarbalz, L. Bittner, L. Guidi, O. Jahn, S. Dutkiewicz, and C. Bowler. The evolution of diatoms and their biogeochemical functions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1728):20160397, 2017. doi: 10.1098/rstb.2016.0397.
- [8] A. Berman and R. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Classics in Applied Mathematics 9. SIAM, Philadelphia, 1994.
- [9] H. Bethge. *Melosira und ihre Planktonbegleiter*. Salzwasser Verlag, Paderborn, 1 edition, 2012.
- [10] G. Bilcke, L. van Craenenbroeck, A. Castagna, C. M. Osuna-Cruz, K. Vandepoele, K. Sabbe, L. de Veylder, and W. Vyverman. Light intensity and spectral composition drive reproductive success in the marine benthic diatom *Semina robusta*. *Scientific Reports*, 11(1):17560, 2021. doi: 10.1038/s41598-021-92838-0.
- [11] F. Brauer and C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. Springer, New York, 2nd edition, 2010.

- [12] V. A. Chepurnov, D. G. Mann, W. Vyverman, K. Sabbe, and D. B. Danielidis. Sexual reproduction, mating system, and protoplast dynamics of *Seminavis* (Bacillariophyceae). *Journal of Phycology*, 38(5):1004–1019, 2002. doi: 10.1046/j.1529-8817.2002.t01-1-01233.x.
- [13] V. A. Chepurnov, D. G. Mann, K. Sabbe, and W. Vyverman. Experimental Studies on Sexual Reproduction in Diatoms. volume 237 of *International Review of Cytology*, pages 91–154. Academic Press, 2004. doi: 10.1016/S0074-7696(04)37003-8.
- [14] V. A. Chepurnov, D. G. Mann, P. von Dassow, P. Vanormelingen, J. Gillard, D. Inzé, K. Sabbe, and W. Vyverman. In search of new tractable diatoms for experimental biology. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 30(7):692–702, 2008. doi: 10.1002/bies.20773.
- [15] D. D’Alelio, M. R. d’Alcalà, L. Dubroca, D. Sarno, A. Zingone, and M. Montresor. The time for sex: A biennial life cycle in a marine planktonic diatom. *Limnology and Oceanography*, 55(1):106–114, 2010. doi: 10.4319/lo.2010.55.1.0106.
- [16] N. A. Davidovich. Factors controlling the size of initial cells in diatoms. *Russian Journal of Plant Physiology*, 41:220–224, 1994.
- [17] N. A. Davidovich and S. S. Bates. Sexual reproduction in the pennate diatoms *Pseudo-nitzschia multiseriata* and *P. pseudodelicatissima* (Bacillariophyceae). *Journal of Phycology*, 34(1):126–137, 1998. doi: 10.1046/j.1529-8817.1998.340126.x.
- [18] N. A. Davidovich, J.-L. Mouget, and P. Gaudin. Heterothallism in the pennate diatom *Haslea ostrearia* (Bacillariophyta). *European Journal of Phycology*, 44(2):251–261, 2009. doi: 10.1080/09670260802710301.
- [19] N. A. Davidovich, I. Kaczmarska, and J. M. Ehrman. Heterothallic and homothallic sexual reproduction in *Tabularia fasciculata* (Bacillariophyta). *Fottea*, 10(2):251–266, 2010. doi: 10.5507/fot.2010.016.
- [20] N. A. Davidovich, R. Gastineau, P. Gaudin, and O. I. Davidovich. Sexual reproduction in the newly-described blue diatom, *Haslea karadagensis*. *Fottea*, 12(2):219–229, 2012. doi: 10.5507/fot.2012.016.
- [21] O. Davidovich, N. Davidovich, and J.-L. Mouget. The effect of temperature on vegetative growth and sexual reproduction of two diatoms from the genus *Haslea* Simonsen. *Russian Journal of Marine Biology*, 44(1):8–13, 2018. doi: 10.1134/S1063074018010030.
- [22] O. Diekmann, S. van Gils, S. Verduyn Lunel, and H.-O. Walther. *Delay Equations*. Applied Mathematical Sciences 110. Springer-Verlag, New York, 1995.

