Calcul intensif, algèbre linéaire exacte et applications

Clément PERNET
LIG, Grenoble Université

1ères journées du GDR Calcul,
10 novembre 2009

Introduction

Calcul formel, calcul algébrique, calcul exact:

- Objets mathématiques **exacts**: $\sqrt{2} \neq 1,414$
- Manipulation d'expressions mathématiques
  - dérivations, primitive
  - résolutions d'équations avec paramètre,
  - simplifications formelles, ...

Introduction

Calcul formel, calcul algébrique, calcul exact:

- Objets mathématiques **exacts**: $\sqrt{2} \neq 1,414$
- Manipulation d'expressions mathématiques
  - dérivations, primitive
  - résolutions d'équations avec paramètre,
  - simplifications formelles, ...

Domaines de calcul de base

- $\mathbb{Z}, \mathbb{Q}$ $\Rightarrow$taille variable
- $\mathbb{Z}_p, \mathrm{GF}(p^k)$ $\Rightarrow$arithmétique propre
- $K[X]$ pour $K = \mathbb{Z}, \mathbb{Z}_p, ...$
- ...

Introduction

Idée reçue: Calcul formel = `Maple`

- Outil pédagogique
- Utile pour simplifier les formules, ... quand ça marche
- Lent, inadapté au calcul intensif
- les complexités sont catastrophiques

Introduction

Idée reçue: Calcul formel = `Maple`

- Outil pédagogique
- Utile pour simplifier les formules, ... quand ça marche
- Lent, inadapté au calcul intensif
- les complexités sont catastrophiques

Et pourtant:

- applications demandeuses de calcul intensif
- bibliothèques/logiciels spécialisés et efficaces
- ameliorations en cours à tous les niveaux: algorithmiques et implantations

Plan

Plan

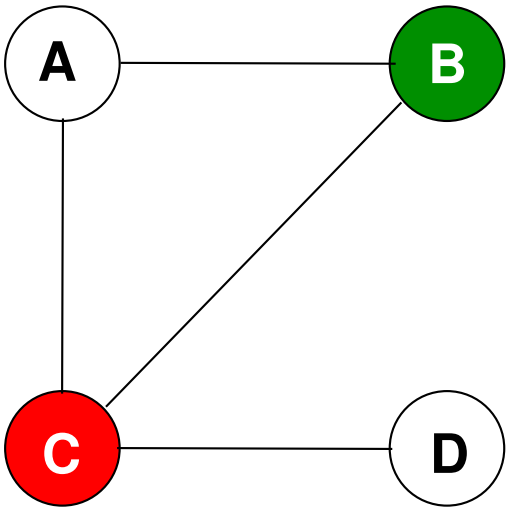
Applications: Calcul mathématiques

Aide à la recherche en Mathématiques:

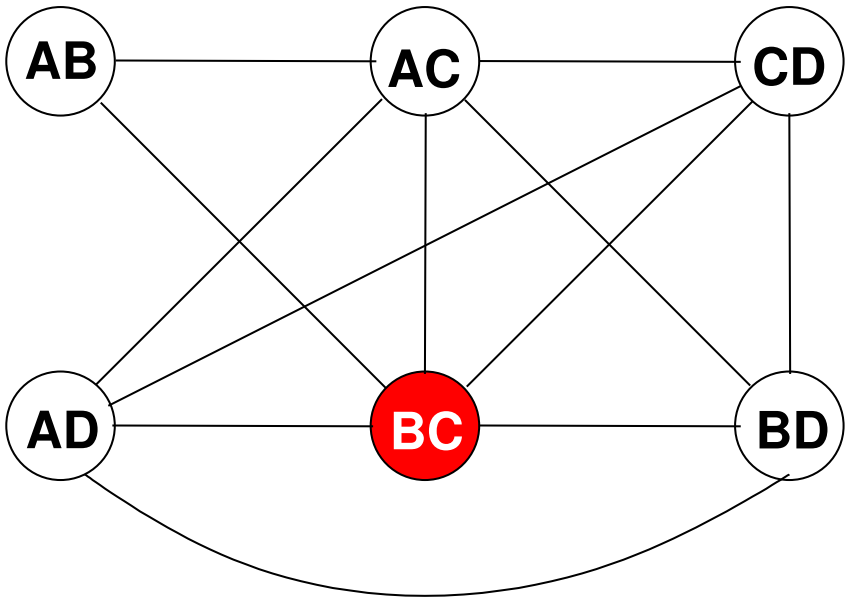
- simplification de formules
- résolution d'équations sous forme analytique
- calcul avec des objets sophistiqués (formes modulaires, groupes abéliens, ...)
- mathématiques expérimentales, tests de conjectures
- calcul de tables: *bestiaires* d'objets mathématiques

