INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *Projet POLKA*

*Polynômes, Combinatoire, Arithmétique*

*Nancy*

THÈME 2B

# Table des matières

POLKA *est un projet de l'UMR 7503 Loria, commun à l'INRIA, au CNRS, à l'Université Henri Poincaré Nancy 1, à l'Université Nancy 2, et à l'Institut National Polytechnique de Lorraine.*

# 1 Composition de l'équipe

**Responsable scientifique**

Grégory Kucherov [CR INRIA, responsable scientifique à partir du 1/2/00]

**Responsable permanent**

Paul Zimmermann [DR INRIA, responsable permanent depuis le 1/2/00]

**Assistante de projet**

Geneviève Grisvard-Pierrelée [AGT INRIA]

**Personnel INRIA**

Guillaume Hanrot [CR]

Vincent Lefèvre [CR, depuis le 1/10/00]

Fabrice Rouillier [CR, mis en disposition au LIP6 depuis le 1/1/00]

**Personnel CNRS**

Jean-Luc Rémy [CR]

Gilles Schaeffer [CR]

**Personnel université**

Isabelle Debled-Rennesson [Maître de conférences IUFM de Lorraine]

Jocelyne Rouyer [Maître de conférences UHP]

**Chercheur doctorant**

Luc Rolland [bourse co-financée Région Lorraine et INRIA]

**Chercheur post-doctorant**

Vincent Bardey [du 1/4/00 au 30/9/00]

**Stagiaires**

Sébastien Briais [ENS Lyon, du 5/6/00 au 17/7/00]

Mariana-Liliana Ibanescu [Université de Iaşi, Roumanie, du 15/6/00 au 31/8/00 et du 15/10/00 au 15/12/00]

Emmanuel Jeandel [ENS Lyon, du 1/6/00 au 15/7/00]

Vianney Miard [UHP, du 1/7/00 au 31/7/00]

**Chercheurs invités**

Alain Goupil [du 1/7/00 au 31/7/00]

Roman Kolpakov [poste d'accueil de spécialistes, depuis le 1/9/00]

Mostefa Mesmoudi [du 16/3/00 au 16/4/00 et du 4/9/00 au 15/9/00]

**Collaborateur extérieur**

Jean-Charles Faugère [LIP6, Paris]

# 2 Présentation et objectifs généraux

L'objectif général du projet POLKA est de développer des méthodes systématiques traitant certaines classes bien définies de problèmes algorithmiques, et d'implanter ces méthodes de façon efficace pour traiter des applications en grandeur nature. Parmi les nombreux domaines couverts, trois axes nous intéressent particulièrement.

### 2.0.1 Arithmétique.

Cet axe de recherche, initialement motivé par la participation à une action de recherche coopérative sur le calcul fiable, a pris une importance croissante au sein du projet. Les deux sous-axes visés sont d'une part une sémantique rigoureuse de l'arithmétique des nombres flottants en précision arbitraire, complétée par une implantation, d'autre part la conception et le développement d'algorithmes pour la résolution d'équations diophantiennes, et leurs applications à divers problèmes mathématiques.

#### 2.0.2 Polynômes.

C'est un domaine dans lequel les membres du projet ont acquis une certaine expérience au cours des dix dernières années, notamment grâce aux projets européens PoSSo et Frisco. Cette compétence dans la résolution des systèmes polynomiaux a porté ses fruits dans des domaines applicatifs comme les filtres de compression d'images ou la robotique, et nous poursuivons dans cette seconde voie, notamment avec une thèse pluridisciplinaire automatique-informatique débutée en 1998. Parallèlement, des recherches théoriques à long terme sont entreprises en vue d'étendre les méthodes actuelles à des systèmes ayant un nombre infini de solutions.

#### 2.0.3 Structures combinatoires.

Nos études dans ce domaine portent sur les propriétés combinatoires et algorithmiques de structures discrètes, telles que mots, arbres, graphes, cartes, polyominos. Le but consiste en général à mettre au point des algorithmes efficaces, qui sont souvent basés sur des études théoriques profondes de propriétés combinatoires de ces structures. Nous développons également des logiciels expérimentaux basés sur des algorithmes issus de nos travaux. Le champ d'applications visé de cet axe est la bioinformatique, domaine dans lequel les modèles discrets apparaissent de façon naturelle et essentielle.

Polka est projet-coordinateur de l'action de recherche coopérative Remag (Recherche et Extraction de Motifs pour l'Analyse Génomique) et participe à d'autres projets et actions de recherche, nationaux et régionaux, sur le thème de la bioinformatique. Dans l'axe *arithmétique*, Polka participe à l'action de recherche coopérative AOC (Arithmétique des Ordinateurs Certifiée).

## 3 Fondements scientifiques

### 3.1 Arithmétique

Les nombres (entiers ou flottants, en précision arbitraire) sont pour nous un moyen de montrer des résultats exacts dans le domaine du calcul symbolique. Sur les nombres entiers, nous étudions essentiellement des équations diophantiennes algébriques. Notre démarche consiste à identifier des classes d'équations résolubles, puis à développer des algorithmes efficaces pour les résoudre. Ces algorithmes combinent l'utilisation de bornes supérieures sur la taille des solutions, et des recherches exhaustives. La difficulté principale – une fois la borne supérieure obtenue, ce qui peut nécessiter une étude théorique non négligeable – consiste à réduire la taille des solutions (en utilisant des arguments numériques), mais aussi à mener à bien beaucoup plus efficacement la recherche exhaustive finale.

En ce qui concerne les nombres flottants, notre objectif à long terme est de développer une bibliothèque portable et fiable de calcul sur de tels nombres en précision arbitraire. Par « fiable » nous entendons entre autres le calcul exact des différents modes d'arrondi (au plus proche, vers zéro, moins et plus l'infini) comme c'est le cas pour les `double` en C grâce à la norme IEEE-754. Il s'agit aussi de développer ou d'améliorer des algorithmes efficaces de calcul

sur les nombres flottants. Ces algorithmes sont bien sûr basés sur les algorithmes calculant sur les entiers, mais des méthodes spécifiques peuvent être développées.

### 3.2 Polynômes

L'étude des zéros réels des systèmes polynomiaux constitue un lien direct entre le calcul formel et les applications. Notre but est de fournir des outils performants pour des applications en vraie grandeur. Ce travail comporte plusieurs aspects :
(i) un aspect théorique en géométrie réelle pour l'élaboration d'algorithmes effectifs : méthodes de comptage avec ou sans contraintes, simplification des systèmes, isolation des racines, traitement des hypersurfaces ;
(ii) un aspect algorithmique pour l'optimisation pratique des méthodes effectives : diminution du nombre d'opérations arithmétiques, contrôle de la taille des données en cours de calcul ;
(iii) un aspect informatique pour l'implantation efficace de ces méthodes sur un grand nombre de plates-formes.

### 3.3 Structures combinatoires

Les travaux de cet axe peuvent être classifiés en trois volets.

Le premier volet concerne l'étude de *motifs* dans les mots et les arbres. Ici, notre spectre d'études est large – allant d'études théoriques sur la décidabilité et la complexité de problèmes liés aux motifs, en passant par des recherches sur des problèmes «classiques» de la combinatoire du mot, jusqu'au développement et à l'implantation d'algorithmes efficaces de recherche de motifs dans les mots et leur application à l'analyse de séquences génomiques.

Une autre voie de recherche concerne la reconstruction d'ensembles discrets bidimensionnels à partir de leurs projections horizontale et verticale. Ces travaux font appel à des outils de géométrie discrète qui sont en lien étroit avec les propriétés combinatoires des mots, comme la reconnaissance de segments de droites discrètes.

Enfin les propriétés des structures discrètes aléatoires sont étudiées sous différents aspects. Les méthodes énumératives, qui font appel aux combinatoires aussi bien bijective, asymptotique qu'algébrique, servent à aborder théoriquement des questions de pertinence de motifs extraits, ou plus classiquement d'aléa discret *i.e.* d'analyse d'algorithmes en moyenne et de propriétés statistiques de structures discrètes. Le développement d'algorithmes de génération aléatoire permet de venir compléter les résultats théoriques par l'expérimentation.

## 4 Domaines d'applications

### 4.1 Robots parallèles

Notre travail s'articule autour des diverses étapes de conception d'un mécanisme à architecture mécanique parallèle dotée d'une unité de commande dans le but d'obtenir un manipulateur et une commande fiables.

Dans ce domaine, quatre familles de problèmes surgissent : la conception pure du mécanisme en tant que tel, la réalisation des algorithmes de commande, la validation de la faisabilité des tâches d'usinage et le diagnostic des performances d'un robot réel.

Les manipulateurs que nous étudions sont des robots parallèles généraux : les hexapodes sont des mécanismes complexes constitués de six chaînes cinématiques souvent identiques, d'une base (corps rigide fixe comprenant six joints ou articulations) et d'une nacelle (corps rigide mobile contenant six autres joints).

Notre but principal est de produire un environnement de calcul fiable pour permettre une étude hors-ligne précise et efficace du comportement de n'importe quel robot parallèle. L'enjeu principal est de savoir mesurer l'impact sur la trajectoire d'un robot des différentes erreurs apparaissant aux diverses étapes de la commande :

- approximation de la trajectoire théorique par un outil de CAO ;
- effets de la discrétisation ;
- limites matérielles (capteurs, cartes) ;
- stratégies de suivi de trajectoire ;
- calcul numérique.

La conception de robots parallèles pour machines-outils ou la chirurgie nécéssite des travaux tels que : l'étude des postures, les modes d'assemblage, le calibrage, la propagation d'erreurs. Ces activités nécessitent l'établissement de modèles géométriques directs (MGD) et inverses (MGI).

