

Article

Not peer-reviewed version

---

# Modeling Properties of Diatoms with Fibonacci Growth Using Lindenmayer Systems

---

[Thomas Harbich](#)\*

Posted Date: 24 April 2025

doi: 10.20944/preprints202504.2004.v1

Keywords: diatom; MacDonald-Pfitzer rule; Lindenmayer system; Fibonacci growth; delayed cell division; chain-forming diatoms; size sequence; fractal



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

# Modeling Properties of Diatoms with Fibonacci Growth Using Lindenmayer Systems

Thomas Harbich

Independent Researcher; Am Brudenrain, Weissach im Tal, Germany; mail@thomas-harbich.de

**Abstract:** In most cases, the sizes of the daughter cells of diatoms follow the MacDonald-Pfitzer rule, whereby in many species all diatoms divide once in each generation. In contrast, there are division schemes in which the smaller or larger daughter cell is delayed in its division by one generation and therefore lead to Fibonacci growth. Several properties of diatoms, especially in chain-like colonies, that exhibit such delayed division can be modelled by Lindenmayer systems. These include, above all, the size and orientation of the diatoms. Certain sequences of properties, such as the differences in size indices of neighboring diatoms, are aperiodic and represent self-similar fractal structures. For the division schemes studied, explicit solutions can be found for the number of diatoms of a certain size in each generation. For the experimental differentiation of the division schemes in a diatom chain, in addition to the observation of the division processes over several generations, methods are available that only require the analysis of the structure of a sufficiently large sample. This includes the investigation of the differences in the sizes of neighboring diatoms, the orientations of the diatoms and the frequencies of size indices in a culture.

**Keywords:** diatom; MacDonald-Pfitzer rule; Lindenmayer system; Fibonacci growth; delayed cell division; chain-forming diatoms; size sequence; fractal

---

## 1. Introduction

Diatoms are an ecologically important class of unicellular algae that can be found in almost all aquatic systems and in terrestrial environments. Diatoms have a hydrated silica exoskeleton consisting of two halves of different sizes, the hypotheca and the epitheca. The epitheca overlaps the hypotheca. Each theca consists of a valve, which varies in shape depending on the species, and a number of associated silicate girdle bands. A detailed description can be found in Round, Crawford and Mann [1].

Diatoms reproduce predominantly by vegetative division. As early as 1869, Macdonald [2] and Pfitzer [3] recognized in the diatoms they studied that during such cell division, the epitheca of the mother cell becomes the epitheca of the larger daughter cell and the hypotheca of the mother cell becomes the epitheca of the smaller daughter cell. For each theca of the mother cell, a slightly smaller theca is formed. This results in one cell the size of the mother cell and one smaller cell. This rule has been confirmed in many species (e.g. [4]), but exceptions are also known [5]. When a minimal size is reached, sexual reproduction occurs, resulting in one or more large initial cells. There are also species that can restore their size by vegetative enlargement (see [6] for a review).

Macdonald and Pfitzer assume in their considerations that there is a uniform doubling time. The larger and smaller daughter cells divide simultaneously and therefore the number of diatoms in each generation doubles. Such a division behavior of a species, which follows the MacDonald-Pfitzer rule, is referred to here as the P-model.

In the following, the generation is numbered with the index  $n$ , whereby the development should begin with a single diatom at  $n = 0$ . In the literature there are also generation counts starting with 1. In the context of cell size reduction for the morphological category of pennates (bilaterally symmetrical), the size of a diatom is understood to be the apical length. In the morphological category of centrics (radially symmetrical), the diameter of the circumcircle can be used. This length is

characterized by a size index  $k$ . The value  $k = 0$  corresponds to the cell of maximum size. With each reduction step,  $k$  is incremented until the maximum value  $k_{max}$  is reached, i.e. the size at which a cell enlargement is required. Using the notation  $C_p(n, k)$  for the number of diatoms with the size index  $k$  in the  $n$ -th generation in the P-model it follows from Macdonald-Pfitzer's rule for  $k > 0$ :

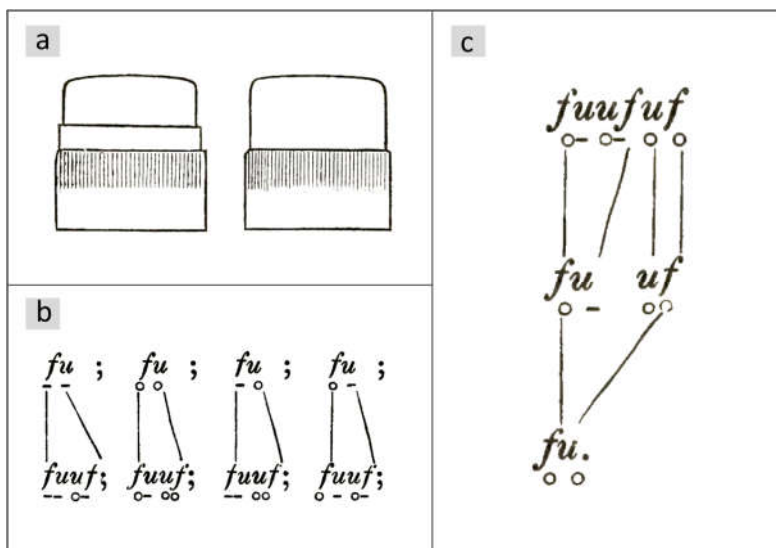
$$C_p(n + 1, k) = C_p(n, k) + C_p(n, k - 1). \quad (1)$$

It is known from Pascal's triangle that:

$$C_p(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (2)$$

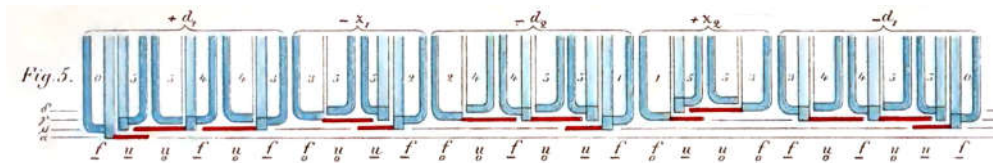
Without taking mortality processes into account and for diatom sizes above the threshold for the necessity of cell enlargement, the size classes are given by binomial coefficients. An early study applying this result to diatoms can be found in Richter [7].

Otto Müller observed a delayed vegetative division in certain diatoms in the chain-forming *Ellerbeckia arenaria* (Moore ex Ralfs) R.M.Crawford 1988 (basonym *Melosira arenaria* D.Moore, see [8]). He names the epithecae with the letter  $f$ , the hypothecae with  $u$ , so that a diatom in a chosen horizontal position is named by a pair  $fu$  or  $uf$ . In his publications from 1863 [9] and 1864 [10], he distinguishes between valves with and without thickening for the species under consideration (Figure 1a) and places a small circle under the letter if the valve has no thickening and a line otherwise, so that the combinations  $\overset{f}{-} \overset{u}{\circ}, \overset{f}{\circ} \overset{u}{-}, \overset{f}{-} \overset{u}{-}, \overset{f}{\circ} \overset{u}{\circ}$  and their mirror patterns result. For details of the morphology, see [11].



**Figure 1.** Drawings from the fundamental work by Otto Müller [9]. (a) Valves with and without thickening. (b) Rules for the shape of valves with respect to the presence of thickenings. The next generation is placed below the previous generation. (c) Applying the division rules to the diatom below results in the generations above. Conversely, generation 2 can also be derived from generation 0. In this case, the diatom  $\overset{u}{\circ} \overset{f}{\circ}$  takes on an interim state and divides one generation later.

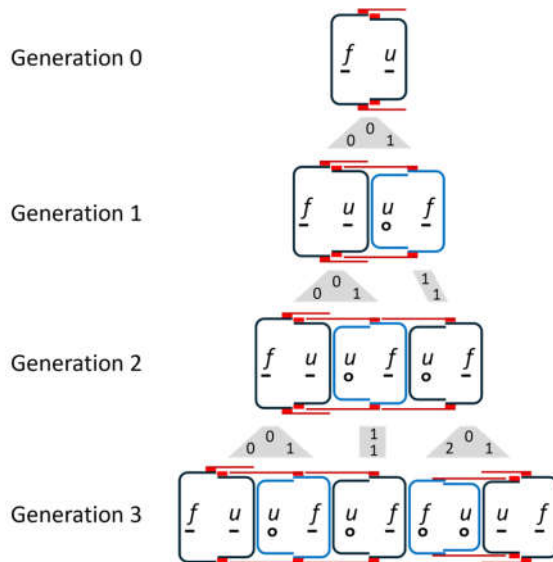
Figure 2 shows an illustration of a chain from [10]. The girdle bands are marked by red lines. Each is connected to the larger valve of the diatom and points to a neighboring thickening.



**Figure 2.** Colony of *Ellerbeckia arenaria* according to Otto Müller [10]. The black horizontal lines indicate the valve sizes.

If two diatoms in the chain have their larger valves next to each other, their girdle bands will point away from each other. There are no girdle bands that cover the valves of the two groups at their point of contact, which makes them visually separate. Otto Müller was able to show that this divides a chain into groups of two diatoms (sets of twins) or three diatoms (sets of triplets). In Figure 2, these groups are combined by horizontal brackets placed above the chain.

In all cell divisions, Otto Müller observed the transitions shown in Figure 1b. The valves of the parental cell keep their structure. For details, please refer to [9,10,12]. Regarding the temporal development, it is essential that there is a delayed division, which Müller shows exemplarily for a chain of three diatoms (Figure 1c). According to the rule in Figure 1b, he deduces from this sequence back to a single starting cell  $f_u$ , whereby the smaller daughter cell  $u_f$ , which has inherited the hypotheca ( $u_o$  becomes  $f_o$ ) from the mother cell, skips a division. In the next generation step, it divides again. According to Müller, such a skipping of one division of the smaller daughter cell occurs in all structures shown in Figure 1b. This implies asymmetry in the division behavior. Such behavior is referred to below as the M-model. Based on the illustration in [13], Figure 3 shows the first three subsequent generations of the development of a diatom chain starting with  $f_u$ .



**Figure 3.** Sequence of generations according to the model of O. Müller. It is striking that each generation is the beginning of the next.

