

La suite logistique et le chaos

Daniel PERRIN

À cause du clou, le fer fut perdu.
À cause du fer, le cheval fut perdu.
À cause du cheval, le cavalier fut perdu.
À cause du cavalier, la bataille fut perdue.
À cause de la bataille, la guerre fut perdue.
À cause de la guerre, la liberté fut perdue.
Tout cela pour un simple clou
Benjamin Franklin
(**Almanach du pauvre Richard, 1758**)

Je remercie René Cori de m'avoir donné l'occasion de m'intéresser de plus près à la suite logistique, qui me fascine depuis longtemps. Je rappelle toutefois que je ne suis nullement un spécialiste du sujet et que j'ai été très impressionné par la vertigineuse production mathématique autour de ce thème. Heureusement, j'ai bénéficié, pour préparer mon exposé, des compétences de Sylvie Ruette que je remercie vivement¹.

L'un des aspects qui m'intéresse dans cette question est son côté expérimental. À cet égard, je remercie les concepteurs de Cabri et ceux de la calculatrice TI Voyage 200 pour la qualité du matériel qu'ils mettent à la disposition des mathématiciens. J'ai aussi utilisé, pour certains calculs plus compliqués, les logiciels Maple et MuPAD.

1 Modélisation

1.1 Historique et sources

¹Je remercie aussi François Sauvageot (resp. Étienne Ghys) pour m'avoir indiqué la référence à Coppel (resp. à Lyubich). Par ailleurs, le livre [Buzzi] fournit un panorama élémentaire de ces questions.

1.1.1 Introduction

Le type même de problèmes qui est à l'origine de la théorie des systèmes dynamiques est celui de l'évolution d'une population en fonction du temps. Le mot population est à prendre ici en un sens très large. Il peut aussi bien s'agir d'une population humaine, qu'animale, des victimes d'une épidémie, d'un ensemble de molécules, de particules, etc. Les modèles dont je vais parler sont des modèles² **déterministes**. Cela signifie qu'ils sont régis par une loi bien déterminée, qui doit permettre, en théorie, de décrire leur évolution à partir d'un état initial connu. On sait depuis Poincaré que, malgré cette hypothèse restrictive, le comportement de ces modèles peut être excessivement compliqué, en particulier à cause de la sensibilité du système aux conditions initiales. C'est la fameuse phrase prêtée au météorologiste Edward Lorenz en 1972 : *Le battement d'ailes d'un papillon au Brésil peut-il provoquer une tornade au Texas ?* qui a donné naissance à beaucoup de spéculations hasardeuses sur le fameux effet papillon³. Dans ce qui suit nous essayerons de manière très modeste, sur l'équation très particulière $u_{n+1} = \mu u_n(1 - u_n)$, de montrer comment se manifeste le chaos et de discuter un point qui semble essentiel d'un point de vue philosophique : le chaos est-il générique ?

1.1.2 Euler, Malthus

Les premiers travaux connus sur ces questions sont ceux d'Euler (1707-1783) (*Recherches générales sur la mortalité et la multiplication du genre humain*). Dans ce travail, qui date de 1760, Euler calcule notamment la population d'une ville ou d'une province pour une certaine année. Voici ce qu'Euler dit *loc. cit.* p. 157 : *Car si nous posons le nombre de tous les vivans l'année prochaine = nM , celui des vivans à présent étant = M , il faut tirer la valeur de n de l'équation trouvée*

$$\frac{M}{N} = 1 + \frac{(1)}{n} + \frac{(2)}{n^2} + \frac{(3)}{n^3} + \frac{(4)}{n^4} + \&c.$$

Avec des notations plus modernes⁴, on a la formule suivante :

$$p_n = q_n \left(1 + \frac{\pi_1}{\lambda} + \frac{\pi_2}{\lambda^2} + \cdots + \frac{\pi_{100}}{\lambda^{100}} + \cdots \right)$$

où p_n est la population de la ville à l'année n , q_n le nombre de naissances de cette même année, π_k la probabilité (calculée à partir de tables, voir

²Je ne parlerai pas du tout ici de modèles aléatoires.

³Voir l'excellent article de Nicolas Witkowski, *Alliage* numéro 22, 1995.

⁴Euler note (k) ce que je note π_k et n ce que je note λ .

[Euler] p. 152 ou ci-dessous Annexe 7.1) de vivre k années dans la région donnée et $\lambda = q_{n+1}/q_n$ le rapport entre le nombre de naissances de l'année $n + 1$ et de l'année n , supposé indépendant de n . Le nombre de naissances à l'année $n - k$ est donc $\frac{q_n}{\lambda^k}$, ce qui donne la formule ci-dessus. Euler arrête généralement la sommation à 100, considérant qu'au-delà la probabilité de survie est négligeable⁵. Avec cette formule, il est clair que le rapport entre la population à l'année $n + 1$ et celle à l'année n est égal au rapport des nombres de naissances donc à λ . Ainsi, on a la formule $p_{n+1} = \lambda p_n$, qui conduit à une suite géométrique⁶.

L'idée d'un accroissement exponentiel de la population est reprise et développée en 1798 par Thomas Robert Malthus (1766-1834). Son analyse conduit à modéliser la population humaine comme une suite géométrique, tandis que la capacité de production se comporterait, elle, comme une suite arithmétique. La distorsion entre les deux, à terme, le conduit à une proposition de limitation des naissances qui est depuis attachée à son nom.

1.1.3 Verhulst

Le modèle malthusien est remis en cause vers 1840 par Pierre François Verhulst (1804-1849) qui propose un modèle dit logistique⁷ qui prend en compte la limitation de la population. Le principe est simple : l'accroissement de la population n'est proportionnel à la population que pour les petites valeurs de celle-ci. Lorsqu'elle croît, des facteurs limitants apparaissent⁸ (place

⁵Les chiffres de mortalité sont assez frappants. Par exemple la probabilité d'atteindre l'âge canonique de 61 ans – le mien – n'est que de 26%.

⁶On a donc un accroissement (ou une diminution) exponentiel selon la position de λ par rapport à 1.

⁷Il semble bien que c'est Verhulst qui, le premier, applique le mot logistique à l'équation en question (voir le travail de Bernard Delmas sur Verhulst). Le sens courant de ce mot fait référence au transport et au ravitaillement des armées, mais il me semble qu'il faut le prendre ici en un sens plus ancien signalé dans le Littré, qui concerne la pratique des quatre opérations usuelles de l'arithmétique, voir François Hetman *Le langage de la prévision*.

⁸Le point de départ de mon intérêt pour la question est un exercice proposé par une étudiante de CAPES, dans lequel une population de canards en Scandinavie suivait une loi en 6^n . Il est clair que ce type de modèle est absurde. Le seul contre-exemple que je connaisse est celui de la population des lapins en Australie. En effet, 27 lapins furent introduits en 1859 dans le pays. N'ayant pas de prédateurs et une place quasiment infinie, ils se sont développés de manière exponentielle (ils étaient 22 millions 6 ans plus tard), dévorant une bonne partie de la végétation. L'introduction de renards comme prédateurs n'ayant pas été concluante, la seule façon d'enrayer cette invasion a été de répandre la myxomatose, avec un résultat foudroyant. Cette méthode est évidemment à employer avec précaution, un médecin français d'Eure-et-Loir l'ayant utilisée en 1952 pour se débarrasser des lapins qui ravageaient son jardin a répandu l'épidémie dans toute la France. Ceux qui,

ou quantité de nourriture disponible, etc.) qui font qu'il y a une population maximale M . Verhulst postule alors que l'accroissement de la population x est proportionnel à la quantité $x(M - x)$. Ce modèle lui permet de donner en 1837 une prévision de la population de la France en 1930 de 40 millions, prévision somme toute raisonnable puisque la population de la France⁹, en réalité, est de 41,5 millions en 1931.

1.1.4 Aujourd'hui

Le modèle logistique de Verhulst est encore utilisé aujourd'hui dans nombre de questions (en démographie, biologie¹⁰, médecine, etc.) et il est assez probant¹¹. On utilise d'ailleurs en statistique une régression logistique, analogue à la régression linéaire pour modéliser les évolutions de populations qui suivent des courbes "en S". Ce modèle est important car il permet, avec une fiabilité relativement bonne, de prévoir l'évolution future de la population.

Nous allons voir qu'il y a deux modèles logistiques, selon qu'on fait varier le temps de manière continue ou discrète. Le modèle à temps continu conduit à une équation différentielle très simple, on en connaît parfaitement les solutions et on sait les interpréter. En revanche, le modèle discret peut mener, pour certaines valeurs des paramètres, à des comportements beaucoup plus compliqués, voire chaotiques. Cette étude est relativement récente et elle est partie des applications¹². On peut citer un biologiste, Robert May (1974), un physicien, Mitchell Feigenbaum¹³ (1975), etc. Les problèmes mathématiques qu'elle pose sont absolument passionnants¹⁴ et constituent le sujet de cet exposé.

1.2 Équations différentielles, équations aux différences finies

comme moi, n'ont plus 20 ans depuis longtemps, s'en souviennent sans doute.

⁹Cela étant, il y a eu la saignée de la guerre de 14, que Verhulst ne pouvait prévoir. Sans cela, son chiffre serait sans doute moins bon.

¹⁰Voir le problème de Bac 2003!

¹¹Pour des alternatives au modèle logistique en géographie, voir [Langlois-Daude].

¹²Un excellent panorama de la question se trouve dans le livre de Cvitanovic, *Universality in chaos*. On y verra que le chaos apparaît dans de multiples situations : dynamique des fluides, chimie, météorologie, etc.

¹³Feigenbaum a découvert bon nombre de phénomènes sur la suite logistique en utilisant uniquement une calculatrice de poche (programmable tout de même).

¹⁴Je cite [Lyubich2] : *Still, it turns out that this family exhibits extremely rich properties and sends important messages to the bigger world.*

1.2.1 Les principes

Notre objectif est de modéliser une population p qui dépend du temps de manière déterministe, c'est-à-dire avec une loi parfaitement définie. Les deux cas essentiels sont les suivants :

- Le temps est continu, $t \in \mathbf{R}$, et on a une **fonction** $p(t)$, à valeurs **réelles**¹⁵. Dans ce cas, on suppose que si on connaît la population au temps t , on la connaît au temps $t + dt$, avec dt “infinitésimal”. Cela revient à se donner la dérivée¹⁶ $p'(t) = \frac{p(t + dt) - p(t)}{dt}$ en fonction de $p(t)$.

- Le temps est discret, $n \in \mathbf{N}$ (n désigne un nombre de minutes, d'heures, voire d'années) et on a une **suite** p_n là encore à valeurs **réelles**. Dans ce cas, on se donne l'accroissement $p_{n+1} - p_n$ en fonction de p_n (ce qui revient à se donner p_{n+1}).

On ne s'intéresse ici qu'aux cas où la population suit un modèle dit **autonome** au sens où elle ne dépend que d'elle-même et non du temps. On aura donc une équation de la forme $p' = f(p)$ ou $p_{n+1} = f(p_n)$, la fonction f étant **indépendante du temps**.

Le modèle discret s'applique¹⁷ notamment aux populations dont les générations ne se chevauchent pas. C'est le cas de nombreuses populations d'insectes¹⁸ des zones tempérées où les parents sont morts lorsque les chrysalides nouvelles éclosent.

1.1 Remarque. J'ai choisi de parler essentiellement ici d'évolution de populations, mais il y a bien d'autres domaines dans lesquels on trouve des systèmes dynamiques du type $p_{n+1} = f(p_n)$ et qui mènent aussi à des situations chaotiques. J'en cite quelques-uns pour mémoire :

- En génétique : p_n est la fréquence d'un gène au temps n .
- En épidémiologie : p_n est la proportion de la population infectée au temps n .
- En économie : n est la quantité de marchandises et p_n le prix.
- En sciences sociales, par exemple dans l'étude de la propagation des rumeurs : p_n est la proportion de gens connaissant la rumeur.

¹⁵Ce point est discuté au paragraphe suivant.

¹⁶Je note les choses comme les physiciens.

¹⁷Par ailleurs, il y a deux façons naturelles de passer d'un modèle continu à un modèle discret. L'une est la méthode de discrétisation d'Euler, bien connue, l'autre consiste à utiliser ce qu'on appelle l'application de “premier retour” inaugurée par Poincaré, voir [Collet-Eckmann].

¹⁸Il y a plusieurs exemples d'espèces de cigales qui présentent des cycles de vie de 13, voire 17 ans.

1.2.2 Un paradoxe ?

Attention, nous considérerons toujours que la population p_n envisagée est un nombre **réel** et non pas un entier. Cela peut sembler paradoxal¹⁹, mais ce n'est pas essentiel. Je vois plusieurs arguments en ce sens.

- Quand on parle de l'effectif d'une population nombreuse, par exemple des bactéries, même s'il s'agit d'un nombre entier, le plus souvent cette donnée ne résulte pas d'un comptage de toute la population, mais d'une moyenne entre données statistiques, qui fournit, au moins, un nombre rationnel.

- Une façon un peu spacieuse de se débarrasser des fonctions à valeurs entières est de remplacer le nombre par la masse. Je dis que c'est spacieux car un physicien nous dira que tout est discret si on se ramène aux particules.

- Les phénomènes sont-ils aussi discrets qu'ils en ont l'air ? Exemple : l'accroissement d'une unité de la population humaine, ça s'appelle une grossesse et, si mes lointains souvenirs sont bons, ça dure 9 mois pendant lesquels on passe, plus ou moins continûment, de 0 à 1.

- Le point principal, en fin de compte, c'est que cette extrapolation aux réels donne des résultats plausibles : c'est l'ordre de grandeur de la population qui compte et on peut retrouver des entiers en prenant la partie entière des valeurs obtenues. Je signale cependant un problème mathématique sérieux à cet égard. Il y a deux façons de retrouver des entiers si on a une suite $p_{n+1} = f(p_n)$, avec un point de départ p_0 entier, mais des p_n réels ensuite. La plus triviale consiste à regarder la suite des parties entières $[p_n]$. Mais, il peut sembler plus plausible de considérer la suite d'entiers q_n avec $q_0 = p_0$ mais $q_{n+1} = [f(q_n)]$ dans laquelle on actualise la valeur à chaque pas. Il n'est cependant pas du tout évident que cette suite soit proche de p_n . Si f est lipschitzienne de rapport $k < 1$, ce qui correspond au cas d'un point fixe attractif, on montre facilement que la différence entre p_n et q_n reste bornée par $1/(1-k)$, mais, sinon, elle peut croître indéfiniment. Le lecteur regardera le cas d'une suite arithmético-géométrique $p_{n+1} = ap_n + b$, avec $a > 1$, pour s'en convaincre.

1.3 Les modèles exponentiels

Dans ces modèles, on postule que l'accroissement de p est proportionnel à p (plus il y a de lapins, plus ils font des petits ...). Les deux variantes sont $p'(t) = kp(t)$ dans le cas continu, ou $p_{n+1} - p_n = kp_n$ dans le cas discret. Cette dernière équation s'écrit encore $p_{n+1} = (k + 1)p_n$. Dans les deux cas

¹⁹Surtout si l'on dit, comme dans l'énoncé du problème de Bac 2003 que le nombre (entier) de bactéries est une fonction dérivable du temps !

la résolution est facile. Si on appelle p_0 la population au temps 0, on trouve $p(t) = p_0 e^{kt}$ dans le cas continu et $p_n = p_0(1+k)^n$ dans le cas discret. Le comportement asymptotique de p est du même ordre :

- Si k est positif, la fonction et la suite tendent vers l'infini.
- Si k est négatif, la fonction et la suite tendent vers 0.

1.4 Le modèle logistique continu

Cette fois, on suppose que la population ne peut croître indéfiniment et on suppose qu'elle admet une borne M . L'idée de Verhulst c'est que l'accroissement de p , donné par la dérivée p' , est "proportionnel" à la fois à la population p , mais aussi à la place disponible $M - p$, d'où l'équation $p' = \mu p(M - p)$. Il est facile d'intégrer cette équation par deux méthodes au moins : le changement de fonction $z = 1/p$ qui conduit à l'équation linéaire $z' = \mu(1 - zM)$ et la décomposition en éléments simples qui mène à $\frac{p'}{p} + \frac{p'}{M-p} = \mu M$ où il ne reste plus qu'à prendre des primitives. Avec l'une ou l'autre méthode, on trouve la solution : $p(t) = \frac{M}{1 + \lambda e^{-\mu M t}}$ et l'étude de cette fonction est très simple. On obtient un graphe comme ci-dessous :

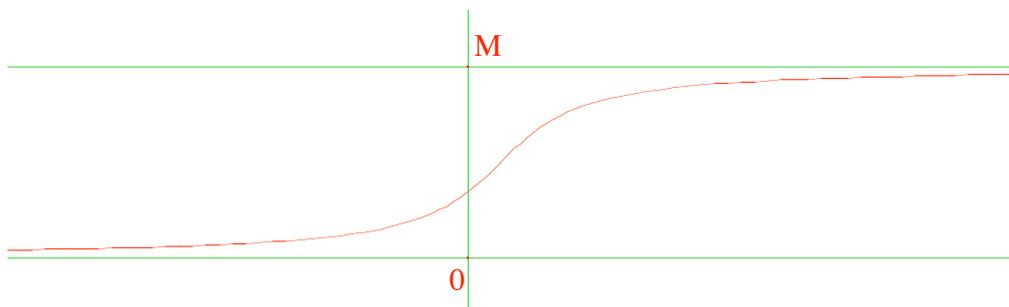


Figure 1: La courbe logistique

1.5 Le modèle logistique discret

On peut utiliser un modèle comme ci-dessus, donnant l'accroissement de population $p_{n+1} - p_n = kp_n(M - p_n)$ ou une analyse en terme de population idéale (voir Annexe 7.2). Dans tous les cas, on obtient une équation²⁰ de degré 2 : $p_{n+1} = \mu p_n - \lambda p_n^2 = \lambda p_n \left(\frac{\mu}{\lambda} - p_n \right) = \mu p_n \left(1 - \frac{\lambda}{\mu} p_n \right)$. Comme p_n doit rester ≥ 0 , on a nécessairement $p_n \leq \mu/\lambda$. Si on appelle u_n le rapport

²⁰On suppose $\mu, \lambda > 0$ en partant de l'idée que la population a tendance à augmenter si elle est petite et au contraire tendance à diminuer si elle est grande.

entre la population p_n et la population maximum μ/λ , on trouve l'équation logistique standard : $u_{n+1} = \mu u_n(1 - u_n)$. On note que cette équation ne dépend plus que de μ .

Comme u_n doit rester compris entre 0 et 1, puisque c'est le rapport entre la population et la population maximale, on vérifie qu'il faut que μ soit compris entre 0 et 4, voir plus loin Figure 3.

1.6 Le continu, c'est plus simple que le discret

Cette affirmation, que Devaney attribue à Birkhoff, est clairement vérifiée par les modèles logistiques. La question que je pose ici est de tenter de comprendre comment et pourquoi.

1.6.1 Côté résolution

Une différence évidente entre les deux types d'équations est l'aspect calculatoire. En effet, on a vu que l'équation logistique continue se résout sans peine. Il n'y a rien de tel avec les suites récurrentes logistiques. La raison est sans doute qu'il est plus facile de calculer une primitive que la somme des termes d'une suite.

1.6.2 Côté dynamique

On s'intéresse à des équations $p' = f(p)$ ou $u_{n+1} - u_n = f(u_n)$. Supposons pour simplifier que f est polynomiale (si on est en temps petit, c'est approximativement vrai en vertu de la formule de Taylor).

Dans le cas continu, $p' = f(p)$, les positions d'équilibre du système correspondent aux zéros de f . En effet, si on a $f(a) = 0$, la fonction $p(t) = a$ est solution de l'équation (puisque on a $p' = f(a) = 0$). Au voisinage de a , on a le développement $f(p) = f(a) + (p - a)f'(a) + o(p - a)$. Si on travaille en temps petit, p est voisin de a et l'équation est voisine de $p' = f'(a)(p - a)$. Si on pose $z = p - a$ on a donc $z(t) = \lambda e^{f'(a)t}$. Si $f'(a)$ est négatif on a une solution stable, sinon une solution instable.

Si f est une fonction polynomiale avec des racines simples, comme la fonction logistique, elle est alternativement croissante puis décroissante, et on a une succession de points d'équilibres qui sont alternativement stables et instables, donc un comportement très simple de p : pour les valeurs du temps situées entre deux racines consécutives a et b de f , p se rapproche de celle des deux positions qui correspond à un équilibre stable. Par exemple, si on a $f(p) = \lambda(p - a_1) \cdots (p - a_n)$, avec $\lambda > 0$ et $a_1 < \cdots < a_n$, on a

$f'(a_j) = \lambda \prod_{i \neq j} (a_i - a_j)$ et la position d'équilibre a_j est stable si et seulement si $n - j$ est impair.

Dans le cas discret, avec la même fonction f , et la relation $u_{n+1} - u_n = f(u_n)$, ce qui compte ce sont encore les zéros de f , c'est-à-dire les points fixes a de $g(x) = x + f(x)$. Mais ce qui gouverne la stabilité, c'est la position par rapport à 1 de la dérivée en a de g : si on a $|g'(a)| < 1$ le point fixe est stable (on dit attractif) si on a $|g'(a)| > 1$ il est instable (ou répulsif). Comme on a $g'(a) = 1 + f'(a)$ on retrouve les conditions $f'(a) > 0$ ou < 0 , mais il apparaît aussi la condition $g'(a) > -1$, donc $f'(a) > -2$.

Cette fois, avec une fonction polynomiale f et des racines a_1, \dots, a_n , on a alternance de valeurs positives et négatives de $f'(a_i)$, mais ces dernières ne sont pas nécessairement attractives (pour cela il faut avoir $f'(a_i) > -2$). Autrement dit, dans ce cas, on peut tout à fait avoir deux points fixes répulsifs de suite. C'est cela qui provoque les phénomènes que nous allons étudier, et notamment la périodicité²¹ et le chaos.

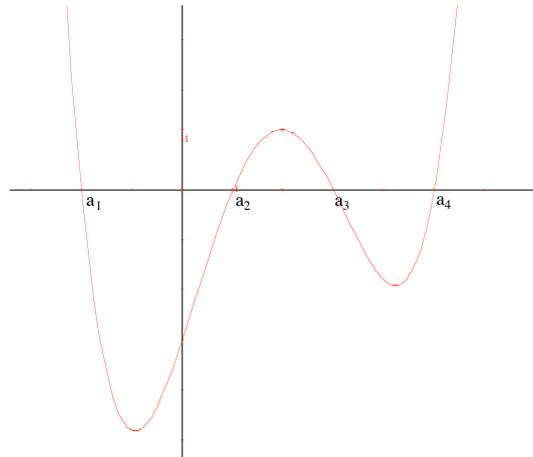


Figure 2: Alternance de points d'équilibre stable et instable

2 La suite logistique : première exploration

Dans tout ce qui suit on pose $I = [0, 1]$.

