Algorithms Seminar 20??–20??,
F. Chyzak (ed.), INRIA, (20??), pp. ??–??.

Available online at the URL
http://algo.inria.fr/seminars/.

Systèmes dynamiques et algorithmique[†]

Viviane Baladi[(a)] *and Brigitte Vallée*[(b)]

[(a)]Institut Mathématique de Jussieu (France) and [(b)]GREYC, Université de Caen (France)

March 18 and 19, 2002

Summary by Frédéric Chazal[‡], *Véronique Maume-Deschamps*[§], *and Brigitte Vallée*[¶]

L'analyse en moyenne d'algorithmes vise à déterminer le comportement « moyen » des algorithmes. Par opposition à la complexité dans le pire des cas, la complexité moyenne d'un algorithme permet d'appréhender les performances de l'algorithme de manière « réaliste ». Il est maintenant classique, en analyse d'algorithmes, de travailler avec un outil essentiel, celui des séries génératrices. Les principales opérations algébriques sur les structures de données ou les algorithmes se traduisent en opérations formelles sur les séries génératrices. Quand les séries génératrices sont vues comme des fonctions de variable complexe, leur singularité dominante permet d'obtenir des renseignements précieux sur le comportement asymptotique moyen de l'algorithme. Cette méthodologie est décrite par exemple dans les livres de Flajolet et Sedgewick[24, 27].

Cependant, quand les algorithmes sont trop « corrélés », cette méthodologie ne peut plus s'appliquer, car les opérations sur les algorithmes ne se traduisent plus aisément en opérations sur les séries génératrices. C'est alors une idée tout à fait naturelle que de considérer un algorithme et l'ensemble de ses données comme un système dynamique. Les données sont alors les particules du système qui sont soumises au « champ » créé par les opérations que leur font subir l'algorithme. À un système dynamique, on associe classiquement, depuis Ruelle, un opérateur appelé opérateur de transfert, ou opérateur de Ruelle, [40, 41] qui permet de décrire l'évolution du système. Cet opérateur dépend d'un paramètre s, est désigné par $\mathbf{H}_s$, et agit sur un espace de fonctions d'une variable.

Opérateur de transfert = opérateur générateur. L'idée originale consiste à détourner l'opérateur de transfert de son usage habituel et à le considérer comme un opérateur « super-générateur », en ce sens qu'il engendre lui-même les séries génératrices associées à l'algorithme. Les opérations sur les algorithmes continuent à se traduire en opérations sur ces opérateurs générateurs. Par ailleurs, aussitôt que le système dynamique possède de « bonnes propriétés », cet opérateur a des propriétés spectrales dominantes : il existe une valeur propre dominante $\lambda(s)$ positive qui est séparée du reste du spectre par un saut spectral. Cette valeur propre dominante joue ainsi un rôle essentiel car c'est elle qui concentre les propriétés essentielles du système. C'est elle qui va jouer le même rôle que la singularité dominante dans le cadre classique des séries génératrices, et va ainsi permettre d'appréhender le comportement asymptotique moyen de l'algorithme, même quand celui-ci est « corrélé ». C'est la philosophie générale (voir Figure 1). De fait, l'opérateur de transfert ne peut

[†]Notes de cours pour le cours donné pendant le groupe de travail ALÉA'02 au Cirm à Luminy (France).

[‡]Université de Bourgogne, B. P. 47870, 21078 Dijon Cedex, France ; email: fchazal@u-bourgogne.fr.

[§]Université de Bourgogne, B. P. 47870, 21078 Dijon Cedex, France ; email: vmaume@u-bourgogne.fr.

[¶]GREYC, Université de Caen, 14032 Caen Cedex, France ; email: brigitte.vallee@info.unicaen.fr.

pas vraiment être utilisé « tel que » en analyse d'algorithmes, il a souvent besoin d'être généralisé, afin d'opérer sur des fonctions de plusieurs variables. Cet opérateur généralisé, désigné par $\mathfrak{H}_s$, étend d'ailleurs dans un sens fort l'opérateur classique $\mathbf{H}_s$, puisqu'il a la même valeur propre dominante.

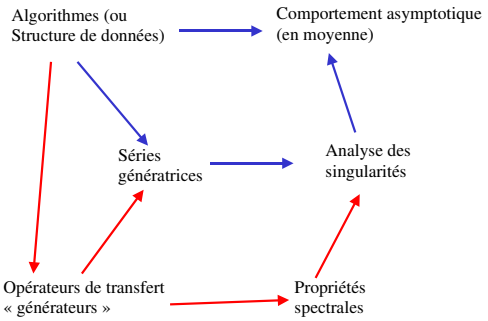

FIG. 1. Analyse classique, analyse dynamique

Les domaines d'application. Cette méthodologie, qu'on appelle « analyse dynamique des algorithmes » s'est installée relativement récemment en analyse d'algorithmes (1995). Elle peut déjà s'appliquer à deux domaines algorithmiques larges, l'algorithmique arithmétique et l'algorithmique du texte. Dans chacun de ces domaines, la méthode prouve son efficacité en permettant de résoudre des problèmes inaccessibles à la méthode classique. La démarche est différente dans les deux domaines : en algorithmique arithmétique, on cherche à analyser des algorithmes existants et utilisés. En algorithmique du texte, il y a une double volonté : on cherche à modéliser le concept de source, qui est le mécanisme sous-jacent à tous les algorithmes de texte, puisque c'est lui qui produit le texte ; on cherche ensuite à analyser les algorithmes quand les textes sont produits sous ce modèle. Bien que les deux domaines soient *a priori* disjoints, il y a de fait un transfert de méthodes de l'un des domaines à l'autre : en algorithmique arithmétique, le concept a été utilisé pour des systèmes dynamiques de plus en plus complexes qui se sont « spontanément » présentés, lors de l'analyse d'algorithmes classiques existant. Ces systèmes qui apparaissent naturellement en algorithmique, apparaissent souvent comme non classiques aux dynamiciens. Il était alors tentant d'utiliser cette expérience pour élargir la possible modélisation dans le contexte de l'algorithmique du texte, et pour généraliser progressivement la définition des sources dynamiques.

Plan. On commence par rappeler, dans la Section 1, les propriétés de base des systèmes dynamiques. Puis, la Section 2 présente les opérateurs qui seront utilisés dans les analyses et qui se situent dans la lignée des opérateurs de transferts des dynamiciens. La Section 3 décrit le cadre d'analyse fonctionnelle nécessaire à l'obtention des propriétés spectrales. Alors, tout est prêt pour décrire l'analyse dynamique, et ce, à travers deux champs d'application : le texte dans la Section 4 et l'arithmétique dans la Section 5.

Ces notes visent à introduire le sujet de l'analyse dynamique des algorithmes, et à donner quelques exemples clés. Elles sont complétées par une bibliographie assez exhaustive. On pourra aussi consulter la page du groupe d'Analyse dynamique à l'adresse

`http://users.info.unicaen.fr/~daireaux/ANADY/index.html` .

Ces notes correspondent à un cours donné par Viviane Baladi et Brigitte Vallée lors des journées annuelles du groupe de travail ALÉA en mars 2002. Les Sections 1 et 3 résument plutôt le cours donné par Viviane, tandis que les Sections 2, 4, 5 sont relatives au cours de Brigitte. Ces notes résument en quelque sorte l'activité du groupe d'Analyse dynamique entre 1995 et ce jour. Brigitte Vallée tient à remercier tous ceux qui ont contribué à ce travail : en tout premier lieu, Philippe Flajolet, mais aussi tous ceux qui font partie ou ont, à un moment ou un autre, fait partie du groupe caennais : (par ordre alphabétique) Ali Akhavi, Jérémie Bourdon, Julien Clément, Benoit Daireaux, Hervé Daudé, Julien Fayolle, Charlie Lemée, Loick Lhote. Un grand merci à Jérémie Bourdon pour le prêt des figures tirées de son mémoire de thèse ..., aux relecteurs attentifs de ce texte et tout particulièrement à l'éditeur de ce volume.

1. Systèmes dynamiques

Ici, on donne la définition des systèmes dynamiques et on insiste sur leurs principales caractéristiques. Le lecteur intéressé à la problématique générale des systèmes dynamiques pourra consulter le livre [4]. Les livres [12, 36] constituent une très bonne introduction élémentaire aux systèmes dynamiques de l'intervalle.

1.1. **Système dynamique.** Un système dynamique (de l'intervalle) est défini par les éléments suivants (voir un exemple Figure 2) :

1. un alphabet $\mathcal{M}$ inclus dans $\mathbb{N}$, fini ou dénombrable.
2. une partition topologique de $I :=]\,0,1\,[$ en intervalles ouverts disjoints I_m, pour $m \in \mathcal{M}$, *i. e.* $\bar{I} = \bigcup_{m\in\mathcal{M}} \bar{I}_m$.
3. une application de codage σ, constante et égale à m sur chaque I_m.
4. une application de décalage $T : I \to I$ inversible et de classe $\mathcal{C}^2$ sur chaque I_m. On désigne par $J_m = TI_m$ l'image par T de l'intervalle I_m, par $h_m : J_m \to I_m$ l'inverse local (appelé encore branche inverse) de T restreint à I_m, et par $\mathcal{H}$ l'ensemble $\mathcal{H} := \{\, h_m \mid m \in \mathcal{M} \,\}$ des branches inverses de T.

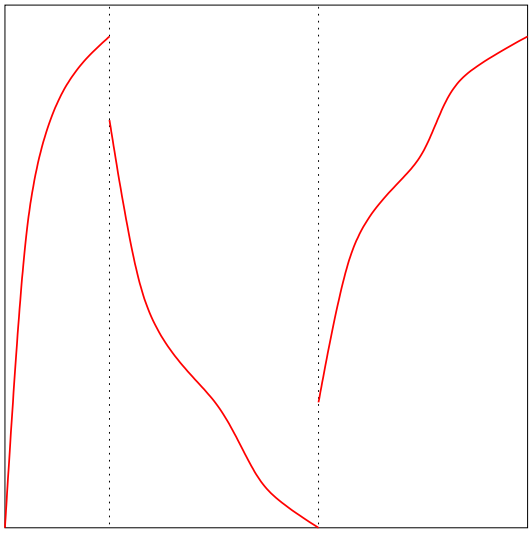

FIG. 2. Exemple de source dynamique avec un alphabet $\mathcal{M}$ de cardinal 3

Il y a plusieurs caractéristiques importantes d'un système dynamique, liées en particulier à la régularité des branches h_m, à leur géométrie (c'est-à-dire à la position des intervalles J_m par rapport aux intervalles I_m), au nombre de branches, fini ou infini, aux propriétés d'expansion du système (le décalage T sera dit expansif s'il existe $\Delta > 1$ pour lequel $\left|T'(x)\right| \geq \Delta > 1$).

La trajectoire (ou l'orbite) d'un élément $x \in I$ est la suite :

$$\mathcal{T}(x) := (x, Tx, \dots, T^k x, \dots).$$

Si on utilise l'application de codage σ, on peut associer au réel x le mot infini $M(x)$ construit sur l'alphabet $\mathcal{M}$,

$$M(x) = \big(\sigma(x), \sigma(Tx), \dots, \sigma(T^k x), \dots\big).$$

On pourra se reporter à la Figure 3 pour un exemple de ces deux notions.

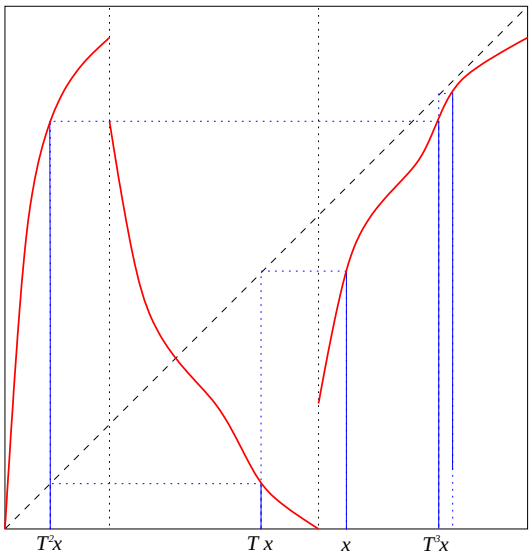

FIG. 3. Une orbite créée par une source dynamique et le mot émis associé *cbac*...

1.2. **Utilisation en algorithmique.** En algorithmique, les systèmes dynamiques interviennent naturellement dans deux types de contextes : les algorithmes arithmétiques et les algorithmes du texte.

Les algorithmes arithmétiques. Un certain nombre d'algorithmes de type « algorithmes d'Euclide » suivent le schéma suivant.

Entrée : $x \in I$
Tant que $x \notin \mathcal{F}$ faire $x := T(x)$
Renvoyer x

Ici, $\mathcal{F}$ désigne l'ensemble des états finaux de l'algorithme. La trace d'une exécution de l'algorithme sur l'entrée x est alors la trajectoire tronquée $\tilde{\mathcal{T}}(x)$ qui s'arrête dès que x entre dans $\mathcal{F}$. Le système associé à la transformation T (qu'on appelle le système sous-jacent à l'algorithme) peut être très varié. Pour cette classe d'algorithmes, le système dynamique de référence est associé à la transformation T défini par

$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

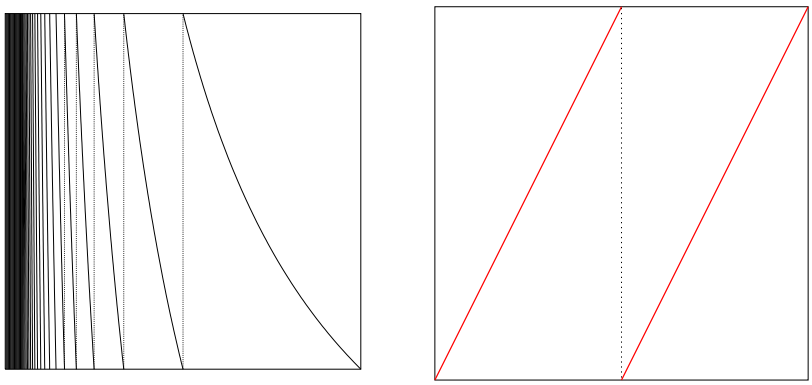

FIG. 4. Les deux systèmes dynamiques de référence

(voir Figure 4 gauche), mais la Section 5 donnera des exemples d'algorithmes « naturels » qui font intervenir des systèmes dynamiques assez complexes.

