INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *Team Preval*

# *Probabilités, modélisation et évaluation de systèmes*

*Rocquencourt*

THEME BIO

Activity Report

2004

# Table of contents

# 1. Team

**Responsable Scientifique**
Guy Fayolle [DR]

**Assistante de projet**
Sylvaine Sperte [jusqu'au 18/12/04]

**Personnel Inria**
Arnaud de La Fortelle [X-Ponts détaché par le ministère de l'Équipement]
Jean-Marc Lasgouttes [CR]
Vadim Malyshev [DR]

**Collaborateurs extérieurs**
Franck Delcoigne [EDF Clamart]
Cyril Furtlehner [contractuel LPTMS à l'Université Paris 11 depuis le 01/03/2004]
Roudolf Iasnogorodski [Université de Saint-Petersbourg]

**Chercheurs doctorants**
Peggy Cénac [Boursière Inria]

# 2. Overall Objectives

## 2.1. Introduction

L'équipe PREVAL rassemble les membres de l'ex-projet MEVAL. De nouveaux axes de recherche sont en cours de définition, qui concernent notamment les grandes structures issues de la physique, des transports et de la biologie. Naturellement, on vise à utiliser et à développer les compétences et les méthodes mathématiques originales développées antérieurement. Ce rapport est donc le prolongement analytique direct des versions précédentes.

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, séquences biologiques, 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 (on parle alors de limite thermodynamique), des phénomènes de propagation du chaos ou de transition de phase apparaissent, comme en physique classique.
- 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 national 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 en uniformisant 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 $\tau_A$ 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 $\tau_A$, ainsi que la vitesse de convergence vers le régime stationnaire.

### 3.2.4. Systèmes dynamiques

**Participants:** Frank Delcoigne, 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 $x/\epsilon$ et $t/\epsilon$) 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 :

**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.

**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:** Franck Delcoigne, 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 $\leq 3$) 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 $O(f^{-1}(N))$.

## 3.4. 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 [28]. 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 $\Gamma$ 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,

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

la constante $\alpha$ 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 : Télécommunications, Transports, Biologie et Cinétique chimique.

## 4.2. Télécommunications

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 citons, par ordre chronologique :

- l'analyse des procédures de liaison dans les systèmes téléinformatiques, principalement le standard HDLC alors très en vogue avec la notion de canal virtuel ;
- l'évaluation de protocoles d'accès et de transmission de données (Aloha, canal satellite Télécom-1, Ethernet, etc.) ;
- les problèmes de multiplexage liés à l'apparition des RNIS (*réseaux numériques à intégration de services*) ;
- allocation de ressources dans les réseaux large bande.

## 4.3. 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, 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 6.11 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.

## 4.4. Modélisation des réseaux chimiques et des structures biologiques

Ce chapitre fait partie des nouveaux axes de recherche de PREVAL. Les résultats partiels déjà obtenus sont décrits dans les sections 6.3 et 6.4.

# 5. Software

## 5.1. Transports

**Participant:** Cyril Furtlehner.

Actuellement, deux prototypes ad hoc ont été développés : l'un étudie des systèmes de véhicules automatiques coopératifs (cf. section 6.11) ; le deuxième simule les déformations de diverses structure aléatoires, de nature biologique ou physique (cf. section 6.5).

## 5.2. 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 6.5) 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 donnés sous formes intégrale.

### *6.2.2. Files d'attente avec service en ordre aléatoire*

**Participant:** Jean-Marc Lasgouttes.

En collaboration avec O. Boxma, S. Foss et R. Núñez Queija, on considère dans [11] une file d'attente à un serveur, le service des clients ayant lieu selon un ordre aléatoire. Pour une large classe de distributions du temps de service, on détermine le comportement de la queue de distribution du temps d'attente. Il est montré que cette quantité, lorsque les arrivées sont de Poisson et les temps de service sont à variation régulière, est proportionnelle à celle du cas PAPS (premier arrivé premier servi), avec une constante qui est donnée explicitement.

