Remerciements

J'ai connu Brigitte Vallée d'abord à travers ses enseignements de qualité à l'ENSI de Caen. C'est elle qui m'a encouragé à faire un DEA, et m'a encadré depuis. Son enthousiasme et sa patience ont su me faire partager sa passion pour la recherche et en particulier pour les problèmes algorithmiques liés à la géométrie des nombres. Sa finesse et sa vivacité, mais aussi sa rigueur et son exigence ont souvent éclairé mes idées, parfois confuses. Aujourd'hui, je connais ma voie et cette certitude réconfortante, je lui en suis redevable. Elle est pour moi non seulement un soutien inestimable, mais aussi une précieuse amie. Merci Brigitte.

Richard Brent, François Morain et Claus Schnorr m'ont fait le grand honneur d'être rapporteurs de cette thèse. Je les remercie très vivement pour le temps qu'ils m'ont accordé: leurs commentaires me furent fort utiles.

Je tiens à exprimer toute ma gratitude envers Philippe Flajolet pour ses points de vue, toujours très enrichissants, pour l'intérêt qu'il porte à mon travail et pour avoir accepté de faire partie de mon jury.

J'adresse mes sincères remerciements à Hervé Daudé pour sa participation à mon jury et aussi pour sa thèse qui a servi comme point de départ à ce travail.

Ma reconnaissance va également à Etienne Grandjean et à Claire Kenyon pour avoir accepté d'être membres de mon jury. Je suis flatté de l'intérêt qu'ils témoignent ainsi à mon travail.

J'ai poursuivi la majeure partie de cette thèse sans allocation. Je tiens à remercier le département d'informatique de l'Université de Caen, et en particulier Etienne Grandjean, pour m'avoir toujours trouvé des vacations. Je remercie également très sincèrement Pierre Vallée, pour les colles en prépa et toute la gymnastique administrative que cela lui imposait parfois. Merci également à Samuel Dellacherie pour les "colles d'informatique" à Rouen. Tous ces cours, nécessaires financièrement et parfois trop nombreux, m'ont en même temps formé au métier d'enseignant–chercheur.

Je remercie tous les membres du GREYC pour m'avoir accueilli au sein du laboratoire et en particulier Julien, Samuel, Samuel et Véronique pour de nombreux échanges d'idées. Je ne saurais oublier les secrétaires du département d'informatique, Françoise, Sylvie et Séverine pour leur gentillesse et leur disponibilité.

Je voudrais aussi remercier tous ceux qui m'ont encouragé et soutenu pendant ce travail, qui m'ont supporté ou qui ont supporté mes absences: mes amis, mes parents, ma famille adoptive, mes beaux–parents et tout particulièrement, Bernard Gille et Gisèle Remteaux, à qui je dois tellement, ma sœur Anne–France,
et surtout mon épouse adorée, Mathilde, à qui je dédie cette thèse.

Table des matières

Chapitre 1

Introduction générale

Ce travail est une étude des algorithmes de réduction des réseaux euclidiens. Un *réseau euclidien* est l'ensemble des combinaisons linéaires entières d'un système libre de p vecteurs de $\mathbb{R}^n$. Les p vecteurs forment alors une base du réseau et leur nombre est la dimension du réseau. Un même réseau peut ainsi être engendré par de multiples bases. *Réduire un réseau* consiste à trouver une base qui engendre le même réseau et qui soit de bonne qualité vis-à-vis de la structure euclidienne: les vecteurs de la base doivent être assez courts et en même temps assez orthogonaux.

Ainsi décrit, ce problème semble certes naturel mais aussi bien étroit. De fait, c'est un problème de base, tout comme l'algorithme d'Euclide dont il est en un sens la généralisation multidimensionnelle. Il s'avère en effet que beaucoup de problèmes divers se formalisent aisément grâce à des réseaux euclidiens, et que la solution du problème posé s'obtient aisément dès que l'on dispose d'une base de bonne qualité. C'est le cas d'un certain nombre de problèmes purement mathématiques, en théorie des nombres, par exemple. Mais, de manière plus inattendue, on retrouve aussi les réseaux euclidiens et leur réduction dans d'autres domaines a priori plus lointains. La cryptographie est un exemple frappant, puisque beaucoup d'attaques intelligentes de cryptosystèmes divers sont fondées sur la réduction d'un réseau sous-jacent au problème.

La réduction des réseaux est un problème ancien qui a déjà intéressé bon nombre de mathématiciens, même si la plupart d'entre eux ne l'ont pas envisagé d'un point de vue que l'on qualifie maintenant d'algorithmique. Les travaux d'Hermite, Minkowski [Min91, Min10], Korkine et Zolotarev ont donné lieu à diverses notions de réduction. Seul Gauss a adopté un point de vue plus constructif et a proposé des algorithmes en petite dimension (2 et 3). Son algorithme de la dimension 2 est maintenant bien compris [Lag82, Val91, DFV97] et sera largement détaillé dans cette thèse (Chapitres 2 et 6) car il constitue la base des algorithmes de réductions efficaces en dimension quelconque.

Alors que la réduction de Minkowski privilégie la recherche de vecteurs courts, la réduction introduite par Korkine et Zolotarev cherche plutôt une base proche d'une base orthogonale. Même si ces deux exigences sont reliées, notamment grâce aux théorèmes de Minkowski, ces notions de réduction s'avèrent différentes, et les réductions introduites plus récemment sont plutôt dans la lignée des travaux de Korkine et Zolotarev. La réduction introduite par Lovász en 1982 peut être considérée comme une version locale plus faible de la réduction de Korkine–Zolotarev. Ce nouveau point de vue sur la réduction a eu des conséquences algorithmiques extrêmement importantes, puisqu'il a permis, en 1982 et pour la première fois, la construction d'un algorithme efficace de réduction en toute dimension: c'est le célèbre algorithme LLL [LLL82], baptisé des initiales de ses inventeurs, H.W. Lenstra, A.K. Lenstra et L. Lovász, qui sert à la fois de point de départ et de référence constante à notre travail.

L'algorithme LLL est un algorithme itératif qui dépend d'un paramètre d'approximation t; il débute par le procédé classique d'orthogonalisation de Gram–Schmidt qui construit la base orthogonalisée de la base initiale. Puis, pendant une phase itérative de translations et d'échanges entre les vecteurs de la base d'entrée, l'algorithme construit une base réduite au sens de Lovász; c'est une base dont chaque vecteur b_i doit vérifier deux conditions principales: la projection de b_i sur le sous-espace engendré par les $i-1$ premiers vecteurs doit être "petite", tandis que la projection de b_i orthogonalement à ce même sous-espace engendré par les $i-1$ premiers vecteurs doit être relativement "grande". La première condition, dite de "propreté", vise à majorer les longueurs des vecteurs de la base, et est assurée par des translations entre vecteurs de la base. La deuxième condition vise à minorer les angles entre les vecteurs de la base, et est assurée par des échanges entre vecteurs de la base. Plus précisément, on cherche à assurer le même minorant s pour tous les rapports des longueurs de deux orthogonalisés consécutifs l_{i+1}/l_i. Ce minorant s est fonction du paramètre d'approximation t:

$$\frac{l_{i+1}}{l_i} > s \quad \text{avec} \quad s^2 = \frac{1}{t^2} - \frac{1}{4}. \tag{1.1}$$

Cet algorithme de réduction représente un très bon compromis entre deux exigences a priori contradictoires, puisque la base réduite jouit de bonnes qualités euclidiennes tout en étant obtenue rapidement. De plus, la qualité de la base de sortie et la complexité de l'algorithme dépendent tous deux du paramètre t, ce qui permet de disposer d'une stratégie algorithmique, qui s'avère efficace dans beaucoup des domaines applicatifs que nous avons déjà mentionnés ci-dessus.

En fait, l'algorithme LLL est un algorithme facile à implanter, très fréquemment utilisé et qui reste pourtant assez "magique" et bien mal compris. Ainsi, l'analyse précise de sa complexité reste très mal connue. Les inventeurs de l'algorithme, Lenstra, Lenstra et Lovász, ont cherché à analyser la complexité dans le pire des cas [LLL82], en majorant le nombre maximum d'itérations. La borne obtenue n'est visiblement pas optimale et n'est plus valable lorsque le paramètre d'approximation t est égal à 1 (valeur qui correspond à la meilleure approximation possible obtenue par l'algorithme LLL). Par ailleurs, Daudé et Vallée [DV94] ont obtenu un premier résultat sur l'analyse en moyenne de cet algorithme. Travaillant dans un modèle simple et naturel, où les vecteurs de la base d'entrée sont choisis de manière uniforme et indépendante dans la boule unité de $\mathbb{R}^n$, ils ont obtenu une majoration du nombre moyen des itérations. Là encore, cette borne ne semble pas optimale.

Notre travail vise à mieux comprendre généralement l'algorithmique de la réduction des réseaux; en particulier, et en tout premier lieu, nous cherchons à élucider le comportement de l'algorithme LLL, et, ce, à travers deux directions principales.

(1) Nous étudions la complexité dans le pire des cas de l'algorithme LLL, pour la valeur optimale $t = 1$ du paramètre d'approximation t.

(2) Nous introduisons de nouvelles notions de réduction, définissons les algorithmes de réduction correspondant et comparons ces réductions avec les algorithmes de réduction classique. Cette étude répond à une double motivation D'abord, la comparaison de ces algorithmes peut permettre de mettre sur pied une stratégie algorithmique de réduction. Ensuite, l'étude des algorithmes proches de l'algorithme LLL peut affiner la compréhension de ce dernier.

La première direction est adoptée dans le dernier chapitre (7) de ce travail. J'y étudie la complexité dans le pire des cas de l'algorithme LLL, pour la valeur optimale $t = 1$ du paramètre d'approximation t. L'algorithme trouve alors la meilleure base réduite possible au sens de Lovász, mais le résultat général obtenu par les inventeurs de l'algorithme ne s'applique pas dans ce cas limite. Lagarias et Odlyzko [LO83] citent des expériences qui semblent montrer qu'en moyenne l'algorithme reste polynomial. À ma connaissance, aucune réponse n'a été apportée jusqu'ici à cette question et le problème restait ouvert, même en petite dimension.

J'étudie l'algorithme LLL optimal lorsque la dimension du réseau est fixée, et, grâce à des arguments combinatoires, je montre qu'il fonctionne dans le pire des cas toujours en temps polynomial par rapport à la taille de la base d'entrée. En particulier, en dimension 3, je montre que le nombre d'itérations reste linéaire en la taille de la base d'entrée, et je donne une borne explicite pour le nombre maximum d'itérations. J'obtiens également une caractérisation précise de la base de sortie.

La motivation de cette étude n'est pas proprement algorithmique: savoir qu'on peut remplacer $1 + \epsilon$ par 1 n'est pas si intéressant en soi. Je pense par contre que c'est un pas vers une meilleure compréhension de l'algorithme. Les méthodes classiques qui permettent de majorer le nombre maximum d'itérations évitent de rentrer dans le détail de l'algorithme. Elles sont basées sur la décroissance d'une quantité entière pendant le déroulement de l'algorithme. Afin d'élucider le mystère "$t = 1$", j'ai décrit précisément le déroulement des phases finales de l'algorithme et envisagé tous les cas. Cette méthode (peut-être trop directe) donne une borne explicite en dimension 3 et montre plus généralement, l'existence d'une borne polynomiale en la taille de la base, quand la dimension du réseau est fixée. Mais elle ne me permet pas d'établir la dépendance exacte du nombre maximal d'itérations de l'algorithme LLL optimal en la dimension du réseau. Cette dernière question reste ainsi ouverte et je reste persuadé que la réponse permettrait d'envisager l'algorithme LLL dans une autre optique. (cf. Conclusion)

Le reste de la thèse (Chapitres 2 à 6) emprunte la deuxième direction. Elle consiste en l'introduction et l'étude de trois nouvelles notions de réduction ainsi qu'à leur comparaison avec les réductions classiques. La plupart des variations autour de l'algorithme LLL visent à obtenir des bases de meilleure qualité, et les algorithmes correspondant sont donc a priori plus lents que l'algorithme LLL. Ici, nous cherchons à affaiblir légèrement les conditions de réduction, afin d'obtenir des algorithmes a priori plus rapides, qui construisent cependant des bases de qualité raisonnable. Nous introduisons ces nouvelles notions de réduction dans le Chapitre 2, et définissons les algorithmes de réduction associés. Tout comme l'algorithme LLL, les trois nouvelles réductions, appelées réduction au sens de Schmidt, de Gram et de Gram–Schmidt, manipulent les rapports des normes des orthogonalisés l_a/l_b en exigeant des conditions plus faibles que la condition (1.1) qui dépendent aussi du paramètre s. Les réductions au sens de Gram et au sens de Gram–Schmidt dépendent également d'un autre paramètre réel $\gamma \in]0,1[$ et lorsque $\gamma = 0$, ces réductions coïncident avec la réduction de Lovász. Les algorithmes correspondant ont la même structure que l'algorithme LLL et fournissent ainsi une classe d'algorithmes de réduction.

Dans ce même chapitre, je fais une première étude comparative de ces algorithmes dans le pire des cas, selon deux critères principaux: la qualité de la base de sortie et la complexité des algorithmes. J'obtiens, pour tous ces algorithmes, la même majoration pour le nombre maximum d'itérations. Je montre également que les nouvelles bases réduites gardent, dans le pire des cas, des propriétés euclidiennes très similaires à celles des bases réduites au sens de Lovász. En particulier, les bases réduites au sens de Gram sont constituées de vecteurs assez orthogonaux et assez courts. Ce sont des bases de bonne qualité, un peu plus faible que celle des bases réduites au sens de Lovász, mais bon nombre d'applications se satisfont de cette qualité un peu moindre. Ainsi, cette étude dans le pire des cas ne permet pas de départager ces nouveaux algorithmes des algorithmes classiques.

C'est pourquoi le reste de notre travail est consacré à une étude comparative en moyenne de ces réductions sur des réseaux aléatoires. Cette étude comparative présente deux intérêts principaux. Le premier est d'ordre plus pratique, puisque nous montrons dans le Chapitre 6 que ces nouveaux algorithmes de réduction sont plus efficaces en moyenne que l'algorithme LLL, sans trop grande dégradation des qualités de la base réduite (Chapitre 2). Le deuxième est plus théorique, puisque la comparaison de ces processus de réduction avec l'algorithme LLL permet d'affiner la compréhension de ce dernier.

Dans le Chapitre 3, je définis tout d'abord le modèle probabiliste que je choisis pour l'étude en moyenne: c'est le modèle le plus simple et le plus naturel où les vecteurs de la base d'entrée sont choisis indépendamment et uniformément dans la boule unité de $\mathbb{R}^n$. En considérant une base aléatoire de ce modèle, Daudé et Vallée ont montré [DV94] que les carrés des normes l_a des vecteurs de la base orthogonalisée suivent une loi Beta [Dav]. Dans la Section 2 du Chapitre 3, on rappelle leur résultat qui est à la base de l'étude en moyenne qui suit, à commencer par une première étude des variables aléatoires qui correspondent aux rapports de deux orthogonalisés différents (l_a/l_b). Puis, j'étudie les bases aléatoires de ce modèle en insistant particulièrement sur les paramètres qui mesurent les qualités euclidiennes d'une base: défauts de longueur, défaut d'orthogonalité; j'obtiens des bornes pour les valeurs moyennes de tous ces paramètres. Ensuite, j'introduis une grandeur qui va jouer un rôle essentiel dans la suite, i.e., le *poids* d'une base qui est, par définition, le rapport entre la dimension du réseau engendré par la base et celle de l'espace ambiant. Je montre alors que la qualité euclidienne moyenne d'une base dépend fortement de son poids. Par exemple, si l'on choisit un nombre fixé de vecteurs dans une boule unité de dimension suffisamment élevée, ces vecteurs forment en moyenne une base orthogonale. Le Chapitre 3 met donc à disposition une échelle de bases aléatoires de qualité variable qui permettra par la suite de classer les algorithmes de réduction.

Tout comme les résultats de l'analyse dans le pire des cas, le Chapitre 3 ne permet pas de départager les comportements des différents algorithmes de réduction. En effet, l'étude de la moyenne des variables aléatoires l_a/l_b n'est plus significative lorsque les indices a et b sont tous deux proches de la dimension de l'espace ambiant, car les variables deviennent trop dispersées. Une étude en distribution de ces variables aléatoire s'est alors imposée.

La méthode-clé de cette étude asymptotique en distribution est la méthode de Laplace d'évaluation d'équivalents pour les intégrales et les sommes [DeB81, SF96]. Le Chapitre 4 y est tout entier consacré. La méthode considère des intégrales de la forme

$$I_n := \int_\Delta F(x,n)dx,$$

où la fonction $x \mapsto F(x,n)$ admet un maximum strict et global sur le domaine Δ. (Si le maximum n'est pas global, on peut toujours décomposer le domaine d'intégration). Pour les intégrales d'une variable, la méthode de Laplace donne la partie principale de I_n, quand n tend vers l'infini, pourvu que quelques hypothèses assez strictes sur la fonction F soient réalisées. Lorsque x est une variable de $\mathbb{R}^p$ $(p > 1)$, les seules généralisations de la méthode qui existent à notre connaissance concernent le cas particulier où le maximum de la fonction F est atteint à l'intérieur du domaine Δ. Or, justement, notre analyse nécessite l'étude de cas plus généraux, où le maximum est atteint sur le bord du domaine Δ (et même souvent en un "coin" de Δ). Dans les cas adaptés à mes besoins, j'obtiens des généralisations précises de la méthode de Laplace à des intégrales doubles (Propositions 8, 9), qui vont permettre de mener à bien les études asymptotiques probabilistes des chapitres suivants.

Dans le Chapitre 5, j'établis le comportement asymptotique de la fonction de répartition et de la densité des variables aléatoires l_a/l_b, quand a et b sont des expressions affines de la dimension de l'espace ambiant. Le choix des expressions affines pour ces indices est justifié puisqu'il permet de mettre en évidence des phénomènes asymptotiques différents selon la localisation des indices a et b, *i.e.* selon les valeurs α, β des pentes de a et de b. Je montre (Théorème 9), que pour les *indices du début*, dont les pentes sont strictement inférieures à 1, la variable aléatoire l_a/l_b, admet asymptotiquement une loi $0-1$, avec une vitesse de convergence exponentielle en la dimension l'espace ambiant. Pour les indices de la fin, de pente égale à 1, la loi asymptotique de la variable l_a/l_b, ne présente plus de discontinuité comme précédemment. Les résultats de ce Chapitre 5 constituent un pas supplémentaire dans la connaissance des propriétés des bases aléatoires du modèle uniforme. Cette étude est essentielle pour la suite. D'une part, elle confirme de façon précise à quel point la moyenne est significative, quand a et b sont des indices du début. D'autre part, pour les indices de la fin, là ou la moyenne n'est plus très significative, elle fournit les probabilités–limites,

$$\lim_{n\to\infty} \Pr\{l_a/l_b < v\}.$$

Le Chapitre 6 constitue la conclusion de l'étude comparative. Dans le cadre du modèle probabiliste choisi, je compare les performances des différents algorithmes de réduction (les algorithmes nouveaux du Chapitre 2 et les algorithmes classiques) sur les bases aléatoires du modèle choisi et je montre que le poids de la base joue un rôle important (le poids est le rapport entre la dimension de la base et celle de l'espace ambiant). Je mets en lumière deux phénomènes complémentaires:

– lorsque le poids est inférieur à un seuil qui dépend du paramètre d'approximation s et que la dimension de l'espace ambiant est assez élevée, alors tous les algorithmes de réduction sont triviaux, puisqu'ils se contentent presque-sûrement de vérifier que les conditions de réductions sont satisfaites. Autrement dit, sur une telle base d'entrée, le nombre d'itérations est quasi-sûrement minimal: pour un réseau de dimension p, ce nombre est $p-1$.

– sur des bases aléatoires *pleines*, i.e. de dimension égale à celle de l'espace ambiant n, les différents algorithmes de réduction se comportent différemment. Les nouveaux algorithmes introduits dans cette thèse gardent le comportement décrit ci-dessus: lorsque n est assez grand, quasi-sûrement le nombre d'itérations est de nouveau minimal (ici $n-1$). En comparaison, j'évalue la probabilité asymptotique qu'une base aléatoire pleine (rendue préalablement propre) soit réduite au sens de Lovász. Les majorants sont fonctions du paramètre d'approximation s et restent petits pour les valeurs habituellement choisis de s. Par exemple, lorsque s et t sont reliés par (1.1), je montre qu'avec une probabilité supérieure à $(1+1/s^2)^{-1/2}$, l'algorithme LLL(t) effectue un nombre d'itérations strictement supérieur à $n-1$.

Ainsi, je réussi à séparer les différents algorithmes de réduction en deux groupes, selon les comportements sur les base aléatoires pleines du modèle uniforme, alors que ni l'étude dans le pire des cas, ni une étude en moyenne grossière ne le permettait.

Enfin, les expérimentations décrites dans le Chapitre 8 confirment d'une part de façon frappante, les résultats de notre étude en moyenne, dans le modèle que nous avons choisi. D'autre part, elles semblent montrer que nos résultats sur les comportements relatifs des différents algorithmes de réduction restent valable dans d'autres modèles, comme le modèle des “réseaux sac–à–dos”.

En conclusion, cette thèse propose donc d'une part des algorithmes de réduction plus rapides en moyenne. D'autre part, elle précise l'analyse, dans le pire des cas et en moyenne, de l'algorithme LLL.

Chapitre 2

Introduction de nouvelles réductions

Ce chapitre introduit tout d'abord les outils de base de la théorie algorithmique des réseaux. Cette introduction sera volontairement brève et on n'y donnera pas les preuves très classiques. D'excellentes références y sont déjà consacrées, tant du point de vue purement mathématique [Cas71, LG87, Sie89, New72] que d'un point de vue plus algorithmique [Kan87, Val89b, Sch89].

Puis après un survol de différentes notions de réductions déjà existantes, trois nouvelles classes de réductions originales sont introduites. Les nouvelles réductions que nous introduisons sont a priori *moins fortes* que celles qui existent déjà. Elles sont obtenues en affaiblissant les conditions exigées par la réduction classique de Lovász (LLL).

Par la suite, nous comparons les qualités des nouvelles base réduites et nous montrons que certaines parmi elles gardent des propriétés similaires aux bases réduites au sens de Lovász.

Enfin nous fournissons et analysons dans le pire des cas, les algorithmes qui construisent ces bases réduites. Pour le pire des cas, nous obtenons la même borne pour toutes les réductions étudiées (la borne habituelle pour le nombre maximum d'itérations de l'algorithme LLL).

Pour les applications, l'intérêt de ces nouvelles réductions viendra de l'analyse en moyenne. Le Chapitre 6 montrera en effet que les nouveaux algorithmes de réductions sont plus rapides en moyenne que l'algorithme LLL classique, alors que nous montrons ici (Section 2.3) que les bases de sortie sont de qualités similaires.

Par ailleurs, les réductions que nous introduisons restent proches de la réduction effectuée par l'algorithme LLL [LLL82]. C'est pourquoi notre analyse permet aussi de mieux comprendre le comportement en moyenne de l'algorithme LLL.

2.1 La réduction des réseaux euclidiens

Cette section formalise les objets manipulés dans cette thèse. Les outils nécessaires à la comparaison des différentes bases réduites et à l'analyse des algorithmes de réductions y sont introduits.

2.1.1 Les réseaux euclidiens.

L'espace $\mathbb{R}^n$ est muni de sa structure euclidienne canonique. Un *réseau euclidien* est un sous–groupe discret de $\mathbb{R}^n$. *La dimension* du réseau est alors la dimension de l'espace vectoriel de $\mathbb{R}^n$ engendré par le réseau. On montre qu'un tel réseau L possède toujours *une base*, c'est à dire une famille libre $b = (b_1, \ldots, b_p)$, tel que le réseau L soit l'ensemble des combinaisons linéaires entières des vecteurs de b.

Réciproquement, on considère une famille libre b de p vecteurs $(b_1, \ldots, b_p)$, $(p \leq n)$ de $\mathbb{R}^n$. L'ensemble des combinaisons linéaires entières de ces p vecteurs est un sous-groupe discret de $\mathbb{R}^n$, donc un réseau euclidien de dimension p, que l'on note $L(b)$:

$$L(b) := \left\{ \sum_{i=1}^{p} x_i b_i : \quad (x_1, \ldots, x_n) \in \mathbb{Z}^n \right\}.$$

Cette deuxième façon de définir un réseau, moins intrinsèque, est algorithmiquement plus utile. Dans la suite, un réseau sera toujours donné par une base. Il nous arrivera de confondre le réseau avec une de ses bases et même par abus de langage, de parler de la dimension d'une base au lieu de son cardinal ou encore de la dimension du réseau qu'elle engendre.

Tout réseau possède une infinité de bases : b et b' sont des bases d'un même réseau si et seulement si la matrice de passage de b à b' est une matrice unimodulaire, i.e. une matrice à coefficients entiers et dont la valeur absolue du déterminant est égale à 1. Il est clair que toutes les bases d'un réseau ont le même nombre de vecteurs et ce nombre est bien la dimension du réseau engendré par ces bases. L'espace $\mathbb{R}^n$ est appelé *l'espace ambiant.*

2.1.2 Les invariants d'un réseau.

On appelle *invariant* d'un réseau toute quantité définie dans un réseau, qui dépend uniquement du réseau et non du choix de la base. La dimension d'un réseau en est un premier exemple. Deux autres types d'invariants sont particulièrement utiles dans la suite.