Calcul mathématique: Théorie des Graphes

Marche aléatoire dans un graphe

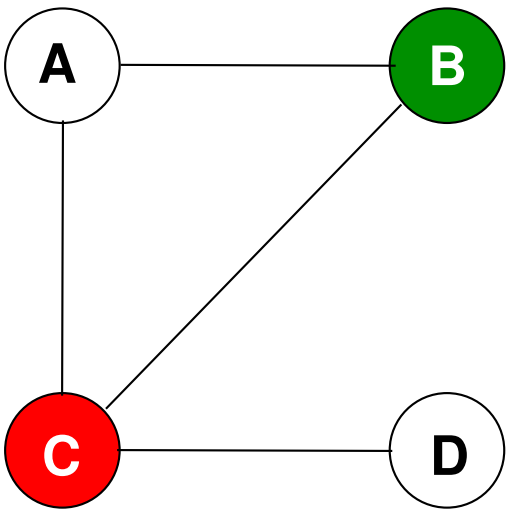


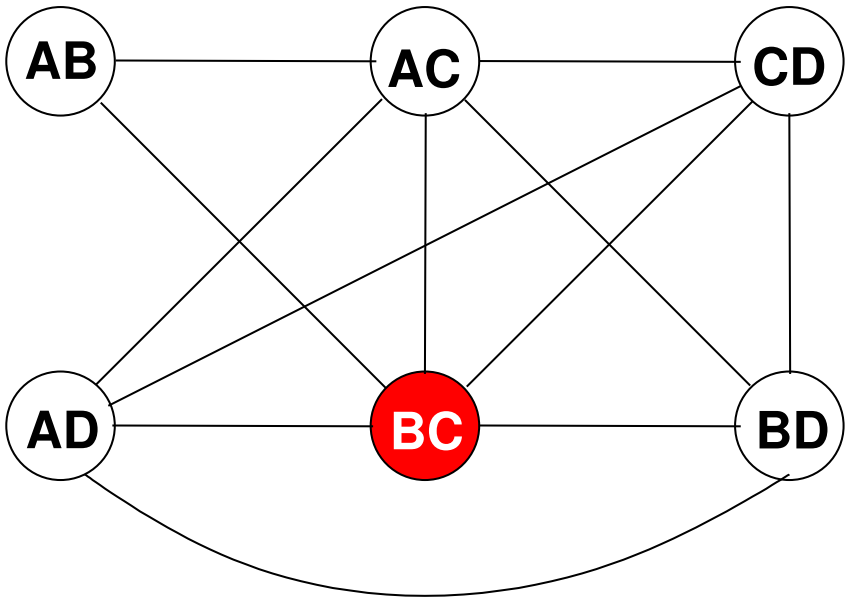
2-marche



Calcul mathématique: Théorie des Graphes

Marche aléatoire dans un graphe



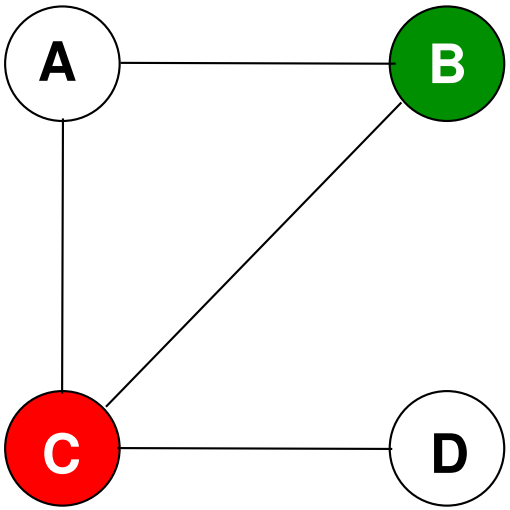


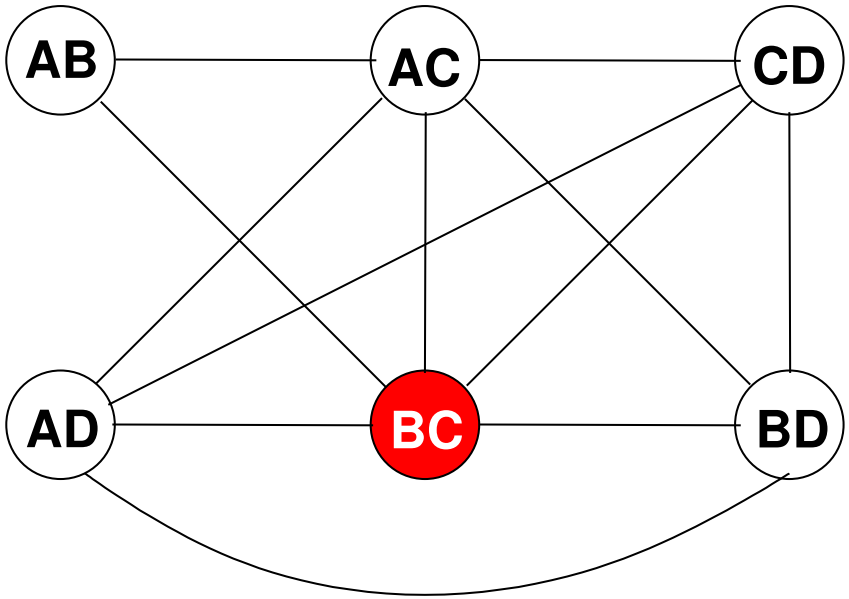
2-marche

- carrés symétriques d'un graphe X : le graphe $X^{\{2\}}$ des $\binom{n}{2}$ paires de sommets
- 2-marche dans $X \equiv$ 1-marche dans $X^{\{2\}}$

Calcul mathématique: Théorie des Graphes

Marche aléatoire dans un graphe





2-marche

- carrés symétriques d'un graphe X : le graphe $X^{\{2\}}$ des $\binom{n}{2}$ paires de sommets
- 2-marche dans $X \equiv$ 1-marche dans $X^{\{2\}}$

Application :

- Modélisation des systèmes Hamiltoniens en mécanique quantique
- Isomorphismes de Graphes

Isomorphisme de Graphes

Problème

Existe-t-il un algorithme polynomial testant si deux graphes sont isomorphes ?