- [23] M. Edlund and R. Bixby. Intra- and inter-specific differences in gametangial and initial sizes in diatoms. Proceedings of the 16th international diatom symposium, pages 169–190. Amvrosiou Press, 2001.
- [24] S. Elaydi. *An Introduction to Difference Equations*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 3rd edition, 2005.
- [25] R. Eppley. Temperature and phytoplankton growth in the sea. *Fish. Bull.*, 70: 1063–1085, 1972.
- [26] P. Fife. *Mathematical Aspects of Reacting and Diffusing Systems*. Lecture Notes in Biomathematics 28. Springer-Verlag, Berlin, 1979.
- [27] T. Fuhrmann-Lieker, N. Kubetschek, J. Ziebarth, R. Klassen, and W. Seiler. Is the diatom sex clock a clock? *Journal of the Royal Society, Interface*, 18(179): 20210146, 2021. doi: 10.1098/rsif.2021.0146.
- [28] J. Gillard, V. Devos, M. J. Huysman, L. De Veylder, S. D’Hondt, C. Martens, P. Vanormelingen, K. Vannerum, K. Sabbe, V. A. Chepurnov, D. Inzé, M. Vuytsteke, and W. Vyverman. Physiological and transcriptomic evidence for a close coupling between chloroplast ontogeny and cell cycle progression in the pennate diatom *Seminavis robusta*. *Plant Physiology*, 148:1394–1411, 2008. doi: 10.1104/pp.108.122176.
- [29] K. Gopalsamy. *Stability and Oscillations in Delay Differential Equations of Population Dynamics*. Kluwer, New York, 1992.
- [30] P. Grindrod. *The Theory and Applications of Reaction-Diffusion Equations: Patterns and Waves*. Oxford Applied Mathematics and Computing Science Series. Clarendon Press, Oxford, 2nd edition, 1996.
- [31] I. Hense and A. Beckmann. A theoretical investigation of the diatom cell size reduction–restitution cycle. *Ecological Modelling*, 317:66–82, 2015. doi: 10.1016/j.ecolmodel.2015.09.003.
- [32] M. Hirsch, S. Smale, and R. Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Academic Press, Amsterdam, 3rd edition, 2013.
- [33] K. E. Holtermann, S. S. Bates, V. L. Trainer, A. Odell, and E. Virginia Armbrust. Mass sexual reproduction in the toxigenic diatoms *Pseudo-nitzschia australis* and *P. pungens* (Bacillariophyceae) on the Washington coast, USA. *Journal of Phycology*, 46(1):41–52, 2010. doi: 10.1111/j.1529-8817.2009.00792.x.
- [34] M. J. Huysman, W. Vyverman, and L. De Veylder. Molecular regulation of the diatom cell cycle. *Journal of Experimental Botany*, 65(10):2573–2584, 2013. doi: 10.1093/jxb/ert387.

- [35] M. J. Huysman, C. Martens, K. Vandepoele, J. Gillard, E. Rayko, M. Heijde, C. Bowler, D. Inzé, Y. Van de Peer, L. De Veylder, and W. Vyverman. genome-wide analysis of the diatom cell cycle unveils a novel type of cyclins involved in environmental signalling. *Genome Biology*, 11:R17, 2020.
- [36] M. J. J. Huysman, A. E. Fortunato, M. Matthijs, B. S. Costa, R. Vanderhaeghen, H. van den Daele, M. Sachse, D. Inzé, C. Bowler, P. G. Kroth, C. Wilhelm, A. Falciatore, W. Vyverman, and L. de Veylder. AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricornutum*). *The Plant cell*, 25(1):215–228, 2013. doi: 10.1105/tpc.112.106377.
- [37] D. H. Jewson. Life cycle of a *Stephanodiscus* sp. (Bacillariophyta). *Journal of Phycology*, 28(6):856–866, 1992. doi: 10.1111/j.0022-3646.1992.00856.x.
- [38] D. H. Jewson. Size reduction, reproductive strategy and the life cycle of a centric diatom. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1277):191–213, 1992. doi: 10.1098/rstb.1992.0056.