1. *Le déterminant* du réseau, noté $d(L)$. C'est le volume du parallélépipède p-dimensionnel construit sur une base quelconque du réseau. On calcule facilement le déterminant à partir du déterminant de Gram d'une base d'un réseau[1],

$$d(L) := |\det(b_1, \ldots, b_n)| = (\det[\langle b_i, b_j \rangle]_{1 \leq i,j \leq n})^{\frac{1}{2}}.$$

2. *les minima successifs* du réseau. Pour un réseau L de $\mathbb{R}^n$ de dimension p, le $i-$ème minimum successif $\Lambda_i(L)$ est le plus petit réel positif t pour lequel il existe i vecteurs linéairement indépendants de L de normes inférieures à t.
   Il est clair que les p nombres $\Lambda_i(L)$ forment une suite croissante. Par définition, il existe un ensemble, non nécessairement unique, de p vecteurs linéairement indépendants de L, appelés aussi minima successifs, notés $\lambda_i(L)$ vérifiant:

$$\text{pour tout } k, 1 \leq k \leq n, \quad |\lambda_k(L)| = \Lambda_k(L).$$

   En particulier, $\lambda_1(L)$ désigne un plus court vecteur non nul de L.

1. La notation $\langle b_i, b_j \rangle$ désigne simplement le produit scalaire usuel de b_i et b_j.

2.1.3 La qualité d'une base et les critères pour la mesurer

Certaines bases d'un réseau sont plus "naturelles" que d'autres: ce sont les bases "assez orthogonales.". Presque toujours, aucune base d'un réseau n'est orthogonale, mais il est important de disposer d'une, dont les vecteurs sont "assez orthogonaux " et en même temps "assez courts".
Pour une base donnée d'un réseau, son *défaut d'orthogonalité* et ses *défauts de longueur* formalisent et mesurent sa qualité du point de vue intuitif expliqué ci-dessus. Soit $b = (b_1, \ldots, b_p)$ une base d'un réseau.

Le *défaut d'orthogonalité* $\rho(b)$, de la base b est le rapport

$$\rho(b) := \frac{\prod_{i=1}^{p} \mid b_i \mid}{d(L)}. \tag{2.1}$$

Le défaut d'orthogonalité compare la base à une base orthogonale: si la base b est orthogonale, alors $d(L) = \prod_{i=1}^{n} \mid b_i \mid$ et donc $\rho(b) = 1$. Sinon, $\rho(b) > 1$ [2]

Le i-éme *défaut de longueur* $\mu_i(b)$ de la base b est le rapport

$$\mu_i(b) := \frac{\mid b_i \mid}{\Lambda_i(L)}. \tag{2.2}$$

Les défauts de longueur comparent une base à une base formée des minima successifs du réseau. Les minima successifs n'engendrent pas nécessairement le réseau, lorsque la dimension dépasse 4. Si un réseau possède une base formée de ces minima successifs, alors tous les défauts de longueur de cette base sont égaux à 1.
Si le premier vecteur de la base b est un plus court vecteur du réseau, alors son premier défaut de longueur est égal à 1. Sinon, $\mu_1(b) > 1$.

Les deux notions de défaut d'orthogonalité et de défauts de longueur sont liées en vertu des résultats d'Hermite et de Minkowski, ce qui montre heureusement que nos deux exigences pour une "bonne base" ne sont pas contradictoires. Hermite a prouvé l'existence d'une constante γ_n appelée constante d'Hermite, définie par:

$$\gamma_n \;=\; \max\{\frac{\Lambda_1(L)}{d(L)^{\frac{1}{n}}} \;/\; L \;\; \text{réseau de rang}\, n\}.$$

Seules les 8 premières valeurs de γ_n sont exactement connues. Pour les autres valeurs, il existe des majorations. Par ailleurs, Minkowski a montré également que

$$d(L) \;\leq \prod\nolimits_{i=1}^{n} \Lambda_i(L) \leq \gamma_n^n \; d(L),$$

ce qui permet de relier défauts de longueur et le défaut d'orthogonalité,

$$1 \;\leq \frac{\rho(b)}{\prod_{i=1}^{n} \mu_i(b)} \leq \gamma_n^n.$$

2. L'écriture 2.4 du défaut d'orthogonalité justifie cette minoration triviale.

On considère un réseau défini par une de ses bases. Les algorithmes qui construisent une base de bonne qualité commencent tous par *orthogonaliser* cette base par le *procédé de Schmidt*. Le résultat de ce procédé est une base orthogonale qui génère le même espace vectoriel que la base b, mais qui n'est pas en général une base du réseau engendré par la base b. Plus précisément, ce procédé associe à une base $b = (b_1, b_2, ..., b_p)$ la base orthogonale $b^* = (b_1^*, b_2^*, ...b_p^*)$ et la matrice $m = (m_{ij})$ qui exprime la base b dans b^*, en opérant comme suit :

$$\begin{aligned} b_1^* &= b_1 \\ b_i^* &= b_i - \sum_{j<i} m_{ij} b_j^* \quad \text{où } m_{ij} \text{ est défini par le relation} \quad m_{ij} = \frac{(b_i \mid b_j^*)}{\mid b_j^* \mid^2}, \end{aligned}$$

ce qui permet de calculer facilement les b_i^* et les m_{ij} par récurrence sur i. La matrice m est donc une matrice triangulaire inférieure possédant une diagonale de 1.

Notation 1. Dans la suite de cette thèse, la base orthogonalisée de la base b sera notée toujours $b^* = (b_1^*, b_2^*, ..., b_n^*)$. La matrice de passage de b à b^* sera notée $m = (m_{ij})$ et la quantité l_i désignera la norme du $i$ème orthogonalisé:

$$l_i = \mid b_i^* \mid .$$

Dans l'étude des problèmes liés aux réseaux, les grandeurs l_i sont d'une importance primordiale. Par exemple, le déterminant $d(L)$ du réseau s'exprime simplement en fonction des l_i,

$$d(L) = \prod_{i=1}^{p} l_i. \tag{2.3}$$

Il en est de même pour le défaut d'orthogonalité (2.1) d'une base b,

$$\rho(b) = \prod_{i=1}^{p} \frac{\mid b_i \mid}{l_i}. \tag{2.4}$$

La relation précédente justifie le nom de défaut d'orthogonalité, en remarquant que le rapport $l_i / \mid b_i \mid$ est le sinus de l'angle θ_i entre le vecteur b_i et le sous espace engendré par les $i-1$ premiers vecteurs de la base. Plus ces sinus sont proches de 1, plus la base est proche d'une base orthogonale.

2.1.4 La taille d'une base d'entrée

Pour évaluer l'efficacité en temps des algorithmes qui construisent de bonnes bases dans un réseau, il est nécessaire de préciser la taille de la donnée.

Le réseau L de dimension p sera toujours donné par une base quelconque $b = (b_1, b_2, ..., b_p)$ exprimée dans la base canonique de $\mathbb{R}^n$. Les vecteurs de la base b sont a priori à coordonnées rationnelles. Pour simplifier, on travaillera toujours dans des réseaux entiers, i.e. tous les vecteurs du réseau sont à coordonnées entières. Les raisonnements se généralisent facilement à des réseaux pseudo-entiers (à déterminant entier).

Si l'on désigne par M le maximum des $\mid b_i \mid$ et par b_{ij}, les coordonnées du i-ème vecteur b_i de la base d'entrée, il est clair que quelque soit la norme choisie (euclidienne ou infinie),

$$\forall (i,j) \in \{1..n\}^2, \quad \mid b_{ij} \mid \leq \mid b_i \mid \leq M.$$

Chaque vecteur de la base d'entrée est alors codé avec un nombre de bits en $O(n \log M)$ et la base initiale a une taille en $O(np \log M)$, qui correspond à la taille de la donnée.
Il existe donc trois paramètres d'entrée à considérer:

1. La dimension de l'espace ambiant n.
2. La dimension du réseau p.
3. La taille de la base initiale, M.

Un cas particulier important qui arrive fréquemment dans les applications est lorsque la dimension p du réseau est la même que celle de l'espace ambiant n. Quand un algorithme qui travaille sur de tels réseaux est polynomial en la taille de son entrée, il est d'une part polynomial en $\log M$, où $M = \max_{(i,j)} |b_{i,j}|$, pour la dimension n fixée, et d'autre part polynomial en n.

2.1.5 La réduction d'une base: point de vue algorithmique

Réduire un réseau consiste à y trouver une *bonne base* appelée aussi *une base réduite*. La réduction des réseaux, qui est aussi le propos général de cette thèse est un problème ancien. Qu'il s'agisse de la recherche d'un vecteur court dans un réseau ou de manière équivalente la recherche des minima entiers d'une forme quadratique, la question a déjà intéressé bon nombre de mathématiciens. La *géométrie des nombres* qui désigne le domaine est née avec les travaux de Minkowski [Min91, Min10]. Ses deux théorèmes fondamentaux exhibent des conditions suffisantes sur le volume d'un ensemble de $\mathbb{R}^n$, convexe et symétrique par rapport à l'origine, pour qu'il contienne des points à coordonnées entières. Comme un réseau est toujours l'image de $\mathbb{Z}^p$ par une transformation linéaire bijective τ, y rechercher un plus court vecteur est exactement équivalent à une condition minimale sur le volume d'un ellipsoïde ($\tau^{-1}(\mathcal{B}_p)$, où $\mathcal{B}_p$ est la boule unité de $\mathbb{R}^p$) contenant un point à coordonnées entières, différent de l'origine.
Par la suite, Minkowski introduit une notion de base réduite (qui porte son nom), comme étant la base formée des vecteurs les plus courts possibles d'un réseau qui engendrent exactement tout le réseau: $(b_1, \ldots, b_p)$ est réduite au sens de Minkowski si pour tout $1 \leq i \leq p$, b_i est un plus court vecteur qui forme avec $(b_1, \ldots, b_{i-1})$ un système libre qui puisse se compléter en une base du réseau.

Alors que Minkowski privilégiait la recherche de vecteurs courts, Korkine et Zoletarev cherchaient surtout une base proche (la plus proche possible à ce jour [LLS86, Sch87, Sch94]) d'une base orthogonale. Précisons qu'en général, il n'existe pas de base orthogonale dans un réseau et que les exigences sur les normes des vecteurs et l'orthogonalité de la base sont reliées, notamment grâce aux théorèmes de Minkowski. Ces remarques étant faites, une base $(b_1, \ldots, b_p)$ est réduite au sens de Korkine-Zolotarev si elle vérifie deux conditions. Premièrement, parmi tous les vecteurs v tels que le système libre $(b_1, \ldots, b_{i-1}, v)$ puisse être complété en une base du réseau, le vecteur b_i est choisi de telle sorte que sa projection orthogonalement au sous-espace engendré par $(b_1, \ldots, b_{i-1})$ soit la plus courte possible. Deuxièmement, la base doit être *propre*. De façon informelle, les coefficients de la matrice de passage de la base b vers son orthogonalisée de Gram-Schmidt b^*, doivent être petits. Une définition précise de la propreté (2.5) est rappelée dans la prochaine section.

Au premier abord, la réduction au sens de Korkine-Zolotarev paraît moins intuitive que celle de Minkowski. Elle s'est pourtant avérée plus intéressante pour ce qui est des problèmes algorithmiques. C'est en effet sa version *locale* qui a conduit Lenstra, Lenstra et Lovász en 82 [LLL82] à introduire pour la première fois, un algorithme efficace de réduction en dimension quelconque. Leur célèbre algorithme LLL est le point de départ de cette thèse.

Avant de parler de l'algorithme LLL, disons simplement qu'au début du siècle, on ne s'occupait guère des problèmes de construction de ces bases, ni a fortiori de la rapidité de telles constructions. Seul Gauss a donné un algorithme efficace en petite dimension. Il cherchait les minima des formes quadratiques binaires et ternaires. L'algorithme de Gauss, maintenant connu sous de nombreux aspects algorithmiques [Lag82, Val91, DFV97], sera détaillé dans cette thèse (Chapitres 2 et 6) car il constitue la base des algorithmes de réductions efficaces en dimension quelconque. Sinon, les seuls algorithmes connus pour construire les bases réduites citées, [HJ85] pour les bases réduites de Minkowski et [Kan83] pour celles de Korkine–Zolotarev, sont exponentiels en la taille de l'entrée. Lorsqu'un réseau est donné par une base quelconque, il n'existe à ce jour aucun algorithme *assez rapide* (polynomial en la taille de l'entrée) pour construire des bases de si bonne qualité.

D'ailleurs, de tels algorithmes polynomiaux n'existent probablement pas: les deux réductions citées exigent en particulier que le premier vecteur de la base réduite soit un plus court vecteur du réseau. Van Emde Boas [vEB81] a montré en 1981 que la recherche d'un vecteur le plus court dans un réseau, connu sous le nom de *Shortest Vector*, est NP-complet, pour la norme infinie. Mais sa preuve échoue pour la norme euclidienne. Beaucoup plus récemment, en 1997, Ajtai [Ajt97] a montré que le problème est NP-complet, par réduction polynomiale non–déterministe. Pour montrer la NP-complétude, Ajtai a besoin d'une conjecture non–triviale de la théorie des nombres. Précisons pour terminer, que d'une part *Shortest Vector* est communément admis comme étant NP-complet et d'autre part que le problème plus difficile de construire une base réduite de Minkowski, est lui, NP-complet.

Les derniers travaux d'Ajtai ont renouvelé l'intérêt pour les études de complexités théoriques des problèmes liés aux réseaux. Par exemple, les problèmes d'approximation liés aux réseaux sont amplement étudiés actuellement [Ajt97, ABSS97, GG98, Mic98, BS98].

La réduction des réseaux a de nombreuses applications en théorie algorithmique des nombres et en cryptographie et la qualité exigée pour la base réduite dépend de l'application considérée. Mais les applications n'ont en général pas besoin de bases *aussi fortement* réduites que celles de Minkowski ou de Korkine-Zolotarev. Pour un paramètre réel $t \in [1, 2]$, l'algorithme LLL(t) est un algorithme d'approximation qui au terme d'une phase itérative construit une base réduite au sens de t-Lovász: c'est une base $(b_1, \ldots, b_p)$ qui est propre et dont la projection de chaque b_i orthogonalement à $(b_1, \ldots, b_{i-1})$ est ,– non le plus court possible, comme c'est la cas pour une base réduite de Korkine-Zolotarev –, mais plus court, à un facteur t près, que la projection de b_{i+1} orthogonalement à $(b_1, \ldots, b_{i-1})$. L'idée de cette double approximation, passage à une condition locale, puis l'introduction du paramètre d'approximation t était, du point de vue constructif, une *idée–clé*, puisque l'algorithme LLL(t) obtient en temps polynomial une base réduite au sens de t-Lovász. Ainsi, pour la première fois, la résolution effective de problèmes très variés en théorie algorithmique des nombres ou en cryptographie devint possible: factorisation de polynômes à coefficients entiers[LLL82, Sch84], approximations diophantiennes[Lag80], programmation linéaire entière [Len83, Kan83], attaques intelligentes de différentes cryptosystèmes [Lag83, Sch95, VGT88], etc. Il est important de noter que la plupart de ces applications utilisent uniquement un seul vecteur assez court du réseau.

Nous allons définir dans cette thèse d'autres réductions, a priori moins fortes encore. Nous montrons toutefois que les nouvelles bases réduites jouissent de propriétés similaires aux bases réduites au sens de Lovász, tout en étant obtenues en moyenne beaucoup plus rapidement (chapitre 6). Toutes ces réductions sont définies précisément dans la prochaine section.

2.2 Présentation de différentes notions de réduction

Dans cette section, nous définissons trois nouvelles classes de réduction, appelées réduction au sens de Schmidt, au sens de Gram et au sens de Gram–Schmidt. Ces nouvelles réductions sont obtenues en relaxant les conditions exigées pour une base réduite au sens de Lovász. En conséquence,

1. Les nouvelles bases réduites ont des qualités similaires aux bases réduites au sens de Lovász, comme le précisera la Section 2.3.
2. Les algorithmes de réduction gardent la même structure que l'algorithme LLL. Seuls les tests de sortie sont remplacés par des tests moins forts. Nous montrons alors que le nombre maximum d'itérations de tous ces algorithmes de réduction est borné polynomialement, exactement de la même façon que dans le cas de l'algorithme LLL. (Section 2.4).
3. En moyenne, les nouveaux algorithmes de réduction sont plus rapides. (Chapitre 6)

En dimension 2, toutes les notions de réductions (y compris celles de Minkowski et de Korkine-Zolotarev) sont les mêmes et il existe un algorithme dû à Gauss, qui permet d'obtenir la base réduite en temps polynomial. Un bref rappel de cet algorithme permet de mieux comprendre les réductions ultérieures qui le généralisent à une dimension quelconque.
Après une description succincte de la réduction en dimension 2, cette section définit quelques différentes bases réduites.

2.2.1 La réduction en dimension 2 : l'algorithme de Gauss

C'est une généralisation en dimension 2 de l'algorithme d'Euclide centré (construction du p.g.c.d. de 2 entiers par une suite de divisions euclidiennes centrées.) Il permet de trouver en temps polynomial, une base réduite d'un réseau de dimension 2. Dans un tel réseau, la base réduite est formée par des vecteurs minima successifs. Voici le schéma de l'algorithme de Gauss:

Algorithme de Gauss

Donnée : une base (b_1, b_2) d'un réseau $\mathcal{L}$.
Résultat : une base minimale du réseau $\mathcal{L}$.

Répéter

- **Échanger** b_1 et b_2 pour avoir $|b_1| \leq |b_2|$.
- **Translater** b_2 parallèlement à b_1 pour le rendre le plus court possible:
  Plus précisément, m étant la matrice de passage exprimant la base réduite b, dans sa base orthogonalisée b^*,
  $m = \begin{pmatrix} 1 & 0 \\ m_{2,1} & 1 \end{pmatrix}$, choisir l'entier r le plus proche de $m_{2,1}$ et faire $b_2 := b_2 - r\ b_1$.

jusqu'à ce que $\mid b_2 \mid \geq \mid b_1 \mid$.

L'étape de translation de cet algorithme se généralise simplement à une dimension quelconque, comme le montre la paragraphe suivant.

2.2.2 La *propreté* d'une base

Pour réduire une base, il paraît naturel d'essayer de diminuer les coefficients de la matrice m, sans que le réseau $L(b)$ et la base b^* associée soient modifiés. Ainsi, la base b sera "assez proche" de la base orthogonalisée b^* associée. Le principe de cette opération consiste à généraliser la phase de translation de l'algorithme de Gauss à des dimensions plus élevées. L'algorithme de Gauss translate un des vecteurs de la base parallèlement à l'autre pour l'écourter au maximum. Plus généralement, on pourra translater chaque vecteur b_i parallèlement à tous les b_j, pour $j < i$, de telle sorte que tous les coefficients m_{ij} de la matrice m, soient en valeur absolue inférieurs à $1/2$, pour $j < i$.

Une base b de dimension p est *propre* si la matrice de passage m de b à son orthogonalisée b^* vérifie

$$\mid m_{ij} \mid \leq \frac{1}{2}, \text{ pour } 1 \leq j < i \leq p. \tag{2.5}$$

L'algorithme Propre est une suite d'itérations de la phase translation de l'algorithme de Gauss et rend ainsi une base propre.

Algorithme Propre

Donnée : une base $b = (b_1, \ldots, b_p)$ d'un réseau $\mathcal{L}$.
Résultat : une base propre du réseau $\mathcal{L}$.

Pour i allant de 2 à p faire
  Pour j allant de $i - 1$ à 1 faire
    $r_j :=$ entier le plus proche de m_{ij};
    $b_i := b_i - r_j b_j$;

Mais une base propre n'est pas forcément presque orthogonale : il faut non seulement majorer les projections des b_{i+1} sur les sous-espaces H_i formés par les $\{b_j \ : \ j \leq i\}$, (opération qui rend une base propre), mais aussi minorer les angles θ_i entre les b_i et les H_i. Il existe des conditions qui permettent de minorer cet angle et donc de majorer le défaut d'orthogonalité. Ces conditions aboutissent à la construction de bases de qualités variables, selon les majorations qu'elles entraînent sur les défauts de longueur ou sur les défauts d'orthogonalité.

2.2.3 Définitions de quelques bases réduites

Pour assurer la bonne qualité des bases réduites, toutes les notions de réductions introduites imposent à la base réduite d'être propre. Le paragraphe précédent a présenté un algorithme qui rend rapidement une base quelconque propre. Pour être réduite, une base devra toujours satisfaire une deuxième condition, plus difficile à réaliser que la propreté et qui va distinguer les différentes notions de réductions, les unes des autres. Ces deuxièmes conditions vont faire essentiellement intervenir les rapports entre les normes de deux orthogonalisés. On rappelle que l_i désigne la norme du i-ème orthogonalisé.

Définition 1. Soient deux paramètres réels s et t reliés par le relation suivante,

$$s^2 = \frac{1}{t^2} - \frac{1}{4} \text{ avec } s^2 \leq \frac{3}{4} \text{ ou encore } t > 1. \tag{2.6}$$

Soient une base $b = (b_1, b_2, \ldots, b_p)$ d'un réseau L, un indice fixé $k \in \{1, \ldots, p\}$ et un paramètre réel $\gamma \in]0, 1[$. Pour chaque indice i, $1 \leq i \leq p-1$, nous considérons les conditions suivantes [3]

$$t\text{-Lovász} \quad : \quad \left(l_{i+1}^2 + m_{i+1,i}^2 l_i^2\right) / l_i^2 \ \geq \ 1/t^2 \tag{2.7}$$

$$s\text{-Siegel} \quad : \quad l_{i+1}/l_i \ \geq \ s \tag{2.8}$$

$$(s,\gamma)\text{-Gram} \quad : \quad l_{i+1}/l_i \ \geq \ s^{i^\gamma} \tag{2.9}$$

$$(s,k)\text{-Schmidt} \quad : \quad \begin{cases} \text{pour } i > k, & l_i/l_k \geq s^{i-k}, \\ \text{et} & \\ \text{pour } i < k, & l_k/l_i \geq s^{k-i} \end{cases} \tag{2.10}$$

$$(s,k,\gamma)\text{-Gram-Schmidt} \quad : \quad (2.9) \quad \text{et} \quad (2.10). \tag{2.11}$$

Soit $\mathcal{C}_i$ une des conditions précédentes pour un indice i fixé. La base b est $\mathcal{C}$-réduite si elle est propre et si elle satisfait la condition $\mathcal{C}_i$, pour tout i, tel que $1 \leq i \leq p-1$.

La condition (2.7) est introduite par Lovász [LLL82]. Elle est utilisée dans la version originale de l'algorithme LLL. La condition s-Siegel est une conséquence immédiate de (2.7) et de la propreté (2.5).
En effet (2.7) équivaut à

$$(l_{i+1}/l_i)^2 \geq 1/t^2 - m_{i+1,i}^2$$

3. Pour les conditions de (s,k)-Schmidt l'indice i varie dans $\{1, \ldots, p\}$.

Si la base est propre $m_{i+1,i}^2 < 1/4$ et la relation (2.6) conduit alors à la condition de s-Siegel. Dans les raisonnements, on utilise presque toujours la condition de s-Siegel à la place de celle de t-Lovász.

L'analyse en moyenne (Chapitre 3) de la variable aléatoire (l_{i+1}/l_i) dans un modèle bien choisi montre qu'en moyenne les rapports (l_{i+1}/l_i) tendent tous vers 1 avec la dimension n de l'espace ambiant,

$$\lim_{n\to\infty} \mathrm{E}_n[l_{i+1}/l_i] = 1.$$

Mais leur variance est une fonction croissante de i. Elle reste en $\mathrm{O}(1/n)$, quand i est de le forme αn avec $\alpha < 1$. Mais lorsque i tend vers n, l'écart-type $\sigma_n[l_{i+1}/l_i]$ devient du même ordre de grandeur que la moyenne. Les rapports (l_{i+1}/l_i) sont alors de plus en plus dispersés, et les conditions de Siegel ont d'autant moins de chance d'être satisfaites. Il semble donc plus raisonnable d'exiger pour le rapport (l_{i+1}/l_i) une condition qui s'affaiblit, quand l'indice i grandit. On peut par exemple, exiger un minorant qui décroît avec l'indice i. C'est dans cet esprit que nous avons introduit ci-dessus, la condition de (s,γ)-Gram. C'est une condition plus faible que la condition de s-Siegel quand $\gamma > 0$ qui se confond exactement avec elle lorsque $\gamma = 0$.

La condition de (s,k)-Schmidt est elle aussi plus faible que celle de s-Siegel. Alors qu'avec (s,γ)-Gram, la condition de s-Siegel est relaxée pour tous les indices i, avec (s,k)-Schmidt, on décide de garder (exactement) les conséquences de s-Siegel seulement pour un vecteur fixé de la base: le k-ème. On remarque que la condition de (s,k)-Schmidt peut être notée de façon plus résumée par

$$\frac{l_{\max(i,k)}}{l_{\min(i,k)}} \geq s^{|i-k|}. \tag{2.12}$$

Il sera souvent intéressant de fixer k à 1.

Enfin la condition composée (s,k,γ)-Gram-Schmidt est un compromis des deux précédentes. On garde (exactement) les conséquences de s-Siegel pour le k-ème vecteur de la base et ces conséquences sont fortement relaxées au fur et à mesure que l'indice i s'éloigne de l'indice fixé k. Les conditions de (s,k)-Schmidt et s-Siegel ne peuvent pas être comparées pour un indice i fixé. Toutefois, une base b est s-Siegel réduite si et seulement si elle est (s,k)-Schmidt réduite pour tout indice $k \in \{1,\dots p\}$. Le lemme suivant compare les conditions précédentes *plus localement*. Il sera particulièrement utile lors de l'étude des algorithmes de réductions.