On s'intéresse aussi au comportement de la distribution du temps d'attente en trafic fort pour une file $M/G/1$. On obtient des résultats non seulement dans le cas où la variance du temps de service est finie, mais aussi quand sa distribution est à variation régulière avec variance infinie.

## 6.3. Réseaux chimiques

**Participants:** Guy Fayolle, Vadim Malyshev.

Les réseaux chimiques peuvent être vus comme des marches aléatoires non-homogènes en espace, avec sauts non bornés. Ils correspondent souvent à des objets mathématiques plus complexes que ceux issus des réseaux de télécommunications.

Dans [18], il s'agit de réseaux ayant un nombre fini de types de molécules, sous divers changements d'échelle temporel ou spacial (taux de réaction). On donne des conditions suffisantes d'existence de points fixes pour certaines équations de Boltzmann correspondant à des niveaux d'énergie discrets.

Dans [21], en collaboration avec S. Pirogov (IPPI), on introduit de nouveaux modèles de réseaux de cinétique chimique avec redistribution d'énergie, lorsqu'il y plusieurs types de molécules. Les réactions ont lieu par paires. Actuellement, les principaux résultats concernent les cas où les mesures invariantes ont une forme produit. Grâce à une interprétation probabiliste, certaines de ces mesures sont calculées explicitement et on montre qu'elles sont effectivement atteintes par convergence. Dans le cas d'une seule classse de molécules, on retrouve l'équation de Boltzmann et une situation analogue à celle du fameux modèle de M. Kac (1956).

Dans [25], on propose un modèle microscopique de thermodynamique chimique, avec des réactions lentes ou rapides. La redistribution d'énergie est analysée sous différentes échelles de temps, pour des systèmes pouvant comporter un nombre infini de particules.

## 6.4. Biologie moléculaire

### 6.4.1. Analyse de séquences ADN

**Participants:** Peggy Cenac, Guy Fayolle, Jean-Marc Lasgouttes.

La CGR (de l'anglais 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 $\mathbb{R}^d$, $d \geq 1$. Il s'agit de voir comment une telle mesure (dont on sait qu'elle détermine la séquence de manière unique) peut être utilisée pour comparer deux chaînes d'ADN de façon pertinente.

Dans [23], on montre que, pour une sequence d'entrée donnée stationnaire ergodique, la CGR converge vers une mesure limite $\pi$, dont on caractérise la dimension de Hausdorff dans le cas markovien.

Si on suppose que la loi de la séquence est Markovienne d'ordre $m$, il est possible de donner une forme explicite pour la transformée de Fourier de $\pi$, qui dans le cas i.i.d. $m = 0$ s'écrit

$$\varphi(t) = \prod_{k=0}^{\infty} g(4^{-n}t),$$

$g(.)$ étant elle-même une transformée de Fourier connue. Le comportement asymptotique de $\varphi$, qui donne des indications sur $\pi$, a été obtenu. Dans le cas $m > 0$, on trouve une forme produit similaire, mais $g$ est alors une matrice d'ordre $4^m$.

D'autre part, on donne une relation simple verifiée par $\pi$ dans les cas i.i.d. et markovien. Cette relation conduit à un test statistique de structure de séquence, plus précisément étudié dans [12].

Enfin, on montre comment la CGR permet de généraliser la notion de *dinucleotide relative abundance*, qui est utilisée pour définir une signature génomique.

### 6.4.2. 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*, DST sur l'ensemble des suffixes de la séquence inversée. L'arbre avec étiquettes numérotées est équivalent à la CGR (section 6.4.1), alors que l'arbre sans étiquettes correspond à une CGR *normalisée*.

Des résultats sur la hauteur, la profondeur d'insertion et le profil existent pour des DST construits sur des suites de séquences indépendantes, toutes de même loi. Il s'agit de généraliser ces résultats pour notre représentation.