Il est possible de résoudre le MGI de manière explicite (distances entre les points d'attache des chaînes cinématiques à la base et à la nacelle connaissant la position absolue de la base et de la nacelle). C'est par conséquent sur la résolution du modèle géométrique direct (MGD) que portent nos efforts (calcul des coordonnées absolues des points d'attache de la plateforme connaissant la position et la géométrie de la base, la géométrie de la plateforme ainsi que les distances entre les points d'attache des chaînes cinématiques à la base et à la nacelle).

Il est facile de voir que résoudre le MGD revient à résoudre un système algébrique. On sait depuis longtemps que ce problème admet un nombre fini de solutions (au plus 40). C'est donc naturellement que nous mettons à profit les travaux et solutions logicielles de calcul exact développées dans Polka pour résoudre ce problème. Les premiers résultats de résolution de MGD avec Gb et Realsolving sur des robots parallèles ont été présentés pour les robots Manta et Kanuk par Luc Rolland au cours d'un séminaire lors d'une conférence de l'ASME [21].

Nous élaborons un simulateur de commande, construit autour du logiciel généraliste MuPAD qui permet de faire du calcul exact et formel, faisant appel aux serveurs de calcul que nous développons en collaboration avec l'université de Paris VI (Gb, RS), mais également à Scilab pour la simulation du comportement numérique des méthodes « temps réel », UDX servant de protocole de communication.

Le simulateur devrait intégrer, à terme, les outils numériques d'analyse et de visualisation de robots parallèles de type Hexapode proposés par le projet INRIA Saga : VisuRobo, XWorkspace et Xjpdraw.

**Évaluation et optimisation de commandes** Une grande variété de stratégies et d'algorithmes de commande ont été implantés pour piloter un mécanisme robotisé. Dans les faits,

plusieurs études de trajectoire ont été réalisées pour obtenir des trajectoires nominales. Les résultats permettent de comparer un grand nombre de stratégies de commande et aussi d'évaluer les impacts des fonctions d'interpolations associées.

Nous avons comparé les résultats nominaux et simulés en calculant les positions réelles obtenues et les précisions de position et d'orientation.

La commande a été configurée en implantant une poursuite de trajectoire selon diverses gestions de réglage : en position, en vitesse ou les deux en cascade. Les sous-programmes pour les implanter ont été écrits et testés.

La mise en place des calculs de discrétisation en implantant diverses fonctions usuelles d'interpolation a été effectuée.

Des méthodes de diagnostic des fichiers produits par le niveau de commande de trajectoire ont été développées et implantées dans des logiciels spécifiques. Ces fichiers de commande sont destinés au pilotage des consignes des asservisements. L'étude des résultats a permis d'analyser la commande pour le compte de la société CMW et un rapport détaillé de l'impact de leur stratégie de commande incluant leur mode de synchronisation a été rédigé. Il est possible de sélectionner les stratégies, les fonctions d'interpolation et les temps de réponse de la commande et des asservissements à partir des résultats de simulation obtenus.

**Précision et sources d'erreurs.** Rattachés aux programmes d'analyse de trajectoire avec commande sélectionnée par le concepteur, des calculs de précision ont été mis en place. Des études ont permis d'obtenir des résultats intéressants appliqués au robot parallèle théorique SSM de CMW.

À la demande de la société CMW, plusieurs sources d'imprécisions ont été identifiées et cernées au niveau de la CAO, des capteurs des jambes, de la mesure de la position des articulations de la nacelle et de la commande comme telle.

Le programme d'analyse de l'impact de l'imprécision des mesures des capteurs des jambes (actionneurs des chaînes cinématiques) a été réalisé et un rapport détaillé a été fourni à la société CMW. Il en est de même pour l'impact de l'imprécision de la mesure de position des articulations de la nacelle.

**Problèmes d'usinage et diagnostic lors de mise en service.** Des méthodes de diagnostic des fichiers de code G produits par les logiciels de CAO ont été réalisés. Ceci a permis d'analyser l'état des données calculées par le fichier de CAO dans un cas concret fourni par la société CMW. Dans ce cas, on a pu déterminer que le logiciel de CAO produisait des points de consigne qui possédaient des irrégularités pouvant entraîner des erreurs importantes au niveau de la commande.

Un fichier de points de CAO optimal a été rédigé automatiquement et programmé en MuPAD satisfaisant les exigences d'erreur de corde tout en évitant le stockage en mémoire d'un trop grand nombre de points.

Il en est de même pour les résultats des logiciels hors-ligne de gestion de trajectoire en amont de l'unité de commande. En effet, ces logiciels utilisés par CMW font des calculs de pré-traitement dont celui du MGI et ont été évalués. Leurs erreurs en mode MGI ont été estimées.

**Intégration de nombreux logiciels performants.** Divers outils sont développés au sein de POLKA, au sein du projet SAGA (INRIA Sophia-Antipolis) et aussi d'autres équipes de recherche, qui serviront à l'élaboration du simulateur.

Du LORIA, les applications MFPR, UDX et RS venant de POLKA sont utilisées. De l'INRIA, on ajoute l'environnement de calcul numérique SciLab et les outils numériques d'analyse et de visualisation de robots parallèles de type Hexapode de SAGA : VisuRobo, XWorkspace et Xjpdraw.

De l'université de Paris 6, on utilise l'application FGb et de l'université de Paderborn, l'environnement de calcul formel MuPAD.

## 4.2 Bioinformatique

**Approche méthodologique.** L'analyse de génomes reste un domaine d'application privilégié pour nos recherches en algorithmique combinatoire. Les travaux que nous menons dans cette direction sont de types différents. Certains de ces travaux visent à fournir aux biologistes des outils génériques pour pouvoir analyser des séquences de génomes. Ces outils peuvent fournir de nouvelles fonctionnalités, absentes dans les méthodes ou logiciels existants, ou peuvent « simplement » être plus rapides que ces méthodes ou logiciels, et par conséquent peuvent permettre de passer à une échelle plus grande par rapport aux volumes de données traitées. Cependant, ces outils peuvent aller plus loin, en proposant de nouvaux concepts qu'ils permettent de découvrir ou de nouveaux modèles informatiques pour les phénomènes biologiques étudiés. Nos travaux sur les algorithmes d'analyse de séquences et les outils qui en résultent, `mreps` et `grappe`, sont des exemples de tels outils que nous détaillons ci-dessous.

Un autre type d'activité consiste à intervenir sur des problèmes ponctuels qui nous sont posés par des biologistes dans le cadre des collaborations auxquelles nous participons. Ici, nous essayons de construire un modèle mathématique du problème en question et de développer de nouvelles approches informatiques. Ci-dessous nous décrivons deux actions de ce type que nous avons entreprises cette année : une en collaboration avec des chercheurs de l'Institut de Génétique et de Biologique Moléculaire et Cellulaire (IGBMC) à Strasbourg, et l'autre en collaboration avec le Laboratoire d'Enzymologie et de Génie Génétique de l'Université Henri Poincaré de Nancy.

**Outils pour l'analyse de génomes.** Nos travaux sur les algorithmes et logiciels d'analyse de séquences sont décrits dans les parties correspondantes de ce rapport (voir les sections 6.3, 5.4, 5.5). Ici, nous nous concentrons sur l'aspect bioinformatique de ces travaux.

Dans les séquences d'ADN les répétitions sont souvent porteuses d'informations biologiques. Par exemple, elles peuvent témoigner d'événements d'évolution et donc porter des informations phylogéniques. Les répétitions successives (en tandem) apparaissent dans les *satellites* – éléments importants dans l'étude du *polymorphisme* de génomes, qui a des applications diverses, comme dans la compréhension des maladies génétiques par exemple. Il est donc très important de pouvoir rechercher les répétitions de façon rapide et exhaustive. Le logiciel `mreps` de recherche de répétitions (sections 6.3, 5.4) est caractérisé par la possibilité de traiter en quelques secondes des séquences de génomes entiers (des centaines de milliers ou des millions de paires de bases) tout en préservant son caractère exhaustif, c'est-à-dire sa capacité de détecter *toutes*

les répétitions successives dans la séquence. Cela est devenu possible grâce à la notion de répétitions maximales, qui, comme nous l'avons démontré, représentent d'une façon compacte toutes les répétitions en tandem et d'autre part, peuvent être trouvées rapidement. Déjà les premières expériences avec ce logiciel ont permis d'obtenir des résultats intéressants [20] : par exemple, dans le premier chromosome de la levure nous avons détecté deux très longues répétitions (un fragment de 135 nucléotides se répétant exactement plus de trois fois) apparaissant respectivement au début et à la fin du chromosome et de plus, ces deux répétitions sont complémentaires (au sens de la relation de complémentarité entre nucléotides) l'une par rapport à l'autre.

Un autre logiciel que nous développons, `grappe` (section 5.5), est lui-aussi issu de nos travaux théoriques. Il permet de rechercher rapidement des motifs dits « à trous » dans les séquences de génomes. Les motifs de ce type ont une signification particulière en génomique, car ils modélisent les sites de fixation de protéines aux séquences nucléotidiques (ADN/ARN). Ce mécanisme biologique intervient aux étapes clefs de l'expression du génome – transcription et épissage – d'où l'importance potentielle pour un biologiste de pouvoir rechercher rapidement ces motifs. La particularité de `grappe` est qu'il est capable de rechercher un grand nombre de motifs en parallèle. Il est aussi capable de rechercher des motifs avec erreurs ou des lettres alternatives, ce qui le rapproche des besoins des biologistes.

**Calcul de score pour l'alignement local de séquences.** Dans le cadre de la collaboration avec l'équipe de bioinformatique de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) à Strasbourg, nous avons travaillé sur un problème de score rencontré dans le logiciel `Ballast`. Différentes approches ont été envisagées, en particulier avec F. Plewniak et O. Poch à l'occasion de visites de G. Schaeffer dans cet institut. À la suite de ces discussions, nous avons entrepris d'adapter des méthodes d'évaluation par matrice positionelle (*Position Specific Scoring Matrix*) au problème.

