INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Team Preval

Probabilités, modélisation et évaluation de systèmes

Rocquencourt

THEME BIO

Activity Report

2006

Table of contents

1. Team

Responsable Scientifique
Guy Fayolle [DR, HdR]

Assistante de projet
Nicole Loza [jusqu'à mars 2006]
Thi-Thanh-Van Tran [à partir d'avril 2006]

Personnel Inria
Cyril Furtlehner [Ancien élève de l'ENS, Post-doc dans le projet Tao depuis octobre 2006]
Arnaud de La Fortelle [X-Ponts détaché par le ministère de l'Équipement]
Jean-Marc Lasgouttes [CR]

Collaborateurs extérieurs
Bernard Bercu [Université Bordeaux 1]
Roudolf Iasnogorodski [Université de Saint-Petersbourg]
Vadim Malyshev [Université de Moscou]

Chercheurs doctorants
Peggy Cénac [Boursière Inria jusqu'à septembre 2006]

2. Overall Objectives

2.1. Introduction

L'équipe PREVAL rassemble depuis 2001 les membres de l'ex-projet MEVAL. De nouveaux axes de recherche ont été explorés, qui concernent notamment les grandes structures issues de la physique, des transports et de la biologie. Naturellement, on suit le fil conducteur des compétences et des méthodes mathématiques originales développées antérieurement. Ce rapport s'inscrit donc dans le prolongement analytique direct des versions précédentes. Enfin, il est utile de signaler que PREVAL sera dissous au 31 décembre 2006 pour permettre la présentation d'une structure de recherche rénovée.

Le but essentiel des travaux est de parvenir à une bonne compréhension du comportement de systèmes aléatoires, d'origines variées (informatique, télécommunications, transports, physique, biologie), par la mise en œuvre de méthodes quantitatives d'évaluation. Deux moyens d'approche sont proposés :

- principalement, l'élaboration et la résolution de modèles mathématiques ;
- à titre complémentaire, la simulation.

Les compétences acquises par l'équipe couvrent largement les domaines liés à la *modélisation stochastique*, point n'étant besoin de rappeler que l'utilisation de la théorie des probabilités s'avère fructueuse dans nombre de branches de la science moderne (physique, économie, sociologie, etc.). L'évolution de la technologie (parallélisme, vitesse, modularité), la complexité et la taille sans cesse croissantes des systèmes ont eu des conséquences importantes :

- d'abord une demande plus forte d'analyse prévisionnelle de performance, afin d'assister les choix de conception ;
- ensuite, un impact considérable sur la théorie (réseaux de files d'attente, graphes et complexes aléatoires, analyse de séquences biologiques, systèmes de particules, etc).

Les recherches menées sont éclectiques et touchent à des domaines d'applications variés : réseaux téléinformatiques et de transport, physique statistique, réseaux de neurones, graphes et structures aléatoires. L'analyse macroscopique (temporelle et spatiale) de ces divers objets conduit (presque) inexorablement à étudier des processus stochastiques, souvent physiquement pertinents (temps de séjour dans un système, répartition de clients ou de particule, régime stationnaire, etc.). Les thèmes abordés comportent à la fois des aspects méthodologiques et des modèles particuliers. La démarche scientifique est toujours, à partir de problèmes concrets, de proposer des méthodes de portée générale, permettant d'obtenir des résultats quantitatifs sur l'efficacité, la stabilité ou les paramètre de contrôle de telle ou telle structure.

2.2. Axes de Recherche

Depuis une vingtaine d'années, des théories originales sont développées au sein de l'équipe dans les domaines d'expertise suivants.

- Réseaux et marches aléatoires : résolution d'équations fonctionnelles de plusieurs variables complexes ; classification des chaînes de Markov dans des polyèdres avec frontières à l'aide de systèmes dynamiques équivalents ; méthodes de construction de semi-martingales pour déterminer les conditions de stabilité des systèmes.
- Grands systèmes : lorsque la taille ou le volume d'un système augmente, et on parle alors de *limite thermodynamique*, des phénomènes de propagation du chaos ou de transition de phase apparaissent, comme en physique statistique classique. Le passage du microscopique au macroscopique demande aussi d'analyser la *limite hydrodynamique* des processu mis en jeu, après changement d'échelle ad hoc.
- Grammaires et complexes aléatoires : on bâtit de nouveaux ponts théoriques entre la physique et l'informatique (chaînes aléatoires, énumération, gravité quantique).
- Grandes déviations : il s'agit d'estimer les probabilités d'événements rares, qui parfois n'en sont pas moins cruciaux ! (ouragan, réacteur nucléaire divergeant, réseau internet en panne). Une exploitation fouillée de la notion d'entropie permet d'obtenir des résultats généraux.

2.3. Relations Internationales et Industrielles

- PREVAL entretient des relations suivies avec les universités de Berkeley, Bielefeld, Cambridge, EURANDOM, Columbia, Moscou, Novosibirsk, Monterey, San Diego.
- L'équipe est également partenaire de IMARA, programme d'expérimentation de nouvelles technologies pour le transport routier.

3. Scientific Foundations

3.1. Préambule

Keywords: *Chaîne de Markov*, *factorisation*, *file d'attente*, *fonction analytique*, *fonction de Lyapounov*, *grammaire*, *grande déviation*, *limite thermodynamique*, *marche aléatoire*, *modélisation*, *neurone*, *processus stochastique*, *protocole*, *réseau*, *semi-martingale*, *système dynamique*, *transition de phase*.

Le ciment existant entre les diverses activités s'est constitué à partir des champs scientifiques suivants : marches aléatoires (méthodes analytiques, classification), réseaux et grands systèmes, propagation du chaos et physique statistique. On donne ci-après un aperçu de quelques domaines d'expertise du projet, mettant en exergue les techniques originales qui y ont été inventées.

3.2. Réseaux et marches aléatoires dans $\mathbf{Z}_+^n$

Si on choisit l'exemple des systèmes de télécommunications ou de transport, la plupart d'entre eux se laissent modéliser de façon assez réaliste par un ensemble de stations de service, où les serveurs sont les ressources (logiques ou physiques) du système considéré et où les entités circulant entre les stations représentent les requêtes, messages, programmes partageant ces ressources. À un réseau formé de n stations avec plusieurs classes de clients, on pourra souvent associer un processus Markovien vectoriel

$$X(t) = (X_1(t), X_2(t), ..., X_n(t)),$$

$X_i(t)$ décrivant la configuration des clients à la station i au temps t. L'étude de ces processus nécessite, principalement, des outils de deux natures :

- probabiliste, pour trouver les conditions d'existence de régimes stationnaires ;
- analytique, lorsqu'on veut calculer des distributions, estimer des vitesses de convergence ou avoir des équivalents.

Pour des raisons physiques évidentes de positivité des quantités mises en jeu, il appert que les marches aléatoires dans $\mathbf{Z}_+^n$ sont isomorphes à des familles de réseaux comportant n sites. La difficulté majeure provient de l'existence des frontières naturelles.

3.2.1. Méthodes analytiques

Participants: Guy Fayolle, Roudolf Iasnogorodski, Vadim Malyshev.

