Mathematical Biology

CrossMark

Golden ratio and phyllotaxis, a clear mathematical link

François Bergeron[1] (iD) · **Christophe Reutenauer[1]**

Abstract
Exploiting Markoff's theory for rational approximations of real numbers, we explicitly link how hard it is to approximate a given number to an idealized notion of growth capacity for plants which we express as a modular invariant function depending on this number. Assuming that our growth capacity is biologically relevant, this allows us to explain in a satisfying mathematical way why the golden ratio occurs in nature.

1 Introduction

Among the frequently mentioned mathematical notions that occur in natural phenomena, surely Fibonacci numbers F_n:

$$F_0 = 1, \ F_1 = 2, \ F_2 = 3, \ F_3 = 5, \ F_5 = 8, \ \ldots,$$

with $F_n = F_{n-1} + F_{n-2}$, and the golden ratio $\varphi = (1 + \sqrt{5})/2$, rate close to the very top in the broad public media. It is perhaps both the simplicity of their definition and their ties to beautiful patterns (such as the photo[1] in Fig. 1) that make them especially appealing to a general audience. This fascination for the interplay between Fibonacci numbers and nature apparently goes back at least to Kepler, with some earlier allusions by da Vinci. It is also tantalizing that they are nicely related by the fact that quotient

---

[1] Photo: Richard Sniezko—US Forest Service.

✉ François Bergeron
  Bergeron.Francois@uqam.ca

  Christophe Reutenauer
  Reutenauer.Christophe@uqam.ca

[1] Département de mathématiques, Université du Québec à Montréal, Montreal, Canada

of successive Fibonacci numbers are the "best" rational approximations of the golden ratio

$$\frac{1 + \sqrt{5}}{2} \simeq \frac{F_{n+1}}{F_n},$$

so that their joint story has both aesthetic appeal and some intellectual surprise. In many social occasions, mathematicians are at risk of being asked for an explanation of why the golden ratio and Fibonacci numbers should play such a nice role, often placing them in somewhat of a quandry since some part of the answer must necessarily involve an understanding of some biologico-physical law that underlies the phenomenon considered. Indeed, it stands to reason that some optimization of an advantageous trait must be behind the appearance of the patterns observed, as the cornered mathematician is bound to try to underline. If more knowledgeable about it, he/she may underline that the golden ratio is characterized by the fact that it is one of the hardest (we will explain how below) numbers to approximate by a rational number, and that this must be why it occurs in the alleged optimization involved. Albeit, this is somewhat incomplete since no explicit tie is established between a biological law and the mathematical fact referred to.

Our objective in this paper is to explain how to directly link this notion of "hard to approximate" to one of the abstract models of plant growth considered by some phyllotaxis researchers (see van Iterson 1907; Okabe 2012b). In fact, there is a lot of literature and interesting work pertaining to mathematical aspects of phyllotaxis, and a very nice broad historical overview (but slightly outdated) of the plentiful and varied efforts along these lines may be found in Adler (1974). A more recent compendium of relevant references may be found in Refahi et al. (2016). Noteworthy from our perspective are the more recent work of Atela et al. (2002), Bacher (2014), Couder and Douady (1996a, b, c), Douady (1998), Leigh (1983), Marzec and Kappraff (1983). Atela et al. (2002) is given a rigorous mathematical analysis of a model of plant pattern formation from the point of view of dynamical systems, explaining the occurrence of Fibonacci numbers in terms of fixed points and bifurcation patterns. Notwithstanding

this, we could not find in the literature a truly satisfying direct mathematical link between the hard to approximate property of the golden ration and some abstract mode of growth of plants, with a precise mathematical formulation of the nature of this direct tie. This work does propose such a formulation, but we make no claim that our model has been validated from the point of view of Biology. We leave this to be checked by the experts in the field. As discussed in Refahi et al. (2016) it seems that "stochastics models" may be given a better biological understanding, but our point of view is mathematical. As far as we could find in the literature, our use of hyperbolic geometry together with Markov's theory to explain the appearance of the golden ratio in phyllotaxis is original. However, our end result (Theorem 1) for the golden ratio is already present in work of Ridley (1986, Thm 4), without our more general framework that allows for a deeper understanding.

We start by recalling how the notion of hard to approximate by a rational number has been beautifully developed by Markoff[2] in two seminal papers (Markoff 1879, 1880) that appeared in 1879 and 1880. His theory is nicely presented in a recent book of Aigner (2013), where more details may be found. Following Markoff's tack, we associate to each irrational number x its Lagrange number, denoted $L(x)$. This is the supremum of the set of real numbers L such that there are infinitely many rational approximations p/q of x for which we have the inequality

$$\left| x - \frac{p}{q} \right| < \frac{1}{L\, q^2}.$$

Part of Markoff's theory says that $L(x) = \sqrt{5}$, if x is equivalent[3] to the golden ratio; and that $L(x) \geq \sqrt{8}$ for any other real number. In other words, any number x, not equivalent to the golden ratio, affords infinitely many rational approximations for which

$$\left| x - \frac{p}{q} \right| < \frac{1}{\sqrt{8}\, q^2},$$

whereas this is not so for the golden ratio. It is in this precise sense that the golden ratio (and its equivalents) is considered hardest to approximate. Markoff's theory goes on to give a very nice filtration of real numbers with respect to how easier they become to approximate, once some relevant subsets are removed. He shows that there is a sequence of Lagrange numbers L_n, generalizing $\sqrt{5}$ and $\sqrt{8}$ above, of the form

$$L_n = \sqrt{9 - \frac{4}{m_n^2}},$$

with the m_n's integers that are now called Markoff (or Markov) numbers. The first Markoff numbers are

---

[2] This is the same Markov as in the well-known Markov chains theory; who used this surname spelling in his French publications.

[3] Here, a number is considered to be **equivalent** to the golden ratio if its continued fraction expansion only contains 1 after a certain rank.