**Prédiction de la localisation des sites de fixation des protéines SR.** Dans le cadre de la collaboration avec le Laboratoire d'Enzymologie et de Génie Génétique à l'université Henri Poincaré de Nancy, des travaux ont été réalisés sur la prédiction de la localisation des sites de fixation des protéines SR. Vincent Bardey, docteur en biologie, a effectué un post-doctorat de 6 mois au sein de l'équipe sur ce sujet. L'objectif de ce travail était la modélisation des sites de fixation des protéines SR en utilisant des matrices de score et les résultats de différentes études expérimentales.

Une première étape du travail a consisté à mettre en place une banque de séquences de gènes du génome humain pour lesquels le phénomène d'épissage alternatif a été observé. De plus, les introns et les exons ont été séparés et classés en plusieurs catégories. Ensuite, nous avons essayé d'obtenir des motifs qui caractériseraient les sites de fixation de protéines SR. Pour cela, nous avons choisi d'utiliser le logiciel MEME[BE95], en nous basant sur des données de séquences publiées issues d'expériences biologiques (méthode SELEX). Selon ces expériences, ces séquences sont supposées contenir un signal commun qui régit la fixation de protéines

[BE95] T. Bailey, C. Elkan, «Unsupervised learning of multiple motifs in biopolymers using expectation maximization», *Machine Learning 21*, 1995, p. 51–80.

SR. Ensuite, nous envisagions de rechercher les motifs obtenus dans la banque de données constituée à la première étape.

Cependant, nous n'avons pas réussi à obtenir, à la deuxième étape, des motifs qui seraient discriminants pour les séquences fixant les protéines SR. Cela peut être dû à plusieurs raisons. Une raison possible est que les données de séquences sont en réalité non fiables. Une autre raison, plus profonde et plus probable, serait que le processus de fixation ne peut pas être caractérisé par un motif simple au sens du logiciel MEME. Par exemple, une seule protéine SR peut avoir plusieurs signaux correspondants, ou bien un seul signal mais qui n'est pas constitué d'une seule séquence de nucléotides mais peut-être de plusieurs séquences séparées par des espaces. La difficulté réside donc dans la modélisation du problème et un progrès éventuel passe tout d'abord par une meilleure compréhension du phénomène du point de vue biologique.

En résumé, nous pensons avoir tiré de ce travail des leçons précieuses qui pourraient nous permettre de développer des idées originales et nouvelles à l'avenir. Nous envisageons donc de poursuivre ce travail en collaboration avec des biologistes.

### 4.3 Tomographie discrète

L'un des problèmes de la tomographie discrète consiste à reconstruire un ensemble fini de points à partir de rayons X, un rayon X fournissant le nombre de points présents sur un axe donné. Tous les rayons parallèles à un nombre fini de directions sont considérés.

Les applications de la tomographie discrète sont nombreuses en particulier en compression de données, en traitement d'images ou encore en imagerie médicale pour l'aide aux diagnostics médicaux en radiographie. De nouvelles motivations sont apparues plus récemment dans le domaine de l'étude des matériaux. Les techniques développées sont utilisées pour reconstruire des structures cristallines à partir de données obtenues par microscopie électronique. Ceci permet de mettre en relief des propriétés des cristaux au niveau atomique. Des collaborations dans ces directions sont à envisager.

## 5 Logiciels

### 5.1 UDX

UDX (Universal Data eXchange) est un nouveau protocole de communication binaire implanté par F. Rouillier et B. Wallrich[1] dont le procédé de base a été breveté en juillet 1999 (Dispositif d'échanges de données entre matériels informatiques, numéro de dépôt 99 08172, auteurs J.-C. Faugère - UPMC[2], F. Rouillier - INRIA).

Les solutions logicielles FGb et RS (voir module RealSolving) sont conçues pour permettre une distribution des calculs sur un réseau de machines hétérogènes. Afin de permettre l'échange efficace d'un large volume d'informations, l'échange des données doit se faire en mode binaire.

1. ingénieur de recherche au LORIA.
2. Université Pierre et Marie-Curie.

Le problème habituel à résoudre pour mettre au point ou utiliser un protocole de communication binaire est de déterminer les permutations d'octets nécessaires pour transformer le codage des données scalaires (entiers, flottants) du format d'un processeur donné en un autre format. Par exemple, les octets du codage d'un entier 32 bits sur processeur SPARC sont ordonnés à l'inverse de ceux du codage interne sur processeur PENTIUM.

Sur ces problèmes usuels viennent se greffer des problèmes spécifiques dus aux objets manipulés en calcul formel, tels que le codage des entiers multi-précision.

Tous ces paramètres rendent quasi impossible la maintenance d'un code efficace transmettant en mode binaire des entiers multi-précision avec les outils usuels (XDR, PVM, MPI, etc.).

UDX est un nouveau protocole de communication binaire qui *détecte* automatiquement et dynamiquement le codage interne des données scalaires (entiers, flottants machine) et des entiers multi-précision sans contenir une seule ligne de code dépendant de l'architecture considérée (et ne faisant évidement appel à aucune implantation dépendant de l'architecture telle que XDR par exemple). Ceci facilite sa maintenance ainsi que celle des programmes l'utilisant.

À noter qu'aucun format d'échange n'est fixé *a priori* dans UDX : le protocole le déterminera également dynamiquement. En particulier, à la différence des protocoles usuels, aucune manipulation et aucune connaissance n'est nécessaire à l'utilisateur pour optimiser la communication. Par exemple, sauf intervention de l'utilisateur, les protocoles usuels utilisent le codage interne des processeurs SPARC comme format d'échange, ce qui ralentit fortement la communication entre deux machines identiques mais dont le codage interne est différent du codage SPARC, compte-tenu des permutations d'octets effectuées. Dans ce genre de situation, UDX détecte automatiquement et dynamiquement la compatibilité des processeurs et n'effectue aucune permutation lors des échanges.

## 5.2 RealSolving et RS

La majeure partie des algorithmes développés dans POLKA et concernant les systèmes polynomiaux sont intégrés dans le logiciel RS (*Real Solutions*) déposé à l'APP (Agence de Protection des Programmes).

RS est le successeur de REALSOLVING, logiciel implanté lors du projet POSSO, maintenu pendant le projet FRISCO mais dont le développement s'est arrêté en 2000.

Les principales nouveautés sont liées aux interfaces. Nous avons mis au point un protocole ASCII commun aux logiciels FGB développé par J.-C. Faugère et RS, mais l'essentiel du travail a été le développement de l'interface *socket-stream* (binaire) qui est à l'origine du protocole UDX. Les logiciels RS et FGB disposent désormais d'interfaces homogènes, solides et performantes permettant d'une part l'exploitation intensive de la distribution des calculs sur un réseau de machines hétérogènes et d'autre part une utilisation transparente à partir de logiciels généralistes tels MUPAD.

Un travail conséquent a par ailleurs été réalisé en collaboration avec le groupe MUPAD pour l'implantation d'un programme de démonstration utilisant l'ensemble des logiciels MUPAD, RS, FGB, connectés grâce au protocole UDX, pour la résolution de quelques problèmes simples en robotique.

### 5.3 mpfr

Le développement de la bibliothèque MPFR, initié fin 1998, se poursuit. La version publique actuelle porte le numéro 1.0, et corrige quelques erreurs tout en intégrant de nombreuses nouvelles fonctionnalités. Depuis un an, MPFR est passé sous licence publique GNU, et est maintenant distribué par GNU conjointement avec la bibliothèque GMP. En parallèle, MPFR a été déposé à l'APP en mars 2000 sous le numéro 2000.000.10800, et apparaît sur la page des logiciels libres de l'INRIA [3].

La bibliothèque MPFR est une bibliothèque de calcul flottant en précision arbitraire avec arrondi exact. La version actuelle implante les quatre opérations de base $(+, -, \times, /)$, la racine carrée, l'exponentielle, le logarithme, la moyenne arithmético-géométrique, et les fonctions sinus et cosinus. Elle intègre de plus une implantation générique des séries hypergéométriques due à E. Jeandel [23], basée sur la méthode *binary splitting*. Cela permet, non seulement d'obtenir une nouvelle implantation de l'exponentielle et du logarithme, asymptotiquement plus efficace, mais d'obtenir aussi un grand nombre de fonctions de façon automatique (fonctions trigonométriques *e.g.*).

### 5.4 mreps

`mreps` est un logiciel de recherche de répétitions maximales dans un texte, basé sur l'algorithme que nous avons récemment proposé [10] (voir la section 6.3). Le résultat de cet algorithme, à savoir l'ensemble de toutes les répétitions maximales, caractérise d'une façon exhaustive toute la structure répétitive du texte. D'autre part, cet algorithme s'exécute en temps linéaire, ce qui le rend particulièrement attractif pour les applications à grande échelle, telles que l'analyse de séquences de génomes entiers par exemple.

Une première version de `mreps` a été écrite en collaboration avec Mathieu Giraud lors de son stage d'été en 1999 [4]. Des tests sur des séquences d'ADN (de l'ordre de centaines de milliers de paires de base) ont été faits et des répétitions intéressantes ont été trouvées. Les résultats de ce travail ont été présentés à la conférence JOBIM [20], première conférence française de bioinformatique, et, sous forme d'un poster, à la *Second International Conference on Bioinformatics of Genome Regulation and Structure (BGRS'2000)* qui a eu lieu à Novissibirsk (Russie) en août 2000.