Ce thème est fondamental et en perpétuel approfondissement. Lorsque $n = 2$, les sauts étant d'amplitude 1, la détermination de la mesure invariante équivaut souvent à résoudre une équation fonctionnelle de la forme

$$Q(x, y)\pi(x, y) + q(x, y)\pi(x) + \tilde{q}(x, y)\tilde{\pi}(y) + \pi_{00}q_0(x, y) = 0.$$

Il s'agit de trouver des fonctions $\pi(x, y), \pi(x), \tilde{\pi}(y)$, holomorphes dans les régions $|x|, |y| < 1$ et continues dans $|x|, |y| \leq 1$. Ici, $Q, q, \tilde{q}, q_0$ sont des polynômes. En se plaçant sur la courbe algébrique $Q(x, y) = 0$ (en général elliptique), deux approches fondamentales ont été développées, conduisant (par factorisation ou uniformisation) à des formes explicites des solutions.

- Dans [8] est introduit le *groupe* $\mathcal{G}$ de la marche aléatoire, dont l'étude, via les automorphismes de Galois et après uniformisation de la courbe algébrique, permet un calcul effectif sous forme de série.
- Dans [3], on se ramène à la résolution d'un problème aux limites de type Riemann-Hilbert-Carleman, pour une fonction d'une variable, posé dans le plan complexe et dont la forme la plus simple s'énonce ainsi : *trouver une fonction analytique dans un domaine simplement connexe $\mathcal{D}$ et satisfaisant une condition sur la frontière $\delta\mathcal{D}$.*

Vers la fin des années 70, cette méthodologie a été reprise avec fruit dans certains laboratoires étrangers (CWI, Université de Newcastle) et elle fait toujours l'objet d'études immédiates ou à long terme (Bell Laboratories, IBM Yorktown Heights, universités d'Ottawa, de Bordeaux, INRIA, etc.). Le livre [4], écrit par les trois auteurs précités, présente une synthèse des principaux résultats, qui est esquissée ci-dessous.

- Liaison avec la géométrie algébrique et le prolongement analytique des fonctions inconnues.
- Lorsque le groupe $\mathcal{G}$ est d'ordre fini, détermination des conditions nécessaires et suffisantes pour que les solutions soient rationnelles ou algébriques.
- Lorsque $\mathcal{G}$ est d'ordre infini, un problème de factorisation est formulé, dont *l'indice*, calculé analytiquement, donne les conditions d'ergodicité (i.e. pour que les fonctions cherchées n'aient pas de pôle dans le disque unité) ; la représentation des fonctions inconnues est donnée sous forme intégrale, à l'aide de la fonction $\wp$ de Weierstrass.

- Pour le genre 0 (problème posé sur la sphère de Riemann), expression des solutions au moyen des fonctions automorphes.
- Comportement asymptotique des probabilités stationnaires, par une analyse fine de type méthode du col + résidus sur la surface de Riemann sous-jacente.

Toujours en dimension 2, lorsque les sauts à l'intérieur du quart de plan sont non bornés selon les directions cardinales Est, Nord, Nord-Est, certaines généralisations ont été effectuées par J. W. Cohen, vers 1980 à l'université d'Utrecht. Par contre, il n'existe actuellement aucune technique de portée générale à partir de la dimension 3.

Le potentiel des méthodes évoquées ci-dessus est loin d'être épuisé. On citera par exemple :

- la détermination de la *vitesse intrinsèque* de convergence (i.e. invariante par rapport aux perturbations dans un domaine fini), pour des chaînes de Markov dans $\mathbf{Z}_+^2$;
- la description analytique de certaines *frontières de Martin* ;
- des équations fonctionnelles rencontrées en *gravité quantique* ;
- plusieurs problèmes de comptage et d'énumération combinatoire.

3.2.2. Classification et théorie constructive des chaînes de Markov dans $\mathbf{Z}_+^n$

Si les conditions de non explosion ou de non saturation des réseaux ont un intérêt pratique évident, leur caractérisation à l'aide de formules explicites résiste encore souvent à l'analyse, même pour de petites dimensions ($n = 2$ ou 3). On a vu ci-dessus, qu'il est difficile de disposer de la forme analytique de la mesure invariante, voire très vite impossible, sauf dans des cas bien répertoriés, tel celui des réseaux éponymes (Jackson, BCMP), dits à *formes produit* car leur mesure invariante est factorisable. Cependant, il existe maintenant une approche très générale pour obtenir les conditions d'ergodicité de marches aléatoires fortement homogènes dans $\mathbf{Z}_+^n$ ou dans des domaines similaires, qui donne une explication structurelle globale de la situation, en ramenant le problème de l'ergodicité d'une marche dans $\mathbf{Z}_+^n$ à une série de problèmes en dimension $n-1$, par construction de semi-martingales et de systèmes dynamiques (paragraphes suivants). Ces travaux fondamentaux ont reçu un cadre, avec la parution de l'ouvrage [7], dans lequel sont consignés des résultats originaux obtenus depuis une vingtaine d'années, sur la classification détaillée en termes d'ergodicité, de récurrence et de transience. Ce champ de recherches très vivant constitue l'un des points d'ancrage du projet.

3.2.3. Techniques de martingales

Participants: Guy Fayolle, Roudolf Iasnogorodski, Vadim Malyshev.

L'idée est de construire des *fonctions de Lyapounov* (FL) idoines, via des techniques de semi-martingales, la difficulté principale étant due à la présence de frontières. Cette approche a été décisive pour la résolution de nombreux problèmes (cf. rapports d'activité antérieurs) :

Fonctions de Lyapounov pour les réseaux En exhibant une FL linéaire par morceaux pour les fameux réseaux de Jackson, on a pu retrouver les équations de conservation de flux liées à l'existence d'un régime stationnaire. Le phénomène intéressant est qu'il est possible d'appliquer cette fonction à des politiques de service ou d'arrivées en groupes, alors que la mesure invariante pour ce système est hors de portée (D. Botvitch et A. Zamyatin à l'Université de Moscou).

Dérives nulles En dimension 2 et 3, lorsque les sauts moyens sont nuls à l'intérieur du cône positif, on a introduit de nouvelles classes de FL, non linéaires (en particulier de la forme $Q^\delta(x, y, z)$, Q étant une forme quadratique ; mais aussi des objets plus complexes faisant intervenir des caractéristiques géométriques). La classification complète des processus peut ainsi être obtenue et il existe des liens directs avec les diffusions.

Stabilité Une série de recherches sur l'intégrabilité de certaines fonctionnelles des temps d'atteinte τ_A d'ensembles compacts A, pour des chaînes de Markov, en collaboration avec S. Aspandiiarov (Université Paris 5) et M. Menshikov. Les moments $E\tau_A^p$ jouent en effet un rôle important dans les théorèmes limites concernant ces chaînes. En liaison avec la théorie du mouvement brownien réfléchi, on donne notamment la valeur critique p_0 maximale, telle que $E\tau_A^p < \infty$, $\forall p < p_0$, lorsque l'espace d'états est $\mathbf{Z}_+^2$. Là encore, les critères donnés reposent sur la construction explicite de semi-martingales et permettent d'étudier le comportement fin de la queue de distribution de τ_A, ainsi que la vitesse de convergence vers le régime stationnaire.