Fig. 2 Cylindric plant
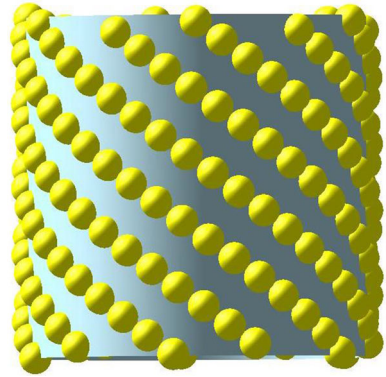
$$1, 2, 5, 13, 29, 34, 89, 169, 194, 233, 433, 610, 985, 1325, \ldots$$

To each Lagrange number (< 3), there corresponds a finite number of explicit families of numbers (all having the same continued fraction expansion after some rank, for a given family) to be excluded, so that all other numbers satisfy the inequality

$$\left| x - \frac{p}{q} \right| < \frac{1}{L_n q^2}.$$

In trying to understand how to tie the hard to approximate property of the golden ratio to plant growth, we consider the following model. The "plant" is considered to be cylindrical, with buds growing successively on an upward helix at regular intervals (see Figs. 2 and 3). The horizontal length of these intervals is measured by the **divergence** x in terms of the "angle" between two successive buds. This is expressed as a proportion of a complete turn (expressed in radians), with the actual angle equal to $2x\pi$. It is often informally stated that for best plant growth, x must be not only irrational but in fact an irrational that is hardest as possible to approximate. Our purpose here is to exploit Markoff theory to justify this last statement making use of a model suggested by van Iterson (1907, page 24) [and used in Ridley (1986)] that suggests what one could consider as an optimization parameter in plant growth. More explicitly, we consider a specific function $f_x(y)$ that measures how "good" a growth scheme is, with given divergence x, where y corresponds to varying height differences between successive buds. We show that $f_x(y)$ is "globally optimal" (that is for all y) if and only if x is equivalent to the golden ratio. From a mathematical perspective, the function $f(x, y) = f_x(y)$ is both sound and with elegant properties. Noteworthy among these is the fact that it is invariant under the Modular Group, when considered as a function of the complex number $x + iy$, linking the problem to an hyperbolic geometry point of view. In fact this plays a key role in the proof of our main result.

Further interesting mathematical work related to phyllotaxis may be found in the work of Adler (1974), Atela et al. (2002), Coxeter (1972), Leigh (1983), Marzec and
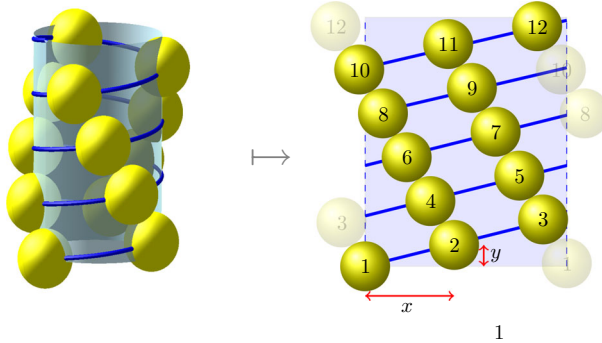
Fig. 3 Buds on a cylindrical trunc, and unfolded version

Kappraff (1983), Okabe (2012a, b); as well as in the papers collected in *Symmetry in Plants* (Jean and Barabé 1998).

2 A mathematical model based on the area around a bud

As sketched above, we consider a spiral **growth scheme** on the cylinder to be specified by the pair of numbers (x, y), with x the divergence angle between successive buds, and y the height difference between these buds, as illustrated in Fig. 3. To introduce a measure of how good a growth scheme (x, y) is, Iterson suggested that one should surround each bud by the largest-area disk (pictured as spheres in Fig. 3, only for aesthetic reasons) so that no two disks overlap. Thus the diameter of circles is the shortest possible distance between two buds. Heuristically put, one considers here that an optimal growth scheme for a plant would be to aim at sprouting the maximal number of buds with a minimal use of resources (here measured by disk-covering-area). Hence, for a given growth scheme, the proportion of area of the trunk covered by the aforementioned disks is considered to measure how capacious the growth scheme is.

Unfolding the cylinder (and periodically repeating horizontally the pattern of buds) we get a lattice $\mathcal{L}_{xy}$ in the plane which is "generated" by the vectors $(1, 0)$ (implicitly assuming that the circumference of the cylinder is equal to 1), and (x, y). More explicitly, we have

$$\mathcal{L}_{xy} := \{\alpha\,(1, 0) + \beta(x, y) \mid \alpha, \beta \in \mathbb{Z}\},$$

with buds placed at each points of $\mathcal{L}_{xy}$.

Following Iterson, as mentioned above, we surround each point of $\mathcal{L}_{xy}$ by a disk whose diameter $d = d(x, y)$ is the smallest distance between two points of the lattice. The parallelogram with sides u and v (for any basis u, v of $\mathcal{L}_{xy}$) is said to be a **fundamental region** for the lattice, and $\mathbb{R} \times \mathbb{R}$ is tiled by $\mathcal{L}_{xy}$ translates of this fundamental region. The area of said region is given by the absolute value of the determinant whose row are the vectors u and v. It is easy to see that this is equal to

Fig. 4 Disk inscribed in the fundamental region
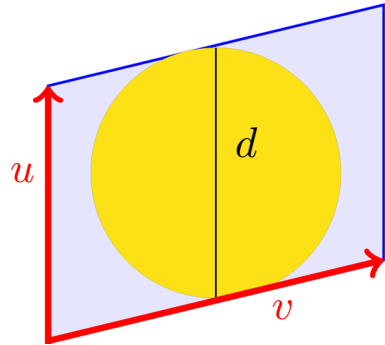
y. Indeed, this area does not depend on the choice of basis, hence we may choose the basis $\{(1, 0), (x, y)\}$, and calculate the area as being

$$\det \begin{pmatrix} 1 & 0 \\ x & y \end{pmatrix} = y.$$

Up to a translation we may assume that the disks originally surrounding each point of $\mathcal{L}_{xy}$ are drawn with center in the middle of each of the translates of the fundamental region, as illustrated in Fig. 4. Thus the measure how well the disks cover the plane corresponds to the ratio of area of one of the disks (of radius $d(x, y)/2$) with respect to the area of one copy of the fundamental region, in formula this gives $\pi \, d(x, y)^2/(4 \, y)$.