²¹Il y a un domaine où la périodicité est la norme et où l'apparition du chaos est porteuse de difficultés, c'est la physiologie. On consultera le livre [Glass-Mackey] pour voir apparaître des phénomènes chaotiques, peut-être modélisables par des équations logistiques, dans le fonctionnement des stimulateurs cardiaques (pacemakers), domaine dans lequel la périodicité est éminemment souhaitable.

À partir de maintenant, on étudie la suite logistique introduite ci-dessus. Il s'agit de la suite (u_n) , définie par récurrence par une valeur initiale $u_0 \in I$ et par la formule $u_{n+1} = \mu u_n(1 - u_n)$ avec $0 < \mu \leq 4$. De manière plus savante, on parle aussi de l'étude du système dynamique défini par la fonction $f(x) = \mu x(1 - x)$, c'est-à-dire de l'étude de la suite des itérés $x, f(x), f(f(x)), f^n(x)$ pour $x \in I$ (on a noté $f^n = f \circ f \circ \dots \circ f$, n fois). L'ensemble de tous ces transformés est l'**orbite** de x . Ce qu'il faut garder en tête, lorsqu'on pense système dynamique, c'est qu'on va faire varier le x , ou le u_0 , ainsi que le paramètre μ , l'objectif étant d'étudier le comportement asymptotique d'un x générique²², pour un μ générique. Il y a d'ailleurs deux sens possibles à ce mot : celui des topologues, pour lequel générique signifie "sur un ouvert dense" et celui des ergodiciens, pour lequel il veut dire "hors d'un ensemble de mesure nulle". Nous reviendrons sur ce point plus loin.

Il faut comprendre aussi que les calculs explicites des itérés de f deviennent très vite compliqués. De plus, le fait que la dynamique soit parfois chaotique fait que les calculs approchés sont rapidement suspects. C'est ce qui conduit à préférer un traitement qualitatif et géométrique.

2.1 La fonction f et ses points fixes

2.1.1 La fonction f

Soit μ un nombre réel appartenant à $]0, 4]$. On pose $f_\mu(x) = \mu x(1 - x)$. Lorsqu'il n'y a pas de risque de confusion, on note même $f = f_\mu$. On a $f'(x) = \mu(1 - 2x)$. La dérivée est nulle au point $1/2$, positive auparavant, négative après. Le graphe de f est une parabole dont le sommet est le point $(1/2, \mu/4)$. On note qu'on a $f(0) = f(1) = 0$. L'image de I est le segment $[0, \mu/4]$. En particulier, comme on a $\mu \leq 4$, l'intervalle I est stable.

2.1.2 Points fixes et convergence : quelques rappels

2.1 Proposition. Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction de classe C^1 .

1) Si une suite récurrente $u_{n+1} = f(u_n)$ converge vers un point l c'est un point fixe de f (i.e., on a $f(l) = l$).

Dans la suite de cette proposition, l désigne un point fixe de f .

2) Si on a $|f'(l)| > 1$ (**point fixe répulsif**), une suite récurrente $u_{n+1} = f(u_n)$ ne peut converger vers l que si elle est constante et égale à l à partir d'un certain rang.

3) Si on a $|f'(l)| < 1$ (**point fixe attractif**), il existe un intervalle J contenant l , stable par f tel que pour tout $u_0 \in J$, la suite récurrente

²²Les ergodiciens disent plutôt typique.

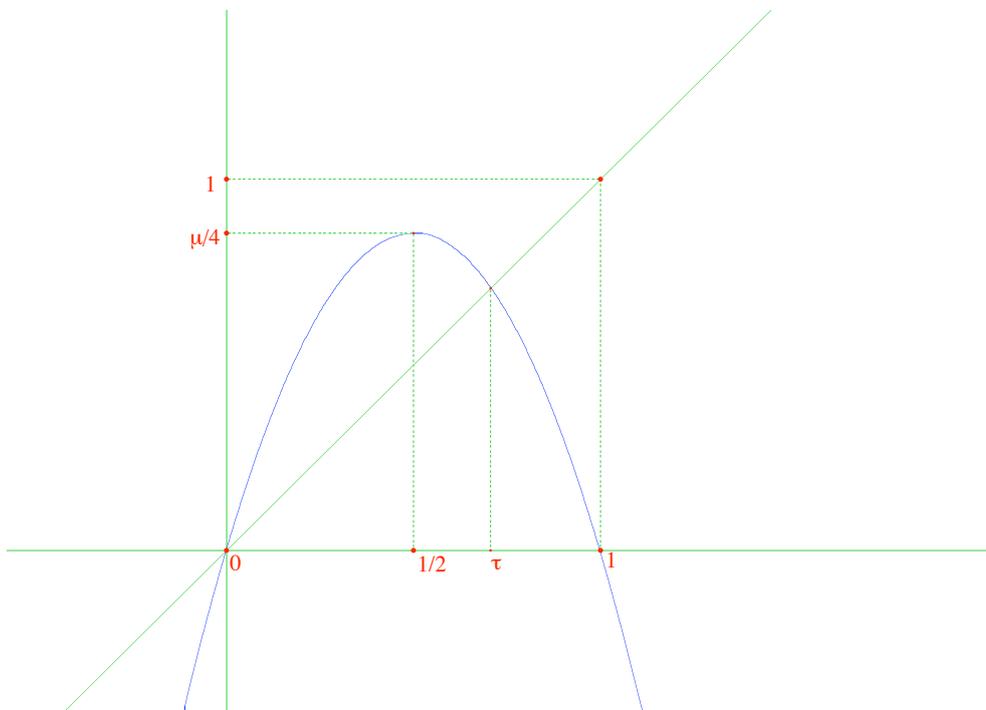


Figure 3: Le graphe de f_μ

$u_{n+1} = f(u_n)$ associée converge vers l et on a une inégalité $|u_n - l| \leq k^n |u_0 - l|$ avec $0 < k < 1$. Le point l sera dit **super-attractif** si $f'(l)$ est nul. Dans ce cas, la convergence de la suite (u_n) est rapide²³.

4) Si on a $|f'(l)| = 1$ (**point fixe parabolique**), la convergence de la suite est possible, mais non assurée, et si elle a lieu, elle est lente (en $n^{-\alpha}$).

5) Si $f'(l)$ est > 0 , la suite est monotone à partir d'un certain rang (convergence **en escalier**). Au contraire, si $f'(l)$ est < 0 les suites des termes pairs et impairs sont monotones de monotonies opposées. On parle d'une convergence **en escargot**.

Démonstration. C'est essentiellement l'inégalité des accroissements finis.

2.1.3 Les points fixes de f

Les points fixes de f sont 0 et $\tau = \frac{\mu - 1}{\mu} = 1 - \frac{1}{\mu}$. Ce dernier point est dans I et non nul si et seulement si on a $\mu > 1$.

Notons qu'on a $f'(0) = \mu > 0$ et $f'(\tau) = 2 - \mu$. Il en résulte que 0 est un point fixe attractif si et seulement si μ est < 1 . Il est répulsif pour $\mu > 1$

²³Précisément quadratique : $|u_n - l| \leq k^{2^n}$ avec $0 < k < 1$

et douteux pour $\mu = 1$. Du côté de τ , on note déjà que $f'(\tau)$ est positif pour $\mu < 2$, nul pour $\mu = 2$ et négatif au-delà. Le point τ est attractif pour $1 < \mu < 3$, douteux pour $\mu = 1$ ou 3 , répulsif pour $\mu > 3$. En particulier, au-delà de 3 , les deux points fixes de f sont répulsifs.

2.1.4 Tomber sur les points fixes ?

Lorsque les points fixes sont répulsifs, la suite (u_n) ne peut converger que si elle tombe sur un point fixe. Le point fixe 0 a pour antécédents 0 et 1 et 1 n'est dans l'image de f que pour $\mu \geq 4$. En ce qui concerne τ , en revanche, ses antécédents sont τ et $1/\mu$ et on peut continuer en cherchant les antécédents de $1/\mu$ grâce aux deux fonctions réciproques partielles de f :

$$g_1(y) = \frac{\mu - \sqrt{\mu^2 - 4\mu y}}{2\mu} \quad \text{et} \quad g_2(y) = \frac{\mu + \sqrt{\mu^2 - 4\mu y}}{2\mu},$$

définies sur $[0, \mu/4]$ et respectivement à valeurs dans $[0, 1/2]$ et $[1/2, 1]$. On voit que les points $g_1(1/\mu)$ et $g_2(1/\mu)$ existent dès que $\mu \geq 2$. De plus, on peut chercher des antécédents des points ainsi obtenus en itérant les fonctions g_i . Le lecteur montrera la proposition suivante :

2.2 Proposition.

1) On suppose $\mu \geq 2$. On définit par récurrence une suite (v_n) de points de $[0, 1/2]$ en posant $v_0 = 1/\mu$ et $v_{n+1} = \frac{\mu - \sqrt{\mu^2 - 4\mu v_n}}{2\mu}$. On a $f(v_{n+1}) = v_n$, de sorte que si on prend $u_0 = v_n$, on a $u_p = \tau$ pour $p > n$. La suite (v_n) décroît et converge vers 0 .

2) Soit $\alpha \sim 3.67857351043$ la racine différente de 2 du polynôme $\mu^4 - 4\mu^3 + 16$. On suppose $\mu > \alpha$. On pose $w_0 = 1/\mu$ et on définit par récurrence une suite (w_n) de points qui sont dans $[1/2, \mu/4]$ (pour $n \geq 1$) en posant $w_{n+1} = \frac{\mu + \sqrt{\mu^2 - 4\mu w_n}}{2\mu}$. On a $f(w_{n+1}) = w_n$, de sorte que si on prend $u_0 = w_n$, on a $u_p = \tau$ pour $p > n$. La suite (w_n) converge en escargot vers τ .

2.3 Remarque. Il y a beaucoup d'autres valeurs qui tombent sur τ , obtenues à partir de $1/\mu$ en appliquant alternativement g_1 et g_2 , mais il faut prendre garde au domaine de définition de ces fonctions.

2.2 Les cas de convergence : $0 < \mu \leq 3$

Dans ce qui suit, on montre le théorème suivant :

2.4 Théorème. *On suppose $0 < \mu \leq 3$. Alors, pour toute valeur initiale $u_0 \in I$, la suite (u_n) est convergente.*

En réalité, ce résultat est une conséquence évidente du théorème de Coppel 3.4. En effet, on vérifie que, pour $\mu \leq 3$, la fonction f_μ n'admet pas de point de période 2 (voir 3.8) et 3.4 montre alors que toute suite récurrente de premier terme $u_0 \in I$ converge. Cela étant, nous en donnons une preuve directe, inspirée de [Devaney], en distinguant les cas.

2.2.1 Le cas $0 < \mu < 1$

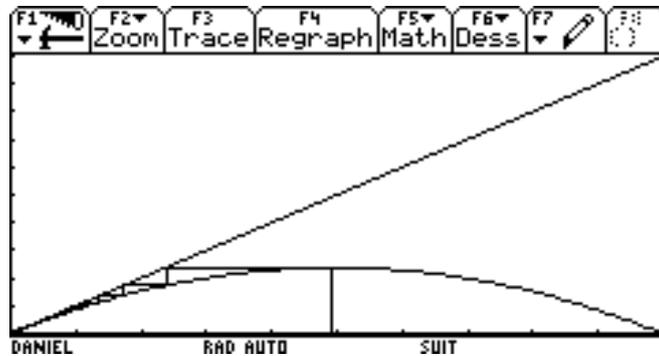


Figure 4: La convergence pour $\mu < 1$

Il y a un unique point fixe dans I , le point 0, qui est attractif. Le lecteur prouvera sans peine le théorème suivant :

2.5 Théorème. *On suppose $0 < \mu < 1$. Soit $u_0 \in I$.*

- 1) *Pour $u_0 = 0$ la suite est constante. Pour $u_0 = 1$ elle est constante et égale à 0 à partir du rang 1.*
- 2) *On suppose $u_0 \in]0, 1/2]$. La suite (u_n) décroît et converge géométriquement vers 0.*
- 3) *On suppose $u_0 \in]1/2, 1[$. On a $u_1 \in]0, 1/2]$ et on est ramené au cas 2).*

2.2.2 Le cas $\mu = 1$

Pour $u_0 \in I$, on montre encore que la suite converge vers 0, mais la convergence est en $1/n$.

2.2.3 Le cas $1 < \mu < 2$

Le cas $1 < \mu < 2$ se ramène au cas $0 < \mu < 1$ en vertu du lemme suivant :

2.6 Lemme. On pose $\nu = 2 - \mu$. La fonction $f(x) = \mu x(1 - x)$ est conjuguée de $g(x) = \nu x(1 - x)$ par la fonction h définie par $h(x) = \frac{\mu - 1 - \mu x}{\mu - 2}$.

Démonstration. On peut vérifier la formule $hf = gh$ ou noter simplement que h fixe le point $1/2$ et qu'elle envoie 0 sur $1 - \frac{1}{\nu}$ et $1 - \frac{1}{\mu}$ sur 0 . Pour un résultat plus général voir annexe 7.1.

La formule de conjugaison ne concerne des paramètres $\mu > 0$ que dans l'intervalle $]0, 2[$. Dans ce cas la dynamique associée à f est la même que celle associée à g , précisément :

2.7 Théorème. On suppose $1 < \mu < 2$ (donc $0 < \tau < 1/2$). Soit $u_0 \in I$.

- 1) Pour $u_0 = 0$ ou $u_0 = \tau$, la suite est constante. Pour $u_0 = 1$ (resp. $u_0 = 1/\mu$), elle est constante et égale à 0 (resp. τ) à partir du rang 1.
- 2) On suppose $u_0 \in]0, \tau[$ (resp. $u_0 \in]\tau, 1/2[$). La suite (u_n) est croissante (resp. décroissante) et converge vers τ .
- 3) On suppose $u_0 \in]1/\mu, 1[$ (resp. $u_0 \in]1/2, 1/\mu[$). Alors on a $u_1 \in]0, \tau[$ (resp. $]\tau, 1/2[$) et on est ramené au cas 2).

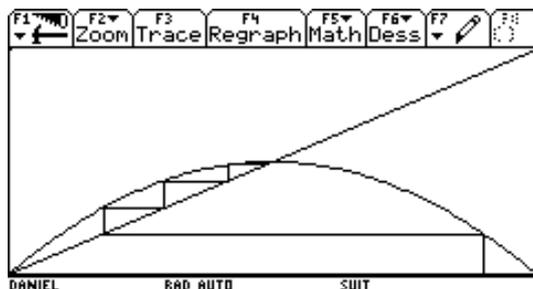


Figure 5: La convergence pour $1 < \mu < 2$

2.2.4 Le cas $\mu = 2$

Cette fois, on a $\tau = 1/2$ et $f'(\tau) = 0$ de sorte que τ est un point super-attractif.

2.2.5 Le cas $2 < \mu < 3$

Le point τ est encore attractif, mais avec une dérivée négative. On a donc, au moins au voisinage de τ , un comportement “en escargot”.

2.8 Proposition. *Soit $u_0 \in]0, 1[$. La suite (u_n) converge vers τ .*

Démonstration. C’est le point le plus délicat. La preuve que nous en donnons est essentiellement celle du livre de Devaney.

On pose $\hat{\tau} = 1/\mu$ et on a $f(1/\mu) = \tau$. Avec l’hypothèse sur μ , on a $0 < \hat{\tau} < 1/2 < \tau < 1$. Posons $J = [\hat{\tau}, \tau]$ et montrons que $f^2(J)$ est contenu dans $[1/2, \tau]$. Comme f atteint son maximum en $1/2$, il est clair qu’on a $f(J) = [\tau, \mu/4]$ avec $\tau < \mu/4$. On a donc $f^2(J) = f([\tau, \mu/4]) = [f(\mu/4), \tau]$ et on a $f(\mu/4) < \tau$ car f est décroissante sur $[1/2, 1]$. Il reste à montrer que $f(\mu/4)$ est bien $\geq 1/2$. Cela équivaut à $-\mu^3 + 4\mu^2 - 8 \geq 0$ et une rapide étude de cette fonction l’atteste.

On en déduit d’abord que, si u_0 est dans J , la suite u_{2n} converge vers τ . En effet, on a alors $u_2 \in [1/2, \tau]$ et cet intervalle est stable par f^2 . Comme f^2 est croissante sur cet intervalle, la suite u_{2n} est croissante et majorée par τ . Elle converge vers un point fixe de f^2 et on vérifie que ce ne peut être que τ , voir 3.8.

On en déduit ensuite que (u_n) converge vers τ . En effet, on sait que la suite u_{2n} converge vers τ et la suite u_{2n+1} aussi, comme on le voit en appliquant f .

Supposons maintenant $u_0 < \hat{\tau}$ et montrons qu’il existe k tel que u_k soit dans J . Cela montrera que la suite converge vers τ en vertu du premier cas. Comme on a $f(x) > x$ pour $x < \tau$, on a $u_0 < u_1$ et la suite u_n croît tant qu’elle reste $< \tau$. Si elle est majorée par $\hat{\tau}$, elle converge, mais cela ne peut être que vers τ et c’est absurde. Sinon, soit $k - 1$ le dernier indice tel que $u_{k-1} < \hat{\tau}$. On a alors $\hat{\tau} \leq u_k = f(u_{k-1}) < f(\hat{\tau}) = \tau$ car u_{k-1} et $\hat{\tau}$ sont dans $[0, 1/2]$ sur lequel la fonction f est croissante.

Supposons enfin $u_0 \in]\tau, 1[$. Comme f est décroissante, on a $u_1 \in f(]\tau, 1[) =]0, \tau[$ et on est ramené à l’un des cas précédents.

2.9 Remarque. Pour $\mu = 3$, la suite converge vers τ (sauf si u_0 est nul) en vertu de Coppel.

3 Le doublement de période

Pour $\mu > 3$, les points fixes de $f = f_\mu$ sont tous deux répulsifs. Dans ce cas, cf. 2.1, une suite récurrente associée à f ne peut converger vers l’un

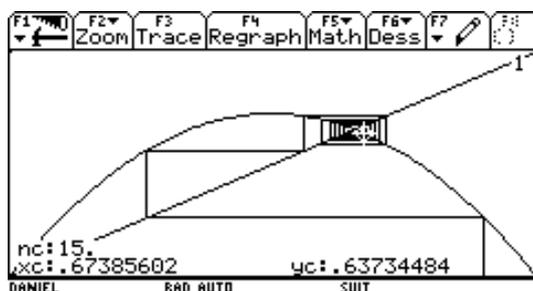


Figure 6: La convergence pour $2 < \mu < 3$

de ces points fixes que si elle est stationnaire. Comme f est polynomiale, cela ne peut concerner qu'un nombre dénombrable de points de départ u_0 , de sorte que la plupart des suites (u_n) associées à f vont maintenant être divergentes. Cependant, la manière dont s'opère cette divergence va se révéler très diverse. Le premier phénomène qui apparaît est l'existence de points périodiques attractifs (ou de cycles attractifs) à commencer par des points de période 2. En effet, ceux-ci sont un point de passage obligé en vertu du théorème de Coppel que nous prouvons maintenant et qui éclaire beaucoup la situation.

3.1 Apparition des 2-cycles : le théorème de Coppel

3.1.1 Cycles

3.1 Définition. Soit $I = [a, b]$ un intervalle et soit $f : I \rightarrow I$ une fonction continue. Un point $x \in I$ est dit **n -périodique** pour f s'il vérifie $f^n(x) = x$ et pas $f^k(x) = x$ pour $0 < k < n$. Dans ce cas, on dit encore que l'ensemble $x_0 = x, x_1 = f(x), \dots, x_k = f^k(x), \dots, x_{n-1} = f^{n-1}(x)$ est un **n -cycle** pour f .

On note $Df^n(x)$ ou $(f^n)'(x)$ la dérivée de f^n au point x .

3.2 Proposition-Définition. Soit $f : I \rightarrow I$ une fonction de classe C^1 et soit x_0, x_1, \dots, x_{n-1} un n -cycle pour f . On a, pour tout i , $Df^n(x_i) = f'(x_0)f'(x_1) \cdots f'(x_{n-1})$. On dit que le n -cycle est **attractif** si cette quantité est < 1 en valeur absolue. Si tel est le cas, et si u_0 est suffisamment voisin de l'un des points du cycle, il existe un entier $k \in [0, n - 1]$ tel que les suites $u_{np}, u_{np+1}, \dots, u_{np+n-1}$ convergent respectivement, quand p tend vers l'infini, vers $x_k, x_{k+1}, \dots, x_{k+n-1}$ (les indices étant pris modulo n). Autrement dit, les termes de la suite se rapprochent alternativement des n valeurs des points du cycle.

Le cycle sera dit **super-attractif** si l'on a $Df^n(x_i) = 0$. La convergence de la suite (u_n) vers le cycle est alors une convergence rapide.

Démonstration. C'est encore l'inégalité des accroissements finis, appliquée cette fois à f^n .

Le lemme suivant donne une méthode pour construire des valeurs de μ qui donnent des cycles super-attractifs :

3.3 Lemme. *On suppose qu'on a $f = f_\mu$. Un n -cycle a_1, \dots, a_n relatif à f (i.e. vérifiant $a_{i+1} = f(a_i)$ avec les indices pris modulo n) est super-attractif si et seulement s'il contient $1/2$. Les valeurs de μ qui donnent des n -cycles super-attractifs sont obtenues en résolvant l'équation polynomiale en μ , $f_\mu^n(1/2) = 1/2$.*

Démonstration. En effet, on a $Df^n(a_i) = 0 = f'(a_1) \cdots f'(a_n)$ et le seul point où f' s'annule est $1/2$.

3.1.2 Le théorème de Coppel

3.4 Théorème. *Soit $I = [a, b]$ un intervalle compact et soit $f : I \rightarrow I$ une fonction continue. On suppose que f n'admet pas de 2-cycle. Alors, pour tout $x_0 \in I$, la suite récurrente $x_{n+1} = f(x_n)$ converge.*

3.5 Remarque. La réciproque de ce résultat est évidemment vraie car, si f admet un 2-cycle : $f(x) = y$ et $f(y) = x$ avec $x \neq y$, la suite récurrente définie par $x_0 = x$ est divergente.

Démonstration. Voir [Coppel] pour des précisions.