Lemme 1. *Soient deux paramètres réels t, s reliés par la relation (2.6). Soit une base de cardinal p, $(b_1,\dots,b_p)$. Pour un indice fixé $i \in \{1,\dots,p-1\}$,*

1. *si la condition de s-Siegel n'est pas satisfaite et $|m_{i+1,i}| \leq 1/2$, alors la condition de t-Lovász ne l'est pas non plus.*
2. *si la condition de (s,γ)-Gram n'est pas satisfaite, alors la condition de s-Siegel ne l'est pas non plus.*
3. *pour $i \geq k+1$, si la condition de (s,k)-Schmidt est vérifiée pour un indice i mais pas pour un indice $i+1$, alors la condition de s-Siegel n'est pas vérifiée pour l'indice i.*
4. *pour $i \leq k-1$, si la condition de (s,k)-Schmidt est vérifiée pour un indice $i+1$ mais pas pour un indice i, alors la condition de s-Siegel n'est pas vérifiée pour l'indice i.*

Preuve. Les points 1. et 2. découlent directement de définitions.
3. Pour $i \geq k+1$, si une base est (s,k)-Schmidt réduite pour un indice i mais pas pour un indice $i+1$, cela signifie:

$$l_i \geq s^{i-k}\, l_k \quad \text{et} \quad l_{i+1} < s^{i+1-k}\, l_k.$$

Donc $l_{i+1} < sl_i$, et la condition de s-Siegel n'est pas vérifiée pour l'indice i.

4. De même, pour $i \leq k-1$, si une base est (s,k)-Schmidt réduite pour un indice $i+1$ mais pas pour un indice i, alors

$$l_k \geq s^{k-i-1} l_{i+1} \quad \text{et} \quad l_k < s^{k-i} l_i.$$

D'où, là encore $l_{i+1} < sl_i$, ce qui signifie que la condition de s-Siegel n'est pas vérifiée pour l'indice i.
Remarquons pour finir que pour $i = k$, la condition de Schmidt est systématiquement satisfaite et pour $i \in \{k-1, k+1\}$, les conditions de Schmidt et de Siegel se confondent. □

2.3 Étude comparative des bases réduites

Le but de cette section est de classer les réductions définies précédemment en considérant la qualité des différentes bases réduites. À cet effet, deux lemmes seront utiles. Le Lemme 2 est un résultat classique et très utile qui relie des grandeurs invariantes, les minima successifs Λ_i d'un réseau aux longueurs des orthogonalisés d'une base quelconque.

Lemme 2. *Soit b une base quelconque d'un réseau de dimension p. Alors, avec les notations habituelles,*

$$\Lambda_1(L) \;\geq\; \min_{1 \leq i \leq p} l_i,$$

et plus généralement pour tout $1 \leq i \leq p$,

$$\Lambda_i(L) \;\geq\; \min_{i \leq j \leq p} l_j.$$

Preuve. Tout vecteur non nul x du réseau se décompose dans la base b sous la forme

$$x := \textstyle\sum_{i=1}^{i_0}\ \alpha_i\ b_i,$$

où $i_0 > 0$ est le dernier indice i pour lequel $\alpha_i \neq 0$.
Le vecteur x s'exprime aussi dans la base orthogonale b^* et par construction de cette base,

$$x \;=\; \sum_{i \leq i_0} \alpha_i \left(\sum_{j \leq i} m_{i,j}\ b_j^* \right) = \sum_{i \leq i_0} \alpha'_i b_i^*, \quad \text{avec} \quad \alpha'_{i_0} = \alpha_{i_0}.$$

Comme α_{i_0} est un entier non nul,$|\alpha_{i_0}| \geq 1$ et $|x| \geq |b_{i_0}^*| = l_{i_0}$ et c'est vrai en particulier pour un plus court vecteur non nul du réseau.

Plus généralement, soient i vecteurs indépendants $x_1, x_2, ..., x_i$ du réseau qui se décomposent sous la forme

$$x_j \;=\; \sum_{k=1}^{p} \alpha_{j,k}\; b_k, \quad \text{pour} \quad 1 \leq j \leq i.$$

Si $k(j)$ désigne le plus grand indice k pour lequel $\alpha_{j,k} \neq 0$, en changeant éventuellement la numérotation des vecteurs initiaux, on peut toujours supposer la suite des $k(j)$ croissante. Sous cette dernière hypothèse, on a nécessairement $k(i) \geq i$, pour tout i. Sinon les vecteurs $x_1, x_2, ..., x_i$ appartiendraient tous au sous-espace H_{i-1}, engendré par $(b_1, b_2, ..., b_{i-1})$, ce qui est impossible puisque les b_i sont indépendants.
Le vecteur x_i vérifie alors

$$|x_i| \;\geq\; l_{k(i)} \;\geq \min_{i \leq j \leq p} l_j.$$

Par suite, $\sup(|x_1|, |x_2|, ..., |x_i|) \geq \min_{i \leq j \leq p} l_j$, et en particulier, $\Lambda_i(L)$ vérifie l'inégalité annoncée. □

Par construction de la base b^* à l'aide du procédé d'orthogonalisation de Schmidt, le premier vecteur de b est le même que le premier vecteur de b^*. On considère alors une base b, $\mathcal{C}$-réduite, où $\mathcal{C}$ est une des conditions de la Définition 1. Le Lemme suivant et son Corollaire précisent les relations qui existent entre les normes de certains vecteurs de la base b et certains vecteurs de son orthogonalisée. Nous adaptons ici les résultats déjà existants pour la réduction de Siegel au cas de la réduction de Gram que nous avons introduite.

Lemme 3. *Toute base $b = (b_1, \ldots, b_p)$, (s, γ)-Gram réduite, vérifie les propriétés suivantes:*

1. $l_j \leq (1/s)^{(i-1)^\gamma (i-j)}\; l_i$ *pour* $1 \leq j \leq i \leq p$.
2. $|\; b_i \;| \leq (1/s)^{(i-1)^{1+\gamma}} l_i$ *pour* $1 \leq i \leq p$.
3. $|\; b_j \;| \leq (1/s)^{(i-1)^{1+\gamma}} l_i$ *pour* $1 \leq j \leq i \leq p$.

Preuve. Pour (1), on utilise la condition de (s, γ)-Gram, pour tous les indices j inférieurs à un indice fixé i,

$$\frac{l_i}{l_j} = \prod_{k=j}^{i-1} \frac{l_{k+1}}{l_k} > (s)^{(i-1)^\gamma + (i-2)^\gamma + \cdots + (j)^\gamma} > (s)^{(i-j)(i-1)^\gamma}$$

Pour (2), on exprime le vecteur b_i dans la base orthogonale b^*. La norme $|b_i|$ vérifie alors

$$|b_i|^2 = l_i^2 + \sum_{j<i} m_{ij}^2 l_j^2.$$

Comme la base b est propre,

$$|\; b_i \;|^2 \leq l_i^2 + \frac{1}{4} \sum_{j<i} l_j^2$$

on utilise alors la condition (1) ,

$$| b_i |^2 \leq l_i^2 \; [\, 1 + \frac{1}{4} \sum_{j<i} \left(\frac{1}{s}\right)^{2(i-1)^\gamma (i-j)} \,].$$

$$\text{Or,} \quad \sum_{j<i} \left(\frac{1}{s}\right)^{2(i-1)^\gamma (i-j)} = \left(\frac{1}{s}\right)^{2(i-1)^\gamma} \frac{(1/s)^{2(i-1)^\gamma (i-1)} - 1}{(1/s)^{2(i-1)^\gamma} - 1},$$

$$\text{et} \quad \frac{(1/s)^{2(i-1)^\gamma}}{4\left((1/s)^{2(i-1)^\gamma} - 1\right)} \leq 1 \quad \text{car} \quad s^2 \geq \frac{3}{4}.$$

$$\text{On en déduit que} \quad | b_i |^2 \leq (1/s)^{2(i-1)^{1+\gamma}} l_i^2.$$

La propriété (3) est une conséquence de (2) et de (1). □

Corollaire 1. *Si b est s-Siegel réduite, alors les propriétés* (1), (2) *et* (3) *du lemme 3 sont vraies avec* $\gamma = 0$.

Si une base est (s, k)*-Schmidt réduite, alors elle vérifie la propriété* (2) *du lemme 3, pour* $i = k$ *avec* $\gamma = 0$.

2.3.1 Majoration des défauts de longueur

Les conditions de (s, γ)-Gram et de (s, k)-Schmidt sont toutes deux des conditions moins fortes que la condition classique de Siegel. Le théorème suivant montre bien la différence entre ces deux conditions. Quand une base est (s, k)-Schmidt–réduite, on ne peut majorer que son k-ème défaut de longueur. Autrement dit par rapport à la réduction au sens de Siegel on perd les majorations sur tous les autres défauts et *on garde (exactement) les conséquences de la réduction au sens de s-Siegel, seulement pour un vecteur fixé de la base: le k-ème.* Pour la condition de (s, γ)-Gram, la condition de s-Siegel est relaxée pour tous les indices i, et on retrouve des majorations de tous les défauts de longueur, moins fortes que celles déduites de la condition de s-Siegel. Le cas $\gamma = 0$ permet d'ailleurs de retrouver exactement les majorations classiques, induites par la réduction de Siegel ou celle de Lovász (LLL).

Théorème 1. *Soit une base b d'un réseau de dimension p.*

1. *Si b est (s, k)-Schmidt réduite alors son k-ème défaut de longueur est majoré,*

$$\mu_k(b) \leq \left(\frac{1}{s}\right)^{p-1}.$$

2. *Si b est (s, γ)-Gram réduite alors tous ses défauts de longueur sont majorés,*

$$\mu_i(b) \leq \left(\frac{1}{s}\right)^{(p-1)^{1+\gamma}}, \ \textit{pour tout } i \textit{ dans } \{1, 2, ..., p\}.$$

Preuve. Le lemme 2 montre l'existence d'un indice $i \geq k+1$ pour lequel $\Lambda_k(L) \geq l_i$. La condition (s,k)-Schmidt d'indice i assure alors la minoration

$$\Lambda_k(L) \geq l_i \geq s^{i-k}\ l_k \geq s^{p-k} l_k.$$

Par ailleurs, le k-ème vecteur b_k d'une base (s,k)-Schmidt réduite vérifie $\mid b_k \mid \leq (1/s)^{k-1} l_k$ pourvu que $s^2 \leq 3/4$. Finalement, on obtient la majoration annoncée,

$$\mu_k = \frac{\mid b_k \mid}{\Lambda_k(L)} \leq \left(\frac{1}{s}\right)^{p-1}$$

Pour les bases (s,γ)-Gram réduites, la propriété (3) du Lemme 3 permet de conclure. □

2.3.2 Majoration du défaut d'orthogonalité

Si l'on désigne par H_i le sous-espace engendré par les i premiers vecteurs de b, et par θ_i, l'angle que forme b_i avec la plan H_i, alors:

$$\sin\theta_i = \frac{l_i}{\mid b_i \mid}$$

Majorer le défaut d'orthogonalité revient à minorer les sinus des angles θ_i. Les conditions de (s,γ)-Gram, s-Siegel et $t-$Lovász permettent de telles minorations. Il semblerait que les conditions de Schmidt (à elles seules) ne le permettent pas. Plus précisément,

Théorème 2. *Le défaut d'orthogonalité d'une base b, (s,γ)-Gram réduite d'un réseau de dimension p est majoré comme le montre la relation suivante.*

$$\rho(b) \leq \left(\frac{1}{s}\right)^{\frac{p^{2+\gamma}}{2+\gamma}}$$

Remarque. En prenant $\gamma = 0$, dans la proposition précédente, on obtient une majoration du défaut d'orthogonalité des bases réduites au sens de Siegel ou Lovász. Plus précisément, si une base b est s-Siegel réduite alors son défaut d'orthogonalité est majoré par $(1/s)^{p(p-1)/2}$.

Preuve. Si une base est (s,γ)-Gram réduite, alors les quantités $(\mid b_i \mid / l_i)$ sont majorées en vertu du lemme précédent. Le défaut d'orthogonalité (2.4) est alors majoré à son tour,

$$\rho(b) = \prod_{i=1}^{p} \frac{\mid b_i \mid}{l_i} \leq \left(\frac{1}{s}\right)^{\sum_{i=1}^{p}(i-1)^{\gamma}} = \left(\frac{1}{s}\right)^{\frac{p^{2+\gamma}}{2+\gamma}}.$$

La dernière majoration s'obtient en comparant la somme à une intégrale.
En particulier, si la base est s-Siegel réduite,

$$\rho(b) = \prod_{i=1}^{p} \frac{\mid b_i \mid}{l_i} \leq \prod_{i=1}^{p} \left(\frac{1}{s}\right)^{(i-1)} = \left(\frac{1}{s}\right)^{\frac{p(p-1)}{2}}$$

□

En conclusion, nous avons disséqué la réduction au sens de s-Siegel, trop riche, en deux types de conditions plus faibles.

1. Les réductions au sens de (s, k)-Schmidt aboutissent à une majoration du k-ème défaut de longueur et cette majoration est exactement la même que celle obtenue pour la réduction de Siegel. Mais il semble qu'il n'y ait de majoration ni pour les autres défauts de longueur, ni pour le défaut d'orthogonalité.
2. Les réductions au sens de (s, γ)-Gram permettent de majorer tous les défauts de longueur et également le défaut d'orthogonalité. Mais ces majorations sont moins fines que celles obtenues pour la réduction de Siegel. (On rappelle que quand $\gamma = 0$, on retrouve exactement la condition de s-Siegel.)

Les réductions "composées" de (s, k, γ)-Gram–Schmidt sont intéressantes, car elles permettent de garder la majoration d'un défaut de longueur (par exemple le premier) au même titre que la réduction de Siegel. Tous les autres défauts sont également majorés, mais de façon moins fine que par la réduction de Siegel.

On verra dans le Chapitre 6 que les réductions que nous avons introduites sont en moyenne plus rapides que les réductions classiques au sens de Siegel (ou Lovász). En pratique, la réduction d'une base d'un réseau est utilisée à une fin précise. Par exemple, la factorisation de polynômes à coefficients entiers était l'application visée par Lovász, Lovász et Lenstra, pour introduire leur algorithme, LLL. Notre propos est de profiter de la rapidité moyenne des algorithmes de réduction que nous introduisons, à chaque fois que l'application visée n'utilise pas pleinement toutes les propriétés euclidiennes des bases LLL–réduites. Le Chapitre suivant montre qu'une grande partie des applications se contentent d'un vecteur assez court dans le réseau (majoration de premier défaut de longueur).

2.4 Les algorithmes de réduction

Le premier algorithme efficace de réduction est dû à Lovász, Lenstra et Lenstra [LLL82]. L'algorithme LLL[4] est un algorithme d'approximation qui dépend d'un paramètre réel $t \in]1, 2[$. Il commence par l'orthogonalisation de la base initiale par le procédé classique de Gram-Schmidt. Puis, pendant une phase itérative, la base initiale est modifiée de telle sorte que les rapports des normes de deux orthogonalisés successifs soient minorés par une constante dépendant du paramètre d'approximation t (2.7, 2.8),

$$\frac{l_{i+1}}{l_i} \geq s \qquad \text{avec } s^2 = \frac{1}{t^2} - \frac{1}{4} \leq \frac{3}{4}. \tag{2.13}$$

Même si la base initiale est déjà réduite, pour le vérifier, il faut au minimum p itérations. C'est pourquoi, pour réduire une base p-dimensionnelle, l'algorithme LLL exige au moins p itérations.

Nous généralisons l'algorithme LLL à deux classes d'algorithme de réduction qui gardent la même structure mais dont les tests varient par rapport à ceux de l'algorithme original.

4. Le nom de l'algorithme LLL correspond aux initiales de ses inventeurs.

2.4.1 Première classe d'algorithmes de réduction

Lorsque $\mathcal{C}$ désigne une des conditions suivantes de la Définition 1:

- t-Lovász (2.7), s-Siegel (2.8), (s,γ)-Gram (2.9)
- $(s,1)$-Schmidt (2.10), $(s,\gamma,1)$-Gram–Schmidt (2.11), i.e. avec l'indice fixé k égal à 1,

voici l'algorithme générique de $\mathcal{C}$-réduction:

Algorithme de $\mathcal{C}$-réduction $\mathcal{A}(\mathcal{C})$

Donnée : une base p-dimensionnelle quelconque b d'un réseau L de $\mathbb{R}^n$.
Résultat : une base $\mathcal{C}$-réduite du réseau L.
Initialisation: Orthogonalisation de Gram-Schmidt:
Calculer la base orthogonalisée b^* , la matrice de passage m de b à b^*
et la suite $l = (l_1, \ldots, l_p)$ des normes des orthogonalisés.
$i := 1$;
Tant que $i < p$ répéter

1. **Translater** b_{i+1} parallèlement à b_i:
   Calculer r_i, l'entier le plus proche de $m_{i+1,i}$;
   Faire $b_{i+1} := b_{i+1} - r_i b_i$;
2. **Tester la condition** $\mathcal{C}_i$
   Si oui :
   Translater b_i parallèlement à b_k pour $k < i$, à l'aide de l'algorithme Propre$(b_1, \ldots, b_i)$;
   $i := i + 1$;
   Sinon :
   Échanger b_i et b_{i+1};
   Remettre à jour le triplet (b^*, m, l).
   Si $i \neq 1$ alors $i := i - 1$;

Pendant l'exécution de l'algorithme $(\mathcal{A}(\mathcal{C}))$, l'indice i varie dans l'ensemble $\{1, \ldots, n\}$ et effectue une marche aléatoire commandée par le résultat des tests. Lorsque la variable i prend une valeur $j \in \{1, \ldots, p-1\}$, la base $(b_1, \ldots, b_j)$ est déjà réduite. Puis, la condition $\mathcal{C}_j$ est testée. Si le résultat du test est positif, l'algorithme Propre$(j+1)$ est exécuté et $(b_1, \ldots, b_{j+1})$ est alors réduite. Donc, i est incrémenté et vaut $j+1$. Sinon, b_j et b_{j+1} sont échangés. À ce moment, rien ne garantit que la base $(b_1, \ldots, b_j)$ reste réduite. La variable i est alors décrémentée. L'algorithme recalcule b^*, m et l, comme on le décrira ci-dessous, puis translate le nouveau b_j parallèlement à b_{j-1} et teste $\mathcal{C}_{j-1}$. Ainsi, la variable i peut éventuellement redescendre à la valeur 1. Enfin, lorsque i vaut p, la base entière est réduite et l'algorithme termine.

Par ailleurs, quand la condition $\mathcal{C}_i$ est vérifiée ni b^* ni les l_i ne sont modifiés; seule la $i+1$ème ligne de la matrice est remplacée par une combinaison linéaire de la $(i+1)$-ème ligne et des lignes précédentes, à l'aide de la procédure Propre$(i+1)$.

Si en revanche, le test est négatif, alors l'échange des vecteurs b_i et b_{i+1} modifie les deux vecteurs b_i^* et b_{i+1}^* et il faut les recalculer, tout comme les colonnes i et $i+1$ de la matrice m. Soit $(\hat{b}^*, \hat{m}, \hat{l})$ le nouveau triplet à calculer après un échange. Les relations

$$\hat{m}_{i+1,i} = m_{i+1,i} \ \frac{l_i^2}{l_{i+1}^2 + m_{i+1,i}^2 l_i^2},$$

$$\text{et} \qquad \begin{pmatrix} \hat{b}_{i+1}^* \\ \hat{b}_i^* \end{pmatrix} = \begin{pmatrix} -\hat{m}_{i+1,i} & 1 - \hat{m}_{i+1,i} m_{i+1,i} \\ 1 & m_{i+1,i} \end{pmatrix} \begin{pmatrix} b_{i+1}^* \\ b_i^* \end{pmatrix},$$

permettent de calculer facilement les colonnes i et $i+1$ de la matrice m et les éléments l_i et l_{i+1} de la liste l selon les formules ci-dessous:

$$\hat{m}_{i,j} = m_{j,i} \hat{m}_{i+1,i} + m_{j,i+1} \frac{\mid b_{i+1}^* \mid^2}{\mid b_i^* \mid^2},$$

$$\hat{m}_{j,i+1} = m_{j,i} - m_{j,i+1} m_{i+1,i}.$$

Donc, la variable i désigne à un moment donné le plus grand indice j pour lequel on a vérifié que la base $(b_1, b_2, ..., b_j)$ est réduite. A ce moment, la matrice formée par les i premières lignes et les i premières colonnes de la matrice m, ainsi que les i premiers l_k ont la forme souhaitée. L'indice i croît dès lors que la base $(b_1, b_2, ..., b_i)$ est réduite. Or, concernant les réductions aux sens de (s,k)-Schmidt ou (s,k,γ)-Gram-Schmidt avec $k > 1$, les variations de i ne peuvent plus garder la même signification. En effet, on ne peut plus commencer par construire un vecteur b_1 qui convienne (comme dans les algorithmes précédents), car si après translation de b_1 parallèlement à b_k et test négatif de la condition de (s,k)-Schmidt ($l_k < (1/s)^{k-1} l_1$), on échangeait b_1 et b_k, cela modifierait la base orthogonalisée de façon trop importante($b_1^*, b_2^*, ..., b_k^*$ seraient tous modifiés) et les arguments classiques tenus pour prouver que l'algorithme est polynomial en la taille de son entrée seraient mis en défaut. C'est pourquoi, afin de construire une base (s,k)-Schmidt-réduite, il est plus astucieux de construire d'abord un vecteur b_{k-1} convenable. C'est l'objet du prochain paragraphe.

2.4.2 Deuxième classe d'algorithmes de réduction

Lorsque $\mathcal{C}$ désigne les conditions de (s,k)-Schmidt ou (s,k,γ)-Gram-Schmidt, (2.10) ou (2.11) avec l'indice fixé $k > 1$, l'algorithme générique de réduction $(\mathcal{B}(\mathcal{C}))$ garde la même structure que l'algorithme $(\mathcal{A}(\mathcal{C}))$, mais la variation de l'indice i prend un sens différent.
Lorsque $i > k$ tout se passe comme précédemment: disposant des vecteurs $b_1, b_2, ..., b_i$ convenables, on cherche à construire un vecteur b_{i+1} qui vérifie la condition de (s,k)-Schmidt.
En revanche, quand $i < k$, on est amené à adopter la démarche inverse: disposant des vecteurs $b_i, b_{i+1}, ..., b_k$ convenables, on cherche à construire un vecteur b_{i-1} qui vérifie la condition de (s,k)-Schmidt. Autrement dit, dans ce cas l'avancement de la réduction se traduit par une décroissance de l'indice i.
Voici donc l'algorithme qui permet de traduire ces considérations:

Procédure (s,k)**-Bloc-inf-réduction**

Donnée : une base b d'un réseau L de dimension k de $\mathbb{R}^n$.
Résultat : une base b de L réduite au sens de (s,k)-Schmidt .

Initialisation: $i := k-1$;
Tant que $i > 0$ **répéter**
1. **Translater** b_{i+1} parallèlement à b_i:
   Calculer r_i, l'entier le plus proche de $m_{i+1,i}$;
   Faire $b_{i+1} := b_{i+1} - r_i b_i$;
2. **Tester la condition** $\mathcal{C}_i$
   Si oui :
   Propre $(b_i, \ldots, b_k)$;
   $i := i-1$;
   Sinon :
   Échanger b_i et b_{i+1};
   Remettre à jour le triplet (b^*, m, l);
   Si $i \neq k-1$ alors $i := i+1$;

Algorithme de $\mathcal{C}$**-réduction** $(\mathcal{B}(\mathcal{C}))$

Donnée : une base b d'un réseau L de dimension p de $\mathbb{R}^n$.
Résultat : une base b de L réduite au sens de (s,k)-Schmidt .

Initialisation: 1. Orthogonalisation de Gram-Schmidt;
2. (s,k)-Bloc-inf-réduction de $(b_1, b_2, \ldots b_k)$;
3. Si $k \neq p$ faire $i := k$;

Tant que $i > 0$ **répéter**
1. **Translater** b_{i+1} parallèlement à b_i:
   Calculer r_i, l'entier le plus proche de $m_{i+1,i}$;
   Faire $b_{i+1} := b_{i+1} - r_i b_i$;
2. **Tester la condition** $\mathcal{C}_i$
   Si oui :
   Propre $(b_1, \ldots, b_i)$;
   $i := i+1$;
   Sinon :
   Échanger b_i et b_{i+1};
   Remettre à jour le triplet (b^*, m, l);
   Si $i \neq k$ alors $i := i-1$;
   Sinon si $k \neq 1$ alors (s,k)-Bloc-inf-réduction de $(b_1, b_2, \ldots b_k)$
   $i := k$;

Pour finir ce paragraphe, nous allons montrer un Lemme très utile. Déjà souligné par [LLL82], ce Lemme a été également utilisé par [DV94]. Il permet de montrer par exemple que les défauts de longueur de la base de sortie des algorithmes de réduction sont majorés au moins aussi bien que ceux de la base d'entrée. Autrement dit, les algorithmes de réduction *ne dégradent pas* la qualité d'une base.