Simplifying by a scalar multiple, we define the measure of "capacity" of a growth scheme as the quotient $d(x, y)^2/y$, considering as above that this capacity is directly correlated to the proportion of area covered by disks. For a fixed divergence x, we will study the behavior of the function $y \mapsto d^2/y$ and show, using Markoff theory, that the upper limit of the minima of this function is largest when x is the golden ratio or an equivalent number.

3 Growth capacity is invariant under the modular group

Let us first straightforwardly reformulate our construction above in terms of Poincaré's half-plane model of hyperbolic geometry, and its completion:

$$\mathbb{H} := \{\omega \in \mathbb{C} \; ; \; \text{Im}(\omega) > 0\}, \quad \text{and} \quad \overline{\mathbb{H}} := \mathbb{H} \cup \mathbb{R} \cup \{\infty\}.$$

Each point (x, y) (with $y > 0$) is considered here as the point $\omega := x + iy$ in $\mathbb{H}$. In this manner, we will consider points of $\mathbb{H}$ as encoding growth schemes. To each such growth scheme $\omega \in \mathbb{H}$, we associate the lattice $\mathcal{L}_\omega := \mathbb{Z} + \mathbb{Z}\omega$. This is the additive subgroup of $\mathbb{C}$ generated by 1 and ω; and $d(\omega)$ is the minimal distance between two points of this lattice. Just as in our previous formulation, we have

$$d(\omega) = \min\{ \, |\alpha + \beta \, \omega| \; ; \; \alpha, \beta \in \mathbb{Z}, (\alpha, \beta) \neq (0, 0)\}.$$
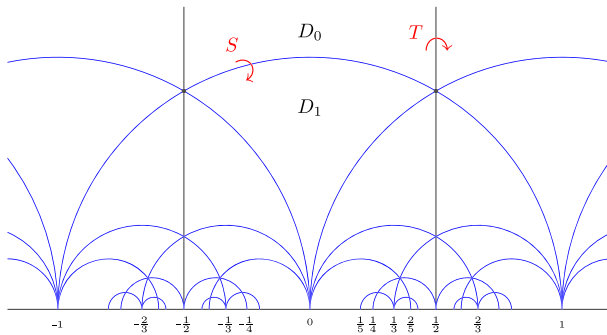
Fig. 5 Tiling of hyperbolic plane

Following our discussion of the previous section, we reformulate the **growth capacity** function $f : \mathbb{H} \to \mathbb{R}$ as

$$f(\omega) := \frac{d(\omega)^2}{\text{Im}(\omega)}. \tag{1}$$

It may very well be that this function has already been considered, together with Proposition 1 below, but we could not find its trace in the literature.

We first recall basic facts about the action of the modular group $\text{PSL}_2(\mathbb{Z})$ on $\overline{\mathbb{H}}$. Elements g of $\text{PSL}_2(\mathbb{Z})$ are 2×2 matrices of determinant 1 with coefficients in $\mathbb{Z}$, with g identified with $-g$. The action $\text{PSL}_2(\mathbb{Z}) \times \overline{\mathbb{H}} \to \overline{\mathbb{H}}$ is defined as

$$g \cdot \omega = \frac{a\,\omega + b}{c\,\omega + d}, \qquad \text{for} \qquad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{PSL}_2(\mathbb{Z}),$$

with $g \cdot \infty := a/c$ and $g \cdot (-d/c) = \infty$, when $c \neq 0$; and $g \cdot \infty := \infty$ otherwise. As is well-known, the modular group is generated[4] by the two functions $T : \omega \mapsto \omega + 1$, and $S : \omega \mapsto -1/\omega$, with relations

$$S^2 = \text{Id}, \qquad \text{and} \qquad (ST)^3 = \text{Id}.$$

A very classical decomposition of the space $\mathbb{H}$, with respect to this action of the modular group, is obtained by considering all images under group elements of the fundamental region

$$D_0 = \{\omega \in \mathbb{C} \,;\; -1/2 \leq \text{Re}(\omega) \leq 1/2, \quad \text{and} \quad |\omega| \geq 1\}.$$

This results is a tiling of $\mathbb{H}$, partly shown in Fig. 5, with D_1 being the image of D_0 under S (which sends ∞ to 0).

Proposition 1 *The function f is invariant under the modular group* $\text{PSL}_2(\mathbb{Z})$*, that is* $f(g \cdot \omega) = f(\omega)$ *for all* $g \in SL_2(\mathbb{Z})$ *and* $\omega \in \mathbb{H}$.

---

[4] See for instance Serre (1970), Theorem 2 of chapter VII.

Proof It is clearly sufficient to show that f is invariant for T and S. It is evident in the first case, since the lattice generated by 1 and ω coincides with the lattice generated by 1 and $\omega + 1$ on one hand; and on the other because the imaginary parts of ω and $\omega + 1$ are equal. The second case proceeds as follows. Observe first that elements of the lattice $\mathcal{L}(-1/\omega)$ may be written as multiples of $1/\omega$ by elements of $\mathcal{L}(\omega)$:

$$\alpha + \beta \left(\frac{-1}{\omega} \right) = \frac{1}{\omega} (\alpha \, \omega - \beta).$$

Hence, the module of $\alpha + \beta \, (-1/\omega)$ is equal to that of $\alpha \omega - \beta$ (which lies in $\mathcal{L}(\omega)$) divided by $|\omega|$. Since this links all elements of $\mathcal{L}(-1/\omega)$ to a corresponding element of $\mathcal{L}(\omega)$, it follows that $d(-1/\omega) = d(\omega)/|\omega|$. On the other hand,

$$\mathrm{Im} \left(\frac{-1}{\omega} \right) = \mathrm{Im} \left(\frac{-\bar{\omega}}{\omega \bar{\omega}} \right) = \frac{\mathrm{Im}(\omega)}{|\omega|^2}.$$

Thus

$$f \left(\frac{-1}{\omega} \right) = \frac{d(\omega)^2/|\omega|^2}{\mathrm{Im}(\omega)/|\omega|^2} = f(\omega),$$

which concludes the proof.   □

Proposition 2 *If ω lies in D_0 or any of its horizontal translates $D_0 + n = T^n(D_0)$, for $n \in \mathbb{Z}$, then $f(\omega) = 1/\mathrm{Im}(\omega)$.*