La preuve repose sur le lemme suivant :

3.6 Lemme. *Avec les notations du théorème, on suppose que f n'admet pas²⁴ de 2-cycle. Soit $c \in I$ et supposons $f(c) > c$ (resp. $f(c) = c$, resp. $f(c) < c$). Alors on a $f^n(c) > c$ (resp. $f^n(c) = c$, resp. $f^n(c) < c$) pour tout $n \in \mathbf{N}^*$.*

Démonstration. (du lemme) Le cas $f(c) = c$ est évident. Supposons qu'on ait $f(c) > c$. Je me contente de montrer qu'on a aussi $f^2(c) > c$, voir [Coppel] pour le reste. Sinon, on a $f^2(c) < c$ (en effet, $f^2(c) = c$ implique $f(c) = c$ et c'est absurde). Comme on a $f^2(a) \geq a$, il existe un point fixe de f^2 dans $[a, c[$. Prenons le plus grand et appelons le d . C'est aussi un point fixe de f . Dans $]d, c[$ il n'y a plus de point fixe de f^2 , ni de f . Par conséquent f et f^2 , dans cet

²⁴Cela signifie que, si on a $f^2(x) = x$ on a déjà $f(x) = x$.

intervalle, sont “comme en c ”, autrement dit, pour $d < x < c$, on a $f(x) > x$ et $f^2(x) < x$. On choisit alors e avec $d < e < c$ vérifiant $d < e < f(e) < c$. C’est possible car on a $f(d) = d < c$ donc encore $f(e) < c$ pour e voisin de d . Comme $f(e)$ est aussi entre d et c on a $f(f(e)) = f^2(e) > f(e) > e$, mais c’est une contradiction.

On peut alors finir la preuve du théorème de Coppel. Soit $x_0 \in I$ quelconque. S’il existe n tel que $x_{n+1} = x_n$, la suite est stationnaire donc convergente. On peut donc supposer qu’on a $x_n \neq x_{n+1}$ pour tout n . Si on a, disons, $x_n < x_{n+1}$ à partir d’un certain rang, la suite est croissante et majorée, donc convergente. On peut donc supposer qu’il existe une infinité d’entiers n tels que $x_n < x_{n+1}$ et une infinité tels que $x_n > x_{n+1}$. Appelons (y_n) (resp. (z_n)) la sous-suite des premiers (resp. des seconds). Ces deux sous-suites ont pour réunion la suite (x_n) . Si on a $y_p = x_n$, on a $x_{n+1} > x_n$ et, en vertu du lemme, $x_{n+k} > x_n$ pour tout $k > 0$. En particulier, on a $y_{p+1} > y_p$, de sorte que la suite (y_p) est croissante. On montre de même que la suite (z_q) est décroissante. Comme ces suites sont confinées dans $[a, b]$, elles convergent respectivement vers y et z . Comme la suite des z_q est infinie, il existe une infinité d’indices n tels que x_n soit un y_p mais x_{n+1} un z_q . Appelons (w_p) la sous-suite des y_p en question. Elle converge aussi vers y et, comme $f(w_p)$ est une sous-suite des z_q , elle converge vers z . On en déduit $f(y) = z$. Le même raisonnement (avec d’autres sous-suites, bien entendu) donne $f(z) = y$, donc $f^2(y) = y$. Mais, comme f n’a pas de 2-cycle, on en déduit $f(y) = y = z$ et les deux sous-suites (y_p) et (z_q) ont même limite. Comme elles épuisent (x_n) , cette dernière suite converge vers y .

3.7 Corollaire. *Soit $f : I \rightarrow I$ une fonction de classe C^1 . On suppose que les points fixes de f sont en nombre fini et répulsifs et que les fibres²⁵ de f sont finies. Alors, f admet un 2-cycle.*

Démonstration. En effet, sinon, toutes les suites récurrentes convergent vers des points fixes. Comme ceux-ci sont répulsifs, les suites sont constantes à partir d’un certain rang. Mais les points qui tombent sur un point fixe sont en nombre dénombrable et c’est absurde.

3.2 Les points fixes de f^2

3.2.1 Détermination des cycles d’ordre 2

Vu le théorème de Coppel, les points fixes de f^2 vont jouer un rôle essentiel :

²⁵C’est-à-dire les ensembles $f^{-1}(\{x\})$.

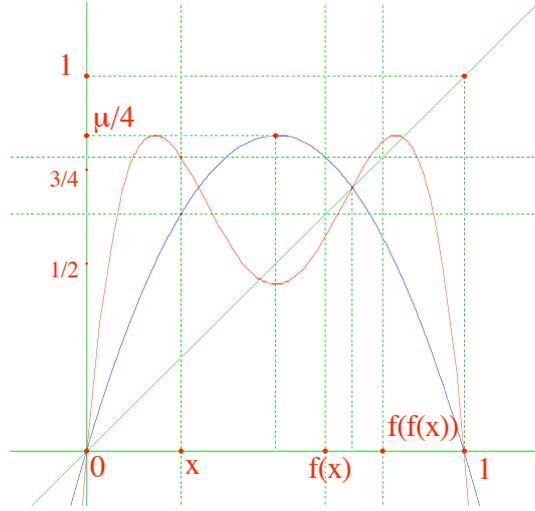


Figure 7: Les points fixes de f^2

3.8 Proposition. On a $f^2(x) - x = -x(\mu x - \mu + 1)(\mu^2 x^2 - \mu(\mu + 1)x + \mu + 1)$. Les points fixes de f^2 autres que 0 et τ sont les racines de l'équation $\mu^2 x^2 - \mu(\mu + 1)x + \mu + 1 = 0$. Il n'y en a que pour $\mu > 3$. Ces points fixes sont alors :

$$x_1 = \frac{1}{2} + \frac{1}{2\mu} - \frac{\sqrt{(\mu+1)(\mu-3)}}{2\mu} \quad \text{et} \quad x_2 = \frac{1}{2} + \frac{1}{2\mu} + \frac{\sqrt{(\mu+1)(\mu-3)}}{2\mu}.$$

Ces points sont dans I , précisément, le point x_2 est toujours $> 1/2$, tandis que x_1 est $< 1/2$, $= 1/2$ ou $> 1/2$, selon que μ est $> 1 + \sqrt{5}$, $= 1 + \sqrt{5}$ ou $< 1 + \sqrt{5}$.

Démonstration. Le calcul est facile. Le discriminant du trinôme du second degré est $\mu^2(\mu+1)(\mu-3)$, d'où le résultat (on note que pour $\mu = 3$ la racine double n'est autre que τ et que c'est le seul cas où τ est racine du trinôme).

3.9 Remarque. On voit, sur cet exemple, la première manifestation du phénomène de **bifurcation**. Pour $\mu < 3$ il n'y a pas de 2-cycle, tandis qu'il y en a pour $\mu > 3$. Lorsque μ tend vers 3 par valeurs supérieures, on voit, sur l'expression des racines ci-dessus, que x_1 et x_2 tendent vers toutes deux vers le point fixe $2/3$ de f_3 .

3.2.2 Attractivité des cycles d'ordre 2

On cherche maintenant à savoir si le 2-cycle (x_1, x_2) est attractif, c'est-à-dire si les termes d'une suite récurrente vont se rapprocher alternativement

de x_1 et x_2 . Cela signifie que la suite des termes pairs u_{2p} (resp. des termes impairs u_{2p+1}) converge par exemple vers x_1 (resp. x_2). Comme ces deux suites sont des suites récurrentes associées à f^2 , il s'agit de voir si les points fixes x_1, x_2 de f^2 sont attractifs²⁶ pour f^2 , donc si la dérivée de f^2 en x_i est plus petite que 1 en valeur absolue, voir 3.2 :

3.10 Proposition. *On suppose $\mu \geq 3$. Soit σ l'un quelconque des points fixes de f^2 ($\neq 0, \tau$). On a $(f^2)'(\sigma) = -\mu^2 + 2\mu + 4$. Cette quantité est positive ou nulle pour $3 \leq \mu \leq 1 + \sqrt{5} \simeq 3,236$ et elle est alors ≤ 1 (et même < 1 pour $\mu > 3$). On a donc un cycle d'ordre 2 attractif. Elle est nulle pour $\mu = 1 + \sqrt{5}$: le cycle d'ordre 2 (formé de $1/2$ et de $(1 + \sqrt{5})/4$) est alors superattractif. Elle est négative pour $\mu > 1 + \sqrt{5}$ et elle est alors > -1 (donc avec un cycle d'ordre 2 attractif) pour $\mu < 1 + \sqrt{6} \simeq 3,449$.*

Démonstration. Si $\bar{\sigma}$ est l'autre point du cycle, on a $(f^2)'(\sigma) = f'(\bar{\sigma})f'(\sigma)$ et un petit calcul de fonctions symétriques des racines conclut.

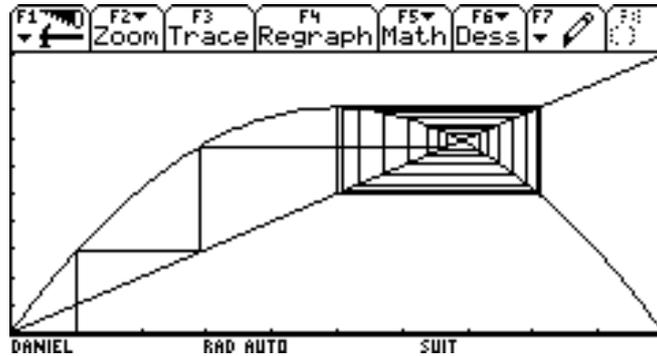


Figure 8: Cas d'un 2-cycle attractif

3.11 Corollaire. *On suppose $3 < \mu < 1 + \sqrt{6}$ et $u_0 \neq 0, 1, \tau, 1/\mu$. Soit (u_n) la suite récurrente associée à u_0 . Alors, les suites (u_{2p}) et (u_{2p+1}) convergent vers les points fixes de f^2 .*

Démonstration. Pour être sûr qu'il n'y a pas d'exceptions, voir le paragraphe suivant.

²⁶En vertu de 3.7, ils le sont tant qu'il n'y a pas de cycles d'ordre 4.

3.3 Les points fixes de f^4

Lorsque le cycle d'ordre 2 cesse d'être attractif, le théorème de Coppel, appliqué à f^2 , montre qu'il y a un cycle d'ordre 2 pour f^2 , donc d'ordre 4 pour f . À partir de maintenant, on a à résoudre des équations algébriques de plus en plus compliquées et on va utiliser la calculatrice.

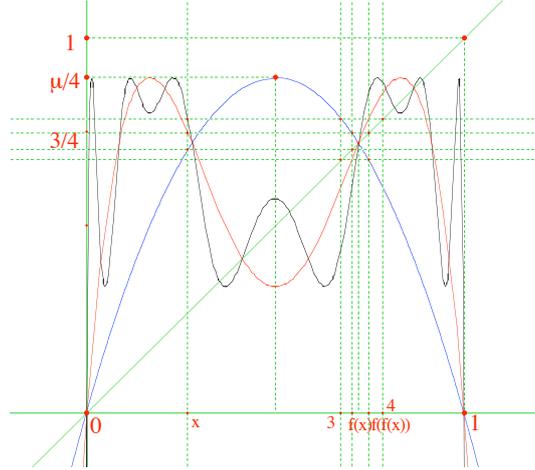


Figure 9: Les graphes de f , f^2 et f^4

3.12 Proposition. Pour $\mu \leq 1 + \sqrt{6}$, f^4 n'a pas de point fixe (autre que ceux de f et f^2). Pour $\mu > 1 + \sqrt{6} = \mu_2$, f admet un cycle d'ordre 4 unique et ce cycle est attractif tant que l'on a $\mu < \mu_3 \sim 3,544090$. Pour $\mu = \nu_2 \sim 3,49856169933$, le cycle d'ordre 4 est super-attractif.

Démonstration. On commence par calculer (avec la calculatrice ou un logiciel de calcul formel, sauf si l'on a des insomnies) le polynôme qui donne les 4-cycles. Il s'agit de $f^4(x) - x$, divisé par $f^2(x) - x$. On obtient un polynôme $Q_\mu(x)$ de degré 12 en x :

$$\begin{aligned} & \mu^{12}x^{12} - 6\mu^{12}x^{11} + 3\mu^{11}(5\mu + 1)x^{10} - \mu^9(20\mu^3 + 15\mu^2 + 1)x^9 \\ & + 3\mu^9(5\mu^3 + 10\mu^2 + \mu + 2)x^8 - 2\mu^8(3\mu^4 + 15\mu^3 + 6\mu^2 + 7\mu + 1)x^7 \\ & + \mu^6(\mu^6 + 15\mu^5 + 18\mu^4 + 17\mu^3 + 10\mu^2 + 1)x^6 - \mu^6(3\mu^5 + 12\mu^4 + 12\mu^3 + 18\mu^2 + \mu + 4)x^5 \\ & + \mu^5(3\mu^5 + 5\mu^4 + 14\mu^3 + 4\mu^2 + 6\mu + 2)x^4 - \mu^3(\mu^6 + 4\mu^5 + 5\mu^4 + 4\mu^3 + 5\mu^2 + 1)x^3 \\ & + \mu^3(2\mu^4 + \mu^3 + 4\mu^2 + \mu + 2)x^2 - \mu^2(\mu^3 + \mu^2 + \mu + 1)x + \mu^2 + 1. \end{aligned}$$

Il n'est pas évident de montrer les assertions de la proposition. Avec une calculatrice ou un ordinateur, on peut calculer les racines de Q_μ de manière approchée lorsque μ est fixé. Par exemple, pour $\mu = 3,45$, on trouve quatre racines : 0,852428087575, 0,433991655785, 0,847467675925, et 0,445967663053 qui forment un cycle d'ordre 4 attractif²⁷.

On peut alors trouver, expérimentalement, les valeurs μ_2 et μ_3 . Pour montrer que la valeur limite μ_2 est exactement $1 + \sqrt{6}$, il faut comprendre le phénomène de bifurcation (ou de ramification) illustré par la figure 10. Décrivons le phénomène. Pour $\mu < \mu_2$ il n'y a pas de 4-cycle, en revanche, il y en a pour $\mu > \mu_2$. Dans ce cas, appelons $x_1(\mu) < x_2(\mu) < x_3(\mu) < x_4(\mu)$ un tel cycle. Lorsque μ tend vers μ_2 , x_1 et x_2 tendent vers un même point y_1 et x_3 et x_4 vers un même point y_2 , (y_1, y_2) étant un 2-cycle de f_{μ_2} . Autrement dit, pour $\mu = \mu_2$, le polynôme Q_μ admet en facteur le polynôme $F_\mu(x) = \mu^2 x^2 - \mu(\mu + 1)x + \mu + 1$ qui donne les 2-cycles.

Si l'on admet que les choses se passent bien ainsi, on trouve la valeur $1 + \sqrt{6}$: on divise Q_μ par F_μ , le reste est égal à $-\mu^2 + 2\mu + 5$ dont l'unique racine positive est $\mu = 1 + \sqrt{6}$. En fait, plus précisément, le polynôme F_μ est au carré dans Q_μ , avec comme quotient :

$$\begin{aligned} & \mu^8 x^8 + (-4\mu^8 + 2\mu^7)x^7 + (6\mu^8 - 5\mu^7 + \mu^6)x^6 + \\ & (-4\mu^8 + 3\mu^7 - 3\mu^5)x^5 + (\mu^8 + \mu^7 - 3\mu^6 + 9\mu^5 - 6\mu^4)x^4 + \\ & + (-\mu^7 + 2\mu^6 - 8\mu^5 + 10\mu^4 - 3\mu^3)x^3 + (2\mu^5 - 4\mu^4 + 2\mu^3 + 6\mu^2)x^2 \\ & + (-6\mu^2 + 12\mu)x + \mu^2 - 5\mu + 6. \end{aligned}$$

Pour voir qu'il n'y a pas de 4-cycle dans le cas $\mu = 1 + \sqrt{6}$, il reste à trouver les racines de ce polynôme²⁸ lorsque μ est égal à $1 + \sqrt{6}$. On vérifie qu'il n'y en a pas.

Pour trouver un cycle super-attractif, on applique 3.3 en résolvant en μ l'équation $Q_\mu(1/2) = 0$. On trouve la valeur de l'énoncé et une autre, 3,960270272187 que l'on reverra.

Enfin, il reste à calculer la valeur $\mu_3 \sim 3,544090$. C'est le plus difficile. Une technique est d'étudier le comportement de la suite et de repérer selon les valeurs de μ si elle admet un 4-cycle limite (pour $\mu < \mu_3$) ou un 8-cycle limite (pour $\mu > \mu_3$), mais c'est assez imprécis à cet endroit de bifurcation.

²⁷Il est commode (et facile) d'écrire un petit programme qui permet d'afficher les r termes qui suivent le millième (par exemple).

²⁸En fait, le mieux pour faire ces calculs est de conjuguer f_μ à $x^2 - 5/4$ (voir 7.1). Le calcul des racines du polynôme qui correspond à Q_μ est plus facile dans ce cas sur la calculatrice et on trouve qu'il n'a que deux racines qui sont les points du 2-cycle.

En effet, lorsque μ se rapproche de μ_3 par valeurs supérieures, le 8-cycle tend vers un 4-cycle et il faut aller très loin pour savoir si les 8 valeurs que l'on examine sont celles d'un cycle d'ordre 8 ou si elles vont se rapprocher pour donner un 4-cycle.

Une autre méthode consiste à repérer la dérivée en le cycle d'ordre 4 et à voir quand elle atteint la valeur -1 .

Par exemple, avec $\mu = 3,544089$, en partant de la valeur $x = 0,8841$ donnée par l'étude de (u_n) , on trouve, par approximations successives une valeur de l'un des points du 4-cycle à 10^{-12} près : $a = 0,88404930742571$ et, pour cette valeur on a $Df^4(a) = -0,999969$.

3.4 La zone des cycles attractifs d'ordre 2^n : $3 < \mu < \mu_\infty$

Le phénomène que nous venons de voir apparaître pour les 2-cycles et les 4-cycles continue de plus belle lorsque μ croît. On parle de **cascade** de doublement de périodes. Ce phénomène a été mis en évidence vers 1978 par Feigenbaum, qui a formulé à son sujet un certain nombre de conjectures, prouvées ensuite par plusieurs auteurs (notamment par Lanford, Collet, Eckmann, Campanino, Epstein, etc.). Il n'est pas évident de trouver un énoncé clair et complet, pour la suite logistique, dans la littérature²⁹. Voilà ce que j'ai pu rassembler :

3.13 Théorème (Feigenbaum et al.). *On considère $f_\mu(x) = \mu x(1-x)$ pour $0 \leq \mu \leq 4$.*

Soit μ_n la borne inférieure des μ tel que f_μ admette un cycle d'ordre 2^n .

0) On a $\mu_0 = 0$, $\mu_1 = 3$, $\mu_2 = 1 + \sqrt{6}$, $\mu_3 \sim 3,544090$, $\mu_4 \sim 3,564407$.

1) La suite (μ_n) est strictement croissante, majorée par 4, elle converge vers un réel $\mu_\infty \sim 3,5699456$ et la différence $\mu_\infty - \mu_n$ est équivalente, quand n tend vers l'infini, à K/δ^n où K est une constante et où δ est la constante de Feigenbaum :

$$\delta \sim 4,66921166091029906.$$

2) Pour $\mu_\infty > \mu > \mu_n$, f_μ admet³⁰ un unique 2^n -cycle, qui est attractif tant que μ est $< \mu_{n+1}$ et répulsif pour $\mu > \mu_{n+1}$.

3) Plus précisément, la dérivée $Df^n(x_i)$ en l'un des points du cycle d'ordre 2^n décroît de 1 à -1 quand μ croît de μ_n à μ_{n+1} . En particulier, il existe un unique nombre réel $\nu_n \in]\mu_n, \mu_{n+1}[$ tel que le 2^n -cycle correspondant soit super-attractif. La suite ν_n tend vers μ_∞ et on a $\mu_\infty - \nu_n \sim K'/\delta^n$. Soit

²⁹Pour la suite logistique, le plus clair est sans doute le livre [Robinson]. Pour la méthode de "renormalisation" proposée par Feigenbaum, l'article [Collet-Eckmann-Lanford] est le plus abordable.

³⁰Mais elle n'en admet pas pour $\mu = \mu_n$.

α_n le point de ce cycle, différent de $1/2$, et le plus proche de $1/2$, et posons $d_n = \frac{1}{2} - \alpha_n$. Alors, on a $\lim_{n \rightarrow +\infty} \frac{d_n}{d_{n+1}} = -\alpha$ avec $\alpha \sim 2,502907875$.

4) Pour $\mu < \mu_\infty$, f_μ n'a pas de cycle d'ordre p si p n'est pas une puissance de 2

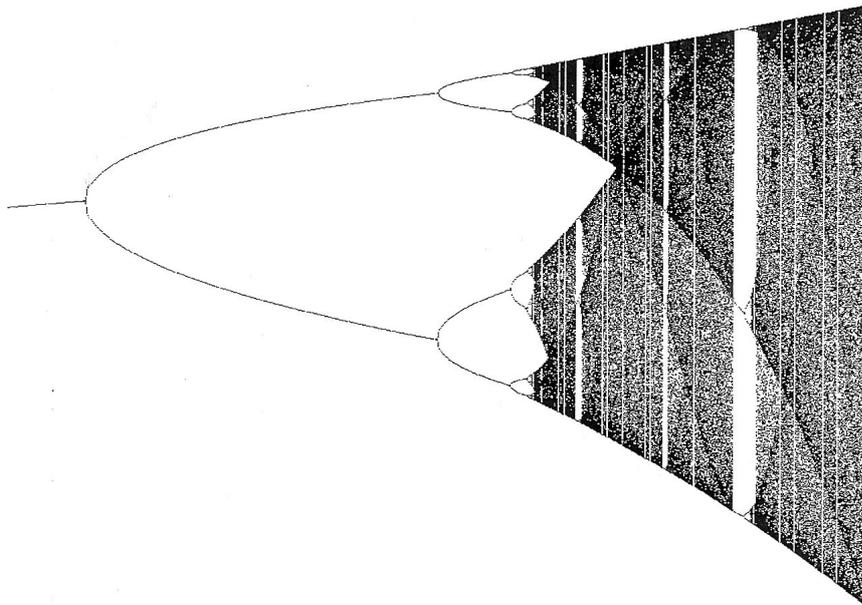


Figure 10: La cascade de doublement de périodes

Démonstration. Il n'est pas question de prouver ici le résultat de Feigenbaum. Je me contenterai de montrer les choses faciles, et d'abord un lemme :

3.14 Lemme. Soit g une fonction continue définie sur un intervalle de \mathbf{R} . Si g admet un 2-cycle, elle admet un point fixe. Si elle admet un 2^n -cycle avec $n \geq 1$, elle admet des 2^k -cycles pour tout k avec $0 \leq k \leq n - 1$.

Démonstration. Si g échange a et b avec $a \neq b$, la fonction $g(x) - x$ change de signe sur $[a, b]$, donc s'annule. La deuxième assertion en résulte par récurrence.

Maintenant, il est clair que la suite μ_n est croissante³¹. En effet, soit $\epsilon > 0$ et soit μ vérifiant $\mu_{n+1} < \mu < \mu_{n+1} + \epsilon$ tel que f_μ admette un 2^{n+1} -cycle. Alors, par le lemme, f_μ admet aussi un 2^n -cycle donc on a $\mu \geq \mu_n$. Comme cela vaut pour tout ϵ on a bien $\mu_{n+1} \geq \mu_n$.

Nous verrons dans la section suivante que f_4 admet des cycles de tous ordres, ce qui montre que la suite des μ_n est majorée par 4, donc converge.

³¹Mais un peu moins clair qu'elle l'est strictement.

