Dangbo *Décembre 2016*

Institut de Mathématiques et de Sciences Physiques (IMSP)
http://imsp-benin.com/home/
Centre d'Excellence Africain en Sciences Mathématiques et Applications
http://imsp-benin.com/ceasma

Diophantine approximation with applications to dynamical systems

Michel Waldschmidt

Université Pierre et Marie Curie (Paris 6) France

http://www.imj-prg.fr/~michel.waldschmidt/

Abstract

Dynamical systems were studied by Henri Poincaré and Carl Ludwig Siegel, who developed the theory of celestial mechanics. The behavior of a holomorphic dynamical system near a fixed point depends on a Diophantine condition.

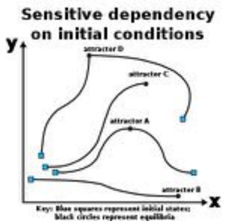
Along these lines, we give a survey on Diophantine approximation, culminating with the subspace theorem of Wolfgang Schmidt.

Celestial mechanics

Classical mechanics :
Sir Isaac Newton
(1643 – 1727)

- The solar system
- The three body problem
- Two body problem

Edward Norton Lorenz (1917 – 2008)



In chaos theory, the *butterfly effect* is the sensitive dependency on initial conditions in which a small change at one place in a deterministic nonlinear system can result in large differences in a later state. The name of the effect, coined by Edward Lorenz, is derived from the theoretical example of the formation of a hurricane being contingent on whether or not a distant butterfly had flapped its wings several weeks earlier.

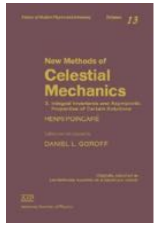
Lorenz's *butterfly effect*

Two states differing by imperceptible amounts may eventually evolve into two considerably different states. If, then, there is any error whatever in observing the present state — and in any real system such errors seem inevitable — an acceptable prediction of an instantaneous state in the distant future may well be impossible. In view of the inevitable inaccuracy and incompleteness of weather observations, precise very-long-range forecasting seems to be nonexistent.

Lorenz's description of the butterfly effect followed in 1969.

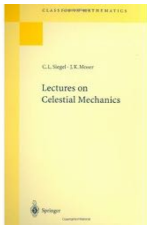
However, recent research shows that *complex systems may not behave like systems with fewer parameters.*

Henri Poincaré (1854 – 1912)


```
Credit Photos:
http://www-history.mcs.st-andrews.ac.uk/history
```


Carl Ludwig Siegel (1896 – 1981)




Dynamical System : iteration of a map

Consider a set X and map $f : X \to X$. We denote by f^2 the composed map $f \circ f : X \to X$.

More generally, we define inductively $f^n : X \to X$ by $f^n = f^{n-1} \circ f$ for $n \geq 1$, with f^0 being the identity.

The orbit of a point $x \in X$ is the sequence

$$(x, f(x), f^2(x), \dots)$$

of elements of X.

Fixed points, periodic points

A fixed point is an element $x \in X$ such that $f(x) = x$. A fixed point is a point, the orbit of which has one element x.

A periodic point is an element $x \in X$ for which there exists $n \geq 1$ with $f^n(x) = x$. The smallest such n is the length of the period of x, and all such n are multiples of the period length. The orbit

$$\{x, f(x), \dots, f^{n-1}(x)\}$$

has n elements.

For instance, a fixed point is a periodic point of period length 1.

Example 1 : endomorphism of a vector space

Take for X a finite dimensional vector space V over a field K and for $f : V \to V$ a linear map.

A fixed point of f is an element $x \in V$ such that $f(x) = x$, hence, it is nothing else than an eigenvector with eigenvalue 1.

A periodic point of f is an element $x \in V$ such that there exists $n \geq 1$ with $f^n(x) = x$, hence, f has an eigenvalue λ with $\lambda^n = 1$ (root of unity).

If V has dimension d and if we choose a basis of V, then to f is associated a $d \times d$ matrix A with coefficients in K.

Associated matrix

When f is the linear map associated with the $d \times d$ matrix A, then, for $n \geq 1$, f^n is the linear map associated with the matrix A^n.