Proof For $\omega = x + iy \in D_0$, elements of the lattice $\mathbb{Z} + \mathbb{Z}\omega$ are of the form $\alpha + \beta \omega = \alpha + \beta \, x + i\beta \, y$, and

$$\begin{aligned} |\alpha + \beta\omega|^2 &= (\alpha + \beta \, x)^2 + (\beta \, y)^2 \\ &= \alpha^2 + 2x \, \alpha\beta + (x^2 + y^2)\beta^2. \end{aligned}$$

Note that $2 \, |x| \leq 1$, so that $(-2) \, |x| \geq -1$, and we get

$$\begin{aligned} \alpha^2 + 2x \, \alpha\beta + (x^2 + y^2)\beta^2 &\geq \alpha^2 + (x^2 + y^2)\beta^2 - 2|x \, \alpha\beta| \\ &\geq \alpha^2 - |\alpha\beta| + \beta^2 \\ &= |\alpha|^2 - |\alpha||\beta| + |\beta|^2. \end{aligned}$$

For α and β in $\mathbb{Z}$, the quadratic form $\alpha^2 - \alpha\beta + \beta^2$ only takes positive integral values, since its discriminant is -3. Its minimum value, for α, β not both 0, is thus 1. It follows that the minimum value of $|\alpha + \beta\omega|^2$, under the same conditions for α, β, is also 1. Thus we have shown that $d(\omega)^2 = 1$, and we get the announced formula for $f(\omega)$ in this case. When $\omega \in n + D_0$, the result also holds since both f and the imaginary part of ω are invariant under horizontal translations. This completes our proof.   □

The previous result implies that f is bounded above by $2/\sqrt{3}$, since this is the maximal value of f in the fundamental domain D_0.

Fig. 6 Cusp of triangle

p/q
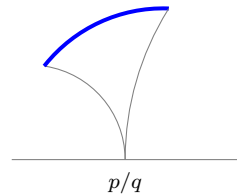
Corollary 1 *The function f is continuous.*

Proof Clearly the restriction of f to D_0 is continuous. For g in the modular group, the restriction of f to $g \cdot D_0$ is also continuous, since this restriction maps $\omega \in g \cdot D_0$ to

$$f(\omega) = f(g^{-1}\omega) = \frac{1}{\text{Im}(g^{-1}\omega)} \tag{2}$$

in view of the invariance of f under the modular group, and by Proposition 2, knowing that $g^{-1} \cdot \omega \in D_0$. But g^{-1} is continuous, hence f is continuous on $g D_0$. We know that $\mathbb{H}$ is the union of the $g D_0$, for g running over $\text{PSL}_2(\mathbb{Z})$ [see Serre (1970) Theorem 1 of Chapter VII]. Moreover, at most three of these images contain any given point (see Fig. 5). It follows that f is continuous at these finitely covered points, and f is continuous everywhere. Thus showing the overall assertion. □

4 Geometrical interpretation of growth capacity

Let us now consider how f behaves for $\omega = x + iy \in \mathbb{H}$, with x fixed. Proposition 2 takes care of all cases when $y > 1$ (at least), and the interesting behavior is thus when y becomes smaller and smaller. To better see this, we consider $y = 1/t$, hence the function that sends t to $f(x + i/t)$. Figure 10, illustrates how this function behaves for some fixed x. Once again we consider the tiling of $\mathbb{H}$ made out of the regions $g \cdot D_0$. Each of these is an **hyperbolic triangle**, with exactly one of its vertices in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ (the regions $n + D_0$ are those for which this vertex is at ∞). This special vertex is said to be the **cusp** of the triangle and, except for the cases $n + D_0$, it is located at some rational number p/q. The **base** of the triangle is the edge opposite to the cusp. See Fig. 6 above for an illustration of such a triangle and its cusp, with the base of the triangle thickly drawn (as is also the case in upcoming figures).

Exploiting the propositions of the previous section, we may give an elegant geometrical interpretation of the function $f(\omega)$. Indeed, it follows from Proposition 2 that $f(\omega)$ is constant along an horizontal line $\text{Im}(\omega) = 1/d$, for $d \leq 1$, since the line is then entirely contained in the translates $D_0 + n$ for $n \in \mathbb{Z}$. By general principles of inversive geometry, the image of this line under the modular group transformation

$$\omega \mapsto \frac{p\,\omega + p'}{q\,\omega + q'}, \quad \text{for} \quad g = \begin{pmatrix} p & p' \\ q & q' \end{pmatrix} \in \text{PSL}_2(\mathbb{Z}),$$

is a circle tangent to the real axis at $p/q = g \cdot \infty$. Its radius is equal to $r = d/(2q^2)$, and hence its center is $p/q + i\,r$. Indeed, we have

$$g \cdot (x + i/d) = \frac{(px + p')\,(qx + q')\,d^2 + pq}{(qx + q')^2\,d^2 + q^2} + i\,\frac{d}{(qx + q')^2\,d^2 + q^2}$$

which evaluates to $p/q + i\,d/q^2$ at $x = -q'/q$. Since this is the point diametrically opposed to p/q, perforce the diameter of the circle is its y-coordinate, hence our formula.

On the other hand, from Proposition 2 we deduce that

$$f(x + i/t) = (x\,q - p)^2\,t + q^2/t. \tag{3}$$

by applying (2) to $\omega = x + i/t$ in $g \cdot D_0$, using $pq' - qp' = 1$, via the calculation

$$\begin{aligned}
f(x + i/t) &= \frac{1}{\mathrm{Im}(g^{-1}(\omega))} \\
&= \mathrm{Im}\left(\frac{q'\,\omega - p'}{-q\,\omega + p}\right)^{-1} \\
&= \mathrm{Im}\left(\frac{(q'\,\omega - p')(-q\,\overline{\omega} + p)}{(-q\,\omega + p)(-q\,\overline{\omega} + p)}\right)^{-1} \\
&= \mathrm{Im}\left(\frac{(-q'\,q\,x^2 + x + (p\,p' - q\,q'/t^2) + i/t}{(x\,q - p)^2 + q^2/t^2}\right)^{-1} \\
&= (q\,x - p)^2\,t + q^2/t,
\end{aligned}$$

as announced. As it happens, this last right-hand side affords the following simple geometrical interpretation.