Isomorphisme de Graphes

Problème

Existe-t-il un algorithme polynomial testant si deux graphes sont isomorphes ?

Piste [Royle et al. 2006] : le spectre d'une puissance symétrique du graphe détermine sa classe d'isomorphisme

Isomorphisme de Graphes

Problème

Existe-t-il un algorithme polynomial testant si deux graphes sont isomorphes ?

Piste [Royle et al. 2006] : le spectre d'une puissance symétrique du graphe détermine sa classe d'isomorphisme

Expérimentations : puissances symétriques de familles de graphes fortement réguliers

Isomorphisme de Graphes

Problème

Existe-t-il un algorithme polynomial testant si deux graphes sont isomorphes ?

Piste [Royle et al. 2006] : le spectre d'une puissance symétrique du graphe détermine sa classe d'isomorphisme

Expérimentations : puissances symétriques de familles de graphes fortement réguliers

- $k = 2$: faux (preuve [Royle et al. 2006])
- $k = 3$: vrai jusqu'à 29 sommets (70 cas, $n = 3654$)
- $k = 3$: vrai jusqu'à 36 sommets (36 510 cas, $n = 7140$)

$\Rightarrow$588 heures CPU

Cryptographie

Sécurité $\equiv$ Difficulté à résoudre un problème

Pour RSA $\Rightarrow$factorisation de grands entiers

Cryptographie

Sécurité $\equiv$ Difficulté à résoudre un problème

Pour RSA $\Rightarrow$factorisation de grands entiers

- Cribles (quadratique, de corps de nombre,...)

Repose sur:

- arithmétique de grands entiers
- pgcd
- Résolution d'un grand système linéaire creux dans $\mathbb{Z}_2$.

Cryptographie

Sécurité $\equiv$ Difficulté à résoudre un problème

Pour RSA $\Rightarrow$factorisation de grands entiers

- Cribles (quadratique, de corps de nombre,...)

Repose sur:

- arithmétique de grands entiers
- pgcd
- Résolution d'un grand système linéaire creux dans $\mathbb{Z}_2$.

Example (RSA-640:)

*Cassé en 2005 en 4,5 mois (*263 500 *heures CPU)*

- 36 000 000 *colonnes*
- 7×10^9 *coeffs non nuls*

Autres domaines

Bio-info: Distribution de motifs dans des chaînes aléatoires d'ADN:

- Grand systèmes creux, $n \approx 60\,000$, $mnz = 20 \times 60\,000$,
- Coefficients: polynômes bivariés.

Autres domaines

Bio-info: Distribution de motifs dans des chaînes aléatoires d'ADN:

- Grand systèmes creux, $n \approx 60\,000$, $mnz = 20 \times 60\,000$,
- Coefficients: polynômes bivariés.

Topologie algorithmique:

- Analyse sémantique massive de textes
  - Graphes d'occurrence de termes dans des documents
  - Calcul d'homologie persistante sur le graphe $\Rightarrow$sémantique
- Analyse de données 3D (simulations et expériences)
  - Homologie $\Rightarrow$extraire l'information géométrique

Plan

Quelques ingrédients pour le calcul exact

Arithmétique flottante: `float, double`

- Privilégiée par les architectures: `fma`, SSE, GPU, ...
- Utilisée pour:
  - Corps finis: `NTL`, `LinBox`,...
  - Réduction de réseaux: `fpLLL`
- Calcul exact sur la mantisse uniquement, ou approximation controllée

Quelques ingrédients pour le calcul exact

Arithmétique flottante: `float, double`

- Privilégiée par les architectures: `fma`, SSE, GPU, ...
- Utilisée pour:
  - Corps finis: `NTL`, `LinBox`,...
  - Réduction de réseaux: `fpLLL`
- Calcul exact sur la mantisse uniquement, ou approximation controllée

Réduction à des routines de base

Multiplication entier/polynôme: Karatsuba, Toom-Cook, FFT

Multiplication de matrices: BLAS, Strassen,

$\Rightarrow$algorithmes de réduction récursifs par blocs

Solutions logicielles pour le calcul algebrique

Bibliothèques spécialisées

corps finis: NTL, Givaro, Lidia, ...

entiers multiprécision: GMP, MPIR

polynômes: NTL, Givaro, zn_poly ...

Solutions logicielles pour le calcul algebrique

Bibliothèques spécialisées

corps finis: NTL, Givaro, Lidia, ...

entiers multiprécision: GMP, MPIR

polynômes: NTL, Givaro, zn_poly ...

Intergiciels

- génériques
- focalisés sur les algorithmes

Solutions logicielles pour le calcul algebrique

Bibliothèques spécialisées

corps finis: NTL, Givaro, Lidia, ...

entiers multiprécision: GMP, MPIR

polynômes: NTL, Givaro, zn_poly ...