3.2.4. *Systèmes dynamiques*

Participants: Guy Fayolle, Jean-Marc Lasgouttes, Vadim Malyshev.

Comme il est montré dans [9] et [7], il est toujours possible d'associer à une marche aléatoire fortement homogène dans $\mathbf{Z}_+^n$, un système dynamique linéaire par morceaux, dont le champ de vecteurs associé (vitesses) ne dépende que des faces du domaine considéré. Cette propriété est équivalente à *l'approximation d'Euler* en physique statistique : on effectue des changements d'échelle identiques en temps et en espace (disons x/ϵ et t/ϵ) pour se placer dans le contexte de la loi des grands nombres. [Il est utile de remarquer d'emblée que, si certaines vitesses sont nulles, la situation se complique, car elle impose de mélanger la normalisation d'Euler avec celle du théorème central limite (diffusions) et la plupart des problèmes restent alors ouverts]. Ces méthodes, plus profondes que celles dites des *approximations fluides*, sont utilisées et développées pour la résolution de modèles très divers :

3.2.4.1. *Réseaux à une classe de clients*

Il s'agit principalement des systèmes dits à *polling* (scrutin), où V serveurs partagent leur puissance entre N stations, et des réseaux de files d'attente classiques avec routage Markovien. Un point agréable ici est qu'on peut souvent donner une caractérisation complète du champ de vecteurs, à l'aide de systèmes d'équations linéaires. Pour une large classe de politiques de service, ont été trouvées les conditions de stabilité des systèmes de polling à V serveurs homogènes avec routage Markovien : la condition nécessaire est donnée par un argument direct classique, l'obtention de la condition suffisante reposant sur l'étude d'un modèle fluide associé.

3.2.4.2. *Chaînes aléatoires en interaction*

On propose diverses lois de stabilisation pour l'évolution stochastique de chaînes modifiables à leurs extrémités gauche ou droite. Cette représentation à base de chaînes est commode et permet de dégager des liens entre des domaines apparemment disjoints : réseaux multiclasses, marches aléatoires sur des groupes (libres non commutatifs, modulaires, graphes, etc.), théorie des champs quantiques en dimension 3, réseaux de neurones, etc. En collaboration avec des chercheurs du LLRS (Univ. de Moscou), a été amorcée la construction d'une théorie décrivant les interactions mutuelles d'un nombre quelconque de chaînes. Elle englobe en particulier les réseaux de files multi-classes de type PAPS (premier arrivé, premier servi) ou DAPS (dernier arrivé, premier servi), et généralise le cadre des marches aléatoires homogènes classiques. Une vision synthétique des objets ci-dessus est donnée dans la section 3.5.

3.3. Grands systèmes aléatoires

Participants: Guy Fayolle, Jean-Marc Lasgouttes, Vadim Malyshev.

Schématiquement, on peut dire que la résolution de modèles raisonnablement réalistes (ce vocable étant bien sûr subjectif !) n'est possible que pour les dimensions extrêmes, c'est dire faibles (disons ≤ 3) ou très grandes. Dans ce dernier cas, on parle de systèmes en *limite thermodynamique*, i.e. dont la taille (volume, graphe associé, nombre de nœuds, etc.), représentée par un paramètre N, augmente indéfiniment. S'il y a clairement une analogie avec les systèmes de particules rencontrés en physique, de nouvelles difficultés surgissent par suite de la non homogénéité des composants (sites). Les questions essentielles sont liées aux phénomènes suivants :

3.3.1. Propagation du chaos

Elle existe dans un réseau si, par définition, tout p-uplet de nœuds se comporte, lorsque $N \to \infty$, comme un ensemble de p nœuds indépendants : on peut alors obtenir des renseignements quantitatifs sur la dynamique, l'état stationnaire et les vitesses de convergence. Pour les systèmes à fort degré de symétrie, les techniques utilisées sont issues de la mécanique statistique et dites *à champ moyen* : asymptotiquement, chaque site aura tendance à évoluer comme un processus de saut Markovien non homogène en temps, dont les taux de transition sont déterminés par calcul de la mesure empirique des actions dues aux autres sites, laquelle devient en fait déterministe. Une des difficultés provient de la rencontre d'équations différentielles non linéaires. Nous avons démontré la propagation du chaos pour diverses classes de réseaux (BCMP, *polling*). Quand la dissymétrie est totale, la plupart des problèmes restent encore largement ouverts.

3.3.2. Condensation et transition de phase

Considérons un réseau fermé du type standard BCMP, non symétrique, comportant M clients et N noeuds. On veut trouver des fonctions $M = f(N)$ conduisant, lorsque $N \to \infty$, à un *bon* comportement, autrement dit à une bonne utilisation des ressources disponibles. Cette problématique a de multiples origines : gestion de parcs de véhicules en libre service, réseaux informatiques, etc. À l'aide de facettes fines du théorème central limite, on montre dans [6] que, pour des conditions initiales indépendantes, il y a propagation du chaos, mais différentes situations peuvent se produire :

- les files restent toutes uniformément bornées ;
- certaines files, en nombre peut-être infini, se saturent : on parle alors de *condensation*.

On a ainsi très logiquement introduit une classe de fonctions $f(N)$, dites *critiques*, non nécessairement linéaires, différenciant les zones de *stabilité* et les zones de *saturation*. En outre, la vitesse de convergence vers l'état chaotique est d'ordre $O(f^{-1}(N))$.

3.3.3. Réseaux de files M/G/1 et réseaux markoviens généraux

On analyse des réseaux généraux de grande taille, pour lesquels la mesure invariante n'a pas de forme explicite. Il s'agit de déterminer les conditions de propagation du chaos, en dynamique ou en régime stationnaire, les techniques étant d'ailleurs fort différentes dans les deux cas.

Le premier candidat étudié est le réseau de Jackson ouvert, dans lequel les distributions des temps de service sont arbitraires. En fait, la possible existence du chaos s'avère a priori intimement liée à l'approximation poissonnienne des flux totaux arrivant à chaque file d'attente. Mais rien n'est ici évident, car il est connu que dans la plupart des réseaux à forme produit les flux internes ne sont pas de Poisson, même si chaque file se comporte comme une simple $M/M/1$ avec distribution géométrique.

Une approche basée sur l'analyse spectrale de générateurs infinitésimaux des processus semble prometteuse. Elle a déjà été testée avec succès dans [5]. On étend les méthodes *champ moyen* à certaines fonctionnelles de chaînes de Markov à espace d'états non dénombrable et ne possédant pas de symétrie. Des liens normaux apparaissent avec le théorème central limite pour des variables faiblement dépendantes.

3.4. Réseaux de neurones

Participant: Vadim Malyshev.

Globalement, on distingue trois classes de réseaux de neurones : ceux de *Hopfield*, les réseaux *biologiques* et le *perceptron*. Si des progrès concernant l'état stationnaire ont été accomplis par diverses équipes ces quinze dernières années, la dynamique de ces modèles, beaucoup plus délicate, reste encore assez mystérieuse. Sur ce dernier point, plusieurs études significatives ont été menées au sein de l'équipe.