Les algorithmes du texte. Le système dynamique intervient ici fortement car c'est lui qui produit le texte. Plus précisément, on considère le modèle probabiliste suivant : on se donne une densité sur I et on étudie l'ensemble des mots de $\mathcal{M}^{\mathbf{N}}$ de la forme

$$M(x) = \big(\sigma(x), \sigma(Tx), \dots, \sigma(T^k x), \dots\big)$$

lorsque $x \in I$ est choisi suivant la densité f. Le système dynamique de référence (voir Figure 4 droite) est alors associé à la transformation T définie par

$$T(x) := 2x - \lfloor 2x \rfloor$$

qui donne lieu aux suites de chiffres binaires indépendants et équiprobables. La Section 4 donnera des exemples d'analyse d'agorithme de texte, quand le texte est produit par une source dynamique.

1.3. **Première caractéristique des systèmes dynamiques : la géométrie des branches.** La géométrie du système décrit la position des intervalles $J_m := TI_m$ par rapport aux intervalles I_m. Elle permet de caractériser l'ensemble $\mathcal{S}_m$ successeur du symbole m, formé de tous les symboles qui peuvent être émis après le symbole m. La géométrie du système donne ainsi un premier accès à la corrélation entre les symboles successifs émis.

Système complet. On dira que le système est *complet* si pour tout $m \in \mathcal{M}$, l'intervalle J_m est l'intervalle I tout entier. Tous les symboles de l'alphabet $\mathcal{M}$ sont possiblement émis après tout symbole m et donc $\mathcal{S}_m = \mathcal{M}$ pour tout symbole m. Ces systèmes-là sont (dans un sens à préciser) les moins corrélés.

Système markovien. Pour ces systèmes, l'ensemble $\mathcal{S}_m$ des symboles émis après un symbole m ne dépend que de m, et non de ce qui s'est passé avant. Par définition, et dans le cas d'un alphabet fini, on dit qu'un système est *markovien* si tout intervalle $J_m := TI_m$ est réunion finie d'intervalles I_ℓ. Plus précisément, pour tout $m \in \mathcal{M}$, il existe un sous ensemble $\mathcal{L}_m \subset \mathcal{M}$ tel que

$$J_m = \bigcup_{\ell \in \mathcal{L}_m} I_\ell,$$

et dans ce cas, on a $\mathcal{S}_m = \mathcal{L}_m$. La Figure 5 donne un exemple où

$$J_1 = I_1 \cup I_2, \qquad J_2 = I_2 \cup I_3, \qquad J_3 = I.$$

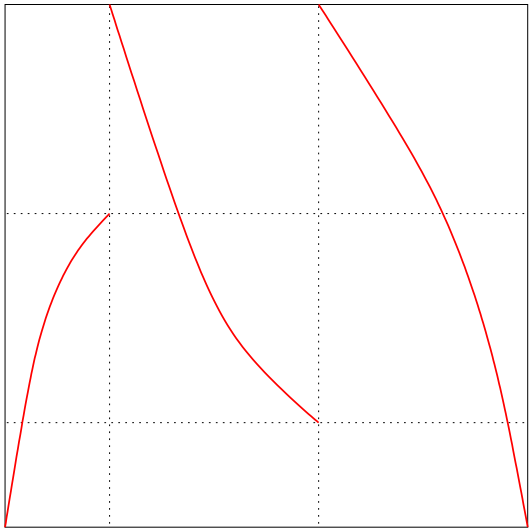

FIG. 5. Une source dynamique markovienne

Dans le cas d'un alphabet infini, il faut être un peu plus précis. On dit qu'un système est markovien s'il existe une partition finie de I en intervalles K_ℓ ($\ell \in \mathcal{L}$ et $\mathcal{L}$ finie) telle que

1. tout intervalle J_m est réunion (nécessairement finie) d'intervalles K_ℓ, pour $\ell \in \mathcal{L}_m$;
2. tout intervalle K_ℓ est réunion (en général non finie) d'intervalles I_m, pour $m \in \mathcal{M}_\ell$.

Un élément ℓ de $\mathcal{L}$ joue un rôle similaire à celui d'un état dans une chaîne de Markov. Pour deux états k et ℓ, on désigne par $\mathcal{M}_{k|\ell}$ l'ensemble des symboles de $\mathcal{M}$ qui permettent de passer de l'état ℓ à l'état k,

$$\mathcal{M}_{k|\ell} := \{\, m \in \mathcal{M} \mid I_m \subset K_\ell \quad \text{et} \quad K_k \subset J_m \,\} = \{\, m \in \mathcal{M} \mid m \in \mathcal{M}_\ell \quad \text{et} \quad k \in \mathcal{L}_m \,\}.$$

La matrice sous-jacente au système dynamique est la matrice booléenne P dont le coefficient $p_{k,\ell}$ est défini par

$$p_{k,\ell} = 1 \qquad \text{si et seulement si} \qquad \mathcal{M}_{k|\ell} \neq \emptyset. \tag{1.1}$$

Elle décrit les transitions possibles entre symboles, et le cas particulier où P est une matrice irréductible est important, puisqu'il traduit une propriété de mélange entre les symboles. (Une matrice irréductible est une matrice dont tous les coefficients sont positifs et qui possède une puissance dont tous les coefficients sont strictement positifs.)

Parfois, la partition de départ $(I_m)_{m\in\mathcal{M}}$ ne donne pas lieu à un système markovien, mais il se peut qu'un raffinement de la partition y donne lieu. La définition plus générale d'un système markovien est finalement la suivante : on construit, à partir de l'ensemble $\mathcal{S}$ des extrémités des intervalles I_m de la partition initiale, les ensembles

$$\mathcal{S}^{[p]} := \bigcup_{i=1}^{p} T^i(\mathcal{S}) \ ; \tag{1.2}$$

le système est markovien si la suite des $\mathcal{S}^{[p]}$ débute par un premier terme $\mathcal{S}^{[1]}$ fini et est stationnaire.

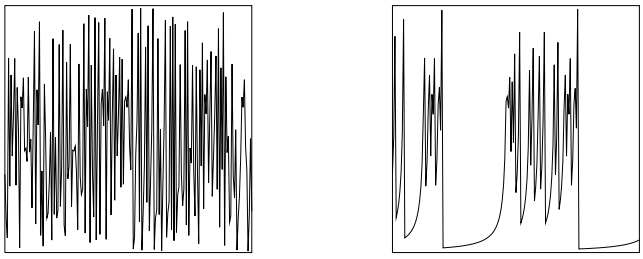

FIG. 6. À gauche, orbite chaotique ; à droite, orbite avec intermittence

Système non markovien. Dans ce cas, les symboles qui peuvent être émis à un moment donné ne peuvent être caractérisés en ne considérant qu'une partie bornée de l'histoire précédente : ce sont les systèmes les plus complexes.

1.4. **Importance du caractère expansif.** Rappelons qu'un système est *expansif* si le nombre $\Delta := \inf \left|T'(x)\right|$ est strictement plus grand que 1. La grandeur $\delta := 1/\Delta$ est le coefficient de contraction des branches inverses, et toute branche inverse h de T vérifie $\left|h'(x)\right| \leq \delta$. À première vue, le caractère expansif du décalage (ou, de manière équivalente, le caractère contractant des branches inverses) n'apparaît pas essentiel. Pour se persuader de l'importance de ce facteur, il suffit de comparer le comportement des orbites de deux systèmes : l'un est associé à un décalage T pour lequel T^2 est expansif ; l'autre est « presque » expansif, puisqu'il existe un point fixe indifférent x_0 (*i. e.* un point x_0 pour lequel $T(x_0) = x_0$, $\left|T'(x_0)\right| = 1$), alors que tous les autres points vérifient $\left|T'(x)\right| > 1$ (voir Figure 6). Dans le premier cas, la trajectoire est chaotique ; dans l'autre, elle présente des phénomènes d'intermittence, et quand la trajectoire s'approche de ce point fixe indifférent, elle s'en éloigne à grand peine ... Ces deux systèmes créeront une algorithmique vraiment différente, le premier donnant lieu à un algorithme rapide, et le second, qui perd beaucoup de temps près de son point fixe, donnant lieu à un algorithme lent. Nous reviendrons à cette situation dans les paragraphes 3.4 et 5.6.

2. Le principal outil de l'analyse dynamique : l'opérateur de transfert et sa descendance

Ici, on définit les principaux opérateurs qui sont les outils privilégiés de l'analyse dynamique. Ils proviennent tous de l'opérateur transformateur de densité, qui est leur ancêtre commun.

2.1. **Opérateur transformateur de densité.** Nous venons de décrire comment la possibilité d'émettre à un instant donné tel ou tel symbole était liée à la géométrie du système. Maintenant, nous nous posons une question plus fine : avec quelle probabilité un symbole — s'il peut être émis — va-t-il être émis ? Cette question est très liée à la manière dont le décalage T déforme les mesures sur l'intervalle I. Plus précisément, la densité de probabilité sur I évolue lorsqu'on itère la transformation de décalage T, et c'est l'opérateur *transformateur de densité*, désigné par $\mathbf{H}$, qui quantifie ce phénomène. Pour une densité initiale f, on désigne par $\mathbf{H}[f]$ la densité après une itération de T. On a ainsi :

$$\mathbf{H}[f](x) = \sum_{m \in \mathcal{M}} \left|h'_m(x)\right| f \circ h_m(x)\, 1_{J_m}(x), \tag{2.1}$$

où 1_A représente la fonction indicatrice de l'ensemble A. Informellement, si f est la densité initiale, la densité en un point x, après une itération, est apportée par tous les antécédents possibles de x.

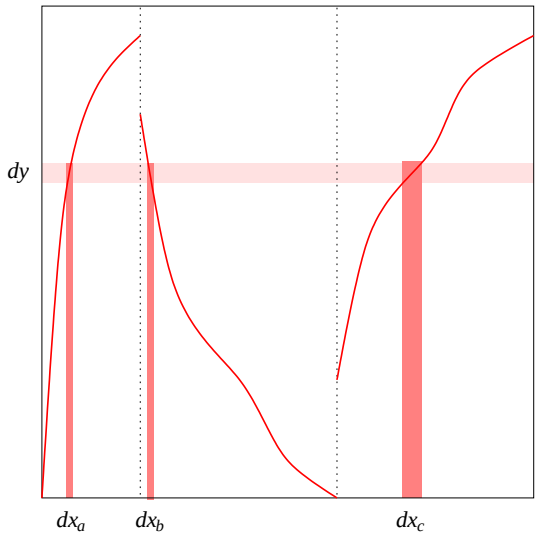

FIG. 7. L'évolution de la densité

L'antécédent de x provenant de la branche d'indice m existe si x appartient à J_m, et dans ce cas, il apporte la densité $f \circ h_m(x)$ distordue par le terme $\left|h'_m(x)\right|$ (lié à la formule de changement de variable). La composante $\mathbf{H}_{[m]}$ de l'opérateur relative au symbole m

$$\mathbf{H}_{[m]}[f](t) := \left|h'_m(t)\right| f \circ h_m(t)\, 1_{J_m}(t)$$

désigne ainsi la contribution apportée par la branche d'indice m (voir Figure 7).

C'est cette distorsion possible par le facteur $\left|h'_m(x)\right|$ qui va constituer le deuxième facteur de corrélation. Si les branches sont affines, avec donc une dérivée constante, cette distorsion n'existera pas. Pour une géométrie de branches fixée, ce sont donc les systèmes dynamiques à branches affines qui seront les moins corrélés. À l'opposé, ceux dont les branches ont une dérivée seconde grande (en valeur absolue) donneront lieu à des sources fortement corrélées. En particulier, c'est plutôt la dérivée de $x \mapsto \log\left|h'(x)\right|$ qui va intervenir, et la *condition de distorsion bornée*,

$$\exists c > 0,\ \forall x \in I,\ \forall h \in \mathcal{H}, \qquad \left|h''(x)\right| \leq c\left|h'(x)\right|, \tag{2.2}$$

toujours vérifiée lorsque le nombre de branches est fini, intervient de manière fréquente.

Le k-ième itéré de l'opérateur $\mathbf{H}$ a aussi une forme très simple; grâce à la propriété de multiplicativité des dérivées de fonctions composées, il s'exprime comme une somme qui fait intervenir tous les mots w de $\mathcal{M}^k$,

$$\mathbf{H}^k[f](x) = \sum_{w \in \mathcal{M}^k} |h'_w(x)|\, f \circ h_w(x)\, 1_{J_w}(x). \tag{2.3}$$

Ici, pour un mot w de $\mathcal{M}^k$ de la forme $w := m_1 m_2 \dots m_k$, la notation h_w désigne la *branche inverse* de T^k de la forme $h_w := h_{m_1} \circ \cdots \circ h_{m_k} \in \mathcal{H}^k$ et J_w désigne l'intervalle de définition de la branche h_w.

Cas particulier des systèmes complets et markoviens. Comme nous le verrons plus loin, la présence des fonctions indicatrices apporte un certain nombre de complications. Le cas le plus simple est donc celui des systèmes complets où ces fonctions indicatrices n'existent pas.

Dans le cas d'un système markovien, quitte à travailler avec une matrice d'opérateurs, on peut faire « disparaître » ces fonctions indicatrices, en procédant comme suit : à une fonction f définie

sur I, on associe la suite (finie) des fonctions f_ℓ, où f_ℓ est la restriction de f à l'intervalle K_ℓ. Au lieu de faire agir l'opérateur $\mathbf{H}$ sur f, et de considérer le transformé $g := \mathbf{H}[f]$, on considère qu'il agit sur la suite $\tilde{f}$ des f_ℓ et on désigne par g_k la k-ième composante de $\tilde{g}$ (*i. e.* la restriction de g à K_k) On a clairement

$$g_k = \sum_{\ell\in\mathcal{L}} \sum_{m\in\mathcal{M}_{k|\ell}} \mathbf{H}_{[m]}[f_\ell],$$

de sorte que $\mathbf{H}$ est maintenant (à conjugaison près) une matrice d'opérateurs, désignée par $\tilde{\mathbf{H}}$, de dimension $|\mathcal{L}| \times |\mathcal{L}|$ dont le coefficient situé en position (k,ℓ) est l'opérateur

$$\tilde{\mathbf{H}}_{k,\ell} := \mathbf{H}_{[k\,|\ell]} = \sum_{m\in\mathcal{M}_{k|\ell}} \mathbf{H}_{[m]}\ ;$$

En remplaçant ainsi l'égalité $g := \mathbf{H}[f]$ par l'égalité $\tilde{g} := \tilde{\mathbf{H}}[\tilde{f}]$, on a supprimé toutes les fonctions indicatrices ...