En collaboration avec B. Chauvin, S. Ginouillac et N. Pouyanne, on a étudié le cas où la séquence est i.i.d, markovienne, ou faiblement mélangeante. 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 relations

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

où $h$ est l'entropie de la séquence

$$h = \lim_{n\to\infty} -\frac{1}{n}\, \mathbb{E}[\log p(w_1 ... w_n)] \,.$$

Ce résultat permet d'envisager des applications statistiques pour estimer la loi de génération d'une séquence.

### 6.4.3. Mathématiques et biologie moléculaire

**Participant:** Vadim Malyshev.

En collaboration avec S. Pirogov de l'IPPI, une revue générale des outils et domaines mathématiques concernés par la biologie moléculaire est en cours d'écriture. Elle devrait déboucher, sous forme d'un livre, sur une vision globale des problèmes avec une tentative d'unification selon les vues des auteurs.

### 6.4.4. Réseaux d'expression génique

**Participant:** Cyril Furtlehner.

Des expériences sur des puces à ADN, permettant d'obtenir des données d'expression génique sous forme de séries temporelles, sont en cours de réalisation au laboratoires de biologie moléculaire de l'ENS-Ulm (équipe de Claude Jacq). L'obtention de corrélations à partir de ces données, qui concernent les 6000 gènes de la levure, permet en principe d'établir le réseau d'expression génique responsable du fonctionnement de la cellule. Nous collaborons au niveau de la modélisation et du traitement statistique des données en développant un logiciel permettant d'automatiser la comparaison des séries temporelles, pour en extraire un graphe de corrélations. Le but est d'élaborer des modèles à base de physique statistique pour aider à comprendre la nature et les caractéristiques du réseau génique sous-jacent.

## 6.5. Enroulements de marches aléatoires et modèles d'exclusion couplés

**Participants:** Guy Fayolle, Cyril Furtlehner.

L'étude d'une marche aléatoire évoluant dans le plan et soumise à des déformations stochastiques a été publiée dans [14]. Selon les taux de déformations, en particulier suivant la valeur d'un paramètre $\eta$ responsable d'une brisure de symétrie gauche-droite, la distribution des enroulements de la marche est modifiée et le système présente trois phases différentes : enroulée, tendue et vitreuse. Une tranformation explicite est proposée, qui conduit à considérer le système comme résultant de deux processus d'exclusion couplés : les particules du premier se déplacent dans un environnement déterminé par la distribution des particules du second et vice-versa. Chaque composante peut alors être vue comme un modèle d'exclusion inhomogène. Pour toute marche périodique fermée, un changement d'échelle permet de montrer une convergence en loi (ou presque sûre sur un espace de probabilité enrichi) vers une courbe donnée par un système de deux équations différentielles stochastiques. La partie déterministe s'exprime explicitement à l'aide d'intégrales elliptiques. La dynamique est analysée par une approche formelle de type limite fluide, qui met en lumière un système d'équations de Burgers.

Dans [20], on généralise [14] (qui concernait de fait des enroulements de marches aléatoire définies sur un réseau carré), en considérant des marches en dimensions 2 et 3 et les liens avec des systèmes couplés de particules avec exclusion. On détermine les conditions d'équilibre pour ces systèmes, ainsi que la forme de l'énergie libre à partir de laquelle s'exprime la mesure invariante. La caractérisation des transitions de phases et l'identification des états métastables sont effectuées à partir d'un calcul de période de fonctions hyper-elliptiques (ou plus générales). Une méthode est mise en oeuvre pour avoir accès au régime hydrodynamique et hors-équilibre de ces modèles, en termes de systèmes d'équations différentielles couplées de type Burgers.