Le modèle de Hopfield (1982), très répandu en mécanique statistique, est caractérisé par le nombre de nœuds N, qui peut tendre vers l'infini, et le nombre d'images p. Lorsque p est arbitraire, la dynamique à température finie est analogue à celle d'une marche aléatoire évoluant dans une union de polyèdres, avec discontinuités aux frontières (comme dans les réseaux de communications !). Les propriétés temporelles locales et globales ont été abordées en analysant les limites fluides obtenues à l'aide de changements d'échelle appropriés. Il a ainsi été montré que le temps pour atteindre un point fixe, à température zéro, est presque sûrement uniformément borné en p, lorsque $p \to \infty$. En ce qui concerne la région des températures finies, il faudra trouver les répartitions spatiales des minima locaux d'énergie et les distributions correspondantes.

La seconde catégorie de réseaux intervient fréquemment en biologie (modèle dit HOURGLASS) pour analyser le fonctionnement de cellules cérébrales. En fait, ils sont assez proches de certains réseaux de files d'attente et ont été étudiés par des techniques d'approximations fluides, conjointement avec l'IPPI de Moscou et l'Université de Lund en Suède. Il s'agissait principalement de décrire des attracteurs, ainsi que leur structure de Gibbs, pour divers types de connexions (asymétriques, aléatoires, d'inhibition et d'excitation, en dimension 2 aux températures extrêmes).

Enfin, la dernière classe de réseaux, analysée en collaboration avec le centre de physique théorique de Luminy, est plus connue des ingénieurs et apparaît comme une généralisation du *perceptron* multi-couches, proposé par Rosenblatt dans les années 60. Le point de vue adopté part d'une représentation sous forme de réseaux *feedforward*, i.e. sans réaction, dont la topologie est relativement simple.

3.5. Grammaires aléatoires, grammaires de graphes, grammaires quantiques

Participant: Vadim Malyshev.

Cette activité vise à développer les liens entre l'informatique et la physique théorique moderne, au delà des réseaux. Un mot dans un alphabet peut être vu comme une file d'attente, dont seules les extrémités sont modifiées de façon aléatoire, les symboles jouant le rôle des clients. Une généralisation immédiate consiste à admettre qu'arrivées et départs de clients puissent avoir lieu à un endroit arbitraire dans la file : il devient alors possible de décrire nombre de nouvelles applications, grâce à la notion de grammaire aléatoire (connue en informatique) et, plus généralement, de *grammaire de graphe*. Ces grammaires conduisent à des constructions géométriques (complexes aléatoires) maintenant très populaires en théorie de la gravité quantique ; elles participent à la modélisation de processus rencontrés en informatique (e.g. machine de Turing, fractales, graphes aléatoires) ; elles décrivent également l'évolution de longues séquence d'ADN en biologie. La dynamique et le comportement stationnaire des objets considérés sont analysés sous un angle nouveau, dans une série d'articles, notamment [23]. Les méthodes probabilistes utilisées sont fortement liées à la physique statistique (perturbations d'opérateurs, changements d'échelle, processus à interaction locale, limite thermodynamique).

On traite également des grandes grammaires de graphes multi-dimensionnelles, selon deux lignes directrices :

- d'abord, en faisant ressortir les propriétés qualitatives de ces objets : croissance asymptotique du nombre de composantes connexes, degré de compacité locale, chaos et phases topologiques, etc.
- Ensuite, en définissant la limite thermodynamique, plus exactement les aspects de dynamique en grappes (*clusters*) sur des configurations infinies. En particulier, on discute des différences entre grammaires de graphes et champs de Gibbs sur les treillis, en introduisant la notion de *complexe infini statistiquement homogène*.

Il est important de noter qu'existe une correspondance naturelle entre grammaires quantiques et systèmes à spins quantiques.

3.6. Complexes aléatoires

Participants: Guy Fayolle, Vadim Malyshev.

Les *complexes* interviennent dans différents domaines : topologie, géométrie, physique, etc. Il existe plusieurs points de vue.

- En mathématiques discrètes, il s'agit surtout de l'énumération de certaines classes de complexes bi-dimensionnels.
- En probabilités, la croissance de certains complexes fait appel à des processus de Markov, parfois non linéaires.
- Enfin, en physique, on s'intéresse à la gravité quantique, plus précisément aux conditions stochastiques laissant invariante la distribution stationnaire de cette gravité. Partant de l'étude théorique de la dynamique, il est possible de donner des résultats sur l'existence et les propriétés de fonctions de corrélations locales en limite thermodynamique.

3.7. Grandes déviations

Participants: Guy Fayolle, Arnaud de La Fortelle.

La théorie des grandes déviations s'intéresse principalement aux événements *rares*. Par exemple, pour un processus $(X_n, n \geq 1)$ prenant ses valeurs dans un espace métrique E et dont les moyennes $\hat{S}_n(\omega) = \frac{1}{n}\sum_{i=1}^{n} X_i(\omega)$ admettent une limite (e.g. les suites de variables i.i.d. ou les chaînes de Markov), on étudie le comportement asymptotique de $\mathsf{P}(\hat{S}_n \in \Gamma)$, où Γ est un borélien de l'espace E. De façon plus générale, il s'agit d'analyser le comportement de certaines familles de distributions dépendant d'un paramètre. Partant du travail fondamental d'Ellis en mécanique statistique, qui considère la mesure empirique d'ordre 2

$$L_n(\omega) = \frac{1}{n}\left(\sum_{i=1}^{n} \delta_{X_i(\omega),X_{i+1}(\omega)} + \delta_{X_n(\omega),X_1(\omega)}\right),$$

il a été montré dans [2] un théorème de Sanov généralisé, valable pour toute chaîne de Markov à espace d'états discret, irréductible mais non nécessairement ergodique. Plus précisément, L_n vérifie le *principe faible de grandes déviations* (PGD) : pour tout ouvert O (resp. fermé K) dans l'ensemble $M_s(E^2)$ des mesures équilibrées, on a

$$\liminf_{n\to\infty} \frac{1}{n} \ln \Pr[L_n \in O] \geq -\inf_{A\in O} H(A\|P),$$
$$\limsup_{n\to\infty} \frac{1}{n} \ln \Pr[L_n \in K] \leq -\inf_{A\in K} H(A\|P),$$

où la fonctionnelle d'action H est l'entropie relative. On retrouve le principe de moindre action : *pour atteindre un état exceptionnel, un système emprunte la trajectoire de moindre information*. Ce résultat est une amélioration de l'état de l'art, car habituellement les PGD soit supposent la finitude de E, soit imposent sur X une forte condition d'uniformité, qui exclut d'importantes classes de chaînes, notamment les réseaux de files d'attente à sauts bornés. Plusieurs conséquences intéressantes ont été mises en évidence. Ainsi,

$$\limsup_{n\to\infty} \frac{1}{n} \ln P^n(x,y) = -\inf_{A\in M_s(E^2)} H(A\|P) = -\alpha, \quad \forall (x,y),$$

la constante α étant nulle pour tout système ergodique. L'équation ci-dessus montre le lien existant entre le PGD et la frontière de Martin (i.e. l'ensemble des fonctions harmoniques), laquelle, dans le cas de transience, rend compte de la façon dont s'effectue la dérive vers l'infini. La plupart des résultats ci-dessus sont généralisables au temps continu et aux processus semi-markoviens, grâce à la notion de générateur empirique développée dans [10]. Il est en outre conjecturé, et vérifié dans plusieurs cas, que le PGD permet généralement d'accéder à l'information sur les queues des distributions stationnaires.