Depuis, nous poursuivons le développement de ce logiciel. Cette année, dans le cadre d'un stage de Mariana-Liliana Ibanescu, de nouvelles fonctionnalités y ont été ajoutées, et en particulier la possibilité de traiter des fichiers en format Fasta – le format standard de représentation de séquences nucléotidiques. Nous sommes en train de mettre en place une interface Web pour pouvoir faire tourner `mreps` via Internet [5]. Dans le cadre du stage de licence de Vianney Miard, nous avons commencé à développer une nouvelle interface Java pour visualiser les répétitions trouvées par `mreps`, travail que nous envisageons de poursuivre prochainement. Nous envisageons d'autres extensions de `mreps` ; en particulier, nous projetons d'intégrer dans `mreps` l'algorithme de recherche de répétitions à espace fixe, proposé dans [19]. Nous avons également

3. `http://www.inria.fr/valorisation/logiciels/index.fr.html`
4. `http://www.loria.fr/projets/REMAG/mreps.10.tar`
5. `http://www.loria.fr/\string~kucherov/SOFTWARE/MREPS/`

entrepris des travaux visant à développer des méthodes de recherche de *répétitions approchées*, c'est-à-dire des répétitions successives d'un facteur admettant un certain taux de différence entre les copies répétées. Pouvoir trouver ces répétitions aurait des applications importantes à l'analyse de génomes (voir la section 4.2).

### 5.5 grappe

`grappe` est un logiciel qui recherche dans un texte plusieurs motifs simultanément, chacun étant composé d'une suite de fragments (mots) séparés par des espaces de longueur *a priori* non-bornée. La naissance de ce logiciel, il y a quelques années, est due à la découverte d'un nouvel algorithme efficace pour rechercher simultanément un ensemble de motifs du type $v_1 * v_2 \ldots * v_m$ où les $v_i$ sont des mots de l'alphabet et $*$ est un symbole spécial (espace non borné) qui peut être substitué par n'importe quelle sous-chaîne[KR97]. L'algorithme de `grappe` est basé sur la structure de données appelée DAWG (*Directed Acyclic Word Graph*) qui peut être vue comme un automate fini qui change dynamiquement au cours de son travail.

Le point fort de `grappe`, par comparaison avec les logiciels de cette famille tels que `egrep` ou `agrep`, est qu'il permet de rechercher rapidement un grand nombre de ces motifs dans de longs textes. En 1999, nous avons repris le développement de `grappe` en le généralisant d'abord aux espaces de longueur *bornée*, c'est-à-dire spécifiée par un intervalle de valeurs. Cette année, grâce au stage de Sébastien Briais, une nouvelle version (version 3) de `grappe` a vu le jour[6]. Cette version, beaucoup plus stable par rapport à la version précedente, inclut aussi de nouvelles options, telles que la possibilité de traiter le fichier texte par enregistrements (*records*), paramétrer le caractère de fin d'enregistrement, compter le nombre d'enregistrements contenant un motif recherché, etc. De plus, une procédure d'installation standard à l'aide de `configure` a été mise en place. Elle prévoit deux versions de `grappe` – une version générique et une version spécialisée pour traiter des séquences d'ADN. Enfin, une page Web du logiciel a également été mise en place[7].

`grappe` a été déposé à l'APP ; il est également diffusé à partir de la page des logiciels développés à l'INRIA[8].

## 6 Résultats nouveaux

### 6.1 Algèbre commutative et résolution de systèmes polynomiaux

**Participants** : Luc Rolland, Fabrice Rouillier, Paul Zimmermann.

John Abbott (Univ. de Gênes, Italie), Victor Shoup (IBM Zurich, Suisse) et P. Zimmermann ont proposé un nouvel algorithme pour l'étape de recombinaison lors de la factorisation d'un polynôme de $\mathbf{Z}[x]$ [11]. Il s'agit ici de trouver quels produits de facteurs $f_i \bmod p^k$, $1 \leq i \leq n$,

6. `http://www.loria.fr/\string~kucherov/SOFTWARE/grappe-3.0/grappe-3.0.tar.gz`
7. `http://www.loria.fr/\string~kucherov/SOFTWARE/grappe-3.0/grappe-3.0-en.html`
8. `http://www.inria.fr/valorisation/logiciels/index.fr.html`

[KR97] G. Kucherov, M. Rusinowitch, «Matching a Set of Strings with Variable Length Don't Cares», *Theoretical Computer Science 178*, 1997, p. 129–154.

vont donner de « vrais » facteurs de $\mathbf{Z}[x]$ ; l'algorithme de Zassenhaus utilisé jusqu'alors a une complexité de $O(2^n)$. Le nouvel algorithme proposé est grosso modo en $O(2^{n/2})$, grâce à l'utilisation de tables stockées en mémoire. Cet algorithme est implanté dans la bibliothèque NTL. Il est à noter qu'en juillet, Mark van Hoeij a proposé une nouvelle méthode utilisant l'algorithme LLL de réduction des réseaux, de meilleure complexité asymptotique[vH].

P. Aubry (LIP6), F. Rouillier et M. Safey (LIP6) proposent une nouvelle génération d'algorithmes permettant de traiter le cas général des systèmes de dimension positive (pour calculer au moins un point par composante semi-algébriquement connexe).

L'idée de base est de regarder les points critiques de la fonction *distance à un point* qui définissent un point au moins sur chaque composante connexe. Le but est de produire un ensemble fini de systèmes zéro-dimensionnels (admettant un nombre fini de solutions) dont les zéros définissent au moins un point par composante (semi-algébriquement) connexe de la variété.

Nous avons montré que, pour une formulation bien choisie, les sous-variétés de la variété initiale dont l'étude nécessite traditionnellement une déformation infinitésimale du système (par exemple les ensembles de points singuliers) pour se ramener à la résolution d'un système zéro-dimensionnel, sont des sous-variétés de la variété initiale de dimension strictement inférieure.

Comme nous pouvons donner explicitement les systèmes d'équations polynomiales définissant ces cas pathologiques, ceci donne une méthode récursive naturelle n'utilisant aucune déformation infinitésimale.

Les choix arbitraires (changements *génériques* de coordonnées) rendant les algorithmes probabilistes ont tous été bannis (et remplacés par des choix dans une famille finie de solutions, dont l'une au moins est bonne, accompagnés d'une stratégie de vérification a posteriori). L'utilisation d'arithmétiques lourdes a été évitée (pas de déformations infinitésimales - tous les calculs s'effectuent dans le corps de base) et aucune astuce à effet pratique désastreux (élévations au carré, changements de coordonnées) n'est utilisée, quitte à éventuellement sacrifier les bornes supérieures de complexité théorique (ce qui d'ailleurs n'est pas prouvé).

Ce résultat (article soumis) est illustré par un prototype dont l'implantation, pour l'instant rudimentaire, offre un niveau de performances incontestablement bon puisqu'elle inverse la tendance vis-à-vis de la CAD (meilleur algorithme du moment capable de faire ce type de calcul en pratique) : pour notre base d'exemples, mise à part de rares exceptions, notre méthode ne prend que quelques secondes alors que le calcul de la CAD est impossible. La table suivante donne une série de temps de calcul pour 3 versions de notre méthode (algorithmes 2, 3 et 4) obtenus avec des implantations très primaires ainsi que la version la plus aboutie de la CAD (fonction QEPCAD du logiciel SACLIB écrit en C). À noter toutefois que l'algorithme 4 ne permet que de décider si une variété admet ou non des points réels mais cette méthode est en

[vH] M. van Hoeij, «Factoring polynomials and the knapsack problem», prépublication, 29 août 2000, `http://www.math.fsu.edu/~hoeij/papers.html`.

cours de généralisation.

| **Système** | **Algorithme 2** | **Algorithme 3** | **Algorithme 4** | **CAD** |
|---|---|---|---|---|
| Vermeer | 62.36 | 3.32 | <0.01 | 43 |
| Wang | 1.37 | 1.37 | 0.13 | $\infty$ |
| Euler | 0.01 | 0.01 | <0.01 | failed(*872043*) |
| Neural | 1.02 | 1.02 | 0.44 | 0.9 |
| Butcher | 1.7 | 1.7 | 1.7 | $\infty$ |
| Buchberger | < 0.01 | <0.01 | <0.01 | failed(*991324*) |
| DiscPb | 0.2 | 0.2 | 0.02 | $\infty$ |
| Donati | 11609 | 10 | 0.04 | 0.6 |
| Hairer2 | $\infty$ | $\infty$ | 23.03 | failed(*872043*) |
| Prodecco | $\infty$ | $\infty$ | 286 | $\infty$ |
| F633 | $\infty$ | $\infty$ | 5700 | $\infty$ |
| F744 | $\infty$ | $\infty$ | 40 | $\infty$ |
| F855 | $\infty$ | $\infty$ | 5664 | $\infty$ |

où (failed($n$) signifie que la méthode à échoué après avoir évalué le nombre minimal de cellules qu'il faudrait étudier ($n$)).

Une chose très encourageante est la marge de progrès dont nous disposons. En effet, l'injection de méthodes de calcul très récentes (en particulier la décomposition en idéaux premiers mise au point par J.-C. Faugère) permet d'atteindre un excellent niveau de performances. Par exemple, le temps de calcul pour l'exemple `F855` descend à 26 secondes !

Enfin, il est à noter que ce premier prototype permet déjà d'appréhender des problèmes à caractère industriel (les exemples `FXXX` sont issus de problèmes d'optimisation de bancs de filtres).

## 6.2 Arithmétique

**Participants** : Guillaume Hanrot, Emmanuel Jeandel, Vincent Lefèvre, Jean-Luc Rémy, Paul Zimmermann.

Les travaux accomplis sur les équations diophantiennes sont dans le prolongement des idées développées par G. Hanrot dans sa thèse[Han97] et ses travaux antérieurs. En particulier, l'article écrit par ce dernier en collaboration avec Yuri Bilu et Paul Voutier [3], et mettant un point final à une conjecture célèbre datant du siècle dernier portant sur les diviseurs de nombres de la forme $\alpha^n - \beta^n$, a été accepté pour publication ; ce résultat repose sur l'utilisation d'idées techniques nouvelles pour mener à bien les calculs intensifs nécessaires. On peut en particulier signaler le travail [9] paru récemment dans *Mathematics of Computation*.