The valves of newly formed cells that will not divide in the next step, i.e. are in an interim state, are indicated by blue valves. The grey shapes between the generations visualize the mapping of mother cells to daughter cells. The numbers in these figures are the size indices of mother (top) and daughter cells (bottom), assuming that development began with an initial cell. Due to the skipping of divisions, the number of diatoms does not grow exponentially. This will be discussed in more detail later.

The third generation consists of a set of triplets followed by a set of twins, each with overlapping girdle bands. When drawing the development for the following generations, it becomes clear that the entire chain consists of these two groups.

A surprisingly different observation was made by Laney, Olson, and Sosik [14] on *Ditylum brightwellii* using time-lapse imaging. They report that the daughter cell that inherits the hypothecal frustule divides in each generation. However, the division of the larger daughter cell is delayed in comparison. This model is subsequently denoted as the L-model. In an idealized approach, in which exactly one division is skipped, the L-model and the M-model form a pair of opposites regarding division behavior. For further details, particularly on the temporal behavior of the development, see [13,15].

The structure of a chain-like colony whose diatoms divide according to the McDonalds-Pfizer rule, in particular the size sequence of the diatoms and the number of diatoms of a certain size at a given generation obviously depend on the model of division. In this work, the properties of such chains and the distribution of sizes will be studied for the presented models. The background to this question is the possibility of investigating the division behavior of a species under consideration. Direct evidence of delayed division requires an observation period on the order of several generations. It should be borne in mind that the question arises in the case of chain-like diatoms, other diatoms living in colonies and individual diatoms. Motility complicates observation. On the other hand, easily testable criteria require the fulfilment of certain conditions, in particular sufficient synchronicity of division processes and the negligibility of death processes within the examined sample.

## 2. Materials and Methods

### 2.1 Synchronicity

Chronological sequences of patterns, such as those shown in Figure 1c and Figure 3, are referred to as generations in the literature cited (e.g. [13]), although not all diatoms divide in the time interval between these snapshots but skip a doubling period. This convenient counting method will also be used here. However, restrictions and limitations of this representation should be pointed out.

A description of the processes, as given by O. Müller [9,10], is implicitly based on the assumption of equal doubling times, whereby the doubling time is understood as the period between successive vegetative cell divisions. In the case of cells that skip a division, the double interval is assumed. It is known that doubling times and thus also the generation time, which is defined as the mean doubling time, depend on environmental parameters such as light intensity, nutrient concentration or temperature (see e.g. [16]). Even under constant environmental conditions, the doubling time is a random variable that has a natural statistical range of variation and can be described by a density function [17]. In the case of the M-model and L-model, there are therefore two distributions, one for division without delay and one for division with delay.

The stochastic character of the doubling times leads to a temporal drifting apart of the divisions in a culture. When starting a clonal culture with a single cell, after many divisions you will encounter a wide variety of division stages occurring simultaneously. For the applicability of the given considerations and the visualization of the generations of a chain, it is essential that the width of the distribution, e.g. measured as the standard deviation, is small compared to the generation time. The smaller this ratio is, the longer a real colony as shown in Figure 2 can grow, which corresponds to the theoretical structure resulting from the stepwise application of the rules in Figure 1b. Observations and quantitative considerations on the synchronicity of divisions in diatom chains can be found in [18,19].

It should be mentioned that cell divisions of diatoms can be synchronized by varying environmental parameters periodically. This can be achieved by periodic changes between phases of high and low light intensity [20,21,22,23]. Periodic addition of silicate to colonies that are kept under silicate deprivation can also be used for synchronization [24,25]. Provided they can be successfully

applied, these methods offer the fundamental possibility of studying chain-like diatom colonies over many generations.

## 2.2 Lindenmayer systems

A Lindenmayer system (L-system) is a mathematical formalism developed by Aristid Lindenmayer that is particularly suited to modeling the growth processes of plant development [26,27]. It is a triple consisting of:

- Alphabet
- Replacement rules (rewriting rules or productions)
- Initial sequence  $\omega$  (axiom)

The alphabet is a set of letters. The term "letter" is broadly defined, as a letter can also contain indices and thus numbers.

The formalism starts with the initial string  $\omega$ , a sequence of letters of the alphabet. The replacement rules are applied to this string. By replacing individual parts (individual characters, substrings), a new string of letters of the alphabet is obtained. The application of the replacement rules is continued iteratively, so that a sequence of such strings is created. In the following, a generation of a chain-like diatom colony is simply understood as the associated string.

Koster and Lindenmeyer successfully used an L-system to describe the positions of heterocysts in a filamentous colony of the cyanobacterium *Anabaena catenula* [28]. As a one-dimensional system is described here, the term "one-dimensional L-system" is used. L-systems allow the treatment of structures of plants with complex branching, which enables the generation of computer-generated images of plants [29]. In the context of diatoms, a one-dimensional L-system can be used to determine the position of connection points in zig-zag-shaped chains of *Diatoma vulgare* [19]. Of importance for the study presented here is the fact that the size sequence of diatoms in a chain-like colony can be modeled by an L-system. This approach is described by Ussing et al. [30]. The authors assume that Macdonald-Pfitzer's rule is applicable and that all cells divide synchronously. The resulting sequence was analyzed in [18].

It is remarkable that Otto Müller [9] defined an alphabet with the abbreviations  $\overset{f}{\_}$ ,  $\overset{u}{\_}$ ,  $\overset{f}{\circ}$ ,  $\overset{u}{\circ}$  and specified substitution rules (Figure 1b). These can be understood not only at the level of pairs of letters of an alphabet, but also as replacement rules for single frustules. The only missing extension for a complete description as an L-system is one that allows the delayed division shown in Figure 1b to be included in the calculation scheme.

## 3. Mathematical treatment

### 3.1 The Müller model as a Lindenmayer system

#### 3.1.1 Alphabet

Whereas Müller bases his alphabet on the frustule as the smallest structure, here a single diatom is represented by one letter of the alphabet  $A_M$  in order to achieve a uniform representation for all the division models. This letter should have the form  $W_j^k$ , where  $W \in \{L, l, R, r\}$ . When the larger valve is on the left, it is labeled L or l; when the larger valve is on the right, it is labeled R or r. For modeling the delayed division of the smaller daughter cell, a distinction must be made between cells that can divide immediately and cells in the interim state. Cells that can divide within a doubling time are labeled with capital letters (L or R), cells in the interim state with lower case letters (l or r). In addition to the four possibilities for diatoms shown in Figure 1b, their mirror image forms (large valve on the right) must also be considered. As the size sequence of the diatoms in chains is to be studied, the size index of their larger frustule is added to each letter as a superscript number  $k$ . Thus,  $L_j^k$  should denote a diatom whose left valve is larger than its right valve, which can divide immediately and has the size index  $k$ . The index  $j$  is used to assign a structure to the diatom as shown in Figure 1b. Diatoms in the interim state, but otherwise with the same characteristics, are named  $l_j^k$ .

The mirror-image diatoms are correspondingly  $R_j^k$  and  $r_j^k$ . The assignment to the structures can be found in Table 1.

**Table 1.** Assignment of the letters of the alphabet to the structure according to Müller [9] with extensions with respect to size index and division behavior.

Structure	Diatoms immediately able to divide	Diatoms in an interim state
$\begin{array}{c} f \ u \\ - \ - \end{array}$	$L_1^k$	$l_1^k$
$\begin{array}{c} f \ u \\ - \circ \end{array}$	$L_2^k$	$l_2^k$
$\begin{array}{c} f \ u \\ \circ \ - \end{array}$	$L_3^k$	$l_3^k$
$\begin{array}{c} f \ u \\ \circ \ \circ \end{array}$	$L_4^k$	$l_j^k$
$\begin{array}{c} u \ f \\ - \ - \end{array}$	$R_1^k$	$r_1^k$
$\begin{array}{c} u \ f \\ \circ \ - \end{array}$	$R_2^k$	$r_2^k$
$\begin{array}{c} u \ f \\ - \ \circ \end{array}$	$R_3^k$	$r_3^k$
$\begin{array}{c} u \ f \\ \circ \ \circ \end{array}$	$R_4^k$	$r_4^k$

Although the size index has a maximum value  $k_{\max}$  in nature, this is not relevant for the following considerations, so that the alphabet  $A_{SOV}$  is defined by the set:

$$A_{SOV} = \{W_j^k \mid W \in \{L, l, R, r\}, j \in \{1,2,3,4\}, k \in \mathbb{N}_0\} \quad (3)$$

The alphabet takes into account size index (S), orientation (O) and valve structure (V). By removal of information like the structure index or size index, simplified alphabets can be obtained for limited considerations.

### 3.1.2 Replacement rules

When forming a generation, each letter of the previous generation is replaced by one or two letters of the alphabet. Consider the replacement rule of  $L_1^k$  as an example.

$$L_1^k \rightarrow L_1^k r_2^{k+1} \quad (4)$$

$L_1^k$  has the structure  $\begin{array}{c} f \ u \\ - \ - \end{array}$ . The size and structure of its larger daughter cell is identical according to Figure 1b, so that it is described as  $L_1^k$  like the mother cell. The smaller daughter cell exhibits the structure  $\begin{array}{c} u \ f \\ \circ \ - \end{array}$  and has structure index 2 according to Table 1. Its larger valve is on the right, and it is in the interim state, which is why it is written with a lowercase letter. Its size index is incremented because it is the smaller daughter cell. It divides with a delay of one generation and is thus described by  $r_2^{k+1}$ . The following applies:

$$r_j^k \rightarrow R_j^k \quad (5)$$

The memory is modeled by the intermediate step. All other productions are determined by an analogous procedure. For the diatoms with the large valve on the right, the replacement rules can be found by mirroring. All replacement rules are summarized below.

$$\begin{aligned}
L_1^k &\rightarrow L_1^k r_2^{k+1} \\
L_2^k &\rightarrow L_1^k r_4^{k+1} \\
L_3^k &\rightarrow L_3^k r_2^{k+1} \\
L_4^k &\rightarrow L_3^k r_4^{k+1} \\
R_1^k &\rightarrow l_2^{k+1} R_1^k \\
R_2^k &\rightarrow l_4^{k+1} R_1^k \\
R_3^k &\rightarrow l_2^{k+1} R_3^k \\
R_4^k &\rightarrow l_4^{k+1} R_3^k \\
l_j^k &\rightarrow L_j^k \\
r_j^k &\rightarrow R_j^k
\end{aligned} \tag{6}$$

As these replacement rules exactly determine the next generation and the neighborhood of a diatom is not included in the rules, this L-system is deterministic and context-free and is therefore referred to as a DOL-system.