To compute A^n, we write the matrix A as a conjugate to either a diagonal or a Jordan matrix

$$A = P^{-1}DP,$$

where P is a regular $d \times d$ matrix. Then, for $n \geq 0$,

$$A^n = P^{-1}D^nP.$$

Diagonal form

If D is diagonal with diagonal $(\lambda_1, \dots, \lambda_d)$, then D^n is diagonal with diagonal $(\lambda_1^n, \dots, \lambda_d^n)$ and

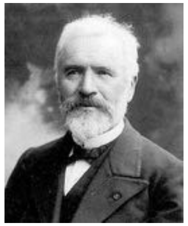
$$A^n = P^{-1} \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_d^n \end{pmatrix} P.$$

`Exercise` : compute A^n for $n \geq 0$ and for each of the two matrices

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Camille Jordan (1838 – 1922)

If A cannot be diagonalized, it can be put in Jordan form with diagonal blocs like



$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

For instance, for $d = 2$,

$$A = P^{-1}DP \quad \text{with} \quad D = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

and

$$A^n = P^{-1} \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} P.$$

Example 2 : holomorphic dynamic

Our second and main example of a dynamical system is with an open set $\mathcal{V}$ in $\mathbf{C}$ and an analytic (=holomorphic) map $f : \mathcal{V} \to \mathcal{V}$. The main goal will be to investigate the behavior of f near a fixed point $z_0 \in \mathcal{V}$. So we assume $f(z_0) = z_0$.

The local behavior of the dynamics defined by f depends on the derivative $f'(z_0)$ of f at the fixed point.

If $|f'(z_0)| < 1$, then z_0 is an attracting point.

If $|f'(z_0)| > 1$, then z_0 is a repelling point.

The most interesting case is $|f'(z_0)| = 1$

Conjugate holomorphic maps

We wish to mimic the situation of an endomorphism of a vector space : in place of a regular matrix P, we introduce a local change of coordinates. Let $\mathcal{D}$ be the open unit disc in $\mathbf{C}$ and $g : \mathcal{D} \to \mathcal{D}$ an analytic map with $g(0) = 0$. We say that f and g are conjugate if there exists an analytic map $h : \mathcal{V} \to \mathcal{D}$, with $h'(z_0) \neq 0$, such that $h(z_0) = 0$ and $h \circ f = g \circ h$.

$$\begin{tikzcd} z_0 \in \mathcal{V} \arrow[r, "f"] \arrow[d, "h"'] & \mathcal{V} \ni z_0 \arrow[d, "h"] \\ 0 \in \mathcal{D} \arrow[r, "g"'] & \mathcal{D} \ni 0 \end{tikzcd} \qquad \begin{array}{c} f(z_0) = z_0 \\ \\ g(0) = 0 \end{array}$$

Local behavior

Assume $f : \mathcal{V} \to \mathcal{V}$ and $g : \mathcal{D} \to \mathcal{D}$ are conjugate : there exists $h : \mathcal{D} \to \mathcal{D}$, with $h'(z_0) \neq 0$ and $h \circ f = g \circ h$.

From $h'(z_0) \neq 0$, one deduces that h is unique up to a multiplicative nonzero factor.

Further,

$$h \circ f^2 = h \circ f \circ f = g \circ h \circ f = g \circ g \circ h = g^2 \circ h$$

and, by induction, $h \circ f^n = g^n \circ h$ for all $n \geq 0$.

Linearization of germs of analytic diffeomorphisms of one complex variable

Lemma. *If f is conjugate to the homothety $g(z) = \lambda z$, then $\lambda = f'(z_0)$.*
Hence, in this case, f is conjugate to its linear part. One says that f is *linearizable*.

`Proof`. Take the derivative of $h \circ f = g \circ h$ at z_0 :

$$h'(z_0)f'(z_0) = \lambda h'(z_0)$$

and use $h'(z_0) \neq 0$.

Johann Samuel König (1712 – 1757)

Define $\lambda = f'(z_0)$.

Theorem (Königs and Poincaré). *For $|\lambda| \notin \{0, 1\}$, f is linearizable.*

For $\lambda = 0$ and $z_0 = 0$, f has a zero of multiplicity $n \geq 2$ at 0 and is conjugate to $z \mapsto z^n$ (A. Böttcher).

We are interested in the case $|\lambda| = 1$. It was conjectured in 1912 by E. Kasner that f is always linearizable. In 1917, G.A. Pfeiffer produced a counterexample. In 1927, H. Cremer proved that in the *generic* case, f is not linearizable.

The case $|\lambda| = 1$

Assume $|\lambda| = 1$. Write $\lambda = e^{2i\pi\theta}$. The real number θ is the rotation number of f at z_0.

In 1942, C.L. Siegel proved that if θ satisfies a Diophantine condition, then f is conjugate to the rotation $z \mapsto e^{2i\pi\theta} z$.