Lemme 4. *Soit $b^0 = (b_1^0, \dots, b_p^0)$ une base proposée en entrée à un des algorithmes de réduction précédemment introduits , a et A les réels positifs définis par*

$$A := \max_{1\le i\le p} l_i^0 \quad et \quad a := \min_{1\le i\le p} l_i^0.$$

Les longueurs l_i des orthogonalisés de toutes les bases intermédiaires construites pendant le déroulement des algorithmes de réduction $\mathcal{A}(\mathcal{C})$ ou $\mathcal{B}(\mathcal{C})$ vérifient

$$a \le l_i \le A, \quad \forall i \in \{1, \dots, p\}.$$

Preuve. Pendant le déroulement d'un des algorithmes génériques de réduction, lorsqu'un test de la condition $\mathcal{C}_i$ est positif, les normes des orthogonalisés l_i restent inchangées.
Si le test de $\mathcal{C}_i$ est négatif, alors b_{i+1} et b_i sont échangés. Seuls l_{i+1} et l_i seront modifiés. On note $(\hat{l}_i, \hat{l}_{i+1})$ le nouveau couple (l_i, l_{i+1}). Par construction,

$$\hat{l}_i = |b_{i+1}^* + m_{i+1,i} b_i^*| = \sqrt{l_{i+1}^2 + m_{i+1,i}^2 l_i^2} \ge l_{i+1}$$

De plus, comme le test de s-Siegel d'indice i (avec $s^2 < 3/4$) est négatif et grâce à la propreté,

$$\hat{l}_i < l_i \sqrt{s^2 + 1/4} < l_i.$$

En projetant les deux vecteurs échangés b_i et b_{i+1}, orthogonalement au sous–espace vectoriel engendré par $b_1, \dots, b_{i-1}$, on constate aisément que

$$\hat{l}_{i+1} \hat{l}_i = l_{i+1} l_i.$$

On en déduit respectivement

$$\hat{l}_{i+1} > l_{i+1} \quad \text{et} \quad \hat{l}_{i+1} \le l_i.$$

On en déduit qu'une quantité qui majore (resp. minore) les longueurs des orthogonalisés de la base d'entrée, majore (resp. minore) celles de toutes les bases construites au cours de l'algorithme de réduction au sens de s-Siegel. Le résultat reste valable pour tous les autres algorithmes de réduction précédemment définis, car le Lemme 1 montre que pendant l'exécution des algorithmes génériques $\mathcal{A}(\mathcal{C})$ ou $\mathcal{B}(\mathcal{C})$, lorsque $\mathcal{C}_i$ n'est pas satisfaite, la condition de s-Siegel pour l'indice i n'est pas satisfaite non plus. □

Nous allons montrer dans la prochaine Section, que les algorithmes de réduction présentés terminent tous au bout d'un nombre d'itérations qui est toujours polynomial en la taille de la base d'entrée. On rappelle que la taille de l'entrée des algorithmes de réduction est en $\mathrm{O}(np \log M)$.

2.4.3 Les nombres maximums d'itérations et complexités dans le pire des cas

Ci-dessous, nous adaptons la démonstration originale de [LLL82], au cas des généralisations que nous avons introduites précédemment.

Théorème 3. *Soient deux paramètres réels s et t liés par (2.6) et une condition $\mathcal{C}$ parmi celles introduites en Définition 1. Le nombre maximum d'itérations K de l'algorithme générique de $\mathcal{C}$-réduction ($\mathcal{A}$ et $\mathcal{B}$), lorsqu'il opère sur une base initiale $(b_1, \ldots, b_p)$, entière vérifie les trois inégalités suivantes*[5]

$$K \leq p(p-1)\log_t(\sqrt{n}M_1) + p - 1, \quad \textit{où} \quad M_1 := \max |b_i|_\infty.$$

$$K \leq p(p-1)\log_t M_2 + p - 1, \quad \textit{où} \quad M_2 := \max |b_i|_2.$$

$$K \leq p(p-1)\log_t \frac{A}{a} + p - 1, \quad \textit{où} \quad A := \max l_i \quad \textit{et} \quad a := \min l_i.$$

Preuve. Comme la base initiale $(b_1, \ldots, b_p)$ est entière et que $|b_i|^2 < nM^2$ pour tout $i \leq p$, tous les *réseaux commençants* L_j engendrés par $(b_1, \ldots, b_j)$, $(j \leq p)$ ont des déterminants entiers, notés d_j. Alors, la quantité

$$D := \prod_{j=1}^{p-1} d_j^2 = \prod_{j=1}^{p-1}\prod_{k=1}^{i} l_k^2,$$

est également entière et vérifie toujours

$$1 \leq D \leq M_2^{p(p-1)} \quad , \quad 1 \leq D \leq (\sqrt{n}M_1)^{p(p-1)} \quad \text{et} \quad a^{p(p-1)} \leq D \leq A^{p(p-1)} \tag{2.14}$$

La dernière inégalité est fondée sur le fait que pendant une exécution de l'algorithme la quantité a croît et la quantité A décroît. (Lemme 4)
La preuve du théorème repose alors sur la décroissance de la quantité D, pendant une exécution de l'algorithme générique de $\mathcal{C}$ réduction.
On suppose d'abord que $\mathcal{C}$ désigne la condition de t-Lovász (2.7) et on considère les deux résultats possibles du test de la condition $\mathcal{C}_i$.
Si le résultat est *positif*, la quantité D n'est pas modifiée, car les réseaux commençants L_j pour $j \leq i$ restent inchangés.
Si le résultat est *négatif*, seul le réseau L_i est modifié. En effet,

$$d_i^2 = \prod_{k=1}^{i} l_k^2 = d_{i-1}^2 l_i^2,$$

et après l'échange le i-ème orthogonalisé, b_i^* est remplacé par $(b_{i+1}^* + m_{i+1,i}b_i^*)$. Comme la condition $\mathcal{C}_i$ (2.7) n'est pas satisfaite, la quantité d_i^2 décroît alors d'un facteur de $1/t^2$. La quantité D décroît du même facteur car les autres d_j ne sont pas modifiés.

5. Les maxima et minima sont toujours pris pour i variant dans $\{1, \ldots, p\}$. Par ailleurs, pour un vecteur u, $|u|_2$ désigne sa norme euclidienne et $|u|_\infty$ sa norme infinie.

Supposons maintenant que $\mathcal{C}$ désigne la condition de (s,k)-Schmidt. Pendant l'exécution de l'algorithme de réduction,

1. lorsque $i \geq k$ et que l'algorithme teste la condition $\mathcal{C}_i$, la base $(b_1, \ldots, b_{i-1})$ est déjà réduite et en particulier la condition $\mathcal{C}_{i-1}$ est satisfaite. Á ce moment, si le test de $\mathcal{C}_i$ est négatif, la condition $\mathcal{C}_i$ n'est pas satisfaite. L'assertion (3) du lemme 1 montre alors que la condition de s-Siegel d'indice i n'est pas non plus satisfaite.
2. lorsque $i \leq k$ et que la procédure (s,k)-Bloc-inf-réduction teste la condition $\mathcal{C}_i$, par construction la condition $\mathcal{C}_{i+1}$ est déjà satisfaite. À ce moment, si le test de $\mathcal{C}_i$ est négatif, d'après l'assertion (4) du lemme 1, la condition de s-Siegel d'indice i n'est pas satisfaite non plus.

Par ailleurs, quelle que soit la condition $\mathcal{C}$ de la Définition 1, par construction avant le test de $\mathcal{C}_i$, le coefficient subdiagonal $m_{i+1,i}$ vérifie

$$|m_{i+1,i}| \leq 1/2.$$

Enfin, d'après les assertions (2) et (1) du lemme 1, à chaque fois que le résultat d'un test $\mathcal{C}_i$ est négatif, la quantité D décroît d'un facteur de $1/t^2 = s^2 + 1/4$.
Le nombre d'itérations K de l'algorithme de $\mathcal{C}$ réduction est exactement le nombre de tests de la condition $\mathcal{C}$. Par ailleurs, les nombres K_- et K_+ de fois de tests avec un résultat négatif, respectivement positif, vérifient

$$K_+ \leq K_- + p - 1 \quad \text{et} \quad K = K_+ + K_- < 2K_- + p - 1.$$

Or la décroissance systématique de la quantité D, à l'occasion de chaque test au résultat négatif entraîne successivement

$$2K_- \leq p(p-1)\log_t(\sqrt{n}M_1) \quad , \quad 2K_- \leq p(p-1)\log_t M_2 \quad , \quad 2K_- \leq p(p-1)\log_t \frac{A}{a}.$$

□

2.5 Conclusion

Grâce aux algorithmes cités, on peut disposer dans un réseau de bases réduites de qualité variable. On distingue deux groupes:

- Les réductions aux sens de t-Lovász, s-Siegel, (s,γ)-Gram et (s,k,γ)-Gram-Schmidt: tous les défauts de longueur, et également le défaut d'orthogonalité sont majorés.
- Les réductions aux sens de (s,k)-Gram-Schmidt: seul le k-ème défaut de longueur est majoré.

En particulier, pour une base réduite au sens de $(s,1,\gamma)$-Gram-Schmidt, le premier vecteur possède les mêmes propriétés que le premier vecteur d'une base t-Lovász réduite. Les autres vecteurs ont des propriétés euclidiennes affaiblies, mais similaires.

Mais malgré des tests d'arrêt différents et donc la qualité variable des bases de sortie, pour toutes les versions de l'algorithme LLL, le nombre d'itérations dans le pire des cas est majoré de la même façon.

En résumé, dans ce Chapitre, nous avons introduis de nouvelles réductions et montré que *dans le pire des cas, nos algorithmes ne sont pas moins efficaces que l'algorithme LLL original*, puisque le nombre maximum d'itérations des algorithmes introduits est le même que celui de l'algorithme LLL.

C'est l'étude en moyenne (Chapitre 6) qui permettra de distinguer le comportement de ces différents algorithmes. Par exemple, la réduction au sens de $(s, 1, \gamma)$-Gram-Schmidt est en moyenne plus rapide que l'algorithme LLL original. On montrera alors qu' *en moyenne, nos algorithmes sont plus efficaces que l'algorithme LLL original.*

Chapitre 3

Les réseaux aléatoires

Dans ce chapitre nous définissons le modèle probabiliste que nous choisissons pour l'étude en moyenne. Nous commencerons par définir les bases et les réseaux aléatoires. Puis par une première étude en moyenne des caractéristiques classiques des bases du modèle, je *quantifie* les qualités euclidiennes de ces bases en fonction du rapport entre la dimension de la base et la dimension de l'espace ambiant. Pour bien éclaircir cette quantification, j'introduis les notions d'indices affines et de bases incomplètes (Définitions 2 et 3).

Le modèle probabiliste choisi est avant tout le plus simple et le plus naturel: les p vecteurs d'une base aléatoire seront choisis indépendamment et uniformément dans la boule unité de $\mathbb{R}^n$ $(p \leq n)$. Les raisons qui motivent le choix d'un modèle continu et de taille maximale 1 sont expliquées dans la première Section.

Nous avons déjà noté que les normes des orthogonalisés et plus particulièrement les rapports de ces normes interviennent de manière essentielle dans les définitions des processus de réductions. En effet, toutes les réductions que nous avons introduites dans le Chapitre 2, s'expriment à l'aide de ces rapports. Dans ce chapitre, on commence par rappeler (Section 2), un résultat important dû à Hérvé Daudé et à Brigitte Vallée [DV94] concernant la distribution des normes des orthogonalisés des bases aléatoires.

Puis, nous étudions en moyenne le rapport entre les normes de deux de ces orthogonalisés. Plus précisément, nous choisissons de considérer deux vecteurs orthogonalisés de la base aléatoire dont *les indices sont des fonctions affines de la dimension* n *de l'espace ambiant*. Ce type d'indice sera d'ailleurs appelé indice affine. Ce deuxième choix sera amplement justifié a posteriori. Il nous permettra en effet de mettre en évidence des phénomènes asymptotiques différents (Chapitre 5, Théorème 9) selon la localisation des indices, i.e. la pente des fonctions affines. Nous définirons les notions d'indices affines de début ou de fin, en considérant cette localisation. Le calcul de la moyenne et de la variance des rapports entre deux orthogonalisés d'indices affines et de leurs développements asymptotiques fournit alors une première estimation de la répartition de ces rapports. (Section 3). Les résultats obtenus seront précisés (Chapitre 5) et utilisés (Chapitre 6) pour classer les différents algorithmes de réduction précédemment introduits selon leurs comportements en moyenne.

Dans un deuxième temps, nous étudions les moyennes des défauts de longueur et du défaut d'orthogonalité des bases de notre modèle. Notre étude (Section 4) fournit des bornes asymptotiques, pour la moyenne de tous ces défauts et nous montrons qu'en moyenne ces défauts dépendent du rapport entre la dimension de la base et celle de l'espace ambiant. Plus précisément, lorsque la dimension n de l'espace ambiant tend vers l'infini, la qualité d'une base se dégrade à mesure que ce rapport s'approche de 1. La limite de ce rapport quand n tend vers l'infini, (elle existe toujours dans les cas que nous étudions) sera appelée *le poids* d'une base et conformément à notre choix d'indice (affine), nous considérons particulièrement les bases aléatoires dont le cardinal est une expression affine de n. Notre étude de la Section 4 nous fournit une échelle de bases aléatoires de qualité variable (en fonction du poids de la base). C'est grâce à cette échelle que nous arrivons ensuite à comparer et à distinguer (Chapitre 6) les comportements en moyenne des différents algorithmes de réductions que nous avons introduits. On rappelle que notre étude dans le pire des cas du Chapitre 2 (Théorème 3) ne permettait pas cette distinction.

3.1 Définition du modèle probabiliste

Tout d'abord, on remarque que tous les algorithmes de réduction définis dans les chapitres précédents se comportent de manière identique sur une base d'entrée b et sur sa transformée par rapport à une similitude quelconque de $\mathbb{R}^n$. Ceci justifie de travailler sur des bases de longueur maximum 1.
Par ailleurs, le choix d'un cadre continu permet d'utiliser les outils provenant de l'analyse mathématique. Des méthodes classiques basées sur le comptage du nombre de points à coordonnées entières à l'intérieur d'une sphère en fonction du volume de la sphère permettent de généraliser les résultats au cadre discret. Pour de plus amples détails se reporter par exemple à la Section 3 de [DV94].
C'est pourquoi on opte ici pour le modèle le plus naturel et le plus simple, où p vecteurs $(b_1, b_2, ..., b_p)$ $(p \leq n)$ sont choisis de façon aléatoire, indépendamment et uniformément dans la boule unité $\mathcal{B}_n$ de $\mathbb{R}^n$. Ce modèle est suffisamment simple pour y mener les calculs à termes et avoir des résultats sur les comportments en moyenne des algorithmes. Mais, il ne correspond pas, a priori, aux applications pratiques de la réduction. Nous avons effectué une série d'expérimentations (Chapitre 8) qui montrent toutefois que nos résultats sur les comportements relatifs des différents algorithmes de réduction restent valables dans d'autres modèles provenants des applications réelles de la réduction (modèle "sac–à–dos").
Revenons à notre modèle, *le modèle uniforme*. La mesure d'un espace de dimension $p-1$ de $\mathbb{R}^n$ étant nul, le système $b = (b_1, b_2, ..., b_p)$ ainsi choisi est presque sûrement libre. Dans toute la suite, un tel système est appelé *une base aléatoire* de cardinal p et le réseaux euclidien qu'il engendre est *un réseau aléatoire* de dimension p. Par abus de langage, on parlera souvent de base aléatoire de dimension p.
Il est alors clair que pour tout i, tel que $1 \leq i \leq p$, la fonction de répartition et la moyenne de la norme de b_i vérifie

$$\Pr\{|b_i| < x\} = x^n \quad \text{et} \quad \mathrm{E}(|b_i|) = \frac{n}{n+1}.$$

La base orthogonalisée de b (par le procédé de Schmidt) est notée $b^* = (b_1^*, b_2^*, ..., b_p^*)$ et les normes des orthogonalisés $l_1, l_2, ..., l_p$ sont alors également des variables aléatoires prenant leurs valeurs dans $[0, 1]$.

Pour clore cette section, on rappelle les définitions et les propriétés élémentaires des intégrales eulériennes. Pour plus de détails on peut par exemple consulter [WW27].
L'intégrale Beta (ou eulérienne de première espèce) relative aux deux paramètres réels et strictement positifs p et q est définie par

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\, dx = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}}\, dt. \tag{3.1}$$

Pour $x > 0$, la fonction Gamma (ou encore l'intégrale eulérienne de seconde espèce) du paramètre x est définie par

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du.$$

La propriété principale de la fonction Gamma est: $\Gamma(x+1) = x\Gamma(x)$. Cette relation permet de prolonger la fonction à $\mathbb{R}\backslash\mathbb{Z}_-$. La fonction Gamma est définie plus généralement sur $\mathcal{C}\backslash\mathbb{Z}_-$. Par ailleurs, en intégrant directement, on obtient quelques valeurs remarquables de la fonction Gamma:

$$\Gamma(1) = 1, \quad \Gamma(n) = (n-1)! \quad \text{et} \quad \Gamma(1/2) = \sqrt{\pi}.$$

La formule de Stirling en donne une bonne approximation, même lorsque x est petit,

$$\Gamma(x+1) = x^x\ e^{-x}\sqrt{2\pi x}\left(1 + \frac{1}{12x} + \frac{1}{288x^2} + \frac{\alpha}{x^3}\right), \text{ avec } \quad \lim_{x\to\infty} \alpha = 0.$$

L'intégrale $B(p,q)$ s'exprime à l'aide de la fonction Gamma,

$$B(p,q) = \frac{\Gamma(p)\ \Gamma(q)}{\Gamma(p+q)}. \tag{3.2}$$

On appelle *intégrale Beta incomplète* la quantité $B(p,q,u)$ définie par

$$B(p,q,u) = \int_0^u x^{p-1}\ (1-x)^{q-1}\ dx, \quad \text{où} \quad 0 \le u \le 1, \tag{3.3}$$

et p et q deux réels positifs. Par définition, la variable aléatoire X à valeurs dans $[0,1]$ suit une loi Beta de paramètres p et q si sa fonction de répartition vérifie

$$\Pr\{X \le u\} = \frac{B(p,q,u)}{B(p,q)} \quad \text{pour} \quad 0 \le u \le 1. \tag{3.4}$$

C'est une loi bien connue en théorie de probabilité. En particulier, si X suit une loi Beta de paramètres p et q, alors sa moyenne et sa variance sont données par les relations

$$\mathrm{E}(X) = \frac{p}{p+q} \quad \text{et} \quad \rho^2(X) = \frac{pq}{(p+q)^2(p+q+1)} \tag{3.5}$$

À titre d'exemple, la loi Beta apparaît naturellement dans l'étude des statistiques d'ordre [Dav]. Un bref rappel de ce résultat classique nous semble utile. Lorsque p variables aléatoires $X_1, \ldots, X_p$ sont rangeés par ordre croissant sous la forme

$$X_{(1)} \le X_{(2)} \le \cdots \le X_{(p)},$$

alors la variable aléatoire $X_{(i)}$ est appelée la i-ème statistique d'ordre.

Proposition 1. *Soient p variables aléatoires indépendantes $X_1, \ldots, X_p$, uniformément distribuées dans l'intervalle $[0,1]$. Alors la i-ème statistique d'ordre $X_{(i)}$ suit une loi Beta de paramètres $(i, n-i+1)$.*

Preuve. Par définition, la fonction de répartition $F_i(x)$ de $X_{(i)}$ vérifie

$$\begin{aligned} F_i(x) &= \Pr\left\{X_{(i)} \leq x\right\} \\ &= \Pr\{\text{au moins, } i \text{ variables parmi } X_1, \ldots, X_p \text{ sont inférieures à } x\} \\ &= \sum_{j=i}^{n} \binom{n}{j} x^j (1-x)^{n-j}, \end{aligned}$$

car chaque terme de la somme est la probabilité qu'exactement j variables parmi $X_1, \ldots, X_p$ sont inférieures à x. Le résultat annoncé s'en déduit alors, grâce à l'identité combinatoire classique,

$$\sum_{j=i}^{n} \binom{n}{j} x^j (1-x)^{n-j} = \frac{B(i, n-i+1, x)}{B(i, n-i+1)}.$$

□

3.2 La distribution des normes des orthogonalisés

Dans le cadre du modèle probabiliste précédemment défini, Daudé et Vallée [DV94] ont établi les résultats suivants sur les variables l_i.

Théorème 4. *[DV94] Soient p vecteurs $(b_1, b_2, ..., b_p)$ pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Alors les longueurs l_i $(i \leq p)$ de leurs orthogonalisés par le procédé de Gram-Schmidt forment une suite de variables aléatoires indépendantes. La densité $f_{i,n}(u)$ de la variable aléatoire l_i est donnée sur l'intervalle $[0,1]$ par:*

$$f_{i,n}(u) = \frac{2}{B(\frac{n-i+1}{2}, \frac{i+1}{2})} u^{n-i} (1-u^2)^{\frac{i-1}{2}}.$$

Preuve. Il est clair que la longueur l_i du i-ème vecteur orthogonalisé b_i^*, ne dépend pas des vecteurs $(b_{i+1}, \ldots, b_n)$. La probabilité $\Pr\{l_i < u\}$ s'exprime en fonction des probabilités conditionnelles

$$\Pr_{(c_1,\ldots,c_{i-1})}\{l_i < u\} := \Pr\{l_i < u \mid (b_1, \ldots, b_{i-1}) = (c_1, \ldots, c_{i-1})\},$$

et se met sous la forme,

$$\Pr\{l_i < u\} = \int_{(\mathcal{B}_n)^{i-1}} \Pr_{(c_1,\ldots,c_{i-1})}\{l_i < u\}\ d\Pr\{(b_1, \ldots, b_{i-1}) = (c_1, \ldots, c_{i-1})\}. \qquad (3.6)$$

Par ailleurs, le système $(b_1, \ldots, b_{i-1})$ est uniformément réparti dans $(\mathcal{B}_n)^{i-1}$. De plus, la probabilité conditionnelle de la relation précédente est indépendante du système fixé $(c_1, \ldots, c_{i-1})$ de $(\mathcal{B}_n)^{i-1}$. En effet, si l'on se fixe un tel système, on peut toujours l'orthonormaliser par le

procédé de Schmidt, puis le compléter pour obtenir une base orthonormée de $\mathbb{R}^n$. Alors, pour que $l_i < u$, il faut et il suffit que les coordonnées $(x_1, \dots, x_n)$ de b_i dans la base orthonormée ainsi construite, vérifient quelques contraintes simples à exprimer.
Quel que soit le système $(c_1, \dots, c_{i-1})$ fixé,

$$\Pr_{(c_1,\dots,c_{i-1})}\{l_i < u\} = \frac{\Pr\{\sum_{j=1}^n x_j^2 \leq 1 \text{ et } \sum_{j=i}^n x_j^2 \leq u^2\}}{\Pr\{\sum_{j=1}^n x_j^2 \leq 1\}}. \tag{3.7}$$

Ainsi d'après la relation (3.6), la probabilité $\Pr\{l_i < u\}$ ne dépend que de n, i et u et en fixant arbitrairement un système de $i-1$ premiers vecteurs, son expression est donnée par (3.7): si ν désigne le volume d'un ensemble de $\mathbb{R}^n$ et ω_k le volume de la boule unité de dimension k,

$$\Pr_{(c_1,\dots,c_{i-1})}\{l_i < u\} = \frac{1}{\omega_n}\int_{v=0}^{u} \nu\left(\{(x_1,\dots,x_n) : \sum_{j=1}^{i-1} x_j^2 \leq \sqrt{1-v^2} \text{ et } \sum_{j=i}^{n} x_j^2 = v^2\}\right) dv.$$

En remarquant que l'ensemble des $(x_i, \dots, x_n)$ qui vérifient $\sum_{j=i}^n x_j^2 = v^2$, définit une sphère de dimension $n-i+1$ dans $\mathbb{R}^n$, l'expression de la densité de l_i se déduit par dérivation de la relation précédente,

$$\frac{d\Pr\{l_i < u\}}{du} = (n-i+1)\frac{\omega_{n-i+1}\omega_{i-1}}{\omega_n}(1-u^2)^{\frac{i-1}{2}} u^{n-i}$$

L'expression de ω_k à l'aide de la fonction Γ [WW27],

$$\omega_k = \pi^{k/2}\Gamma(1+k/2),$$

la relation $\Gamma(1+x) = x\Gamma(x)$ et (3.2) permettent alors de conclure. □

Le corollaire suivant montre que les carrés des normes des orthogonalisés suivent alors une loi Beta (3.4).

Corollaire 2. *[DV94] Les variables aléatoires l_i^2 forment une suite de variables indépendantes suivant une loi Beta (3.4) de paramètres $\left(\frac{n-i+1}{2}, \frac{i-1}{2}\right)$. La moyenne et la variance de l_i^2 sont alors données par les relations,*

$$\mathrm{E}(l_i^2) = \frac{n-i+1}{n+2} \quad \textit{et} \quad \rho^2(l_i^2) = \frac{2(n-i+1)(i+1)}{(n+2)^2(n+4)}$$

Preuve. La moyenne et la variance d'une variable suivant une loi Beta sont rappelées en (3.5). □

Les procédés de réduction précédemment cités du réseau engendré par $(b_1, b_2, ..., b_n)$ manipulent tous les rapports $l_{b(n)}/l_{a(n)}$, où $a(n)$ et $b(n)$ sont des fonctions à valeurs dans $\{1 \dots n\}$. Par exemple, le cas où $b(n) := a(n) + 1$ est important dans la réduction au sens de s-Siegel. Le cas où $a(n) := 1$ concerne la réduction au sens de $(s, 1)$-Schmidt, etc.