3.15 Remarques.

1) Ce qui est extraordinaire dans le théorème de Feigenbaum, c'est qu'il vaut pour beaucoup d'autres familles de fonctions f que la famille logistique, voir [Collet-Eckmann-Lanford]. En ce sens, les constantes δ et α sont des constantes **universelles**.

2) Si μ est dans $]\mu_n, \mu_{n+1}[$, la fonction admet des cycles d'ordre 2^k pour tout $k \leq n$, mais d'aucun autre ordre en vertu du théorème de Sarkovsky 5.4.

3) En chaque μ_n on a une **bifurcation** : on passe d'un régime à un autre. Précisément, pour f^{2^n} on passe d'un point fixe attractif à un point fixe répulsif en passant par un point "parabolique" (avec une dérivée égale à -1 ici). On parle d'une bifurcation du genre selle-nœud.

4) Une conséquence du résultat c'est qu'on a $\lim_{n \rightarrow +\infty} \frac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} = \delta$ et la même formule avec ν_n . La convergence est assez rapide. Si on calcule ν_3 à partir de ν_1 et ν_2 avec cette formule, on trouve 3,5547, ce qui ne semble pas loin de la vérité.

3.16 Corollaire. *Pour $\mu < \mu_\infty$, toute suite récurrente associée à f converge vers un point fixe ou un cycle. Le bassin d'attraction du cycle attractif d'ordre 2^n (i.e. l'ensemble des points dont les itérés convergent vers ce cycle) est partout dense.*

Démonstration. Si on a $\mu_n \leq \mu < \mu_{n+1}$, il n'y a pas de cycle d'ordre 2^{n+1} pour f , donc pas de cycle d'ordre 2 pour f^{2^n} . En vertu du théorème de Coppel, cela montre que toute suite récurrente associée à f^{2^n} converge, donc que la suite associée à f converge vers un cycle d'ordre 2^i avec $i \leq n$. De plus, comme le cycle d'ordre 2^n est le seul à être attractif, les autres donnent des suites périodiques stationnaires, en nombre dénombrable, de sorte que les points dont l'orbite tend vers le 2^n -cycle sont partout denses.

Plutôt que de poursuivre l'exploration dans la zone $\mu_\infty \leq \mu < 4$, qui est la zone vraiment difficile, nous allons sauter directement au cas $\mu = 4$ qui présente le double avantage d'être facile et spectaculaire.

4 Le cas $\mu = 4$ et le chaos

4.1 Définir le chaos ?

Dans ce paragraphe, on pose $I = [a, b]$ avec $a < b$.

4.1.1 Les orbites denses

Il y a de nombreuses notions de chaos, différentes selon les auteurs (chaos au sens de Devaney, de Li-Yorke, etc.). Une excellente référence sur le sujet est le (projet de) livre de Sylvie Ruelle [Ruelle]. J'adopte dans ce qui suit la définition suivante :

4.1 Définition. Soit $f : I \rightarrow I$ une fonction continue. La dynamique associée à f sera dite **chaotique** s'il existe $u_0 \in I$ tel que la suite récurrente (u_n) associée à f et u_0 soit partout dense dans I .

4.2 Remarque. Si cette propriété est vraie, les u_0 dont l'orbite est dense sont eux-mêmes denses dans I . En effet, si u_0 convient, tous les u_n de la suite associée à u_0 conviennent aussi (l'orbite de u_n est dense, car si l'on retire un nombre fini de points à une partie dense dans un intervalle, elle reste dense).

Une propriété équivalente est la suivante :

4.3 Définition. Soit $f : I \rightarrow I$ une fonction continue. On dit que f est **topologiquement transitive** ou simplement **transitive** si étant donnés deux ouverts non vides U, V de I il existe un entier p tel que $f^p(U) \cap V \neq \emptyset$.

4.4 Proposition. Avec les notations précédentes, la dynamique est chaotique si et seulement si f est transitive.

Démonstration. S'il existe un point dont l'orbite est dense, il est clair que f est transitive. En effet, il y a alors une infinité de termes u_n dans tout ouvert non vide, de sorte que si on en prend un dans U il y a une de ses images qui est dans V .

Réciproquement, si f est transitive, soit $(\omega_n)_{n \in \mathbf{N}}$ une base dénombrable d'ouverts de I (par exemple les intervalles ouverts³² contenus dans I à extrémités rationnelles). Pour un ω_n on considère l'ouvert $\Omega_n = \cup_{p \in \mathbf{N}} f^{-p}(\omega_n)$. Cet ouvert est partout dense dans I . En effet, si U est un ouvert de I , la transitivité montre qu'il existe p tel que $f^p(U) \cap \omega_n \neq \emptyset$, ou encore $U \cap f^{-p}(\omega_n) \neq \emptyset$, de sorte que U rencontre Ω_n . En vertu du théorème de Baire, l'intersection des Ω_n est encore partout dense. Mais, cette intersection est formée de points dont les orbites sont partout denses. En effet, si u est dans cette intersection on a :

$$\forall n \in \mathbf{N}, \quad \exists p_n \in \mathbf{N}, \quad f^{p_n}(u) = u_{p_n} \in \omega_n.$$

Cela signifie que la suite (u_n) associée à u a des points dans tous les ω_n : elle est dense.

³²Ou semi-ouverts s'ils sont au bord.

4.1.2 Chaos, sensibilité et densité des points périodiques

Le chaos au sens précédent implique deux autres propriétés fondamentales. Donnons d'abord une autre définition :

4.5 Définition. Soit $f : I \rightarrow I$ une fonction continue. On dit que f est **sensible aux conditions initiales** ou simplement **sensitive**, s'il existe $\epsilon > 0$ tel que, pour tout $x \in I$ et pour tout $\eta > 0$, il existe $y \in I$ et $n \in \mathbf{N}$ avec $|x - y| < \eta$ et $|f^n(x) - f^n(y)| > \epsilon$.

Le ϵ est une sorte d'erreur incompressible, à terme, quelle que soit la précision du point de départ.

On a alors le théorème suivant :

4.6 Théorème. Soit $f : I \rightarrow I$ une fonction continue. On suppose que la dynamique associée est chaotique. Alors :

- 1) l'ensemble des points périodiques de f est partout dense dans I ,
- 2) f est sensitive.

Démonstration. Voici l'idée de la preuve de 1), voir [Ruelle] 2.1.16 pour des précisions. On suppose qu'il y a un intervalle $[c, d] \subset I$, avec $c < d$ qui ne contient aucun point périodique. Comme la dynamique est chaotique, il existe un $x \in]c, d[$ dont l'orbite est dense, donc un entier p tel que $x < f^p(x) < d$. De même il existe un y avec $c < f^q(y) < y < x$. On montre alors que, pour tout $m > 0$, on a $f^{mp}(x) > x$. Sinon, il y a un intervalle $[x, f^{mp}(x)]$ qui augmente³³ quand on applique f^p et, par le lemme de l'intervalle gourmand 5.2, il y a un point périodique. Le même raisonnement montre qu'on a $f^{nq}(y) < y$ pour tout $n > 0$. Mais alors, on a $f^{pq}(y) < y < x < f^{pq}(x)$ et le lemme de l'intervalle gourmand conclut.

Pour 2), voir par exemple [Ruelle] 2.2.3.

4.7 Remarque. La conjonction des trois propriétés (transitivité, sensibilité, densité des points périodiques) est ce qu'on appelle le **chaos au sens de Devaney**. Devaney justifie sa définition en caractérisant un système dynamique chaotique par trois éléments :

- Il est imprévisible. C'est la sensibilité aux conditions initiales : même si on bouge à peine les données initiales, au bout d'un certain nombre d'itérations, le système sera très différent. C'est le fameux effet papillon dont on a tant parlé. Bien entendu, un tel comportement rend extrêmement difficile (même pour un ordinateur) le calcul de $f^n(x)$ pour un x donné, puisque la moindre erreur d'arrondi a des conséquences catastrophiques, voir 4.16 ci-dessous.

³³C'est un joli exercice pour le lecteur.

- Il est indécomposable. Cela signifie qu'on ne peut pas le partager en deux intervalles indépendants et stables. C'est, bien entendu, la transitivité qui interdit cela.

- Il n'est pas aléatoire. En effet, il a des éléments de régularité, à savoir les points périodiques, qui sont partout denses.

4.8 Remarque. Attention, le chaos, s'il impose la densité des points périodiques, n'implique pas qu'il en existe de toutes les périodes. En particulier l'exemple de Ruelle, voir 6.5.1, est chaotique sur un intervalle mais n'a pas de points de période 3.

4.1.3 Le chaos et les cycles

Le résultat suivant donne une condition nécessaire de chaos :

4.9 Proposition. *Soit $I = [a, b]$ un intervalle compact et soit $f : I \rightarrow I$ une fonction de classe C^1 . Si la dynamique de f est chaotique, il n'y a pas de cycles attractifs pour f .*

Démonstration. Supposons qu'il existe un point $\alpha \in I$, périodique d'ordre $n \geq 1$ et attractif. Soient $\alpha = \alpha_0, \alpha_1, \dots, \alpha_{n-1}$ les images de α . Comme α est un point fixe attractif de f^n , il existe un intervalle $J = [\alpha - \eta, \alpha + \eta]$, avec $\eta > 0$, tel que, pour tout $x \in J$, $f^{pn}(x)$ tende vers α quand p tend vers $+\infty$ (voir 2.1). Il en résulte que la suite $f^{pn+k}(x)$ converge vers α_k pour tout $k = 0, 1, \dots, n-1$. Les valeurs d'adhérence de la suite $f^i(x)$, pour $x \in J$, sont donc exactement les points du cycle. Mais, on sait qu'il existe dans J un point dont l'orbite est partout dense, donc admet tous les points de l'intervalle I comme valeurs d'adhérence, et c'est une contradiction.

4.2 Chaos et transitivité pour $\mu = 4$

On pose à nouveau $I = [0, 1]$.

4.2.1 La remarque fondamentale

Elle consiste à noter qu'on a, pour $\theta \in \mathbf{R}$, $\sin^2 2\theta = 4 \sin^2 \theta \cos^2 \theta = 4 \sin^2 \theta (1 - \sin^2 \theta)$. On a donc, avec $f(x) = 4x(1-x)$, la formule :

$$(*) \quad \sin^2 2\theta = f(\sin^2 \theta).$$

On utilisera cette remarque sous la forme suivante :

4.10 Proposition. *Tout nombre $x \in [0, 1]$ s'écrit de manière unique sous la forme $x = \sin^2 \theta$, avec $\theta \in [0, \pi/2]$. On a les formules $f(x) = f(\sin^2 \theta) = \sin^2(2\theta)$ et, plus généralement, $f^n(x) = f^n(\sin^2 \theta) = \sin^2(2^n \theta)$.*

4.11 Remarque. Si l'on pose $g(\theta) = 2\theta$ et $h(\theta) = \sin^2 \theta$, on a $h \circ g = f \circ h$, de sorte que la formule (*) ressemble à la formule de conjugaison $f = hgh^{-1}$. Mais, attention, h n'est pas un homéomorphisme (c'est une fonction de période π). On parle à ce propos de semi-conjugaison.

En revanche, f est vraiment conjuguée de l'application "tente", très proche de l'application de doublement de l'angle, voir aussi 6.10 :

4.12 Proposition. *On considère l'application $t : [0, 1] \rightarrow [0, 1]$ définie par $t(x) = 2x$ pour $0 \leq x \leq 1/2$ et $t(x) = 2 - 2x$ pour $1/2 \leq x \leq 1$. Soit φ l'application de $[0, 1]$ dans lui-même définie par $\varphi(x) = \sin^2 \frac{\pi x}{2}$. Alors φ est un homéomorphisme, de réciproque $\varphi^{-1}(x) = \frac{2}{\pi} \text{Arcsin} \sqrt{x}$, et on a la formule : $f = \varphi t \varphi^{-1}$.*

Démonstration. Il suffit de vérifier la formule : $\sin^2 \frac{\pi t(x)}{2} = \sin^2 \pi x$ pour tout $x \in [0, 1]$.

4.2.2 La transitivité

4.13 Théorème. *Il existe $a \in I$ tel que l'orbite de a par f_4 soit partout dense dans I . Autrement dit, la dynamique de f_4 est chaotique sur I .*

Démonstration. Tout repose sur la fabrication d'un nombre magique m de $]0, 1[$, dit nombre de Champernowne³⁴, dont le développement en base 2 est le suivant :

$$m = 0,01000111000001010100111011101110000 \dots,$$

autrement dit, le nombre obtenu en mettant bout à bout, après la virgule, toutes les suites possibles formées d'un chiffre, puis deux, puis trois, etc. choisis parmi 0 et 1. La propriété de ce nombre c'est que pour tout $n \in \mathbf{N}^*$, toute suite $a_1 \dots a_n \in \{0, 1\}^n$ apparaît dans m . On pose $a = \sin^2 m\pi$. Alors, quel que soit $x \in [0, 1]$, et quel que soit $\epsilon > 0$, il existe N tel que $|x - f^N(a)| < \epsilon$. En effet, on écrit $x = \sin^2 \alpha$ avec $\alpha \in [0, \pi/2]$ et on calcule le développement

³⁴David Champernowne, étudiant à Cambridge, établit en 1933 les propriétés de l'analogie de m en base 10.

en base 2 de $\alpha/\pi = 0, a_1 \dots a_n \dots$, en choisissant n tel que $2^{-n} < \epsilon/\pi$. La suite $a_1 \dots a_n$ apparaît entre $N + 1$ et $N + n$ dans le développement de m :

$$m = 0, b_1 \dots b_N a_1 \dots a_n b_{N+n+1} \dots$$

On a alors $2^N m = b_1 \dots b_N, a_1 \dots a_n b_{N+n+1} \dots = p + (0, a_1 \dots a_n b_{N+n+1} \dots)$, où p est un entier. On en déduit la formule :

$$2^N m \pi = p\pi + \pi \times (0, a_1 \dots a_n b_{N+n+1} \dots),$$

de sorte que la différence $|2^N m \pi - \alpha - p\pi|$ est $< \pi 2^{-n} < \epsilon$. Comme la fonction \sin^2 est lipschitzienne de rapport 1, on a aussi $|\sin^2 2^N m \pi - \sin^2 \alpha| < \epsilon$ soit $|f^N(a) - x| < \epsilon$.

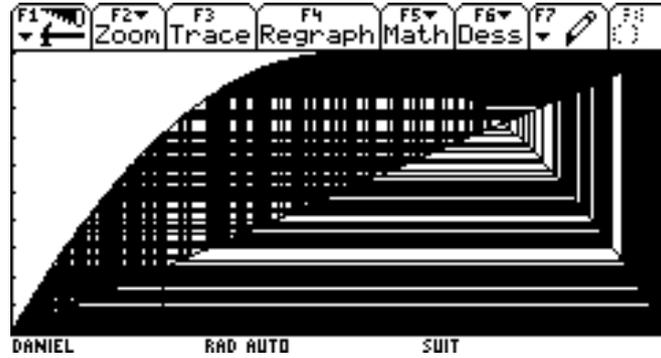


Figure 11: La suite logistique est chaotique pour $\mu = 4$

4.14 Remarque. Il est facile de programmer sur la calculatrice une fonction $champer(x, n, p)$ qui fournisse, avec une part d'aléatoire, un point u_0 dont le p -ième itéré par f_4 approche x à 2^{-n} près. On commence par calculer $a = \text{Arcsin}(\sqrt{x})/\pi$, puis la partie entière de $2^n a$, soit b , puis $c = \frac{alea(2^p)}{2^p} + \frac{b}{2^{n+p}}$ (où $alea(2^p)$ est un nombre entier aléatoire compris entre 1 et 2^p) et enfin $u_0 = \sin^2(c\pi)$. Le p -ième itéré est proche de $\sin^2 \frac{b\pi}{2^n}$ donc de $\sin^2 a\pi$ c'est-à-dire de x . Par exemple, avec $x = 1/\sqrt{2}$ et $p = 12$, le nombre $u_0 = 0,75618379413453$ donne x au 12-ième coup à $2^{-20} \sim 10^{-6}$ près.

4.3 Les autres oripeaux du chaos pour $\mu = 4$: la sensibilité

4.3.1 La sensibilité

On a vu que la transitivité implique la sensibilité. Dans le cas de f_4 , il est facile de prouver celle-ci directement :

4.15 Proposition. *Soit $J =]x - \epsilon, x + \epsilon[$ un intervalle ouvert contenu dans I . Il existe un entier n tel que $f^n(J) = I$. En particulier, il existe y dans J tel que $|f^n(x) - f^n(y)| \geq 1/2$, ce qui montre la sensibilité aux conditions initiales³⁵.*

Démonstration. On utilise la formule (*). On écrit $x - \epsilon = \sin^2 \alpha$ et $x + \epsilon = \sin^2 \beta$, avec $0 \leq \alpha < \beta \leq \pi/2$. Pour n assez grand, on a $2^n(\beta - \alpha) > 2\pi$, de sorte que l'intervalle $]2^n\alpha, 2^n\beta[$ contient un intervalle de la forme $[k\pi, (k+1)\pi]$. Pour cet n on a $f^n(J) \supset I$. En effet, si z est dans I on peut l'écrire $z = \sin^2 \theta$ avec $\theta \in [k\pi, (k+1)\pi]$. On a alors $\theta = 2^n\varphi$ avec $\alpha < \varphi < \beta$, donc $z = f^n(y)$ avec $y = \sin^2 \varphi \in J$.

Si on choisit³⁶ z tel que $|z - f^n(x)| \geq 1/2$, on a bien réalisé la condition de sensibilité.

4.16 Remarque. La sensibilité aux conditions initiales rend très problématique le calcul des valeurs de u_n pour u_0 donné, même pour une calculatrice ou un ordinateur. Le lecteur pourra s'en convaincre en pratiquant quelques petites expériences :

- Pour la fonction f_4 , le point $x = \sin^2 \frac{\pi}{384}$ vérifie $f^7(x) = 3/4 = \tau$ et il reste en ce point puisque τ est fixe par f . C'est bien ce qu'on vérifie avec la calculatrice, à ceci près qu'à partir de $n = 16$, on voit la suite dérailler et devenir chaotique (regarder soit le graphe, soit la tabulation). La raison est que la calculatrice prend pour u_0 une valeur approchée et que la petite erreur commise au départ donne une grande erreur au bout de quelques itérations. Le même type de phénomène se produit avec tous les points périodiques, cf. 4.18.

- Tous ? Non, car un point périodique³⁷ résiste vaillamment. Pour $u_0 = \sin^2 \frac{\pi}{17}$ on a un cycle d'ordre 4 et c'est ce que dit la Voyage 200, même si on l'oblige à calculer jusqu'à $n = 2000$. J'ai cru d'abord que la calculatrice était très intelligente. En effet, dans ce cas, le nombre donné est constructible (puisque le polygone à 17 côtés l'est), c'est-à-dire qu'il s'écrit avec des racines carrées, ce qui m'a conduit à penser que la calculatrice faisait le calcul de

³⁵Par exemple, si l'on part d'un intervalle de largeur $1/10$ il suffit de cinq itérations pour remplir tout $[0, 1]$.

³⁶C'est toujours possible, même si $f^n(x) = 1/2$.

³⁷L'orbite de ce point est la seule stable sur la TI Voyage parmi les points de période ≤ 6 .

manière formelle³⁸. Cette belle théorie s'effondre car la calculatrice trouve encore un cycle d'ordre 4 si on lui donne u_0 de manière approchée :

$$u_0 = 0,033763885297822.$$

En regardant les arrondis des itérés de cette valeur, on tombe sur un cycle d'ordre 4 (pour la calculatrice), formé des valeurs :

$$\begin{aligned} v_0 &= 0,033763885297826, & v_1 &= 0,13049554138968, \\ v_2 &= 0,45386582026838 & \text{et} & & v_3 &= 0,99148654984195. \end{aligned}$$

En étudiant de près le calcul de ces itérés avec *MuPAD* (qui calcule exactement avec ce nombre de décimales), on s'aperçoit que la calculatrice ne trouve un cycle que parce qu'elle commet une erreur dans le calcul de $f(v_2)$ qui devrait donner $v_3 = 0,99148654984196(232)\dots$ (le dernier chiffre de la calculatrice est donc faux), valeur qui conduirait à une orbite partout dense³⁹. Si l'on regarde de plus près le calcul de $f(v_2)$ on voit que l'erreur provient d'une erreur d'arrondi. Il semble que la machine calcule d'abord $4v_2$. Comme elle ne garde que 14 chiffres significatifs et que $4v_2$ commence par 1,81..., elle néglige le chiffre des 10^{-14} , commettant une erreur de 2×10^{-14} qui, en multipliant par $1 - v_2 \sim 0,54$ donne une erreur de plus de 10^{-14} . D'ailleurs, si l'on oblige la calculatrice à faire le calcul dans l'ordre $4 \times (v \times (1 - v))$, elle donne un résultat correct !

• Une expérience très proche de cette dernière remarque consiste à rentrer, sur la TI Voyage 200, les deux suites $u_{n+1} = 4u_n(1 - u_n)$ et $v_{n+1} = 4v_n - 4v_n^2$, avec le même point de départ 0,23 par exemple. Ce sont évidemment les mêmes, pourtant, les arrondis sont différents et on voit ces deux suites diverger dès le 11-ième coup et, au 48-ième coup, différer de plus de 0,5. Question : l'une des valeurs de $f^{48}(0,23)$ est-elle exacte ? laquelle ? Réponse : elles sont toutes deux fausses. Si on effectue ces deux calculs avec MuPAD et 100 décimales, on obtient les mêmes décimales jusqu'à la 93-ième, ce qui semble indiquer que ce résultat est fiable. Or, si on compare ce que donne la Voyage et ce que donne MuPAD, on a, pour $n = 48$, sur TI :

$$u_{48} = 0,99455195753 \quad \text{et} \quad v_{48} = 0,3879228736$$

et sur MuPAD $u_{48} = v_{48} = 0,8436149560$.

Avec la suite (v_n) , le point de départ $u_0 = \sin^2 \frac{\pi}{17}$ conduit à une suite chaotique.

³⁸Ce qui pouvait être corroboré par le fait que lorsqu'on lui rentre une valeur trigonométrique, elle peut lui faire subir une transformation. Par exemple si l'on rentre $\sin^2 \frac{5\pi}{17}$, elle écrit $\cos^2 \frac{7\pi}{34}$.

³⁹D'ailleurs, si l'on part de $\sin^2(4\pi/17)$ ou $\sin^2(8\pi/17)$ on obtient des orbites denses !

4.4 Les points périodiques pour $\mu = 4$

Comme pour la sensibilité, la densité des points périodiques est une conséquence du chaos, mais il est facile de la prouver directement.

4.4.1 Détermination des points périodiques

Le lemme de base est purement trigonométrique :

4.17 Lemme. Soient $\theta, \varphi \in \mathbf{R}$. On a $\sin^2 \theta = \sin^2 \varphi$ si et seulement si φ est égal, modulo π , à θ ou $-\theta$.