2.2. **Opérateur transformateur de densité, intervalles fondamentaux et probabilités fondamentales.** Si w est un mot fini, on désigne par p_w la probabilité qu'un mot produit par la source commence par w.

Associons à un mot w de longueur finie k la branche inverse h_w ; l'intervalle $h_w(I)$ est alors l'ensemble des réels x pour lesquels le mot $M(x)$ débute par le préfixe w : c'est ce que nous appelons *l'intervalle fondamental* associé au mot w, et que nous désignons par I_w ; pour un mot réduit à un symbole m, c'est exactement l'intervalle I_m de la partition initiale. Considérons une densité de probabilité f sur l'intervalle I. La mesure de l'intervalle $I_w = h_w(I)$ est exactement la probabilité p_w et

$$p_w := \int_{h_w(I)} f(t)\,dt = \int_I |h'_w(t)|\, f\circ h_w(t)\, 1_{J_w}(t)\,dt.$$

La composante de l'opérateur $\mathbf{H}^k$ relatif à la branche h_w, désignée par $\mathbf{H}_{[w]}$ et définie par

$$\mathbf{H}_{[w]}[f](t) := \left|h'_w(t)\right| f\circ h_w(t)\, 1_{J_w}(t) \tag{2.4}$$

permet donc d'exprimer la probabilité p_w, via la relation

$$p_w = \int_I \mathbf{H}_{[w]}[f](t)\,dt, \tag{2.5}$$

de sorte que cet opérateur $\mathbf{H}_{[w]}$ peut être considéré comme l'opérateur « générateur » de la probabilité p_w. De plus, la concaténation ww' entre deux mots se traduit par la *propriété de composition*

$$\mathbf{H}_{[ww']} = \mathbf{H}_{[w']} \circ \mathbf{H}_{[w]}, \tag{2.6}$$

qui est essentielle car elle permet de généraliser la propriété multiplicative

$$p_{ww'} = p_w p_{w'}$$

qui n'est vérifiée que par les sources sans mémoire.

2.3. **Sources classiques simples : sources sans mémoire, chaînes de Markov.** Pour une géométrie donnée, les systèmes dynamiques les plus simples sont ceux dont les branches sont affines.

Une source sans mémoire est modélisée par un système dynamique complet à branches affines, initialisé avec la densité uniforme. La Figure 8 donne un exemple de modélisation possible d'une source sans mémoire qui produit trois symboles suivant les probabilités $1/2, 1/6, 1/3$.

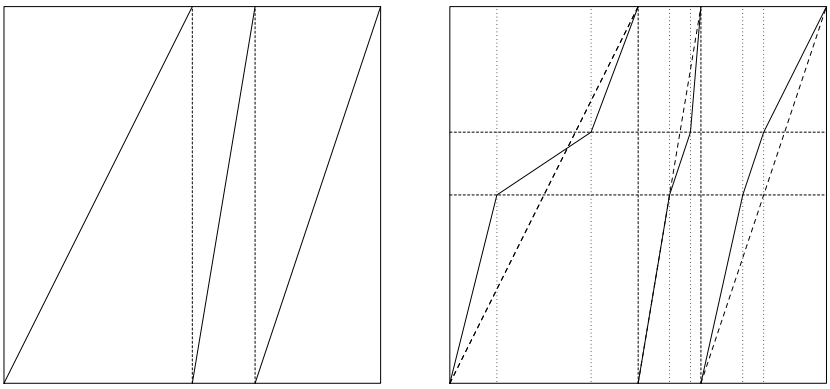

FIG. 8. Une source sans mémoire, une chaîne de Markov

Une chaîne de Markov est modélisée par un système dynamique markovien à branches affines, initialisé avec une densité constante sur chaque K_ℓ. La Figure 8 montre un exemple de modélisation d'une chaîne de Markov d'ordre 1.

Une chaîne de Markov d'ordre k s'obtient en cassant en morceaux les branches (affines) d'une chaîne de Markov d'ordre $k-1$. La Figure 8 montre comment on peut passer du cas $k=0$ au cas $k=1$. C'est pour cela, que, informellement du moins, un système général markovien peut être considéré comme une limite de chaînes de Markov d'ordre de plus en plus élevé.

2.4. **L'opérateur de transfert.** Dans l'étude des systèmes dynamiques, il est très utile de généraliser l'opérateur transformateur de densité $\mathbf{H}$ (défini en (2.1) en lui adjoignant un paramètre s. On obtient alors *l'opérateur de transfert*, désigné par $\mathbf{H}_s$ et défini par

$$\mathbf{H}_s[f](x) = \sum_{m\in\mathcal{M}} \left|h'_m(x)\right|^s f\circ h_m(x)\, 1_{J_m}(x). \tag{2.7}$$

Ici, l'ajout du paramètre s permettra de relier cet opérateur à des séries génératrices et plus précisément à des séries génératrices de Dirichlet.

Comme en (2.3), le k-ième itéré de l'opérateur $\mathbf{H}_s$ a aussi une forme très simple, et s'exprime comme une somme qui fait intervenir tous les mots w de $\mathcal{M}^k$,

$$\mathbf{H}_s^k[f](x) = \sum_{w\in\mathcal{M}^k} \left|h'_w(x)\right|^s f\circ h_w(x)\, 1_{J_w}(x). \tag{2.8}$$

Nous aurons besoin des composantes de tels opérateurs, et nous désignerons par $\mathbf{H}_{s,[w]}$ l'opérateur associé à la branche h_w et défini par

$$\mathbf{H}_{s,[w]}[f](t) := \left|h'_w(t)\right|^s f\circ h_w(t)\, 1_{J_w}(t).$$

Remarquons cependant que cet opérateur, qui vérifie une propriété de composition analogue à (2.6),

$$\mathbf{H}_{s,[ww']} = \mathbf{H}_{s,[w']}\circ\mathbf{H}_{s,[w]} \tag{2.9}$$

ne permet pas d'exprimer simplement la quantité p_w^s.

L'opérateur qui fait intervenir l'ensemble $\mathcal{M}^\star$ de tous les mots (finis) produits par la source est alors la somme de tous les itérés k-ième de l'opérateur définis en (2.8) : c'est ce que nous appelons le quasi-inverse ou l'étoile,

$$(\mathbf{1}-\mathbf{H}_s)^{-1} := \sum_{k\geq 0} \mathbf{H}_s^k, \tag{2.10}$$

et qui jouera un rôle si important dans la suite ...

2.5. **Pondération de l'opérateur de transfert.** Dans les applications aux algorithmes (et tout particulièrement aux algorithmes arithmétiques), on désire souvent pondérer chaque branche du décalage par une quantité qui mesure le coût de l'algorithme associé quand l'exécution « passe par » la branche. Ce coût peut dépendre de manière assez variée de la branche, mais, très souvent, comme nous le verrons dans les applications, ce coût est « additif », et le coût total d'une exécution est la somme des coûts dûs à l'emprunt de chaque branche. On remplace alors chaque opérateur composant $\mathbf{H}_{s,[w]}$ par un opérateur pondéré par un coût c,

$$\mathbf{H}^{[c]}_{s,u,[w]} := u^{c(h_w)}\, \mathbf{H}_{s,[w]},$$

et l'additivité des coûts montre que la propriété de composition se prolonge aux opérateurs pondérés.

2.6. **Opérateur de transfert généralisé.** Il est nécessaire ici de considérer des sources dynamiques complètes ou markoviennes. Commençons par le cas complet. Les quantités p^s_w s'expriment alors en fonction de l'opérateur de transfert généralisé, appelé encore opérateur sécant. Si F désigne la fonction de répartition de f, la quantité p^s_w s'exprime comme

$$p^s_w = \big|F\circ h_w(0) - F\circ h_w(1)\big|^s,$$

et fait donc intervenir la valeur de la fonction $F\circ h_w$ en les deux points $x=0$ et $x=1$. C'est pourquoi on introduit un opérateur de transfert $\mathfrak{H}_{s,[w]}$ qui agit sur des fonctions de deux variables en utilisant la « sécante » de la branche h_w (d'où son nom d'opérateur sécant)

$$\mathfrak{H}_{s,[w]}[\Phi](u,v) := \left|\frac{h_w(u)-h_w(v)}{u-v}\right|^s \Phi\big(h_w(u), h_w(v)\big),$$

ce qui résout le problème puisque

$$p^s_w = \mathfrak{H}_{s,[w]}[L^s](0,1) \qquad \text{avec} \qquad L(x,y) = \left|\frac{F(x)-F(y)}{x-y}\right|. \tag{2.11}$$

La multiplicativité de la « sécante » permet de prouver la propriété de composition

$$\mathfrak{H}_{s,[ww']} = \mathfrak{H}_{s,[w']}\circ \mathfrak{H}_{s,[w]},$$

qui, comme en (2.6) généralise la relation $p^s_{ww'} = p^s_w p^s_{w'}$.

Les opérateurs qui généralisent respectivement $\mathbf{H}_s$, ses itérés $\mathbf{H}^k_s$ et son quasi-inverse $(\mathbf{1}-\mathbf{H}_s)^{-1}$ sont alors les opérateurs $\mathfrak{H}_s$, $\mathfrak{H}^k_s$ et $(\mathbf{1}-\mathfrak{H}_s)^{-1}$ définis par

$$\mathfrak{H}_s := \sum_{m\in\mathcal{M}} \mathfrak{H}_{s,[m]}, \qquad \mathfrak{H}^k_s = \sum_{w\in\mathcal{M}^k} \mathfrak{H}_{s,[w]}, \qquad (\mathbf{1}-\mathfrak{H}_s)^{-1} = \sum_{w\in\mathcal{M}^\star} \mathfrak{H}_{s,[w]}. \tag{2.12}$$

Ce formalisme peut se transporter aisément dans le cas d'une source markovienne : la matrice $\mathfrak{H}_s$ a pour coefficient l'opérateur $\mathfrak{H}_{s,[k|\ell]}$.

2.7. **Problèmes à longueur fixée, ou à longueur quelconque.** Comme le montrent les relations (2.8), (2.10) et (2.12), les k-ième itérés des opérateurs font intervenir l'ensemble $\mathcal{M}^k$ des mots de longueur k et les quasi-inverses l'ensemble $\mathcal{M}^\star$ de tous les mots finis. Si on travaille sur des problèmes à taille fixée (longueur des textes fixée pour les algorithmes de texte, nombre d'itérations fixé pour les algorithmes arithmétiques), c'est donc le comportement asymptotique de ces k-ième itérés qu'on utilisera (pour $k\to\infty$). Si le problème fait intervenir toutes les tailles possibles, les opérateurs adéquats seront les opérateurs quasi-inverses, et on s'intéressera à leurs singularités.

Pour une matrice M, le comportement asymptotique de M^k ou les singularités de $(\mathrm{Id}-M)^{-1}$ sont très liés aux propriétés spectrales de la matrice M, et en particulier aux propriétés spectrales

dominantes (correspondant aux valeurs propres ayant le plus grand module). Nous sommes donc conduits à étudier l'analogue, mais en dimension infinie.

3. Analyse fonctionnelle et propriétes spectrales

Cette section est dédiée à l'étude des propriétés spectrales des opérateurs de transfert. Un livre de référence est celui de V. Baladi [2].

Pour un opérateur $\mathbf{L}$ qui agit sur un espace de Banach $\mathcal{F}$, le spectre $\operatorname{Sp}\mathbf{L}$ de $\mathbf{L}$ est l'ensemble des nombres complexes z pour lesquels $\mathbf{L} - z\mathbf{1} : \mathcal{F} \to \mathcal{F}$ n'est pas inversible. Un élément z de $\operatorname{Sp}\mathbf{L}$ est une valeur propre si $\mathbf{L} - z\mathbf{1}$ n'est pas injective. En dimension finie, le spectre d'une matrice est l'ensemble de ses valeurs propres. L'espace sur lequel agit l'opérateur est fondamental car le spectre d'un opérateur dépend beaucoup de l'espace sur lequel il opère. (Plus l'espace est « gros », plus il contient de possibles fonctions propres, et plus le spectre est lui-même « gros ».) Ainsi, un opérateur peut avoir de « bonnes » propriétés spectrales sur un espace et de moins bonnes sur un autre. Le choix de cet espace est fondamental et constitue généralement un des points délicats de l'analyse.

3.1. **Critères de choix pour l'espace fonctionnel.** Ce choix résulte en général d'un compromis : On veut que l'espace fonctionnel $\mathcal{F}$ soit suffisamment « gros » pour que l'opérateur de transfert $\mathbf{H}_s$ opère sur $\mathcal{F}$ (*i. e.* $\mathbf{H}_s[\mathcal{F}] \subset \mathcal{F}$). Mais on veut aussi qu'il ne soit pas trop gros pour que le spectre reste discret (formé de points isolés), ou du moins que la partie « supérieure » du spectre reste discrète.

Ce choix va dépendre des caractéristiques du système dynamique. Il sera dicté en tout premier lieu par la géométrie du système, et modulé par la régularité des branches. Dans la formule (2.7), apparaissent les fonctions caractéristiques 1_{J_m}. En fonction de la géométrie du système, ces fonctions caractéristiques peuvent introduire des discontinuités, et $\mathbf{H}_s[f]$ peut être discontinue même si f est très régulière.

1. Si le système est complet, les opérateurs $\mathbf{H}_s$ n'introduisent pas de discontinuités et on peut travailler sur des espaces de fonctions régulières (fonctions C^r sur I, fonctions analytiques, etc.) adaptés à la régularité des branches h_m.
2. Si le système est markovien, les opérateurs $\mathbf{H}_s^p$ introduisent des discontinuités uniquement au bord des K_ℓ et on peut travailler sur des espaces de fonctions régulières sur chacun des K_ℓ, ayant donc un nombre fini de discontinuités.
3. Enfin, si le système n'est pas markovien, on introduit à chaque itération de nouvelles discontinuités, de sorte que l'ensemble des discontinuités introduites est dénombrable et peut être dense dans I. On est alors conduit à travailler sur l'espace des fonctions à variation bornée.

3.2. **Le bon comportement désiré.** On considère d'abord le cas où $s = 1$. L'opérateur étudié est donc le transformateur de densité $\mathbf{H}$.