Dans la suite, nous allons considérer le cas où a et b sont des fonctions affines de n. Ce choix nous permet de recouvrir une large panoplie d'indices dans les réseaux. L'ensemble des indices considérés est alors assez dense pour exhiber des phénomènes probabilistes différents selon la localisation de l'indice par rapport à la dimension n de l'espace ambiant ou encore selon la pente de la fonction affine qui représente l'indice. D'autre part, en optant pour les fonctions affines, on arrive à mener les calculs asymptotiques à terme. En particulier ce choix est bien adapté à l'utilisation de la méthode de Laplace de recherche d'équivalents pour les intégrales (Chapitre 4). Remarquons enfin que de nombreuses autres *fonctions indices* peuvent être traitées, par majoration ou minoration, à partir de notre étude.

Définition 2. Soit un réel θ élément de $[0,1]$ et une constante entière r. La quantité

$$f(n) := \theta n + r$$

est appelée *un indice affine* si et seulement si $f(n)$ est élément de $\{1,...,n\}$.
Un indice affine est appelé *un indice affine du tout début* si et seulement si la pente θ de la fonction affine f est nulle, $\theta = 0$.
Un indice affine est appelé *un indice affine du début* si et seulement si la pente θ de la fonction affine f est strictement inférieure à 1, $\theta < 1$.
Un indice affine est appelé *un indice affine de la fin* si et seulement si la pente θ de la fonction affine f est égale à 1, $\theta = 1$.
Lorsqu'il n'y a pas d'ambiguïté, on omettra l'adjectif *affine*: on parlera alors *d'indice du tout début, du début, de la fin.*

Dans toute la suite (à une exception près), les quantités $a(n)$ et $b(n)$ seront des indices[1] et sont définis par

$$\left\{ \begin{array}{lcl} a(n) & := & \alpha n + i \\ b(n) & := & \beta n + j \end{array} \right. \quad \text{avec} \quad (\alpha, \beta) \in [0,1]^2 \text{ et } (i,j) \in \mathbb{Z}^2 \tag{3.8}$$

Notation 2. Les densités et fonctions de répartition des variables l_b/l_a seront notées respectivement $g_{n,(a,b)}(v)$ et $G_{n,(a,b)}(v)$ et les indices a, b et n seront omis, lorsqu'il n'y aura pas d'ambiguïté:

$$G_{n,(a,b)}(v) = \Pr\{\frac{l_{b(n)}}{l_{a(n)}} < v\}$$

$$g_{n,(a,b)}(v)dv = \frac{dG_{n,(a,b)}}{dv}dv = \Pr\{v < \frac{l_{b(n)}}{l_{a(n)}} < v + dv\}$$

Lorsque $a(n) \neq b(n)$, les variables aléatoires $l_{a(n)}$ et $l_{b(n)}$ sont indépendantes d'après le Théorème 4. La densité du couple $(l_{a(n)}, l_{b(n)})$ est alors $f_{a,b}(x,y) = f_a(x)f_b(y)$ et

$$G(v) = \Pr\{\frac{y}{x} < v\} = \int_{\Delta} f_a(x)f_b(y)\ dx\ dy \tag{3.9}$$

$$\text{où } \Delta = \{(x,y) \in [0,1]^2 : \frac{y}{x} < v\}$$

1. Les résultats énoncés seront toujours vrais pour $(a(n), b(n)) \in [1,n]^2$, mais n'auront un sens que lorsque $(a(n), b(n)) \in \{1, \ldots, n\}^2$.

En distinguant les deux cas $v \leq 1$ et $v \geq 1$, l'expression exacte de la densité $g(v)$ s'obtient aisément en dérivant la fonction de répartition $G(v)$. Par exemple, lorsque $v \leq 1$:

$$g(v) = \int_0^1 x f_a(x) f_b(vx) dx. \tag{3.10}$$

Dans l'état, ces expressions restent en général trop complexes pour être exploitées. Par ailleurs, les applications classiques et en particulier cryptographiques utilisent des réseaux à grande dimension et c'est aussi le cadre général de l'étude de la complexité. Tout ceci justifie l'étude asymptotique de la distribution des variables $l_b(n)/l_a(n)$.
Une première approche est l'étude en moyenne de ces variables. Complétée par le calcul de la variance, cette approche fournit une première approximation du comportement de ces variables aléatoires. C'est l'objet de la Section suivante.
Les résultats de l'étude asymptotique en moyenne seront confirmés et précisés par l'étude asymptotique en distribution. L'outil principal de cette étude est une généralisation à la dimension 2 de *la méthode de Laplace* de calcul d'équivalents pour les intégrales dépendant d'un paramètre. Le chapitre 4 rappelle cette méthode et en énonce une généralisation pour les intégrales doubles prises sur des domaines suffisamment simples du plan. Le chapitre 5 exhibe les comportements asymptotiques de la fonction de répartition G, puis de la densité g en utilisant la méthode de Laplace. Les résultats obtenus sont appliqués à l'étude des comportements des algorithmes de réduction sur les réseaux aléatoires (Chapitre 6).

3.3 Étude en moyenne des variables aléatoires l_a/l_b

Cette Section calcule les asymptotiques des moyennes et variances des variables aléatoires l_b/l_a. Puis, grâce à l'inégalité de Bienaymé-Chebytchev, nous donnons une première estimation asymptotique de ces variables. L'étude ultérieure en distribution du chapitre 5 confirmera tous ces résultats.

3.3.1 Préliminaires

La moyenne de la variable aléatoire $l_{b(n)}/l_{a(n)}$ est notée $\mathrm{E}_n(l_b/l_a)$, sa variance $\sigma_n^2(l_b/l_a)$.
Soit une fonction réelle quelconque h de la variable aléatoire réelle X, de densité $f(x)$. La moyenne $\mathrm{E}[h(X)]$, de la variable aléatoire $h(X)$, lorsqu'elle est définie, est donnée par la relation suivante.

$$\mathrm{E}[h(X)] = \int_{\mathbb{R}} h(x) f(x) dx. \tag{3.11}$$

La variance $\sigma_n^2(X)$ est le moment centré d'ordre 2,

$$\sigma_n^2(X) := (X - \mathrm{E}[X])^2 = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 \,. \tag{3.12}$$

L'inégalité de Bienaymé-Chebytchev majore la probabilité que la variable aléatoire s'écarte de sa moyenne en fonction de la variance:

$$\forall \lambda \in R_+ \;\; \Pr\{|X - \mathrm{E}[X]| > \lambda\, \sigma(X)\} \;<\; \frac{1}{\lambda^2}. \tag{3.13}$$

Pour les calculs asymptotiques, il sera souvent fait appel à l'approximation

$$\frac{\Gamma\left(x+\frac{1}{2}\right)}{\Gamma\left(x\right)} = \sqrt{x}\left(1-\frac{1}{8x}+\frac{1}{128x^2}+\mathrm{O}\left(\frac{1}{x^3}\right)\right), \tag{3.14}$$

qui s'obtient simplement à l'aide de la formule de Stirling. La relation (3.14) est par exemple très utile lors de certaines analyses de phénomènes de marches aléatoires.

3.3.2 La moyenne

Soient deux fonctions $a(n)$ et $b(n)$ prenant leurs valeurs dans $\{1,\ldots,n\}$. Le Théorème 4 montre que, aussitôt que $a(n) \neq b(n)$, les deux variables aléatoires l_a et l_b sont indépendantes. La moyenne $\mathrm{E}_n(l_b/l_a)$ est alors $\mathrm{E}_n(l_b)\mathrm{E}_n(1/l_a)$.
Le Théorème 4 donne également les densités de l_a et de l_b. La définition d'une moyenne (3.11) appliquée au cas de l_a avec $h(x) = 1/x$, puis au cas de l_b avec $h(x) = x$ permet d'obtenir l'expression exacte de $E_n(\frac{l_b}{l_a})$ qui est donnée par l'équation suivante.

$$\mathrm{E}_n\left(\frac{l_b}{l_a}\right) = \frac{B(\frac{n-b+2}{2},\frac{b+1}{2})\ \ B(\frac{n-a}{2},\frac{a+1}{2})}{B(\frac{n-b+1}{2},\frac{b+1}{2})\ \ B(\frac{n-a+1}{2},\frac{a+1}{2})} = \frac{\Gamma(\frac{n-b+2}{2})}{\Gamma(\frac{n-b+1}{2})}\frac{\Gamma(\frac{n-a}{2})}{\Gamma(\frac{n-a+1}{2})}\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+3}{2})}\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})},$$

où $B(p,q)$ est une intégrale Beta, dont la définition est rappelée en (3.1). La dernière égalité s'obtient simplement en exprimant les intégrales Beta à l'aide de la fonction Gamma comme en (3.2).

Quand on s'intéresse aux rapports qui interviennent dans la réduction au sens de Siegel, on choisit $b(n) = a(n)+1$ et l'équation ci-dessus se simplifie. En particulier, la moyenne de la variable aléatoire l_{a+1}/l_a ne dépend pas de la fonction a. Donc quelque soit la fonction a de la dimension n (en particulier non-affine), la relation (3.14) donne une approximation de la valeur moyenne du rapport,

$$\mathrm{E}\left(\frac{l_{a+1}}{l_a}\right) = 1-\frac{1}{2n}+\frac{5}{8n^2}+\mathrm{O}\left(\frac{1}{n^3}\right). \tag{3.15}$$

Plus généralement, toujours en utilisant l'approximation de Stirling (3.14), on obtient les développements asymptotiques suivants pour $\mathrm{E}_n(l_b/l_a)$ en fonction de la dimension de l'espace ambiant.

Proposition 2. *Soient n vecteurs choisis uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Lorsque a et b sont des indices affines de n définis en (3.8) de pentes respectives α et β, la moyenne de la variable aléatoire l_b/l_a vérifie une des relations suivantes, en fonction de la position des indices a et b, i.e. les valeurs des pentes α et β.*

1. *Lorsque a et b sont des indices du début, i.e. $(\alpha,\beta) \in [0,1[^2$,*

$$\mathrm{E}_n\left(\frac{l_{\beta n+j}}{l_{\alpha n+i}}\right) = \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}\left(1+\frac{2i(1-\beta)-2j(1-\alpha)+\alpha+\beta-2\alpha\beta}{4(1-\alpha)(1-\beta)}\,\frac{1}{n}+O\left(\frac{1}{n^2}\right)\right)$$

En particulier, lorsque $\alpha = \beta$, et $i \neq j$, $\mathrm{E}_n\left(\frac{l_{\alpha n+j}}{l_{\alpha n+i}}\right) = 1 + \frac{\alpha + i - j}{2(1-\alpha)}\,\frac{1}{n} + O\left(\frac{1}{n^2}\right).$

2. *Lorsque a et b sont des indices du tout début, i.e. $(\alpha, \beta) = (0, 0)$,*

$$\mathrm{E}_n\left(\frac{l_j}{l_i}\right) = 1 + \frac{i-j}{2}\,\frac{1}{n} + \frac{5 + 4i^2 - (i+j-1)^2}{8}\,\frac{1}{n^2} + O\left(\frac{1}{n^3}\right).$$

3. *Lorsque a et b sont des indices de la fin, i.e. $(\alpha, \beta) = (1, 1)$,*

$$\mathrm{E}_n\left(\frac{l_{n-j}}{l_{n-i}}\right) = \frac{\Gamma(\frac{j+2}{2})\Gamma(\frac{i}{2})}{\Gamma(\frac{j+1}{2})\Gamma(\frac{i+1}{2})}\,\left(1 - \frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right).$$

Remarque 1. Les trois dernières relations concernent le cas où la différence $b(n) - a(n)$ est une constante (indépendante de n). Si $b(n) - a(n) = 1$, on retrouve naturellement le développement (3.15). Plus généralement, pour a et b deux indices du début, quand $b(n) - a(n)$ est une constante, la moyenne $\mathrm{E}(_b/l_a)$ tend toujours vers 1. Mais quand $a := n-i$ et $b := n-j$ sont deux indices de la fin, cette moyenne ne tend plus vers 1, mais vers une valeur proche de $\sqrt{(j+1)/i}$.

Remarque 2. Lorsque $b := n - j$ est un indice de la fin et $a := \alpha n + i$ est un indice du début, alors

$$\mathrm{E}_n\left(\frac{l_{n-j}}{l_{\alpha n+i}}\right) \sim \sqrt{\frac{j}{1-\alpha}}\,\frac{1}{\sqrt{n}}.$$

Le prochain paragraphe montrera que dans ce cas, la moyenne et l'écart–type sont de même ordre de grandeur.

3.3.3 La variance

Les mêmes considérations que pour le calcul de la moyenne $\mathrm{E}_n(l_b/l_a)$ permettent de calculer l'expression exacte de $\mathrm{E}_n(l_b^2/l_a^2)$. Si $a(n) \geq n - 1$, cette dernière moyenne diverge. Sinon, grâce encore au théorème 4 et après simplification,

$$\mathrm{E}_n(\frac{l_b^2}{l_a^2}) = \frac{n}{n+2}\,\frac{n-b+1}{n-a+1}.$$

On en obtient alors facilement un développement asymptotique en fonction de la dimension n du réseau, pour toutes les fonctions a et b considérées. Puis en appliquant la relation (3.12), et en utilisant les asymptotiques de la moyenne, trouvées dans le paragraphe précédent, les expressions asymptotiques de la variance sont facilement calculées. Dans les résultats qui suivent, les trois premiers cas concernent les cas où $a(n) = \alpha n + i$ et $b(n) = \beta n + j$, comme précédemment. Le dernier cas est utile uniquement pour la réduction au sens de Siegel. Ce cas permettra de préciser le nombre d'indices de la fin pour lesquels la condition de s-Siegel n'est pas satisfaite.
La proposition suivante donne les variances asymptotiques des variables (l_b/l_a).

Proposition 3. *Soient n vecteurs choisis uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Lorsque a et b sont des indices affines de n définis en (3.8), la variance de la variable aléatoire l_b/l_a vérifie une des relations suivantes, selon la position des indices a et b, i.e. les valeurs des pentes α, β de ces indices.*

1. *Lorsque a et b sont des indices du tout début, i.e. $(\alpha, \beta) = (0, 0)$,*
$$\sigma_n^2\left(\frac{l_j}{l_i}\right) = \frac{i+j+2}{2n^2} + O\left(\frac{1}{n^3}\right).$$
2. *Lorsque a est un indice du début, i.e. $\alpha \neq 1$,*
$$\sigma_n^2\left(\frac{l_b}{l_a}\right) = \frac{\alpha+\beta-2\alpha\beta}{2(1-\alpha)^2}\,\frac{1}{n} + O\left(\frac{1}{n^2}\right).$$
Le premier terme de l'expression asymptotique de la variance ne dépend pas de i et de j. Dans le cas particulier où $\alpha = \beta$ (et $i \neq j$),
$$\sigma_n^2\left(\frac{l_{\alpha n+j}}{l_{\alpha n+i}}\right) = \frac{\alpha}{(1-\alpha)}\,\frac{1}{n} + O\left(\frac{1}{n^2}\right).$$
3. *Lorsque a et b sont des indices de la fin, i.e. $(\alpha, \beta) = (1, 1)$,*
$$\sigma_n^2\left(\frac{l_{n-j}}{l_{n-i}}\right) = \frac{j+1}{i-1} - \left(\frac{\Gamma(\frac{j+2}{2})\Gamma(\frac{i}{2})}{\Gamma(\frac{j+1}{2})\Gamma(\frac{i+1}{2})}\right)^2 + O\left(\frac{1}{n}\right).$$
En particulier, pour $j = i-1 \geq 1$, l'expression ci-dessus se simplifie et
$$\sigma_n^2\left(\frac{l_{n-i+1}}{l_{n-i}}\right) = \frac{1}{i-1} + O\left(\frac{1}{n}\right).$$

Lorsque a et b sont des indices non–affines. Plus précisément, quand $a(n)$ vérifie
$$a(n) := n - f(n) \quad \textit{avec} \quad \lim_{n\to\infty} f(n) = \infty \quad \textit{et} \quad f(n) = o(n) \tag{3.16}$$
et $b(n) := a(n) + 1$,
$$\sigma_n^2\left(\frac{l_{b(n)}}{l_{a(n)}}\right) = \frac{1}{f(n)} + O\left(\frac{1}{f(n)^2}\right). \tag{3.17}$$

Remarque 3. Dans la relation (3.16), la fonction $f(n)$ peut par exemple être choisie la fonction logarithme, itérée k fois.

Il est important de noter que la variance de l_b/l_a tend vers 0 avec n, à l'exception du cas où l'indice a est un indice de la fin (a *très proche* de la dimension de l'espace ambiant n, i.e. la pente α est égale à 1).

En particulier, pour $b = a + 1$, l'expression asymptotique de la variance croît en fonction de la pente commune α de a et de b, et lorsque cette pente est égale à 1, la variance tend vers une limite non nulle. Autrement dit, lorsque l'indice a tend vers n, les variables aléatoires l_{a+1}/l_a sont de plus en plus dispersées par rapport à leur moyenne qui d'après (3.15), est toujours asymptotiquement égale à 1.

Plus généralement, lorsque a ou b est un indice de la fin, la moyenne et l'écart–type de la variable l_a/l_b sont de même ordre de grandeur, autrement dit la moyenne n'est plus très significative.

3.3.4 Évaluation approximative de la fonction de répartition

Dans les paragraphes précédents, nous avons décrit les comportements asymptotiques de l'espérance et de la variance des variables aléatoires l_b/l_a. Ici, l'inégalité de Bienaymé-Chebytechev (3.13) permet d'en déduire une évaluation grossière de la répartition asymptotique de ces variables. Comme dans les paragraphes précédents, on distingue quatre cas selon la position des indices a et b dans l'intervalle $\{1, \dots, n\}$.

Proposition 4. *Soient n vecteurs $b_1, b_2, ..., b_n$ pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Soient deux indices affines de n, a et b définis en (3.8) et de pentes respectives α et β. Alors, selon la position des indices a et b, i.e. les valeurs de α et de β, la variable aléatoire l_b/l_a vérifie une des relations suivantes.*

1. a et b des indices du tout début, i.e. $(\alpha, \beta) = (0, 0)$. Pour $n > n_0(i, j)$,

$$\Pr\left\{\left|\frac{l_j}{l_i} - 1\right| < \frac{\sqrt{i+j+2}}{\sqrt{n}}\right\} > 1 - \frac{1}{n}. \tag{3.18}$$

2. a et b des indices du début et $(\alpha, \beta) \neq (0, 0)$, i.e. $(\alpha, \beta) \in]0, 1[^2$. Pour $n > n_0(\alpha, \beta)$,

$$\Pr\left\{\left|\frac{l_{\beta n+j}}{l_{\alpha n+i}} - \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}\right| < \frac{\sqrt{\alpha+\beta-2\alpha\beta}}{(1-\alpha)} \frac{1}{n^{\frac{1}{4}}}\right\} > 1 - \frac{1}{\sqrt{n}}. \tag{3.19}$$

3. a et b des indices de la fin, i.e. $(\alpha, \beta) = (1, 1)$ et $j = i - 1 \geq 1$. Pour $n > n_0(i)$.

$$\Pr\left\{\left|\frac{l_{n-i+1}}{l_{n-i}} - 1\right| < \frac{1}{2}\right\} > 1 - \frac{4}{i-1}. \tag{3.20}$$

Lorsque a et b sont des indices non–affines avec $a(n) := n - f(n)$, vérifiant la condition (3.16) et $b(n) := a(n) + 1$, alors pour $n > n_0(f)$,

$$\Pr\left\{\left|\frac{l_{n-f(n)+1}}{l_{n-f(n)}} - 1\right| < \frac{1}{\sqrt{f(n)}}\right\} > 1 - \frac{1}{f(n)}. \tag{3.21}$$

Remarque. Toutes les vitesses de convergence vont être améliorées par une étude précise de la fonction de répartition asymptotique (Chapitre 5).

Preuve. Toutes les relations énoncées sont obtenues en utilisant les expressions asymptotiques de la moyenne et de la variance de la variable aléatoire considérée. Ces expressions ont été calculées dans les paragraphes précédents. Puis, on a choisi à chaque fois dans l'inégalité de Bienaymé-Chebytechev (3.13), une valeur particulière pour λ. Pour obtenir les relations (3.19), (3.18), (3.20), (3.21), il suffit de choisir pour λ respectivement $n^{1/4}$, $n^{1/2}$, $\sqrt{(i-1)/4}$, $\sqrt{f(n)}$.
Pour simplifier l'écriture, les valeurs choisies pour λ ne sont pas *optimales*. De toute façon, la suite de notre travail précisera ces résultats grâce à la méthode de Laplace. □

En conclusion, la variable aléatoire l_b/l_a relative à deux indices du début tend asymptotiquement vers sa moyenne. Ce dernier point est à l'origine des nouvelles réductions définies dans le chapitre 2.

La réduction de s-Siegel exige que tous les rapports l_{a+1}/l_a soient *à la constante s près* proches de leurs moyennes. Or pour un indice a de la fin, la moyenne et l'écart–type du rapport l_{a+1}/l_a sont *de même ordre de grandeur*. Autrement dit, pour un tel indice, ce rapport est beaucoup trop dispersé pour que la condition de s-Siegel soit satisfaite avec une probabilité satisfaisante. Afin d'avoir des algorithmes de réduction plus rapides en moyenne, notre idée est d'affaiblir fortement la contrainte sur le rapport l_b/l_a à mesure que les indices s'approchent de la fin (on accepte alors que ces rapports s'écartent de leur moyenne). Dans cet esprit, nous avons introduit

1. les réductions aux sens de (s, k)-Schmidt (2.10) où les rapports manipulés sont de la forme l_a/l_k et on cherche à les minorer par une quantité qui décroît avec a. Aussitôt que l'indice fixé k est suffisamment petit par rapport à la dimension de l'espace ambiant n, le chapitre 6 en montrera l'efficacité.
2. les réductions aux sens de (s, γ)-Gram (2.9) où les rapports manipulés l_{a+1}/l_a sont les mêmes que dans la réduction au sens de s-Siegel (2.8), mais là encore on cherche à minorer ces rapports par une quantité qui décroît avec a, de sorte à ce que ces minorations soient assez probablement réalisées, même quand l'indice a est proche de n. L'analyse du chapitre 6 montrera là encore l'efficacité de l'algorithme de (s, γ)-Gram réduction par rapport à l'algorithme de réduction au sens de s-Siegel (l'algorithme LLL original).
3. les réductions aux sens de (s, k, γ)-Gram-Schmidt, compromis des deux précédentes. Les bases réduites gardent presque les mêmes propriétés euclidiennes que les bases réduites au sens de t-Lovász, et elles sont obtenues plus rapidement en moyenne.

3.4 Les propriétés euclidiennes des bases aléatoires

Dans cette section, nous allons étudier les propriétés euclidiennes des bases aléatoires. Nous nous intéressons d'abord aux probabilités de réalisation des différentes conditions de réduction, puis aux caractéristiques plus générales de ces bases en rapport avec la théorie des réseaux: les défauts de longueur et le défaut d'orthogonalité. Nous donnons des majorations sur les moyennes de tous ces défauts pour les bases aléatoires du modèle choisi. Nous montrons en particulier que ces défauts sont liés au rapport entre la dimension du réseau et celle de l'espace ambiant.

3.4.1 Probabilité de réalisation d'une condition $\mathcal{C}_{\mathrm{m}}$, pour un indice affine du début

L'évaluation de la section précédente sur la répartition des rapports l_b/l_a fournit des informations sur la probabilité asymptotique que pour un indice affine m, la condition $\mathcal{C}_m$ soit satisfaite:

Proposition 5. *Soit une base aléatoire $b = (b_1, \dots b_n)$. Soit m un indice affine du début et $\mathcal{C}$ une des conditions suivantes introduites par la définition 1:*

(i) La condition de s-Siegel (2.8) *ou t-Lovász*[2] (2.7)

2. En toute rigueur, $\mathcal{C}$ n'est pas la condition de t-Lovász. Dans ce cas, il faut supposer en plus que les

(ii) La condition de (s, γ)-Gram

(iii) La condition de (s, k)-Schmidt (2.10) *ou (s, k, γ)-Gram–Schmidt* (2.11), *avec k un indice affine du début.*

Alors, lorsque n tend vers l'infini, la probabilité que la condition $\mathcal{C}_m$ soit satisfaite tend vers 1:

$$\Pr\{\textit{La condition } \mathcal{C}_m \textit{ est satisfaite}\} \to 1, \quad \textit{lorsque } n \to \infty.$$

Les relations suivantes, précisent les vitesses de convergence dans tous les cas.
Lorsque i et j sont deux indices du tout début vérifiant $i < j$, il existe $n_0(i, j)$, tel que pour tout $n > n_0(i, j)$,

$$\Pr\left\{\frac{l_j}{l_i} > s^{j-i}\right\} > 1 - \frac{i+j+2}{(1-s^{j-i})^2}\,\frac{1}{n^2}. \tag{3.22}$$

Lorsque $a := \alpha n + i$ et $b := \beta n + j$ sont deux indices distincts du début vérifiant $\alpha \le \beta$ et $i \le j$, il existe $n_0(a, b)$, tel que pour tout $n > n_0(a, b)$,

$$\Pr\left\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} > s^{(\beta-\alpha)n+(j-i)}\right\} > 1 - \left(\frac{\alpha}{1-\alpha} + \frac{\beta}{1-\beta}\right)\left(1 - \frac{\sqrt{1-\alpha}}{\sqrt{1-\beta}}\, s^{(\beta-\alpha)n+(j-i)}\right)^{-2} \frac{1}{n} \tag{3.23}$$

Remarque 4. Les relations (3.22) et (3.23) montrent que les vitesses de convergence sont polynomiales en la dimension n de l'espace ambiant. Notre étude plus précise du Chapitre 6 montrera que ces vitesses de convergence sont *exponentielles* en n.