On en déduit le calcul des points périodiques :

4.18 Proposition. Les points $u_0 \in [0, 1]$ qui sont tels que la suite (u_n) soit périodique de période ≥ 1 sont les points $u_0 = \sin^2 \frac{k\pi}{2^n \pm 1}$, avec $n \in \mathbf{N}^*$ et $0 \leq k \leq 2^{n-1}$ (resp. $0 \leq k < 2^{n-1}$) si le signe est $+$ (resp. si le signe est $-$). La période de la suite associée à un tel point est égale à n sauf s'il existe un entier $p < n$ et un entier $l \leq 2^{p-1}$ tel que l'on ait $\frac{k}{2^n \pm 1} = \frac{l}{2^p \pm 1}$. Si la fraction $\frac{k}{2^n \pm 1}$ est irréductible, ce cas ne peut se produire.

Démonstration. Posons $u_0 = \sin^2 \theta$ avec $\theta \in [0, \pi/2]$. Dire que la suite admet la période n signifie qu'on a $f^n(u_0) = u_0$, c'est-à-dire $\sin^2(2^n \theta) = \sin^2 \theta$. On a ainsi, en vertu du lemme 4.17, $2^n \theta = \pm \theta + k\pi$. On a donc bien $u_0 = \sin^2 \frac{k\pi}{2^n \pm 1}$.

4.19 Remarque. Le cas de diminution de la période peut effectivement se produire. Par exemple avec $u_0 = \sin^2 \frac{15\pi}{2^8-1}$ on a la période 4 à cause de la formule $\frac{15}{2^8-1} = \frac{1}{2^4+1}$. De même avec $\sin^2 \frac{13\pi}{2^6+1}$ qui est de période 2. En revanche, dans le cas de $\sin^2 \frac{3\pi}{2^7+1} = \sin^2 \frac{\pi}{43}$ on a la période 7 bien que la fraction soit réductible car 43 n'est pas égal à une puissance de 2 à ± 1 près. La décomposition des nombres de la forme $2^n \pm 1$ n'est pas chose évidente puisqu'on y trouve les nombres de Fermat et ceux de Mersenne, voir annexe 7.5.

4.20 Proposition. Les points périodiques sont partout denses dans $[0, 1]$. Si (p_n) est une suite d'entiers tendant vers $+\infty$, on peut même préciser que les points admettant les p_n comme périodes sont denses dans $[0, 1]$. C'est le cas de la suite $p_n = 2^n$.

Démonstration. Soit $x \in [0, 1]$ que l'on écrit $x = \sin^2 \theta_0$ et soit $\epsilon > 0$. Il s'agit de montrer que l'intervalle $J =]x - \epsilon, x + \epsilon[$ contient un point périodique. Comme la fonction $\theta \mapsto \sin^2 \theta$ est lipschitzienne de rapport 1 sur $[0, \pi/2]$, il suffit de trouver un θ correspondant à un point périodique tel que $|\theta - \theta_0| < \epsilon$. Si on choisit n assez grand pour que $\frac{\pi}{2^{n+1}}$ soit $< 2\epsilon$, il y a un k tel que $\frac{k\pi}{2^{n+1}}$ soit dans l'intervalle $]\theta_0 - \epsilon, \theta_0 + \epsilon[$ et le point périodique $\sin^2 \frac{k\pi}{2^{n+1}}$ est alors dans J .

4.4.2 Même si l'on ne naît pas périodique, on peut le devenir à force de persévérance !

La proposition suivante détermine les points "ultimement" périodiques, c'est-à-dire ceux qui tombent, à terme, sur des points périodiques :

4.21 Proposition. *Les points $u_0 \in [0, 1]$ qui sont tels que la suite (u_n) soit périodique à partir du rang q sont les points $u_0 = \sin^2 \left(\frac{k\pi}{2^q(2^n \pm 1)} \right)$, avec $0 \leq k \leq 2^{q+n-1}$ si le signe est + (resp. $0 \leq k < 2^{q+n-1}$ si le signe est -).*

On notera que, pour $n = 1$, on obtient les points qui deviennent fixes à partir d'un certain rang. Par exemple, la suite de départ $u_0 = \sin^2 \frac{17\pi}{384}$ stationne à $3/4$ à partir du 7-ième coup.

4.4.3 La répulsivité des cycles périodiques

On montre d'abord une relation trigonométrique :

4.22 Lemme. *Soit $\theta \in \mathbf{R}$. Pour $n \in \mathbf{N}^*$, on a la formule :*

$$\sin 2^{n+1}\theta = 2^{n+1} \sin \theta \cos \theta \cos 2\theta \cdots \cos 2^n \theta = 2^n \sin 2\theta \cos 2\theta \cos 4\theta \cdots \cos 2^n \theta.$$

Démonstration. Récurrence sur n .

4.23 Proposition. *Soit x un point de période n pour $f = f_4$. On a l'égalité $|(f^n)'(x)| = 2^n$, de sorte que le n -cycle défini par x est répulsif⁴⁰.*

Démonstration. La dérivée de f est $f'(x) = 4(1 - 2x)$. En $x = \sin^2 \varphi$ on a donc $f'(x) = 4 \cos 2\varphi$. Le cycle est formé de $x = x_0 = \sin^2 \theta$, $x_1 = \sin^2 2\theta$, ..., $x_{n-1} = \sin^2 2^{n-1}\theta$ et on a $\sin^2 2^n \theta = \sin^2 \theta = x$ et aussi $|\sin 2^{n+1}\theta| = |\sin 2\theta|$. Les dérivées en chaque point sont les quadruples des cosinus de 2θ , ..., $2^n \theta$. En vertu de 3.2, la dérivée de f^n est le produit de ces dérivées. La formule trigonométrique ci-dessus, après multiplication par 2^n et simplification par les sinus, donne le résultat.

⁴⁰Bien entendu, on sait que les cycles sont répulsifs en vertu de 4.9, mais la proposition précise la dérivée.

4.24 Remarque. On notera, comme conséquence du résultat précédent, la superbe identité, avec $\epsilon = \pm 1$:

$$\cos \frac{2k\pi}{2^n + \epsilon} \times \cos \frac{4k\pi}{2^n + \epsilon} \times \cdots \times \cos \frac{2^j k\pi}{2^n + \epsilon} \times \cdots \times \cos \frac{2^n k\pi}{2^n + \epsilon} = -\frac{\epsilon}{2^n}.$$

5 La zone des tempêtes : $\mu_\infty \leq \mu \leq 4$

Cette fois, fini de rire, nous abordons la zone vraiment difficile, celle où les théorèmes deviennent hors de prix. Le premier résultat essentiel (et très joli) est dû à Sarkovsky⁴¹ et il explique dans quel ordre doivent apparaître les cycles. En particulier il confirme que les cycles d'ordre une puissance de 2 doivent venir en premier. Il montre aussi le rôle éminent⁴² des cycles de période 3. On discutera ensuite vraiment de ce qui se passe dans la zone en question. On verra qu'on y trouve entremêlées des valeurs de μ pour lesquelles on a un cycle attractif et d'autres qui mènent au chaos et on essaiera de se faire une opinion sur la brûlante question : le chaos est-il accidentel ou structurel ?

5.1 Le théorème de Sarkovsky

Il s'agit d'un remarquable résultat qui explique l'ordre d'apparition des points fixes. Nous avons déjà rencontré une partie de ce résultat à propos du doublement de période.

Dans tout ce qui suit, I désigne un intervalle compact non réduit à un point.

5.1.1 Variante faible : le rôle des 3-cycles

5.1 Théorème. *Soit $f : I \rightarrow I$ une fonction continue. On suppose que f admet un point de période 3. Alors, f admet des points périodiques de toutes les périodes $n \in \mathbf{N}^*$.*

La démonstration utilise deux lemmes.

⁴¹En vérité, même si je ne parle pas l'ukrainien couramment, je crois qu'il faudrait écrire ce nom Sharkovsky, voire Šarkov'skiĭ, mais, je ne sais pourquoi, une intuition subliminale m'a conduit à l'orthographier ainsi.

⁴²Qui a conduit Li et Yorke, qui redécouvraient le théorème de Sarkovsky, à ce titre racoleur et largement inexact : *Period three implies chaos*.

5.1.2 Les lemmes du gourmand et de l'ajusté

5.2 Lemme. (Lemme de l'intervalle gourmand⁴³) Soit $g : I \rightarrow I$ une fonction continue. On suppose qu'il existe un intervalle $[a, b] \subset I$ vérifiant $g([a, b]) \supset [a, b]$. Alors, g admet un point fixe dans $[a, b]$.

Démonstration. Il existe $c, d \in [a, b]$ tels que $g(c) = a$ et $g(d) = b$. On a alors $g(c) - c \leq 0$ et $g(d) - d \geq 0$, de sorte que $g(x) - x$ s'annule entre c et d en vertu du théorème des valeurs intermédiaires, donc aussi entre a et b .

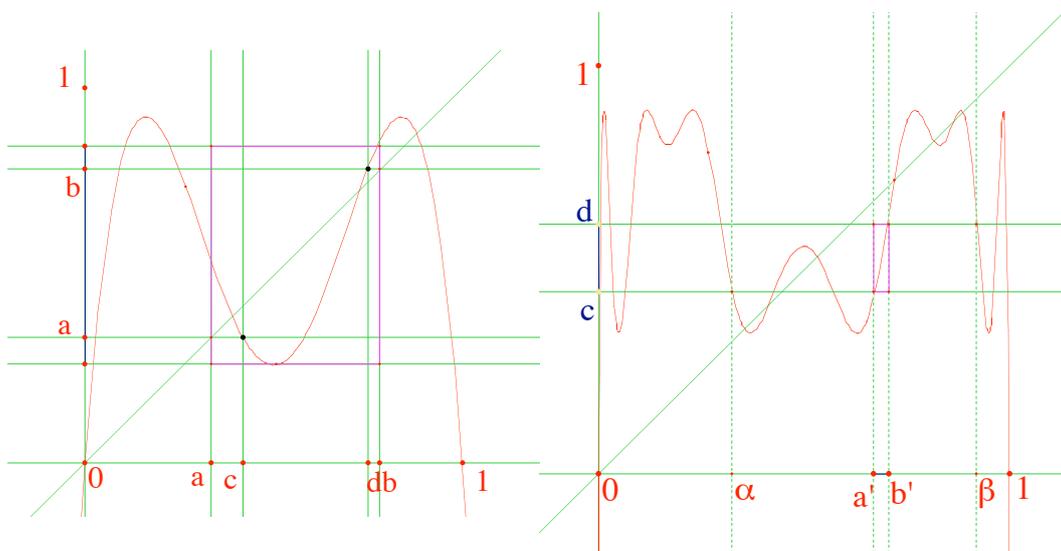


Figure 12: Les lemmes de l'intervalle gourmand et de l'intervalle ajusté

5.3 Lemme. (Lemme de l'intervalle ajusté) Soit $f : [a, b] \rightarrow \mathbf{R}$ une fonction continue et soit $[c, d]$ un intervalle contenu dans $f([a, b])$. Il existe $[a', b'] \subset [a, b]$ tel que l'on ait $f([a', b']) = [c, d]$.

Démonstration. Soient $\alpha, \beta \in [a, b]$ tels que $f(\alpha) = c$ et $f(\beta) = d$ et supposons par exemple $c < d$ et $\alpha < \beta$. Soit $a' = \sup\{x \in [\alpha, \beta] \mid f(x) = c\}$. Ce sup existe car on a $f(\alpha) = c$ et, comme f est continue, donc $f^{-1}(\{c\})$ fermé, on a $f(a') = c$. On en déduit $a' < \beta$. On considère alors $b' = \inf\{x \in [a', \beta] \mid f(x) = d\}$. Il existe car on a $f(\beta) = d$. On a $f(b') = d$ et $a' < b'$.

Je dis alors que $[a', b']$ convient. Il est clair que son image par f contient $[c, d]$ par le théorème des valeurs intermédiaires. Si on avait, disons, $x \in [a', b']$ avec $f(x) < c$, on aurait $a' < x < b'$ et, comme on a $f(b') = d > c$, il existerait y entre x et b' avec $f(y) = c$. Mais alors, a' ne serait pas le sup des z tel que $f(z) = c$ dans $[\alpha, \beta]$. De même, si on a x tel que $f(x) > d$, on trouve un y entre a' et x tel que $f(y) = d$ et cela contredit la définition de b' .

⁴³Ou encore, de l'intervalle qui a les yeux plus gros que le ventre.

5.1.3 Sarkovsky : l'existence d'un cycle d'ordre 2

On suppose qu'on a un 3-cycle a, b, c et, par exemple⁴⁴, qu'on a $a < b < c$ avec $f(a) = b$, $f(b) = c$, $f(c) = a$. Si on pose $I_0 = [a, b]$ et $I_1 = [b, c]$, on a $f(I_0) \supset I_1$ et $f(I_1) \supset I_0 \cup I_1$.

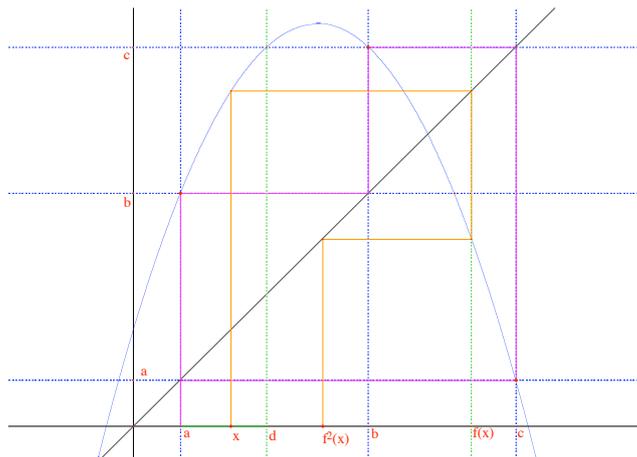


Figure 13: L'intervalle $[a, d]$ s'envoie sur $[b, c]$. Comme on a $f([b, c]) = [a, c]$ et que $[a, c]$ contient $[a, d]$, cet intervalle est gourmand pour f^2 , donc contient un point fixe pour f^2 .

Montrons déjà l'existence d'un 2-cycle. Le lemme de l'intervalle ajusté montre qu'il existe un intervalle $J \subset I_0$ qui s'envoie exactement sur I_1 . Comme on a $f(I_1) \supset I_0 \supset J$, donc $f^2(J) \supset J$, le lemme de l'intervalle gourmand montre que J contient un point fixe x pour f^2 . Comme on a $x \leq b$ et $f(x) \geq b$, ce point n'est pas fixe par f (la seule possibilité serait $x = b$, mais b n'est pas fixe).

5.1.4 Le cas général

Soit n un entier > 3 (on pourra penser à $n = 5$). Il s'agit de montrer que f admet un n -cycle. On définit par récurrence une suite d'intervalles A_i comme suit. On pose $A_0 = I_1 = [b, c]$. Comme on a $f(I_1) \supset I_1$, le lemme de l'intervalle ajusté montre qu'il existe un intervalle $A_1 \subset A_0$ avec $f(A_1) = A_0 = I_1$. On construit ainsi les A_i pour $i \leq n - 2$, de telle sorte qu'on ait $A_{i+1} \subset A_i$ et $f(A_{i+1}) = A_i$. On en déduit qu'on a $f^i(A_i) = A_0 = I_1$ pour $i \leq n - 2$.

Le pas suivant consiste à noter que, comme on a $f(I_1) \supset I_0$, donc $f^{n-1}(A_{n-2}) \supset I_0$, il existe un intervalle $A_{n-1} \subset A_{n-2}$ tel que l'on ait $f^{n-1}(A_{n-1}) =$

⁴⁴Il y a un autre cas de figure que le lecteur traitera lui-même.

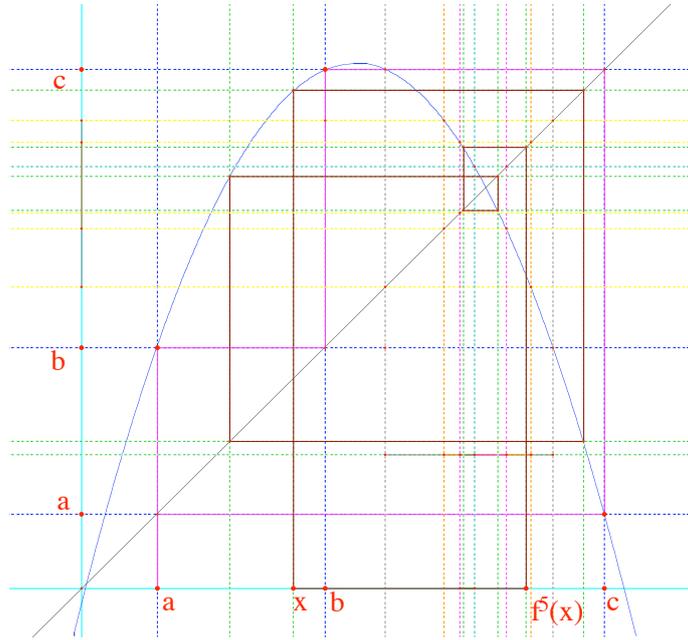


Figure 14: On a des intervalles emboîtés : *vert* \subset *rose* \subset *orange* \subset *gris* \subset $[b, c]$ avec $f(\textit{gris}) = [b, c]$, $f(\textit{orange}) = \textit{gris}$, $f(\textit{rose}) = \textit{orange}$. Comme $f([b, c]) \supset [a, b]$, on a sélectionné *vert* pour que $f^4(\textit{vert}) = [a, b]$. Alors, on a $a : f^5(\textit{vert}) = f([a, b]) \supset [b, c] \supset \textit{vert}$, donc *vert* est gourmand pour f^5 et contient un point fixe.

I_0 et donc $f^n(A_{n-1}) \supset I_1$. Comme on a $A_{n-1} \subset A_{n-2} \subset \dots \subset A_0 = I_1$, l'intervalle A_{n-1} est gourmand pour f^n , de sorte qu'il existe un point $p \in A_{n-1}$ fixe par f^n .

Il reste à montrer que p est bien de période n . Sinon, il existe $m < n$ avec $f^m(p) = p$. Considérons le point $f^{n-1}(p)$. Il est dans I_0 . Supposons d'abord qu'il est différent de b . Alors, on a $f^{n-m-1}(p) = f^{n-m-1}(f^m(p)) = f^{n-1}(p)$. Mais, comme on a $n - m - 1 \leq n - 2$, l'image de p par f^{n-m-1} est dans I_1 et c'est absurde.

Si on a $f^{n-1}(p) = b$, on a $f^n(p) = f(b) = c = p$. Mais, comme n est > 3 , on a $n - 2 > 1$ et p est dans A_1 . On devrait donc avoir $a = f(p) \in A_0 = I_1$ et c'est absurde.

5.1.5 Le théorème général de Sarkovsky

Il s'agit du résultat suivant :

5.4 Théorème. *On considère l'ordre suivant sur les entiers positifs :*

$$3 \succ 5 \succ 7 \succ 9 \dots \succ 2 \cdot 3 \succ 2 \cdot 5 \dots \succ 2^2 \cdot 3 \succ 2^2 \cdot 5 \dots \succ 2^n \succ \dots \succ 2^3 \succ 2^2 \succ 2 \succ 1.$$

Alors, si une fonction $f : I \rightarrow I$ admet un point de période n , elle admet des points de période p pour tous les $p \prec n$ dans l'ordre de Sarkovsky.

Démonstration. Voir par exemple [Ruelle] ou [Devaney].

5.5 Remarque. Voici une description analytique de l'ordre de Sarkovsky.

Soient $p, p', q, q' \in \mathbf{N}$, on a $2^p(2q+1) \prec 2^{p'}(2q'+1)$ dans les cas suivants :

- Si on a $q = 0$ et $q' > 0$.
- Si on a $q = q' = 0$ et $p < p'$.
- Si on a q et $q' > 0$ et $p > p'$.
- Si on a $p = p'$ et $0 < q' < q$.

Si le théorème de Sarkovsky est un résultat remarquable, il n'épuise pas les questions combinatoires sur le sujet : par exemple, pour chaque partie cofinale⁴⁵ de l'ordre de Sarkovsky, existe-t-il une valeur μ qui admet exactement cette partie comme ensemble de périodes ?

Voici une conjecture sur ce thème :

5.6 Conjecture. Soit n un entier de la forme $2^p(2q+1)$ avec $q > 0$ et soit α_n la borne inférieure des $\mu \in [0, 4]$ tel qu'il existe un n -cycle pour f_μ . Alors, les α_n sont **strictement** dans l'ordre de Sarkovsky, c'est à dire $\alpha_n < \alpha_{n'}$ si et seulement si $p' < p$ ou si $p = p'$ et $2q'+1 < 2q+1$. Pour p fixé, quand q tend vers l'infini, la suite décroissante des $\alpha_{2^p(2q+1)}$ tend vers une valeur β_p et la suite décroissante des β_p tend vers μ_∞ .

Le théorème de Sarkovsky donne l'inégalité large $\alpha_n \leq \alpha_{n'}$, la décroissance de $\alpha_{2^p(2q+1)}$ par rapport à q (donc sa convergence puisque ce sont des nombres ≥ 0) et la décroissance des β_p (donc sa convergence vers $\beta \geq \mu_\infty$). La question non évidente (sur laquelle je ne sais rien) est l'existence d'une zone non vide entre μ_∞ et β où n'existeraient que des 2^n -cycles (répulsifs).

5.2 Les points fixes de f^3

Pour les calculs qui suivent, je donne deux versions, l'une avec le polynôme $f_\mu(x) = \mu x(1-x)$, l'autre, un peu plus simple, avec son conjugué $g_a(x) = x^2 - a$, avec $a = \frac{\mu(\mu-2)}{4}$, la conjugaison se faisant par $h(x) = -\mu x + \mu/2$, voir Annexe 7.3.

⁴⁵C'est-à-dire l'ensemble des $n \prec n_0$ pour un n_0 fixé.

5.2.1 Le polynôme qui donne les points fixes

5.7 Proposition. *Les points fixes de f_μ^3 non points fixes de f_μ sont les racines du polynôme :*

$$P_{3,\mu}(x) = \mu^6 x^6 - \mu^5(3\mu+1)x^5 + \mu^4(\mu+1)(3\mu+1)x^4 - \mu^3(\mu^3+5\mu^2+3\mu+1)x^3 \\ + \mu^2(2\mu^3+3\mu^2+3\mu+1)x^2 - \mu(\mu^3+2\mu^2+2\mu+1)x + \mu^2 + \mu + 1.$$

Les points fixes de g_a^3 non points fixes de g_a sont les racines de :

$$Q_{3,a}(x) = x^6 + x^5 - (3a-1)x^4 - (2a-1)x^3 + (3a^2-3a+1)x^2 + (a^2-2a+1)x - a^3 + 2a^2 - a + 1.$$

Démonstration. Il s'agit de calculer $(f_\mu^3(x) - x)/(f_\mu(x) - x)$ afin d'éliminer les points fixes de f . La Voyage fait très bien ce genre de choses.