Sur un espace fonctionnel adéquat $\mathcal{F}$, les propriétés

(P1) la valeur 1 est valeur propre simple dominante unique de $\mathbf{H}$,

(P2) il y a un saut spectral : le reste du spectre de $\mathbf{H}$ est contenu dans un disque de rayon strictement inférieur à 1,

entrainent un certain nombre de conséquences. Tout d'abord, il existe alors un disque Γ du plan complexe, de frontière γ, qui contient comme seul point du spectre la valeur 1. De plus, l'opérateur

$\mathbf{P}$ défini par

$$\mathbf{P} := \frac{1}{2i\pi} \int_\gamma (z\mathbf{1} - \mathbf{H})^{-1}\, dz$$

est le projecteur sur le sous-espace propre dominant, et l'opérateur $\mathbf{H}$ se décompose en $\mathbf{H} = \mathbf{P} + \mathbf{N}$ où $\mathbf{N}$ est un opérateur dont le spectre est le même que celui de $\mathbf{H}$, excepté la valeur 1. Le rayon spectral de $\mathbf{N}$ est ainsi strictement inférieur à 1. Enfin, on a aussi $\mathbf{H}^k = \mathbf{P} + \mathbf{N}^k$ de sorte que

$$(\mathbf{1} - z\mathbf{H})^{-1} = \frac{\mathbf{P}}{1-z} + \mathbf{R}(z), \tag{3.1}$$

avec une fonction reste $\mathbf{R}$,

$$\mathbf{R}(z) := (\mathbf{1} - z\mathbf{N})^{-1} - \mathbf{P} = \sum_{k\geq 1} z^k (\mathbf{H}^k - \mathbf{P})$$

qui décrit les corrélations du système dynamique. De plus, le projecteur $\mathbf{P}$ s'exprime en fonction de la fonction propre dominante ϕ, normalisée par $\int_I \phi(u)\, du = 1$, sous la forme :

$$\mathbf{P}[f](t) = \phi(t) \int_I f(u)\, du.$$

Si, de plus, la condition (P3) suivante est satisfaite,

(P3) l'application $s \mapsto \mathbf{H}_s$ est analytique sur un voisinage de $s = 1$,

la théorie de la perturbation s'applique alors [34] et montre l'existence de fonctions $s \mapsto \lambda(s)$, $s \mapsto \mathbf{P}_s$, $s \mapsto \mathbf{N}_s$ analytiques dans un voisinage de $s = 1$. Ici, $\lambda(s)$ est la valeur propre dominante de $\mathbf{H}_s$, $\mathbf{P}_s$ est le projecteur sur le sous-espace propre dominant et $\mathbf{N}_s$ est un opérateur dont le rayon spectral est strictement inférieur à $\big|\lambda(s)\big|$. La décomposition $\mathbf{H}_s^k = \lambda(s)^k \mathbf{P}_s + \mathbf{N}_s^k$ perdure et finalement, la décomposition spectrale

$$(\mathbf{1} - \mathbf{H}_s)^{-1} = \frac{\mathbf{P}_s}{1-\lambda(s)} + \mathbf{N}_s(\mathbf{1} - \mathbf{N}_s)^{-1} \tag{3.2}$$

montre que $(\mathbf{1} - \mathbf{H}_s)^{-1}$ possède un pôle d'ordre 1 en $s = 1$, dont le résidu est $-\lambda'(1)\mathbf{P}$. Cette dernière valeur $-\lambda'(1)$ est l'entropie du système dynamique, comme nous le verrons plus loin.

3.3. **Compacité et quasi-compacité.** La propriété (P1) est une propriété de type Perron–Frobenius : elle est liée à des propriétes de forte positivité. Rappelons que la propriété (P1) est vérifiée pour une matrice M stochastique qui a une puissance k-ième dont tous les coefficients sont strictement positifs.

La propriété (P2) est toujours vraie en dimension finie, car le spectre est alors fini. Plus généralement, la validité de (P2) est assurée aussitôt que le spectre de $\mathbf{H}$ est discret, ou, du moins, aussitôt que la partie « supérieure » du spectre est discret.

Les opérateurs *compacts* sont les opérateurs qui, en dimension infinie, ressemblent le plus aux opérateurs de la dimension finie. Leur spectre est discret à ceci près qu'un point d'accumulation est possible en 0, et la validité de (P2) est alors assurée. Mais, on ne peut pas toujours trouver un espace fonctionnel $\mathcal{F}$ sur lequel l'opérateur $\mathbf{H}$ soit compact, et l'on ne peut donc toujours assurer que la totalité du spectre soit discret. On considère alors la propriété de quasi-compacité, plus générale. Le rayon spectral $R(\mathbf{L})$ d'un opérateur $\mathbf{L}$ est la borne supérieure des modules des éléments du spectre $\operatorname{Sp}\mathbf{L}$, de sorte que $\operatorname{Sp}\mathbf{L} \subset \{\, \lambda \;\big|\; |\lambda| \leq R(\mathbf{L}) \,\}$. Le rayon spectral essentiel $R_e(\mathbf{L})$ d'un opérateur $\mathbf{L}$ est le plus petit réel $r > 0$ pour lequel tout élément λ de $\operatorname{Sp}\mathbf{L}$ ayant un module $|\lambda| > r$ est une valeur propre isolée et de multiplicité finie. Pour un opérateur compact, on a $R_e(\mathbf{L}) = 0$. Un opérateur

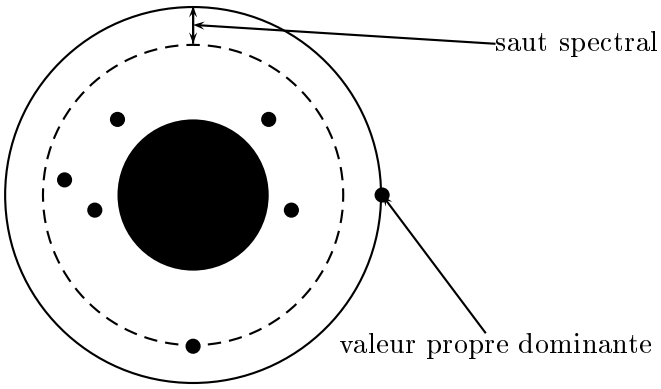

FIG. 9. Saut spectral

pour lequel $R_e(\mathbf{L}) < R(\mathbf{L})$ est appelé *quasi-compact*. Son spectre se décompose en deux parties, une partie supérieure discrète et une partie inférieure qui peut être quelconque (voir Figure 9).

3.4. **Des espaces fonctionnels adéquats.** L'espace fonctionnel où les propriétés (P1), (P2) et (P3) sont vérifiées dépend des caractéristiques du système. Nous donnons ici quelques exemples d'espaces fonctionnels adaptés à certaines classes de systèmes dynamiques.

Type 1 : Systèmes complets (ou markoviens) avec branches uniformément holomorphes et contractantes. Ce sont d'abord les systèmes complets, bien décrits dans [38], qui vérifient ce qui suit :

> Il existe un disque complexe $\mathcal{V}$ sur lequel toutes les branches inverses $h \in \mathcal{H}$ se prolongent en des fonctions holomorphes sur $\mathcal{V}$, envoyant $\mathcal{V}$ strictement dans lui-même, (*i. e.* $h\left(\bar{\mathcal{V}}\right) \subset \mathcal{V}$) et contractantes (*i. e.* $\left|h'(z)\right| \leq \delta_h < 1$ avec la série $\sum_h \delta_h^\alpha$ convergente pour un réel $\alpha < 1$).

Dans ce cas, l'opérateur $\mathbf{H}$ agit sur l'espace $\mathcal{A}_\infty(\mathcal{V})$ des fonctions holomorphes définies sur $\mathcal{V}$ et continues sur $\bar{\mathcal{V}}$. Comme tous les opérateurs composants (qui sont des opérateurs de « composition » de la forme $f \mapsto f \circ h$) y sont compacts, l'opérateur $\mathbf{H}$ y est aussi compact. Un théorème dû à Krasnoselsky [35] généralise les résultats à la Perron–Frobenius et prouve que (P1) est aussi vérifiée ; (P3) est également vérifiée sans problème, par perturbation analytique, dès que $\Re(s) > \alpha$, ce pour un certain $\alpha > 1$.

Si de plus, le système a une distorsion bornée, les propriétés citées ci-dessus se généralisent à l'opérateur $\mathfrak{H}_s$ (voir [14, 45]), à condition de le faire opérer sur l'espace $\mathcal{B}_\infty(\mathcal{V})$ des fonctions holomorphes définies sur $\mathcal{V} \times \mathcal{V}$ et continues sur $\bar{\mathcal{V}} \times \bar{\mathcal{V}}$.

On peut aussi considérer la version « markovienne » du début de la condition précédente (on reprend les notations des paragraphes 1.3 et 2.1) :

Pour tout k et tout ℓ de $\mathcal{L}$, il existe un disque complexe $\mathcal{V}_k$, voisinage de K_k sur lequel toutes les branches inverses $h \in \mathcal{H}_{[k|\ell]}$ ont leurs restrictions à K_k qui se prolongent en des fonctions holomorphes sur $\mathcal{V}_k$, envoyant $\mathcal{V}_k$ strictement dans $\mathcal{V}_\ell$.

Cette dernière condition assure que chaque opérateur $\mathbf{H}_{[k|\ell]}$ a de bonnes propriétés de compacité et de positivité. Si, de plus, la matrice de transition P définie en (1.1) est irréductible et apériodique, alors l'opérateur matriciel a toutes les bonnes propriétés souhaitées.

Type 2 : Systèmes à géométrie quelconque, contractants. Cas du nombre de branches fini. Dans ce cas (voir [17]), l'espace fonctionnel adapté est l'espace $BV(I)$ des fonctions à variation bornée sur l'intervalle I. Cet espace est un espace de Banach dense dans $\mathcal{L}^1(I)$ dont la boule unité est précompacte dans $\mathcal{L}^1(I)$. L'opérateur $\mathbf{H}$ agit sur $BV(I)$ et le théorème suivant [32] permet de montrer sa quasi-compacité.

Théorème. *Soit* $\mathbf{L}$ *un opérateur qui agit sur* $\mathcal{L}^1$. *Supposons qu'il existe deux suites* (r_n) *et* (t_n) *de nombres positifs pour lesquelles, pour tout* $n \geq 1$, *et pour tout* $f \in BV(I)$, *on a*

$$\left\|\mathbf{L}^n[f]\right\|_{BV} \leq r_n \|f\|_{BV} + t_n \|f\|_1. \tag{3.3}$$

Alors l'opérateur $\mathbf{L}$ *est borné sur* $BV(I)$ *et son rayon spectral essentiel vérifie*

$$R_e(\mathbf{L}) \leq r := \liminf_{n\to\infty} (r_n)^{1/n}.$$

On applique le théorème en montrant que r peut être choisi égal au coefficient de contraction $\delta < 1$ et que l'opérateur $\mathbf{H}$ a une valeur propre égale à 1.

Type 3 : Cas du nombre infini de branches. Systèmes à géométrie pseudo-markovienne, contractants, à distorsion bornée. Quand le nombre de branches est infini, ce qui arrive très souvent dans les applications arithmétiques, on peut aussi travailler sur $BV(I)$, à condition d'exiger des propriétés supplémentaires pour le système dynamique. En particulier (voir [8, 13]), on exige que le système ait une distorsion bornée, et aussi qu'il ne soit pas trop différent d'un système markovien. Dans le cas d'un système markovien, l'ensemble $\mathcal{S}^{[p]}$, défini en (1.2), et formé des extrémités des intervalles J_w associés à l'ensemble $\{\, w \mid |w| \leq p \,\}$ est fini pour tout p. Là, on lui laisse la possibilité d'être infini, mais on exige que les intervalles J_w, quand ils sont non vides, ne soient pas trop petits, *i. e.*

$$\ell_p := \inf \left\{\, |J_w| \;\middle|\; J_w \neq \emptyset,\ |w| \leq p \,\right\} > 0.$$

C'est une condition qui a été donnée au départ par Rychlick. Dans ces conditions, les propriétés (P1), (P2) et (P3) sont vérifiées pour l'opérateur $\mathbf{H}_s$ agissant sur $BV(I)$.

3.5. **La méthode d'induction.** Dans tout le paragraphe précédent, le système était supposé expansif. On peut traiter relativement aisément des systèmes complets où la condition d'expansion est seulement violée en un point, et qui sont « presque expansifs » avec seulement un point indifférent (voir paragraphe 1.4). Dans ce cas, il y a une seule « mauvaise » branche (*i. e.* non expansive), et on va la grouper avec des bonnes branches, pour tenter d'améliorer son comportement. Supposons que cette branche soit la branche correspondant au symbole a, et corresponde donc à un intervalle I_a.

Considérons le système dynamique (J, U) où l'intervalle J est $J := I \setminus I_a$ et le décalage U est défini par le premier retour à J : pour $x \in J$, on désigne par $n(x)$ le plus petit entier pour lequel $T^{n(x)} \in J$, et on pose $U(x) := T^{n(x)}(x)$. Ce système dynamique est appelé le système induit. La partition fondamentale sur J est maintenant formée des intervalles fondamentaux de l'ancien système de la forme

$$I_w \quad \text{avec} \quad w \in \mathcal{N} := \big(\mathcal{M} \setminus \{a\}\big)\{a\}^\star,$$

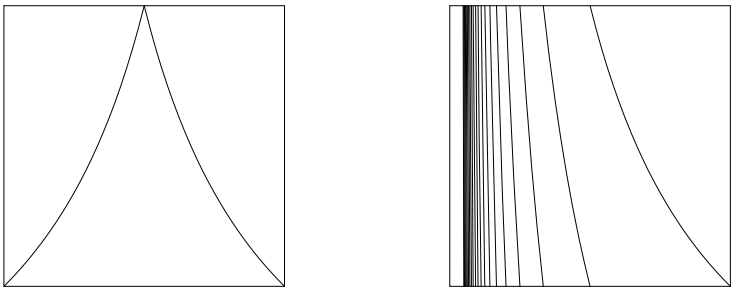

FIG. 10. Un système dynamique et son système induit

et le nouvel alphabet $\mathcal{N}$ est ainsi infini.