## 6.6. Arbres aléatoires

**Participants:** Guy Fayolle, Jean-Marc Lasgouttes.

Dans [15], en collaboration avec M. Krikun, on donne le taux de croissance et les conditions d'ergodicité pour une famille d'arbres aléatoires avec adjonction d'arcs suivant un processus de Poisson et suppression des feuilles avec un taux $\mu$, les nœuds pouvant appartenir à des classes différentes. Les résultats principaux mettent en vedette le fameux nombre $e$. On obtient une classification complète du processus selon les valeurs du facteur d'intensité $\rho = \lambda/\mu$ : ergodicité si $\rho \leq e^{-1}$ et transience si $\rho > e^{-1}$. Un phénomène de transition de phase apparaît : dans l'espace des paramètres, la région correspondant habituellement à la récurrence nulle n'existe pas. Cette situation est rare pour des chaînes de Markov à espace d'états dénombrable. La démonstration est essentiellement analytique et revient à étudier l'existence d'une équation intégrale non linéaire. On prouve aussi un théorème de type ergodique pour la hauteur $H$ de l'arbre qui, lorsque $t \to \infty$, croît presque sûrement linéairement à une vitesse obtenue analytiquement. On analyse assez finement certaines grandeurs naturelles (nombre de nœuds, distribution de $H$ dans le cas ergodique).

## 6.7. Partage de bande passante dans les réseaux de télécommunications

**Participants:** Franck Delcoigne, Arnaud de La Fortelle.

Ce thème était issu d'une action contractuelle avec France Télécom R&D (cf. [5]). Typiquement, on considère un réseau parcouru par un ensemble de routes de longueur fixée, dans lequel la bande passante disponible sur chaque canal est partagée entre les connexions actives suivant la politique dite du « min » : chaque lien partage sa capacité équitablement entre les différents appels qu'il gère, de façon totalement décentralisée ; en outre, un appel est servi au taux offert par le lien le plus chargé qu'il utilise.

Dans [13], sous l'angle des grandes déviations (cf. section 6.10), on généralise des résultats précédemment obtenus sur les réseaux en étoile. Il apparaît que les méthodes proposées fonctionnent sous des hypothèses satisfaites par des topologies et des classes de protocoles générales : partage de processeur généralisé, allocation-$\alpha$ de bande passante, politiques « min » et « équité max-min »).

## 6.8. Systèmes en limite thermodynamique

### *6.8.1. Synchronisation*

**Participant:** Vadim Malyshev.

En collaborarion avec A. Manita (Université de Moscou), on étudie des modèles où interviennent des synchronisation temporelles entre $N$ éléments (cellules, processeurs, etc.), qui travaillent à des vitesses différentes. Dans [16][17][26], trois types de résultats sont obtenus : convergence $t \to \infty$ à $N$ constant ;

système limite d'équations aux dérivées partielles lorsque $N \to \infty$ à $t$ fixé ; bornes uniformes sur la dispersion de vitesses en s'appuyant sur des techniques de graphes aléatoires.

### *6.8.2. Réseaux de files M/G/1 et réseaux markoviens généraux*

**Participants:** Guy Fayolle, Jean-Marc Lasgouttes.

Nous avons poursuivi l'analyse de 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], à l'occasion des travaux sur le partage de bande passante évoqués dans la section 6.7. 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.

## 6.9. Grammaires aléatoires, grammaires quantiques

**Participant:** Vadim Malyshev.

Les grammaires aléatoires existent depuis longtemps en informatique, mais les études concernant leur limite thermodynamique et leur comportement stationnaire sont très embryonnaires. Ces dernières années, on a analysé de façon détaillée les grammaires « context-free » dans la région sur-critique, i.e. lorsque la longueur du mot croît exponentiellement. Diverses statistiques et limites sont calculées, relatives notamment à la taille des sous-mots, lorsque $t \to \infty$.

Quant à l'introduction des grammaires quantiques, le but est double :

- contribuer au développement de l'informatique quantique ;
- introduire rigoureusement des modèles simples dans la physique moderne (spécialement pour la gravité quantique).