4. Application Domains

4.1. Introduction

Si la section précédente mettait l'accent sur la méthodologie, en liaison avec les processus stochastiques, il faut souligner que l'évolution des problèmes que nous traitons reste souvent guidée par les développements technologiques. Le mot « réseau » évoque aujourd'hui immanquablement les technologies de l'information, mais il couvre une problématique beaucoup plus large. En effet, l'information véhiculée peut avoir des contenus sémantiques aussi variés que le sont communications téléphoniques, transactions boursières, calculs scientifiques, marchandises, énergie, etc. « *Timeo hominem unius libri* », disait Saint-Thomas d'Aquin : de fait peu de domaines sont, par essence, foncièrement étrangers à la modélisation. Nous en énumérons trois, dont l'importance relative varie au fil des temps et qui ont de fortes dépendances mutuelles : Transports, Télécommunications, Biologie et Cinétique chimique.

4.2. Transports

Ce champ, très actuel et riche en questions difficiles, s'est inséré de façon effective dans nos activités sous l'impulsion du programme PRAXITÈLE (achevé en 1997), qui visait à la conception de nouveaux modes de déplacement en zone urbaine, basés sur l'utilisation de petits véhicules en libre service (cf. sections 6.7 et 7.1). La continuation de PRAXITÈLE est l'action IMARA sur la route automatisée, à laquelle participe l'équipe.

Ainsi, la conception de réseaux de transport constitués de véhicules automatiques (cybercars) *coopératifs* est devenue un enjeu majeur pour les prochaines années et la commission européenne en a fait un axe de recherche stratégique. Des questions fondamentales touchant au comportement de ces systèmes (contrôle centralisé ou distribué, exigences en termes de communications, etc) sont encore mal maîtrisées.

Des réflexions récentes ont mis l'accent sur des liens avec la mécanique statistique, ceci étant vrai aussi bien d'un point de vue macroscopique (gestion de flotte) que microscopique (comme l'étude fine du trafic routier). Lorsque la taille ou le volume des structures précitées augmente (ce qui correspond par exemple à un passage en limite thermodynamique), on souhaite établir des classifications selon divers critères quantitatifs (performance, débit, etc.) et qualitatifs (stabilité, comportement asymptotique, transition de phase, complexité, etc.). Nous avons entamé dans cet esprit une collaboration avec des chercheurs du Laboratoire de Physique Théorique et Modèles Statistiques (CNRS/Paris XI) et du CEA (Saclay).

4.3. Télécommunications

Il s'agit d'un terrain de prédilection pour l'application des outils présentés dans le paragraphe 3.2. De nombreuses études ont été réalisées au sein du projet dès le début des années 1980, certaines dans un cadre contractuel avec le CNET, au moment de l'essor des réseaux à commutation de paquets. Nous citerons, par ordre chronologique :

- l'évaluation de protocoles d'accès et de transmission de données (ALOHA, canal satellite, ETHERNET, etc) ;
- divers problèmes de multiplexage et allocation de ressources dans les réseaux large bande.

4.4. Modélisation des réseaux chimiques et des structures biologiques

Ce chapitre fait partie des axes de recherche explorés par PREVAL ces dernières années. On pourra consulter les récents rapports d'activité, ainsi que la section 6.3.

5. Software

5.1. ModelBench

Participant: Cyril Furtlehner.

ModelBench est un environnement graphique destiné à visualiser des effets physiques, à faire la démonstration d'algorithmes et à obtenir des résultats numériques. En particulier ont été implémentés :

- un modèle de véhicules automatiques coopératifs (cf. rapport d'activité 2004) ;
- un modèle stochastique de déformations de structure aléatoires unidimensionnelles, de nature biologique ou physique (cf. section 6.4) ;
- un modèle de prédiction de trafic (cf. section 6.7) ;
- des processus d'exclusion multi-types modélisant l'écoulement du trafic routier.

5.2. MyCGR

Participant: Peggy Cénac.

MyCGR est une bibliothèque et une collection d'outils permettant d'utiliser la CGR sur des séquences d'ADN pour implémenter les méthodes et résultats décrits la section 6.3.1. MyCGR est distribuée sous licence Cecill sur le site http ://mycgr.inria.fr/ et a été développé en collaboration avec Maxence Guesdon (Miriad, Gallium).

5.3. Publicité

À titre d'information, on rappelle que le projet, par l'intermédiaire de M. Badel, a longtemps servi de soutien logistique pour la conception et le développement de QNAP2. Ce logiciel permet la résolution et l'estimation de réseaux de serveurs et de files d'attente très généraux, par calculs mathématiques ou par simulation. Il est intégré dans l'environnement MODLINE, produit réalisé et commercialisé par la société SIMULOG.

6. New Results

6.1. Préambule

Les points énumérés ci-dessous doivent être plongés dans le contexte des sections 3. et 4., bien que les chronologies ne soient pas identiques, certains recouvrements étant par ailleurs inévitables. On a souhaité faire ressortir quelques directions stratégiques, dont ne font plus directement partie les applications réseaux de télécommunications.

6.2. Méthodes analytiques

6.2.1. Équations de type Lotka-Volterra

Participants: Guy Fayolle, Cyril Furtlehner.

L'étude de la convergence des trajectoires de marches aléatoires soumises à des déformations (voir section 6.4) conduit, de façon assez surprenante, à des problèmes variationnels sur des équations de Lotka-Volterra. Il s'agit en particulier, lorsque les courbes dans le plan des phases sont fermées, d'analyser la sensibilité de la période à des conditions initiales données sous forme intégrale. De nouveaux exemples de ces phénomènes sont mis en lumière dans [18], à propos du système connu sous le sigle ABC.

6.3. Biologie moléculaire

6.3.1. Chaos Game Representation et séquences biologiques

Participants: Peggy Cénac, Guy Fayolle, Jean-Marc Lasgouttes.

La CGR (*Chaos Game Representation*) est une méthode qui permet de faire correspondre, à une séquence donnée, une mesure empirique sur l'intervalle $[0,1]$ ou sur un autre ensemble compact de $\mathbb{R}^d$, $d \geq 1$. Dans [21], on montrait que, pour une séquence d'entrée donnée stationnaire ergodique, la CGR converge vers une mesure limite π, dont on caractérisait la dimension de Hausdorff dans le cas markovien.

On donnait aussi une relation simple vérifiée par π dans les cas i.i.d. et markovien, qui conduisait :

- d'une part à une généralisation de la notion de *dinucleotide relative abundance* utilisée pour les signatures génomiques ;
- d'autre part à l'élaboration de tests statistique pour comprendre la structure qualitative d'une séquence aléatoire de symboles dans un alphabet fini (indépendance, chaîne de Markov d'ordre m, etc). Contrairement aux tests classiques, les statistiques employées s'appuient sur toute l'histoire de la séquence. Les tests de structure étaient étendus dans [20] à des collections de partitions, afin d'éviter le choix rigide d'une partition unique. Ils ont été appliqués à des séquences génomiques de plusieurs espèces, tant pour les parties codantes que non codantes.