If we start the development with the axiom  $\omega = L_j^0$  and denote the string of the  $n$ th generation by  $G_{j,M}^n$ , this generation sequence results for  $j = 1$ :

$$\begin{aligned}
G_{1,M}^0 &= \omega = L_1^0 \\
G_{1,M}^1 &= L_1^0 r_2^1 \\
G_{1,M}^2 &= L_1^0 r_2^1 R_2^1 \\
G_{1,M}^3 &= (L_1^0 r_2^1 R_2^1) (l_4^2 R_1^1) \\
G_{1,M}^4 &= (L_1^0 r_2^1 R_2^1) (l_4^2 R_1^1) (L_4^2 l_2^2 R_1^1) \\
G_{1,M}^5 &= (L_1^0 r_2^1 R_2^1) (l_4^2 R_1^1) (L_4^2 l_2^2 R_1^1) (L_3^2 r_4^3) (L_2^2 l_2^2 R_1^1) \\
&\dots
\end{aligned} \tag{7}$$

From generation 3 on, groups of twins and groups of triplets are indicated by parentheses. If one starts with a size index greater than 0, all size indices in all subsequent generations must be increased by this starting value according to the substitution rules.

$G_{1,M}^5$  corresponds to Figure 2, which indicates the same starting value. In this generation, all the diatom structures according to Figure 1b appear for the first time. As the L-system is context-free, each diatom develops into a structure in the same generation sequences that would have resulted if it had been chosen as an axiom. For example, from  $L_3^2$  in  $G_{1,M}^5$ , a subtree arises in subsequent generations, as it would have formed with the axiom  $L_3^2$ . As mentioned, this can also be mapped to the axiom  $\omega = L_3^0$  by lowering the size indices by the value 2.

### 3.1.3 Introduction of operators and alternative algorithm for calculating generations

In the following, an equivalent formulation is chosen for the replacements [18]. The productions are expressed by an operator  $P$ . The substitution rules assign the string  $P(g)$  to the letter  $g$  of the alphabet. Parentheses are not required for an operator. They are used here to indicate the range of the operator. The first substitution rule in (6) can be written as:

$$P(L_1^k) = L_1^k r_2^{k+1} \tag{8}$$

Thus the  $n$ th generation of the axiom  $\omega = L_j^0$  is expressed by  $G_{j,M}^n = P^n(L_j^0)$ . As the L-system is not context-sensitive and therefore each character is replaced separately, the following is valid when applying  $P$  to a string  $G$ :

$$P(G) = P(g_1 g_2 g_3 \dots g_j) = P(g_1) P(g_2) P(g_3) \dots P(g_j) \tag{9}$$

Furthermore, the operator  $S$  is introduced, which mirrors the string including its elements.

$$S(G) = S(g_1 g_2 g_3 \dots g_j) = S(g_j) S(g_{j-1}) \dots S(g_3) S(g_2) S(g_1) \tag{10}$$

To mirror the characters of the alphabet, use the structures defined in Table 1 or the verbal description. The following applies for all  $j$ :

$$\begin{aligned}
S(L_j^k) &= R_j^k \\
S(R_j^k) &= L_j^k \\
S(r_j^k) &= l_j^k \\
S(l_j^k) &= r_j^k
\end{aligned} \tag{11}$$

It is easy to convince yourself of the commutativity [18]:

$$S P = P S \tag{12}$$

If you change the view of a diatom chain from top view to bottom view, you will see the mirrored chain. If you determine the next generation by applying the replacement rules, you get the mirrored next generation. The same result is obtained if you first determine the next generation and then change the view, i.e. mirror it.

Finally, an operator  $I$  is defined which increments the size indices of all diatoms in a string:

$$I(G) = I(g_1 g_2 g_3 \dots g_j) = I(g_1) I(g_2) I(g_3) \dots I(g_j) \tag{13}$$

Whereby:

$$\begin{aligned}
I(L_j^k) &= L_j^{k+1} \\
I(R_j^k) &= R_j^{k+1} \\
I(r_j^k) &= r_j^{k+1} \\
I(l_j^k) &= l_j^{k+1}
\end{aligned} \tag{14}$$

As already stated, the boundary  $k < k_{\max}$  is not explicitly considered. These commutator relationships are valid as the operator  $I$  only affects the size indices:

$$I P = P I \tag{15}$$

$$I S = S I \tag{16}$$

### 3.1.4 Reformulation of the algorithm

The operators make it possible to reformulate the algorithm for calculating subsequent generations.

**Theorem 1.** For  $n > 1$  and  $\omega = L_j^0$  the following applies:

$$\begin{aligned}
G_{1,M}^{n+1} &= G_{1,M}^n IS G_{2,M}^{n-1} \\
G_{2,M}^{n+1} &= G_{1,M}^n IS G_{4,M}^{n-1} \\
G_{3,M}^{n+1} &= G_{3,M}^n IS G_{2,M}^{n-1} \\
G_{4,M}^{n+1} &= G_{3,M}^n IS G_{4,M}^{n-1}.
\end{aligned} \tag{17}$$

**Proof of Theorem 1.** By applying the substitution rules, the context independence and the commutator relations, we obtain the following for  $n > 1$ :

$$\begin{aligned}
G_{1,M}^{n+1} &= P^{n+1}(L_1^0) = P^n(P(L_1^0)) = P^n(L_1^0 r_2^1) = P^n(L_1^0) P^n(r_2^1) = \\
&P^n(L_1^0) P^{n-1}(R_2^1) = P^n(L_1^0) P^{n-1}(IS(L_2^0)) = \\
&P^n(L_1^0) IS P^{n-1}(L_2^0) = G_{1,M}^n IS G_{2,M}^{n-1}.
\end{aligned} \tag{18}$$

The  $G_{j,M}^{n+1}$  for  $j \neq 1$  are calculated in the same way.  $\square$

Each generation can be created by recursion from a previous generation and a second previous generation, without explicitly applying the replacement rules to each character of the string. The calculation scheme for calculating  $G_{j,M}^{n+1}$  for  $n > 1$  consists of concatenating the strings  $G_{j,M}^{n-1}$  and

IS  $G_{j,M}^{n-1}$  according to equations (18). The generations resulting from the axiom  $R_j^0$  are obtained by mirroring.

As Otto Müller [9,10] points out, the lengths of the generations are given by the Fibonacci sequence. Defining the Fibonacci sequence by  $f_0 = 0$ ,  $f_1 = 1$  and  $f_{n+2} = f_{n+1} + f_n$  then the number of letters in  $G_{j,M}^n$  equals  $f_{n+2}$  because of the lengths of the first two generations and equations (17). Note that the I and S operators do not change the length of a string.

### 3.1.5 Restriction to orientation and size index

The division rule, according to which the smaller daughter cell divides with a delay, is not fundamentally linked to a specific structure with thickenings. This information is now omitted and only the size indices and orientations of the diatoms are considered. The index describing the thickenings is thus dropped and the alphabet is reduced to:

$$A_{SO} = \{W^k \mid W \in \{L, l, R, r\}, k \in \mathbb{N}_0\}. \quad (19)$$

The replacement rules become:

$$\begin{aligned} L^k &\rightarrow L^k l^{k+1} \\ R^k &\rightarrow l^{k+1} R^k \\ l^k &\rightarrow L^k \\ r^k &\rightarrow R^k. \end{aligned} \quad (20)$$

In the generation sequence (7), only the subscripts have to be removed. Instead of the four equations in (17), we find for  $n > 1$ :

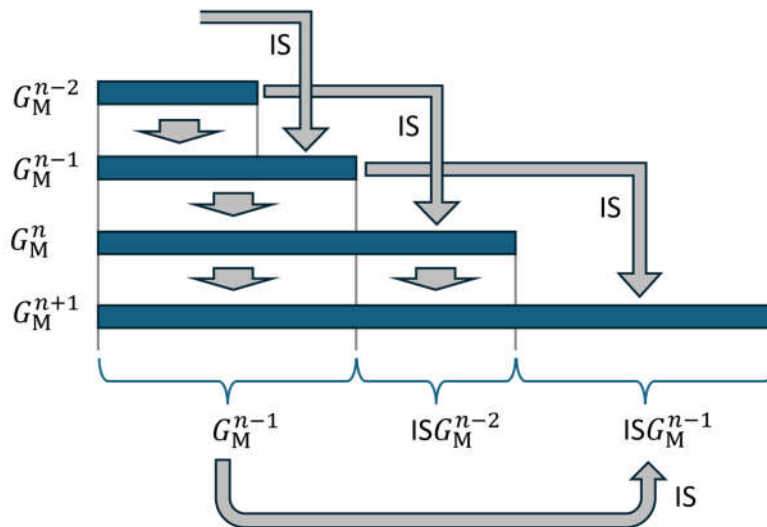
$$G_M^{n+1} = G_M^n \text{ IS } G_M^{n-1}. \quad (21)$$

A generation is obtained by concatenating the previous generation with the mirrored and incremented second-last generation. Each generation is the beginning of the next.

It should be remembered that the axiom  $L^0$  was used as axiom here. If you want to get the generation sequence for the axiom  $R^0$ , you just have to mirror  $G_M^n$ , because  $P^n(R^0) = P^n(SL^0) = SP^n(L^0) = SG_M^n$ . If you want to obtain an expression for the sequence of generations for the axiom  $R^0$  analogous to (21), apply the mirror operator to both sides of this equation. Using the notation  $H_M^n = P^n(R^0) = SG_M^n$ , you get  $H_M^{n+1} = SG_M^{n+1} = S(G_M^n \text{ IS } G_M^{n-1}) = S(\text{IS } G_M^{n-1}) S G_M^n = I(S(H_M^{n-1})) H_M^n$ . Each generation is the end of the next.

A higher size index in the axiom  $L^k$  ( $k > 0$ ) causes a corresponding increase in all indices, because  $P^n(L^k) = P^n(l^k L^0) = l^k P^n(L^0)$ . If you want to start the development with an interim state, only one additional step, i.e. one generation time is required to achieve the described generation sequence. In this sense, the alternative description of the L-system (21) is complete. The same procedure can be followed for the other L-systems in this work if you want to change the axiom.