Intergiciels

- génériques
- focalisés sur les algorithmes

Logiciels généralistes, haut-niveau

- Maple, Mathematica, MuPad, ... (propriétaire)
- Sage, Pari, Maxima, ... (libre)

Solutions logicielles pour le calcul algebrique

Bibliothèques spécialisées

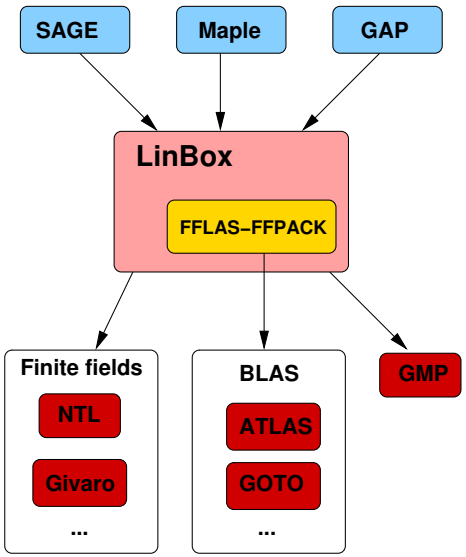
corps finis: NTL, Givaro, Lidia, ...

entiers multiprécision: GMP, MPIR

polynômes: NTL, Givaro, zn_poly ...

Intergiciels

- génériques
- focalisés sur les algorithmes



Logiciels généralistes, haut-niveau

- Maple, Mathematica, MuPad, ... (propriétaire)
- Sage, Pari, Maxima, ... (libre)

Sage: logiciel libre (GPL) de mathématiques

- calcul formel, numérique, géométrie, statistiques, ...

```
----------------------------------------------------------------------
| Sage Version 4.2, Release Date: 2009-10-24                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: R.<x>=Pol<tab>
Polyhedron                          PolynomialQuotientRingElement
Polynomial                          PolynomialRing
PolynomialQuotientRing
sage: R.<x>=PolynomialRing(GF(2))
sage: P=x^2+1
sage: P.parent()
Univariate Polynomial Ring in x over Finite Field of size 2 (using NTL
sage: P.factor()
(x + 1)^2
```

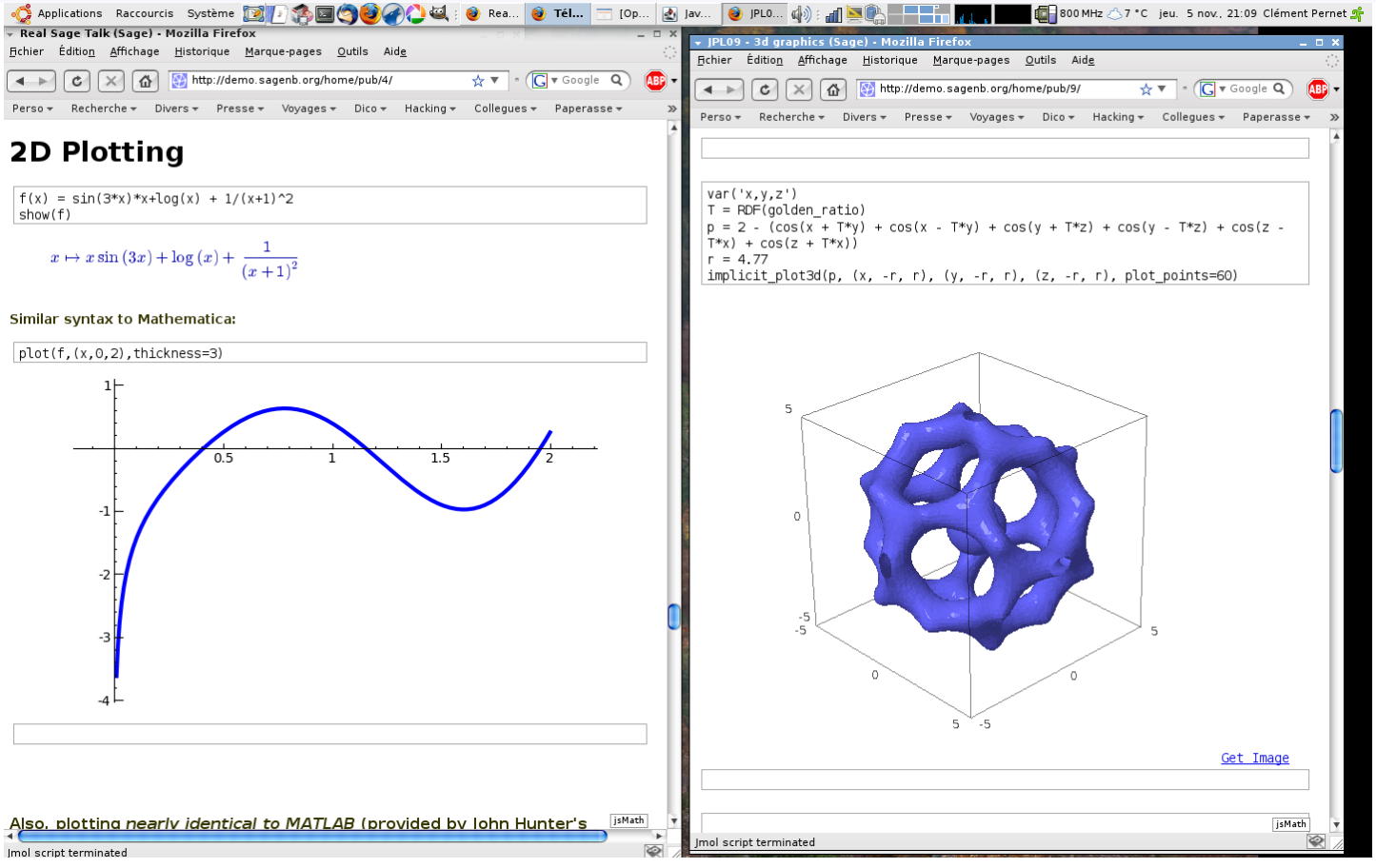
- Linux, MacOS X, Solaris, Windows (en cours de portage)
- x86, x86_64, PPC,