Proposition 3 *For any $\omega \in \mathbb{H}$, let $g \in \mathrm{PSL}_2(\mathbb{Z})$ be such that $g^{-1} \cdot \omega$ lies in some $D_0 + n$ (for $n \in \mathbb{Z}$), and let $p/q := g \cdot \infty$. Then, the value of the function f at ω is equal to $d\,q^2$, with d being the diameter of the circle which is tangent to the real axis at p/q, and which passes through the point ω.*

Proof In terms of real coordinates, the equation of the circle considered (represented in Fig. 7) is $(x - p/q)^2 + (y - r)^2 - r^2 = 0$. Multiplying both sides by q^2/y, this may be written as $(q\,x - p)^2/y + q^2 y - d\,q^2 = 0$, with $d = 2\,r$. Thus, with $y = 1/t$, we get

$$d\,q^2 = (q\,x - p)^2 t + q^2/t = f(x + i/t),$$

thus showing our assertion. □

The following proposition will help us tie our growth capacity measure to how well or not the number x may be approximated by a rational number. To this end, we first clarify the domain on which formula (3) applies, for a fixed value of x. Since the

Fig. 7 Interpretation of $f(\omega)$ as a radius (up to a scalar multiple)
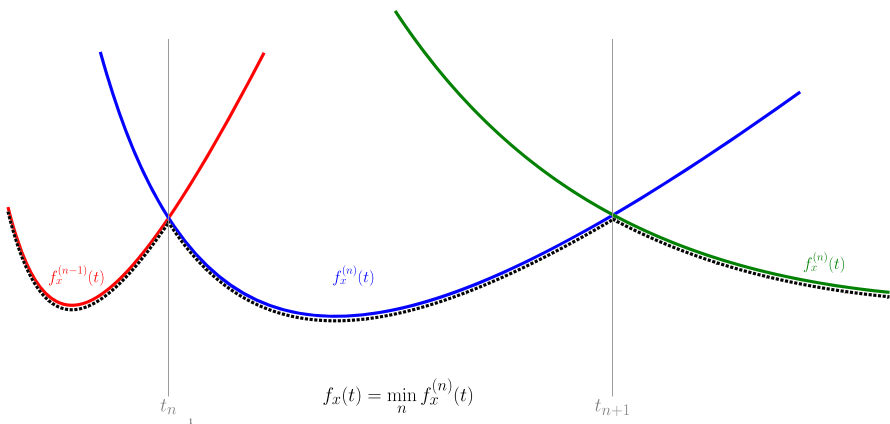

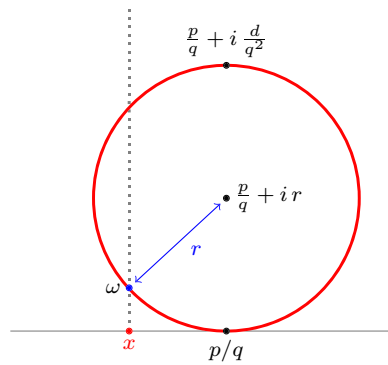

$$f_x(t) = \min_n f_x^{(n)}(t)$$

Fig. 8 The function f_x is obtained by gluing pieces of successive functions $f_x^{(n)}(t)$

right-hand side of (3) is a smooth convex function of t, and f is globaly continuous, it results that

$$f_x := t \mapsto f(x + i/t)$$

is a piecewise smooth convex function between some local maxima, where it is not derivable. More precisely, we have an increasing sequence of real numbers $t_n = t_n(x)$

$$t_1 < t_2 < \cdots < t_{n-1} < t_n < \cdots$$

such that the function f_x is (locally) given by the formula

$$f_x^{(n)}(t) := (x\, q_n - p_n)^2\, t + q_n^2/t. \tag{4}$$

This is to say that $f_x(t) = f_x^{(n)}(t)$, when $t_n \leq t \leq t_{n+1}$. Observe that $f_x^{(n)}(t)$ makes sense for all $t > 0$, and Fig. 8 illustrates how f_x is obtained by gluing pieces of successive functions $f_x^{(n)}(t)$, for increasing values of n. We will see later that p_n/q_n is the nth Hermite convergent of x. This will imply that

$$f_x^{(n-1)}(t) < f_x^{(n)}(t), \quad \text{if} \quad t < t_n,$$

and

$$f_x^{(n-1)}(t) > f_x^{(n)}(t), \quad \text{if} \quad t > t_n;$$

hence t_n is a local maximum of f_x. We may thus write

$$f_x(t) = \min_n f_x^{(n)}(t),$$

with the minimum taken over n, for any fixed t. Continuity of f forces $f_x^{(n)}$ to agree with $f_x^{(n-1)}$ at $t_n = t_n(x)$, hence

$$(q_n x - p_n)^2 t_n + q_n^2/t_n = (q_{n-1} x - p_{n-1})^2 t_n + q_{n-1}^2/t_n.$$

Solving this equality for t_n gives

$$t_n := \sqrt{\frac{q_n^2 - q_{n-1}^2}{(q_{n-1} x - p_{n-1})^2 - (x q_n - p_n)^2}}, \tag{5}$$

and we may then calculate directly that

$$f_x(t_n) = t_n \frac{p_n/q_n + p_{n-1}/q_{n-1} - 2 x}{q_n/q_{n-1} - q_{n-1}/q_n}.$$

Proposition 4 *The local minima of the function f_x, from $\mathbb{R}^*$ to $\mathbb{R}$, are the numbers $2 |q_n(q_n x - p_n)|$, and these are achieved at $t_0 = |q_n/(q_n x - p_n)|$.*