Des travaux communs à Yann Bugeaud (Université de Strasbourg) et G. Hanrot tirent profit d'un aspect initialement marginal d'un article écrit en commun avec Yuri Bilu, qui dans certaines circonstances permet d'obtenir des résultats frappants de façon « élémentaire ». Cette méthode permet, en particulier, d'améliorer les résultats connus sur l'équation de Catalan $x^p - y^q = 1$, ce qui a donné lieu à la rédaction d'un travail [5] à paraître.

[Han97] G. Hanrot, *Résolution effective d'équations diophantiennes : algorithmes et applications*, Thèse de doctorat, Université de Bordeaux I, avril 1997.

Par ailleurs, un travail de G. Hanrot, en collaboration avec N. Saradha et T. Shorey, du Tata Institute à Bombay et utilisant lourdement les techniques de résolution effective d'équations diophantiennes a permis de progresser dans l'étude d'une conjecture d'Erdős et Selfridge, qui cherche à déterminer quand un produit d'entiers consécutifs, duquel on peut omettre au plus un terme, est presque une puissance pure. Ce travail [8] paraîtra prochainement aux *Acta Arithmetica*.

En ce qui concerne l'arithmétique des ordinateurs, G. Hanrot et P. Zimmermann ont décrit – avec M. Quercia – un algorithme permettant de calculer directement la partie médiane du produit d'un nombre de $2n$ mots par un nombre de $n$ mots. Cela permet d'accélérer les calculs dans la division et la racine carrée lorsque ces dernières sont évaluées par la méthode de Newton. Une publication reprenant ces résultats a été soumise à *AAECC* [22].

L'article détaillant la factorisation de l'entier « RSA-155 », première clé RSA de 512 bits ayant été factorisée (août 1999) est paru dans les actes de la conférence Eurocrypt'2000 [15] ; P. Zimmermann avait participé à la phase de crible de cette factorisation. À noter qu'une seconde clé RSA de 512 bits a été factorisée en octobre 2000, à savoir celle contenue dans le $10^{\text{e}}$ problème du livre *The Code Book* de Simon Singh[9] ; cette fois-ci la seconde phase, qui consiste essentiellement en de l'algèbre linéaire sur d'énormes matrices creuses sur $\mathbf{F}_2$, a été faite sur un quadri-processeur Compaq, et non plus sur un Cray comme pour RSA-155. Il n'est donc plus nécessaire de disposer de super-calculateurs pour casser de telles clés.

Dans le cadre de l'action de recherche coopérative « Arithmétique des Ordinateurs Certifiée» (AOC), P. Zimmermann a effectué une preuve de ses implantations de division et racine carrée rapides dans la bibliothèque GNU MP [25]. Ces implantations seront intégrées dans la prochaine version de GMP. En collaboration avec Yves Bertot, Laurence Rideau et Loïc Pottier du projet LEMME, une preuve formelle de l'algorithme de racine carrée a été réalisée à l'aide du logiciel Coq et de l'interface Pcoq.

**Étude d'une autre suite autoconstruite.** À l'occasion d'un stage d'initiation à la recherche, pour des élèves d'ESIAL, nous avons mis en évidence, d'abord expérimentalement puis théoriquement, le comportement asymptotique d'une nouvelle suite autoconstruite. Cette suite, notée $G_2(n)$, satisfait la relation de récurrence : $|G_2^{-1}(n)| = G_2^2(n)$. où la notation $F^{-1}(n)$ désigne l'image inverse de n par une fonction F. L'objectif était de déterminer si $G_2(n)$ était ou non asymptotiquement équivalente à la fonction $2^{1/4}n^{1/2}$. Les calculs ont mis en évidence que, tout en étant bien de l'ordre de grandeur de cette fonction, $G_2$ ne lui était pas équivalente, mais oscillait autour d'elle. En nous servant des valeurs calculées pour certaines grandes valeurs de $n$, ainsi que d'un argument théorique semblable à celui utilisé il y a quelques années pour la fonction de Golomb, nous avons obtenu une preuve formelle de ce résultat. Cette preuve reste à rédiger.

## 6.3 Algorithmique combinatoire

**Participants** : Vincent Bardey, Sébastien Briais, Isabelle Debled-Rennesson, Liliana-Mariana Ibanescu, Roman Kolpakov, Grégory Kucherov, Mostefa Mesmoudi,

9. Cf `http://www.simonsingh.com/cipher.htm`

Jean-Luc Rémy, Jocelyne Rouyer, Gilles Schaeffer.

**Répétitions dans les mots.** Un article décrivant l'ensemble des travaux sur les répétitions maximales, que nous avons effectués en 1998-1999, a été accepté pour publication dans la revue *Journal of Discrete Algorithms* et doit y paraître cette année [10].

Rappelons que dans ces travaux nous avons étudié les répétitions dites *maximales* dans les mots. Le résultat principal que nous avons obtenu affirme que le nombre de répétitions maximales dans les mots de longueur $n$ est borné par une fonction linéaire en $n$. Qui plus est, la somme des exposants de toutes les répétitions maximales dans un mot de longueur $n$ est elle aussi bornée linéairement en $n$. Ces résultats combinatoires ont eu, en particulier, des conséquences algorithmiques importantes ; ils ont permis de démontrer que toutes les répétitions maximales peuvent être trouvées en un temps linéaire en la longueur du mot. L'algorithme résultant a été implanté dans un logiciel appelé `mreps`, dont les derniers développements sont décrits dans la section 5.4.

L'étude des répétitions dans les mots a été poursuivie cette année ; nous avons proposé un nouvel algorithme efficace recherchant dans un mot toutes les *répétitions à espace fixe*. Plus précisément, il s'agit de toutes les paires d'occurrences d'un même facteur, séparées par une distance constante $d$ spécifiée par avance. Notre algorithme permet de rechercher toutes ces répétitions en temps $O(n \log d + S)$, où $S$ est leur nombre dans le mot et $n$ est la longueur du mot. Cet algorithme améliore la borne $O(n \log n + S)$ obtenue dans un article récent[BLPS99] pour un problème plus général, dans lequel l'espace entre les deux copies du facteur répété est borné par un intervalle de valeurs qui dépendent éventuellement de la taille de ce facteur. Il est intéressant de noter que notre algorithme est basé sur les mêmes techniques que celles que nous avons utilisées dans la recherche de répétitions maximales [10], à savoir sur les fonctions de Main et Lorentz et sur une factorisation spéciale de mot, appelée la $s$-factorisation. Cela montre que ces techniques sont puissantes car elles s'appliquent à toute une série de problèmes de recherche de répétitions. D'autre part, ces techniques sont simples à implanter et par conséquent notre algorithme a un avantage de simplicité de réalisation par rapport à celui mentionné plus haut, ce dernier étant basé sur des techniques algorithmiques complexes et à coût d'implantation élevé. Nous envisageons donc d'intégrer prochainement notre algorithme dans le logiciel `mreps` (section 5.4). Enfin, notre algorithme permet également de rechercher au même coût les paires d'occurrences d'un même facteur, séparées par un mot donné (au lieu d'une distance donnée).

Une publication [7] décrivant des travaux antérieurs est parue cette année dans la revue *Algorithmica*.

**Aléa discret (i) Marches dans le plan.** L'an dernier, de surprenants résultats d'algébricité avaient été obtenus avec Mireille Bousquet-Mélou du LaBRI (Bordeaux) pour les séries génératrices de marches aléatoires dans le plan qui évitent une barrière. Un résumé de ces travaux a été publié dans les actes du colloque *Mathematics and Computer Science* [14] à Versailles. (Notons que M. Bousquet-Mélou y a aussi consacré deux exposés invités, au colloque

[BLPS99] G. Brodal, R. Lyngsø, C. Pedersen, J. Stoye, « Finding maximal pairs with bounded gap », *in : Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, M. Crochemore, M. Paterson (éditeurs), *Lecture Notes in Computer Science, 1645*, Springer-Verlag, 1999.

*Classical Combinatorics* à Philadelphie et aux rencontres *Self interacting random processes* à Oberwolfach.)

Nos résultats concernent un modèle de barrière assez spécifique (une demi-droite horizontale), mais ont suscité un intérêt particulier, d'une part au sein de la communauté de combinatoire énumérative à cause des formules particulièrement simples obtenues, et d'autre part au sein de la communauté probabiliste car en général les marches contraintes donnent seulement lieu à des études asymptotiques et il y a peu de cas connus non triviaux d'algébricité des séries génératrices. Ceci nous a amenés à rechercher le modèle le plus général sous lequel nos résultats d'algébricité tiennent. C'est dans ce cadre que nous avons placé la version étendue de notre article.

**Aléa discret (ii) Combinatoire algébrique.** La collaboration avec Alain Goupil du LaCIM (Université du Québec à Montréal) et Dominique Poulalhon du LIX (École Polytechnique) a bénéficié du séjour d'A. Goupil à Nancy comme professeur invité au mois de juillet 2000. Nos preuves des conjectures de Jacob Katriel ont fait l'objet d'une communication à la conférence internationale FPSAC'00 à Moscou, dont les actes ont été publiés par Springer [17]. Ces résultats portent sur la décomposition de caractères du groupe symétrique en termes de certaines fonctions puissances, appelées puissances des contenus. Cette année, nous avons pu aborder l'extension de ces travaux à une famille de polynômes, les polynômes de Jack, qui, en temps que spécialisations des polynômes de Macdonald, sont actuellement au centre de nombreux travaux en combinatoire algébrique. Dans le même temps, à l'aide d'expériences de calcul formel, nous avons construit quelques premières conjectures sur le relèvement de nos décompositions dans l'algèbre de Hecke du groupe symétrique. Enfin nous continuons notre travail sur le problème de l'énumération et de la construction des solutions d'équations dans le groupe symétrique ; des résultats asymptotiques ont été obtenus qui se sont naturellement intégrés au travail [PS]. Des contacts se développent autour de ce thème avec Igor Pak au MIT.