- [39] I. Kaczmarska, A. Pouličková, S. Sato, M. B. Edlund, M. Idei, T. Watanabe, and D. Mann. Proposals for a terminology for diatom sexual reproduction, auxospores and resting stages. *Diatom Research*, 28(3):263–294, 2013. doi: 10.1080/0269249X.2013.791344.
- [40] W. Kelley and A. Peterson. *Difference Equations: An Introduction with Applications*. Academic Press, Boston, 1991.
- [41] M. Kimmel and D. Axelrod. *Branching Processes in Biology*. Interdisciplinary Applied Mathematics. Springer-Verlag, New York, 2002.
- [42] M. Kloster, A. S. Rigual-Hernández, L. K. Armand, G. Kauer, T. W. Trull, and B. Beszteri. Temporal changes in size distributions of the southern ocean diatom *Fragilariopsis kerguelensis* through high-throughput microscopy of sediment trap samples. *Diatom Research*, 34(3):133–147, 2019. doi: 10.1080/0269249X.2019.1626770.
- [43] S. R. Laney, R. J. Olson, and H. M. Sosik. Diatoms favor their younger daughters. *Limnology and Oceanography*, 57(5):1572–1578, 2012. doi: 10.4319/lo.2012.57.5.1572.
- [44] P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33(3):183–212, 1945. doi: 10.1093/biomet/33.3.183.
- [45] E. Lewis. *Network Models in Biology*. Biomathematics 7. Springer-Verlag, Berlin, 1977.
- [46] W. M. Lewis. The diatom sex clock and its evolutionary significance. *The American Naturalist*, 123(1):73–80, 1984. doi: 10.1086/284187.

- [47] J. D. Macdonald. I.—On the structure of the Diatomaceous frustule, and its genetic cycle. *Annals and Magazine of Natural History*, 3(13):1–8, 1869. doi: 10.1080/00222936908695866.
- [48] D. G. Mann. Why didn’t Lund see sex in *Asterionella*? A discussion of the diatom life cycle in nature. *Algae and the Aquatic Environment*, pages 383–412. 1988.
- [49] D. G. Mann. Size and Sex. *The Diatom World*, pages 145–166. Springer Netherlands, Dordrecht, 2011. doi: 10.1007/978-94-007-1327-7_6.
- [50] M. Mizuno and K. Okada. Seasonal change in the distribution of cell size of *Cocconeis suctellum* var. *ornata* (Bacillariophyceae) in relation to growth and sexual reproduction. *J. Phycol.*, 21:547–553, 1985.
- [51] J. Moisan, T. Moisan, and M. Abbott. Modelling the effect of temperature on the maximum growth rates of phytoplankton populations. *Ecol. Model.*, 153: 197–215, 2002.
- [52] C. M. Moore, M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, S. L. Jaccard, T. D. Jickells, J. La Roche, T. M. Lenton, N. M. Mahowald, E. Marañón, I. Marinov, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, A. Tsuda, and O. Ulloa. Processes and patterns of oceanic nutrient limitation. *Nature Geoscience*, 6(9):701–710, 2013. doi: 10.1038/ngeo1765.
- [53] G. F. O. Müller. Das Gesetz der Zelltheilungsfolge von *Melosira* (*Orthosira*) *arenaria* Moore. *Berichte der Deutschen Botanischen Gesellschaft*, v. 1:35–44, 1883.
- [54] G. F. O. Müller. Die Zellhaut und das Gesetz der Zelltheilungsfolge von *Melosira* (*Orthosira* Thwaites) *arenaria* Moore. *Jahrbücher für wissenschaftliche Botanik*, 14:232–290, 1884.
- [55] S. Nagai and T. Manabe. Auxospore formation of a giant diatom, *Coscinodiscus wailesii* (Bacillariophyceae), in culture. *Bulletin of Plankton Society of Japan (Japan)*, 1994.
- [56] F. Nipkow. Über das Verhalten der Skelette planktischer Kieselalgen im geschichteten Tiefenschlamm des Zürich- und Baldeggersees. *Zeitschrift für Hydrologie*, 4(1):71–120, 1928. doi: 10.1007/BF02485761.