Preuve. Lorsque E'_n est un minorant asymptotique de la moyenne $\mathrm{E}_n(l_b/l_a)$ (Proposition 2) et σ'_n un majorant asymptotique de l'écart–type $\sigma_n(l_b/l_a)$ (Proposition 3), l'inégalité de Bienaymé–Chebytchev montre que

$$\text{pour tout } \lambda > 0, \quad \Pr\left\{\mathrm{E}'_n - \lambda\sigma'_n < \frac{l_b}{l_a}\right\} > 1 - \frac{1}{\lambda^2}.$$

Les relations (3.22) et (3.23) s'obtiennent ensuite en choisissant λ tel que $\mathrm{E}'_n - \lambda\sigma'_n$ soit supérieur à s^{j-i}, respectivement asymptotiquement supérieur à $s^{(\beta-\alpha)n+(j-i)}$.

Les conditions de s-Siegel, (s, γ)-Gram.
Lorsque m est un indice du tout début (i.e. m est une constante et ne dépend pas de n), il suffit de choisir $j := m + 1$ et $i := m$ dans la relation (3.22). De même, si m est un indice du début de la forme $\omega n + i$ avec $0 < \omega < 1$, on choisit $\alpha = \beta := \omega$ et $j := i + 1$ dans (3.23) pour déduire que la condition de s-Siegel est quasi-sûrement satisfaite, lorsque n tend vers l'infini. Le résultat reste vrai pour la condition de Gram qui est une conséquence de celle de Siegel (Lemme 1).

La condition de (s, k)-Schmidt.

---

coefficients sub-diagonaux $m_{i+1,i}$ de la matrice m de passage de b à b^* sont en valeur absolue inférieurs à 1/2. Pour l'étude de l'algorithme de t-Lovász réduction (LLL), cette hypothèse n'est pas importante, car l'algorithme effectue toujours le nécessaire pour que $|\, m_{i+1,i} \,| \le 1/2$ avant d'effectuer le test de t-Lovász, pour un indice i.

Lorsque m et k sont des indices du tout début (i.e. ne dépendent pas de n), c'est encore une simple conséquence de (3.22) en prenant $j := \max(m, k)$ et $i := \min(m, k)$.
Supposons maintenant que $m := \omega n + i$ avec $0 \leq \omega < 1$, et $k := \theta n + j$ avec $0 \leq \theta < 1$. On note alors $\beta := \max(\omega, \theta)$ et $\alpha := \min(\omega, \theta)$ [3]. En excluant le cas $\omega = \theta = 0$ qui est déjà traité, les hypothèses sur ω et θ entraînent $0 < \alpha < 1$. La relation (3.23) permet alors de conclure. □

La proposition précédente montre déjà combien les qualités euclidiennes d'une base aléatoire dépendent du rapport entre sa dimension et celle de l'espace ambiant. Les Sections suivantes vont confirmer ce fait. Ce rapport qu'on appelera le *poids* d'une base permet de définir les notions de base *pleine, complète, incomplète et triviale* de la façon suivante.

Définition 3. Soit b une famille libre de p vecteurs de $\mathbb{R}^n$.
Si p est un indice du tout début, alors b est appelée *une base triviale.*
Si p est un indice du début, alors b est appelée *une base incomplète.*
Si p est un indice de la fin , alors b est appelée *une base complète.*
Enfin, si $p = n$, b est appelée *une base pleine.* Sinon, b est *non–pleine.*
La limite quand n tend vers l'infini du rapport p/n, lorsqu'elle existe, est appelé *le poids* de la base b.

Le poids d'une base triviale est nulle, celui d'une base incomplète est une constante de l'intervalle $[0, 1[$. Le poids d'une base complète est égale à 1.
Les nouvelles notions de bases introduites à partir de celles d'indices affines (Définition 2) interviennent de façon naturelle dans la suite du chapitre. Par exemple, une conséquence immédiate de la proposition précédente est que si une base aléatoire est incomplète, i.e. de dimension $p := \omega n + i$ avec $\omega < 1$, alors pour tout indice affine $m \leq p$, la probabilité que la condition $\mathcal{C}_m$ soit satisfaite tend vers 1 avec n, quelle que soit la condition envisagée.

Corollaire 3. *Soit une base aléatoire incomplète de cardinal p. Pourvu que la dimension de l'espace ambiant soit assez grande, pour tout indice affine m, $m \leq p$, la condition $\mathcal{C}_m$ est quasi-sûrement satisfaite avec une vitesse de convergence en $1 - O_m(1/n)$.*

Une autre conséquence est la probabilité de réduction des bases triviales. Ces bases sont en effet de cardinal fini. Pour une telle base, il existe un nombre fini d'indices et ces indices sont tous des affines du tout début.
Corollaire 4. *Soit une base aléatoire b triviale dont les vecteurs sont choisis indépendamment et uniformément dans la boule unité de dimension n. Quand n tend vers l'infini, la probabilité que b soit réduite (au sens de toutes les réductions précédemment définies) tend vers 1 en $O(1/n^2)$.*
Pour un entier m fixé, le Corollaire 3 montre que pour tout $\varepsilon > 0$, il existe $n_0(\varepsilon, m)$ tel que pour toutes les bases aléatoires de dimensions supérieures à $n_0(\varepsilon, m)$,

$$\Pr\{\text{La condition } \mathcal{C}_m \text{ est satisfaite}\} > 1 - \varepsilon.$$

Il est clair que la fonction n_0 peut toujours être choisie strictement croissante en fonction de m. Elle admet donc une fonction réciproque strictement croissante. Autrement dit,

3. En changeant éventuellement les noms de i et de j, on peut toujours supposer que $i \leq j$.

Corollaire 5. *Pourvu que n soit assez grand, il existe pour chaque n, un rang m_0 tel que toutes les conditions $\mathcal{C}_m$, pour $m \leq m_0$ soient quasi-sûrement satisfaites avec une vitesse de convergence en $O(1/n^2)$.*

Remarque. Pour un indice de la fin $n - i$, on ne peut pas conclure que la probabilité que $\mathcal{C}_{n-i}$ est satisfaite tende vers 1. En effet, dans ce cas la moyenne et l'écart–type du rapport considéré sont de même ordre de grandeur et l'étude en moyenne n'est pas assez précise. Par exemple, la relation (3.20) n'est intéressante que pour les grandes valeurs de i.
Cependant, la relation (3.21) fournit une évaluation du nombre d'indices m *de la fin*, pour lesquels la condition de s-Siegel n'est pas quasi-certainement satisfaite. Il s'agit donc des indices m, tels que $n - f(n) \leq m \leq n - 1$, avec une fonction f qui remplit les conditions suivantes introduites en (3.16),

$$\lim_{n\to\infty} f(n) = \infty \quad et \quad f(n) = o(n).$$

En conclusion, l'étude en moyenne ne permet pas de classer les conditions de réduction précédemment définies. Le Chapitre 6 d'une part affine les résultats de la Proposition 5 (les vitesses de convergence ne vont plus être polynomiales mais exponentielles) et d'autre part classe les différentes conditions de réduction, en considérant les indices de la fin. Par exemple, pour la condition de s-Siegel, on montrera que pour tout s fixé, cette probabilité tend vers une constante qui dépend de s et qui est strictement inférieure à 1.

Dans la suite du chapitre, la base $b = (b_1, \ldots, b_p)$ est une base aléatoire dont les vecteurs sont choisis indépendamment et uniformément dans la boule unité de $\mathbb{R}^n$ $(p \leq n)$. Nous allons montrer qu'une telle base possède en moyenne de très bonnes propriétés euclidiennes. Plus précisément sa qualité est d'autant meilleure que le poids de la base est faible (la dimension p de la base est petite devant la dimension n de l'espace ambiant). Pour des raisons techniques qui seront expliquées par la suite, on distingue le cas des bases aléatoires pleines (de dimension n) des autres bases aléatoires.

3.4.2 Les défauts de longueur des bases aléatoires non–pleines

On rappelle que le k-ème $(k \leq p)$ défaut de longueur de la base $b = (b_1, \ldots, b_p)$ est défini par (2.2):

$$\mu_k(b) := \frac{|\ b_k\ |}{\Lambda_k(L)}, \qquad \text{où } \Lambda_k(L) \text{ est le } k\text{-ème minimum du réseau.}$$

Le Lemme 2 du Chapitre 2 fournit une majoration du k-ème $(k \leq p)$ défaut de longueur μ_k d'une base en fonction des normes des orthogonalisés. Comme par hypothèse, les vecteurs b_k sont tous de normes inférieures à 1,

$$\exists i \in \{k, \ldots, p\} : \quad \mu_k \leq 1/l_i.$$

Nous allons "probabiliser" la relation précédente pour en déduire d'une part une minoration de la fonction de répartition de μ_k et d'autre part, une majoration de sa moyenne.

$$\Pr\{\mu_k > v\} \leq \sum_{i=k}^{p} \Pr\{\frac{1}{l_i} > v\}. \tag{3.24}$$

$$\mathrm{E}[\mu_k] \leq \sum_{i=k}^{p} \mathrm{E}\left[\frac{1}{l_i}\right]. \tag{3.25}$$

La fonction de répartition

Pour minorer la fonction de répartition, il suffit de majorer la partie droite de l'inégalité (3.24). C'est l'objet du lemme suivant.

Lemme 5. *Soit un réel v vérifiant $v > \sqrt{n}$. Avec les notations habituelles*[4] ,

$$\sum_{i=k}^{p-1} \Pr\{l_{i+1} \leq 1/v\} \leq 2\frac{(p-k+1)n^{\frac{n}{2}}(1/v)^{n-p+1}}{(p-1)^{\frac{p-1}{2}}(n-p+1)^{\frac{n-p+1}{2}}} \leq e^{\frac{2}{3n}}\sqrt{\frac{e}{\pi}}(p-k+1)\sqrt{n}\frac{1}{v}.$$

Preuve. Le Théorème 4 permet d'évaluer la fonction de répartition de l_i,

$$\Pr\{l_{i+1} \leq 1/v\} = \frac{2}{B(\frac{n-i}{2}, \frac{i+2}{2})}\int_{x=0}^{1/v} (1-x^2)^{\frac{i}{2}} x^{n-i-1}\, dx.$$

Dans l'intégrale précédente, la quantité $(1-x^2)$ est trivialement majorée par 1 et l'intégrale Beta du dénominateur s'exprime à l'aide de la fonction Γ comme en (3.2). Après simplification, la fonction de répartition de l_i vérifie

$$\Pr\{l_{i+1} \leq 1/v\} \leq \frac{(1/v)^{n-i}\Gamma\left(\frac{n+2}{2}\right)}{\Gamma\left(\frac{i+2}{2}\right)\Gamma\left(\frac{n-i+1}{2}\right)}.$$

Puis, l'approximation classique de Stirling de la fonction Γ [WW27],

$$x^x e^{-x}\sqrt{2\pi x} \leq \Gamma(x+1) \leq x^x e^{-x}\sqrt{2\pi x}e^{1/12x}, \tag{3.26}$$

et le fait que

$$\text{pour tout } i \in \{k \dots p-1\} \subset \{1 \dots n-1\}, \quad i(n-i) \geq n-1,$$

montrent que la probabilité $\Pr\{l_i \leq 1/v\}$ est majorée par

$$\frac{e^{1/6n}}{\sqrt{\pi}}\sqrt{\frac{n}{n-1}}\frac{n^{\frac{n}{2}}}{i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}}\,(1/v)^{n-i}.$$

Or, pour $p \leq n$, l'étude des variations de la fonction $i \mapsto i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}\, v^{n-i}$ montre que la fonction est décroissante sur $i \in [0, p-1]$ et que son minimum est atteint en $i := p-1$, pourvu que $v > \sqrt{n}$. Plus précisément, le minimum de la fonction est atteint en $i_0 := nv^2/(1+v^2)$ et pour $v > \sqrt{n}$, $i_0 > n-1$. Ainsi, chaque terme de la somme précédente est majoré par le dernier terme et des développements asymptotiques élémentaires permettent alors d'obtenir le majorant du Lemme. □

4. Comme par construction de la base orthogonalisée $l_1 =\mid b_1 \mid$, le lemme 2 fournit a priori une majoration plus précise pour le premier défaut de longueur, $\mu_1 \leq l_1/l_i$, dans l'énoncé du Lemme 5, lorsque $k := 1$, on pourrait plutôt chercher à majorer la quantité

$$\sum_{i=1}^{p-1} \Pr\{\frac{l_{i+1}}{l_1} \leq \frac{1}{v}\}.$$

Mais cette précision ne change pas le majorant final.

De la relation (3.24) et du lemme 5, on déduit un minorant pour la fonction de répartition du k-ème défaut de longueur:

Théorème 5. *Soit une base aléatoire $b = (b_1, \ldots, b_p)$ ($p \leq n$) et μ_k, son k-ème défaut de longueur. Alors, lorsque $v > \sqrt{n}$,*

$$\Pr\{\mu_k > v\} \leq 2(p - k + 1)\sqrt{n}\frac{1}{v}. \tag{3.27}$$

Le théorème précédent est utile quand v est *assez grand* par rapport à la dimension n de l'espace ambiant. En particulier, lorsque la dimension p du réseau aléatoire est égale à n, et que l'on considère le premier défaut de longueur ($k = 1$), pour que la majoration ne soit pas triviale, v doit être choisi supérieur à $2n^{3/2}$.

Par ailleurs, les vecteurs de la base aléatoire sont choisis indépendamment et uniformément dans la boule unité de $\mathbb{R}^n$ et sont donc en moyenne de même norme. Comme les minima successifs λ_k forment par définition une suite croissante, et que $\mu_k :=| b_k | /\lambda_k$, il est alors naturel que le majorant donné par le théorème 5 soit de plus en plus *fin*, à mesure que k croît. D'ailleurs dans le cadre du modèle choisi, les défauts de longueur désignent en fait le rapport de la taille de la base (ici 1) par les minima successifs.

La moyenne

Nous allons utiliser la relation (3.25),

$$\mathrm{E}[\mu_k] \leq \sum_{i=k}^{p} \mathrm{E}\left[\frac{1}{l_i}\right].$$

Cependant, il faut considérer à part, le cas où $p = n$ (base aléatoire pleine) car la moyenne de $1/l_n$ diverge. Pour $i \leq n - 1$, la moyenne de $1/l_i$ se calcule aisément comme dans la section précédente (Proposition 2) et

$$\mathrm{E}[\mu_k] \leq \frac{\Gamma\left(\frac{n+2}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} \sum_{i=n-p}^{n-k} \frac{\Gamma\left(\frac{i}{2}\right)}{\Gamma\left(\frac{i+1}{2}\right)}.$$

Puis, la relation (3.14) entraîne[5]

$$\mathrm{E}[\mu_k] \leq \sqrt{n+1} \sum_{i=n-p-1}^{n-k-1} \frac{1}{\sqrt{i}}.$$

Finalement, en majorant cette dernière somme par une intégrale, pour $p \leq n - 2$, on obtient

$$\mathrm{E}[\mu_k] \leq 2\sqrt{n+1}\left(\sqrt{n-k-1} - \sqrt{n-p-2}\right). \tag{3.28}$$

5. La majoration (3.14) des termes dans la somme qui suit n'est valable que pour $i \geq 2$. Si $p = n-1$, il faut séparer le premier terme de la somme du reste (le premier terme vaut alors $\sqrt{\pi}$). En première approximation, cela ne modifie pas l'ordre de grandeur du résultat final.

Cherchant une évaluation asymptotique de $\mathrm{E}[\mu_k]$ selon les différents cas, on obtient les résultats suivants. La relation (3.28) et sa version *asymptotique* détaillée par le théorème 6 montrent avant tout qu'en moyenne les défauts de longueur d'une base aléatoire sont *au plus* linéaires en la dimension de l'espace ambiant: tous les défauts de longueur sont inférieurs à $2n$.

Théorème 6. *Soit un entier $p \leq n-1$. La moyenne $\mathrm{E}[\mu_k]$ du k-ème défaut de longueur des bases aléatoires de cardinal p choisies dans la boule unité de $\mathbb{R}^n$ vérifie une des majorations suivantes, selon la dimension p du réseau engendré.*

1. *Les bases aléatoires triviales (la dimension p ne dépend pas de la dimension de l'espace ambiant n): pour tout $k \in \{1, \ldots, p\}$,*

$$\mathrm{E}[\mu_k] < (p-k+1) + \left(\frac{p^2-k^2}{4} + \frac{3p-2k}{2} + \frac{5}{4}\right)\frac{1}{n} + O\left(\frac{1}{n^2}\right)$$

2. *Les bases aléatoires incomplètes et non triviales (la dimension p est de la forme $p := \omega n$ avec $0 < \omega < 1$). Pour tout indice du début k, tel que $k := \theta n$ avec $0 \leq \theta \leq \omega$,*

$$\mathrm{E}[\mu_k] < 2\left(\sqrt{1-\theta} - \sqrt{1-\omega}\right)n + \left(\sqrt{1-\theta} - \sqrt{1-\omega} + \frac{2}{\sqrt{1-\omega}} - \frac{1}{\sqrt{1-\theta}}\right) + O\left(\frac{1}{n}\right).$$

3. *Les bases aléatoires complètes de dimension $p := n - j$ où j est une constante entière vérifiant $j \geq 2$.*
   Si k est un indice du début, $k := \theta n$ avec $0 \leq \theta < 1$,

$$\mathrm{E}[\mu_k] < 2\sqrt{1-\theta}\, n - 2\sqrt{j-2}\sqrt{n} + \left(\sqrt{1-\theta} - \frac{1}{\sqrt{1-\theta}}\right) + O\left(\frac{1}{\sqrt{n}}\right).$$

   Si k est un indice de la fin, $k := n - r$, avec ici $r \geq j$,

$$\mathrm{E}[\mu_k] < 2\left(\sqrt{r-1} - \sqrt{j-2}\right)\sqrt{n} + O\left(\frac{1}{\sqrt{n}}\right).$$

4. *Les bases aléatoires complètes de dimension $p = n - 1$.*
   Si k est un indice du début, $k := \theta n$ avec $0 \leq \theta < 1$,

$$\mathrm{E}[\mu_k] < 2\sqrt{1-\theta}\, n + \sqrt{\pi/2}\,\sqrt{n} + \left(\sqrt{1-\theta} - \frac{1}{\sqrt{1-\theta}}\right) + O\left(\frac{1}{\sqrt{n}}\right).$$

   Si k est un indice de la fin, $k := n - r$,avec ici $r \geq 1$,

$$\mathrm{E}[\mu_k] < 2\left(\sqrt{r-1} + \sqrt{\pi/8}\right)\sqrt{n} + O\left(\frac{1}{\sqrt{n}}\right).$$

De nouveau, on remarque que pour les bases de cardinal p, le majorant donné pour la moyenne du k-ème défaut de longueur μ_k décroît avec l'indice k. C'est naturel puisque les vecteurs d'une base aléatoire du modèle adopté sont en moyenne de même taille, que les minima successifs λ_k forment par définition une suite croissante, et que $\mu_k := \mid b_k \mid / \lambda_k$.

Pour les bases aléatoires triviales, le majorant de μ_k décroît de p à 1.
Pour les bases aléatoires incomplètes, l'ordre du majorant décroît de n à une constante.
Enfin, pour les bases aléatoires complètes, l'ordre du majorant décroît de n à $\sqrt{n}$.

Voici à titre d'exemple une conséquence du point (1) du théorème. Si on choisit p vecteurs (p fixé) uniformément et indépendamment dans la boule unité $\mathcal{B}_n$ de dimension suffisamment élevée, alors le vecteur le plus court du réseau qu'ils engendrent, est en moyenne, au plus p fois plus court que la taille de la base. *Le p-ème minimum du réseau engendré est asymptotiquement de la taille de la base.*

Par ailleurs, le Théorème précédent montre que les défauts de longueur des bases incomplètes se détériorent à mesure que le nombre de vecteurs de la base s'approche de la dimension n de l'espace ambiant. C'est donc bien le poids de la base, ou encore le rapport entre la dimension de la base aléatoire et celle de l'espace ambiant qui influence notablement la qualité de la base aléatoire. Par exemple, supposons que l'on dispose d'une base aléatoire de cardinal p et que l'on choisisse encore p vecteurs (pour obtenir une base de cardinal $2p$). Naturellement dans le nouveau réseau, il apparaît des vecteurs plus courts que dans l'ancien.
Si la base aléatoire initiale (ou la base aléatoire “ doublée”) est une base triviale, alors d'après le premier point du théorème, en moyenne, dans le nouveau réseau il n'existe pas de vecteur qui soit plus de deux fois plus court que dans l'ancien réseau.
En revanche, si la base aléatoire initiale est incomplète, mais non triviale, ($p := \omega n$, avec $0 < \omega < 1/2$ pour que l'expérience soit possible), en moyenne il peut exister des vecteurs du nouveau réseau dont la norme est abaissée d'un facteur plus grand que 2 (par exemple, pour $\omega = 1/4$, jusqu'à un facteur de $2,18$).

Les chapitres suivants confirmeront la détérioration de la qualité d'une base lorsque son poids tend vers 1.
Notons enfin, que cette dégradation est perceptible au sein même des bases complètes, par exemple lorsque l'on passe d'une base de cardinal $n-2$ à une base de cardinal $n-1$. Dans les résultats des points 3 et 4 du Théorème précédent, les termes principaux sont les mêmes, mais les deuxièmes termes, en $\mathrm{O}(\sqrt{n})$, témoignent de cette dégradation.

3.4.3 Le défaut d'orthogonalité des bases aléatoires non–pleines

L'expression du défaut d'orthogonalité ρ d'une base est rappelée en (2.4) et comme tous les vecteurs d'une base aléatoire b (de dimension $p \le n$) sont de norme inférieure à 1,

$$\rho(b) := \prod_{i=1}^{p} \frac{|b_i|}{l_i} \le \prod_{i=1}^{p} \frac{1}{l_i}. \tag{3.29}$$

La relation précédente et l'indépendance des variables l_i (Théorème 4) fournissent une majoration pour la moyenne du défaut d'orthogonalité $\mathrm{E}[\rho]$ des bases aléatoires non–pleines [6], c'est à dire de cardinal $p \le n-1$:

$$\mathrm{E}[\rho] \le \prod_{i=1}^{p} \mathrm{E}[\frac{1}{l_i}].$$

6. Là encore, le cas des bases pleines (de cardinal $p = n$) est à traiter à part. Dans ce cas, on cherchera à évaluer la moyenne de $1/\rho$.

En utilisant les moyennes des $1/l_i$, un calcul de routine montre que

$$\mathrm{E}[\rho] \leq \left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})} \right)^p \frac{\Gamma(\frac{n-p}{2})}{\Gamma(\frac{n}{2})}.$$

Puis grâce aux approximations de Stirling (3.14 et 3.26) [7]

$$\mathrm{E}[\rho] \leq e^{\frac{p}{2} + \frac{1}{6(n-p-2)}} \left(\frac{n-p-2}{2} \right)^{\frac{n-p-1}{2}} \left(\frac{n-2}{2} \right)^{-\frac{n-1}{2}} \left(\frac{n+1}{2} \right)^{\frac{p}{2}}.$$

Le théorème suivant résume les différents cas en fonction de la dimension p de la base aléatoire.

Théorème 7. *La moyenne* $\mathrm{E}[\rho]$ *du défaut d'orthogonalité des bases aléatoires non–pleines choisies dans la boule unité de* $\mathbb{R}^n$ *vérifie une des majorations suivantes, selon la dimension* $p \leq n-1$ *de la base.*

1. *Les bases aléatoires triviales (la dimension* p *est une constante qui ne dépend pas de la dimension* n *de l'espace ambiant).*

$$\mathrm{E}[\rho] < 1 + \left(\frac{5}{4}p^2 + 4p + \frac{1}{6} \right) \frac{1}{n} + O\left(\frac{1}{n^2} \right).$$

2. *La dimension* $p \leq n-1$ *est de la forme* $p := \omega n$ *avec* $0 < \omega \leq 1$.

$$\mathrm{E}[\rho] < \left(e^{\frac{\omega}{2}} (1-\omega)^{\frac{1-\omega}{2}} \right)^n e^{\frac{\omega}{2} + \frac{1}{6(1-\omega)n-12}}.$$

La première assertion montre que lorsque la dimension de l'espace ambiant tend vers l'infini, une base aléatoire triviale est en moyenne orthogonale. Ce résultat explique (en le précisant) le premier point du Théorème 6. Dans un réseau engendré par une base orthogonale, les minima successifs sont les vecteurs de la base. La majoration du Théorème 6 sur le premier défaut de longueur d'une base triviale est à comprendre comme une majoration de la moyenne du rapport

$$\frac{\max_{1 \leq i \leq p} |b_i|}{\min_{1 \leq i \leq p} |b_i|}.$$

Par ailleurs, la quantité qui majore le défaut d'orthogonalité croît là encore avec la dimension p du réseau aléatoire pour atteindre $e^{n/2}$ lorsque $p = n-1$. La Figure 3.1 illustre la dégradation de la qualité moyenne d'une base aléatoire, lorsque son poids croît.

3.4.4 Le cas particulier des bases aléatoires pleines

Pour des raisons surtout techniques, les bases pleines sont traitées à part. Comme la moyenne de $1/l_n$ n'est pas définie, ici on ne peut pas mener tout à fait les mêmes calculs que précédemment. Pour contourner cette difficulté technique, on considère les quantités inverses des grandeurs classiques, défauts de longueur et d'orthogonalité. Il est clair que ces *grandeurs inverses* sont tout aussi représentatives de la qualité des bases.