5.2.2 Le premier cycle d'ordre 3

5.8 Proposition. *La fonction g_a admet des points de période 3 pour $7/4 \leq a \leq 2$. La fonction f_μ admet des points de période 3 pour $1 + \sqrt{8} \leq \mu \leq 4$ (et donc aussi des points périodiques⁴⁶ de toutes les périodes en vertu de Sarkovsky). On a $1 + \sqrt{8} \sim 3,82842712475$.*

Démonstration. Il suffit d'établir le résultat pour g_a , l'autre s'obtient par conjugaison. On commence par noter que, pour $a = 0$, $Q_{3,0}(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$ est le polynôme cyclotomique Φ_7 qui n'a pas de racines réelles. Au contraire, on a vu que f_4 admet des points de période 3, donc aussi g_2 . Le polynôme Q , vu comme polynôme en x et a définit une courbe algébrique C du plan des (x, a) et il s'agit de déterminer sa projection P sur la partie ≥ 0 de l'axe des a . Cette projection est fermée. En effet, si on considère la courbe projective \overline{C} définie par le polynôme homogénéisé $\overline{Q}(x, a, t)$, elle est compacte, donc sa projection sur l'axe est fermée. Si un point a est dans l'adhérence de la projection affine sans être dans la projection c'est donc qu'il est image d'un point à l'infini. Mais, on voit que \overline{C} n'a pas d'autre point à l'infini que $x = t = 0$, donc pas de tel point correspondant à une valeur finie de a .

Soit a_0 un point frontière de P . Alors, en tout point (x, a_0) de C , la tangente est horizontale. Sinon, le théorème des fonctions implicites montre que l'on peut exprimer a en fonction de x sur C sur un voisinage ouvert de a_0 et a_0 est intérieur à P .

⁴⁶On voit apparaître, dans le cas particulier de l'ordre 3, un théorème dû à Douady et Hubbard (Annales de l'ENS 1985) : lorsque μ augmente, il peut seulement apparaître de nouvelles périodes, mais elles ne disparaissent pas.

On cherche donc les points de C à tangente horizontale. Pour cela, il faut éliminer x entre Q et sa dérivée partielle par rapport à a . Je l'ai fait avec Maple, on obtient un polynôme en a :

$$-16384a^7 + 94208a^6 - 208896a^4 + 247296a^3 + 131712a^2 - 48020a + 1680$$

qui admet une seule racine réelle : $7/4$, de multiplicité 3.

Comme P contient 2, le lemme ci-dessous montre alors que la projection P est l'intervalle $[7/4, +\infty[$.

5.9 Lemme. *Un fermé F de \mathbf{R} dont la frontière est réduite à un point a est égal soit à $\{a\}$, soit à $[a, +\infty[$ soit à $] - \infty, a]$.*

Démonstration. Si F n'est pas réduit à a , supposons qu'il contienne $b > a$. Le point b est alors intérieur à F . Soit G l'ensemble des $x > b$ tels que $[b, x] \subset F$. Posons $d = \sup G$. Si d est fini, il est dans F (puisque F est fermé), donc intérieur à F (puisque $> a$). Mais alors, il y a un intervalle $[b, d + \epsilon]$ inclus dans F et c'est absurde. On raisonne de même à gauche de b pour montrer que $c = \inf G$ est égal à a .

5.10 Exemple. Pour $\mu = 3,82843$, le polynôme admet 6 racines dans $[0, 1]$. On repère facilement que trois d'entre elles forment un 3-cycle attractif (il suffit de calculer les termes de la suite de point de départ $u_0 = 0,5$ jusqu'à 1000 et on voit apparaître ce cycle à 10^{-7} près). Je les note $a_1 = 0,159769931604$, $a_2 = 0,513941845577$, $a_3 = 0,9563633394772$.

Les autres sont $b_1 = 0,160087452099$, $b_2 = 0,514768628867$ et $b_3 = 0,956272485874$. Ils forment un cycle répulsif⁴⁷.

5.11 Remarques.

1) On peut trouver la valeur $\mu_3 = 1 + \sqrt{8} = 1 + 2\sqrt{2}$ d'une autre manière, si on a l'intuition de la genèse des 3-cycles. Comme on l'a vu ci-dessus, quand on est au-delà de la valeur limite μ_3 , on a deux cycles d'ordre 3, l'un attractif, l'autre répulsif. Ce qu'on imagine, c'est que, lorsque μ tend vers μ_3 , ces deux cycles vont se coller en un seul 3-cycle, et comme l'un est répulsif et l'autre attractif, le 3-cycle limite sera parabolique, i.e. avec une dérivée 1.

Or, si on a un cycle x_1, x_2, x_3 , et si on pose $P(x) = a(x-x_1)(x-x_2)(x-x_3)$, la dérivée en un quelconque des points du cycle $Df^3(x_i)$ est :

$$\mu^3(1-2x_1)(1-2x_2)(1-2x_3) = \frac{(2\mu)^3}{a}P(1/2).$$

⁴⁷Il y a au plus un 3-cycle attractif, cf. 5.14.

Si les points du 3-cycle sont tous racines doubles de $P_{3,\mu}$, on aura, de même $(Df^3(x_i))^2 = 2^6 P_{3,\mu}(1/2)$. Si la situation est bien telle qu'on l'envisage, μ_3 sera donc racine de l'équation $2^6 P_{3,\mu}(1/2) = 1$. Or on a :

$$P_{3,\mu}(1/2) = -\frac{\mu^6}{64} + \frac{3\mu^5}{32} - \frac{\mu^4}{16} - \frac{3\mu^3}{8} + \frac{\mu^2}{4} + \frac{\mu}{2} + 1$$

et on retrouve $\mu = 1 + \sqrt{8}$.

On vérifie effectivement que, pour cette valeur de μ , le polynôme $P_{3,\mu}(x)$ est un carré parfait, précisément, c'est le carré de :

$$(25 + 22\sqrt{2})x^3 - (42 + 35\sqrt{2})x^2 + (21 + 14\sqrt{2})x - (3 + \sqrt{2})$$

comme on le voit en calculant les coefficients de $ax^3 + bx^2 + cx + d$ de proche en proche : $a = \mu^3$, $b = -\frac{\mu^2(3\mu+1)}{2}$, $c = \frac{\mu(3\mu+1)(\mu+3)}{8}$, $d = \frac{\mu^3-7\mu^2-5\mu-5}{16}$ et pour que le carré de d soit $\mu^2 + \mu + 1$ il faut avoir $\mu = 1 + 2\sqrt{2}$.

2) Là encore, les calculs sont beaucoup plus faciles avec la version $g_a(x) = x^2 - a$. On a $g'_a(x) = 2x$. Pour un cycle d'ordre 3, x_1, x_2, x_3 , formé des racines du polynôme $Q(x) = \alpha(x - x_1)(x - x_2)(x - x_3)$, la dérivée $Df^3(x_i)$ vaut $-8x_1x_2x_3$, c'est-à-dire⁴⁸ $-8Q(0)/\alpha$. Si on calcule avec le polynôme $Q_{3,a}$ (en pensant que c'est un carré pour la valeur limite de a), on a $(Df^3(x_i))^2 = 64Q_{3,a}(0) = 64(-a^3 + 2a^2 - a + 1)$ et il s'agit de voir quand cette quantité vaut 1. On a donc l'équation :

$$-64a^3 + 128a^2 - 64a + 63 = 0$$

et on vérifie que $7/4$ en est racine (et que c'est la seule).

Dans le cas général, la quantité $64(-a^3 + 2a^2 - a + 1)$ est le produit des deux dérivées correspondant aux deux cycles d'ordre 3. Elle est décroissante, vaut 1 en $a = 7/4$, elle passe ensuite par 0, puis -1 et finit par tendre vers $-\infty$.

3) Revenons au paramétrage en μ pour cette discussion. On rappelle que le produit des deux dérivées en les deux 3-cycles est égal à :

$$2^6 P_{3,\mu}(1/2) = -\mu^6 + 6\mu^5 - 4\mu^4 - 24\mu^3 + 16\mu^2 + 32\mu + 64.$$

On trouve ainsi la valeur $\mu = 3,83187405528$ qui donne une dérivée nulle et correspond à un 3-cycle super-attractif et la valeur $\mu = 3,83528309011$ qui donne une dérivée -1 (qui correspond à la limite d'existence⁴⁹ d'un cycle attractif).

⁴⁸On n'oubliera pas que par la conjugaison, $1/2$ s'envoie sur 0.

⁴⁹Attention, comme on regarde le produit des deux dérivées, ce qui est clair c'est que si ce produit est < 1 en valeur absolue il y a un cycle attractif, mais s'il est > 1 , on ne peut pas conclure. D'ailleurs, pour $\mu = 3,835228309012$, il semble bien qu'il y ait un 3-cycle attractif.

5.3 Les résultats fondamentaux : Jakobson, Graczyk-Swiatek, Lyubich

Il y a beaucoup de résultats, tous très difficiles⁵⁰, qui décrivent le comportement des f_μ pour μ voisin de 4. J'en ai choisi trois que j'espère emblématiques. Les deux premiers, d'une certaine façon, vont en sens inverse, l'un de l'autre :

5.12 Théorème (Jakobson, 1981). *L'ensemble C (comme chaotique) des $\mu \in [3, 4]$ tels que f admette une dynamique chaotique⁵¹ sur l'intervalle stable minimum (voir 7.4 et 7.5) est de mesure de Lebesgue positive. Plus précisément, si λ désigne la mesure de Lebesgue, on a :*

$$\lim_{\epsilon \rightarrow 0} \frac{\lambda(C \cap [4 - \epsilon, 4])}{\epsilon} = 1.$$

Voir [Jakobson].

5.13 Théorème (Hyperbolicité, Graczyk-Swiatek, Lyubich, 1997). *L'ensemble H (comme hyperbolique) des $\mu \in [3, 4]$ tels que f admette un cycle attractif est un ouvert dense.*

Voir [Lyubich1] ou [Graczyk-Swiatek1,2] si l'on souhaite se frotter à une démonstration vraiment difficile.

5.14 Remarques.

- 1) Que l'ensemble H soit ouvert est clair, mais c'est bien la seule chose.
- 2) Si f admet un cycle attractif, on montre deux faits complémentaires :
 - Il y en a un seul. Cela résulte d'une propriété un peu bizarre (la dérivée schwarzienne négative), voir [Devaney].
 - Pour u_0 en dehors d'un ensemble de mesure nulle⁵², la suite $u_{n+1} = f(u_n)$ converge vers le cycle attractif.

Enfin, le théorème suivant, qui clôt, pour l'essentiel, la question, affirme qu'on est – presque – obligatoirement dans l'un ou l'autre de ces cas (voir [Lyubich3]) :

5.15 Théorème (Lyubich, 2002). *Presque tous les $\mu \in [0, 4]$ correspondent à une dynamique hyperbolique ou chaotique, autrement dit l'ensemble $[0, 4] - (H \cup C)$ est de mesure de Lebesgue nulle.*

⁵⁰Il y a des controverses de paternité sur les résultats cités et leurs preuves. J'espère ne pas avoir commis d'impair.

⁵¹Précisément, pour ces valeurs de μ , f admet mesure de probabilité absolument continue invariante et ergodique, ce qui implique le comportement chaotique.

⁵²Mais peut-être non dénombrable, voir Li-Yorke ou Devaney.

5.16 Remarques.

1) Il est clair qu'on a $C \cap H = \emptyset$ en vertu de 4.9 et Lyubich montre que leur réunion est "presque" tout. Les résultats ci-dessus ne sont – heureusement – pas incompatibles. Le théorème d'hyperbolicité dit que H est dense, ce qui montre que C est d'intérieur vide. Mais on sait bien qu'il y a des ensembles de Cantor d'intérieur vide et de mesure > 0 . Le C doit donc être nécessairement un ensemble totalement discontinu.

2) Le lecteur se demande peut-être ce qu'il y a hors de C et de H . Pas grand-chose répond Lyubich. En fait, il se passe des choses plus compliquées avec notamment des Cantor attractifs, voir [De Melo-Van Strien].

3) C'est ici que la réflexion philosophique sur le chaos devient intéressante et qu'elle oppose topologues et ergodiciens : un ensemble de mesure > 0 , mais d'intérieur vide peut-il être considéré comme générique ? Je laisse le lecteur méditer sur cette grave question. Pour ma part, j'aurais tendance à penser comme les topologues qu'un ensemble générique doit être ouvert et dense (pour qu'une petite modification du paramètre ne change pas la dynamique). En ce sens, le chaos serait non générique. Mais d'autres, depuis Kolmogorov, et notamment des physiciens, mesurent la généricité en termes de mesures. Je me garderai de prendre une position trop catégorique sur ce sujet.

5.4 Une approche expérimentale du théorème d'hyperbolicité

Dans ce paragraphe, fidèle à certaines de mes prises de position en faveur d'une approche expérimentale des mathématiques, voir [Perrin], je donne des indices permettant de comprendre pourquoi le théorème d'hyperbolicité peut être vrai⁵³.

5.4.1 L'idée de base

On a déjà rencontré à plusieurs reprises l'idée d'utiliser les itérés du point critique $1/2$ de f pour montrer l'existence de cycles super-attractifs. On part de la constatation suivante :

⁵³Le prouver est une autre paire de manches. Les démonstrations publiées, qu'elles soient de Graczyk-Swiatek ou de Lyubich ont entre cinquante et cent pages dans des revues de tout premier plan et sont incroyablement difficiles. Je cite l'un des protagonistes reviewant un papier de l'autre : *A simple attempt of going through the statements of the paper named "theorems", and presumably representing the main steps of the proof, reveals that most of them cannot be understood by an expert without extensive reading through earlier parts of the paper.*

5.17 Proposition-Définition. *On pose $P_n(\mu) = f_\mu^n(1/2) - 1/2$. C'est un polynôme de degré 2^n en μ et si μ est une racine de ce polynôme, $1/2$ fait partie d'un n -cycle super-attractif pour f_μ .*

Démonstration. Le fait que P_n soit un polynôme de degré 2^n est immédiat par récurrence. Si μ annule ce polynôme, on a $f_\mu^n(1/2) = 1/2$, de sorte que $1/2$ fait partie d'un cycle d'ordre divisant n . Comme la dérivée de f est nulle en $1/2$, ce cycle est super-attractif (voir 3.2).

5.4.2 Trouver des cycles super-attractifs : les équations

En vertu de 5.17, il suffit, pour trouver des n -cycles super-attractifs de résoudre l'équation en $\mu : P_n(\mu) = 0$. Le problème est que ce polynôme devient très vite compliqué. Pour $n = 2$, le calcul est immédiat et il n'y a que la racine $1 + \sqrt{5}$. Le cas $n = 3$ a été traité en 5.11.3 et on a trouvé une unique valeur de μ donnant un 3-cycle super-attractif : $\mu = 3,83187405528$.

La cas $n = 4$ a été vu en 3.12, le polynôme $P_{4,\mu}(x)$ étant alors appelé Q_μ . La résolution de l'équation algébrique $Q_\mu(1/2) = 0$ a donné deux valeurs, la valeur $\nu_4 \sim 3,49856169933$ de la cascade de Feigenbaum et une autre : $\mu = 3,960270272187$. Cela signifie qu'il y a deux plages de μ où les 4-cycles sont attractifs : celle attendue, juste après $1 + \sqrt{6}$, avec l'apogée en ν_4 , mais aussi une deuxième zone de super-attractivité⁵⁴ vers $\mu = 3,96$.

La méthode directe permet aussi de traiter le cas $n = 5$. Cependant, le polynôme $P_{5,\mu}(x)$ est assez compliqué (de degré 31 en x et en μ) et la calculatrice met plusieurs minutes pour résoudre l'équation $P_{5,\mu}(1/2) = 0$. On trouve trois valeurs :

$$3,99029126214, 3,90577240955; 3,73874882305.$$

Pour ces valeurs on a bien des 5-cycles attractifs, qui ne passent pas exactement par $0,5$ à cause des erreurs d'arrondi et dont la dérivée n'est pas nulle (mais voisine de 0). Par exemple, pour la deuxième valeur, le cycle passe par $0,500617$ avec une dérivée $-0,22$.

Pour trouver des 7-cycles attractifs, il suffit de modifier un peu la méthode en rentrant la fonction f_μ^7 dans la calculatrice et en cherchant un μ tel que $f^7(1/2) = 1/2$ par une méthode de calcul approché (par exemple simplement un balayage). Par ce procédé, j'ai obtenu ainsi $3,9221934 < \mu < 3,9221935$. Avec la valeur de gauche, on a un 7-cycle passant presque par $1/2$ (il passe par $0,499999853047$), avec une dérivée $0,00000048748$. Il faut de la patience

⁵⁴Dans un premier temps, j'avais cru que les cycles d'ordre 2^n n'étaient plus jamais attractifs après μ_∞ . La source – bien imprudente – de cette idée fautive est le cas du 2-cycle, qui cesse d'être attractif pour $\mu > 1 + \sqrt{6}$ et ne le redevient jamais.

pour déterminer le bassin d'attraction du cycle. Avec $0,51$ comme u_0 , le cycle est stabilisé au 1000-ième coup, avec $u_0 = 0,52$ il est stabilisé au 2000-ième.

On trouve aussi, pour $\mu = 3,8008$, un 8-cycle attractif dont l'un des points est $0,936217614607$. On vérifie que l'intervalle :

$$J = [0,9362176146; 0,93621761461]$$

est stable par f^8 et que la dérivée de f^8 sur J est à peu près égale à $-0,392$.

5.4.3 Suites de mixage

Le calcul direct devenant rapidement inextricable, on va utiliser un outil combinatoire, facile à mettre en évidence sur une calculatrice, qui va permettre de montrer l'existence de valeurs du paramètres admettant des cycles super-attractifs (mais qui ne donne pas explicitement le paramètre μ). On pose $f = f_\mu$.

5.18 Définition. Soit $x \in [0,1]$. On appelle **itinéraire** de x , et on note $S(x)$, la suite formée des symboles $0,1$ ou C dont le n -ième terme est 0 si l'on a $f^n(x) < 1/2$, 1 si l'on a $f^n(x) > 1/2$ et C si l'on a $f^n(x) = 1/2$. Autrement dit, le n -ième terme de la suite est⁵⁵ $[2f^n(x)]$ (sauf si ce nombre vaut 1 auquel cas c'est C).

On appelle **suite de mixage**⁵⁶ l'itinéraire $S(f(1/2))$ du maximum $f(1/2) = \mu/4$. On la note $K(f)$.

5.19 Exemple. Pour $\mu = 4$, on a $f(1/2) = 1$, $f(1) = 0$ et $f(0) = 0$. La suite de mixage est donc $1,0,0,\dots,0,\dots$. Pour μ entre 0 et 2 la suite est constamment nulle, pour μ entre 2 et 3 elle est constamment égale à 1 .

5.4.4 Montrer l'existence de cycles super-attractifs

Le résultat fondamental est le suivant :

5.20 Proposition. Soient α, β vérifiant $3 < \alpha < \beta < 4$. On suppose que les suites de mixages $K(\alpha)$ et $K(\beta)$ ne sont pas égales. Alors, il existe $\mu \in]\alpha, \beta[$ tel que f_μ admette un cycle super-attractif.

⁵⁵On note $[x]$ la partie entière de x .

⁵⁶On dit kneading sequence en anglais, mixage est ma proposition de traduction, mais j'ai vu aussi la version "suite de pétrissage" dans une note CRAS de Gillot.

Démonstration. Dire que les suites ne sont pas égales c'est dire qu'on a, par exemple, $s_n = 0$ et $t_n = 1$, ou encore $f_\alpha^n(1/2) < 1/2$ et $f_\beta^n(1/2) > 1/2$, soit $P_n(\alpha) < 0$ et $P_n(\beta) > 0$. On conclut par les valeurs intermédiaires.

5.21 Corollaire. *Pour tout nombre premier $n \geq 2$ il existe un $\mu \in]3, 4[$ tel que f_μ admette un n -cycle⁵⁷ super-attractif.*

Démonstration. En effet, on a $K(3)_n = 1$ et $K(4)_n = 0$.

Ce qui précède montre que le théorème cherché est une conséquence du résultat suivant :

5.22 Théorème. *Soient $\mu_\infty < \alpha, \beta < 4$ deux réels⁵⁸ distincts. Alors on a $K(\alpha) \neq K(\beta)$, ou, à défaut, il existe $\alpha', \beta' \in [\alpha, \beta]$ qui vérifient cette condition.*

Que cette assertion soit vraie résulte de [Graczyk-Swiatek] (mais c'est tout à fait non trivial). On peut s'en convaincre expérimentalement⁵⁹. Ainsi, on voit que les suites de mixage $K(3, 91)$ et $K(3, 92)$ diffèrent pour $n = 8$, de sorte qu'on doit avoir un point attractif de période 8 pour un paramètre μ compris entre ces valeurs. Entre 3, 9157863 et 3, 9157864 les suites diffèrent au 26-ème coup. On a donc quelque part un point de période 26 super-attractif entre les deux⁶⁰. Plus près de μ_∞ , j'ai fait l'expérience entre 3, 57 et 3, 571 et ça marche au 49-ième coup.

5.4.5 Un résultat du côté de 4

La méthode élémentaire utilisant la suite de mixage que j'ai indiquée ci-dessus ne donne certes pas le théorème d'hyperbolicité, mais elle permet toutefois de montrer le résultat facile suivant. Ce résultat n'est pas évident *a priori*, puisqu'on sait par Jakobson que pour μ tendant 4, les valeurs de μ qui mènent au chaos tendent à être de mesure 1 :

5.23 Proposition. *Soit a un réel avec $3 < a < 4$. Il existe μ dans l'intervalle $]a, 4[$ tel que f_μ admette un cycle super-attractif. Autrement dit, il y a des valeurs de μ arbitrairement proches de 4 pour lesquelles f_μ est hyperbolique.*

⁵⁷Si n n'est pas premier ce cycle est *a priori* d'ordre diviseur de n seulement.

⁵⁸Le résultat est évidemment faux pour $\alpha, \beta < \mu_\infty$.

⁵⁹Mais certaines expériences ne sont pas convaincantes. Par exemple, les suites de mixages de 3, 835 et 3, 836 (resp. 3, 835 et 3, 8355) sont égales jusqu'au terme d'indice 8000 (resp. 40000). On voit que les choses ne sont pas simples.

⁶⁰Un calcul par approximation successives montre que μ est voisin de 3, 91578630770274. Avec cette valeur MuPad donne des valeurs de u_{500} et u_{526} qui coïncident jusqu'à la 40-ième décimale. Cette valeur ne correspond pas tout à fait à un cycle super-attractif, mais elle en est suffisamment proche pour qu'il y ait un cycle d'ordre 26 attractif.