Proof Assume that f_x is given by formula (4) in the segment $t_n \leq t \leq t_{n+1}$, and observing that this is a convex function, therefore the minimum occurs when

$$\frac{d}{dt} f_x^{(n)} = -q_n^2/t^2 + (q_n x - p_n)^2 = 0,$$

hence when t is equal to $t_0 := |q_n/(q_n x - p_n)|$, and the corresponding value

$$f_x^{(n)}(t_0) = 2 |q_n(q_n x - p_n)|$$

is the announced minimum.   □

Geometrically, this minimum occurs when the circle of Proposition 3 is tangent to the vertical line whose points have real part equal to x.
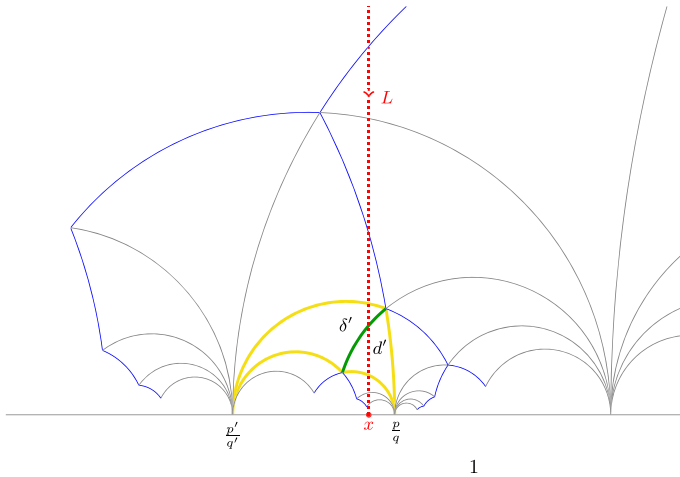
Fig. 9 Traveling down from infinity along the line $\mathrm{Re}(\omega) = x$

5 Global behavior of growth capacity

We will now see that the global behavior of f_x may be revealed using interesting properties of Hermite's *approximation theory* (Hermite 1916) for real numbers. We will exploit this to understand what singles out the golden ratio as a champion from the point of view of the associated growth capacity function.

To this end, we borrow on Humbert's approach [see Humbert (1916)] to Hermite's theory. Consider a point traveling down a vertical hyperbolic line of abscissa x, going from ∞ to 0. In other words, these are the points of the form $x + i/t$, with t going from 0 to ∞. The point successively traverses hyperbolic triangles $g \cdot D_0$ as illustrated in Fig. 9, whose cusps (at $p/q \neq \infty$) are by definition (Humbert 1916, page 82 , or Jacobs 2014) **Hermite convergents** of x. These convergents satisfy

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{\sqrt{3}\, q^2},$$

and the two vertices lying on the base of the hyperbolic triangles in question are on the (real plane) circle of equation

$$(x - p/q)^2 + (y - r)^2 = r^2, \quad \text{with} \quad r = \frac{1}{\sqrt{3}\, q^2}. \tag{6}$$

Let $x = [a_0, a_a, a_2, \ldots]$ be the continued fraction expansion of x (positive), with $a_i \in \mathbb{N}$. Recall that its (classical) **convergents** are the rational numbers
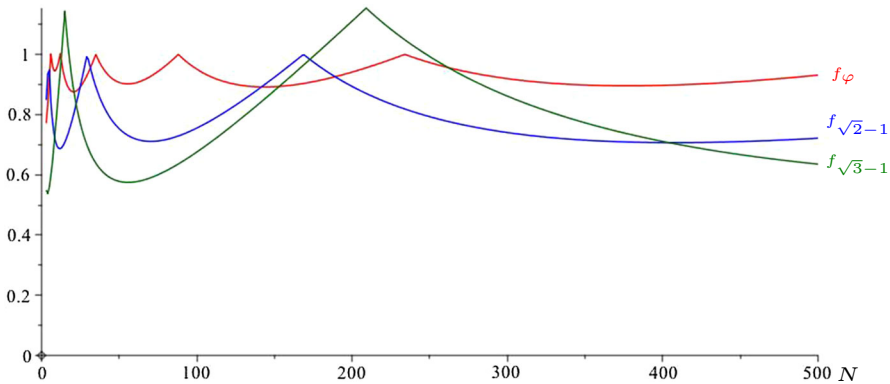
Fig. 10 Three growth capacity functions

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}.$$

for $n \in \mathbb{N}$. Successive Hermite convergents appear as a subsequence of the sequence of all classical convergents of x, see Humbert (1916, page 95). One property that characterizes some of the Hermite convergents of x goes as follows. If $a_{n+1} \geq 2$, then $[a_0, a_1, \ldots, a_n]$ is an Hermite convergent of x, see Humbert (1916, page 96). Moreover, if p'/q' and p/q are two consecutive Hermite convergents, then $p'q - pq' = \pm 1$, see Humbert (1916, page 84). This is not an exhaustive set of property if one intends to characterized the p_n/q_n, and we refer to *loc. cit.* for the necessary details. Just to illustrate, the first convergents of $\sqrt{7} - 1$ are:

$$\frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{23}{14}, \frac{28}{17}, \frac{51}{31}, \frac{79}{48}, \frac{367}{223}, \frac{446}{271}, \frac{813}{494}, \ldots$$

whereas among these the only Hermite convergents are:

$$\frac{2}{1}, \frac{5}{3}, \frac{28}{17}, \frac{79}{48}, \text{ and } \frac{446}{271}.$$

Figure 10 illustrates that a local minimum occurs in each of the region associated to an Hermite convergent [in which $f_x(t)$ may be calculated using formula (3)], with three different values of x.

We now want to establish that the golden ratio (and equivalent numbers) gives the best growth scheme. As before, for a real x, let $x = [a_0, a_1, a_2, \ldots]$ be its continued fraction expansion. We assume that this expansion is infinite, which is to say that x is

irrational. Once again denote by $[a_0, \ldots, a_n] = p_n/q_n$ its n-th convergent, expressed as an irreducible fraction. It is well known [Adler 1974, (1.15) section 1.4] that for all $n \geq 1$, we have

$$\left| x - \frac{p_n}{q_n} \right| = \frac{1}{\lambda_n(x) \, q_n^2},$$

with $\lambda_n(x) = [a_n, \ldots, a_1]^{-1} + [a_{n+1}, a_{n+2}, \ldots]$. Equivalently, $|q_n(q_n \, x - p_n)| = 1/\lambda_n(x)$. Moreover, the upper limit of the $\lambda_n(x)$, as n goes to ∞, is precisely the **Lagrange number** of x mentioned earlier, and it is denoted by $L(x)$, see Aigner [2013, (1.15), Proposition 1.22 and Definition 1.7]. From Markoff's theory, we know that $L(x) = \sqrt{5}$ for x equal to the golden ratio, or equal to any number whose continued fraction expansion contains only 1 starting from some rank. For any other number, $L(x) \geq \sqrt{8}$ (*loc.cit.*). From this we get the following, after proving an auxiliary lemma.