Il y a une autre manière d'induire, un peu différente, en restant dans l'intervalle I, et en remplaçant la partition initiale par la partition formée des anciens intervalles fondamentaux de la forme

$$I_w \quad \text{avec} \quad w \in \mathcal{Q} := \{a\}^\star\big(\mathcal{M} \setminus \{a\}\big).$$

C'est celle-là qu'on utilisera plutôt en algorithmique, et qui remplace l'alphabet $\mathcal{M}$ initial par l'alphabet $\mathcal{Q}$. La Figure 10 représente un système dynamique (à gauche) et son système dynamique induit associé (on induit ici par rapport à la première branche, car c'est elle qui possède un point indifférent).

Grâce aux propriétés de dictionnaire dues à la propriété de composition (2.9), l'opérateur de transfert $\tilde{\mathbf{H}}_s$ du système dynamique induit fait intervenir l'opérateur de transfert $\mathbf{H}_s$ et l'opérateur $\mathbf{A}_s := \mathbf{H}_{s,[a]}$ relatif au symbole a sous la forme

$$\tilde{\mathbf{H}}_s = \sum_{k\geq 0} (\mathbf{H}_s - \mathbf{A}_s)\mathbf{A}_s^k = (\mathbf{H}_s - \mathbf{A}_s)(\mathbf{1} - \mathbf{A}_s)^{-1}. \tag{3.4}$$

Puisque le nouveau décalage regroupe une suite de « mauvaises » branches avec une « bonne » branche, le nouveau système dynamique sera expansif, et le quasi-inverse $(\mathbf{1} - \tilde{\mathbf{H}}_s)^{-1}$ vérifiera souvent des propriétés de type (3.2). Alors, la relation $\mathcal{M}^\star = \mathcal{Q}^\star\{a\}^\star$, qui se traduit par une relation entre les deux quasi-inverses,

$$(\mathbf{1} - \mathbf{H}_s)^{-1} = (\mathbf{1} - \mathbf{A}_s)^{-1}(\mathbf{1} - \tilde{\mathbf{H}}_s)^{-1} \tag{3.5}$$

permet de « revenir » au quasi-inverse initial, en y intégrant les propriétés de la « mauvaise » branche.

4. **Analyse dynamique des algorithmes du texte**

Le comportement de tout algorithme qui travaille sur du texte est très influencé par la manière dont le texte est produit. Il y a d'abord un premier fait qui est vrai pour une source $\mathcal{S}$ quelconque :

1. L'ensemble des probabilités $\{\, p_w \mid w \in \mathcal{M}^\star \,\}$, ou plus généralement, pour un complexe s, l'ensemble des quantités $\{\, p_w^s \mid w \in \mathcal{M}^\star \,\}$ joue un rôle essentiel dans l'analyse des algorithmes du texte, lorsque le texte est produit par une source quelconque $\mathcal{S}$.

L'intérêt des sources dynamiques provient du caractère explicite de ces probabilités, que nous avons décrit dans la Section 2 :

2. Pour une source dynamique, les probabilités p_w s'expriment en fonction des composantes de l'opérateur transformateur de densité (voir Section 2.2).
3. Pour une source dynamique complète (ou markovienne), les quantités p_w^s s'expriment en fonction de l'opérateur de transfert généralisé (voir l'opérateur sécant de la Section 2.6).

Nous allons maintenant décrire quelques exemples d'application de ces trois faits.

4.1. **Les problèmes de mots qui font intervenir des langages.** Un langage $\mathcal{L}$ défini sur l'alphabet $\mathcal{M}$ est un sous-ensemble de $\mathcal{M}^\star$. À un langage $\mathcal{L}$, on associe classiquement la série génératrice

$$L(z) := \sum_{w\in\mathcal{L}} p_w z^{|w|} \tag{4.1}$$

où la variable z « marque » la taille $|w|$ du mot w. Cette série génératrice s'avère essentielle dans l'analyse des proriétés du langage $\mathcal{L}$.

Pour une source sans mémoire, la propriété de multiplicativité des probabilités permet de traduire les opérations sur les langages en opérations sur les séries génératrices associées. Ce n'est plus possible dès que la source garde « de la mémoire ». On remplace alors, dans la série génératrice du langage définie en (4.1), la probabilité p_w par l'opérateur générateur $\mathbf{H}_{[w]}$ défini en (2.4), et on obtient ce qu'on appelle l'opérateur générateur du langage $\mathcal{L}$ défini par

$$\mathbf{L}(z) := \sum_{w\in\mathcal{L}} \mathbf{H}_{[w]} z^{|w|}. \tag{4.2}$$

La propriété de composition (2.6) sur les opérateurs permet de traduire les opérations sur les langages en opérations sur les opérateurs générateurs associés. Grâce à (2.5), on peut alors revenir à la série génératrice par la relation

$$L(z) = \int_I \mathbf{L}(z)[f](t)\,dt. \tag{4.3}$$

Exemple d'application : les motifs généralisés. (Le cadre est celui des sources de type 2 ou 3 de la Section 3.4). On pourra se reporter à [10] pour plus de précisions.

Un motif généralisé $\mathcal{L}$ est une suite finie de langages construits sur le même alphabet $\mathcal{M}$, de la forme $\mathcal{L} := (\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_r)$. Chacun des langages $\mathcal{L}_i$ est de longueur finie. On dit que le motif $\mathcal{L}$ apparaît dans le texte $T \in \mathcal{M}^\star$ si le texte contient comme sous-séquence un élément $\ell = (\ell_1, \ell_2, \dots, \ell_r)$ de $\mathcal{L}$. Dans ce cas, T est de la forme

$$T = w_0\ell_1 w_1 \ell_2 \dots w_i \ell_i w_{i+1} \dots w_r \ell_r w_{r+1} \qquad \text{avec} \qquad w_i \in \mathcal{M}^\star \quad \text{et} \quad \ell_i \in \mathcal{L}_i.$$

Cette notion de motif généralisé recouvre beaucoup de problèmes de recherche de motifs, tout particulièrement les motifs cachés, qui apparaissent naturellement dans des contextes divers (bio-informatique, détection d'intrusions) et a déjà été étudiée dans le contexte des sources sans mémoire [26].

L'ensemble de toutes les *occurrences* du motif généralisé $\mathcal{L}$ est alors la collection $\rho(\mathcal{L})$ (avec répétitions) donnée par concaténation,

$$\rho(\mathcal{L}) = \mathcal{M}^\star \times \mathcal{L}_1 \times \mathcal{M}^\star \times \mathcal{L}_2 \times \cdots \times \mathcal{M}^\star \times \mathcal{L}_r \times \mathcal{M}^\star. \tag{4.4}$$

Cette opération ρ transforme une suite finie de langages en une collection de mots (par opposition à un langage qui est un ensemble de mots, une collection est un multi-ensemble de mots), et dans la collection $\rho(\mathcal{L})$, un texte T est présent autant de fois qu'il contient d'occurrences de $\mathcal{L}$. Pour un texte T de longueur n, on désigne par $\Omega_n(\mathcal{L}, T)$ le nombre d'occurrences de $\mathcal{L}$ dans T, et la remarque

précédente permet de montrer que la série génératrice des espérances coïncide exactement avec la série génératrice $L(z)$ de la collection $\rho(\mathcal{L})$,

$$L(z) := \sum_{w\in\rho(\mathcal{L})} p_w z^{|w|} = \sum_{n\geq 1} \mathbf{E}\big[\Omega_n(\mathcal{L},T)\big]\, z^n.$$

Grâce aux règles de transfert citées précédemment, l'opérateur générateur $\mathbf{L}(z)$ de la collection $\rho(\mathcal{L})$ s'écrit facilement en fonction des opérateurs générateurs $\mathbf{L}_i(z)$ des langages et de l'opérateur $(\mathbf{1}-z\mathbf{H})^{-1}$ associé au langage $\mathcal{M}^\star$,

$$\mathbf{L}(z) = (I-z\mathbf{H})^{-1}\circ \mathbf{L}_r(z)\circ (I-z\mathbf{H})^{-1}\circ\cdots\circ \mathbf{L}_1(z)\circ (I-z\mathbf{H})^{-1}. \tag{4.5}$$

Cet opérateur contient $r+1$ occurrences du quasi-inverse $(I-z\mathbf{H})^{-1}$, qui « apportent » chacune un pôle en $z=1$. Elles sont « mélangées » avec les opérateurs $\mathbf{L}_i(z)$ des langages $\mathcal{L}_i$ qui sont des polynômes en z (et n'apportent pas de pôles). Via la relation (4.3), on caractérise alors aisément les singularités de la série $L(z)$ et on obtient ainsi le résultat suivant :

Proposition. *Le nombre moyen $\mathbf{E}\big[\Omega_n(\mathcal{L},T)\big]$ d'occurrences du motif généralisé $\mathcal{L}$ dans un texte de longueur n produit par une source dynamique de type 2 ou 3 vérifie :*

$$\mathbf{E}\big[\Omega_n(\mathcal{L},T)\big] = \binom{n+r}{r}\pi(\mathcal{L}) + \binom{n+r-1}{r-1}\pi(\mathcal{L})\big(C(\mathcal{L})-N(\mathcal{L})\big) + O(n^{r-2}).$$

Ici, $\pi(\mathcal{L})$ est le poids total du motif

$$\pi(\mathcal{L}) := \prod_{i=1}^{r} p(\mathcal{L}_i)$$

où, pour une collection $\mathcal{M}$, on pose

$$p(\mathcal{M}) := \sum_{w\in\mathcal{M}} p_w,$$

et $N(\mathcal{L})$ est sa longueur moyenne. Le coefficient $C(\mathcal{L})$ décrit la corrélation entre deux composantes successives du motif et s'exprime en fonction de l'opérateur $\mathbf{R}$ défini en (3.1).

À l'aide des mêmes techniques, utilisées cette fois pour des collections associées aux doubles occurrences, on peut avoir accès à la variance du nombre d'occurrences. On démontre ainsi un phénomène de concentration autour de la valeur moyenne [10].

4.2. **Les grandeurs fondamentales d'une source (cas d'une source de type 1).** Pour plus de précisions, on peut consulter [45]. Les séries de Dirichlet des probabilités fondamentales font intervenir les quantités p_w^s et sont définies par

$$\Lambda_k(s) := \sum_{|w|=k} p_w^s, \qquad \Lambda(s) := \sum_{k\geq 0}\Lambda_k(s) = \sum_{w\in\mathcal{M}^\star} p_w^s. \tag{4.6}$$

La plupart des grandeurs fondamentales associées à la source $\mathcal{S}$ s'expriment à l'aide de ces séries. Nous en donnons quatre exemples.

Entropie. L'entropie $h(\mathcal{S})$ de la source satisfait à la relation

$$h(\mathcal{S}) := \lim_{k\to\infty}\frac{-1}{k}\sum_{|w|=k} p_w \log p_w = \lim_{k\to\infty}\frac{-1}{k}\left(\frac{d}{ds}\Lambda_k(s)\right)\Big|_{s=1}.$$

Probabilité de coïncidence. La coïncidence $C(x,y)$ entre les deux mots $M(x)$ et $M(y)$ pour deux réels x et y tirés indépendamment selon une même loi est la longueur du plus long prefixe commun. La probabilité pour que $M(x)$ et $M(y)$ aient le même préfixe de longueur k est donc la probabilité de l'événement $\big[C(x,y] \geq k\big]$. Cet événement se produit si (et seulement si) les deux réels x et y appartiennent à un même intervalle fondamental I_w de profondeur k (voir Section 2.2). On a ainsi

$$\mathbf{P}\big[C(x,y) \geq k\big] = \sum_{|w|=k} p_w^2,$$

et la probabilité de coïncidence $c(\mathcal{S})$ vérifie la relation

$$c(\mathcal{S}) := \lim_{k\to\infty} \left(\sum_{|w|=k} p_w^2 \right)^{1/k} = \lim_{k\to\infty} \Lambda_k(2)^{1/k}.$$

Équirépartition des mots de longueur k. On cherche à décrire les probabilités possibles de tous les mots de longueur k. Plus précisément, on veut décrire la distribution de l'ensemble

$$\mathcal{P}_k := \{\, p_w \mid w \in \mathcal{M}^k \,\}.$$

On définit sur $\mathcal{M}^k$ une variable aléatoire ℓ_k par $\ell_k(w) := \log p_w$, et on veut analyser la distribution de la variable ℓ_k. Un outil important pour l'analyse d'une variable aléatoire X est la série génératrice des moments,

$$M(X)(s) := \mathbf{E}\big[\exp(sX)\big] = \sum_{n\geq 0} \frac{s^n}{n!}\, \mathbf{E}[X^n].$$

Ici, la série génératrice des moments de la variable ℓ_k, désignée par $M_k(s)$, vérifie

$$M_k(s) := \mathbf{E}[p_w^s] = \sum_{w\in\mathcal{M}^k} p_w p_w^s = \Lambda_k(1+s). \tag{4.7}$$

Nombre de préfixes assez probables. La quantité $B(\rho)$ désigne le nombre de préfixes w dont la probabilité est au moins égale à ρ ($\rho \to 0$). Un outil principal est ici une transformation intégrale, la transformée de Mellin (voir [25]), que nous utiliserons aussi en Section 4.3. La transformée de Mellin de la fonction B est reliée à la fonction $\Lambda(s)$, via la relation

$$\Lambda(s) = s \int_0^\infty B(x) x^{s-1}\, dx.$$

Dans les quatre exemples, les grandeurs caractéristiques $h(\mathcal{S})$, $c(\mathcal{S})$ et les fonctions B et $M_k(s)$ s'expriment donc en fonction des séries de Dirichlet $\Lambda_k(s)$ et $\Lambda(s)$ définies en (4.6).

Transcription algébrique. Dans le cas des sources dynamiques complètes (ou markoviennes), et grâce à la relation (2.11), les séries de Dirichlet (4.6) ont une autre expression en fonction de l'opérateur de transfert sécant,

$$\Lambda_k(s) = \mathfrak{H}_s^k[L^s](0,1) \quad \text{et} \quad \Lambda(s) = (\mathbf{1} - \mathfrak{H}_s)^{-1}[L^s](0,1) \tag{4.8}$$

où L est aussi définie en (2.11).