The recursive structure in equation (21) enables a representation in which  $G_M^{n+1}$  only depends on  $G_M^n$ . Using equation (21),  $G_M^n = G_M^{n-1} \text{ IS } G_M^{n-2}$  applies. As  $G_M^{n-1}$  forms the first part of  $G_M^n$  with the length  $f_{n+1}$ , this string can also be taken from  $G_M^n$  when calculating  $G_M^{n+1}$  according to (21). This is shown graphically in Figure 4.



**Figure 4.** Four generations are represented by horizontal blue bars. Each generation is formed by copying the previous generation followed by the mirrored and incremented second last generation. This formation is indicated by the grey arrows.

This means that  $G_M^{n+1}$  can be determined from  $G_M^n$  alone. This is indicated by the grey arrow at the bottom of the figure.

As can be seen in generation  $n + 1$ , the sequence can also be regarded as consisting of three strings, with the middle one from generation  $n - 2$  being created by applying IS. The whole generation  $n + 1$  is mirror symmetric if the size indices are neglected, except for this middle string. This is mentioned in [13] for the example of the 5th generation shown in the paper.

### 3.1.6 Size indices

According to the alphabet, each generation contains information about size indices and orientations, including interim states. One can go a step further and focus on orientations including the interim states or on size indices. Restricting to size indices, each generation can be written as  $a_1^n, a_2^n, a_3^n, \dots, a_{f_{n+2}}^n$  with  $a_i^n \in \mathbb{N}_0$ . The size index  $a_j^n$  is the  $j$ -th element in the generation  $n$ . The alphabet is reduced to  $A_S = \mathbb{N}_0$ . Expression (21) is still applicable. The mirror operator only changes the order, but not the values. In terms of size indices, (21) can be rewritten as

$$a_i^{n+1} = \begin{cases} a_i^n & \text{for } i = 1, 2, \dots, f_{n+2} \\ a_{f_{n+3}-i+1}^{n-1} + 1 & \text{for } i = f_{n+2} + 1, \dots, f_{n+3} \end{cases} \quad (22)$$

According to the note above, for  $i > f_{n+2}$ , the values of generation  $n - 1$  contributing to generation  $n + 1$  can also be taken from generation  $n$ .

### 3.1.7 Sequence of differences and fractal structure

As the absolute assignment of a measured size to a size index is not straightforward, the difference sequence of the size indices is of particular importance. This is discussed in more detail below. The difference sequence of a generation  $D_M^n$  is defined for  $n \geq 1$  by a sequence of integers  $d_j^n$

$$d_i^n = a_{i+1}^n - a_i^n, \quad (23)$$

where  $i$  takes the values  $1, 2, \dots, f_{n+2} - 1$ . The string  $D_M^n$  has the length  $f_{n+1} - 1$ . Using (22) yields for  $n > 2$ :

$$d_i^{n+1} = \begin{cases} d_i^n & \text{for } i = 1, 2, \dots, f_{n+2} - 1 \\ 1 & \text{for } i = f_{n+2} \\ -d_{f_{n+3}-i}^{n-1} & \text{for } i = f_{n+2} + 1, \dots, f_{n+3} - 1 \end{cases} \quad (24)$$

As can be quickly verified, the difference between the first value of the second subsequence and the last value of the first subsequence is always equal to 1. Again, the self-similarity of the sequences allows the differences from generation  $n - 1$  that contribute to generation  $n + 1$  to be taken from generation  $n$  for  $i > f_{n+2}$ . In compact notation, (24) can be written as follows for  $n > 2$

$$D_M^{n+1} = D_M^n \mathbf{1} \overline{SD_M^{n-1}}, \quad (25)$$

where the upper bar is meant to indicate the inversion of all signs. The first five generations read:

$$\begin{aligned} D_M^1 &= 1 \\ D_M^2 &= 1, 0 \\ D_M^3 &= 1, 0, \mathbf{1}, -1 \\ D_M^4 &= 1, 0, 1, -1, \mathbf{1}, 0, -1 \\ D_M^5 &= 1, 0, 1, -1, 1, 0, -1, \mathbf{1}, 1, -1, 0, -1 \end{aligned} \quad (26)$$

The values 1 for  $i = f_{n+2}$  are shown in bold.

When looking at the difference sequence, it is striking that only the values -1, 0 and 1 occur in it. The proof that this applies to all generations can be done by induction (for the proof technique, please refer to textbook [31]).  $D_M^3$ , in which all stated values occur, is taken as the base case. The induction step uses the fact that the range of values {-1, 0, 1} is not changed when applying the iteration described in (24) and (25). It should be noted that for division rules without delayed division (P-model) only the differences -1 and -1 occur, but not the value 0 [18]. There are no neighboring diatoms of the same size in the P-model.

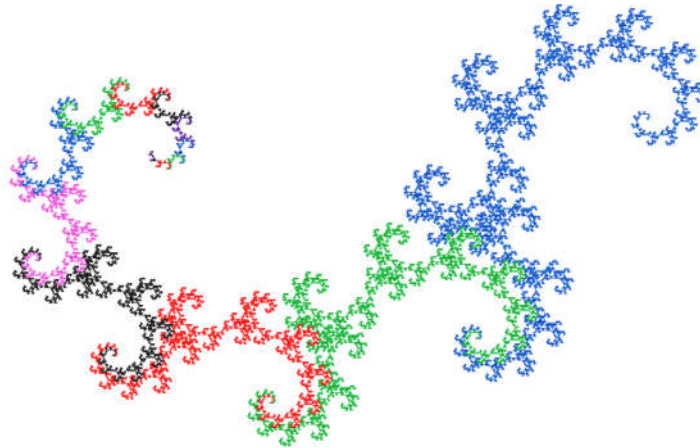
The sequences according to equation (21) cannot become periodic as the generation index  $n$  increases, because the maximum size index increases continuously as a result of incrementing. Appendix A shows that the sequences of differences do not become periodic either.

It is known that self-similarity and fractal structures are typical for Lindenmayer systems [29,24]. In a P-model, the sequence of differences can be visualized as a dragon curve [18]. To do this, a line with a unit length is first drawn in the manner of a turtle graphic. A rotation is then performed for each element of the sequence, the size of which depends on the value of the element, and then a step is taken forwards by the same unit of length. The dragon curve is obtained by performing a 90° rotation in a clockwise direction for the value +1 and a 90° counterclockwise rotation for the value -1. The following applies to the generations of size differences  $D_P^n$  in the P-model [18]:

$$D_P^{n+1} = D_P^n \mathbf{1} \overline{SD_P^n}. \quad (27)$$

A curve segment of the same length  $\overline{D_P^n}$  is added to the existing curve  $D_P^n$  after rotation by 90° (value 1). There is self-similarity, in particular  $D_P^n$  is similar to  $D_P^{n+1}$ . For details on the dragon curve, please refer to Tabachnikov [32].

In the case of the M-model, there are three different values for size differences. An appealing visualization of the sequence can be obtained by proceeding similarly to the dragon curve but choosing an angle of 120° and not performing a rotation for the value 0, i.e., appending just one step of one unit of length in the same direction. The result is shown in Figure 5.



**Figure 5.** Visualization of the difference in size indices of the diatoms according to Müller's model. 20 generations were used for the image. The curve starts at the top left, where the color changes are close together.

Each sequence  $D_p^n$  is assigned a different color. The size of the curves does not double from generation to generation but follows the Fibonacci sequence. For the next generation, curve segments of different lengths are combined. Each generation is similar to every other generation, which can be seen by scaling them to the same size and aligning them in the same direction.

### 3.1.8 Orientations

The orientation and the interim states can be described by a simple L-system and can be computed independently of the size indices. We restrict ourselves to the alphabet  $A_0 = \{l, L, r, R\}$ . For example,  $L$  can be chosen as the axiom. The substitution rules follow directly from (20) to give  $L \rightarrow Lr$ ;  $R \rightarrow lR$ ;  $l \rightarrow L$ ;  $r \rightarrow R$ . Equation (21) can still be used, only the increment operator is no longer needed, i.e.  $J_M^{n+1} = J_M^n S J_M^{n-1}$ , where the letter  $J$  has been chosen for the generations of orientations. The mirroring is done according to the rules (11), where the size index is irrelevant.

The sequence of orientations is not periodic, which does not immediately result from what has been said so far. The proof follows the scheme of the proof for the sequence of differences (Appendix A). The number of occurrences of a letter of the alphabet in generation  $n$  is not given by a Fibonacci number. Therefore, for a generation  $n$ , consider the number  $u_n$  of all lowercase letters ( $l, r$ ) and the number  $U_n$  of all uppercase letters ( $L, R$ ). The operator  $S$  does not replace lowercase letters with uppercase letters, i.e. this number is invariant when  $S$  is applied. From  $J_M^{n+1} = J_M^n S J_M^{n-1}$  we can see that these numbers have Fibonacci growth. Taking into account the initial conditions, we obtain  $u_n = f_n$  and  $U_n = f_{n+1}$ . The ratio  $u_n/U_n$  converges to an irrational number for  $n \rightarrow \infty$ , which would not be the case if the sequence were periodic.

The sequence as seen in (7) can be visualized by a fractal curve if one chooses a rotation angle and assigns certain multiples of this angle to the characters of the alphabet. With a P-model there is an alternating sequence of right and left orientations [18]. Aperiodicity and a fractal structure therefore do not occur.