Dans [11] on présente de nouveaux arbres taxonomiques obtenus à partir des généralisations d'abondances relatives en dinucléotides à l'aide de la CGR. On met en évidence un apport d'information de la CGR par rapport aux méthodes de comptages de mots classiques pour la comparaison de ces signatures génomiques.

Ces méthodes ont donné lieu au développement du logiciel MyCGR (section 5.2).

6.3.2. Représentation en arbre quaternaire de séquences ADN

Participant: Peggy Cénac.

À partir d'une séquence de lettres, on construit un arbre digital de recherche (*Digital Search Tree*) sur l'ensemble des suffixes de la séquence inversée. L'arbre avec étiquettes numérotées est équivalent à la CGR, alors que l'arbre sans étiquettes correspond à une CGR *normalisée*.

Des résultats sur la hauteur et la profondeur d'insertion existent lorsque les séquences à placer dans l'arbre sont indépendantes les unes des autres. Il s'agit de généraliser ces résultats pour notre représentation, où les séquences à placer dans l'arbre sont fortement dépendantes.

En collaboration avec B. Chauvin, S. Ginouillac et N. Pouyanne (UVSQ), on étudie les cas de séquences i.i.d ou markoviennes. On montre en particulier que la profondeur d'insertion D_n dans l'arbre d'une lettre à l'instant n, ainsi que la longueur M_n d'une branche choisie aléatoirement dans l'arbre à l'instant n, vérifient les convergence en probabilité

$$\lim_{n\to\infty} \frac{D_n}{\log n} = \lim_{n\to\infty} \frac{M_n}{\log n} = \frac{1}{h},$$

où h est l'entropie de la séquence. De plus, on établit également les convergences presque sûres pour les longueurs des branches les plus courtes et les plus longues. Ces convergences presque sûres peuvent aussi s'interpréter comme des résultats de convergence sur les longueurs de plus longues répétitions d'une lettre dans une séquence. Ces travaux ont été publiés dans [16].

6.3.3. Convergence des moments dans le TLC pour martingales vectorielles et applications statistiques

Participants: Peggy Cénac, Guy Fayolle.

Soit (ξ_n) une suite de variables indépendantes et de même loi, avec $\mathbb{E}[\xi_n] = 0$ et $\mathbb{E}[\xi_n^2] = \sigma^2$. Soit $\Sigma_n = \xi_1 + \cdots + \xi_n$.

Le theorème de la limite centrale presque sûre (TLCPS) dit que, pour toute fonction h continue bornée, on a avec probabilité 1

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} h\left(\frac{\Sigma_k}{\sqrt{k}}\right) = \int_{\mathbb{R}} h(x)dG(x),$$

où G est la mesure gaussienne $\mathcal{N}(0,\sigma^2)$. Ce theorème a aussi été démontré dans un cadre de martingales.

En collaboration avec B. Bercu, on cherche à généraliser ces résultats à un cadre vectoriel, pour des fonctions h non bornées. Dans le rapport [15], on établit de nouvelles propriétés de convergence presque sûre de transformées de martingales vectorielles. On montre en particulier que, sous certaines conditions de régularité du processus croissant et sous certaines conditions de moments sur la martingale, il y a convergence des moments normalisés de tout ordre pair dans le TLCPS. On applique ces résultat aux modèles de régression linéaire, en particulier l'autorégressif, ainsi qu'à certains processus de branchement, ce qui permet d'obtenir des propriétés asymptotiques sur les erreurs d'estimation et de prédiction.

6.4. Enroulements de marches aléatoires et modèles d'exclusion

Participants: Guy Fayolle, Cyril Furtlehner.

Depuis trois ans, plusieurs études ont été réalisées, qui portaient sur des marches aléatoires évoluant dans le plan ou dans l'espace et soumises à diverses déformations stochastiques (cf. rapports d'activité 2004 et 2005).

L'article [18] s'inscrit dans ce cadre. On y étudie la limite continue de systèmes diffusifs unidimensionnels, qui décrivent en particulier les déformations stochastiques de courbes discrètes, codées de différentes façons. Ces systèmes constituent des cas particuliers de réactions-diffusions. Un formalisme fonctionnel général est élaboré pour traiter la limite hydrodynamique (cf. section 6.5.1). On s'intéresse également au régime stationnaire, les processus n'étant pas nécessairement réversibles et pouvant alors donner lieu à des états de type non-Gibbs. Un lien est établi entre les propriétés récursives à l'origine des solutions matricielles et les cycles dans un graphe d'états, en introduisant des courants de boucles par analogie avec des circuits électriques. En outre, à l'aide de l'approche fonctionnelle précitée, on établit un lien entre les constantes de structure impliquées dans ces relations de récurrence au niveau discret et les constantes qui apparaissent dans les systèmes de type Lotka-Voltera décrivant la limite fluide des états stationnaires. Finalement, à partir d'un schéma itératif, on obtient le Lagrangien permettant de rendre compte des fluctuations de courants de particules. L'équation de Hamilton-Jacobi qui en découle –et dont on extrait la fonctionnelle de grande déviation– est résolue dans le cas réversible, permettant de retrouver un résultat établi par ailleurs. Une version améliorée et compactée de [18] a donné lieu à [12].

6.5. Physique statistique

6.5.1. Limite hydrodynamique

Participants: Guy Fayolle, Cyril Furtlehner.

Il s'agit d'un projet global consacré aux déformations stochastiques de courbes (cf. section 6.4). Les problèmes sont essentiellement posés en termes de processus d'exclusion, le but final étant d'obtenir des équations hydrodynamiques pour ces systèmes, après les changements d'échelle adéquats. L'étude [22] revisitait le système ASEP sur le tore $\mathbf{Z}/N\mathbf{Z}$. La suite habituelle des mesures empiriques convergeait en probabilité, lorsque $N \to \infty$, vers une mesure déterministe, unique solution faible d'un problème de Cauchy. Le point crucial est l'analyse d'une famille d'opérateurs paraboliques, avec calcul au sens des variations, les variables étant de fait les valeurs prises par une fonction sur un ensemble de points possiblement infini. La méthode proposée présente quelques traits nouveaux et nous sommes en train de l'appliquer à divers systèmes d'exclusion multi-types, dont le fameux modèle ABC [17].

6.6. Grandes déviations : résultats généraux

Participant: Arnaud de La Fortelle.

Les années précédentes, on a traité plusieurs cas de réseaux de transport ou de télécommunications, essentiellement équivalents à des marches aléatoires maximalement homogènes dans $\mathbf{R}^n$ ou $\mathbf{Z}^n$.