Démonstration. On considère $P_n(\mu)$. On a $P_n(4) = -1/2$ pour $n \geq 2$. En effet, on a $f(1/2) = 1$, donc $f^2(1/2) = 0$ et on reste indéfiniment en 0 : $f^n(1/2) = 0$ pour $n \geq 2$, ou encore $P_n(4) = -1/2$. S'il existe un $n \geq 2$ tel que $P_n(a) > 0$, il y a une racine μ entre a et 4 et on a gagné. Sinon, cela signifie que, pour tout $n \geq 2$, $u_n = f^n(1/2)$ reste $< 1/2$. Mais alors la suite u_n est croissante et majorée par $1/2$, donc converge, ce qui est absurde puisque les points fixes de f sont 0 et $\tau > 1/2$.

6 Mesures invariantes

Ce paragraphe a pour objectif d'éclairer le lecteur, voire l'auteur lui-même, sur la formulation des théorèmes affirmant l'existence du chaos, tels qu'on les trouve dans la littérature. En effet, ces résultats, comme par exemple le théorème de Jakobson 5.12, sont énoncés le plus souvent en prouvant l'existence d'une mesure invariante ergodique pour l'application f . Nous allons voir que cette situation est bien synonyme de chaos. Les résultats de cette partie seront parfois donnés sous une forme un peu elliptique. On renvoie le lecteur à [Walters] ou [Halmos]⁶¹ pour toutes précisions.

Dans toute cette section on considère un intervalle compact I et une application continue $f : I \rightarrow I$.

6.1 Introduction

6.1.1 Définitions

Pour le confort du lecteur, donnons quelques définitions dont l'intérêt n'apparaîtra que dans un instant :

6.1 Définition. Soit ν une mesure positive sur l'intervalle I .

1) On dit que ν est une mesure **de probabilité** si sa masse totale $\nu(I)$ est égale à 1.

2) On dit que ν est **absolument continue** par rapport (resp. **équivalente**) à la mesure de Lebesgue dx si elle est de la forme $\nu = h(x)dx$ où h est une fonction intégrable ≥ 0 (resp. presque partout > 0).

3) On dit que ν est **invariante** par f si on a, pour tout A mesurable, $\nu(A) = \nu(f^{-1}(A))$.

4) On dit que ν est **ergodique** par rapport à f si tout ensemble A qui vérifie $f^{-1}(A) = A$ est de mesure 0 ou 1.

⁶¹Le petit livre de Halmos est particulièrement agréable.

6.1.2 La problématique : systèmes dynamiques ergodiques

Lorsqu'on étudie les suites récurrentes $u_{n+1} = f(u_n)$, en faisant varier $u_0 \in I$, et que leur comportement n'est pas simple, on peut chercher si ces suites présentent au moins une régularité statistique. L'idée est de regarder si la suite passe plus fréquemment dans telle ou telle partie de I . Cela conduit à la définition suivante :

6.2 Définition. (Fréquence de visite) Soit $x \in I$, $A \subset I$ et $n \in \mathbf{N}$. On note $N(x, A, n)$ le cardinal de l'ensemble :

$$\{k \in \mathbf{N} \mid k < n \text{ et } f^k(x) \in A\}.$$

La fréquence $F(x, A)$ de visite de l'orbite de x dans A est la limite du rapport $N(x, A, n)/n$ quand n tend vers l'infini (si cette limite existe).

Avec cette notion, on peut définir un système dynamique **statistiquement régulier** ou **ergodique**⁶² :

6.3 Définition. Soit $f : I \rightarrow I$ une application continue. On dit que la dynamique associée à f est **ergodique** (ou que f est ergodique) si la fréquence de visite $F(x, A)$ existe et ne dépend (presque) pas du point de départ x , c'est-à-dire si, pour A fixé, la fonction $x \mapsto F(x, A)$ est presque constante au sens de la mesure de Lebesgue. On note alors $\nu(A)$ cette constante.

6.4 Remarques.

- 1) La définition précédente signifie bien que le système considéré se comporte statistiquement de manière régulière : quel que soit le point de départ, ou presque, le système va visiter aussi souvent chaque partie A de I .
- 2) Le mot "presque" est essentiel pour se prémunir contre les points qui ont un comportement particulier. Le lecteur pensera au cas des points périodiques ou stationnaires du cas $\mu = 4$.
- 3) Si f n'est pas ergodique c'est qu'il y a plusieurs zones dans I avec des comportements radicalement différents. C'est le cas par exemple de la fonction $f(x) = \sqrt[3]{x}$ sur $[-1, 1]$ pour laquelle $f^n(x)$ tend vers 1 si x est > 0 et vers -1 si x est < 0 . On voit aisément que la valeur de $F(x, [1/2, 1])$ est 1 si x est > 0 et 0 sinon, et inversement pour $F(x, [-1, -1/2])$. Dans ce cas, on pourra en général décomposer le système en plusieurs systèmes stables et ergodiques (comme ici avec $x \geq 0$ et $x \leq 0$).

⁶²Cette appellation n'est pas homologuée !

6.2 Le lien entre systèmes ergodiques et mesures invariantes : le théorème de Birkhoff

La proposition suivante fait le lien entre la notion de fonction ergodique et celle de mesure invariante :

6.5 Proposition. *On suppose que f est ergodique. Alors, l'application ν définie en 6.3 est une mesure de probabilité sur I , invariante par f et ergodique.*

Démonstration. Montrons déjà que ν est une mesure sur I , de masse totale 1. En effet, si A et B sont disjoints on a $N(x, A \cup B, n) = N(x, A, n) + N(x, B, n)$ pour tous x et n et on en déduit $\nu(A \cup B) = \nu(A) + \nu(B)$. De plus, on a $N(x, I, n) = n$, donc $\nu(I) = 1$.

Cette mesure est invariante par f . En effet, $N(x, f^{-1}(A), n)$ est égal à $N(x, A, n+1)$ à une unité près et il en résulte $\nu(f^{-1}(A)) = \nu(A)$.

Enfin, la mesure est ergodique. En effet, supposons $A = f^{-1}(A)$ et $\nu(A) > 0$. On a aussi $f(A) = A$ et on voit aussitôt que pour une partie stable A et un x de A , on a $F(x, A) = 1$. Comme on a, pour presque tout x de A , $F(x, A) = \nu(A)$, on voit que $\nu(A)$ est bien égal à 1.

6.6 Remarques.

- 1) Si on suppose que presque toutes les suites récurrentes associées à f convergent vers un même point τ , on obtient pour ν la mesure de Dirac associée à τ .
- 2) Si on suppose que presque toutes les suites récurrentes associées à f convergent vers un cycle attractif d'ordre n on obtient une mesure équirépartie sur les n points du cycle.
- 3) Pour le cas chaotique, on a le résultat suivant, qui explique le lien entre les mesures invariantes et le chaos (voir aussi 6.14).

6.7 Corollaire. *Dans la situation de 6.5, si l'on suppose de plus que la mesure ν est équivalente à la mesure de Lebesgue, la dynamique de f est chaotique.*

Démonstration. Soit J un intervalle non vide contenu dans I . Comme ν est absolument continue par rapport à la mesure de Lebesgue, on a $\nu(J) > 0$, donc, pour presque tout x , $F(x, J) > 0$. Cela implique en particulier que l'orbite d'un tel x rencontre J . On en déduit l'existence d'orbites denses en utilisant une base dénombrable d'intervalles ouverts.

La réciproque de 6.5 est vraie, mais plus difficile. C'est une partie d'un célèbre théorème de Birkhoff (1932) :

6.8 Théorème. (Birkhoff) *On suppose que f admet une mesure de probabilité ν invariante et ergodique. Alors, pour presque tout $x \in I$, $F(x, A)$ existe et on a $F(x, A) = \nu(A)$, autrement dit f est ergodique.*

Démonstration. Voir [Halmos] ou [Walters].

6.3 Exemples de mesures invariantes

6.3.1 Un critère

Il est en général très difficile de montrer qu'il existe des mesures invariantes ergodiques. Voici déjà un critère pour l'invariance :

6.9 Proposition.

1) *Pour montrer que ν est invariante par f il suffit de montrer la formule $\nu(A) = \nu(f^{-1}(A))$ lorsque A est un intervalle.*

2) *On suppose que f est la fonction logistique f_μ et que la fonction h vérifie pour tout x , $2h(x) = h(f(x))|f'(x)|$. Alors, la mesure $\nu = h(x)dx$ est invariante par f .*

Démonstration. 1) En effet, on obtient ensuite les boréliens, puis les mesurables.

2) Si $J = [a, b]$ est un intervalle, $f^{-1}(J)$ est formé de deux intervalles symétriques par rapport à $1/2$, $J_1 = [c, d]$ (avec $d \leq 1/2$) et J_2 . On a donc $\nu(f^{-1}(J)) = \nu(J_1) + \nu(J_2)$. Sur l'intervalle J_k , f est un C^1 -difféomorphisme et on peut faire le changement de variables $x = f(t)$. On obtient $\nu(J) = \int_J h(x)dx = \int_{J_1} h(f(t))|f'(t)|dt + \int_{J_2} h(f(t))|f'(t)|dt$. Avec la formule $2h(x) = h(f(x))|f'(x)|$, on en déduit $\nu(J) = 2 \int_{J_1} h(t)dt + 2 \int_{J_2} h(t)dt = 2 \int_{J_1} h(t)dt + 2 \int_{J_2} h(t)dt = \nu(f^{-1}(J))$.

6.3.2 Un exemple : l'application tente

Rappelons, cf. 4.12, que l'application "tente" est l'application $t : [0, 1] \rightarrow [0, 1]$ définie par $t(x) = 2x$ pour $0 \leq x \leq 1/2$ et $t(x) = 2 - 2x$ pour $1/2 \leq x \leq 1$. Son intérêt vient du résultat suivant :

6.10 Théorème. *La mesure de Lebesgue sur $[0, 1]$ est invariante par l'application t et ergodique.*

Démonstration. L'invariance est évidente (on prend $h(x) = 1$ et la formule $2h(x) = h(t(x))|t'(x)|$ est vérifiée car on a $t'(x) = \pm 2$ pour tout x). Il reste à voir l'ergodicité. On considère un ensemble $A \subset I$, mesurable, vérifiant $A = t^{-1}(A)$ et il s'agit de montrer qu'il est de mesure 0 ou 1. Soit χ la

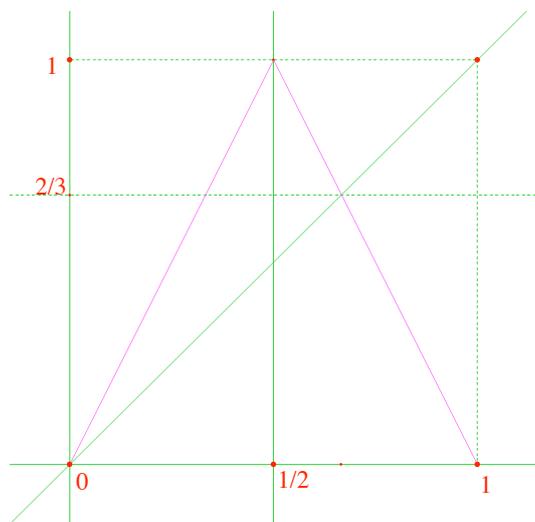


Figure 15: L'application tente

fonction caractéristique de A . Elle vérifie $\chi(t(x)) = \chi(x)$ pour tout $x \in I$. On en déduit, avec $x = 1$, qu'on a $\chi(1) = \chi(0)$. On peut prolonger χ en une fonction périodique de période 1 sur \mathbf{R} , appartenant à $L^2(I)$. Si l'on montre que χ est constante presque partout, on a gagné car A ou $I - A$ sera de mesure nulle.

Montrons que χ vérifie $\chi(x) = \chi(1 - x)$ pour $x \in I$. En effet, on a $\chi(x) = \chi(t(x))$, $\chi(1 - x) = \chi(t(1 - x))$ et le résultat découle de $t(x) = t(1 - x)$. Notons qu'on en déduit que χ est paire. On a alors, pour tout $x \in I$, $\chi(2x) = \chi(x)$. C'est clair pour $x \in [0, 1/2]$. Pour $x \in [1/2, 1]$ on a $\chi(x) = \chi(2 - 2x) = \chi(1 - (2 - 2x)) = \chi(2x - 1)$. Comme χ est de période 1, on a donc $\chi(x) = \chi(2x)$. Il reste à montrer le lemme suivant :

6.11 Lemme. *Soit χ une fonction périodique de période 1, paire, appartenant à $L^2(I)$. On suppose qu'on a, pour tout x , $\chi(x) = \chi(2x)$. Alors χ est constante presque partout.*

Démonstration. Il suffit de montrer que les coefficients de Fourier $\widehat{\chi}(n)$ sont nuls pour $n \neq 0$. Or, on a :

$$\widehat{\chi}(n) = \int_0^1 \chi(x) e^{-2i\pi n x} dx = 2 \int_0^{\frac{1}{2}} \chi(x) e^{-2i\pi n x} dx.$$

On calcule la première intégrale avec le changement de variable $x = 2u$. On trouve $\widehat{\chi}(n) = 2 \int_0^{\frac{1}{2}} \chi(2u) e^{-4i\pi n u} du$ et, en tenant compte de l'invariance, on obtient $\widehat{\chi}(n) = \widehat{\chi}(2n)$ pour tout n . On en déduit par récurrence $\widehat{\chi}(n) = \widehat{\chi}(2^p n)$. Lorsque p tend vers $+\infty$, $2^p n$ tend vers l'infini si n est non nul. Comme les coefficients de Fourier tendent vers 0 à l'infini en vertu de Parseval, on en déduit $\widehat{\chi}(n) = 0$ pour tout n non nul.

6.3.3 Le cas de la fonction logistique pour $\mu = 4$

On a vu en 4.12 que la fonction logistique $f_4(x) = 4x(1-x)$ est conjuguée de l'application tente. On en déduit le théorème :

6.12 Théorème. *On suppose qu'on a $I = [0, 1]$ et $f(x) = 4x(1-x)$. On pose : $h(x) = \frac{1}{\pi\sqrt{x(1-x)}}$.*

La fonction h est intégrable, la mesure $\nu = h(x)dx$ est absolument continue par rapport à la mesure de Lebesgue, invariante par f et ergodique.

Démonstration. Avec les notations de 4.12, posons $\psi(x) = \varphi^{-1}(x) = \frac{2}{\pi} \text{Arcsin } \sqrt{x}$. On vérifie qu'on a $\psi'(x) = h(x)$. Si on dérive la formule $t \circ \psi(x) = \psi \circ f(x)$ on obtient : $t'(\psi(x))h(x) = h(f(x))f'(x)$, soit $2h(x) = h(f(x))|f'(x)|$ en prenant les valeurs absolues. On conclut que la mesure ν est invariante en vertu de 6.9.

Montrons l'ergodicité de ν . Soit $A \subset I$ tel que $f^{-1}(A) = A$. On en déduit $t^{-1}(\varphi^{-1}(A)) = \varphi^{-1}(A)$ par conjugaison. Il en résulte que $\varphi^{-1}(A)$ est de mesure de Lebesgue nulle en vertu de 6.10, donc aussi A et c'est encore vrai pour ν qui est absolument continue par rapport à la mesure de Lebesgue.

Notons au passage qu'on retrouve l'identité de 4.23 :

6.13 Proposition. *Pour tout $x \in]0, 1[$ on a la formule*

$$2^n h(x) = h(f^n(x))|f'(x)||f'(f(x))| \cdots |f'(f^{n-1}(x))| = h(f^n(x))|(f^n)'(x)|.$$

En particulier, si x est un point de période n on a $|(f^n)'(x)| = 2^n$.

Démonstration. On procède par récurrence à partir de la formule d'invariance. Si x est de période n on a $f^n(x) = x$ donc $h(x) = h(f^n(x))$ et on simplifie par cette quantité (non nulle). (On notera que h n'est pas définie en 0 et 1.)

6.4 Une mesure invariante et ergodique impose le chaos

Le point essentiel qui fait l'importance de la notion de mesure invariante ergodique dans la théorie des systèmes dynamiques est le théorème suivant, qui résulte de 6.7 et du théorème de Birkhoff 6.8, mais que nous montrons directement ici.

6.14 Proposition. *On suppose qu'il existe une mesure de probabilité, μ , équivalente à la mesure de Lebesgue, invariante par f et ergodique. Alors,*

l'ensemble R des $x \in I$ dont l'orbite $O(x) = \{f^k(x), k \in \mathbf{N}\}$ est dense dans I est de mesure 1 (et donc non vide). En particulier, la dynamique de f est chaotique.

Démonstration. On considère une base dénombrable $(U_n)_{n \in \mathbf{N}}$ d'ouverts de I . On pose $A = \bigcap_{n \in \mathbf{N}} \bigcup_{k \in \mathbf{N}} f^{-k}(U_n)$. Dire que x est dans R , donc que l'orbite

$O(x)$ est partout dense, c'est dire que tout U_n contient un point de $O(x)$ ou encore qu'il existe $k \in \mathbf{N}$ tel que $f^k(x) \in U_n$. C'est exactement la même chose de dire que x est dans A . On a donc $R = A$ et il s'agit de calculer la mesure de A . Mais, si on pose $V_n = \bigcup_{k \in \mathbf{N}} f^{-k}(U_n)$, on a $f^{-1}(V_n) \subset V_n$.

L'ergodicité de la mesure montre que V_n est de mesure 0 ou 1, voir le lemme ci-dessous. Comme il contient U_n qui est un ouvert non vide, donc de mesure (de Lebesgue ou invariante, c'est pareil) non nulle, on voit que V_n est de mesure 1. L'intersection A des V_n est alors aussi de mesure 1 (en effet, si on appelle Z_n le complémentaire de V_n , il est de mesure nulle et il en est de même de la réunion des Z_n).

6.15 Lemme. *On suppose qu'on a une mesure μ invariante et ergodique et un A tel que $f^{-1}(A) \subset A$. Alors, on a $\mu(A) = 0$ ou 1.*

Démonstration. Posons $B = A - f^{-1}(A)$. L'ensemble B est l'ensemble des $x \in A$ tels que $f(x) \notin A$. Comme la mesure est invariante, on a $\mu(A) = \mu(f^{-1}(A))$ et donc $\mu(B) = 0$. On considère alors $\widehat{B} = B \cup f^{-1}(B) \dots \cup f^{-k}(B) \dots$. Comme la mesure est invariante, les $f^{-k}(B)$ sont de mesure nulle, donc aussi \widehat{B} . On a $\widehat{B} \subset A$. Je dis que $A - \widehat{B}$ vérifie $f^{-1}(A - \widehat{B}) = A - \widehat{B}$. Si x est dans A et pas dans \widehat{B} , il s'agit de voir qu'il en est de même de $f(x)$. Comme x n'est pas dans B , on a $f(x) \in A$. Si $f(x)$ était dans \widehat{B} on aurait un k tel que $f^k(f(x)) \in B$, mais alors x serait dans \widehat{B} . Inversement, si $f(x)$ est dans $A - \widehat{B}$, x est dans $f^{-1}(A)$, donc dans A . S'il était dans \widehat{B} , de deux choses l'une. Ou bien x est dans B , mais alors $f(x)$ n'est pas dans A et c'est absurde. Ou bien il existe $k > 0$ avec $x \in f^{-k}(B)$ donc $f^k(x) \in B$. Mais alors on a $f^{k-1}(f(x)) \in B$ et $f(x) \in \widehat{B}$: contradiction.

On en déduit $\mu(A - \widehat{B}) = 0$ ou 1. Comme on a $\mu(A) = \mu(A - \widehat{B}) + \mu(\widehat{B}) = \mu(A - \widehat{B})$, on a le résultat.

6.5 Un autre exemple de mesure invariante : l'exemple de Ruelle

6.5.1 L'exemple

L'exemple suivant, dû à Ruelle (voir [Ruelle]) est l'un des premiers exemples de dynamique chaotique pour la fonction logistique f_μ avec $\mu < 4$. Le principe est le suivant. On cherche pour quel μ le troisième itéré de $1/2$ tombe sur le point fixe $1 - \frac{1}{\mu}$. On a $f^2(1/2) = f(\mu/4) = \frac{\mu^2(\mu - 4)}{16}$ et il y a deux façons de tomber sur le point fixe au coup suivant : soit on y tombe tout de suite, ce qui mène à l'équation $\mu^4 - 4\mu^3 + 16\mu - 16 = 0 = (\mu - 2)^3(\mu + 2)$, soit on tombe sur $1/\mu$ qui s'envoie sur le point fixe au coup suivant. On obtient alors l'équation $\mu^4 - 4\mu^3 + 16 = 0$ qui admet encore la racine 2 et une autre, racine de l'équation $(\mu - 2)^2(\mu + 2) = 16$. On trouve $\mu = 3,67857351043$. Ruelle a montré alors que f_μ envoie l'intervalle $[1/\mu, 1 - 1/\mu]$ sur $[1 - 1/\mu, \mu/4]$ et que f_μ admet une mesure invariante sur l'intervalle stable minimum (voir Annexe 7.4) $[1/\mu = f(\mu/4), \mu/4]$. Avec $u_0 = 0.4$, le résultat est spectaculaire.

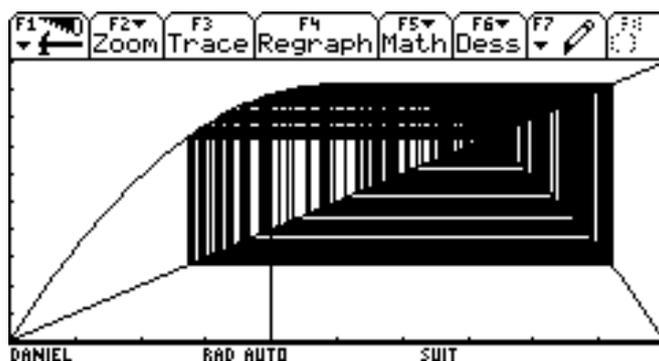


Figure 16: Le chaos de Ruelle

6.5.2 Un exemple analogue

Il n'est pas évident de prouver que la dynamique de l'exemple de Ruelle est chaotique, voir [Ruelle]. En revanche on y parvient⁶³ pour une fonction affine par morceaux telle que $f^3(1/2)$ est le point fixe :

6.16 Proposition. *On considère la fonction f affine par morceaux définie par $f(x) = 2/3 + 2x/3$ pour $x \leq 1/2$ et $f(x) = 2 - 2x$ pour $x \geq 1/2$. Alors, pour tout intervalle $J \subset I$ non réduit à un point, il existe un entier m tel que $\bigcup_{k=0}^m f^k(J) = I$. L'application f est transitive, donc chaotique (cf. 4.4).*

Démonstration. On note ν la mesure de Lebesgue sur I .

⁶³La preuve de ce résultat m'a été soufflée par Sylvie Ruelle.