### 3.1.9 Sets of twins and sets of triplets

Sets of twins and sets of triplets are characterized by free thecae at both ends [9,10,13]. This property is retained when mirroring and incrementing are applied. Since structures of sets of twins and sets of triplets arise from a single cell, every chain that develops according to O. Müller's model consists exclusively of a sequence of these two structures. It is not surprising that this sequence can also be described by an L-system. In contrast to the systems presented so far, the objects here do not stand for a single diatom, but two or three diatoms. It can be seen in (7) that a set of twins becomes a set of triplets in the next generation and then the set of triplets becomes a set of triplets and a set of

twins. The order in which sets of triplets and sets of twins appear depends on the orientation of the set of twins with which the generation sequence begins. Consequently, orientation must also be taken into account in the alphabet. We denote a set of twins in the orientation as shown in (7) for the second generation (structure Lr) with the letter T and its mirror image with  $T_s = ST$ . Correspondingly, a set of triplets as in the third generation (structure LrR) should be denoted by V and its mirror image by  $V_s = SV$ , so that the alphabet is given by

$$A_T = \{T, V, T_s, V_s\}. \quad (28)$$

The production rules for the above specified structures result from  $L \rightarrow Lr$ ;  $R \rightarrow IR$ ;  
 $l \rightarrow L$ ;  $r \rightarrow R$  to give

$$\begin{aligned} T &\rightarrow V \\ V &\rightarrow VT_s \\ T_s &\rightarrow V_s \\ V_s &\rightarrow TV_s. \end{aligned} \quad (29)$$

When starting the development with T, a set of triplets and a set of twins appear after two iterations, whereby the set of twins has the opposite orientation to the parent. To obtain the replacement rules for  $T_s$  and  $V_s$ , it is sufficient to mirror the rules for T and V. If we restrict ourselves to the axiom V without loss of generality (otherwise mirroring and/or executing an iteration step) and label the generations  $Q_M^n = P^n(V)$ , we obtain analogous results to Theorem (21)

$$\begin{aligned} Q_M^{n+1} &= P^{n+1}(V) = P^n(P(V)) = P^n(VT_s) = P^n(V) P^n(T_s) \\ &= Q_M^n P^{n-1}(P(T_s)) = Q_M^n P^{n-1}(V_s) = Q_M^n P^{n-1}(S(V)) = Q_M^n S Q_M^{n-1}. \end{aligned}$$

and thus

$$Q_M^{n+1} = Q_M^n S Q_M^{n-1}. \quad (30)$$

The similarity with the expression for the orientations is due to the fact that the replacement rule  $T \rightarrow V$  does not increase the number of objects and acts as a delay, even if the number of diatoms does not remain the same. The first generations are:

$$\begin{aligned} Q_M^0 &= \omega = V \\ Q_M^1 &= VT_s \\ Q_M^2 &= VT_s V_s \\ Q_M^3 &= VT_s V_s TV_s \\ Q_M^4 &= VT_s V_s TV_s VTV_s \\ Q_M^5 &= VT_s V_s TV_s VTV_s VT_s VTV_s \end{aligned} \quad (31)$$

Again, there is Fibonacci growth with the length of  $Q_M^n$  equal to  $f_{n+2}$ . The number of groups of twins grows with  $f_n$ , the number of groups of triplets with  $f_{n+1}$ , whereby the orientation was not considered, i.e. both orientations were grouped together. The evidence for non-periodicity presented in Appendix A is therefore transferable, considering the quotient of the number of groups of twins and groups of triplet groups.

### 3.1.10 Number of diatoms of the same size

For practical applications, it is important to know how many diatoms with a certain size index  $k$  occur in generation  $n$ . In Müller's model, this number is denoted by  $C_M(n, k)$ . If the development begins with a cell of maximum size ( $k = 0$ ), then  $C_M(0, 0) = 1$  and  $C_M(0, k) = 0$  for  $k > 0$ . In the following generation there is again a cell of maximum size and a cell with the following size index, so that  $C_M(1, 0) = 1$  and  $C_M(1, 1) = 1$ . There are no further cells, thus  $C_M(1, k) = 0$  for  $k > 1$ . For the subsequent generations with  $n > 1$ , a recursion formula can be derived from (21):

$$C_M(n + 1, k) = C_M(n, k) + C_M(n - 1, k - 1) \quad (32)$$

The number of diatoms of size  $k$  in the generation  $n + 1$  has a part, which results from the identical repetition of the previous generation  $n$  and a second part, which originates from the second last

generation  $n - 1$ . As all size indices are hereby incremented, the sizes  $k - 1$  contribute to this. It should be noted that there are no size indices smaller than zero and therefore no contribution to  $k = 0$  from generation  $n - 1$ . Therefore  $C_M(n, -1) = 0$  is defined. For  $k = 0$  one can immediately find that  $C_M(n, 0) = 1$  for any  $n$ .

To find the solution (33), you can iterate (32). For a chosen  $k$  as a function of  $n$ , a sequence of figurate numbers arises, which can be expressed by binomial coefficients.

**Theorem 2.** *The solution of the equation (32) for  $n > 1$  reads:*

$$C_M(n, k) = \binom{n - k + 1}{k} \quad (33)$$

Note that  $C_M(n, k) = 0$  for  $n - k + 1 < k$ , i.e. for  $k > \lfloor (n + 1)/2 \rfloor$ .

**Proof of Theorem 2.** For  $n = 0$  and  $n = 1$ , the correspondence with the initial values is checked by inserting them into the binomial coefficient. Similarly, for  $k = 0$  and any  $n$ , (33) gives the value 1 in accordance with (32). Proof of correctness for  $k > 0$  is obtained by inserting it into the right-hand side of (32).

$$\begin{aligned} \text{RHS} &= \binom{n - k + 1}{k} + \binom{(n - 1) - (k - 1) + 1}{k - 1} \\ &= \binom{n - k + 1}{k} + \binom{n - k + 1}{k - 1} = \binom{n - k + 2}{k} \end{aligned}$$

In the last step, the recurrence relation for binomial coefficients was used. Inserting the binomial coefficient into the left-hand side immediately results in

$$\text{LHS} = \binom{n - k + 2}{k}.$$

And therefore LHS=RHS.  $\square$

It is worth noting that all binomial coefficients can be found in Pascal's triangle. When this is left-justified, all solutions (33) for a chosen  $n$  are located on a diagonal. The sum of all binomial coefficients of a diagonal gives the Fibonacci numbers, as can also be seen from the recursion (32).

### 3.2 The Model according to Laney et al. as a Lindenmayer system

#### 3.2.1 Notes on the model

At first glance, the M-model and L-model look very similar, so you might think that the results would be also very similar. However, this is only partly the case. The operators and methods introduced are fully applicable.

In the following, the same basic and partly simplifying assumptions are made, in particular the existence of synchronous divisions and a division delay of one generation for certain diatoms.

Valve structures and girdle band structures are not considered, so that the investigation is limited to sizes and orientations.

#### 3.2.2 Alphabet and replacement rules

Because of the restrictions on size and orientation, the alphabet  $A_{S0}$  defined in (19) can be used. The replacement rules differ from those shown in (20) because the division of the larger cell is now delayed by one generation:

$$\begin{aligned} L^k &\rightarrow l^k R^{k+1} \\ R^k &\rightarrow L^{k+1} r^k \\ l^k &\rightarrow L^k \\ r^k &\rightarrow R^k. \end{aligned} \quad (34)$$

If the string for the  $n$ th generation at the start value  $L^0$  is denoted by  $G_L^n$ , then the L-system according to (34) yields this generation sequence:

$$\begin{aligned} G_L^0 &= \omega = L^0 \\ G_L^1 &= l^0 R^1 \\ G_L^2 &= L^0 L^2 r^1 \end{aligned} \quad (35)$$

$$\begin{aligned}
 G_L^3 &= l^0 R^1 l^2 R^3 R^1 \\
 G_L^4 &= L^0 L^2 r^1 L^2 L^4 r^3 L^2 r^1 \\
 G_L^5 &= l^0 R^1 l^2 R^3 R^1 l^2 R^3 l^4 R^5 R^3 l^2 R^3 R^1 \\
 &\dots
 \end{aligned}$$

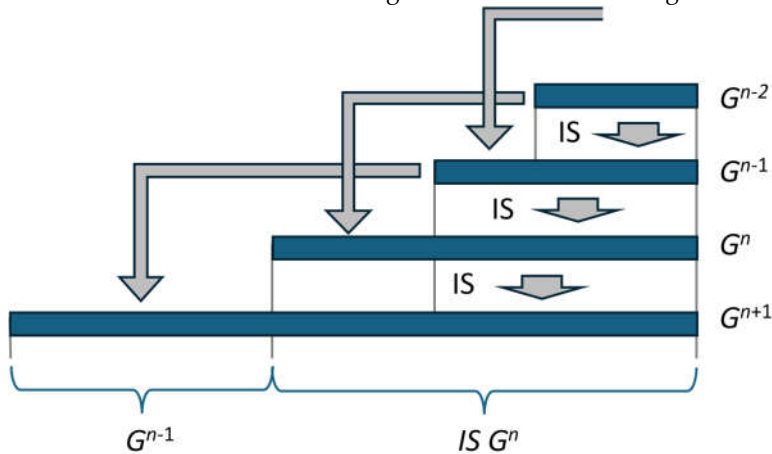
Using the presented method, we find

$$\begin{aligned}
 G_L^{n+1} &= P^{n+1}(L^0) = P^n(P(L^0)) = P^n(l^0 R^1) = P^n(l^0) P^n(R^1) \\
 &= P^{n-1}(L^0) P^n(IS L^0) = P^{n-1}(L^0) IS P^n(L^0) = G_L^{n-1} IS G_L^n
 \end{aligned}$$

and thus

$$G_L^{n+1} = G_L^{n-1} IS G_L^n. \tag{36}$$

The concatenation of the last generation with the second last generation results in a subsequent generation with the sum of the lengths and thus a Fibonacci growth. As in the case of the M-model, the length of  $G_L^n$  is equal to  $f_{n+2}$ . In contrast to  $G_M^n$  and  $G_P^n$ , however, no sequence is created in which the previous one is repeated, neither at the beginning nor at the end, so that the question of periodicity does not arise. The formation of a generation is shown in Figure 6.



**Figure 6.** To obtain a well-structured graphic, the bars representing generations have been right aligned. No part of a previous generation is copied unchanged. This results in a sequence that is not continuously extended.

As in Figure 4, a decomposition into three strings can be given.