Sage: une distribution

- Une distribution des meilleures bibliothèques spécialisées et logiciels libre de mathématique (plus de 70 paquets)

Arithmetique	GMP, MPFR, Givaro, MPFI
Algèbre commutative	PolyBoRi, SINGULAR (libSINGULAR)
Algèbre linéaire	LinBox, M4RI, IML, fpLLL
Cryptosystèmes	GnuTLS, PyCrypto
factorisation entière	FlintQS, ECM
Théorie des groupes	GAP
Combinatoire	Symmetrica, sage-combinat
Théorie des graphes	NetworkX
Théorie des nombres	PARI, NTL, Flint, mwrank, eclib
Calcul numérique	GSL, Numpy, Scipy, ATLAS
Calcul formel	Maxima, Sympy, Pynac
Statistiques	R
Interface utilisateur	Sage Notebook, jsmath, Moin wiki, IPython
Graphiques	Matplotlib, Tachyon, libgd, JMol
Réseau	Twisted
Bases de donnée	ZODB, SQLite, SQLAlchemy, Python pickle
Langage de programmation	Python, Cython (compiled)

Sage: interface graphique



- Intégrée dans un navigateur Web
- typographie LaTeX

- partage de feuilles de calcul
- plots 2D, 3D interactifs
- Applets interactifs

Sage

Une bibliothèque de code source propre

- plus de 1M de lignes de code
- langage Python, Cython (Python compilé)

Un modèle de développement de type académique:

- Plus de 150 contributeurs, dont 50 environ par release
- Tout nouveau code, est proposé, puis reviewé par un référé, avant l'inclusion dans la distribution suivante.
- tests de régression automatisés

Sage

Une interface unifiée vers d'autre systèmes:

```
  ----------------------------------------------------------------------
| Sage Version 4.1.1, Release Date: 2009-08-14                         |
| Type notebook() for the GUI, and license() for information.          |
----------------------------------------------------------------------
sage: pari('factor(x^2-1)')
[x - 1, 1; x + 1, 1]
```

- Maple, Magma, Mathematica, Matlab, Pari, MuPad, Maxima, Octave, Singular, GAP, Axiom, Macaulay2, Pari/GP,...

Plan

Le problème

Clay Math Institute, $1M challenge:

Problème (Conjecture Birch Swinnerton-Dyer)

Une "méthode" pour déterminer si toute équation du type $Y^2 = X^3 + aX + b$ a une infinité de solutions rationnelles

Le problème

Clay Math Institute, $1M challenge:

Problème (Conjecture Birch Swinnerton-Dyer)

Une "méthode" pour déterminer si toute équation du type $Y^2 = X^3 + aX + b$ a une infinité de solutions rationnelles

Expérimentations
$\Downarrow$
infirmatons, ou reformulation
$\Downarrow$
...
$\Downarrow$
résolution

Besoin:

- Bestiaires de formes modulaires
- Calculer des tables les plus grandes possibles

Calcul de formes modulaires

Action des opérateurs de Hecke sur l'espace

- Décomposition de l'espace en invariants de similitude
- Calcul de
  - Polynômes caractéristiques dans $\mathbb{Z}_p$
  - Noyaux dans $\mathbb{Z}$

"Mathematics is the art of reducing any problem to linear algebra" *W. Stein*

Algèbre linéaire exacte: algorithmique

Modèles algorithmiques

- Dense
- Creuse
- Boite-noire

Problématiques:

- Pas d'instabilité
- Mais taille variable

$\Rightarrow$réduire la manipulation de grands entiers au maximum.

Algèbre linéaire exacte: algorithmique

Modèles algorithmiques

- Dense
- Creuse
- Boite-noire

Problématiques:

- Pas d'instabilité
- Mais taille variable

$\Rightarrow$réduire la manipulation de grands entiers au maximum.