7. En toute rigueur, les cas $p = n-1$ ou $n-2$ devraient être traités à part. Ces petites exceptions techniques ne seront pas détaillées ici. Les résultats finaux que nous énonçons restent bien sûr valables même dans ces cas.

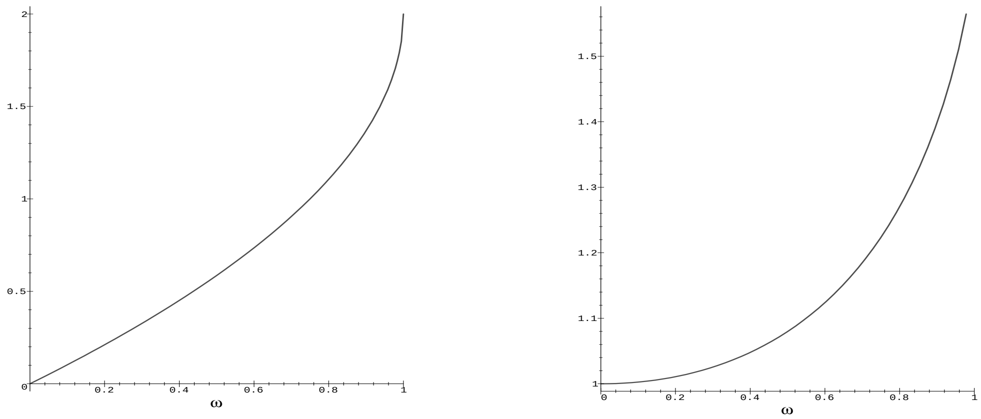

FIG. 3.1 – *Dégradation des bornes sur les moyennes du premier défaut de longueur et du défaut d'orthogonalité en fonction du poids ω de la base aléatoire. La courbe de gauche représente $2(1-\sqrt{1-\omega})$ (Théorème 6). La courbe de droite représente $e^{\omega/2}(1-\omega)^{(1-\omega)/2}$ (Théorème 7).*

Les inverses des défauts de longueur

Le Théorème 5 fournit une minoration de la fonction de répartition des défauts de longueur. Pour $0 < u < 1/\sqrt{n}$, la relation (3.27) entraîne

$$\Pr\{\frac{1}{\mu_k} > u\} \geq e^{\frac{2}{3n}} \sqrt{\frac{e}{\pi}}(n-k+1)\sqrt{n}\, u.$$

En notant $A := e^{2/3n}\sqrt{e/\pi}(n-k+1)\sqrt{n}$, la moyenne de $1/\mu_k$ vérifie alors,

$$\mathrm{E}[\mu_k] = \int_0^1 \Pr\{\frac{1}{\mu_k} > u\} du > \int_0^{1/A} (1 - Au) du = \frac{1}{2A}, \tag{3.30}$$

ce qui donne une évaluation de la moyenne des défauts de longueur.

L'inverse du défaut d'orthogonalité

Tous les vecteurs de la base aléatoire sont de normes inférieures à 1. Comme en (3.29),

$$\frac{1}{\rho(b)} \geq \det(\mathcal{L}) = \prod_{i=1}^{n} l_i.$$

On *probabilise* alors la relation précédente en utilisant les moyennes des l_i. De nouveau, comme les l_i sont indépendants, par un calcul de routine,

$$\mathrm{E}[\frac{1}{\rho}] \geq \left(\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+3}{2})}\right)^n \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{1}{2})}.$$

Les approximations de Stirling (3.14 et 3.26) permettent alors de donner une bonne évaluation de défaut d'orthogonalité.
Le théorème suivant résume le cas des bases aléatoires de dimension n.

Théorème 8. *Les moyennes* $\mathrm{E}[1/\mu_k]$ *et* $\mathrm{E}[1/\rho]$ *des inverses des défauts de longueur et d'orthogonalité des bases aléatoires pleines sont minorées par les relations suivantes.*

$$\mathrm{E}[1/\mu_k] \geq e^{-\frac{2}{3n}} \sqrt{\frac{\pi}{4e}}\ \frac{1}{(n-k+1)\sqrt{n}}.$$

$$\mathrm{E}[1/\rho] \geq \sqrt{2}\, e^{-\frac{n}{2}-1}.$$

Remarque. La minoration de la moyenne $\mathrm{E}[1/\rho]$ correspond exactement au résultat attendu, donné par le Théorème 7 (2). Mais le résultat sur les moyennes $\mathrm{E}[1/\mu_k]$ n'est malheureusement pas aussi précis. La borne attendue en $\mathrm{O}(n)$, s'est transformée en une borne en $\mathrm{O}(n^{3/2})$. Cette dernière dégradation est un problème de calcul: la minoration (3.30) est trop grossière. Je pense que la moyenne des défauts de longueur des bases aléatoires de dimension n reste en $\mathrm{O}(n)$.
On rappelle que Daudé et Vallée ont évalué [DV94] le premier minimum d'une base aléatoire pleine,

$$\Pr\{\Lambda_1 \leq u\} \leq 2u\sqrt{n}, \quad \text{pour tout } u \in [0,1] \quad \text{et} \quad \mathrm{E}(\Lambda_1) \geq \frac{1}{4\sqrt{n}}.$$

Dans cette section, nous avons évalué les principales caractéristiques moyennes des bases aléatoires. Lorsque l'on travaille dans un espace ambiant de dimension n, on note en particulier que les bornes exhibées par les théorèmes 6, 7 et 8 se dégradent à mesure que la dimension de la base aléatoire croît, que le poids de la base aléatoire tend vers 1. Cette dégradation n'est sans doute pas due uniquement au calcul. C'est aussi la qualité d'une base aléatoire qui se détériore avec sa dimension. Nous montrons dans le chapitre 6 que l'algorithme LLL(s) se contente de la qualité des bases incomplètes de dimension $p := \omega n$ avec $\omega \leq \omega_0(s)$. Mais lorsque $p = n - j$ (avec j une constante entière), bien que les bases aléatoires soient de qualité encore très satisfaisante, l'algorithme LLL(s) ne s'en contente plus.

3.5 Conclusion

Ce Chapitre a défini les bases de l'étude aléatoire conduite dans les chapitres suivants, en définissant le modèle probabiliste et en étudiant les caractéristiques intrinsèques de ce modèle.

L'étude en moyenne des rapports des orthogonalisés fournit déjà des résultats sur les probabilités de réductions de certaines bases aléatoires. (Proposition 5, Corollaires 3,4,5). Cette étude en moyenne sera précisée et complétée par une étude en distribution (Chapitres 4 et 5).

L'étude en moyenne des propriétés euclidiennes des bases aléatoires du modèle uniforme, outre son intérêt propre, fournit *une gamme de bases aléatoires de qualité variable.* Même si la portée de cette gamme est limitée à des bases qui sont toutes globalement de bonne qualité, elle nous permettra entre autres de séparer les différents algorithmes de réduction introduits dans le chapitre 2.

Chapitre 4

La méthode de Laplace

Ce Chapitre fournit l'outil principal de l'étude de la distribution asymptotique des rapports l_b/l_a: la méthode de Laplace d'évaluation des intégrales et des sommes [BO78, DeB81, SF96]. L'idée de la méthode de Laplace consiste en l'utilisation, sous des conditions que l'on précisera, d'*une connaissance précise mais locale* du terme général de l'intégrale ou de la somme, pour évaluer le comportement asymptotique de toute l'intégrale ou de toute la somme.
Dans ce chapitre, cette méthode permet surtout d'obtenir des équivalents asymptotiques pour certaines intégrales dépendant d'un paramètre. La méthode de Laplace pour les intégrales est d'abord rappelée dans le cadre usuel de la dimension 1. Puis je propose une généralisation de la méthode de Laplace au cas des intégrales doubles (et un cas particulier d'intégrale triple) dépendant d'un paramètre et prises sur des domaines suffisamment simples du plan. Ce chapitre s'achève sur une explication de l'esprit de cette démarche pour l'évaluation de sommes.
Dans le prochain chapitre, la méthode est appliquée à la recherche des comportements asymptotiques des fonctions de répartition et densités des variables aléatoires l_b/l_a. C'est ainsi que nous verrons apparaître des seuils (ou encore des lois 0-1) lorsque la dimension n de la boule unité de $\mathbb{R}^n$ dans laquelle les vecteurs de la base aléatoire sont choisis tend vers l'infini.

4.1 La méthode de Laplace pour les intégrales

Cette méthode s'applique aux intégrales générales de type:

$$I_n = \int_a^b F(x,n)dx$$

qui convergent. Elle propose une partie principale pour ces intégrales, lorsque n tend vers l'infini, dans le cas où le graphe de $x \mapsto F(x,n)$ admet *"un pic aigu"* quelque part dans l'intervalle $[a,b]$ qui peut être fini ou infini et peut dépendre de n. L'idée fondamentale est que moyennant quelques hypothèses sur F, la contribution d'un voisinage de ce pic dans I_n est équivalente à toute l'intégrale lorsque n est grand. Si au voisinage du maximum de la fonction F, on dispose d'un développement local de F par des fonctions simples, cette idée permet alors d'évaluer toute l'intégrale quand n tend vers l'infini. En définitive, une connaissance précise, *mais locale* de la fonction F aura fourni une estimation sur I_n. Dans un premier temps, on rappelle l'énoncé de deux résultats de la méthode de Laplace en une dimension.

4.1.1 Les intégrales simples

Les deux propositions suivantes, très classiques, sont extrêmement utiles en analyse élementaire. Elles permettent par exemple d'étudier les comportements asymptotiques des fonctions eulériennes. En particulier, la formule de Sirling est obtenue rapidement. Les deux propositions donnent la partie principale des intégrales I_n de la forme

$$I_n = \int_a^b g(x)e^{nh(x)}dx, \tag{4.1}$$

où l'intervalle $]a,b[$ peut être infini et peut éventuellement dépendre de n. Pour la proposition 6, on suppose que a est un nombre réel et que le maximum de la fonction h est atteint en a, i.e. sur un *bord* de l'intervalle $]a,b[$. La proposition 7 correspond au cas où le maximum de la fonction h est atteint *à l'intérieur* de l'intervalle $]a,b[$.

Proposition 6. *Soit l'intégrale I_n de la forme* (4.1) *où a est un nombre réel, $b > a$ est un nombre réel ou $+\infty$. Les fonctions g et h sont supposées réelles, continues par morceaux dans l'intervalle ouvert $]a,b[$, et vérifiant les conditions suivantes:*

1. *I_1 est absolument convergente.*
2. *Il existe un intervalle $[a,\delta_0]$ tel que, pour $0 < \delta < \delta_0$, on ait $h(x) \leq h(\delta)$ pour tout $x \geq \delta$.*
3. *Au voisinage de a, on a les développements asymptotiques suivants, avec $\lambda > -1$, $\rho > 0$, $C \neq 0$, $A > 0$:*

$$g(x) \sim C(x-a)^{\lambda}, \quad h(x) = h(a) - A(x-a)^{\rho} + o((x-a)^{\rho}).$$

Dans ces conditions, on a la partie principale

$$I_n \sim \frac{C}{\rho}\Gamma\left(\frac{\lambda+1}{\rho}\right) e^{nh(a)} \left(\frac{1}{An}\right)^{\frac{\lambda+1}{\rho}}$$

La proposition se limite volontairement à un cas très utile et fréquent.

Proposition 7. *Soit I_n une intégrale de la forme* (4.1). *On suppose que g et h sont des fonctions réelles et deux fois dérivables sur l'intervalle $]a,b[$ (borné ou non), que $\int_a^b |g(x)|e^{nh(x)}dx$ est définie et que h' ne change de signe qu'en un seul point c tel que $a < c < b$, où de plus h atteint un maximum et $g(c) \neq 0$, $h''(c) < 0$. Alors, lorsque n tend vers l'infini,*

$$I_n \sim \sqrt{\pi}g(c)e^{nh(c)}\sqrt{\frac{2}{-nh''(c)}}$$

4.1.2 Les intégrales doubles dépendant d'un paramètre

Soit Δ un polygône convexe du plan et

$$I_n = \int_{\Delta} g(x,y) e^{nh(x,y)}\, dxdy.$$

Nous nous proposons de rechercher un équivalent pour ces intégrales lorsque le paramètre n tend vers l'infini. Le seul cas que nous avons trouvé explicitement traité [DeB81] est celui où le maximum de la fonction h se trouve à l'intérieur du domaine Δ et où la fonction g est continue et non nulle en ce maximum. Or, dans l'étude asymptotique de la distribution de l_a/l_b, tous les autres cas seront rencontrés. En particulier le maximum de la fonction h peut être atteint soit à *l'intérieur*, soit sur *un bord* (un côté et un seul), soit sur un sommet (ou encore *un coin*) du polygône Δ. Pour conclure, 5 hypothèses (citées ci-dessous) seront suffisantes. Les 3 premières résument des conditions générales d'existence d'équivalent. Lorsqu'elles sont remplies, la contribution d'un certain voisinage du maximum de la fonction h est équivalente à toute l'intégrale quand n devient grand. L'hypothèse 4 regroupe les cas où les développements locaux sur les fonctions g et h au voisinage du maximum sont suffisamment simples pour permettre d'évaluer le comportement asymptotique de I_n en terme de fonctions usuelles. Ce sont par ailleurs, ces mêmes cas qui s'avéreront utiles pour l'étude en distribution des variables l_b/l_a. L'hypothèse 5 précise l'hypothèse 4 selon que le maximum de la fonction h se trouve à l'intérieur de Δ, sur un bord de Δ ou en un coin de Δ.

Proposition 8. *Soit I_n une intégrale définie par*

$$I_n = \int_{\Delta} g(x,y) e^{nh(x,y)}\, dxdy, \tag{4.2}$$

où Δ est un domaine du plan, g et h sont deux fonctions de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$.
On considère les cinq hypothèses suivantes $\mathcal{H}_1, \ldots, \mathcal{H}_5$.

$\mathcal{H}_1$. *Il existe $n_0 > 0$ tel que I_{n_0} soit convergente.*

$\mathcal{H}_2$. *La fonction h admet un maximum strict et absolu sur le domaine Δ, soit en (x_0, y_0).*

$\mathcal{H}_3$. *Il existe $c_0 > 0$ tel que*[1]*,*

$$\forall c < c_0, \quad \exists (x_1,y_1) \in \Delta \cap B_{(0,0),c} \quad \forall (x,y) \in \Delta \backslash B_{(0,0),c}, \quad h(x,y) \;\leq\; h(x_1,y_1) < 0.$$

$h(x_1,y_1)$ sera noté $-d$, avec $d > 0$.

$\mathcal{H}_4$. *Il existe deux formes linéaires X et Y des variables $(x-x_0, y-y_0)$ telles qu'au voisinage de (x_0,y_0) et à l'intérieur de Δ,*
(i) la fonction h vérifie la relation suivante , où H_1 et H_2 sont deux polynômes à une variable, de valuations respectives ρ, σ avec $\rho \geq 1$ et $\sigma \geq 1$,

$$h(x,y) - h(x_0,y_0) = \Theta\left(H_1(X) + H_2(Y)\right), \qquad \text{lorsque } (x,y) \longrightarrow (x_0,y_0), \tag{4.3}$$

1. L'hypothèse 3 est une hypothèse technique qui est systématiquement vérifiée lorsque par exemple les fonctions g et h sont différentiables sur le domaine Δ.

(ii) la fonction g vérifie la relation suivante où $C \neq 0$, $\lambda > -1$ et $\mu > -1$ sont des constantes réelles.

$$g(x,y) \sim C X^{\lambda} Y^{\mu}, \qquad \text{lorsque } (x,y) \longrightarrow (x_0,y_0). \tag{4.4}$$

$\mathcal{H}_5$. *Les formes linéaires Y, X et les entiers λ, μ doivent en plus vérifier des conditions supplémentaires selon que le maximum (x_0,y_0) est à l'intérieur, sur un côté ou en un sommet du polygône Δ.*

1. *Si le maximum (x_0,y_0) est atteint à l'intérieur de Δ, alors λ et μ doivent être des entiers pairs positifs.*
2. *Si le maximum (x_0,y_0) est atteint sur un bord de Δ, alors $Y = 0$ (la forme linéaire Y est introduite par l'hypothèse $\mathcal{H}_4$) doit être l'équation du bord (segment de droite) qui contient (x_0,y_0). De plus, λ doit être un entier pair positif.*
3. *Si le maximum (x_0,y_0) est atteint en un sommet de $\Delta(v)$, alors $X = 0$ et $Y = 0$ (les formes linéaires X et Y sont introduites par l'hypothèse $\mathcal{H}_4$) doivent être les équations des droites de la frontière qui contiennent (x_0,y_0).*

Sous les hypothèses $\mathcal{H}_1$ à $\mathcal{H}_5$, lorsque n tend vers l'infini,

$$I_n = \Theta\left(e^{nh(x_0,y_0)}\, n^{-\frac{\lambda+1}{\rho}-\frac{\mu+1}{\sigma}}\right). \tag{4.5}$$

Remarque 5. Le domaine Δ est a priori un domaine quelconque (même infini) du plan. La seule restriction sur Δ concerne le cas où le maximum de la fonction h est atteint sur la frontière de ce domaine. Dans ce cas, au voisinage du maximum, la frontière de Δ est

1. soit un segment de droite et le maximum est donc sur *un bord* de Δ.
2. soit deux segments de droite et le maximum se trouve en leur intersection ou encore sur *un coin* de Δ.

Remarque 6. La conclusion de la proposition précédente est en fait un équivalent précis pour I_n et la constante que nous omettons en utilisant la notation Θ est connue. Nous ne l'avons pas mentionnée, par souci de simplicité de l'écriture, l'énoncé précédent étant déjà assez technique. La valeur de cette constante (dans tous les cas) est donnée dans la preuve détaillée fournie en annexe.

Remarque 7. Les précisions de l'hypothèse $\mathcal{H}_5$ selon l'appartenance du maximum (x_0,y_0) à l'intérieur, à un bord ou à un coin de Δ concernent d'une part les formes linéaires X et Y, d'autre part les valuations λ et μ.

1. Les précisions sur les formes linéaires Y et/ou X introduites en $\mathcal{H}_4$. Elles sont nécessaires pour les preuves que nous présentons en annexe. Je pense que la conclusion (4.5) reste vraie, sans ces précisions. Plus exactement, l'équivalent final donné pour I_n ne devrait alors varier que d'une constante multiplicative.
2. Les précisions sur les nombres λ et/ou μ. Elles sont en réalité bien moins restrictives qu'elles ne paraissent:
   Supposons qu'au voisinage de (x_0,y_0) et à l'intérieur de Δ, la quantité $x - x_0$ (resp. $y - y_0$) puisse être négative et que g vérifie l'équation (4.4). Alors, comme la fonction g est réelle, nécessairement le nombre λ est un rationnel, et sa forme irréductible est de dénominateur impair. Pour que la conclusion (4.5) reste valable, il suffit que dans la forme irréductible du rationnel λ, le numérateur soit pair.

Remarque 8. Dans tous les cas qui nous intéresseront, les valuations ρ, σ des polynômes H_1, H_2 dans l'approximation locale (4.3) de la fonction h, seront des entiers valant 1 ou 2.

En pratique, l'approximation (4.3) est obtenue par un développement de Taylor de la fonction h au voisinage du maximum (x_0, y_0). Si la dérivée partielle de la fonction h en (x_0, y_0) ne s'annule selon aucune direction, le maximum de h n'est atteint en (x_0, y_0) que par la présence des bords et (x_0, y_0) est un coin de Δ. Il est alors aisé de développer localement la fonction h en (x_0, y_0) à l'aide des deux formes linéaires X et Y qui correspondent aux directions des bords qui contiennent (x_0, y_0). On pourra conclure aussitôt que g vérifie (4.4) avec les mêmes formes X et Y.

Dans les autres cas, le lemme suivant est utile.

Lemme 6. *Soit h une fonction de classe $\mathcal{C}^2$ en (x_0, y_0) et dont la dérivée partielle s'annule en (x_0, y_0) selon au moins une direction donnée, notée ici $Y(x - x_0, y - y_0) = 0$. Alors, il existe une forme linéaire X unique de variables $(x - x_0, y - y_0)$ et des réels, A, B, D, tels qu'au voisinage de $(0, 0)$:*

$$h(x, y) - h(x_0, y_0) = AX^2 + DY + BY^2 + o\left((x - x_0)^2 + (y - y_0)^2\right). \tag{4.6}$$

Si la dérivée de la fonction h en (x_0, y_0) s'annule selon une seule direction $(Y = 0)$, alors le maximum de h sur Δ est de nouveau dû à la présence des bords (un ou deux) et le point (x_0, y_0) est soit sur un bord, soit en un coin de Δ. Le lemme 6 permet alors d'obtenir le développement local de la fonction h en (x_0, y_0) à l'aide de la forme linéaire Y (qui correspond nécessairement à la direction d'un bord) et une autre forme X déterminée de façon unique.
Si le maximum est sur un bord, il suffit pour conclure que la fonction g s'écrive au voisinage de (x_0, y_0) comme en (4.4) avec les formes linéaires X et Y préalablement fixées par l'approximation de la fonction h.
Si le maximum est en un coin, il faut en plus que la forme X corresponde à la direction du deuxième bord.

Supposons enfin que le point (x_0, y_0) soit un maximum de nature elliptique.
Si le point (x_0, y_0) est à l'intérieur de Δ, alors la fonction h vérifie (4.3) et les formes linéaires X et Y ne sont pas uniques. On pourra conclure aussitôt qu'il existe un couple de formes X et Y telles qu'au voisinage de (x_0, y_0), les fonctions g et h se développent respectivement, comme en (4.4) et (4.3).
Si le point (x_0, y_0) est sur un bord de Δ, de nouveau les formes linéaires X et Y sont fixées par l'approximation de la fonction h: dans l'écriture (4.3) on fixe une des deux formes linéaires par exemple Y pour qu'elle corresponde à la direction du bord qui contient (x_0, y_0). L'autre forme X est alors uniquement déterminée (Lemme 6). Pour conclure, il faut en plus que g se développe comme en (4.4) avec les formes linéaires Y et X fixées préalablement par l'approximation de la fonction h.
Si le point (x_0, y_0) est en un coin de Δ, il faut déjà que la fonction h admette un développemnt de la forme (4.3) avec les formes X et Y correspondant aux directions des deux bords qui contiennent (x_0, y_0). En pratique, dans l'approximation de la fonction h, on fixe une des deux formes linéaires pour qu'elle corresponde à la direction d'un des deux bords . Il faut alors que l'autre, determinée de façon unique par le lemme 6, corresponde à la direction du

deuxième bord. De plus, g doit avoir un développement de la forme (4.4) avec les formes linéaires Y et X correspondant aux directions des bords.

Il existe un cas très particulier, important dans la suite où nous utiliserons l'équivalent précis pour I_n. D'ailleurs, ce cas ne rentre pas exactement dans le cadre de la proposition précédente. Les hypothèses $\mathcal{H}_4$ et $\mathcal{H}_5$ y sont modifiées pour utiliser un passage en coordonnées polaires. Ce cas fait l'objet de la proposition suivante.

Proposition 9. *Soit I_n une intégrale définie en (4.2) vérifiant les conditions $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$. On suppose en plus que*

1. *Le maximum de le fonction h sur le domaine d'intégration est atteint sur un coin de Δ et que l'angle intérieur de Δ en ce coin est aigu. On note v la tangente de cet angle.*
2. *Il existe des réels $A > 0$, $C \neq 0$, λ et μ , avec $\lambda + \mu > -2$, tels qu'au voisinage de (x_0, y_0) et à l'intérieur de Δ,*

$$\left\{ \begin{array}{lcl} h(x,y) & \sim & h(x_0,y_0) - A((x-x_0)^2 + (y-y_0)^2) \\ g(x,y) & \sim & C\,(x-x_0)^{\lambda}\,(y-y_0)^{\mu} \end{array} \right. . \tag{4.7}$$

Alors, lorsque n tend vers l'infini,

$$G(n) \sim \frac{C}{4}\,\Gamma(\frac{\lambda+\mu+2}{2})\,B\left(\frac{\mu+1}{2}, \frac{\lambda+1}{2}, \frac{v^2}{1+v^2}\right)\,e^{nh(x_0,y_0)}\,\left(\frac{1}{An}\right)^{\frac{\lambda+\mu+2}{2}}$$

où le dernier facteur est une intégrale Beta incomplète, dont la définition est rappelée en (3.3).

Les preuves. Une preuve détaillée est donnée en annexe. □

4.1.3 Les intégrales triples et multiples

La Proposition 8 se généralise en une dimension quelconque fixée, mais les hypothèses deviennent de plus en plus difficiles à réaliser. C'est pourquoi, nous n'énonçons pas de généralisation de cette proposition, d'autant qu'elle ne sera pas utile dans la suite.
En revanche, le chapitre 6 utilisera une généralisation très particulière de la proposition 9 en dimension 3, où les coordonnées polaires en dimension 2 sont remplacées par les coordonnées sphériques.

Proposition 10. *Soit un réel $s \in [0,1]$, un polygône convexe $\Omega(s)$ de l'espace,*

$$\Omega(s) = \left\{(x,y,z) \in [0,1]^3 \;:\quad y \geq sx \quad et \quad z \geq sy\right\}, \tag{4.8}$$

et l'intégrale I_n définie par

$$I_n = \int_{\Omega(s)} g(x,y,z) e^{nh(x,y,z)}\, dxdy. \tag{4.9}$$

On suppose que I_n vérifie les hypothèses $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$ de la proposition 8, ainsi que les hypothèses $\mathcal{H}_4$ et $\mathcal{H}_5$ redéfinies ci-dessous.