### 3.2.3 Size indices

If we restrict ourselves again to the size indices  $a_1^n, a_2^n, a_3^n, \dots, a_{f_{n+2}}^n$  with  $a_i^n \in \mathbb{N}_0$ , it follows from (36):

$$a_i^{n+1} = \begin{cases} a_i^{n-1} & \text{for } i = 1, 2, \dots, f_{n+1} \\ a_{f_{n+3}-i+1}^n + 1 & \text{for } i = f_{n+1} + 1, \dots, f_{n+3} \end{cases} \tag{37}$$

### 3.2.4 Sequence of differences and fractal structure

The difference sequence  $D_L^n$  of a generation for  $n > 1$  will also be given for this model. For this purpose, the differences  $d_i^n = a_{i+1}^n - a_i^n$  with  $i = 1, 2, \dots, f_{n+2}$  are formed using (37). For  $n > 2$  you get

$$d_i^{n+1} = \begin{cases} d_i^{n-1} & \text{for } i = 1, 2, \dots, f_{n+1} - 1 \\ 1 & \text{for } i = f_{n+1} \\ -d_{f_{n+3}-i+1}^n & \text{for } i = f_{n+1} + 1, \dots, f_{n+3} - 1 \end{cases} \tag{38}$$

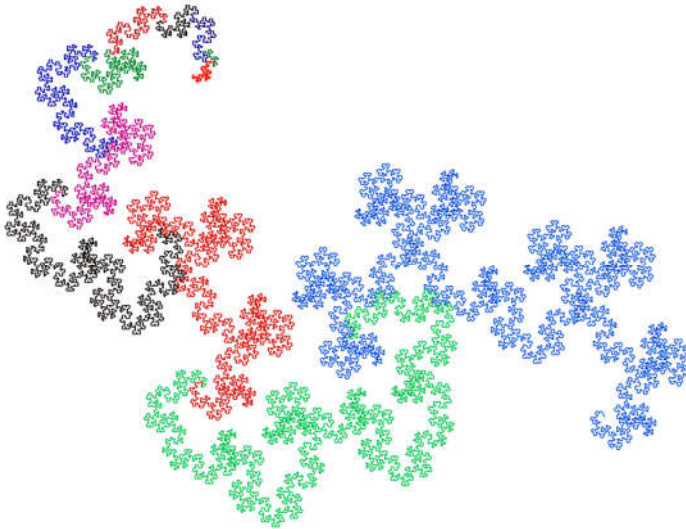
The compact version with the notation used in (25), is as follows

$$D_L^{n+1} = D_L^{n-1} 1 \overline{SD_L^n}. \quad (39)$$

These are the first five sequences:

$$\begin{aligned} D_L^1 &= 1 \\ D_L^2 &= 2, -1 \\ D_L^3 &= 1, \mathbf{1}, 1, -2 \\ D_L^4 &= 2, -1, \mathbf{1}, 2, -1, -1, -1 \\ D_L^5 &= 1, 1, 1, -2, \mathbf{1}, 1, 1, 1, -2, -1, 1, -2 \end{aligned} \quad (40)$$

The values 1 for  $i = f_{n+1}$  are shown in bold. The range of values that differences can take on is larger than in the M-model and contains the values -2, -1, 1, 2. The proof is again by induction. Compared to the representation of the sequence of differences in the M-model, a different choice of angle proves to be favorable. In Figure 7, an angle of rotation of 60 degrees was chosen. With an absolute value of 2, two rotations are performed, i.e. one rotation of 120 degrees. Apart from that, the illustrations correspond.



**Figure 7.** Visualization of the difference in size indices of the diatoms according to the L-model. 20 generations were used for the image. The curve starts at the top left, where the color changes are close together.

Fractal structure and self-similarity are evident.

### 3.2.5 Orientations

If we consider only the orientations, i.e. restrict ourselves to the alphabet  $A_0 = \{l, L, r, R\}$ , the reduced substitution rules by  $L \rightarrow lR$ ;  $R \rightarrow Lr$ ;  $l \rightarrow L$ ;  $r \rightarrow R$ , using the same axiom as for the M-model, lead to a non-periodic sequence  $J_L^n$ . In recursive notation this sequence is given by  $J_L^{n+1} = J_L^{n-1} S J_L^n$ . The first sequences are obtained from (35) by omitting the size indices.

### 3.2.6 Number of diatoms of the same size

The number of diatoms of a given size index  $k$  in generation  $n$  is also to be determined for the L-model. This is denoted by  $C_L(n, k)$ . The initial conditions for the calculation are identical to those for the M-model, i.e.  $C_L(0,0) = 1$ ,  $C_L(1,0) = 1$  and  $C_L(1,1) = 1$ . All other values in generations 0 and 1 are zero. The recursion for  $C_L(n, k)$  for  $n > 1$  is derived from equation (36) and gives

$$C_L(n+1, k) = C_L(n-1, k) + C_L(n, k-1). \quad (41)$$

The diatoms of the same size from generation  $n - 1$ , as well as the diatoms of size index  $k - 1$  from generation  $n$ , contribute to the size index  $k$  in generation  $n + 1$ , since the latter are incremented according to (36). As there are no size indices smaller than zero, we again define  $C_L(n, -1) = 0$ .

To understand equation (41) it is helpful to imagine a 2-dimensional array with square cells and axes for  $n$  and  $k$ . As with a chessboard, color all cells black where  $n + k$  is even and all cells white where  $n + k$  is odd. The cells can now be assigned according to the initial conditions and equation (41). It follows from the recursion that only values from black cells contribute to a black array and only values from white cells contribute to a white array. The system consists of two distinct subsystems. In particular, this means that only the initial conditions from the same subsystem are included in the calculation of the values of a subsystem. In the recursion of the M-model (35), however, each value with  $n > 1$  depends on all initial conditions. These subsystems lead to a case distinction. To derive the solution to (41), consider the black and white marked fields separately. To restrict yourself to the black fields ( $n + k$  even), you can introduce new variables  $p$  and  $q$  by setting  $n = p + q$  and  $k = p - q$ . For the white fields ( $n + k$  odd), choose  $n = p + q + 1$  and  $k = p - q$ . After some transformations you will find the solution (42).

**Theorem 3.** *The solution of the equation (41) for  $n > 1$  reads:*

$$C_L(n, k) = \begin{cases} \binom{\lfloor n+k \rfloor / 2}{k} & \text{for } n+k \text{ even} \\ \binom{\lfloor n+k-1 \rfloor / 2}{k} & \text{for } n+k \text{ odd.} \end{cases} \quad (42)$$

**Proof of Theorem 3.** The proof is done by checking the conditions for the first two generations and inserting them into (41). Let us write (41) in the equivalent form  $C_L(n, k) = C_L(n-2, k) + C_L(n-1, k-1)$  and consider a cell with  $n+k$  even. If we insert (42) into the right-hand side, we obtain

$$\begin{aligned} RHS &= \binom{\lfloor n-2+k \rfloor / 2}{k} + \binom{\lfloor n-1+k-1 \rfloor / 2}{k-1} \\ &= \binom{\lfloor n+k \rfloor / 2 - 1}{k} + \binom{\lfloor n+k \rfloor / 2 - 1}{k-1} \\ &= \binom{\lfloor n+k \rfloor / 2}{k} \end{aligned}$$

This is identical to the left-hand side for  $n+k$  even. In the case of with  $n+k$  odd, the right-hand side gives:

$$\begin{aligned} RHS &= \binom{\lfloor n-2+k-1 \rfloor / 2}{k} + \binom{\lfloor n-1+k-1-1 \rfloor / 2}{k-1} \\ &= \binom{\lfloor n+k-1 \rfloor / 2 - 1}{k} + \binom{\lfloor n+k-1 \rfloor / 2 - 1}{k-1} = \binom{\lfloor n+k-1 \rfloor / 2}{k} \end{aligned}$$

Here, too, one can see the correspondence with the left-hand side for  $n+k$  odd.  $\square$

## 4. Results

### 4.1 Alternatives to long-term observations

When investigating the population dynamics of a diatom species, in particular the question of the validity of the McDonalds-Pfitzer rule arises. If this is fulfilled, which we will assume in the following, a simple division scheme without delay, one of the two models presented, or another scheme could be possible. An obvious option is to use long-term observation, that provides immediate information after a few division processes. This requires a sample that is vital over the observation period. A sufficient supply of nutrients and adequate light exposure must be ensured.

Long-term observations can also provide information without observing the individual division processes. The growth of a culture under consistently good conditions, i.e. in particular without depletion of nutrients and without decreasing exposure to light, can provide an indication of whether exponential or Fibonacci growth is present after measuring the size of the population. Ideally, the culture should be started with a single cell. To measure the population, an automated method using image analysis is recommended. A counting chamber can also be used for this purpose. Alternatively, biomass can be determined if samples are large enough. Whether the species forms chains is not

important. As the Fibonacci sequence grows asymptotically exponentially, differentiation becomes more and more difficult as the number of generations increases. Measurement of the growth rate does not differentiate between the M-model and the L-model.

Other results already presented offer possibilities for analysis, especially on non-living samples and preparations. They therefore do not necessarily require cultivation of the species. The methods are explained below. They mainly refer to chain-shaped colonies. Synchronicity of cell divisions is only to be expected locally in such a chain, as it is lost after a few divisions among its descendants. In the tree structure of the divisions, starting from an initial cell, there are inevitably neighboring cells whose common ancestor dates back so many generations that synchronicity of their divisions is very unlikely. On the other hand, there are many sections that are sufficiently synchronous to allow analysis using the methods described. If we assume that the distribution of doubling times is so narrow that synchronicity over 5 generations is almost always given [18], a chain of at least  $2^5 = 32$  diatoms is created in the P-model, which is likely to be consistent with theory. In one of the models with Fibonacci growth, there are only  $f_7 = 13$  diatoms. Regardless of whether size differences, orientations or sets of twins and sets of triplet structures are examined, a loss of synchronicity must be expected for shorter chains. The shorter the range of a match with a theoretical sequence, the more likely it is that it is a random coincidence and the less informative it is.

#### 4.2 Testing the applicability of the Müller model

A detailed description including the girdle band structure in chain-shaped colonies was only given for the Müller model. Whether this correctly describes the situation in a sample can be checked by placing an existing fragment of a chain in a defined horizontal position and expressing the visible structure as a string using the alphabet  $A_{SOV}$  (3). If we restrict ourselves to the four types of the alphabet and their mirror images, i.e. ignoring the sizes, we obtain a sequence of orientations and thickenings. As it is not known at which point in the generation sequence the observed string occurs for the first time, it must be compared to the theoretical sequence according to (18) for a sufficient length. If there is no match, one must also search for the mirrored string in the theoretical sequence, as the chain has an orientation that is caused by the orientation of the initial sequence. Only a match is relevant, as there can be various reasons why no match is found. The model may be inappropriate, the recording may be incorrect or there may be a loss of synchronization. This statement applies analogously to all comparisons of recorded sequences with theoretical sequences.

The grouping of a chain into sets of twins and sets of triplets is based on the girdle band structure. It is easy to recognize visually. The procedure is analogous, the comparison is based on the alphabet and development as described in section 3.1.9.