Theorem 1 *If x is equal to the golden ratio, or to any number whose continued fraction expansion contains only 1 starting from some rank, then the lower limit of the minima of its growth capacity function is $2/\sqrt{5}$. For any other number x, this limit is $\leq 2/\sqrt{8}$.*

For a given x, let us denote by $H(x)$ the subset of integers n such that p_n/q_n is an Hermite convergents for x. For instance, for $x = \sqrt{7} - 1$, we have

$$H(x) = \{0, 2, 4, 6, 8, \ldots\}.$$

Lemma 1 *The upper limit, as n goes to infinity, of the sequence of all $\lambda_n(x)$, for $n \in N$, is equal to the upper limit of the subsequence $(\lambda_n(x))_{n \in H(x)}$.*

Proof Let p/q be an irreducible fraction, with $q > 0$. Let us set $u = \varepsilon \, q \, (p - q \, x)$ where $\varepsilon = \pm 1$ is chosen so that u is positive. Consider q' the unique integer solution of $p \, q' \equiv \varepsilon \mod q$ with $0 \leq q' < q$. Let p' be such that $p \, q' = \varepsilon + q \, p'$. Then p/q is an Hermite convergent for x if and only if

$$u < \frac{q(q + 2q')}{2(q^2 + qq' + q'^2)},$$

see Humbert (1916, page 95). Observe that, since $q > q'$, we have

$$\frac{q(q + 2q')}{2(q^2 + q \, q' + q'^2)} > \frac{q(q + 2q')}{2(q^2 + q \, q' + q \, q')} = \frac{1}{2}.$$

It follows that a convergent p_n/q_n which is not an Hermite convergent, must be such that $|q_n(q_n \, x - p_n)| = u > 1/2$, and hence $\lambda_n(x) < 2$. Since the upper limit of $\lambda_n(x)$ is greater or equal to $\sqrt{5}$ (> 2), it follows that this limit does not change if we restrict n to be such that p_n/q_n is an Hermite convergent, that is $n \in H(x)$. □

We can now prove the theorem as follows.

Proof (of Theorem 1) Proposition 4 says that the minima are of the form $2|q(q\,x - p)|$, where p/q is an Hermite convergent for x. This Hermite convergent occurs as one of the convergents of the continued fraction of x, say $p/q = p_n/q_n$. By the above formula, this minimum is of the form $2/\lambda_n(x)$, where n is the rank of an Hermite convergent, i.e.: $n \in H(x)$. By the lemma, the lower limit of these numbers is $2/L(x)$, and the corollary follows. $\qquad\square$

6 Further considerations

As we have seen, in instances where growth capacity could be considered to be a good measure from the point of view of phyllotaxis, it gives a clear mathematical indication why one should so often encounter the golden ratio. The theory considered here also suggests that if other growth schemes could occur in exceptional (or extraterrestrial!) instances, then the next most frequent such growth schemes would be tied to the number $1 + \sqrt{2}$ (and equivalents); with variants of the Pell numbers, P_n,

$$1, 2, 5, 12, 29, 70, 169, 408, 985, 2378, 5741, 13860, 33461, 80782, \ldots$$

replacing the Fibonacci numbers (and their own variants). After that would come, in rarer and rarer instances, growth schemes associated to the numbers

$$\frac{11 + \sqrt{221}}{10}, \ \frac{29 + \sqrt{1517}}{26}, \ \ldots$$

For more on this from the point of view of Markoff theory, see Reutenauer (2018, Section 10.2).

Our explanation of the optimality of the golden ratio may be seen to be even more plausible if one considers the average

$$g_x := \limsup_{n \to \infty} \overline{f}_x(n), \qquad \text{with} \qquad \overline{f}_x(n) := \frac{1}{t_{n+1} - t_n} \int_{t_n}^{t_{n+1}} f_x^{(n)}(t)\, dt, \quad (7)$$

as a comparison tool between growth schemes. Rather than only whining from the point of view of a local behavior of minima, g_x gives a global measure that may be even more significant from the biological point of view. For the golden ratio φ, we observe that $2/\sqrt{5} \approx 0.89443 \ (< g_\varphi)$ is an upper bound for g_x, for all x not equivalent to φ. More technically, it may be shown (see "Appendix") that

$$g_\varphi = \frac{1}{2} + \frac{2}{\sqrt{5}} \log(\varphi)$$
$$\approx 0.93041,$$

and that $g_x < g_\varphi$ for all number x not equivalent to φ. For instance,

$$g_{(1+\sqrt{2})} = \frac{1}{2} + \frac{1}{\sqrt{8}}\log(1+\sqrt{2})$$
$$\approx 0.81161.$$

Appendix

We calculate here the limit of the averaging integral (7) in the case of the golden ratio $x = \varphi$, whose Hermite's convergents are the quotients F_{n+1}/F_n. The required calculation is greatly simplified using the simplified expression

$$f_\varphi^{(n)}(t) = \varphi^{-2(n+1)}t + F_n^2\, t. \tag{8}$$

for the function $f_\varphi^{(n)}$, which follows from the fact that $(F_n - F_{n+1}\,\varphi)^2 = \varphi^{-2(n+1)}$. The corresponding minimum occurs at $F_n\,\varphi^{n+1}$, and takes the value

$$2F_n/\varphi^{n+1} \approx 2/\sqrt{5}.$$

These assertions follows from the well known Binet formula

$$F_n = \frac{\varphi^{n+1} - (-1/\varphi)^{n+1}}{\sqrt{5}}.$$

Exploiting that, $0 \leq \left|(-1/\varphi)^{n+1}\right| \ll 1$, we also deduce from it the very good approximation $F_n \approx \varphi^{n+1}/\sqrt{5}$. Thus $f_\varphi^{(n)}(t)$ is very well approximated by $\varphi^{-2(n+1)}\,t + \varphi^{2(n+1)}/(5\,t)$ when n is large enough. We may also calculate that