Exemple: Calcul du déterminant dans $\mathbb{Z}$

Méthode	Complexité
Gauss naïf dans $\mathbb{Q}$	$\mathcal{O}(\exp(n))$
Gauss mod det	$\mathcal{O}(n^6)$
Gauss mod p + TRC	$\mathcal{O}(n^4), \mathcal{O}(n^{\omega+1})$
Lifting p-adique	$\mathcal{O}(n^3), \mathcal{O}^{\sim}(n^{\omega})$

Algèbre linéaire exacte: algorithmique

Le polynôme caractéristique dans $\mathbb{Z}_p$:

1890 $\mathcal{O}\left(n^4\right)$

1937 $\mathcal{O}\left(n^3\right)$

1985 $\mathcal{O}\left(n^\omega \log n\right)$

2007 $\mathcal{O}\left(n^\omega\right)$

Algèbre linéaire exacte: algorithmique

Le polynôme caractéristique dans $\mathbb{Z}_p$:

1890 $\mathcal{O}(n^4)$

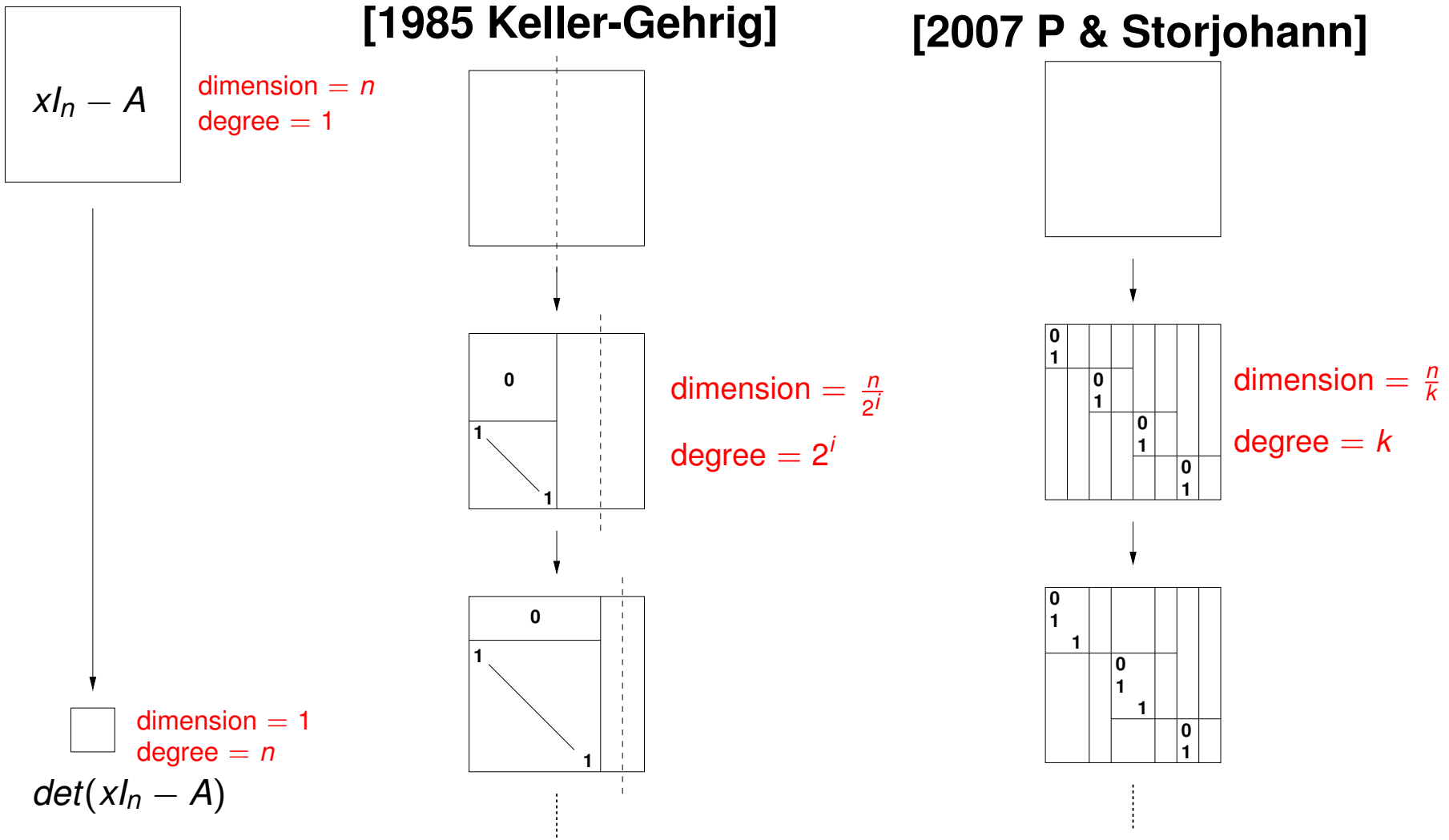