**Aléa discret (iii) Entrelacs alternants.** Au sein de la théorie des nœuds, qui traite de la classification des configurations de lacets fermés dans l'espace, le premier résultat d'énumération significatif a été obtenu très récemment par Sundberg et Thisttlewaithe[ST98], pour la classe des entrelacs alternants. Ce travail repose sur la traduction du problème topologique en un modèle combinatoire sur carte planaire et réutilise intelligemment des résultats anciens de cette théorie. Cependant les auteurs ne parviennent qu'à un encadrement asymptotique et laissent ouvert le problème du comportement asymptotique exact du nombre d'entrelacs alternants. Dans le cadre du DEA « Algorithmique », G. Schaeffer a dirigé au Laboratoire d'Informatique de l'École Polytechnique (LIX) le stage de Sébastien Kunz-Jacques, élève à l'ENS Paris, sur ce sujet. L'analyse de schémas de composition de [2, 12], combinée à une compréhension plus fine du modèle combinatoire, nous a permis de donner une solution complète à ce problème. Un résumé étendu de ce travail [24] est soumis au colloque international FPSAC

[PS] D. Poulalhon, G. Schaeffer, « Factorizations in the symmetric group », 20pp, en cours de révision pour *Discrete maths.*

[ST98] C. Sundberg, M. Thistlethwaite, « The rate of growth of the number of prime alternating links and tangles », *Pacific J. Math. 182*, 2, 1998, p. 329–358.

2001 (Arizona State University). Enfin le problème de la génération aléatoire des structures associées a été étudié à partir des résultats de [2] et un algorithme proposé et implanté.

**Aléa discret (iv) Topologie des cartes aléatoires.** Les cartes aléatoires sous-jacentes aux modèles combinatoires sur cartes, sont l'objet d'études approfondies pour elles-mêmes. Dans cet ordre d'idées, grâce à des simulations avec générateur aléatoire (suivant [2]), une conjecture selon laquelle le diamètre d'une carte de taille $n$ se comporte comme $\Theta(n^{1/4})$ avait été formulée dans la thèse de G. Schaeffer[Sch98]. Des discussions autour de ce thème, en particulier avec Gil Kalaï (Hebrew University of Jerusalem), suggèrent d'étudier plus généralement la croissance du nombre de sommets dans une boule de rayon croissant autour d'un point, en liaison avec le temps de diffusion d'une marche dans une carte aléatoire (questions considérées par des physiciens statisticiens). En collaboration avec Philippe Chassaing, probabiliste de l'IECN à Nancy, nous avons conjecturé l'existence d'un modèle continu du problème, en nous appuyant sur un codage combinatoire en terme d'arbres étiquetés. Ce modèle, très prometteur, s'inscrit dans la théorie actuellement florissante en probabilité des superprocessus.

L'article [6], écrit avec Robert Cori du LaBRI (Bordeaux) et décrivant des travaux antérieurs sur les arbres de description et leur lien avec les codages de cartes planaires a été accepté pour publication dans la revue *Theoretical Computer Science*. Enfin l'article [4] décrivant des travaux avec Mireille Bousquet-Mélou sur l'unification, l'extension et la preuve combinatoire des formules de Tutte et d'Hurwitz pour les constellations est paru dans la revue *Adv. in Applied Math.*

**Aléa discret (v) Coalescence de cols.** La collaboration entamée avec l'équipe de Philippe Flajolet sur l'utilisation de méthodes de cols en présence de coalescence s'est amplifiée tout au long de l'année. Un résumé étendu de nos résultats de l'an dernier avec Cyril Banderier, Philippe Flajolet et Michèle Soria sur les composantes de cartes planaires est paru dans les actes de la conférence ICALP'00 [12].

Cette année, nous avons pu mener à son terme un travail portant sur l'énumération des graphes généraux connexes. Les méthodes de coalescence de cols permettent en effet de retrouver naturellement et de préciser à volonté l'asymptotique des résultats difficiles de Wright sur ce problème. Cette approche avait été déjà envisagée il y a quelques années, mais un argument de convergence manquait alors, argument que nous avons finalement obtenu. Un article avec Philippe Flajolet et Bruno Salvy est en cours de rédaction.

**Algorithmes de génération aléatoire.** Une version de test des algorithmes de génération aléatoire de cartes et graphes planaires issus de [2] est maintenant opérationnelle[10]. Cette première version du programme `genmap` traite pour l'instant seulement une petite partie des familles décrites dans [2], mais a déjà permis de faire quelques expériences probantes. Du point de vue théorique deux nouvelles familles d'objets ont été traitées cette année : les entrelacs avec

10. cf : `http://www.loria.fr/\string~schaeffe/genmap.html`

[Sch98] G. Schaeffer, *Conjugaison d'arbres et cartes combinatoires aléatoires*, thèse de doctorat, Université Bordeaux I, 1998.

S. Kunz-Jacques (voir plus haut), les constellations planaires avec Michel Bousquet et Cédric Chauve. Ce dernier travail a fait l'objet d'un poster à la conférence LaCIM2000 [13].

**Géométrie discrète (i) Détection de la convexité discrète.** Nous avons étudié le problème de la convexité d'un polyomino hv-convexe. Celle-ci se ramène à celle de son bord. Ce bord, qui est une courbe discrète, peut être décomposé en plusieurs segments de droite discrets. Nous avons observé que le bord n'était pas convexe dans le cas particulier où l'analyse de l'un de ses segments s'interrompt sur ce qu'on appelle un point fortement extérieur (par rapport au polyomino). Nous avons alors montré que ce cas particulier était en réalité le seul cas de non-convexité. Nous avons écrit un algorithme qui va être présenté en décembre à la conférence DGCI, à Uppsala [16]. Nous avons montré depuis que cet algorithme était de complexité linéaire en temps.
Un test de convexité discrète peut être ajouté à une méthode de reconstruction de polyominos hv-convexes pour obtenir une méthode de reconstruction de polyominos convexes. Des tests sont en cours pour comparer une méthode utilisant l'algorithme de Chrobak et Dürr [CD99], qui est le meilleur — en temps de calcul — algorithme de reconstruction de polyominos hv-convexes actuellement connu, et une méthode utilisant la programmation logique avec contraintes. L'intérêt de la deuxième méthode est qu'on n'a pas besoin d'avoir reconstruit les polyominos hv-convexes pour tester s'ils sont convexes, les contraintes traduisant la convexité d'un polyomino étant traitées globalement.

**Géométrie discrète (ii) Reconnaissance de morceaux de plan discrets.** Le problème de la reconnaissance consiste à déterminer si un sous-ensemble 26-connexe, borné de $\mathbf{Z}^3$ est ou non un morceau de plan discret naïf. Un plan discret naïf est défini par les quatre entiers $\mu$, a, b et c où (a,b,c) est le vecteur normal du plan dont les points $(x, y, z)$ de $\mathbf{Z}^3$ vérifient $\mu \leq ax + by + cz < \mu + max(|a|, |b|, |c|)$. L'algorithme de reconnaissance existant [DR96] est incrémental et reconnaît des morceaux de plans rectangulaires : les sections parallèles à un plan de coordonnées du morceau à reconnaître sont parcourues en ajoutant les voxels un à un et en décidant par des constructions géométriques simples si l'ensemble déjà parcouru plus le point ajouté est ou non un morceau de plan discret. Un travail a été réalisé en collaboration avec Mostefa Mesmoudi, chercheur invité dans l'équipe, sur la démonstration de l'algorithme de reconnaissance cité précédemment et un algorithme plus performant, de reconnaissance de morceaux de plans de formes quelconques est à l'étude. La rédaction de ces travaux est en cours d'achèvement.

[CD99] M. Chrobak, C. Dürr, «Reconstructing hv-Convex Polyominoes from Orthogonal Projections», *Information Processing Letters*, 1999.

[DR96] I. Debled-Rennesson, J.-P. Reveillès, «Incremental algorithm for recognizing pieces of digital planes», *in : Spie's International Symposium on Optical Science, Engineering, and Instrumentation, Technical Conference, Vision Geometry 5, Denver, USA*, 1996.

# 7 Contrats industriels (nationaux, européens et internationaux)

Notre contrat avec CMW (Constructions Mécaniques des Vosges Marioni) a pris fin en 1999 ; un renouvellement de ce contrat est en cours de négociation.

Des discussions sont en cours avec un nouveau partenaire institutionnel : les Hopitaux de Paris. Il s'agit d'aider des concepteurs de robots appliqués dans le domaine médical à analyser le comportement géométrique et évaluer les performances de ces manipulateurs basés sur des structures parallèles tripodes.

Le contrat pour la conception du lien MuPAD/Scilab va être complété : une version simplifiée du protocole UDX sera utilisée pour les communications entre ces deux logiciels.

# 8 Actions régionales, nationales et internationales

## 8.1 Actions nationales

Polka participe à l'action de recherche coopérative REMAG (Recherche et Extraction de Motifs pour l'Analyse Génomique) qui a démarré en 1999 (coordinateur G. Kucherov). Remag regroupe cinq sites (projets INRIA `{\scPolka}`, `{\scAlgo}` et `{\scAïda}`, Université Marne-la-Vallée et Institut Pasteur) ; son but est de développer de nouveaux algorithmes d'analyse de séquences génomiques[11]. Le thème central des travaux dans le cadre de l'action est la recherche et l'extraction de motifs « à trous », très importants dans la modélisation de certains signaux biologiques, en particulier les promoteurs dans les séquences d'ADN. Cette année, une réunion de l'action s'est tenue en février 2000 (avec participation de I. Debled-Rennesson, G. Kucherov et G. Schaeffer) et une réunion de clôture est prévue pour la fin de l'année 2000.