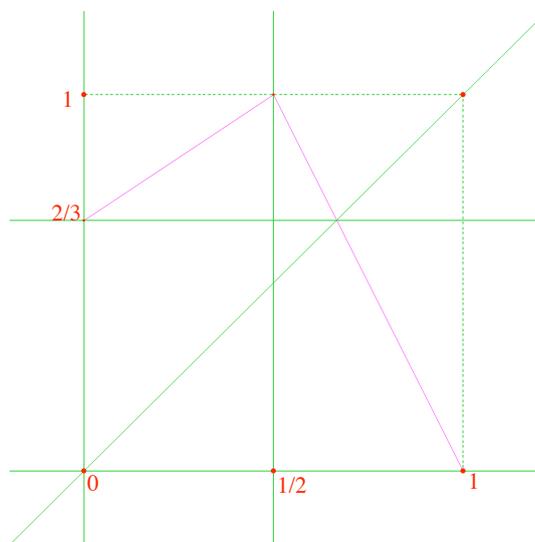


Figure 17: La tente de Ruelle

1) Supposons que $J =]a, b[$ contienne le point fixe $2/3$. Tous les intervalles $f^n(J)$ contiennent aussi $2/3$. Il suffit de montrer qu'il existe un entier n tel que $f^n(J)$ contient 1 (car le suivant contiendra 0 et leur union sera égale à I). Comme on a $f(3/4) = 1/2$ et $f(1/2) = 1$, il suffit même de montrer qu'un itéré de J contient $1/2$ ou $3/4$. Sinon, c'est que tous les $f^n(J)$ sont contenus dans $]1/2, 3/4[$. Mais, sur cet intervalle, on a $f(x) = 2 - 2x$ et f multiplie par 2 les distances. On a donc $\nu(f^n(J)) \geq 2^n \nu(J)$ et cet intervalle ne peut être inclus dans $]1/2, 3/4[$ pour n assez grand.

2) Si l'intervalle J contient 0, ou 1, ou $1/2$, la conclusion de la proposition est encore vraie (car on a $f(1/2) = 1$, $f(1) = 0$ et $f(0) = 2/3$).

3) On raisonne par l'absurde en supposant qu'il existe un intervalle J de longueur > 0 qui ne vérifie pas la conclusion de la proposition. Soit l la borne supérieure des longueurs des contre-exemples. Il existe un contre-exemple J de longueur $\lambda > 3l/4$. Cet intervalle est contenu, soit dans $]0, 1/2[$, soit dans $]1/2, 1[$. Dans ce deuxième cas, $f(J)$ est de longueur 2λ (car f est affine de multiplicateur 2 sur l'intervalle). C'est encore un contre-exemple et on a $2\lambda > 3l/2 > l$, ce qui est absurde. Si au contraire J est contenu dans $]0, 1/2[$, $f(J)$ est de longueur $2\lambda/3$, contenu dans $]2/3, 1[$ et $f^2(J)$ est de longueur $4\lambda/3$ et c'est encore un contre-exemple. Mais on a $4\lambda/3 > l$ et c'est absurde.

6.17 Remarque. On peut montrer que la fonction logistique de l'exemple de Ruelle et la fonction f de la proposition précédente sont conjuguées, ce qui montre que la dynamique de l'exemple de Ruelle est bien chaotique.

7 Annexes

7.1 La table de mortalité utilisée par Euler

❁		152		❁	
(1) = 0,804	(31) = 0,499	(61) = 0,264	(91) = 0,006		
(2) = 0,768	(32) = 0,490	(62) = 0,254	(92) = 0,004		
(3) = 0,736	(33) = 0,482	(63) = 0,245	(93) = 0,003		
(4) = 0,709	(34) = 0,475	(64) = 0,235	(94) = 0,002		
(5) = 0,688	(35) = 0,468	(65) = 0,225	(95) = 0,001		
(6) = 0,676	(36) = 0,461	(66) = 0,215			
(7) = 0,664	(37) = 0,454	(67) = 0,205			
(8) = 0,653	(38) = 0,446	(68) = 0,195			
(9) = 0,646	(39) = 0,439	(69) = 0,185			
(10) = 0,639	(40) = 0,432	(70) = 0,175			
(11) = 0,633	(41) = 0,426	(71) = 0,165			
(12) = 0,627	(42) = 0,420	(72) = 0,155			
(13) = 0,621	(43) = 0,413	(73) = 0,145			
(14) = 0,616	(44) = 0,406	(74) = 0,135			
(15) = 0,611	(45) = 0,400	(75) = 0,125			
(16) = 0,606	(46) = 0,393	(76) = 0,114			
(17) = 0,601	(47) = 0,386	(77) = 0,104			
(18) = 0,596	(48) = 0,378	(78) = 0,093			
(19) = 0,592	(49) = 0,370	(79) = 0,082			
(20) = 0,584	(50) = 0,362	(80) = 0,072			
(21) = 0,577	(51) = 0,354	(81) = 0,063			
(22) = 0,571	(52) = 0,345	(82) = 0,054			
(23) = 0,565	(53) = 0,336	(83) = 0,046			
(24) = 0,559	(54) = 0,327	(84) = 0,039			
(25) = 0,552	(55) = 0,319	(85) = 0,032			
(26) = 0,544	(56) = 0,310	(86) = 0,026			
(27) = 0,535	(57) = 0,301	(87) = 0,020			
(28) = 0,525	(58) = 0,291	(88) = 0,015			
(29) = 0,516	(59) = 0,282	(89) = 0,011			
(30) = 0,507	(60) = 0,273	(90) = 0,008			

7.2 Mise en équation logistique

Voici la justification adoptée en général pour valider le modèle logistique. On postule qu'il y a, pour la population p_n , une valeur idéale, une valeur d'équilibre e et que si la population s'écarte de la valeur d'équilibre, elle tend à y revenir : si p_n est $> e$ (resp. $< e$), on a $p_{n+1} - p_n < 0$ (resp. > 0) et le retour de balancier est d'autant plus fort qu'on s'est plus écarté de e . En termes mathématiques, on postule que le taux d'accroissement $\frac{p_{n+1} - p_n}{p_n}$ est proportionnel à l'écart $p_n - e$, avec un coefficient < 0 :

$$\frac{p_{n+1} - p_n}{p_n} = -k(p_n - e)$$

avec $k > 0$ ⁶⁴. Avec le choix opéré ci-dessus on a donc :

$$p_{n+1} = (ek + 1)p_n - kp_n^2$$

et on retrouve une relation de la forme $p_{n+1} = \mu p_n - \lambda p_n^2$ comme en 1.5 ci-dessus.

7.3 Conjugaison des fonctions du second degré

Le problème est le suivant : on considère l'ensemble E des fonctions $f(x) = ax^2 + bx + c$, de degré 2, c'est-à-dire vérifiant $a \neq 0$. On note f_0 le polynôme x^2 . Sur cet ensemble, on fait opérer le groupe affine de la droite, G , formé des applications $h(x) = \alpha x + \beta$ avec $\alpha \neq 0$. Il opère sur E par conjugaison, donc en posant $h.f = hf h^{-1}$. Un argument de dimension montre que le quotient est de dimension 1. On peut le représenter en utilisant trois types particuliers de polynômes de degré 2 dépendant d'un paramètre : les polynômes $x^2 + \gamma$, les polynômes $1 - \lambda x^2$ avec $\lambda \neq 0$ et enfin les polynômes $\mu x - \mu x^2$ avec $\mu \neq 0$. On note respectivement C, L, M les ensembles de ces polynômes. On a alors le théorème suivant :

7.1 Théorème. *L'ensemble quotient E/G est en bijection avec C . Précisément, si $f(x) = ax^2 + bx + c$ est dans E , il existe α, β, γ uniques tels que $hf h^{-1} = g$ avec $g(x) = x^2 + \gamma$. On a $\alpha = a$, $\beta = b/2$ et $\gamma = ac + \frac{b}{2} - \frac{b^2}{4}$.*

Le polynôme $f(x) = -\lambda x^2 + 1$ est conjugué de $x^2 - \lambda$ par $h(x) = -\frac{1}{\lambda}x$, de sorte que l'ensemble $E/G - \{f_0\}$ est en bijection avec L .

Le polynôme $\mu x - \mu x^2$ est conjugué de $x^2 + \frac{\mu}{2} - \frac{\mu^2}{4}$ par $h(x) = -\mu x + \frac{\mu}{2}$ et les polynômes $\mu x - \mu x^2$ et $\nu x - \nu x^2$ sont conjugués si et seulement si on

⁶⁴On notera que ce n'est pas le seul modèle possible. On aurait pu aussi imposer la formule $p_{n+1} - p_n = -k(p_n - e)$ qui conduit à une suite arithmético-géométrique.

$a \nu = 2 - \mu$. L'ensemble E/G est en bijection avec le quotient de M par l'involution $\mu \mapsto 2 - \mu$, ou encore avec $M^+ = \{f \in M \mid \mu \geq 1\}$.

Démonstration. C'est facile : on se donne f et on résout en α, β, γ l'équation $hf = gh$ avec $g(x) = x^2 + \gamma$. Pour le dernier point on résout $\mu/2 - \mu^2/4 = \nu/2 - \nu^2/4$.

7.2 Remarque. En vérité, le paramétrage par $x^2 + \gamma$ doit être plutôt plus commode. Je n'en veux pour preuve que les valeurs de γ qui correspondent aux valeurs remarquables de μ : $\mu = 4$ donne $\gamma = -2$, $\mu = 2$ donne $\gamma = 0$, $\mu = 3$, correspond à $\gamma = -3/4$, et les valeurs suivantes de γ : $-1, -5/4, -7/4$ correspondent à $\mu = 1 + \sqrt{5}, 1 + \sqrt{6}, 1 + \sqrt{8} = 1 + 2\sqrt{2}$.

7.3 Remarque. Le fait que les fonctions f_μ , pour $\mu \geq 1$ ne soient pas deux à deux conjuguées est l'une des raisons qui fait que la famille logistique est un exemple particulièrement riche, voir [Lyubich2].

7.4 Intervalles stables

On a vu que l'intervalle $I = [0, 1]$ est stable par f_μ . La question est ici l'existence d'autres intervalles $[a, b]$ stables avec $0 \leq a < b \leq 1$.

7.4 Proposition.

1) On suppose $\mu \leq 1$. Les intervalles stables sont les intervalles $[0, b]$ avec $0 < b \leq 1$.

2) On suppose $1 < \mu \leq 2$. Les intervalles stables sont les intervalles $[a, b]$ avec $0 \leq a \leq \tau \leq b \leq 1/2$ et les intervalles $[a, b]$ avec $b \geq 1/2$, $a \leq \tau$ et $a \leq \mu b(1 - b)$.

3) On suppose $2 < \mu \leq 3$. Les intervalles stables sont :

- Les intervalles $[a, b]$ avec $1/2 \leq a \leq \tau$ et

$$(*) \quad \mu a(1 - a) \leq b \leq \frac{1}{2} + \frac{1}{2} \sqrt{\frac{\mu - 4a}{\mu}}.$$

- Les intervalles $[a, b]$ avec $a \leq 1/2 \leq \tau$ et

$$(**) \quad \tau \leq \frac{\mu}{4} \leq b \leq \frac{1}{2} + \frac{1}{2} \sqrt{\frac{\mu - 4a}{\mu}}.$$

4) On suppose $3 < \mu \leq 1 + \sqrt{5}$. Les intervalles stables sont de deux types :

- Les intervalles $[a, b]$ avec $1/2 \leq a \leq x_1 \leq \tau$ où x_1 est le plus petit des points du 2-cycle de f (voir 3.8) et la condition (*) ci-dessus.

- Les intervalles $[a, b]$ avec $a \leq 1/2 \leq \tau$ et la condition $(**)$ ci-dessus.
- 5) On suppose $1 + \sqrt{5} < \mu < 4$. Les intervalles stables sont les intervalles $[a, b]$ avec $a \leq f(\frac{\mu}{4}) = \frac{\mu^2(4-\mu)}{16} < \frac{1}{2}$ et la condition $(**)$ ci-dessus. Parmi ces intervalles, le plus petit est $[f(\frac{\mu}{4}), \frac{\mu}{4}]$.
- 6) On suppose $\mu = 4$. Le seul intervalle stable est I .

Démonstration. On n'oubliera pas que le maximum de f est atteint en $1/2$ et qu'il vaut $\mu/4$. L'image de $[a, b]$ est donc :

- $[f(a), f(b)]$, avec $f(a) \leq f(b)$, si b est $\leq 1/2$,
- $[\text{Min}(f(a), f(b)), \mu/4]$ si on a $a < 1/2 < b$,
- $[f(b), f(a)]$ avec $f(b) \leq f(a)$, si a est $\geq 1/2$.

La condition de stabilité est donc, respectivement, la suivante :

- $a \leq f(a)$ et $f(b) \leq b$,
- $a \leq f(a)$, $a \leq f(b)$ et $\mu/4 \leq b$,
- $a \leq f(b)$ et $f(a) \leq b$.

On notera que les conditions $a \leq f(a)$ et $f(b) \leq b$ sont nécessaires, dans tous les cas.

1) Un intervalle stable contient un point fixe, donc ici 0. On a donc $a = 0$. Si b est $\leq 1/2$ on a $f(b) \leq b$, si b est $> 1/2$, on a $b \geq \mu/4$ (car on a $\mu/4 \leq 1/4$).

Pour $\mu > 1$ on a le point fixe $\tau = 1 - 1/\mu$. Si a est non nul, la condition $a \leq f(a)$ s'écrit $a \leq \tau$. La condition $f(b) \leq b$ s'écrit $b \geq \tau$.

2) Le cas $a \leq \tau \leq b \leq 1/2$ est évident. Pour l'autre, on a la condition $a \leq \tau$ et $f(b) \geq a$ qui donne l'expression annoncée. La condition $\mu/4 \leq b$ est automatique car on a $\mu/4 \leq 1/2$.

3) On traduit les conditions ci-dessus. Le point essentiel est de noter que, pour $b \geq 1/2$, la condition $b \leq \frac{1}{2} + \frac{1}{2}\sqrt{\frac{\mu-4a}{\mu}}$ équivaut à $f(b) \geq a$. On notera que pour $a \leq 1/2$ donné, il y a toujours des valeurs de b convenables.

Traisons à la fois les points 4) et 5). On vérifie aussitôt qu'il n'y a pas d'intervalle stable contenu dans $[0, 1/2]$ (sur cet intervalle on a $f(x) > x$ pour $x > 0$, ce qui empêche $f(b) \leq b$).

Pour les intervalles stables par f contenus dans $[1/2, 1]$ le calcul est le même qu'en 3) et mène à la même inégalité double. La condition pour que l'intervalle définissant b soit non vide est $f^2(a) - a \geq 0$ et, pour $a \leq \tau$, cela signifie $a \leq x_1$, cf. 3.8. Il existe de tels points si et seulement si x_1 est $\geq 1/2$, soit $\mu \leq 1 + \sqrt{5}$.

Enfin, pour le cas $a \leq 1/2 < \tau \leq b$, il suffit de traduire les conditions ci-dessus. La condition pour que l'intervalle de définition de b soit non vide

est $\mu^3 - 4\mu^2 + 16a < 0$. C'est la condition imposée en 5). Elle est inutile dans 4) car on a $\frac{\mu^2(4-\mu)}{16} \geq 1/2$ pour $\mu \leq 1 + \sqrt{5}$.

6) Si J est un intervalle stable non réduit à un point, il contient un point dont l'orbite est dense dans I en vertu de 4.2. On a donc $I = J$.

7.5 Remarque. La conclusion du point 5) de cette proposition est très importante. En effet, on voit que pour $\mu < 4$ on a toujours des intervalles stables $J \subset I$ non triviaux. Cela prouve que, sur I tout entier, la dynamique n'est pas chaotique au sens strict du terme puisque si x est dans J , la suite des $f^n(x)$ reste dans J , donc n'est pas partout dense.

En réalité, toute la dynamique de f se passe dans l'intervalle stable minimum $J = [f(\mu/4), \mu/4]$. En effet, pour tout $x \in I$, on vérifie⁶⁵ que $f^2(x)$ est dans J (car on a $f(x) \leq \mu/4$ et, comme f est décroissante au-delà de $\mu/4$, $f^2(x) \geq f(\mu/4)$).

On voit aussi, pour confirmer ce qui précède, qu'il n'existe pas de mesure invariante absolument continue sur I tout entier. En effet, on aurait alors $\mu(J) = \mu(f^{-2}(J)) = \mu(I)$, d'où $\mu(I - J) = 0$ ce qui est interdit pour une mesure absolument continue. Bref, s'il y a une mesure invariante c'est sur J , ce qui confirme que la dynamique est tout entière dans l'intervalle stable minimum. C'est donc sur cet intervalle stable minimum qu'il faut comprendre les assertions d'existence de mesure invariante ou de chaos du théorème de Jakobson par exemple.

7.5 Période des points dans le cas $\mu = 4$

Rappelons (cf. 4.18) que les points $u_0 \in [0, 1]$ qui sont tels que la suite (u_n) soit périodique de période ≥ 1 sont les points $u_0 = \sin^2 \frac{k\pi}{2^n \pm 1}$, avec $n \in \mathbf{N}^*$ et $0 \leq k \leq 2^{n-1}$ (resp. $0 \leq k < 2^{n-1}$) si le signe est + (resp. si le signe est -).

La question est de savoir si n est la plus petite période du point correspondant. On peut donc supposer $n \geq 2$. On note déjà que si k est nul on a affaire à un point fixe (période 1) et on écarte ce cas. Avec un peu de patience, le lecteur pourra prouver le résultat suivant :

7.6 Théorème. Soit n un entier ≥ 2 . On pose $u_0^+(n, k) = \sin^2 \frac{k\pi}{2^n + 1}$ avec $0 < k \leq 2^{n-1}$ et $u_0^-(n, k) = \sin^2 \frac{k\pi}{2^n - 1}$ avec $0 < k < 2^{n-1}$

⁶⁵En revanche, $f(x)$ n'y est pas forcément. Exemple : $\mu = 3,9$, $\mu/4 = 0,975$, $f(\mu/4) = 0,0950625$, $x = 0,99$, $f(x) = 0,03861$, $f^2(x) \sim 0,144765$.

1) Si n est une puissance de 2, $u_0^+(n, k)$ est un point de période exactement n pour tout k .

2) Si n est premier, $u_0^-(n, k)$ est un point de période exactement n pour tout k .

3) Si k et $2^n + 1$ (resp. $2^n - 1$) sont premiers entre eux, $u_0^+(n, k)$ (resp. $u_0^-(n, k)$) est un point de période exactement n , sauf dans le cas $n = 2$, $k = 1$ où il est de période 1.

4) Si n admet un facteur impair > 1 , il existe k tel que $u_0^+(n, k)$ soit de période $< n$. Précisément, pour tout facteur impair n' de n , distinct de 1, on pose $p = n/n'$, on a alors $2^n + 1 = k(2^p + 1)$ et $u_0^+(n, lk)$ est de période p pour l compris entre 1 et 2^{p-1} .

5) Si n n'est pas premier, il existe k tel que $u_0^-(n, k)$ soit de période $< n$. Précisément, si p est un diviseur de n , avec $p < n$, on a $2^n - 1 = k(2^p - 1)$ et $u_0^-(n, lk)$ est de période p pour $0 < l < 2^{p-1}$. Si $2p$ est un diviseur pair de n , on a $2^n - 1 = k'(2^p + 1)$ et $u_0^-(n, l'k')$ est de période⁶⁶ p pour $0 < l' \leq 2^{p-1}$.

8 Références

La bibliographie sur le sujet est monstrueuse : il y a des centaines d'articles qui tournent autour des systèmes dynamiques associés à un polynôme quadratique. J'ai donc dû faire un choix drastique.

[Buzzi] **Jérôme Buzzi** *Chaos et stabilité* Le pommier (2005).

[Collet-Eckmann] **Pierre Collet** et **Jean-Pierre Eckmann** *Iterated maps on the interval as dynamical systems* Progress in Physics, Birkhäuser (1980).

[Collet-Eckmann-Lanford] **P. Collet ; J.-P. Eckmann ; O.E. Lanford III** *Universal properties of maps on an interval* Commun. Math. Phys. 76, 211-254 (1980).

[Coppel] **William-A. Coppel** *Solutions of equations by iteration* Proceedings of the Cambridge Philosophical Society, Vol. 51, 41-43 (1955).

[Cvitanovic] **Predrag Cvitanović** *Universality in Chaos* Adam Hilger Ltd (1984).

[De Melo-Van Strien] **Wellington de Melo** et **Sebastian van Strien** *One-Dimensional Dynamics* Springer Verlag (1993).

[Devaney] **Robert L. Devaney** *An introduction to chaotic dynamical systems* Addison-Wesley (1989).

⁶⁶Attention, dans tous les cas on affirme que p est une période, pas qu'elle est la plus petite. Mais la procédure permet de se ramener au cas des fractions de dénominateurs $2^p + 1$ ou $2^p - 1$ plus petits que celui de la fraction initiale.

[Euler] **Leonhard Euler** *Recherches generales sur la mortalite et la multiplication du genre humain* Mémoires de l'académie des sciences de Berlin 16, 144-164 (1767) ou Opera Omnia : Series 1, Volume 7, 79 - 100.

[Glass-Mackey] **Leon Glass** et **Michael C. Mackey** *From clocks to chaos : The Rhythms of Life* Princeton University Press (1988).

[Graczyk-Swiatek1] **Jacek Graczyk** et **Grzegorz Swiatek**, *Generic hyperbolicity in the logistic family* Ann. Math. (2) 146, No.1, 1-52 (1997).

[Graczyk-Swiatek2] **Jacek Graczyk** et **Grzegorz Swiatek**, *The real Fatou conjecture* Annals of Mathematics Studies 144, Princeton University Press (1998).

[Halmos] **Halmos Paul R.** *Lectures on ergodic theory* The mathematical society of Japan (1956).

[Jakobson] **M.V. Jakobson** *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps* Commun. Math. Phys. 81, 39-88 (1981).

[Langlois-Daude] **Patrice Langlois** et **Éric Daude**, *Concepts et modélisations de la diffusion géographique*, Revue européenne de géographie, N°364 (2007).

[Lyubich1] **Mikhail Lyubich**, *Dynamics of quadratic polynomials. I, II.* Acta Math. 178, N° 2, 185-297 (1997).

[Lyubich2] **Mikhail Lyubich** *The quadratic family as a qualitatively solvable model of chaos* Notices Am. Math. Soc. 47, N° 9, 1042-1052 (2000).

[Lyubich3] **Mikhail Lyubich**, *Almost every real quadratic map is either regular or stochastic* Ann. Math. (2) 156, N° 1, 1-78 (2002).

[Perrin] **Daniel Perrin** *L'expérimentation en Mathématiques* Petit x, numéro 73 (2007).

[Robinson] **Clark Robinson** *An introduction to dynamical systems* CRC Press, 1995.

[Ruelle] **David Ruelle** *Applications conservant une mesure absolument continue par rapport à dx sur $[0, 1]$* Commun. math. Phys. 55, 47-51 (1977).

[Ruelle] **Sylvie Ruelle** *Chaos for continuous interval maps* (livre en projet)

<http://www.math.u-psud.fr/~ruette/>

[Walters] **Peter Walters** *An Introduction to Ergodic Theory* Springer Verlag, 1982.