1937 $\mathcal{O}(n^3)$

1985 $\mathcal{O}(n^\omega \log n)$

2007 $\mathcal{O}(n^\omega)$

$\Rightarrow$réduction au produit matriciel

Algèbre linéaire exacte: algorithmique



Le produit de matrices : une brique de base

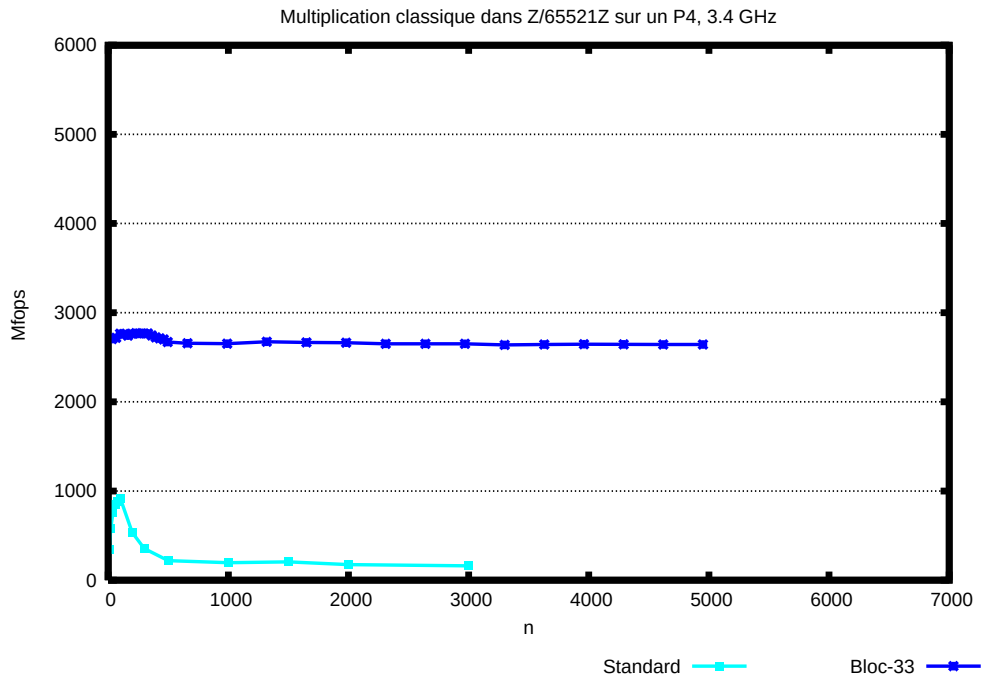
Principe:

- Réduction modulaire différée
- Arithmétique flottante (`fma`, SSE2, ...)

Le produit de matrices : une brique de base

Principe:

- Réduction modulaire différée
- Arithmétique flottante (`fma`, SSE2, ...)
- Optimisation de cache

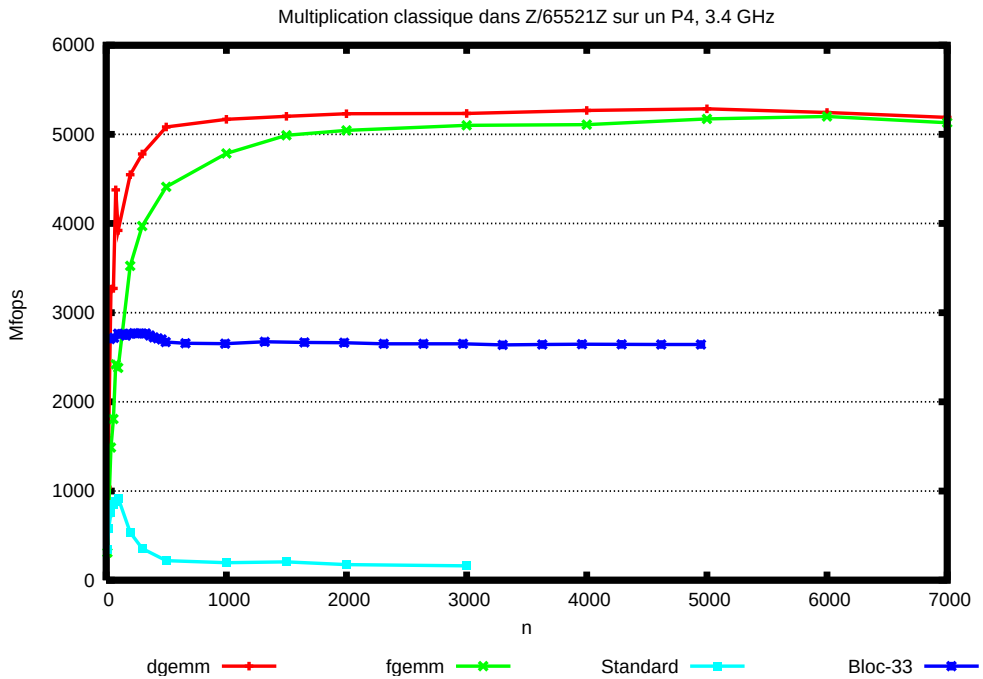


Le produit de matrices : une brique de base

Principe:

- Réduction modulaire différée
- Arithmétique flottante (`fma`, SSE2, ...)
- Optimisation de cache