Polka participe également à l'action de recherche coopérative AOC (Arithmétique des Ordinateurs Certifiée), avec les projets Arénaire (INRIA Rhône-Alpes) et Lemme (INRIA Sophia-Antipolis). Dans le cadre de cette action, une première réunion a été organisée à Sophia-Antipolis en février, une autre à Nancy en mai, et un workshop de deux jours à Lyon en novembre. Dans le cadre de cette action, P. Zimmermann a effectué un séjour d'une semaine au sein du projet Lemme début novembre.

Sur le plan régional, nous participons au *Génopôle Strasbourg-Alsace-Lorraine* (travaux sur l'alignement de séquences de protéines à l'aide du logiciel `Ballast`, voir la section 4.2) ainsi qu'à l'axe « Bioinformatique » du Pôle Régional Scientifique et Technologique « Intelligence Logicielle » (travaux sur les sites de fixation des protéines SR, voir la section 4.2).

Nous faisons partie du projet « Approches multicritères pour la modélisation et l'analyse in silico des génomes », qui s'est constitué en réponse à l'appel d'offres commun CNRS - INRA - INRIA - INSERM et implique 9 laboratoires français d'informatique et de biologie.

L'ACI Blanche proposée par J.C. Faugère et F. Rouillier (algorithmes efficaces pour l'étude des zéros réels des systèmes paramètrés d'égalités et d'inégalités) à été acceptée.

G. Hanrot participe au projet « Cryptocourbes » (utilisation des courbes algébriques et de leurs jacobiennes en cryptographie) mené par F. Morain (LIX, École polytechnique) et accepté dans le cadre de l'ACI Cryptologie.

11. cf `http://www.loria.fr/projets/REMAG/`

G. Hanrot a participé au colloque « Algorithmes en théorie des nombres » au CIRM à Marseille, où il a donné un exposé. Il a également donné un exposé au séminaire de l'équipe de théorie des nombres de l'université de Saint-Étienne, un exposé au séminaire d'informatique fondamentale du LORIA, ainsi qu'un exposé dans le cadre des rencontres LoriaTech. Il a enfin été orateur invité au colloque : « Mathématiques : calculs et formes», qui s'est tenu à l'université Toulouse 2.

G. Kucherov a fait des exposés au séminaire du LIP (Laboratoire de l'Informatique du Parallélisme) à l'École Normale Supérieure de Lyon et au LITA (Laboratoire d'Informatique Théoriques et ses Applications) de l'Université de Metz.

G. Kucherov et G. Schaeffer ont assisté au colloque *Traitement et Analyse des Séquences* à Évry ; G. Kucherov y a fait un exposé.

G. Schaeffer a séjourné à deux reprises (5 jours au total) dans l'équipe de bioinformatique de l'IGBMC à Strasbourg dans le cadre du génopôle. Il participe à l'action de recherche coopérative ALCOPHYS, coordonnée par le projet ALGO. Il a été invité à donner des exposés au séminaire de probabilité de l'IECN et au séminaire du projet ALGO à l'INRIA Rocquencourt ; il a également été invité à présenter ses travaux dans une session « Probabilités et Algorithmes», aux journées nationales MAS de la SMAI, qui réunissent tous les deux ans la communauté française de probabilité et statistique (environ 170 participants cette année). Il a enfin participé en mars 2000 aux rencontres du GDR CNRS ALÉA au CIRM à Marseille, où il a donné un exposé intitulé « Chemins dans le plan coupé».

I. Debled-Rennesson a donné un exposé au séminaire d'informatique fondamentale du LIAFA en février 2000. En mars, elle a fait une intervention dans le cadre du groupe de travail en géométrie discrète qui se réunit à Paris chaque mois. En septembre, elle a fait un exposé aux journées de SPECIF.

I. Debled-Rennesson, M. Giraud, G. Kucherov et G. Schaeffer ont participé à la conférence JOBIM'00 (*Journées Ouvertes : Biologie, Informatique et Mathématiques*), première conférence française de bioinformatique d'audience internationale. M. Giraud et G. Kucherov y ont présenté leur travail.

P. Zimmermann a été invité au séminaire du projet Algorithmes (janvier), et au séminaire d'algorithmique de l'université de Caen (décembre).

## 8.2 Actions internationales

Une proposition de coopération scientifique et technologique franco-australienne autour du thème « cryptographie et calcul formel : expérimentation et implantation» implique, outre des chercheurs du LIX (École polytechnique) et du GRIMM (Université Toulouse 2), deux membres de l'équipe (G. Hanrot et P. Zimmermann). Cette action a pour but de développer des outils pour l'utilisation des courbes elliptiques et des jacobiennes de courbes algébriques en cryptographie.

R. Kolpakov et G. Kucherov ont assisté au *7th International Symposium on String Processing and Information Retrieval (SPIRE)* à La Coruña (Espagne) en septembre 2000 pour y faire un exposé sur leur travaux.

En février 2000, G. Kucherov a rendu visite à la société CERES Inc. à Malibu (États-Unis) où il a fait un exposé sur les travaux de POLKA en bioinformatique et a discuté de la possibilité

de collaboration.

Les travaux de G. Schaeffer (dont certains en collaboration avec d'autres chercheurs) ont donné lieu cette année à des exposés aux conférences et séminaires internationaux suivants : *27th International Colloquium on Automata, Languages, and Programming (ICALP'00)* à Genève (avec C. Banderier, P. Flajolet et M. Soria), *12th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC'00)* à Moscou (avec A. Goupil et D. Poulalhon), *6th Seminar on the Analysis of Algorithms* à Gdansk (contribution invitée), *44*$^{e}$ *Séminaire Lotharingien de Combinatoire* à Otrott, *Colloquium Mathematics and Computer Science* à Versailles (avec M. Bousquet-Mélou).

L. Rolland a présenté ses travaux sur les robots parallèlles Manta et Kanuk au congrès de génie mécanique de l'ASME intitulé *IMECE*.

I. Debled-Rennesson, J.-L. Rémy et J. Rouyer ont participé au 9$^{e}$ colloque DGCI, à Uppsala (Suède) en décembre 2000, où ils ont présenté leurs travaux sur la convexité discrète.

P. Zimmermann a participé à la conférence ISSAC'00 à St-Andrews (Écosse) en août ; il y a présenté ses travaux avec J. Abbott et Victor Shoup.

P. Zimmermann a donné en avril un exposé invité sur la bibliothèque MPFR à la conférence *Real Numbers and Computers'4* (RNC'4) à Dagstuhl (Allemagne).

### 8.2.1 Europe

Des membres de POLKA participent au projet INTAS *Methods, algorithms and software for functional and structural annotation of complete genomes* (projet avec la Russie, incluant des laboratoires d'Allemagne, de France et d'Autriche).

P. Zimmermann et Rob Harley (projet Cristal) ont rendu visite au principal développeur de la bibliothèque GNU MP, Torbjörn Granlund, à Stockholm en janvier ; au cours de cette semaine, ils ont intégré leurs implantations efficaces d'algorithmes dans GMP. Ces implantations sont disponibles dans la version 3.1 de GMP diffusée depuis juillet.

### 8.2.2 Russie et Asie Centrale

Le projet « Algorithmique et combinatoire des structures discrètes » de l'Institut franco-russe A. M. Liapounov d'informatique et de mathématiques appliquées, dont G. Kucherov était responsable du côté français, est arrivé à son terme cette année après trois ans d'existence. Des actes des séminaires franco-russes tenus dans le cadre de ce projet (trois séminaires au total : en juin 1997 et en avril 1998 à Moscou et en février 1999 à Nancy) ont été préparés pour publication à l'Université de Moscou et devraient paraître prochainement. En outre, un rapport d'activités du projet pour les trois ans est également à paraître dans le recueil des rapports d'activités des projets de l'Institut Liapounov. R. Kolpakov, un des principaux collaborateurs du projet du côté russe, a été recruté dans POLKA au 1$^{er}$ septembre 2000 sur un poste INRIA d'accueil de spécialistes.

G. Kucherov a fait un exposé invité à la conférence *Discrete Models in the Theory of Control Systems* qui a été organisée par l'Université de Moscou et s'est tenue à Krasnovidovo (région de Moscou) en juin 2000.

### 8.3 Visites, et invitations de chercheurs

Alain Goupil, professeur associé au LaCIM à l'Université du Québec à Montréal, a été invité au sein de POLKA. Il a bénéficié d'un poste de professeur invité au LORIA pour un mois (juillet 2000) et à cette occasion, Dominique Poulalhon, doctorante au LIX à l'École Polytechnique, est venue travailler une semaine au sein de POLKA.

Mireille Bousquet-Mélou a été invitée une semaine en janvier 2000 au sein de POLKA. À cette occasion, elle a donné un exposé au Séminaire d'Informatique Fondamentale du LORIA et travaillé avec G. Schaeffer sur des problèmes de marches aléatoires.

En mars, Mostefa Mesmoudi, chercheur algérien, a été invité dans l'équipe pour une durée d'un mois et a travaillé avec I. Debled-Rennesson sur les problèmes de reconnaissance de morceaux de plans discrets.

En octobre, Joris van der Hoeven (Université de Paris Sud) a été invité pendant une semaine au sein de POLKA ; il a principalement travaillé avec P. Zimmermann sur l'implantation efficace de son algorithme de division « détendue ».

### 8.4 Animation scientifique

G. Kucherov est membre du comité de programme des conférences JOBIM'2001 [12] (*Journées Ouvertes : Biologie, Informatique et Mathématiques*) et PSI'2001 [13] (*Perspectives of System Informatics*).