Traitement analytique. Dans le cas des sources dynamiques de type 1, les bonnes propriétés spectrales de l'opérateur de transfert sécant induisent un bon comportement des séries de Dirichlet, et tous les résultats vont s'exprimer en fonction de la valeur propre dominante $s \mapsto \lambda(s)$, omniprésente dans ce cadre. Remarquons que $s \mapsto \lambda(s)$ ne dépend que du système dynamique et non pas de la densité (analytique) initiale f choisie. Au voisinage de l'axe réel, la série $\Lambda_k(s)$ se comporte comme une quasi-puissance : il existe $a(s)$ tel que, sur un voisinage complexe d'un point s_0 réel, on ait

$$\Lambda_k(s) \sim a(s)\lambda(s)^k$$

pour $k \to \infty$ uniformément en s sur ce voisinage. Par ailleurs, $\Lambda(s)$ est analytique sur le demi-plan $\Re(s) > 1$, avec un pôle simple en $s = 1$, et pour s au voisinage de 1 sur ce domaine,

$$\Lambda(s) \sim \frac{a(s)}{1-\lambda(s)} \sim \frac{-1}{\lambda'(1)}\frac{1}{s-1}.$$

Dans le cas où la fonction $s \mapsto \lambda(s)$ est périodique, il peut y avoir d'autres pôles régulièrement espacés sur la droite $\Re(s) = 1$. Ce phénomène de périodicité se produit en particulier pour certaines classes de sources simples, mais ce sont essentiellement les seuls cas où il se produit.

On déduit d'abord aisément les deux relations

$$h(\mathcal{S}) = -\lambda'(1), \qquad c(\mathcal{S}) = \lambda(2). \tag{4.9}$$

Des techniques classiques d'analyse (transformée de Mellin, théorème taubérien) permettent d'obtenir le comportement de la fonction B au voisinage de 0. Par exemple, si la fonction $s \mapsto \lambda(s)$ n'est pas périodique, on obtient

$$B(\rho) \sim \frac{-1}{\lambda'(1)\rho} \quad \text{pour} \quad \rho \to 0.$$

Enfin, la série génératrice des moments (4.7) se comporte presque exactement comme la fonction $a(s)\lambda(1+s)^k$, ce comportement étant uniforme en s sur un voisinage de 0. Alors des résultats classiques, dûs en particulier à Hwang [33], montrent que la variable aléatoire ℓ_k suit asymptotiquement (quand $k \to \infty$) une loi gaussienne, avec

$$\mathbf{E}[\ell_k] \sim \lambda'(1)k, \qquad \mathbf{Var}[\ell_k] \sim \big(\lambda''(1) - \lambda'(1)^2\big)k, \tag{4.10}$$

la convergence vers la loi normale étant en $O\big(1/\sqrt{k}\,\big)$. (Là encore, il y a quelques exceptions, essentiellement liées à des sources simples.) Ce résultat est une version forte d'un théorème célèbre en Théorie de l'Information, dû à Shannon–Macmillan–Breiman qui montre que pour de « bonnes sources », l'ensemble $\mathcal{M}^k$ des mots de longueur k se répartit en deux sous-ensembles : les mots probables, qui ont à peu près tous la même probabilité, égale à $\exp\big(-kh(\mathcal{S})\big)$ et un ensemble de mots très peu probables. Le résultat obtenu ici démontre en plus un phénomène de concentration autour de la valeur moyenne.

4.3. **Comportement des arbres dictionnaires (cas des sources de type 1).** Pour plus de précisions, on peut consulter [7, 9, 15, 16, 23].

Une structure de données essentielle dans les algorithmes de traitement du texte est l'arbre digital, ou trie (le mot « trie » est obtenu par contraction des deux mots « tree » et « retrieval »), et ses variations (le patricia-trie et le suffix-trie). Un trie est tout simplement un arbre qui implante un dictionnaire : un dessin suffit à comprendre comment il fonctionne (voir Figure 11). Les nœuds internes servent à diriger la recherche, et ce sont les feuilles qui contiennent les mots du dictionnaire. Il y a en particulier (et par définition) autant de feuilles que de mots dans le trie. Un nœud du trie (interne ou feuille) peut être étiqueté par le chemin qui le lie à la racine. Pour obtenir le patricia-trie

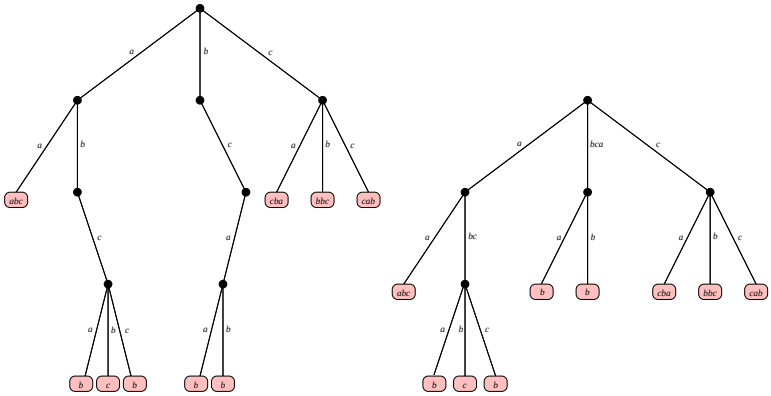

FIG. 11. Un exemple de trie et du patricia-trie associé

associé, on supprime simplement les noeuds internes qui ne sont pas des points de branchement (voir Figure 11).

Les atouts du trie sont sa facilité d'implantation et son dynamisme : il est facile à modifier (insertion, suppression, etc.). L'efficacité de la structure de données associée est liée à la compacité de la forme de l'arbre, qu'on peut quantifier par les paramètres usuels d'un arbre : longueur de cheminement externe, nombre de nœuds (internes) — encore appelé taille —, hauteur, ... Ici, la mesure de la donnée est le nombre n de mots présents dans le dictionnaire.

L'analyse de la structure de trie et des arbres de sa descendance a été largement étudiée dans le cadre des sources classiques. On peut consulter à ce sujet le livre de W. Szpankowski [42]. Ici, nous cherchons à faire l'analyse dans le cadre « dynamique ». Il y a une très grande affinité entre les propriétés du trie et celles de la source dynamique. Un trie construit sur un ensemble de n mots est défini par l'ensemble $X := \{x_1, x_2, \ldots, x_n\}$ des n réels qui ont donné naissance aux n mots. On le désigne par la suite par $T(X)$. Un tel trie est complètement déterminé par les nœuds internes qui sont effectivement présents. Or ces nœuds-là sont étiquetés par les préfixes w pour lesquels l'intervalle fondamental I_w contient au moins deux éléments de X. Pour que les contributions de l'ensemble X dans deux intervalles fondamentaux disjoints I_w et $I_{w'}$ soient indépendants, on est alors conduit à travailler dans un modèle de Poisson : on tire la cardinalité N de l'ensemble X suivant une loi de Poisson de paramètre z,

$$\mathbf{P}[N = n] = e^{-z}\,\frac{z^n}{n!},$$

puis on tire les n réels de l'ensemble X indépendamment suivant une loi de densité f. Alors la variable aléatoire N_w qui mesure la cardinalité de l'ensemble $I_w \cap X$ suit une loi de Poisson de paramètre $p_w z$: et, crac, voilà la probabilité fondamentale p_w qui intervient de nouveau ! La probabilité d'existence du nœud interne n_w d'étiquette w est égale à $\mathbf{P}[N_w \geq 2]$, tandis que la contribution de ce noeud n_w à la longueur moyenne de cheminement externe est $\mathbf{E}[\, N_w \mid N_w \geq 2\,]$.

On obtient ainsi l'expression des valeurs moyennes des deux variables taille, S, et longueur de cheminement externe, P. L'indice z fait référence au paramètre du modèle de Poisson.

$$\mathbf{E}[P_z] = \sum_{w \in \mathcal{M}^\star} p_w z(1 - e^{-p_w z}), \qquad \mathbf{E}[S_z] = \sum_{w \in \mathcal{M}^\star} \left(1 - e^{-p_w z}(1 + p_w z)\right).$$

Ces deux expressions sont des sommes harmoniques, et l'instrument pour étudier le comportement asymptotique de telles expressions (pour $z \to \infty$) est la transformée de Mellin [25]

$$\hat{A}(s) := \int_0^\infty A(x) x^{s-1}\, dx$$

car la transformée d'une somme harmonique $A(z)$ se factorise en un produit de deux facteurs : si

$$A(z) = \sum_{w \in \mathcal{M}^\star} g(p_w z),$$

alors

$$\hat{A}(s) = \hat{g}(s) \sum_{w \in \mathcal{M}^\star} p_w^{-s}.$$

En particulier, les transformées de Mellin des espérances de la taille et de la longueur de cheminement externe font intervenir la série de Dirichlet $\Lambda(s)$ définie en (4.6), et, tout particulièrement son comportement singulier autour de son pôle dominant $s = 1$ (qui s'exprime à l'aide de l'entropie $-\lambda'(1)$).

Par ailleurs, la hauteur de $T(X)$ est au plus égale à k pourvu qu'il n'existe pas de noeuds internes n_w associés à des préfixes de longueur k. Compte tenu du phénomène d'indépendance induit par le modèle de Poisson, on a donc :

$$\mathbf{P}[H_z \le k] = \prod_{w \in \mathcal{M}^k} \mathbf{P}[N_w \le 1] = \prod_{w \in \mathcal{M}^k} e^{-p_w z}(1 + p_w z)$$

de sorte que

$$\log \mathbf{P}[H_z \le k] = -z + \sum_{w \in \mathcal{M}^k} \log(1 + p_w z). \tag{4.11}$$

Supposons dans un premier temps que l'on puisse utiliser dans (4.11), pour tous les couples (w, z) et successivement, les deux approximations suivantes

$$-p_w z + \log(1 + p_w z) \sim -\frac{p_w^2 z^2}{2} \qquad \text{puis} \qquad \sum_{w \in \mathcal{M}^k} p_w^2 = \Lambda_k(2) \sim a\lambda(2)^k.$$

Alors, la valeur moyenne de la hauteur s'écrit

$$\mathbf{E}[H_z] \sim \sum_{k \ge 0} \left(1 - \exp\left(-\frac{az^2}{2}\lambda^k(2)\right)\right)$$

et c'est encore une somme harmonique ! La série de Dirichlet associée,

$$\sum_{k \ge 0} \lambda(2)^{-ks} = \frac{1}{1 - \lambda(2)^{-s}}$$

a un pôle simple en $s = 0$, avec un résidu qui fait intervenir $\big|\log \lambda(2)\big|$.

On peut d'abord rendre rigoureux tout ce qui est dit précédemment. Ensuite, il faut revenir au modèle dit de Bernoulli où le nombre des mots est fixé égal à n. Dans ce cas les paramètres étudiés se notent $S^{[n]}, P^{[n]}, H^{[n]}$. On obtient finalement :

Théorème. *Dans une source dynamique de type 1, les trois paramètres de forme du trie (taille, longueur de cheminement, hauteur) construit sur n mots de la source tirés indépendamment ont pour*

valeur moyenne asymptotique (pour $n \to \infty$) les quantités suivantes qui font intervenir l'entropie et la probabilité de coïncidence,

$$\mathbf{E}\left[S^{[n]}\right] \sim \frac{n}{h(\mathcal{S})}, \qquad \mathbf{E}\left[P^{[n]}\right] \sim \frac{n \log n}{h(\mathcal{S})}, \qquad \mathbf{E}\left[H^{[n]}\right] \sim \frac{\log n}{2\left|\log c(\mathcal{S})\right|}.$$

Dans le cas de sources périodiques, le terme principal de $\mathbf{E}\left[S^{[n]}\right]$ fait intervenir un facteur supplémentaire, qui contient une fonction de n oscillante, avec de faibles amplitudes.

Cette approche est suffisamment robuste pour s'adapter à l'analyse des tries plus compliqués (patricia-tries, tries hybrides) ou pour étudier d'autres paramètres de tries simples (par exemple, la hauteur de pile, qui fait intervenir une source induite au sens du paragraphe 3.4). Le suffix-trie est d'un abord plus complexe : c'est par définition un trie construit sur l'ensemble des suffixes d'un mot, et la propriété d'indépendance entre les mots du dictionnaire n'est plus préservée.

5. **Analyse dynamique des algorithmes arithmétiques**

L'objet de cette section est d'illustrer sur un exemple simple l'utilisation des opérateurs de transfert pour l'analyse d'algorithmes arithmétiques. Nous commençons par traiter le cas de l'algorithme d'Euclide standard du calcul du p. g. c. d. de deux entiers. Le résultat que nous exposons ici n'est pas original, puisque le nombre moyen d'itérations de l'algorithme d'Euclide classique a été déterminé autour de 1970 indépendamment par Heilbronn [30] et Dixon [22]. La méthode décrite est, elle, typique de l'analyse dynamique et peut être facilement généralisée dans de multiples directions (voir paragraphes 5.5 et 5.6).

À partir d'une entrée (v_1, v_0) formée de deux entiers positifs vérifiant $v_1 \leq v_0$ l'algorithme effectue une suite de divisions euclidiennes,

$$v_0 = a_1 v_1 + v_2, \qquad v_1 = a_2 v_2 + v_3, \qquad \dots \qquad v_{k-1} = a_k v_k + 0. \tag{5.1}$$

L'algorithme s'arrête dès qu'apparaît un reste nul. Le coût étudié ici est le nombre k de divisions successives effectuées.

5.1. **Le système dynamique sous-jacent à l'algorithme.** Une étape de l'algorithme remplace une paire (v_1, v_0) par la paire (v_2, v_1) avec

$$\frac{v_2}{v_1} = \frac{v_0}{v_1} - a_1.$$

Si, à la place des paires d'entiers (v_1, v_0), on considère les rationnels de la forme v_1/v_0, la transformation T définie par

$$T(x) = \left\{\frac{1}{x}\right\} := \frac{1}{x} - \left\lfloor\frac{1}{x}\right\rfloor$$

où $\lfloor x \rfloor$ désigne la partie entière de x, exprime (v_2/v_1) en fonction de (v_1/v_0). Le système sous-jacent (voir Figure 12) est complet et l'ensemble des branches inverses de la transformation T est

$$\mathcal{H} = \left\{ h : z \mapsto \frac{1}{z+m} \;\middle|\; m \in \mathbb{N},\ m \neq 0 \right\}.$$

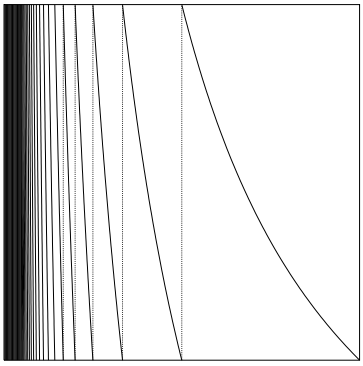

FIG. 12. Le système dynamique euclidien standard

5.2. **Les séries génératrices des coûts.** L'ensemble des entrées possibles de l'algorithme est

$$\tilde{\Omega} = \big\{ (u,v) \;\big|\; 0 \le u \le v \big\},$$

et l'ensemble des entrées de taille N est

$$\tilde{\Omega}_N := \big\{ (u,v) \;\big|\; 0 \le u \le v \le N \big\}.$$

Pour simplifier l'étude, nous travaillons sur des ensembles d'entrées possibles plus restreints,

$$\Omega = \big\{ (u,v) \in \tilde{\Omega} \;\big|\; \operatorname{pgcd}(u,v) = 1 \big\}, \qquad \Omega_N := \big\{ (u,v) \in \tilde{\Omega}_N \;\big|\; \operatorname{pgcd}(u,v) = 1 \big\},$$

formés des entrées pour lesquelles la réponse de l'algorithme est connue à l'avance ..., mais nous reviendrons ensuite aux ensembles $\tilde{\Omega}$, $\tilde{\Omega}_N$ plus « naturels ».