In 1965, A.D. Brjuno relaxed Siegel's assumption.

In 1988, J.C. Yoccoz showed that if θ does not satisfies Brjuno's condition, then the dynamic associated with

$$f(z) = \lambda z + z^2$$

has infinitely many periodic points in any neighborhood of 0, hence, is not linearizable.

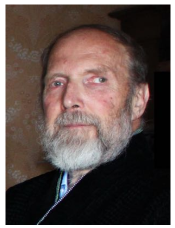
C.L. Siegel, A.D. Brjuno, J.C. Yoccoz

Carl Ludwig Siegel
(1896 – 1981)

Jean-Christophe Yoccoz
(1957 — 2016)

Alexander Dmitrijewitsch Brjuno
(1940 –)

1942



1965

1988

KAM Theory

Andrey Nikolaevich Kolmogorov (1903 – 1987)

Vladimir Igorevich Arnold (1937 – 2010)

Jürgen Kurt Moser (1928 – 1999)

Siegel's Diophantine condition : Liouville numbers

Siegel's Diophantine condition on the rotation number θ is that a rational number p/q with a *small denominator* q cannot be too good of a rational approximation of θ.

The same condition was introduced by Liouville, who proved in 1844 that Siegel's Diophantine condition is satisfied if θ is an algebraic number.

Algebraic numbers

A complex number α is algebraic if there exists a nonzero polynomial $f \in \mathbf{Z}[X]$ such that $f(\alpha) = 0$. The smallest degree of such a polynomial is the degree of the algebraic number α.

For instance $\sqrt{2}$, $i = \sqrt{-1}$, $\sqrt[3]{2}$, $e^{2i\pi a/b}$ (for a and b integers, $b > 0$) are algebraic numbers.

The roots of the quintic polynomial

$$X^5 - 6X + 3$$

are algebraic numbers (but cannot be expressed using radicals).

Transcendental numbers

A number which is not algebraic is transcendental.
The existence of transcendental numbers was not known before 1844, when Liouville produced the first examples, like

$$\xi = \sum_{n \geq 0} \frac{1}{10^{n!}}.$$

The idea of Liouville is to prove a *Diophantine property* of algebraic numbers, namely that rational numbers with small denominators do not produce sharp approximations. Hence, a real number with too good rational approximations cannot be algebraic.
For instance, with the above number ξ and $q = 10^{N!}$,

$$p = \sum_{n=0}^{N} 10^{N!-n!}, \qquad 0 < \xi - \frac{p}{q} < \frac{2}{10^{(N+1)!}} = \frac{2}{q^{N+1}}.$$

Liouville's inequality (1844)

Liouville's inequality. *Let* α *be an algebraic number of degree* $d \geq 2$. *There exists* $c(\alpha) > 0$ *such that, for any* $p/q \in \mathbf{Q}$ *with* $q > 0$,

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^d}$$

Joseph Liouville
(1809 - 1882)

The Diophantine condition of Liouville and Siegel

A real number θ satisfies a *Diophantine condition* if there exists a constant $\kappa > 0$ such that

$$\left|\theta - \frac{p}{q}\right| > \frac{1}{q^\kappa}$$

for all $p/q \in \mathbf{Q}$ with $q \geq 2$.

A real number is a *Liouville number* if it does not satisfy a Diophantine condition.

Generic vs full measure, Baire vs Lebesgue

René Baire
(1874 – 1932)

Henri Léon Lebesgue
(1875 – 1941)

In dynamical systems, a property is satisfied for a *generic rotation number* θ if it is true for all numbers in a countable intersection of dense open sets – these sets are called G_δ sets by Baire who calls *meager* the complement of a G_δ set.
The set of numbers which do not satisfy a Diophantine condition is a generic set. For Lebesgue measure, the set of Liouville numbers (i.e. the set of numbers which do not satisfy a Diophantine condition) has measure zero.

Mathematical genealogy

René Baire (1899)
|
Arnaud Denjoy (1909)
|
Charles Pisot (1938)
|
Yvette Amice (1965)
|
Jean Fresnel (1967)
|
Michel Waldschmidt (1972)

`http://genealogy.math.ndsu.nodak.edu`

Brjuno's condition

In terms of continued fraction, the Diophantine condition (of Liouville and Siegel) can be written

$$\sup_{n\geq 1} \frac{\log q_{n+1}}{\log q_n} < \infty.$$

The condition of Brjuno is

$$\sum_{n\geq 1} \frac{\log q_{n+1}}{q_n} < \infty.$$

If a number θ satisfies the Diophantine condition, then it satisfies Brjuno's condition. But there are (transcendental) numbers which do not satisfy the Diophantine condition, but satisfy Brjuno's condition.