#### 4.3 Analysis of sizes and size differences

There is not necessarily a strict relationship between delayed cell division and the occurrence of thickenings, and the girdle band structure as described by Müller. To differentiate the division models without delay with delay, where the smaller or larger daughter cell can divide with delay, it is useful to examine the size sequence. However, it is difficult to assign a size index to measured absolute sizes. The reasons are

- There is natural variation in size even within the initial cell. This means that at best it is possible to give an interval for the size index in which the measured diatom is likely to fall.
- The decrease in mean size per generation is a non-linear function of time. Size, plotted against the size index, decreases most rapidly in large diatoms [33,34]). This makes classification difficult.
- As this data is not systematically collected, it would need to be collected as a first step.

Although it turns out that incorrectly assigned size indices shifted by a fixed number also produce sequences that exist in the theoretical size sequence, an alternative is to examine differences between sizes and size indices. In this case, no uncertain large numbers appear in the sequences, but only the

small differences. Even with this approach, differences in size must be assigned to differences in size indices.

As mentioned above, in the P-model the size indices between neighboring diatoms differ by only  $\pm 1$  [18]. The frequencies of the length differences are then present in two peaks, with the scatter increasing with sample size due to the non-linear dependence of the size on the size index. To avoid errors due to loss of synchronization, it is important to ensure that the division stages are close together. The value range for the M-model additionally contains zero and, in the L-model the values +2 and -2 are included. This makes it possible to distinguish between the three models presented without analyzing the sequence, but it does not allow us to say whether another model describes the colony better. The advantage of the method lies in the simplicity of its procedure. As can be seen from the chains (26) and (40), all possible values occur even in short chains.

If a longer fragment of a chain-like colony with similar division stages is available, a comparison with the difference sequences according to (25), (27) and (39) is recommended. The significance of a match with one of these sequences increases with the length of the matching sequence. Once you have analyzed a fragment and assigned a difference value of the size indices to the size differences of neighboring diatoms, we look for this pattern or its mirror image in the calculated sequences. This sequence is similar to a part of a fingerprint being assigned to a location in a complete image. Since the difference sequence is self-similar, this assignment is not unique. If a sequence appears for the first time in a particular generation, it will also appear in subsequent generations due to the structure of the recursions. In all recursions, the number of fingerprint matches increases with the generation number, as a particular pattern is repeatedly appended and mirrored.

The feasibility of the method was demonstrated on a *Eunotia* sp., probably *Eunotia glacialis* Meister 1912 in [18]. In this species the length differences for adjacent size indices are around 1  $\mu\text{m}$  and are therefore easy to measure. A colony of 25 diatoms was analyzed and found to be consistent with the P-model theory.

#### 4.4 Analysis of the orientations

As shown, the sequences of sizes and orientations can be considered separately. It is useful to use the orientations of a clonal diatom chain as a fingerprint. The advantage of this is that it is not necessary to measure lengths quantitatively, but only to identify and record which diatoms have the larger valve on the left (*l* and *L*) or right (*r* and *R*). Since it must be assumed that it is not possible to distinguish visually between *l* and *L* or *r* and *R*, these characters are treated as identical when comparing the sample and the theoretical sequence. In a P-model, only the letters *L* and *R* exist anyway. It has been mentioned that they have alternating orientations.

#### 4.5 Analysis of the size distribution

Measuring the lengths of many diatoms and classifying them in a histogram with subsequent comparison with theory takes a special place among the possibilities presented:

- One is not dependent on longer chain-like colonies but can also analyze short chains of a few diatom species, as well as diatoms that separate immediately after division.
- With an increasing number of diatoms, i.e. with a high generation index  $n$  in the sample analyzed, synchronism becomes less and less important. The reason for this is that the curves for the functions  $C_P(n, k)$ ,  $C_M(n, k)$  and  $C_L(n, k)$  are similar as a function of  $k$  for closely spaced  $n$  (see below).

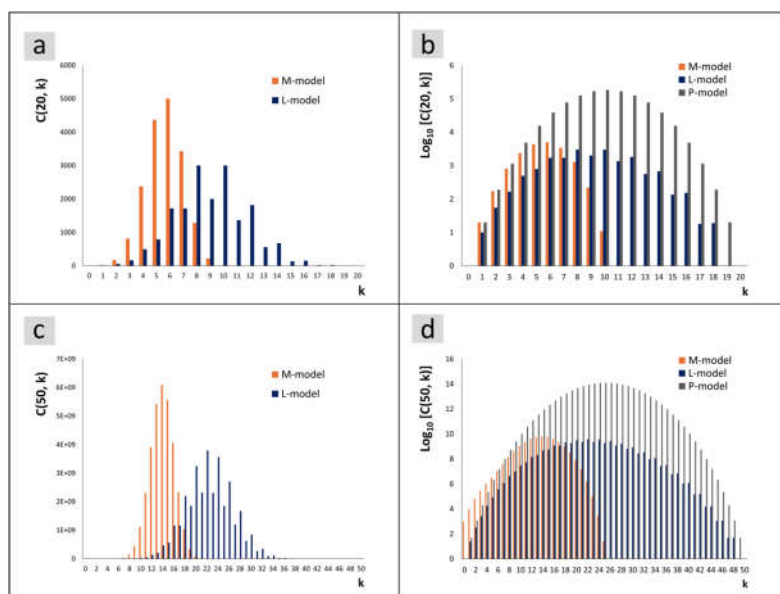
However, there are limitations and challenges:

- A culture started with a single diatom should be used as the basis, otherwise there will be a superposition of several distributions for the respective starting size.
- The theory gives the number of diatoms for a size index, not the size. The comparison of the measurement to the theory therefore requires the assignment of the size to the size index.

To compare the curve shapes, you do not need an exact assignment to a size index, as would be necessary for the comparison with the size sequence of chains, but a function as shown in [33,34]

should be available to equalize the diagrams. If we consider the size indices at a given generation in a model without delay, then the theoretical frequency is given by (2). The binomial coefficients at a generation  $n$ , i.e. at a chosen measurement time, are symmetrical with respect to  $k$ . Apart from the fact that the size index 0 corresponds to the maximum size, this is not equivalent to a histogram of sizes due to the non-linearity mentioned above. However, symmetry is an important distinguishing criterion for the division models. To illustrate the characteristics of the distributions, the number of diatoms in the M-model and L-model are shown in Figure 8a for  $n = 20$  over  $k$ . The two models have the same number of diatoms in each generation, so that they can be visualized at the same scale.

In order to visualize the P-model together with the other models, one could limit oneself to normed values. To emphasize the quantitative difference of the entire population, a logarithmic representation was chosen in Figure 8b for values  $k > 0$  and all three models. The uneven heights in the L-model are characteristic, although they are less noticeable in the logarithmic representation. These discontinuities remain characteristic in later generations. Figure 8c corresponds to Figure 8a but has been calculated for  $n = 50$ . Although the maxima for  $n = 50$  are about 6 orders of magnitude higher than those for  $n = 20$ , the similarity of the bar charts is striking. The corresponding logarithmic representation is shown on the right in Figure 8d.



**Figure 8.** The number of diatoms with size index  $k$  is shown for the models discussed. The number is generally denoted as  $C(n, k)$ . (a) Number of diatoms at  $n = 20$  for the M-model and for the L-model. (b) Decadic logarithm for  $n = 20$  and values  $k > 1$  for all three models. (c) Number of diatoms for  $n = 50$  in the M-model and L-model. (d) Decadic logarithm for  $n = 50$  and values  $k \geq 1$  for all three models.

The position of the maximum in the M-model is clearly shifted towards smaller size indices compared to the other models. However, the delayed formation of the larger daughter cell in the L-model does not lead to a corresponding shift towards larger indices. The maxima of the curves are at slightly lower  $k$  values in the L-model compared to the P-model.

#### 4.6 Maximum size index

As explained at the beginning, no limitation of the size index has been introduced. However, there is a minimum or maximum size index  $k_{max}$  at which cell enlargement becomes necessary. Whether this occurs through sexual reproduction, as is usually the case, or through vegetative reproduction depends on the species. A certain natural range of variation is presumably also given here, so that  $k_{max}$  is not necessarily a well-defined value. In the context of size distributions, it

should merely be noted that a certain size index  $k_{max}$  is reached much earlier with the P-model compared to the M-model. The division behavior not only reduces the reproduction rate but also leads to a higher proportion of diatoms having a large valve size after the same period of development, i.e. the same number of generations. Species that divide according to the M-model significantly delay the onset of cell enlargement. As already discussed with reference to Figure 8, an analogous statement does not apply to the L-model, because the courses of the L-model and P-model are only slightly shifted with respect to each other.

## 5. Discussion

In view of the restriction of this mathematical modeling to the division process, the question arises as to the limitations and significance of the results presented here. Models for the development of diatoms, including their size can have different objectives. Models that are intended to describe the dynamics of the population over a longer period of time must include death processes as well as regularly occurring cell enlargement [13,15,34].

As the focus here is on the possibility of differentiating between the division models, only the growth phase starting from a single diatom is considered. The applicability is limited by the period of Fibonacci growth or, in the case of the P-model, exponential growth. This implicitly means a negligible mortality rate. This statement applies to the study of a fragment of a chain colony as well as to the size distribution of an entire culture.

Most of the concepts presented require a clonal chain-like colony. The need for a sufficient number of synchronous divisions in such samples was pointed out. Looking at the size distribution of a culture that has grown from a single diatom does not have these difficulties, because the characteristics of the distributions are almost independent of the generation. Loss of synchronicity means that the number of generations passed through is not identical but does not change the basic forms of the curves.

If it is not possible to assign a sample to any of the models considered, this may be due to the limitations of the method or its applicability to the sample (e.g. loss of synchronicity), but it could also indicate other possibilities that occur in nature. There is probably no principle that would prevent more than one generation being skipped during cell division, or the larger and smaller daughter cells dividing alternately with a delay. Delays of fractions of a generation time or random processes cannot be ruled out either. For the investigation of so far unknown division rules, direct observation of the processes is the method of choice.

The complexity of the systems studied, in particular the fractal structures, represent emergent properties resulting from the elementary rules of L-systems. They demonstrate the role of fractal structures in nature. However, in the context of the study of occurring patterns, they are less important than locally testable properties, such as the range of values of differences in size indices or shorter sequences of orientations.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The author declares no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

P-model	Cell Division model without delayed divisions
M-model	Cell division model in which the smaller daughter cell divides with a delay.
L-model	Cell division model in which the larger daughter cell divides with a delay.
L-system	Lindenmayer system
LHS	Left Hand Side
RHS	Right Hand Side

## Appendix A

Let  $D_M^n$  be considered for  $n \rightarrow \infty$ . As each sequence always forms the beginning of the following one, an infinite sequence  $D_M$  is created.