Désignons par $C(u,v)$ la fonction de coût correspondant au nombre de divisions effectuées par l'algorithme d'Euclide sur l'entrée $(u,v) \in \Omega$. L'analyse en moyenne du coût C est l'étude du comportement asymptotique de la valeur moyenne du coût C sur Ω_N ou sur $\tilde{\Omega}_N$

$$\mathrm{E}_N[C] = \frac{\sum_{(u,v)\in\Omega_N} C(u,v)}{\sum_{(u,v)\in\Omega_N} 1}, \qquad \tilde{\mathrm{E}}_N[C] = \frac{\sum_{(u,v)\in\tilde{\Omega}_N} C(u,v)}{\sum_{(u,v)\in\tilde{\Omega}_N} 1},$$

lorsque N tend vers ∞. Les séries génératrices de Dirichlet

$$G_1(s) := \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}} = \sum_{v\ge 1} \frac{a_v}{v^{2s}}, \qquad G_C(s) := \sum_{(u,v)\in\Omega} \frac{C(u,v)}{v^{2s}} = \sum_{v\ge 1} \frac{c_v}{v^{2s}}$$

et leurs homologues « tildées »

$$\tilde{G}_1(s) := \sum_{(u,v)\in\tilde{\Omega}} \frac{1}{v^{2s}} = \sum_{v\ge 1} \frac{\tilde{a}_v}{v^{2s}}, \qquad \tilde{G}_C(s) := \sum_{(u,v)\in\tilde{\Omega}} \frac{C(u,v)}{v^{2s}} = \sum_{v\ge 1} \frac{\tilde{c}_v}{v^{2s}}$$

sont relatives aux coûts intervenant au numerateur et au dénominateur. Les relations

$$\tilde{G}_C(s) = \zeta(s) G_C(s), \qquad \tilde{G}_1(s) = \zeta(s) G_1(s)$$

(où $\zeta(s)$ est la série zeta de Riemann) montrent qu'il suffit de travailler sur Ω, comme il était annoncé. Ici, a_v désigne le nombre d'éléments de Ω ayant un dénominateur égal à v et c_v désigne

la somme des coûts associés aux éléments de Ω ayant un dénominateur égal à v. Remarquons que $E_N[C]$ s'exprime en fonction des sommes partielles des coefficients des séries précédentes :

$$E_N[C] = \frac{\sum_{v\le N} c_v}{\sum_{v\le N} a_v}.$$

Le comportement asymptotique des sommes partielles est lié au comportement des fonctions G_1 et G_C via le théorème taubérien suivant [21, 43].

Théorème taubérien. *Soit une série de Dirichlet $F(s)$ à coefficients positifs ou nuls*

$$F(s) = \sum_{n\ge 1} \frac{a_n}{n^{2s}}$$

telle que :

1. *$F(s)$ converge dans un demi-plan $\Re(s) > \sigma > 0$ et est analytique sur $\Re(s) = \sigma$, $s \neq \sigma$,*
2. *il existe $\gamma \ge 0$ tel que $F(s) = A(s)(s-\sigma)^{-\gamma-1} + C(s)$ où A et C sont analytiques en σ et $A(\sigma) \neq 0$.*

Alors, lorsque $N \longrightarrow \infty$,

$$\sum_{n\le N} a_n = \frac{2^\gamma A(\sigma)}{\sigma\Gamma(\gamma+1)} N^{2\sigma} \log^\gamma N \big(1 + \epsilon(N)\big), \quad \textit{avec} \quad \epsilon(N) \longrightarrow 0.$$

Pour appliquer le théorème précédent, il faut exprimer les deux séries de Dirichlet en fonction de l'opérateur de transfert associé au système dynamique. Ce seront alors les propriétés spectrales de cet opérateur qui permettront d'étudier le comportement de G_1 et G_C au voisinage de $\sigma = 1$.

5.3. **Lien entre la fonction de coût et les opérateurs de transfert.** Les couples (u, v) de Ω sur lesquels l'algorithme effectue exactement k divisions sont ceux qui s'écrivent

$$\frac{u}{v} = h(0) \qquad \text{avec} \qquad h = h_1 \circ \cdots \circ h_k \in \mathcal{H}^k.$$

Puisque toutes les branches inverses $h \in \mathcal{H}^\star$ sont des homographies de déterminant 1, la dérivée $h'(z)$ s'exprime simplement en fonction du carré de son dénominateur : pour $(u, v) \in \Omega$ tel que $u/v = h(0)$ avec $h \in \mathcal{H}^\star$, on a $1/v^2 = \big|h'(0)\big|$.

Ceci permet d'exprimer différemment les séries de Dirichlet

$$G_1(s) = \sum_k \sum_{h\in\mathcal{H}^k} \big|h'(0)\big|^s, \qquad G_C(s) = \sum_k k \sum_{h\in\mathcal{H}^k} \big|h'(0)\big|^s. \tag{5.2}$$

Et c'est maintenant que l'opérateur de transfert $\mathbf{H}_s$ associé au système dynamique intervient. Ici, il est appelé opérateur de Ruelle–Mayer et prend la forme bien connue suivante

$$\mathbf{H}_s[f](x) = \sum_{m\ge 1} \left(\frac{1}{m+x}\right)^{2s} f\left(\frac{1}{m+x}\right).$$

La comparaison des relations (2.8) et (5.2) montre que

$$G_1(s) = \sum_{k\ge 0} \mathbf{H}_s^k[\mathbf{1}](0) = (\mathbf{1} - \mathbf{H}_s)^{-1}[\mathbf{1}](0),$$

$$G_C(s) = \sum_{k\ge 1} k\mathbf{H}_s^k[\mathbf{1}](0) = \mathbf{H}_s(\mathbf{1} - \mathbf{H}_s)^{-2}[\mathbf{1}](0).$$

Les séries de Dirichlet des coûts s'expriment donc à l'aide du quasi-inverse de l'opérateur de transfert.

5.4. **Analyse spectrale.** Comme le système dynamique associé est de type 1, l'espace fonctionnel adéquat est l'espace $\mathcal{A}_\infty(\mathcal{V})$ et l'opérateur $\mathbf{H}_s$ est compact et vérifie les propriétés (P1), (P2) et (P3) de la Section 3.2. La quantité $(\mathbf{1} - \mathbf{H}_s)^{-1}[\mathbf{1}](0)$ possède un pôle d'ordre 1 en $s = 1$ dont le résidu est $-1/\lambda'(1)$. Comme on l'a vu en (4.9), la valeur $-\lambda'(1)$ est l'entropie du système dynamique T. Ici, cette entropie fait intervenir des constantes classiques et vaut

$$h = \frac{\pi^2}{6 \log 2} \approx 2.3731.$$

Le théorème taubérien s'applique en $\sigma = 1$, avec $\gamma = 0$ pour G_1 et $\gamma = 1$ pour G_C. Il permet d'obtenir le comportement asymptotique de $\mathrm{E}_N[C]$ et $\tilde{\mathrm{E}}_N[C]$,

$$\mathrm{E}_N[C] \sim \tilde{\mathrm{E}}_N[C] \sim \frac{-2}{\lambda'(1)} \log N = \frac{12 \log 2}{\pi^2} \log N.$$

Cet exemple montre, dans un cas simple, la démarche suivie lors de l'analyse dynamique d'algorithmes arithmétiques, tout à fait conforme au schéma décrit en Figure 1. De fait, l'analyse dynamique a permis d'obtenir de nombreux autres résultats sur les algorithmes euclidiens (autres coûts, autres algorithmes).

5.5. **D'autres coûts.** On peut d'abord s'intéresser à d'autres paramètres, plus précis. L'un d'entre eux est la complexité en bits, désignée par B, qui compte le nombre d'opérations binaires effectuées par l'algorithme. Des méthodes similaires à celles décrites ici (mais plus subtiles ...) permettent d'évaluer la complexité moyenne en bits $E_N[B]$

$$E_N[B] \sim \frac{6 \log^2 2}{\pi^2} \left(2 + \log_2 \prod_{k=0}^{\infty} \left(1 + \frac{1}{2^k} \right) \right) \log_2^2 N.$$

On utilise (voir [1, 49]) à la fois des opérateurs pondérés (voir 2.5) et les dérivés des opérateurs par rapport à la variable s.

On peut aussi se poser beaucoup d'autres questions sur les rationnels, analogues à celles qu'on se pose classiquement sur le développement en fraction continue des nombres réels. Par exemple : quelle est la fréquence d'un chiffre donné dans le développement en fraction continue d'un rationnel ? Pour les réels, on répond à cette question à l'aide des théorèmes ergodiques. Ici, on remplace l'utilisation des théorèmes ergodiques par les théorèmes taubériens, et on peut montrer que vis-à-vis d'une classe très large de paramètres, les rationnels se comportent « en moyenne » comme les réels le font presque sûrement [49].

5.6. **La classe des algorithmes euclidiens.** Il existe toute une classe d'algorithmes d'Euclide, car il y a autant d'algorithmes d'Euclide que de divisions possibles : on peut effectuer des divisions caractérisées par la classe des quotients (quelconques, pairs, impairs), par la position du reste (division par défaut, par excès, centrée, ou plus généralement α-division), par la parité du reste (on peut vouloir un reste impair, qu'on obtient en enlevant les puissances de 2 du reste classique, ce qui se justifie tout particulièrement quand on veut calculer le symbole de Jacobi à l'aide de la loi de réciprocité quadratique). On peut aussi éviter les divisions et les remplacer par des opérations plus simples (soustractions et décalages binaires) : c'est le cas de l'algorithme binaire, de l'algorithme Plus-Moins et des algorithmes binaires généralisés. On peut aussi éviter les divisions entre grands entiers, et les remplacer par des divisions entre des entiers plus petits : c'est le principe de l'algorithme de Lehmer–Euclide.

À ce jour, les méthodes d'analyse dynamique ont permis d'établir un cadre très général où l'on a pu analyser (presque) tous les algorithmes cités. La démarche décrite dans les paragraphes 5.2

et 5.3 se généralise aisément, car, bien que les systèmes dynamiques « euclidiens » puissent être extrêmement divers, ils ont un point commun important : toutes leurs branches sont des homographies. Comme la dérivée d'une homographie s'exprime en fonction du carré de son dénominateur (avec une intervention possible du déterminant qui n'est plus toujours égal à 1), on peut relier les séries de Dirichlet des coûts et les opérateurs de transfert.

Mais la géométrie des branches et les propriétés d'expansion peuvent vraiment varier d'un algorithme à l'autre, et cette classe dite euclidienne regroupe (presque) toute la diversité possible des systèmes dynamiques. En particulier, les algorithmes pseudo-euclidiens (*i. e.* ceux où l'on « enlève » du reste les éventuelles puissances de 2) obligent à travailler avec des systèmes dynamiques probabilistes, où l'on prolonge la valuation dyadique, bien définie sur les rationnels, en une variable aléatoire sur les réels. Les « bons » espaces fonctionnels ne sont pas alors toujours faciles à trouver, et ils peuvent être autres que ceux qui sont décrits en 3.4. L'analyse fonctionnelle devient alors assez délicate, et moins standard. En particulier, dans [46], en utilisant un espace fonctionnel bien adapté à l'algorithme binaire, qui est alors un espace de Hardy, on a pu analyser cet algorithme et répondre ainsi à une conjecture de Brent [5]. L'espace fonctionnel adapté à l'algorithme Plus-Moins [18], lui, reste encore à trouver !

Même pour les systèmes liés à des algorithmes euclidiens plus simples, la présence d'un point indifférent complique aussi l'analyse : il faut alors travailler avec le système dynamique induit (voir paragraphe 3.4), en utilisant des idées de Prellberg et Slawny [39]. C'est le cas des systèmes liés aux algorithmes Par Excès, Pair ou Soustractif. On obtient alors des algorithmes lents avec un nombre d'itérations quadratique (en $\log^2 N$) [48]. Par exemple, la Figure 13 represente six systèmes dynamiques euclidiens ; selon les colonnes, on obtient deux comportements bien différents ; la première colonne (qui contient les systèmes Standard, Impair et Centré) donne lieu à des algorithmes rapides ; la seconde colonne contient les systèmes Par Excès, Pair, et Soustractif qui ont chacun un point indifférent ; elle donne lieu à des algorithmes lents (comme annoncé en 1.4).

La famille des algorithmes japonais est liée à une α-division de la forme $a = bq + r$ avec un reste $r \in]\, b(\alpha - 1), b\alpha\,]$. Elle est représentée Figure 14. Le carré total est le carré $[\,-1, 1\,] \times [\,-1, 1\,]$. Pour obtenir la représentation du système japonais lié au paramètre $\alpha \in [\,0, 1\,]$, on se limite à la fenêtre délimitée par le carré $[\,\alpha - 1, \alpha\,] \times [\,\alpha - 1, \alpha\,]$. Les systèmes dynamiques japonais sont le plus souvent non complets, en général non markoviens (sauf pour des cas très particuliers du paramètre α) et sont associés à des systèmes de type 3 [8].

On peut aussi chercher à analyser des extensions des algorithmes euclidiens en dimension supérieure : l'algorithme de Gauss qui réduit les réseaux en dimension 2 [20], l'algorithme qui calcule le signe d'un déterminant en utilisant deux développements « parallèles » en fraction continue [47]. De manière un peu inattendue, l'analyse de ces deux algorithmes se révèle proche et fait intervenir la grandeur $\lambda(2)$. L'analyse dynamique de l'algorithme LLL, si utilisé en théorie algorithmique des nombres et en cryptographie, reste aussi un problème ouvert à ce jour ...