Elles ont des analogues en termes de systèmes à spins quantiques et sous divers aspects : caractère auto-adjoint de l'Hamiltonien, algèbres $C^*$ et leurs automorphismes, états KMS, lien avec le modèle classique de Lorentz.

Comme il a été brièvement évoqué en section 3.5, les travaux en cours visent à établir des ponts concrets entre ces deux classes d'objets et l'analyse de certaines séquences de bio-chimiques.

## 6.10. Grandes déviations : résultats généraux

**Participants:** Franck Delcoigne, Arnaud de La Fortelle.

Comme il a été évoqué dans la section 3.7, une approche nouvelle, basée sur une décomposition en cycles, avait permis d'étendre le théorème de Sanov aux chaînes de Markov à espace d'états dénombrable, sous la seule hypothèse d'irréductibilité [2] : la mesure empirique d'ordre 2 satisfait un PGD. On obtient un certain nombre de corollaires, parmi lesquels le lemme intégral de Varadhan ou encore, sous certaines conditions, un principe de contraction qui entraîne immédiatement le théorème de Sanov faible pour les mesures empiriques d'ordre 1. La preuve repose sur une agrégation des états et sur des propriétés fines de la fonctionnelle d'action $H$, l'entropie.

Une démarche alternative plus 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 a une méthode pour mieux appréhender la dynamique des systèmes. Cette démarche a permis de montrer 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. On a achevé l'étude [13], dans laquelle on détermine les modes de saturation pour un réseau étoilé de taille finie, non symétrique. Ces résultats théoriques donnent, au prix d'une heuristique encore à l'étude, un algorithme de calcul pour la queue de la distribution stationnaire. En comparant les résultats avec le critère d'ergodicité et avec les simulations, on constate que le critère des grandes déviations est de fait bien meilleur que celui de l'ergodicité. La topologie en étoile n'est pas une contrainte fondamentale et la méthode semble prometteuse.

## 6.11. Réseaux de transport

### *6.11.1. Localisation*

**Participant:** Arnaud de La Fortelle.

Dans [22], à partir de la dynamique d'un modèle de transport, nous calculons à chaque instant la propagation des erreurs dues aux mesures imprécises sur tout les paramètres. Ceci permet, en temps réel, d'avoir des bornes pour les incertitudes de localisation.

### *6.11.2. Simulation numérique*

**Participant:** Cyril Furtlehner.

On a poursuivi l'étude [27] dans laquelle, en liaison avec l'action IMARA, était proposé un modèle de voitures automatiques (*cybercars*) circulant sur un réseau quelconque. Les demandes sont aléatoires et il s'agit d'acheminer les utilisateurs vers des destinations également tirées au hasard suivant des poids prédéfinis. Le modèle est implémenté numériquement sur un simulateur graphique. Un algorithme basé sur l'effet d'un champ d'interaction sur le réseau entre les clients et véhicules permet d'optimiser l'offre en fonction de la demande et de limiter les temps d'attente. Le champ permettant d'orienter les véhicules est l'inverse d'un certain Laplacien discret et il tens à orienter les cybercars vers la demande tout en réduisant l'engorgement. Le simulateur a permis de montrer l'efficacité de l'algorithme sur un grand réseau de 1000 nœuds. Quelques démonstrations théoriques de ces phénomènes sont en cours.

# 7. Contracts and Grants with Industry

## 7.1. Praxitèle, imara

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 des équipes du centre scientifique franco-russe A. M. Lyapounov.

Actuellement, PREVAL participe à l'action IMARA, qui s'inscrit dans le prolongement de PRAXITÈLE et fait partie du consortium français « La Route Automatisée » en coopération avec l'INRETS, le LCPC, l'ENSMP, l'ENPC et l'ENST. Des réflexions récentes ont mis l'accent, du point de vue modélisation, sur des liens possibles avec la mécanique statistique. En outre, A. de La Fortelle participe à temps partiel à l'élaboration et au suivi des contrats européens d'IMARA.