- [57] T. Nishikawa. Cell size reduction and restoration of the diatom *Eucampia zodiacus* in culture. *Plankton and Benthos Research*, 12(2):139–142, 2017. doi: 10.3800/pbr.12.139.
- [58] E. T. Olivieri. Feasibility of estimating phytoplankton size and biomass in fresh and preserved samples from the Benguela current with a Coulter counter.

- South African Journal of Marine Science*, 3(1):99–110, 1985. doi: 10.2989/025776185784461135.
- [59] P. H. Petersen, O. Blaaberg, M. Andersen, Jørgensen, K. Schousboe, and E. Jensen. Graphical interpretation of confidence curves in rankit plots. *Clin. Chem. Lab. Med.*, 42(7):715–724, 2004. doi: 10.1515/CCLM.2004.122.
 - [60] E. Pfitzer. Über Bau und Zellteilung der Diatomeen. *Botanische Zeitung*, 27: 774–776, 1869.
 - [61] L. Roselli, E. Stanca, F. Paparella, A. Mastrolia, and A. Basset. Determination of *Coscinodiscus* cf. *granii* biovolume by confocal microscopy: comparison of calculation models. *Journal of Plankton Research*, 35(1):135–145, 2013. doi: 10.1093/plankt/fbs069.
 - [62] F. Round. Auxospore structure, initial valves and the development of populations of *Stephanodiscus* in Farmoor Reservoir. *Ann. Bot.*, 49:447–459, 1982.
 - [63] J. Sandefur. *Discrete Dynamical Systems*. Clarendon Press, Oxford, 1990.
 - [64] E. Scalco, K. Stec, D. Iudicone, M. I. Ferrante, and M. Montresor. The dynamics of sexual phase in the marine diatom *Pseudo-nitzschia multistriata* (Bacillariophyceae). *Journal of Phycology*, 50(5):817–828, 2014. doi: 10.1111/jpy.12225.
 - [65] R. Schwarz, M. Wolf, and T. Müller. A probabilistic model of cell size reduction in *Pseudo-nitzschia delicatissima* (Bacillariophyta). *Journal of Theoretical Biology*, 258(2):316–322, 2009. doi: 10.1016/j.jtbi.2009.02.002.
 - [66] H. Smith. *An Introduction to Delay Differential Equations with Applications to the Life Sciences*. Texts in Applied Mathematics 57. Springer-Verlag, New York, 2011.
 - [67] J. Smoller. *Shock Waves and Reaction Diffusion Equations*. Grundlehren der Mathematischen Wissenschaften 258. Springer-Verlag, New York, 2nd edition, 1994.
 - [68] H. A. V. Stosch. Manipulierung der Zellgrösse von Diatomeen im Experiment. *Phycologia*, 5(1):21–44, 1965. doi: 10.2216/i0031-8884-5-1-21.1.
 - [69] A. Terzieva and G. Terziev. Model diatom population by branching stochastic processes. *Applied and Computational Mathematic*, 8(1):3–8, 2019. doi: 10.11648/j.acm.20190801.12.
 - [70] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Graduate Studies in Mathematics 140. American Mathematical Society, Providence, 2012.
 - [71] P. Tréguer, C. Bowler, B. Moriceau, S. Dutkiewicz, M. Gehlen, O. Aumont, L. Bittner, R. Dugdale, Z. Finkel, D. Iudicone, O. Jahn, L. Guidi, M. Lasbleiz, K. Leblanc, M. Levy, and P. Pondaven. Influence of diatom diversity on the

- ocean biological carbon pump. *Nature Geoscience*, 11(1):27–37, 2018. doi: 10.1038/s41561-017-0028-x.
- [72] A. Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. B*, 237(641):37–72, 1952. doi: 10.1098/rstb.1952.0012.
- [73] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Texts in Applied Mathematics 2. Springer-Verlag, New York, 2nd edition, 2003.
- [74] R. Williams. Division rates of salt marsh diatoms in relation to salinity and size. *Ecology*, 45(4):877–880, 1964. doi: 10.2307/1934940.