5.7. **Les constantes euclidiennes.** Un certain nombre de constantes qui apparaissent dans l'analyse des algorithmes d'Euclide sont liés à des objets spectraux des opérateurs de transfert, et s'expriment en fonction de la valeur propre dominante $s \mapsto \lambda(s)$. Il s'agit tout particulièrement de l'entropie $-\lambda'(1)$, omniprésente, de la valeur $\lambda''(1)$ qui intervient dans les moments d'ordre 2, et de la valeur $\lambda(2)$ qui intervient dans la coïncidence (voir 4.2). Dans l'algorithme d'Euclide standard, la fonction propre dominante est explicite, et donc, l'entropie l'est aussi. Mais, même dans ce cas-là, les deux autres valeurs ne sont pas explicites. Il s'agit de préciser le statut de la calculabilité de ces constantes, pour les algorithmes d'Euclide généraux (où même l'entropie n'est plus explicite), et de les calculer, si leur statut le permet. Il s'agit aussi de calculer de manière exacte la dimension de

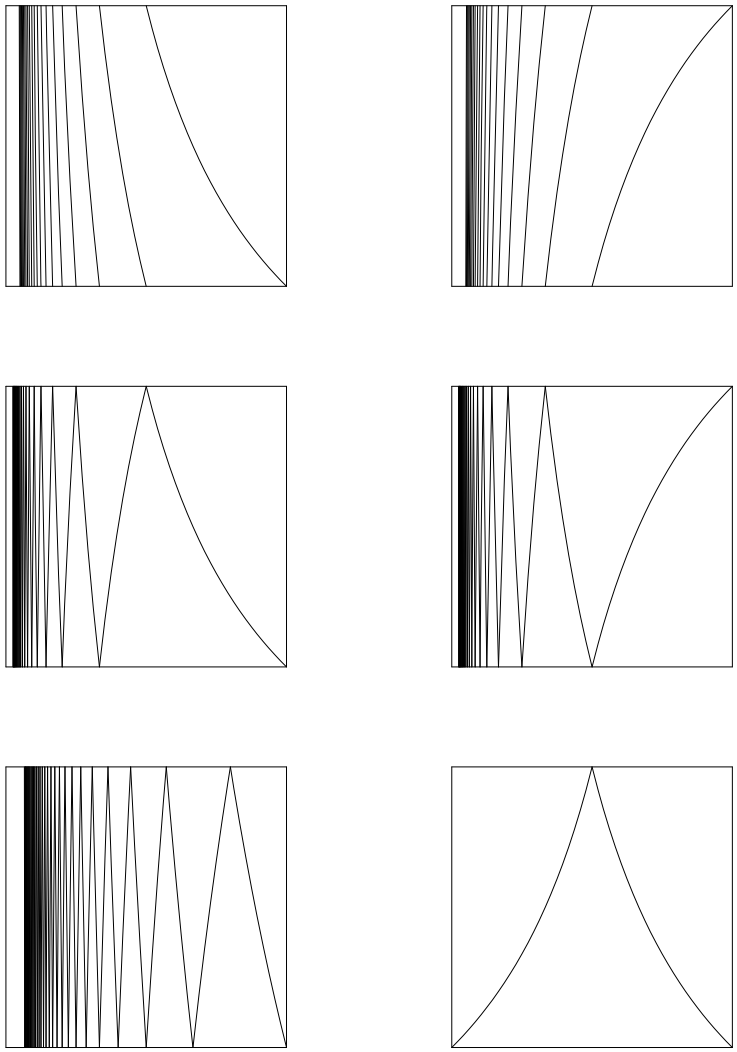

FIG. 13. Les six systèmes euclidiens classiques ; à gauche, les « rapides » : Standard, Impair, Centré ; à droite, les « lents » : Par Excès, Pair, Soustractif

Hausdorff de réels dont les fractions continues sont « contraintes ». On pourra consulter pour plus de détails [28, 29, 37, 44].

6. Un problème encore ouvert : l'analyse en distribution

Ici, nous avons décrit principalement des analyses en moyenne, où l'on cherche à déterminer principalement les valeurs moyennes de certains paramètres, ou plus généralement leurs moments d'ordre supérieur. C'est alors le comportement de l'opérateur de transfert (ou de ses généralisés) autour de la valeur $s = 1$ qui joue un rôle essentiel. Mais le rêve de l'algorithmicien consiste à effectuer une analyse en distribution de ces paramètres (*i. e.* chercher la distribution limite de ces paramètres quand la taille du problème devient grande). Ce sont alors les propriétés de l'opérateur

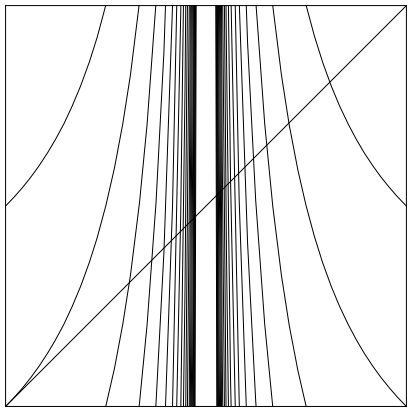

FIG. 14. La famille des systèmes japonais

de transfert à gauche de la droite $\Re(s) = 1$ qui vont intervenir. Plus précisément, une situation favorable est celle où le quasi-inverse $(\mathbf{1} - \mathbf{H}_s)^{-1}$ possède une région sans pôle à gauche de la droite $\Re(s) = 1$. Dans ce cas, on peut espérer obtenir une distribution limite gaussienne pour une certaine classe de paramètres liés au système dynamique. Cela permettrait tout particulièrement d'obtenir une nouvelle preuve, plus simple, du résultat d'Hensley [31] qui montre que le nombre d'itérations de l'algorithme d'Euclide suit une loi asymptotiquement gaussienne. C'est l'objet d'un travail en cours [3].

Références

[1] AKHAVI, A., VALLÉE, B. Average bit–complexity of Euclidean Algorithms, Proceedings of ICALP'00, *Lecture Notes in Computer Science* 1853, pp 373-387, Springer.

[2] BALADI, V. *Positive Transfer operators and decay of correlations,* Advanced Series in non linear dynamics, vol 16, World Scientific (2000).

[3] BALADI, V., VALLÉE, B. Analyse dynamique en distribution, In preparation.

[4] BEDFORD, T., KEANE, M., AND SERIES, C., Eds. *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, Oxford University Press, 1991.

[5] BRENT, R.P. Analysis of the Binary Euclidean algorithm, *Algorithms and Complexity, New Directions and Recent Results*, ed. by J.F. Traub, Academic Press 1976, pp 321–355.

[6] BOURDON, J. Size and Path-Length of Patricia Tries : Dynamical Sources Context., *Random Structures and Algorithms (2001), pp 289–315.*

[7] BOURDON, J. Analyse dynamique des algorithmes : exemples en algorithmique du texte et en algorithmique arithmétique, Thèse de l'université de Caen (2002).

[8] BOURDON, J., DAIREAUX, B., VALLÉE, B. Dynamical analysis of α-Euclidean Algorithms, *Journal of Algorithms* 44 (2002) pp 246–285.

[9] BOURDON, J., NEBEL, M., VALLÉE, B. On the stack-size of general tries, *Theoretical Informatics and Applications* 35 (2001) pp 163–185.

[10] BOURDON, J., VALLÉE, B. Generalized Pattern-Matching statistics *Colloquium on Mathematics and Computer Science : Algorithms, Trees, Combinatorics and Probability*, B. Chauvin et al., ed., Birkhauser Verlag, Trends in Mathematics, 2002, pp 249–265.

[11] BOWEN, R. Invariant measures for Markov maps of the interval, *Commun. Math. Phys.* 69 (1979) 1–17.

[12] BOYARSKY, A. AND GORA, P. *Laws of Chaos, Invariant measures and dynamical systems in one dimension*, Probability and its applications, Birkhauser (1997).

[13] BROISE, A. Transformations dilatantes de l'intervalle et théorèmes limites, *Asterisque* 238, pp 5–109, Société Mathématique de France (1996).

[14] Chazal, F., Maume-Deschamps, V., Vallée, B. Erratum to "Dynamical sources in Information Theory : Fundamentals Intervals and Word Prefixes" by B. Vallée, *Les cahiers du GREYC* 2002 et *les Prépublications du Laboratoire de Topologie de Dijon* (299).

[15] Clément, J. Arbres digitaux et sources dynamiques, Thèse de l'université de Caen (2000).

[16] Clément, J., Flajolet, P., Vallée, B. Dynamical sources in information theory : A general analysis of trie structures, *Algorithmica* (2001), vol 29 (1/2) pp 307–369.

[17] Collet, P. Some ergodic properties of maps of the interval, *Dynamical systems, Proceedings of the first UNESCO CIMPA School on Dynamical and Disordered Systems* (Temuco, Chile, 1991), Hermann (1996).

[18] Daireaux, B. Analyse d'algorithmes du PGCD, Mémoire de DEA, Université de Caen, 2001.

[19] Daireaux, B., Vallée, B. Dynamical analysis of the Lehmer-Euclid Algorithm, *Les cahiers du GREYC* 2003.

[20] Daudé, H., Flajolet, P., Vallée, B. An average-case analysis of the Gaussian algorithm for lattice Reduction, *Combinatorics, Probability and Computing* (1997) 6, pp 397–433.

[21] Delange, H. Généralisation du Théorème d'Ikehara, *Ann. Sc. ENS*, (1954) 71, pp 213–242.

[22] Dixon, J. D. The number of steps in the Euclidean algorithm, *Journal of Number Theory* 2 (1970), 414–422.

[23] Fayolle, J. Paramètres des arbres suffixes dans le cas des sources simples, Mémoire de DEA, Université de Paris VI, 2002.

[24] Flajolet, P. Analytic analysis of algorithms, In Proceedings of the 19th International Colloquium "Automata, Languages and Programming", Vienna, July 1992, W. Kuich, editor,*Lecture Notes in Computer Science* 623, pp 186–210.

[25] Flajolet, P., Gourdon, X., Dumas, P. Mellin transforms and asymptotics : harmonic sums, *Theoretical Computer Science* 144 (1-2), 1995, pp 3-58.

[26] Flajolet, P., Guivarc'h, Y., Szpankowski, W., Vallée, B. Hidden pattern Statistics, Comptes-rendus de ICALP'01, *Lecture Notes in Computer Science* 2076, pp 152–165.

[27] Flajolet, P. and Sedgewick, R. Analytic Combinatorics, Book in preparation, voir aussi les *Rapports de Recherche INRIA* 1888, 2026, 2376, 2956.

[28] Flajolet, P., and Vallée, B. Continued fraction Algorithms, Functional operators and Structure constants, *Theoretical Computer Science* 194 (1998), 1–34.

[29] Flajolet, P., and Vallée, B. Continued Fractions, Comparison Algorithms, and Fine Structure Constants, *Constructive, Experimental et Non-Linear Analysis*, Michel Thera, Editor, Proceedings of Canadian Mathematical Society, Vol 27 (2000), pages 53-82.

[30] Heilbronn, H. On the average length of a class of continued fractions, Number Theory and Analysis, ed. by P. Turan, New-York, Plenum, 1969, pp 87-96.

[31] Hensley, D. The number of steps in the Euclidean algorithm, *Journal of Number Theory* 49, 2 (1994), 142–182.

[32] Hennion H. Sur un théorème spectral et son application aux noyaux lipschitziens, *Proc. Amer. Math. Soc.* 118 (1993) pp 627–634.

[33] Hwang, H-K. Théorèmes limite pour les structures combinatoires et les fonctions arithmétiques, PhD thesis, Ecole Polytechnique, Dec. 1994.

[34] Kato, T. *Perturbation Theory for Linear Operators*, Springer-Verlag (1980).

[35] Krasnoselsky, M. *Positive Solutions of Operator Equations*, P. Noordhoff, Groningen, 1964.

[36] Lasota, A. and Mackey, M. *Chaos, Fractals and Noise ; Stochastic Aspects of Dynamics,* Applied Mathematical Science 97, Springer (1994).

[37] Lhote, L. Modélisation et approximation de sources complexes, Mémoire de DEA, Université of Caen, 2002.

[38] Mayer, D. H. Continued fractions and related transformations, In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, T. Bedford, M. Keane, and C. Series, Eds. Oxford University Press, 1991, pp. 175–222.

[39] Prellberg, T. and Slawny, J. Maps of intervals with indifferent fixed points : Thermodynamic formalism and Phase transitions. *Journal of Statistical Physics* 66 (1992) pp 503–514.

[40] Ruelle, D. *Thermodynamic formalism,* Addison Wesley (1978).

[41] Ruelle, D. *Dynamical Zeta Functions for Piecewise Monotone Maps of the Interval*, vol. 4 of *CRM Monograph Series*, American Mathematical Society, Providence (1994).

[42] Szpankowski, W. *Average Case Analysis of Algorithms on Sequences,* John Wiley and Sons, New York, 2001.

[43] Tenenbaum, G. *Introduction à la théorie analytique des nombres*, vol. 13. Institut Élie Cartan, Nancy, France, 1990.

[44] Vallée, B. Fractions continues à contraintes périodiques, *Journal of Number Theory* 72 (1998) pp 183–235.

[45] Vallée, B. Dynamical sources in information theory : fundamental intervals and word prefixes, *Algorithmica* (2001) vol. 29, pp 262–306.

[46] Vallée, B. Dynamics of the Binary Euclidean Algorithm : Functional Analysis and Operators, *Algorithmica* (1998) vol 22 (4) pp 660–685.

[47] Vallée, B. Algorithms for computing signs of 2×2 determinants : dynamics and average–case analysis, Congrès ESA'97 (5th Annual European Symposium on Algorithms) (Graz, Septembre 97), *Lecture Notes in Computer Science* 1284, pp 486–499.

[48] Vallée, B. Dynamical Analysis of a Class of Euclidean Algorithms, to appear in *Theoretical Computer Science* (2003).

[49] Vallée, B. Digits and Continuants in Euclidean Algorithms. Ergodic Versus Tauberian Theorems, *Journal de Théorie des Nombres de Bordeaux* 12 (2000) pp 531-570.