  $\Rightarrow$repose sur les BLAS existants

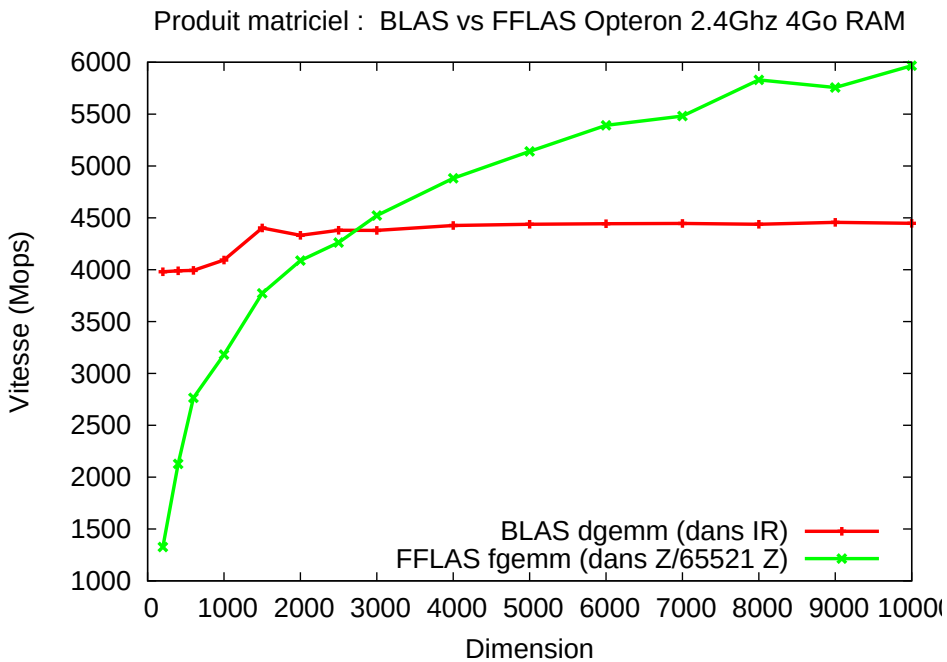


Le produit de matrices : une brique de base

Principe:

- Réduction modulaire différée
- Arithmétique flottante (`fma`, SSE2, ...)
- Optimisation de cache

  $\Rightarrow$repose sur les BLAS existants
- Algorithme sous-cubique (Strassen-Winograd)



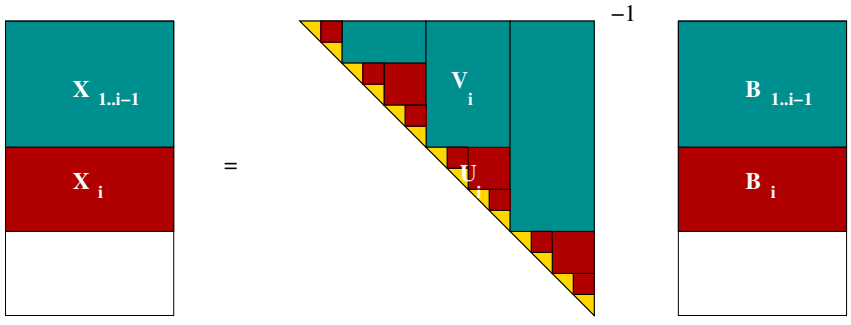
Autres routines d'algèbre linéaire dense

- Réduction au produit de matrice
- Bornes pour la réduction modulaire différée

Autres routines d'algèbre linéaire dense

- Réduction au produit de matrice
- Bornes pour la réduction modulaire différée

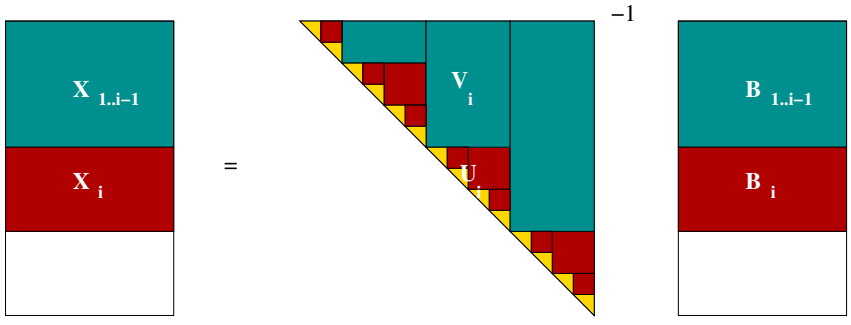
$\Rightarrow$Algorithmes par blocs en cascade



Autres routines d'algèbre linéaire dense

- Réduction au produit de matrice
- Bornes pour la réduction modulaire différée

$\Rightarrow$Algorithmes par blocs en cascade



	n	*1000*	*2000*	*3000*	*5000*	*10 000*
TRSM	$\frac{ftrsm}{dtrsm}$	1,66	1,33	1,24	1,12	1,01
LQUP	$\frac{lqup}{dgetrf}$	2,00	1,56	1,43	1,18	1,07
INVERSE	$\frac{inverse}{dgetrf+dgetri}$	1.62	1.32	1.15	0.86	0.76

Polynôme caractéristique:

n	500	5000	15 000
`LinBox`	0.91s	4m44s	2h20m
`magma-2.13`	1.27s	15m32s	7h28m

Plan

Calcul formel intensif, spécificités et similitudes

Une approche similaire au calcul numérique

- Structuration des bibliothèques
- Approche en algèbre linéaire:
  - calcul flottant
  - noyaux et réductions
  - méthodes itératives

Calcul formel intensif, spécificités et similitudes

Une approche similaire au calcul numérique

- Structuration des bibliothèques
- Approche en algèbre linéaire:
  - calcul flottant
  - noyaux et réductions
  - méthodes itératives

Parallèlisme:

- Pas de parallélisation imposée par le domaine
- Restes Chinois: parallélisation *facile*
- Estimation des coûts plus délicate (arithmétique de coût variable) $\Rightarrow$balancement dynamique de la charge, vol de travail (`Kaapi`)

Conclusion

Interactions

- Calcul certifié, haute précision
- Couplage de méthodes exacte/approchées
- ...