La démarche générale, développée dans [10], se fonde sur la notion de processus *déviant*, où intervient un principe analogue à celui de la loi dite de moindre action : pour atteindre un état exceptionnel, un système emprunte la trajectoire d'information minimale. Un changement de mesure permet de focaliser sur cette trajectoire et H quantifie l'information. Au delà d'une technique de calcul, on dispose d'une méthode pour mieux appréhender la dynamique des systèmes. On a ainsi pu montrer l'existence d'un PGD pour les chaînes en temps continu, à l'aide de la notion de *générateur empirique*, qui est une extension de la mesure empirique classique.

En utilisant ces outils, on s'est intéressé dans [13] au processus de Yule, qui correspond à une marche spatialement non homogène. Le but est d'obtenir des grandes déviations précises, ainsi que des formules de changement de mesure.

6.7. Réseaux de transport

Participants: Cyril Furtlehner, Arnaud de La Fortelle, Jean-Marc Lasgouttes.

On s'intéresse à un problème de forte importance pratique : la reconstruction la prédiction de trafic routier à partir de véhicules traceurs [19], [14], dans le cadre du projet européen REACT.

Le cadre est un celui d'un modèle d'Ising sur un graphe espace-temps dont les sites sont les liens du réseau de base, pris à des instants discrets et connectés entre eux en fonction des corrélations obtenues au niveau des carrefours. Des informations concernant l'intensité du trafic sont collectées par des véhicules traceurs circulant de façon aléatoire sur le réseau routier. On obtient ainsi des données moyennes et des corrélations entre liens proches à dates consécutives. À partir de ces données, il est possible d'identifier les paramètres du modèle d'Ising pour que son approximation de Bethe possède les mêmes corrélations de paires. Cette dernière approximation est un outil classique de la physique théorique.

La partie « reconstruction de trafic » consiste, à partir des données temps réel et des données historiques, en une application de l'algorithme Belief Propagation (BP) de Pearl, qui fournit une approximation des marginales conditionnelles sur les paires de liens. Cette approche est cohérente avec le modèle, puisque l'on sait depuis quelques années que les points fixes de l'algorithme BP sont les minima locaux de l'approximation de Bethe. L'algorithme fonctionne par envoi de messages entre les nœuds, ce qui permet de propager positivement et négativement dans le temps l'information et, ainsi, de reconstruire le trafic passé tout en donnant des prédictions sur le trafic à venir.

Ces travaux ont conduit à l'implémentation du programme `Spika` par A. Chinea Manrique de Lara, ingénieur expert dans l'équipe IMARA.

7. Contracts and Grants with Industry

7.1. Imara, Lara

Le programme Praxitèle, qui avait donné lieu en 1993 à un consortium industriel regroupant l'INRIA, l'INRETS, la CGEA, DASSAULT AUTOMATIQUE, EDF et RENAULT, s'est achevé en Juin 1997. Il a été suivi d'une expérimentation en vraie grandeur dans la région de Saint-Quentin en Yvelines, terminée en 1999 de façon très convaincante. Il s'agissait d'étudier et de promouvoir un nouveau mode de transport public, basé sur l'utilisation de petits véhicules urbains, de préférence électriques, dont le parc serait géré par un système informatique s'appuyant sur un réseau de télécommunications.

La relation contractuelle établie avec PREVAL, issue d'une proposition intitulée *Problèmes de modélisation mathématique pour la conception de réseaux de véhicules électriques à usage collectif*, avait permis de subventionner plusieurs doctorants. Cette thématique a aussi concerné plusieurs équipes du centre scientifique franco-russe A. M. Lyapounov.

Actuellement, PREVAL collabore à l'action IMARA et est partie prenante de l'Action Nationale de Recherche et Développement LaRA en cours de création.

D'un point de vue modélisation, des réflexions récentes ont mis l'accent sur l'utilisation de méthodes issues de la mécanique statistique, par exemple dans le cadre des projets REACT et COM2REACT (voir section 6.7).

En outre, A. de La Fortelle œuvre fortement à l'élaboration et au suivi des contrats européens de l'équipe IMARA, dont il est le responsable adjoint.

8. Other Grants and Activities

8.1. Actions nationales

8.1.1. Relations académiques

Le projet entretient des collaborations plus ou moins étroites avec les centres de recherche suivants :

- Université de Moscou (V. Malyshev) ;
- Université de Saint-Petersbourg (R. Iasnogorodski) ;
- Université de Marne-la-Vallée (J. Diebolt) ;
- Université Bordeaux 1, Institut de Mathématiques de Bordeaux (B. Bercu) ;
- Université Paris 11, LPTMS (A. Comtet et S. Majumdar) ;
- France Télécom R&D, DAC/OAT (J. Roberts) ;
- École polytechnique (C. Graham) ;
- École normale supérieure de Paris (P. Brémaud, B. Derrida, J.-F. Le Gall) ;
- ENSAE (P. Doukhan) ;
- CEA (C. Godrèche et K. Mallick) ;
- ENSMP (C. Laurgeau).

Cette année a vu le renforcement des contacts précédemment amorcés avec le CEA, le LPTMS et l'ENSMP.

8.1.2. Séminaires

Le séminaire « Probabilité, Optimisation, Contrôle » a lieu à Rocquencourt, en synergie avec le projet MAX-PLUS. L'organisation côté PREVAL en est confiée à J.-M. Lasgouttes.

8.1.3. Divers

G. Fayolle est membre du Conseil Scientifique de l'INRIA.

A. de La Fortelle est membre de la commission *Positionnement Statique et Dynamique* au sein du Conseil National de l'Information Géographique.

8.2. Actions internationales

8.2.1. Centre Franco-Russe

G. Fayolle et V. Malyshev sont parmi les membres fondateurs de L'Institut *A. M. Lyapounov*, situé au sein de l'Université de Moscou, officiellement inauguré le 19 Décembre 1993, mais actuellement en total état d'hibernation. Entre autres activités, ils ont été co-responsables du projet *Probabilités et analyse de grands réseaux* et, dans ce cadre, ont organisé plusieurs séminaires internationaux. La partie russe comprenait des équipes du LLRS (*Laboratoire d'analyse des grands systèmes*, Université de Moscou) et du *Laboratoire R. Dobrushin* de l' IPPI (*Institut des problèmes de transmission de l'information*, dépendant de L'Académie des Sciences).

8.2.2. Comités de programmes et d'édition de revues

V. Malyshev fait partie du comité éditorial des revues russes *Problems of Information Transmission* et *Fundamental and Applied Mathematics* ; il est également fondateur et éditeur en chef du journal *Markov Processes and Related Fields*, G. Fayolle en étant membre du comité de rédaction.

G. Fayolle est membre du groupe de travail IFIP WG 7.3, qui comprend une centaine de membres élus et représente diverses communautés scientifiques intéressées par la modélisation des systèmes.

8.2.3. Relations internationales

G. Fayolle et V. Malyshev collaborent avec les universités et centres de recherches de Leuven (Belgique), Bielefeld et Braunschweig (Allemagne), Moscou (Département de probabilités de l'université + IPPI de l'Académie des Sciences), EURANDOM (Pays-Bas), Cambridge et Newcastle (Angleterre).

De longue date, l'équipe maintient divers contacts avec les États-Unis (Berkeley, Columbia, Georgia Inst. of Tech., San Diego, Monterey, AT&T) et avec la Russie (Moscou, Novossibirsk, Saint-Petersbourg).