# 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 Saint-Petersbourg (R. Iasnogorodski) ;
- Université de Marne-la Vallée (J. Diebolt) ;
- Université PARIS 6, Laboratoire de Probabilités (J. Jacod) ;
- Université PARIS 10, Département de Mathématiques (C. Léonard) ;
- Université Paul Sabatier, Laboratoire de Statistiques et Probabilités (B. Bercu) ;
- Université PARIS 11, LPTMS, A. Comtet
- France Télécom R&D, DAC/GTR (J. Roberts) ;
- École polytechnique (F. Dunlop, C. Graham) ;
- École normale supérieure de Paris (P. Brémaud, B. Derrida, J.-F. Le Gall, G. Ruget) ;
- ENSAE (P. Doukhan), CEA (K. Mallick).

### *8.1.2. Séminaires*

Le séminaire hebdomadaire *Probabilité Optimisation Contrôle* a lieu à Rocquencourt, en synergie avec le projet MAX-PLUS. L'organisation, côté PREVAL, en est confiée à A. de La Fortelle.

### *8.1.3. Divers*

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

Arnaud 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 participent aux activités du centre *A. M. Lyapounov*, situé au sein de l'Université de Moscou et officiellement inauguré le 19 Décembre 1993. Ils ont été co-responsables du projet intitulé *Probabilités et analyse de grands réseaux* et ont également organisé plusieurs séminaires. La partie russe comprenait des équipes du LLRS (*Laboratoire d'analyse des grands systèmes*, Université de Moscou) et le *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 ont été diffusés dans quelques uns des principaux colloques concernant le domaine et ont fait 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 deux 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.

V. Malyshev s'est rendu plusieurs fois en Russie (universités de Saint-Pétersbourg et de Moscou, IPPI). En liaison avec divers programmes de coopération internationale, il a aussi visité les universités de Bielefeld (Allemagne) et de Rome, où il a donné des séminaires sur des problèmes mathématiques en biologie moléculaire en cinétique chimique.

## 9.3. Conférences

A. de La Fortelle a participé à la conférence CODATU (Bucarest, 23–25 avril), où il a présenté les perspectives liées aux cybercars. Il a donné un séminaire le 26 avril à Lille (APPA), sur l'impact des systèmes de transport du futur sur le plan écologique. Le 9 juin, il a fait une conférence à l'université de Pau sur le thème *Problèmes de grandes déviations*. Enfin, il a assisté à la rencontre HITACHI–INRIA (Sophia-Antipolis, 19 novembre), qui portait sur les réseaux de télécommunications dans les systèmes de transport du futur.

G. Fayolle a exposé le contenu de l'article [15] au séminaire du département de probabilités de l'université P. Sabatier (Toulouse, 7 mai). Invité par P. Flajolet et le groupe ALEA, il a donné six conférences spéciales à l'université de Versailles-Saint-Quentin (15-16 novembre) sur le contenu du livre [4].

G. Fayolle et C. Furtlehner ont présenté respectivement [21] et [14] au *Third Colloquium on Mathematics and Computer Science*, Vienne 13-17 septembre.

## 9.4. Activités universitaires

G. Fayolle a été membre du jury des épreuves orales du concours d'agrégation externe de mathématiques, du 20 juin au 17 juillet 2004.

## 9.5. Thèses

Guy Fayolle est co-directeur avec B. Bercu (Laboratoire de Statistique et Probabilités, université P. Sabatier) de la thèse de P. Cénac, qui a débuté en décembre 2002.

# 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://www.inria.fr/rrrt/rr-4202.html.

[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.

## Articles in referred journals and book chapters

[11] O. J. BOXMA, S. G. FOSS, J.-M. LASGOUTTES, R. NÚÑEZ QUEIJA. *Waiting time asymptotics in the single server queue with service in random order*, in "Queueing Systems, Theory and Applications", vol. 46, n$^o$ 1, 2004, p. 35-74.