G. Schaeffer est membre du jury du prix de thèse 2000 de SPECIF.

G. Schaeffer est responsable du Séminaire d'Informatique Fondamentale du LORIA.

P. Zimmermann est membre du comité de programme de la conférence ARITH15 qui aura lieu en 2001 (`http://arith15.polito.it/`), et membre du comité scientifique de l'ACI (action concertée incitative) Cryptologie lancée par le ministère de la recherche en juillet 2000 (`http://www.edutel.fr/rt/appel/2000/crypto.htm`). P. Zimmermann est également membre élu de la commission d'évaluation de l'INRIA ; il a été rapporteur des thèses de Vincent Lefèvre (Lyon) et de Jean-Guillaume Dumas (Grenoble), et examinateur de la thèse de Julien Clément (Caen).

### 8.5 Enseignement

G. Hanrot et P. Zimmermann ont dispensé conjointement un cours de cryptographie en troisième année à l'ESIAL de janvier à mars 2000. Ce cours a été reconduit pour l'année 2000-2001, et a eu lieu en novembre-décembre 2000.

G. Hanrot a assuré – en commun avec F. Morain (École polytechnique) le cours de DEA Algorithmique « Théorie algorithmique des nombres ».

G. Kucherov a préparé son mémoire d'habilitation à diriger les recherches [Kuc00].

12. `http://jobim.toulouse.inra.fr/`
13. `http://www.iis.nsk.su/PSI01/index.html`

[Kuc00] G. KUCHEROV, *Motifs dans les mots et les arbres*, Habilitation à diriger des recherches, Université Henri Poincaré Nancy 1, 2000.

G. Kucherov, en commun avec D. Kratsch (Université de Metz), enseigne le module *Algorithmique des structures discrètes* du DEA d'Informatique de l'Université Henri Poincaré de Nancy (filière *Algorithmique Numérique et Symbolique*).

Dans cette même filière, G. Hanrot et P. Zimmermann organisent avec N. Hermann (Prothéo) un module de « Théorie Algorithmique des Nombres, Codage et Cryptographie» ; cette année le module est orienté vers le codage, et les cours sont dispensés par N. Hermann et P. Zimmermann.

G. Schaeffer a encadré le stage de S. Kunz-Jacques (élève à l'ENS Paris), au sein du DEA « Algorithmique ». Ce stage, effectué au Laboratoire d'Informatique de l'École Polytechnique (LIX), a porté sur les entrelacs et les algorithmes de génération aléatoire, et se poursuit par une thèse, en codirection avec Robert Cori (professeur à l'École polytechnique).

G. Hanrot a co-encadré, avec François Morain (LIX), le stage de N. Gurel (élève au DEA Algorithmique), effectué au LIX. Ce stage traitait de l'algorithmique dans les jacobiennes de certaines familles de courbes, avec applications en cryptographie. Ce stage se poursuit par une thèse, toujours au LIX et en codirection avec F. Morain.

J. Rouyer, responsable du module «Découverte de la recherche» en ESIAL deuxième année depuis plusieurs années, a publié un article avec Monique Grandbastien sur cette expérience, article présenté à la conférence IFIP2000 [18].

Fabrice Rouillier a assuré une partie du cours de l'option « calcul formel» du DEA Algorithmique à Jussieu.

# 9 Bibliographie

## Ouvrages et articles de référence de l'équipe

[1] R. Kolpakov, G. Kucherov, «Finding Maximal Repetitions in a Word in Linear Time», *in : Proceedings of the 1999 Symposium on Foundations of Computer Science, New York (USA)*, IEEE Computer Society, p. 596–604, New-York, 17-19 octobre 1999.

[2] G. Schaeffer, «Random sampling of large planar maps and convex polyhedra», *in : Proceedings of the thirty-first annual ACM symposium on theory of computing (STOC'99)*, ACM press, p. 760–769, Atlanta, Georgia, mai 1999.

## Articles et chapitres de livre

[3] Y. Bilu, G. Hanrot, P. Voutier, « Existence of primitive divisors of Lucas and Lehmer numbers (with an appendix by Maurice Mignotte)», *J. Reine Angew. Math.*, à paraître.

[4] M. Bousquet-Mélou, G. Schaeffer, «Enumeration of planar constellations», *Adv. in Appl. Math. 24*, 4, 2000, p. 337–368.

[5] Y. Bugeaud, G. Hanrot, «Un nouveau critère pour l'équation de Catalan», *Mathematika*, à paraître.

[6] R. Cori, G. Schaeffer, « Description trees and Tutte's formulas », *Theoretical Computer Science*, à paraître.

[7] V. Grebinski, G. Kucherov, «Optimal Reconstruction of Graphs under the Additive Model», *Algorithmica 28*, 2000, p. 104–124.

[8] G. Hanrot, N. Saradha, T. N. Shorey, « Almost perfect powers in consecutive integers », *Acta Arith.*, à paraître.

[9] G. Hanrot, «Solving Thue equations without the full unit group», *Math. Comp. 69*, 229, 2000, p. 395–405.

[10] R. Kolpakov, G. Kucherov, «On Maximal Repetitions in Words», *Journal on Discrete Algorithms*, 2000, à paraître, `http://www.loria.fr/publications/2000/A00-R-170/A00-R-170.ps`.

## Communications à des congrès, colloques, etc.

[11] J. Abbott, V. Shoup, P. Zimmermann, « Factorization in $\mathbb{Z}[x]$: the searching phase », *in : Proceedings of ISSAC'2000*, C. Traverso (éditeur), ACM Press, p. 1–7, 2000.

[12] C. Banderier, P. Flajolet, G. Schaeffer, M. Soria, «Planar Maps and Airy Phenomena», *in : Automata, Languages, and Programming*, U. Montanari, J. Rolim, E. Welzl (éditeurs), *LNCS*, 1853, p. 388–402, 2000. Proceedings of the 27th ICALP Conference, Geneva, juillet 2000.

[13] M. Bousquet, C. Chauve, G. Schaeffer, «Énumération et génération aléatoire de cactus m-aires», *in : Colloque LaCIM 2000, Combinatoire, Informatique et Applications*, 2000, `http://www.lacim.uqam.ca/lacim2000`.

[14] M. Bousquet-Mélou, G. Schaeffer, « Counting paths on the slit plane », *in : Mathematics and Computer Science : Algorithms, Trees, Combinatorics and Probabilities*, D. Gardy, A. Mokkadem (éditeurs), *Trends in Mathematics*, Birkhäuser, p. 101–112, 2000. Proceedings of the Colloquium on Mathematics and Computer Science, in Versailles.

[15] S. Cavallar, B. Dodson, A. K. Lenstra, W. Lioen, P. L. Montgomery, B. Murphy, H. te Riele, K. Aardal, J. Gilchrist, G. Guillerm, P. Leyland, J. Marchand, F. Morain, A. Muffett, C. Putnam, C. Putnam, P. Zimmermann, « Factorization of a 512–bit RSA Key », *in : Proceedings of Eurocrypt'2000*, Bruges, mai 2000. 16 pages.

[16] I. Debled-Rennesson, J.-L. Rémy, J. Rouyer-Degli, «Detection of the Discrete Convexity of Polyominoes», *in : DGCI 2000, Uppsala, Suède*, *Lecture Notes in Computer Science*, Springer-Verlag, décembre 2000.

[17] A. Goupil, D. Poulalhon, G. Schaeffer, « Central Characters and Conjugacy Classes in the Symmetric Group », *in : Formal Power Series and Algebraic Combinatorics*, D. Krob, A. Mikhalev, A. Mikhalev (éditeurs), Springer, p. 238–249, 2000. Proceedings of the 12th International Conference, FPSAC'00, Moscow.

[18] M. Grandbastien, J. Rouyer, « Research in informatics engineering curricula », *in : IFIP International Conference on Knowledge Systems and Distributed Intelligence, Beijing-China*, août 2000.

[19] R. Kolpakov, G. Kucherov, « Finding Repeats With Fixed Gap », *in : 7th International Symposium on String Processing and Information Retrieval (SPIRE) , A Coruna, Spain (27 - 29 Septembre, 2000), Coruna, Spain*, IEEE Computer Society, p. 162–168, septembre 2000, `http://www.loria.fr/publications/2000/A00-R-166/A00-R-166.ps`.

[20] G. Kucherov, M. Giraud, « Maximal Repetitions and Application to DNA sequences », *in : Journées Ouvertes : Biologie, Informatique et Mathématiques, Montpellier, France*, G. Caraux, O. Gascuel, M.-F. Sagot (éditeurs), p. 165–172, mai 2000, `http://www.loria.fr/publications/2000/A00-R-165/A00-R-165.ps`.

[21] L. Rolland, « The Manta and the Kanuk: Novel 4-DOF Parallel Mechanisms for Industrial Handling », *in : Proceedings of the IMECE 99 conference*, *DSC-67*, p. 831–844, Novembre 1999.

## Rapports de recherche et publications internes

[22] G. Hanrot, M. Quercia, P. Zimmermann, « Speeding up the Division and Square Root of Power Series », *Research Report n° 3973*, juillet 2000, `http://www.inria.fr/rrrt/rr-3973.html`.

[23] E. Jeandel, « Évaluation rapide de fonctions hypergéométriques », *Rapport technique n° 242*, juillet 2000, 17 pages, `http://www.inria.fr/rrrt/rt-0242.html`.

[24] S. Kunz-Jacques, G. Schaeffer, « The asymptotic number of prime alternating links », *rapport de recherche*, 2000, 12pp, résumé étendu, soumis à FPSAC'01.

## Divers

[25] P. Zimmermann, « A proof of GMP fast division and square root implementations », septembre 2000, 14 pages, `ftp://www.loria.fr/~~zimmerma/papers/proof-div-sqrt.ps.gz`.