9. Dissemination

9.1. Préambule

Les résultats obtenus dans l'équipe sont régulièrement diffusés dans certains des principaux colloques concernant le domaine et ils font l'objet de conférences et d'invitations dans divers centres de recherche. On ne cite que les éléments les plus marquants.

9.2. Visites de laboratoires

G. Fayolle a été invité à séjourner un mois à l'Institut Mittag-Leffler (Djursholm, Suède) dans le cadre d'un programme scientifique sur la modélisation des systèmes. Il a également reçu des invitations des universités de Moscou, de Newcastle et de Cambridge.

9.3. Conférences

A. de La Fortelle a présenté le contenu de [14] à la *Conference on Intelligent Systems*, Moscou, 23-27 octobre.

G. Fayolle a été invité aux séminaires de l'ENS Ulm et de l'IHP.

J.-M. Lasgouttes a fait un exposé sur le thème *Processus de naissance et de mort sur certains arbres aléatoires* au séminaire SAMOS de l'université Paris 1.

9.4. Activités universitaires

G. Fayolle a été membre du jury des épreuves orales du concours d'agrégation externe de mathématiques, du 23 juin au 14 juillet 2006.

J.-M. Lasgouttes intervient dans le Magistère de Finance de l'université Paris 1, où il est chargé du cours « Analyse de Données ».

9.5. Thèses

B. Bercu et G. Fayolle étaient co-directeurs de la thèse de P. Cénac [discipline *Mathématiques appliquées*], qui a été soutenue à l'université Paul Sabatier–Toulouse 3, le 13 juin 2006.

10. Bibliography

Major publications by the team in recent years

[1] S. ASPANDIIAROV, R. IASNOGORODSKI. *General Results on Stationary Measures of Recurrent Countable Markov Chains and their Applications*, in "Bernoulli", 1999.

[2] G. FAYOLLE, A. DE LA FORTELLE. *Large Deviation Principle for Markov Chains in Discrete Time*, in "Problems of Information Transmission", vol. 38, n^{o} 4, 2002.

[3] G. FAYOLLE, R. IASNOGORODSKI. *Two Coupled Processors : The Reduction to a Riemann-Hilbert Problem*, in "Z. Wahrscheinichkeitstheorie verw. Gebiete", vol. 47, 1979, p. 325-351.

[4] G. FAYOLLE, R. IASNOGORODSKI, V. A. MALYSHEV. *Random walks in the Quarter Plane*, Applications of Mathematics, n^{o} 40, Springer-Verlag, 1999.

[5] G. FAYOLLE, J.-M. LASGOUTTES. *Partage de bande passante dans un réseau : approches probabilistes*, 70 pages, Technical report, n^{o} 4202, INRIA, 2001, http://hal.inria.fr/inria-00072420.

[6] G. FAYOLLE, J.-M. LASGOUTTES. *Asymptotics and Scalings for Large Product-Form Networks via the Central Limit Theorem*, in "Markov Processes and Related Fields", vol. 2, n^{o} 2, 1996.

[7] G. FAYOLLE, V. A. MALYSHEV, M. V. MENSHIKOV. *Topics in the constructive theory of countable Markov chains*, Cambridge University Press, 1995.

[8] V. A. MALYSHEV. *Marches aléatoires, équations de Wiener-Hopf dans le quart de plan, automorphismes de Galois*, en russe, Éditions de l'Université de Moscou, 1970.

[9] V. A. MALYSHEV. *Networks and dynamical systems*, in "Adv. Appl. Prob.", vol. 25, 1993, p. 140-175.

[10] A. DE LA FORTELLE. *Contribution à la théorie des grandes déviations et applications*, Ph. D. Thesis, École nationale des ponts et chaussées, novembre 2000.

Year Publications

Doctoral dissertations and Habilitation theses

[11] P. CÉNAC. *Étude statistique de séquences biologiques et convergence de martingales*, Ph. D. Thesis, Inria Rocquencourt et Université Paul Sabatier Toulouse III, June 2006.

Articles in refereed journals and book chapters

[12] G. FAYOLLE, C. FURTLEHNER. *Stochastic Dynamics of Discrete Curves and Multi-Type Exclusion Processes*, in "J. of Statistical Physics", to appear, 2007.

[13] A. DE LA FORTELLE. *Yule process sample path asymptotics*, in "Electronic Communications in Probability", vol. 11, 2006, p. 193-199, http://www.math.washington.edu/~ejpecp/EcpVol11/paper20.abs.html.

Publications in Conferences and Workshops

[14] C. FURTLEHNER, A. DE LA FORTELLE, J.-M. LASGOUTTES. *Belief Propagation Algorithm for Traffic Prediction*, in "Proceedings of the 9 t h Intelligent Systems and Computer Sciences Conference", Faculty of Mechanics and Mathematics, Moscow State University, 2006.

Internal Reports

[15] B. BERCU, P. CÉNAC, G. FAYOLLE. *Convergence des moments dans le théorème de la limite centrale pour les martingales vectorielles*, Technical report, n^{o} 6056, INRIA, December 2006, https://hal.inria.fr/inria-00118982.

[16] P. CÉNAC, B. CHAUVIN, N. POUYANNE, S. GINOUILLAC. *Digital Search Trees and Chaos Game Representation*, accepté pour publication dans ESAIM Prob.& Stat., Technical report, n^{o} 5856, INRIA, March 2006, https://hal.inria.fr/inria-00070170.

[17] G. FAYOLLE, C. FURTLEHNER. *Hydrodynamic Limit of Some Multi-type Exclusion Processes*, to appear, Technical report, INRIA, 2006.

[18] G. FAYOLLE, C. FURTLEHNER. *Stochastic Dynamics of Discrete Curves and Exclusion Processes. Part 2 : Functional Equations and Continuous Descriptions*, Technical report, n^{o} 5808, INRIA, March 2006, https://hal.inria.fr/inria-00070216.

[19] C. FURTLEHNER, A. DE LA FORTELLE, J.-M. LASGOUTTES. *Belief-Propagation Algorithm for a Traffic Prediction System Based on Probe Vehicles*, Technical report, n^{o} 5807, INRIA, January 2006, https://hal.inria.fr/inria-00070217.

References in notes

[20] P. CÉNAC. *Test on the Structure of Biological Sequences via Chaos Game Representation*, in "Statistical Applications in Genetics and Molecular Biology", vol. 4, n^{o} 1, 2005.

[21] P. CÉNAC, G. FAYOLLE, J.-M. LASGOUTTES. *Dynamical systems in the analysis of biological sequences*, Technical report, n^{o} 5351, INRIA, October 2004, http://hal.inria.fr/inria-00070651.

[22] G. FAYOLLE, C. FURTLEHNER. *Stochastic Dynamics of Discrete Curves and Exclusion Processes. Part 1 : Hydrodynamic limit of the ASEP System*, Technical report, n^{o} 5793, INRIA, December 2005, http://hal.inria.fr/inria-00070229.

[23] V. A. MALYSHEV. *Random Grammars*, in "Russian Math. Reviews", vol. 2, 1998, p. 107-134.