[12] P. CÉNAC. *Test on the structure of biological sequences via Chaos Game Representation*, in "submitted to Journal of Computational Biology", 2004.

[13] F. DELCOIGNE, A. DE LA FORTELLE. *Large deviation problems for star networks : the min policy*, in "Annals of Applied Probability", vol. 14, n$^o$ 2, 2004, p. 1006-1028.

[14] G. FAYOLLE, C. FURTLEHNER. *Dynamical Windings of Random Walks and Exclusion Models. Part I : Thermodynamic Limit in* $\mathbf{Z}^2$, in "Journal of Statistical Physics", vol. 114, n$^o$ 1/2, 2004, p. 229-260.

[15] G. FAYOLLE, M. KRIKUN, J.-M. LASGOUTTES. *Birth and Death Processes on Certain Random Trees : Classification and Stationary Laws*, in "Probability Theory and Related Fields", Digital Object Identifier : 10.1007/s00440-003-0311-1, vol. 128, n$^o$ 3, 2004, p. 386-418.

[16] V. MALYSHEV, A. MANITA. *Uniform bounds in time synchronization models*, in "Theory of Probability and Its Applications", to appear, 2004.

[17] V. MALYSHEV, A. MANITA. *Asymptotics in time synchronization model*, in "Asymptotic Combinatorial Methods in Representation Theory and Dynamical Systems", Adv. Math. Sc., to appear, American Math. Society, 2005.

[18] V. MALYSHEV, S. PIROGOV, A. RYBKO. *Random Walks and Chemical Networks*, in "Moscow Math. Journal", vol. 4, n$^o$ 2, 2004.

[19] V. MALYSHEV, A. ZAMYATIN. *Some remarks about high density Fermi systems*, in "Mathematical Notes", to appear, 2005.

## Publications in Conferences and Workshops

[20] G. FAYOLLE, C. FURTLEHNER. *Stochastic Deformations of Sample Paths of Random Walks and Exclusion Models*, in "Mathematics and Computer Science III, Basel/Switzerland", Trends in Mathematics, Birkhäuser Verlag, 2004.

[21] G. FAYOLLE, V. MALYSHEV, S. PIROGOV. *Stochastic Chemical Kinetics with Energy Parameters*, in "Mathematics and Computer Science III, Basel/Switzerland", Trends in Mathematics, Birkhäuser Verlag, 2004.

[22] M. KAIS, S. MORIN, A. DE LA FORTELLE, C. LAUGIER. *Geometrical Model to Drive Vision Systems with Error Propagation*, in "Proceedings of the International Conference on Control, Automation, Robotics and Vision", 2004.

## Internal Reports

[23] P. CÉNAC, G. FAYOLLE, J.-M. LASGOUTTES. *Dynamical systems in the analysis of biological sequences*, Technical report, n$^o$ 5351, INRIA, October 2004, http://www.inria.fr/rrrt/rr-5351.html.

[24] F. DELCOIGNE, A. DE LA FORTELLE. *Large deviations for a class of Markov processes modelling communication networks*, revised version submitted to Annals of Applied Probability, December 2004, Technical report, n$^o$ 4474, INRIA, 2002, http://www.inria.fr/rrrt/rr-4474.html.

[25] V. MALYSHEV. *Microscopic Models for Chemical Thermodynamics*, submitted to J. Stat. Physics, Technical report, n° 5200, INRIA, May 2004, http://www.inria.fr/rrrt/rr-5200.html.

[26] V. MALYSHEV, A. MANITA. *Time synchronization models*, Technical report, n° 5204, INRIA, May 2004, http://www.inria.fr/rrrt/rr-5204.html.

## Bibliography in notes

[27] C. FURTLEHNER. *Stochastic Model of Automatic Traffic Service Controlled by Lattice Coulomb Interactions*, Technical report, n° 4681, INRIA, December 2002, http://www.inria.fr/rrrt/rr-4681.html.

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