Improvements of Liouville's inequality

In the lower bound

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^d}$$

for α real algebraic number of degree $d \geq 3$, the exponent d of q in the denominator of the right hand side was replaced by κ with

- any $\kappa > (d/2) + 1$ by A. Thue (1909),
- $2\sqrt{d}$ by C.L. Siegel in 1921,
- $\sqrt{2d}$ by F.J. Dyson and A.O. Gel'fond in 1947,
- any $\kappa > 2$ by K.F. Roth in 1955.

Thue– Siegel– Roth Theorem

Axel Thue (1863 – 1922)

Carl Ludwig Siegel (1896 – 1981)

Klaus Friedrich Roth (1925 – 2015)

For any real algebraic number α, for any $\epsilon > 0$, the set of $p/q \in \mathbf{Q}$ with $|\alpha - p/q| < q^{-2-\epsilon}$ is finite.

Thue– Siegel– Roth Theorem

An equivalent statement is that, for any real algebraic irrational number α and for any $\epsilon > 0$, *the set of $p/q \in \mathbf{Q}$ such that*

$$q|q\alpha - p| < q^{-\epsilon}$$

is finite.

The conclusion can be phrased :
the set of $(p, q) \in \mathbf{Z}^2$ such that

$$q|q\alpha - p| < q^{-\epsilon}$$

is contained in the union of finitely many lines in $\mathbf{Z}^2$.

Schmidt's Subspace Theorem (1970)

For $m \geq 2$ let $L_0, \ldots, L_{m-1}$ be m independent linear forms in m variables with algebraic coefficients. Let $\epsilon > 0$. Then the set

$$\{\mathbf{x} = (x_0, \ldots, x_{m-1}) \in \mathbf{Z}^m ;$$
$$|L_0(\mathbf{x}) \cdots L_{m-1}(\mathbf{x})| \leq |\mathbf{x}|^{-\epsilon}\}$$

is contained in the union of finitely many proper subspaces of $\mathbf{Q}^m$.

Wolfgang M. Schmidt
(1933 –)

Schmidt's Subspace Theorem

W.M. Schmidt (1970) : *For $m \geq 2$ let $L_0, \ldots, L_{m-1}$ be m independent linear forms in m variables with algebraic coefficients. Let $\epsilon > 0$. Then the set*

$$\{\mathbf{x} = (x_0, \ldots, x_{m-1}) \in \mathbf{Z}^m ; \ |L_0(\mathbf{x}) \cdots L_{m-1}(\mathbf{x})| \leq |\mathbf{x}|^{-\epsilon}\}$$

is contained in the union of finitely many proper subspaces of $\mathbf{Q}^m$.

Example : $m = 2$, $L_0(x_0, x_1) = x_0$, $L_1(x_0, x_1) = \alpha x_0 - x_1$.
Roth's Theorem : *for any real algebraic irrational number α, for any $\epsilon > 0$, the set of $p/q \in \mathbf{Q}$ with $q|\alpha q - p| < q^{-\epsilon}$ is finite.*

Specialization arguments

The proof of Schmidt's Subspace Theorem has an arithmetic nature, the fact that the linear forms have algebraic coefficients is crucial.
The subspace Theorem does not hold without this assumption.

However, there are *specializations arguments* which enable one to deduce consequences without any arithmetic assumption, these corollaries ave valid for fields of zero characteristic in general.
An example is the so–called *Theorem of the generalized S–unit equation.*

The generalized S–unit equation (1982)

Let K be a field of characteristic zero, let G be a finitely generated multiplicative subgroup of the multiplicative group $K^\times = K \setminus \{0\}$ and let $n \geq 2$.
Theorem (Evertse, van der Poorten, Schlickewei). *The equation*

$$u_1 + u_2 + \cdots + u_n = 1,$$

where the unknowns $u_1, u_2, \cdots, u_n$ take their values in G, for which no nontrivial subsum

$$\sum_{i \in I} u_i \qquad \emptyset \neq I \subset \{1, \ldots, n\}$$

vanishes, has only finitely many solutions.