**Theorem.** *The sequence  $D_M$  is not periodic. This means that it does not have the form  $D_M = D_a D_p D_p D_p D_p \dots$  with a possibly empty aperiodic string at the beginning, followed by identical strings  $D_p$ .*

**Proof of Theorem.** Assumption:  $D_M$  is periodic.

Consider the number of objects with the values 0 and 1 in the sequence  $D_M$ .

In  $D_a$  let there be  $e_0$  objects with the value 0 and  $e_1$  objects with the value 1.

In  $D_p$  let there be  $c_0$  objects with the value 0 and  $c_1$  objects with the value 1.

A string  $D_a D_p D_p D_p D_p \dots D_p$ , which consists of  $D_a$  and  $j$  following strings  $D_p$ , has, according to these definitions,  $e_0 + j c_0$  objects of type 0 and  $e_1 + j c_1$  objects of type 1. The limit of the ratio of these values is

$$\lim_{j \rightarrow \infty} \frac{e_1 + j c_1}{e_0 + j c_0} = \frac{c_1}{c_0} \in \mathbb{Q}, \quad (\text{A1})$$

i.e. a rational number.

Now we want to derive the ratio of objects 0 and 1 in  $D_M^n$  using equation (25). Let the number of objects in  $D_M^n$  with the value -1 be  $z_{-1}^n$ , the number of objects with the value 0 be  $z_0^n$  and the number of objects with the value 1 be  $z_1^n$ . From (25) follows for  $n > 2$  with these definitions:

$$z_1^{n+1} = z_1^n + z_{-1}^{n-1} + 1 \quad (\text{A2})$$

$$z_{-1}^{n+1} = z_{-1}^n + z_1^{n-1} \quad (\text{A3})$$

$$z_0^{n+1} = z_0^n + z_0^{n-1}. \quad (\text{A4})$$

The first summand on the right-hand side comes from  $D_M^n$ , the second from  $D_M^{n-1}$ . For object 1, the inserted value according to (25) must be considered in (A2). Together with initial values for  $n \leq 2$ , which result from the axiom, (A2), (A3) and (A4) represent explicit calculation rules. It follows directly from (A4), taking the initial values into account

$$z_0^n = f_{n-1}. \quad (\text{A5})$$

By checking the initial values and inserting them into (A2), (A3), you can prove the correctness of the following solutions:

$$z_1^n = f_n, \quad (\text{A6})$$

$$z_{-1}^n = f_n - 1. \quad (\text{A7})$$

The limit for the proportion of values 1 to the proportion of values 0 is therefore

$$\lim_{n \rightarrow \infty} \frac{z_1^n}{z_0^n} = \lim_{n \rightarrow \infty} \frac{f_n}{f_{n-1}} = \frac{1 + \sqrt{5}}{2} \in \mathbb{I}. \quad (\text{A8})$$

The ratio converges to an irrational number, in contradiction to (A1). The periodicity assumption is therefore disproved.  $\square$

## References

1. Round, F.E.; Crawford, R.M.; Mann, D.G. *The diatoms: biology and morphology of the genera*; Cambridge University Press: Cambridge, UK, 1990.
2. Macdonald, J.D.I.—On the structure of the Diatomaceous frustule, and its genetic cycle. *J. Nat. Hist.* 1869, 3(13), 1-8.
3. Pfitzer, E. Über den Bau und die Zellteilung der Diatomeen. *Bot. Ztg.* 1869, 27, 774–776.

4. Locker, F. Beiträge zur Kenntnis des Formwechsels der Diatomeen an Hand von Kulturversuchen. *Oesterr. Bot. Z.* 1950, 97, 322–2.
5. Wiedling, S. Die Gültigkeit der MacDonald-Pfitzerschen Regel bei der Diatomeengattung Nitzschia. *Naturwissenschaften*, 1943, 31. Jg., Nr. 9, 115-115.
6. Edlund, M.B.; Stoermer, E.F. Ecological, evolutionary, and systematic significance of Diatom life histories. *J. Phycol.* 1997, 33, 897-918.
7. Richter, O. *Zur Physiologie der Diatomeen: (II. Mitteilung): die Biologie der Nitzschia putrida Benecke*. Aus der Kaiserlich-Königlichen Hof-und Staatsdruckerei: Wien, Österreich, 1909
8. *Ellerbeckia arenaria* (Moore ex Ralfs). Available online: [https://diatoms.org/species/48715/ellerbeckia\\_arenaria](https://diatoms.org/species/48715/ellerbeckia_arenaria) (accessed on 11 April 2025).
9. Müller, O. Das Gesetz der Zelltheilungsfolge von *Melosira (Orthosira) arenaria* Moore. *Ber. dt. Bot. Ges.* 1883, 1, 35–44.
10. Müller, O. Die Zellhaut, das Gesetz der Zelltheilungsfolge von *Melosira (Orthosira Thwaites) arenaria* Moore. *Jahrb. Wiss Bot.* 1884, 14, 232-291.
11. Schmid, A.M.M.; Crawford, R.M. *Ellerbeckia arenaria* (Bacillariophyceae): formation of auxospores and initial cells. *Eur. J. Phycol.* 2001, 36(4), 307-320.
12. Jewson, D. H. Size reduction, reproductive strategy and the life cycle of a centric diatom. *Philos. Trans. R. Soc. B.*, 1992, 336(1277), 191-213.
13. Ziebarth, J.; Seiler, W.M.; Fuhrmann-Lieker, T. Size-Resolved Modeling of Diatom Populations: Old Findings and New Insights. In *The Mathematical Biology of Diatoms*, Pappas, J.L. Ed.; Wiley: Hoboken, NJ, USA, 2023, pp. 19-61.
14. Laney, S. R., Olson; R. J.; Sosik, H.M. Diatoms favor their younger daughters. *Limnol. Ocean.* 2012, 57(5), 1572-1578.
15. Fuhrmann-Lieker, T.; Kubetschek, N.; Ziebarth, J.; Klassen, R.; Seiler, W. Is the diatom sex clock a clock? *J. R. Soc. Interface*, 2021, 18(179), 20210146.
16. Eppley, R.W. The growth and culture of diatoms. In: *The Biology of Diatoms*, Werner D. Ed.; Blackwell: Oxford, UK, 1977, pp. 24-64.
17. Engelberg, J.; Hirsch, H.R. (1966) On the Theory of Synchronous Cultures. In: *Cell Synchrony*, Cameron I.L., Podilla G.M., Eds.; Academic Press: New York, USA, pp. 14–37
18. Harbich, T. On the Size Sequence of Diatoms in Clonal Chains In: *Diatom Morphogenesis*, Annenkov, V., Seckbach J., Gordon R., Eds.; Wiley: Hoboken, NJ, USA, 2021, pp. 69-92.
19. Harbich, T. (2023). Pattern Formation in *Diatoma vulgare* Colonies: Observations and Description by a Lindenmayer-System. In *The Mathematical Biology of Diatoms*, Pappas, J.L. Ed.; Wiley: Hoboken, NJ, USA, 2023, pp. 265-290.
20. Lewin, J.C.; Reimann, B.E.; Busby, W.F.; Volcani, B.E. Silica shell formation in synchronously dividing diatoms. In *Cell Synchrony*. Cameron, I.L., Podilla G.M., Eds.; Academic Press: New York, USA, 1966; pp. 169–188.
21. Pirson, A.; Lorenzen, H. Synchronized dividing algae. *Annual Rev. Plant Physiol.*, 1966, 17, 439–58
22. Paasche, E. Marine Plankton Algae Grown with Light-Dark Cycles. I. *Coccolithus huxleyi*. *Physiol. Plant.* 1967, 20, 946–956
23. Paasche, E., Marine plankton algae grown with light-dark cycles. 2. *Ditylum brightwellii* and *Nitzschia turgidula*. *Physiol. Plant.* 1968, 21, 66–77.
24. Busby, W.F.; Lewin, J. Silicate uptake and silica shell formation by synchronously dividing cells of the diatom *Navicula pelliculosa* (Breb.) Hilse. *J. Phycol.* 1967, 3, 127–131.
25. Davis, C.O.; Harrison, P.J.; Dugdale, R.C. Continuous culture of marine diatoms under silicate limitation I. Synchronized life cycle of *Skeletonema costatum*. *J. Phycol.* 1973, 9, 175–80.
26. Lindenmayer, A. Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. *J. Theor. Biol.*, 1968, 18(3), 280-299.
27. Lindenmayer, A. Mathematical models for cellular interactions in development: II. Simple and branching filaments with two-sided inputs. *J. Theor. Biol.* 1968, 18, 300–315.
28. De Koster, C.G.; Lindenmayer, A. Discrete and continuous models for heterocyst differentiation in growing filaments of blue-green bacteria. *Acta Biotheor.* 1987, 36, 249–273

29. Prusinkiewicz, P.; Lindenmayer, A. *The Algorithmic Beauty of Plants*; Springer: New York, USA, 1990.
30. Ussing, A.P.; Gordon, R.; Ector, L.; Buczko, K.; Desnitskiy, A.G.; Vanlandingham, S.L. *The colonial diatom "Bacillaria paradoxa": chaotic gliding motility, Lindenmeyer Model of colonial morphogenesis, and bibliography, with translation of O.F. Müller (1783), 'About a peculiar being in the beach-water'*; A.R.G. Gantner Verlag: Ruggell, Liechtenstein, 2005.
31. Gunderson, D.S.; Rosen, K.H. *Handbook of mathematical induction*. Chapman and Hall/CRC: Boca Raton, FL, USA, 2010.
32. Tabachnikov, S. Dragon curves revisited. *Math. Intell.* 2014, 36, 13–17
33. Schwarz, R.; Wolf, M.; Müller, T. A probabilistic model of cell size reduction in *Pseudo-nitzschia delicatissima* (Bacillariophyta). *J. Theor. Biol.* 2009, 258(2), 316–322.
34. Hense, I.; Beckmann, A. A theoretical investigation of the diatom cell size reduction–restitution cycle. *Ecol. Model.* 2015, 317, 66–82.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.