$$t_n(\varphi) = \sqrt{\frac{F_n^2 - F_{n-1}^2}{\varphi^{-2n} - \varphi^{-2(n+1)}}}$$
$$= \varphi^n\, F_{n-1}\sqrt{((F_n/F_{n-1})^2 - 1)\,\varphi}$$
$$\approx \varphi^{n+1}\, F_{n-1} \approx \varphi^{2(n+1)}/\sqrt{5},$$

from which we get

$$\frac{t_{n+1}^2 - t_n^2}{t_{n+1} - t_n} = t_{n+1} + t_n \approx \varphi^{n+1}(\varphi\, F_n + F_{n-1}) = \varphi^{2(n+1)},$$

as well as

$$\frac{\log(t_{n+1}/t_n)}{t_{n+1} - t_n} \approx \sqrt{5}\,\frac{\log(\varphi^{2n+3}/\varphi^{2n+1})}{\varphi^{2n+3} - \varphi^{2n+1}} = 2\sqrt{5}\,\frac{\log(\varphi)}{\varphi^{2(n+1)}}$$

Hence, applying formula (7), we find that

$$\begin{aligned}
g_\varphi &= \limsup_{n\to\infty} \frac{1}{t_{n+1} - t_n} \int_{t_n}^{t_{n+1}} f_\varphi^{(n)}(t)\,dt \\
&= \limsup_{n\to\infty} \frac{1}{t_{n+1} - t_n} \left[\varphi^{-2(n+1)}\,\frac{t^2}{2} + \frac{\varphi^{2(n+1)}}{5}\log(t) \right]_{t=t_n}^{t=t_{n+1}} \\
&= \limsup_{n\to\infty} \frac{\varphi^{-2(n+1)}}{2}\left(\frac{t_{n+1}^2 - t_n^2}{t_{n+1} - t_n} \right) + \frac{\varphi^{2(n+1)}}{5}\left(\frac{\log(t_{n+1}/t_n)}{t_{n+1} - t_n} \right) \\
&= \frac{1}{2} + \frac{2}{\sqrt{5}}\log(\varphi).
\end{aligned}$$

In the case of $\psi := 1 + \sqrt{2}$, one replaces Fibonacci numbers by Pell numbers, P_n, and uses

$$f_\psi^{(n)}(t) = P_n/t + \psi^{-2n}\,t, \qquad P_n \approx \varphi^n/\sqrt{8}, \qquad \text{and} \qquad t_n \approx \psi^{2n+1}/\sqrt{8},$$

to show that $g_\psi = 1/2 + \log(\psi)/\sqrt{8}$ with a very similar calculation.

References

Adler I (1974) A model of contact pressure in phyllotaxis. J Theor Biol 45:1–79

Aigner M (2013) Markov's theorem and 100 years of the uniqueness conjecture, a mathematical journey from irrational numbers to perfect matchings. Springer, Berlin

Atela P, Golé et C, Hotton S (2002) A dynamical system for plant pattern formation: a rigorous analysis. J Nonlinear Sci 12:641–676

Bacher R (2014) On geodesics of phyllotaxis. Confluentes Math 6(1):3–27

Couder Y, Douady S (1996a) Phyllotaxis as a dynamical self organizing process part I: the spiral modes resulting from time-periodic iterations. J Theor Biol 178:255–274

Couder Y, Douady S (1996b) Phyllotaxis as a dynamical self organizing process part II: the spontaneous formation of a periodicity and the coexistence of spiral and whorled patterns. J Theor Biol 178:275–294

Couder Y, Douady S (1996c) Phyllotaxis as a dynamical self organizing process part III: the simulation of the transient regimes of ontogeny. J Theor Biol 178:295–312

Coxeter HSM (1972) The role of intermediate convergents in Tait's explanation for phyllotaxis. J Algebra 20:167–175

Douady S (1998) The selection of phyllotactic patterns. In: Jean RV, Barabe D (eds) Symmetry in plants. World Scientific, Singapore, pp 335–358

Hermite C (1916) Sur l'introduction des variables continues dans la théorie des nombres. J Reine Angew Math 41:191–216

Humbert G (1916) Sur la méthode d'approximation d'Hermite. J Math Pures App 7ème série 2:79–103

van Iterson G (1907) Mathematische und mikroskopisch-anatomische Studien über Blattstellungen, nebst Betrachtungen über den Schalenbau der Miliolinen, Verlag von Gustav Fischer in Jena

Jacobs B (2014) On Hermite's algorithm. Bachelor's thesis, Utrecht University

Jean RV, Barabé D (1998) Symmetry in plants. World Scientific, Singapore

Leigh EG Jr (1983) The golden section and spiral leaf-arrangement. Trans Conn Acad Arts Sci 44:163–176

Markoff AA (1879) Sur les formes quadratiques binaires indéfinies. Math Ann 15:381–496

Markoff AA (1880) Sur les formes quadratiques binaires indéfinies (second mémoire). Math Ann 17:379–399

Marzec C, Kappraff J (1983) Properties of maximal spacing on a circle related to phyllotaxis and to the golden mean. J Theor Biol 103:201–226

Okabe T (2012a) Systematic variations in divergence angle. J Theor Biol 313:20–41. arXiv:1212.3377

Okabe T (2012b) Geometric interpretation of phyllotaxis transition. arXiv:1212.3112

Refahi Y, Brunoud G, Farcot E, Jean-Marie A, Pulkkinen M, Vernoux T, Godin C (2016) A stochastic multicellular model identifies biological watermarks from disorders in self-organized patterns of phyllotaxis. eLIFE. https://doi.org/10.7554/eLife.14093

Reutenauer C (2018) From Christoffel words to Markoff numbers. Oxford University Press, Oxford

Ridley JN (1986) Ideal phyllotaxis on general surfaces of revolution. Math Biosci 79(1):1–24

Serre J-P (1970) Cours d'arithmétique. Presses universitaires de France, Paris