The generalized S–unit equation (1982)

Jan Hendrick Evertse Alf van der Poorten
Hans Peter Schlickewei



Linear recurrence sequences

Given a field K (of zero characteristic), a sequence $(u_n)_{n\geq 0}$ is a linear recurrence sequence if there exist an integer $d \geq 1$ and elements $a_0, a_1, \ldots, a_{d-1}$ of K such that, for $n \geq 0$,

$$u_{n+d} = a_{d-1}u_{n+d-1} + \cdots + a_1 u_{n+1} + a_0 u_n.$$

Such a sequence $(u_n)_{n\geq 0}$ is determined by the coefficients $a_0, a_1, \ldots, a_{d-1}$ and by the initial values $u_0, u_1, \ldots, u_{d-1}$.

Exponential polynomials

If $\alpha_1, \ldots, \alpha_k$ are the distinct roots of the polynomial

$$X^d - a_{d-1}X^{d-1} - \cdots - a_1 X - a_0$$

and $s_1, \ldots, s_k$ their multiplicities, then one can write

$$u_n = \sum_{i=1}^{k} A_i(n)\alpha_i^n,$$

where $A_1, \ldots, A_k$ are polynomials with A_i of degree $< s_i$.

Hence, a linear recurrence sequence is given by an exponential polynomial. Conversely, a sequence given by an exponential polynomial is a linear recurrence sequence.

Skolem – Mahler – Lech Theorem

The generalized S–unit Theorem yields the following :
Theorem (Skolem 1934 – Mahler 1935 – Lech 1953). *Given a linear recurrence sequence, the set of indices $n \geq 0$ such that $u_n = 0$ is a finite union of arithmetic progressions.*

Thoralf Albert Skolem (1887 – 1963) Kurt Mahler (1903 – 1988) Christer Lech


An *arithmetic progression* is a set of positive integers of the form $\{n_0, n_0 + k, n_0 + 2k, \ldots\}$. Here, we allow $k = 0$.

Another dynamical system

Let V be a finite dimensional vector space over a field of zero characteristic, H an hyperplane of V, $f : V \to V$ an endomorphism (linear map) and x an element in V.

Corollary. *If there exist infinitely many* $n \geq 1$ *such that* $f^n(x) \in H$, *then there is an (infinite) arithmetic progression of* n *for which it is so.*

Idea of the proof

Choose a basis of V. The endomorphism f is given by a square $d \times d$ matrix A, where d is the dimension of V. Consider the characteristic polynomial of A, say

$$X^d - a_{d-1}X^{d-1} - \cdots - a_1 X - a_0.$$

By the Theorem of Cayley – Hamilton,

$$A^d = a_{d-1}A^{d-1} + \cdots + a_1 A + a_0 I_d$$

where I_d is the identity $d \times d$ matrix.

Theorem of Cayley – Hamilton

Arthur Cayley
(1821 – 1895)

Sir William Rowan Hamilton
(1805 – 1865)

Hence, for $n \geq 0$,

$$A^{n+d} = a_{d-1}A^{n+d-1} + \cdots + a_1 A^{n+1} + a_0 A^n.$$

It follows that each entry $a_{ij}^{(n)}$, $1 \leq i, j \leq d$, satisfies a linear recurrence sequence, the same for all i, j.

Hyperplane membership

Let $b_1 x_1 + \cdots + b_d x_d = 0$ be an equation of the hyperplane H in the selected basis of V. Let ${}^t\underline{b}$ denote the $1 \times d$ matrix $(b_1, \ldots, b_d)$ (transpose of a column matrix $\underline{b}$). Using the notation $\underline{v}$ for the $d \times 1$ (column) matrix given by the coordinates of an element v in V, the condition $v \in H$ can be written ${}^t\underline{b}\,\underline{v} = 0$.

Let x be an element in V and $\underline{x}$ the $d \times 1$ (column) matrix given by its coordinates. The condition $f^n(x) \in H$ can now be written

$${}^t\underline{b} A^n \underline{x} = 0.$$

The entry u_n of the 1×1 matrix ${}^t\underline{b} A^n \underline{x}$ satisfies a linear recurrence sequence, hence, the Skolem – Mahler – Lech Theorem applies.

Solution of the exercises

Exercise 1 : compute A_1^n for $n \geq 0$ and

$$A_1 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}.$$

By induction one checks

$$A_1^n = \begin{pmatrix} 1 & 0 \\ 1-2^n & 2^n \end{pmatrix}.$$

The trace of A_1 is 3, the determinant is 2, the characteristic polynomial is $X^2 - 3X + 2 = (X-1)(X-2)$, the linear recurrence is

$$u_{n+2} = 3u_{n+1} - 2u_n.$$

Solution of the second exercise

Exercise 2 : compute A_2^n for $n \geq 0$, where A_2 is the matrix

$$A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

The trace of A_2 is 1, the determinant is -1, the characteristic polynomial is $X^2 - X - 1$, the linear recurrence is

$$u_{n+2} = u_{n+1} + u_n.$$

It follows that

$$A_2^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}$$

where $(F_n)_{n\geq 0}$ is the solution of linear recurrence sequence $F_{n+2} = F_{n+1} + F_n$ given by the initial conditions $F_0 = 0$, $F_1 = 1$.

Leonardo Pisano Fibonacci

The Fibonacci sequence $(F_n)_{n\geq 0}$:

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,$$

$$34,\ 55,\ 89,\ 144,\ 233\dots$$

is defined by

$$F_0 = 0,\ F_1 = 1,$$

$$F_n = F_{n-1} + F_{n-2} \quad (n \geq 2).$$

Leonardo Pisano Fibonacci
(1170–1250)

The online encyclopaedia of integer sequences

$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233,\ 377,\ 610,\ 987,\ 1597,$
$2584,\ 4181,\ 6765,\ 10946,\ 17711,\ 28657,\ 46368,\ 75025,\ 121393,\ 196418,$
$317811,\ 514229,\ 832040,\ 1346269,\ 2178309,\ 3524578,\ 5702887,\ 9227465,\ \dots$

Fibonacci sequence is online
The On-Line Encyclopedia of Integer Sequences

Neil J. A. Sloane

http://www.research.att.com/~njas/sequences/A000045

Diagonalization

The two eigenvalues of the matrix

$$A_1 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$$

are 1 and 2 with eigenvectors $(1,1)$ and $(0,1)$ respectively, so that

$$A_1 = P^{-1}DP$$

with

$$P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

and

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Computation of A_1^n for $A_1 = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$

From $A_1 = P^{-1}DP$ with

$$P^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},$$

we deduce

$$\begin{aligned} A_1^n &= P^{-1}D^nP \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 1-2^n & 2^n \end{pmatrix}. \end{aligned}$$

Diagonalization of A_2

The characteristic polynomial of the matrix

$$A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

is

$$X^2 - X - 1 = (X - \Phi)(X + \Phi^{-1}),$$

where Φ is the Golden Number :

$$\Phi = \frac{1+\sqrt{5}}{2} = 1.618033\ldots, \qquad \Phi^{-1} = \frac{-1+\sqrt{5}}{2} = \Phi - 1$$

and

$$\Phi + \Phi^{-1} = \sqrt{5}.$$

Diagonalization of A_2

The eigenvalues of A_2 are Φ and $-\Phi^{-1}$ with eigenvectors $(1, \Phi)$ and $(1, -\Phi^{-1})$. Hence

$$A_2 = P^{-1}DP$$

with

$$P = \frac{-1}{\sqrt{5}} \begin{pmatrix} -\Phi^{-1} & -1 \\ -\Phi & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 \\ \Phi & -\Phi^{-1} \end{pmatrix}$$

and

$$D = \begin{pmatrix} \Phi & 0 \\ 0 & -\Phi^{-1} \end{pmatrix}.$$

Computation of A_2^n for $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

From

$$
\begin{aligned}
A_2^n &= P^{-1} D^n P \\
&= \frac{-1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \Phi & -\Phi^{-1} \end{pmatrix} \begin{pmatrix} \Phi^n & 0 \\ 0 & (-\Phi)^{-n} \end{pmatrix} \begin{pmatrix} -\Phi^{-1} & -1 \\ -\Phi & 1 \end{pmatrix} \\
&= \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix}.
\end{aligned}
$$

we deduce

$$F_n = \frac{1}{\sqrt{5}} \left(\Phi^n - (-\Phi)^{-n} \right)$$

Fibonacci sequence and the Golden Number

A. De Moivre (1730), L. Euler (1765), J.P.M. Binet (1843) :

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Remark Since $0 < \Phi^{-1} < 1$, the quantity Φ^{-n} is exponentially small, hence F_n is very close to $\frac{1}{\sqrt{5}}\Phi^n$.

De Moivre – Euler – Binet formula

Abraham de Moivre (1667–1754)

Leonhard Euler (1707–1783)

Jacques Philippe Marie Binet (1786–1856)

F_n is the nearest integer to $\frac{1}{\sqrt{5}}\Phi^n$.