$\mathcal{H}_4$. *Le maximum de le fonction h sur le domaine d'intégration est atteint à l'origine*[2] *$(0,0,0)$ qui est un coin de Δ.*

2. Sans perte de généralité et pour simplifier les notations, on suppose que le point (x_0, y_0, z_0) est l'origine du repère

$\mathcal{H}_5$. *Il existe des réels $A > 0$, $C \neq 0$, λ et μ , avec $\lambda + \mu + \nu > -2$, tels qu'au voisinage de $(0,0,0)$ et à l'intérieur de Δ,*

$$\left\{ \begin{array}{lcl} h(x,y,z) & = & h(0,0,0) - A(x^2+y^2+z^2) + o(x^2+y^2+z^2) \\ g(x,y,z) & \sim & C\, x^\lambda\, y^\mu\, z^\nu \end{array} \right. . \tag{4.10}$$

Alors, lorsque n tend vers l'infini,

$$I_n \sim C\ \Gamma(\frac{\lambda+\mu+\nu+2}{2})\ \Upsilon(\lambda,\mu,\nu,s)\ e^{nh(0,0,0)} \left(\frac{1}{An}\right)^{\frac{\lambda+\mu+\nu+2}{2}},$$

$$\text{où} \quad \Upsilon(\lambda,\mu,\nu,s) := \iint_{\substack{\arctan s \le \theta \le \pi/2 \\ 0 \le \phi \le \pi/2 - \arctan(s\,\sin\theta)}} \cos^\lambda\theta\ \ \sin^\mu\theta\ \ \sin^{\lambda+\mu+1}\phi\ \ \cos^\nu\phi\ d\theta\, d\phi.$$

Remarque. En utilisant les coordonnées *hyper–sphériques*, la Proposition précédente se généraliserait à une dimension quelconque fixée.

4.2 La méthode de Laplace pour l'évaluation des sommes

De même que pour les intégrales, afin d'évaluer le comportement asymptotique d'une somme ou d'une série, l'idée de la méthode de Laplace consiste en l'utilisation d'*une connaissance précise mais locale* du terme général (information assez précise mais pas forcément valable sur tout l'intervalle à sommer). Cette méthode permet par exemple l'évaluation de la Q-fonction de Ramanujan [SF96]. En résumé, il s'agit de

1. localiser le maximum du terme général,
2. restreindre l'intervalle à sommer à un *petit* intervalle autour de l'indice pour lequel le maximum du terme général est atteint,
3. sur cet intervalle approximer le maximum et majorer (même grossièrement) les termes qui sont assez loin du maximum.
4. étendre l'intervalle restreint et majorer les nouveaux restes qui apparaissent.

Lorsque cette méthode fonctionne, le résultat est une somme plus simple à manipuler.
Par extension, lorsque l'on dispose d'un majorant non–uniforme pour le terme général, il peut être intéressant de décomposer la somme en plusieurs morceaux et de traiter chaque *sous–somme* séparément. Cette technique est efficace en particulier lorsque l'ordre de grandeur du maximum est *grand* par rapport au nombre de termes à sommer. Nous utiliserons cette démarche dans le chapitre 6 pour évaluer les probabilités asymptotiques de réductions des réseaux alétoires.

4.3 Les preuves

Dans cette Section, nous donnons les preuves des généralisations de la méthode de Laplace que nous avons proposées précédemment. Après de brèves préliminaires, nous montrons la Proposition 8, puis la Proposition 9. Enfin, on donne la preuve du Lemme 6.

4.3.1 Préliminaires

Aux moyens de changements de variables simples, on montre les équivalences suivantes.

– α entier pair positif, $b > 0$, $p > 0$,

$$n \to \infty \ : \qquad \int_{-b}^{b} x^{\alpha}\, e^{-npx^2} dx \ \sim \ \Gamma(\frac{\alpha+1}{2}) \left(\frac{1}{np}\right)^{\frac{\alpha+1}{2}}.$$

– $\alpha > -1$, $b > 0$, $p > 0$,

$$n \to \infty \ : \qquad \int_{0}^{b} x^{\alpha}\, e^{-npx^2} dx \ \sim \ \frac{1}{2}\Gamma(\frac{\alpha+1}{2}) \left(\frac{1}{np}\right)^{\frac{\alpha+1}{2}}.$$

– $\alpha > -1$, $b > 0$, $p > 0$,

$$n \to \infty \ : \qquad \int_{0}^{b} x^{\alpha}\, e^{-npx} dx \ \sim \ \Gamma(\alpha+1) \left(\frac{1}{np}\right)^{\alpha+1}.$$

4.3.2 Preuve de la Proposition 8

Lorsque Π est un domaine du plan, $I(\Pi)$ désigne dans cette preuve

$$\int_{\Pi} g(x,y) e^{nh(x,y)} \ dxdy.$$

On suppose que les deux polynômes H_1 et H_2 définis par l'hypothèse $\mathcal{H}_4$ vérifient

$H_1(X) \sim AX^{\rho}$ et $H_2(Y) \sim BY^{\sigma}$, lorsque $(x,y) \longrightarrow (x_0,y_0)$.

Les signes de A et de B sont déterminés de sorte à ce que (x_0,y_0) soit bien un maximum sur Δ. En multipliant éventuellement par $\exp(-nh(x_0,y_0))$, on peut toujours supposer que $h(x_0,y_0) = 0$. Soit $\mathcal{J}$ le Jacobien de la transformation $(x,y) \mapsto (X,Y)$ et $a \in \{0,1,2\}$ le nombre de côtés du polygône Δ contenant le maximum (x_0,y_0). Enfin, on définit $\Phi(n)$ par

$$\Phi(n) = 2^{-a} \frac{C}{\rho\sigma} \mathcal{J}\, \Gamma\left(\frac{\lambda+1}{\rho}\right) \Gamma\left(\frac{\mu+1}{\sigma}\right) \left(\frac{1}{An}\right)^{\frac{\lambda+1}{\rho}} \left(\frac{1}{Bn}\right)^{\frac{\mu+1}{\sigma}}.$$

Pour tout $\varepsilon > 0$, on peut successivement

1. déterminer un parallélogramme Π, tel que $(x_0,y_0) \in \Pi \subset \Delta$ et un nombre n_1 tels que pour tout $n > n_1$:

$$\left(1 - \frac{\varepsilon}{2}\right) \Phi(n) \le I(\Pi) \le \left(1 - \frac{\varepsilon}{2}\right) \Phi(n) \tag{4.11}$$

2. le parallélogramme Π et l'entier n_1 étant fixés, déterminer un nombre n_2 tel que pour tout $n > n_2$,

$$\int_{\Delta \backslash \Pi} |g(x,y)| e^{nh(x,y)} \ dxdy \le \frac{\varepsilon}{2} \Phi(n) \tag{4.12}$$

On en conclut que pour tout $n > \sup(n_1,n_2)$,

$$(1-\varepsilon)\Phi(n) \ \le \ I_n \ \le \ (1+\varepsilon)\Phi(n)$$

1. Pour tout $\eta \in [0,1]$, il existe des réels $\delta_1, \delta_2 > 0, \omega_1, \omega_2 > 0$ et un parallélogramme Π délimité par les quatre droites d'équations

$$\left\{ \begin{array}{rclcrcl} Y(x-x_0, y-y_0) & = & \delta_1 & ; & Y(x-x_0, y-y_0) & = & \delta_2 \\ X(x-x_0, y-y_0) & = & \omega_1 & ; & X(x-x_0, y-y_0) & = & \omega_2 \end{array} \right. ,$$

tels que $\forall (x,y) \in \Pi$,

$$\left\{ \begin{array}{rcccl} (1-\eta) C X^\lambda Y^\mu & \leq & g(x,y) & \leq & (1+\eta) C X^\lambda Y^\mu, \\ (1+\eta)\left(H_1(X) + H_2(Y)\right) & \leq & h(x,y) & \leq & (1-\eta)\left(H_1(X) + H_2(Y)\right) \end{array} \right. . \tag{4.13}$$

Plus précisément,

1. lorsque le maximum (x_0, y_0) est à l'intérieur du polygône $\Delta(v)$ (et non sur la frontière) alors $\delta_1 = -\delta_2$ et $\omega_1 = -\omega_2$ et les réels λ, μ sont des entiers positifs pairs.
2. lorsque le maximum (x_0, y_0) est sur un bord (et non en un coin) du polygône Δ, alors $Y = 0$ dans (4.3) est l'équation du bord du polygône qui contient (x_0, y_0). Dans ce cas, $\omega_1 = -\omega_2$, $\delta_1 = 0$, le réel λ est un entier pair positif et $\mu > -1$.
3. lorsque le maximum (x_0, y_0) est en un coin du polygône $\Delta(v)$, alors $X = 0$ et $Y = 0$ dans (4.3) sont les équations des deux côtés du polygône dont le point (x_0, y_0) est l'intersection. Dans ce cas, $\lambda > -1$ et $\mu > -1$

Les encadrements 4.13 pour les fonctions g et h fournissent alors des bornes pour l'intégrale $I(\Pi)$. Puis, le changement de variables $(x,y) \mapsto (X,Y)$ transforme le minorant et le majorant de $I(\Pi)$ en produit d'intégrales simples. Alors, grâce à la méthode de Laplace en une dimension (Propositions 6 et 7), pour tout $\varepsilon > 0$, il existe $n_1 > 0$, tel que pour $n > n_1$,

$$\sqrt{1-\varepsilon/2}\, \Phi(n)(1-\eta)(1+\eta)^{-\frac{\lambda+1}{\rho}-\frac{\mu+1}{\sigma}} \leq I(\Pi) \leq \sqrt{1-\varepsilon/2}\, \Phi(n)(1+\eta)(1-\eta)^{-\frac{\lambda+1}{\rho}-\frac{\mu+1}{\sigma}}$$

Pour finir, la continuité des fonctions de η utilisées entraîne la relation (4.11).

2. Les réels $\delta_1, \delta_2 > 0, \omega_1, \omega_2 > 0$ peuvent toujours être choisis de sorte à ce que le parallélogramme Π soit inclus dans la boule de centre (x_0, y_0) et de rayon c_0, défini par l'hypothèse $\mathcal{H}_3$. Donc, il existe $d > 0$, tel que

$$\forall (x,y) \in \Delta \backslash \Pi \qquad h(x,y) \leq -d, \quad \text{ou encore} \quad h(x,y) + d \leq 0.$$

Par la suite, pour $n \geq n_0$ (n_0 défini par l'hypothèse $\mathcal{H}_1$), on a

$$\left| \int_{\Delta \backslash \Pi(\delta,\mu)} g(x,y) e^{nh(x,y)} dxdy \right| \leq e^{-(n-n_0)d} \int_\Delta |g(x,y)| e^{n_0 h(x,y)} dxdy,$$

où la dernière intégrale est convergente en vertu de l'hypothèse $\mathcal{H}_1$ .On en déduit que

$$\int_{\Delta \backslash \Pi(\delta,\mu)} |g(x,y)| e^{nh(x,y)} dxdy \;=\; \mathrm{O}\left(e^{-nd}\right) \text{ avec } d > 0.$$

Pour tout $\varepsilon > 0$, il existe donc $n_2 > 0$, tel que pour $n > n_2$, la relation (4.12) soit vérifiée.$\Box$

4.3.3 Preuve de Proposition 9

Sans perte de généralité, on peut supposer que $(x_0, y_0) = (0,0)$. Un changement de variables simple permet toujours de s'y ramener. En multipliant éventuellement par $e^{h(O,O)}$, on peut aussi supposer que $h(0,0) = 0$. Plutôt que de choisir un parallélogramme suivant les directions des bords, au voisinage du maximum (comme dans la preuve de la Proposition 8), on choisira ici une portion de disque délimitée par les bords de Δ, de centre le point $(0,0)$ et de rayon τ suffisamment petit, pour pouvoir remplacer à l'intérieur de ce disque $g(x,y)$ et $h(x,y)$ par leurs équivalents en $(0,0)$. (Ce disque est noté $\mathcal{D}(\tau)$).
On applique alors la démarche détaillée dans la preuve précédente. Ici, la forme particulière de la fonction h au voisinage de (x_0, y_0) est adaptée à l'utilisation des coordonnées polaires. Plus précisément, on note

$$\Phi(n) \;=\; \frac{1}{2}\Gamma\left(\frac{\lambda+\mu+2}{2}\right)\left(\frac{1}{An}\right)^{\frac{\lambda+\mu+2}{2}} \int_0^{\arctan(v)} \cos^\lambda\theta \sin^\mu\theta d\theta\,,$$

et pour tout $\varepsilon > 0$, on va successivement:

1) mettre en évidence un disque $\mathcal{D}(\tau)$ et un nombre n_1 tels que pour tout $n > n_1$,

$$\left(1-\frac{\varepsilon}{2}\right)\Phi(n) \le \int_{\Delta\cap\mathcal{D}(\tau)} g(x,y)e^{nh(x,y)}\;dxdy \le \left(1-\frac{\varepsilon}{2}\right)\Phi(n)$$

2) Les nombres τ et n_1 étant fixés, déterminer un nombre $n_2 > 0$ tel que pour tout $n > n_2$ on ait:

$$\int_{\Delta\setminus\mathcal{D}(\tau)} |g(x,y)|e^{nh(x,y)}\;dxdy \le \frac{\varepsilon}{2}\Phi(n).$$

(La preuve de 2 est la même que celle de la preuve précédente)

On en conclut que pour tout $n > \sup(n_1, n_2)$, on a:

$$(1-\varepsilon)\Phi(n) \;\le\; G(n) \;\le\; (1+\varepsilon)\Phi(n).$$

1. Pour tout η élément de $[0,1]$, il existe $\tau > 0$ et un disque $\mathcal{D}(\tau)$ tels que pour tout couple (x,y) élément de $\Pi(v,\tau) := \Delta \cap \mathcal{D}(\tau)$, on a à la fois,

$$(1-\eta)x^\lambda y^\mu \;\le\; g(x,y) \;\le (1+\eta)x^\lambda y^\mu,$$

$$\text{et}\quad -(1+\eta)A\left(x^2+y^2\right) \le h(x,y) \le -(1-\eta)A\left(x^2+y^2\right).$$

L'intégrale $I(v,\tau)$ définie par

$$I(v,\tau) \;:=\; \int_{\Pi(v,\tau)} g(x,y)e^{nh(x,y)}dxdy,$$

est alors encadrée comme suit

$$(1-\eta)\int_{\Pi(v,\tau)} x^\lambda y^\mu e^{-nA(1+\eta)\left(x^2+y^2\right)}dxdy \le I(v,\tau) \le (1+\eta)\int_{\Pi(v,\tau)} x^\lambda y^\mu e^{-nA(1-\eta)\left(x^2+y^2\right)}dxdy.$$

Puis, grâce au passage en coordonnées polaires,

$$\begin{cases} x &= r\cos\theta \\ y &= r\sin\theta \end{cases} \quad \text{et} \quad \mathcal{J} \;=\; \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| \;=\; r,$$

les intégrales minorante et majorante précédentes deviennent des intégrales doubles à variables séparées. Par exemple, le minorant s'écrit $(1-\eta)\, I_1\, I_2$, avec

$$I_1 = \int_0^\tau r^{\lambda+\mu+1} e^{-\frac{n}{2}(1+\eta)r^2} dr \qquad \text{et} \qquad I_2 = \int_0^{\arctan(v)} \cos^\lambda \theta \sin^\mu \theta d\theta.$$

Les rappels préliminaires montrent alors que quand n tend vers l'infini,

$$I_1 \sim \frac{1}{2}\Gamma\left(\frac{\lambda+\mu+2}{2}\right)\left(\frac{2}{n}\right)^{\frac{\lambda+\mu+2}{2}} (1+\eta)^{-\frac{\lambda+\mu+2}{2}}.$$

Finalement, lorsque n tend vers l'infini, le minorant et le majorant de $I(v,\tau)$ admettent respectivement comme équivalents:

$$\Phi(n)(1-\eta)(1+\eta)^{-\frac{\lambda+\mu+2}{2}} \qquad \text{et} \qquad \Phi(n)(1+\eta)(1-\eta)^{-(\lambda+\mu+2)}.$$

Donc, $\forall \varepsilon > 0$, $\exists n_1 > 0$, $\forall n > n_1$,

$$\sqrt{1-\varepsilon/2}\,\Phi(n)(1-\eta)(1+\eta)^{-\frac{\lambda+\mu+2}{2}} \leq I(v,\tau) \leq \sqrt{1+\varepsilon/2}\,\Phi(n)(1-\eta)(1+\eta)^{-(\lambda+\mu+2)}.$$

Enfin, en vertu de la continuité des fonctions de η qui interviennent, on peut toujours choisir τ tel que pour n assez grand,

$$(1-\varepsilon/2)\Phi(n) \leq I(v,\tau) \leq (1+\varepsilon/2)\Phi(n).$$

Enfin, utilisant l'égalité $\sin^2(\arctan v) = v^2/(1+v^2)$ et le changement de variable $x := \sin^2\theta$, l'intégrale I_2 s'exprime au moyen d'intégrale Beta incomplète (3.3),

$$\int_0^{\arctan(v)} \cos^\lambda\theta \;\; \sin^\mu\theta \;\, d\theta = \frac{1}{2}B\left(\frac{\mu+1}{2}, \frac{\lambda+1}{2}, \frac{v^2}{1+v^2}\right).$$

□

4.3.4 Preuve de la Proposition 10

La preuve est similaire à la précédente: les coordonnées polaires en dimension 2 sont remplacées ici par les coordonnées sphériques en dimension 3.

4.3.5 Preuve du Lemme 6

Sans perte de généralité, et pour simplifier les notations dans la preuve, on peut toujours supposer que

$$(x_0, y_0) = (0,0) \quad \text{et} \quad h(x_0,y_0) = 0.$$

La fonction h est de classe $\mathcal{C}^2$ à l'origine et on pose

$$p = \frac{\partial h}{\partial x}(0,0) \qquad q = \frac{\partial h}{\partial y}(0,0)$$

$$r = \frac{\partial^2 h}{\partial x^2}(0,0) \qquad s = \frac{\partial^2 h}{\partial x \partial y}(0,0) \qquad t = \frac{\partial^2 h}{\partial y^2}(0,0),$$

et $Y(x,y) = -vx + y$, désignera la direction du segment de la frontière qui contient le maximum $(0,0)$. (Si cette direction est verticale, on pose $Y(x,y) = x$).
On va montrer que la forme linéaire X, et les réels A, B et D existent bien et s'expriment en fonction de p, q, r, s, t, v.
Comme h est de classe $\mathcal{C}^2$ en $(0,0)$, il existe un voisinage de l'origine inclus dans Δ, tel qu'à son intérieur:

$$h(x,y) = h(x,vx) + (y-vx)h'_y(x,vx) + \frac{(y-vx)^2}{2}h''_y(x,vx) + o\left((y-vx)^2\right).$$

En notant $\begin{cases} l_0(x) &= h(x,vx) \\ l_1(x) &= h'_y(x,vx) \end{cases}$,

on a respectivement

$$\begin{aligned} l'_0(x) &= \frac{\partial h}{\partial x}(x,vx) + v\frac{\partial h}{\partial y}(x,vx), \\ l''_0(x) &= \frac{\partial^2 h}{\partial x^2}(x,vx) + 2v\frac{\partial^2 h}{\partial x \partial y}(x,vx) + v^2\frac{\partial^2 h}{\partial y^2}(x,vx). \end{aligned}$$

Mais le point $(0,0)$ est un maximum sur le bord $y = vx$, et par la suite 0 est un maximum pour l_0 et

$$\begin{cases} l_0(0) &= 0 &= h(0,0), \\ l'_0(0) &= p + qv &= 0, \\ l''_0(0) &= r + 2vs + v^2 t &> 0. \end{cases}$$

Donc

$$h(x,vx) = \frac{1}{2}(r + 2vs + v^2 t)x^2 + o(x^2). \tag{4.14}$$

Par ailleurs, $$l'_1(x) = \frac{\partial^2 h}{\partial x \partial y}(x,vx) + v\frac{\partial^2 h}{\partial y^2}(x,vx),$$

et en particulier, $\begin{cases} l_1(0) &= q, \\ l'_1(0) &= s + vt. \end{cases}$

D'où

$$\mathrm{h}'_y(x,vx) = q + (s+vt)x + o(x). \tag{4.15}$$

Et h étant de classe $\mathcal{C}^2$ à l'origine,

$$h''_y(x,vx) = t(1 + \epsilon(x)). \tag{4.16}$$

Les relations (4.14), (4.15) et (4.16) montrent qu'il existe un voisinage de $(0,0)$ inclus dans Δ, tel que pour tout (x,y) appartenant à ce voisinage:

$$h(x,y) = \frac{r + 2vs + v^2 t}{2}x^2 + q(y-vx) + (s+vt)x(y-vx) + \frac{t}{2}(y-vx)^2 + o(x^2+y^2).$$

Et finalement,

$$h(x,y) = \frac{r + 2vs + v^2 t}{2}\left(x + \frac{s+vt}{r + 2vs + v^2 t}(y-vx)\right)^2$$

$$+q(y-vx)+\frac{tr-s^2}{2(r+2vs+v^2t)}(y-vx)^2+o(x^2+y^2).$$

D'où l'existence de la forme linéaire X et des réels A, B et D:

$$X(x,y)=\frac{r+vs}{r+2vs+v^2t}\,x+\frac{s+vt}{r+2vs+v^2t}\,y$$

$$A=\frac{r+2vs+v^2t}{2}\qquad B=\frac{tr-s^2}{(r+2vs+v^2t)}\qquad D=q.$$

Par ailleurs, le jacobien $\mathcal{J}$ de la transformation $(x,y)\mapsto(X,Y)$ est égal à 1.
Les signes de A,B,D sont determinés de sorte à ce que l'origine soit bien un maximum. $\square$

4.4 Conclusion

Ce chapitre a donc fourni la méthode–clé pour l'étude asymptotique en distribution des variables l_a/l_b. Les Propositions de ce chapitre vont en effet permettre de compléter l'analyse en moyenne des variables l_a/l_b que nous avons effectuée dans le Chapitre 3, par une analyse en distribution. D'une part, ils permettront de préciser toutes les vitesses de convergence pour les résultats déjà obtenus. D'autre part, ils s'avéreront *essentiels* pour évaluer la probabilité de réduction d'une base aléatoire pleine au sens de s-Siegel. Le prochain chapitre se base sur les Propositions que nous avons montrées ici pour étudier la distribution asymptotique des rapports l_a/l_b.

Pour finir le chapitre, précisons une limite de la méthode de Laplace, à laquelle nous serons confrontés (on utilisera alors des majorations plutôt que des équivalents):
le domaine d'intégration peut a priori dépendre du paramètre n qui tend vers l'infini, mais si l'aire du domaine tend exponentiellement vers 0 avec n, alors notre généralisation ne permet plus de conclure.

Chapitre 5

Distribution asymptotique des rapports des orthogonalisés. Phénomènes de seuil

Dans ce Chapitre, nous effectuons l'étude de la distribution asymptotique des variables $l_{b(n)}/l_{a(n)}$. Notre étude précise les résultats du Chapitre 3 et met en évidence des phénomènes de seuils dans les réseaux aléatoires. L'outil principal est la méthode de Laplace d'évaluation des intégrales et surtout sa généralisation en dimension deux que nous avons proposée dans le Chapitre précédent. Ici, la méthode de Laplace est appliquée d'abord à l'étude asymptotique de la fonction de répartition $G_{n,a,b}$, puis à celle de la densité $g_{n,a,b}$ des variables $l_{b(n)}/l_{a(n)}$, lorsque les vecteurs des bases considérées ont été choisis uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Les fonctions a et b sont toujours des fonctions affines de n comme définies en (3.8). On rappelle que lorsque $b(n) \neq a(n)$, les variables l_b et l_a sont indépendantes et la densité $f_{a,b}(x, y)$ du couple (l_a, l_b) est de la forme

$$\frac{g(x,y)e^{nh(x,y)}}{\int_{[0,1]^2} g(x,y)e^{nh(x,y)}dxdy},$$

$$\text{avec} \quad g(x,y) = \frac{(1-x^2)^{\frac{i-1}{2}}(1-y^2)^{\frac{j-1}{2}}}{x^i\, y^j}, \tag{5.1}$$

$$\text{et} \quad h(x,y) = \frac{\alpha}{2}\ln(1-x^2) + \frac{\beta}{2}\ln(1-y^2) + (1-\alpha)\ln(x) + (1-\beta)\ln(y). \tag{5.2}$$

L'expression exacte[1] de $G_{n,a,b}(v)$ fait intervenir les polygônes $\Delta(v)$ et $\Delta(\infty)$ définis par

$$\Delta(v) = \{(x,y) \in [0,1]^2 : y/x < v\} \quad \text{et} \quad \Delta(\infty) = [0,1]^2, \tag{5.3}$$

et s'écrit comme le rapport de deux intégrales génériques,

$$G(v) = \frac{I_n(v)}{I_n(\infty)}, \quad \text{où} \quad I_n(v) := \int_{\Delta(v)} g(x,y)e^{nh(x,y)}dxdy. \tag{5.4}$$

1. Lorsqu'il n'y a pas d'ambiguïté on omet les indices.

En dérivant $G(v)$, l'expression exacte de la densité s'obtient aisément. Pour $v \leq 1$, la densité s'écrit sous la forme suivante,

$$g_{n,(a,b)}(v) = \frac{1}{I_n(\infty)} \int_0^1 x g(x, vx) e^{nh(x,vx)} dx. \tag{5.5}$$

Et lorsque $v \geq 1$, la densité a une forme similaire,

$$g_{n,(a,b)}(v) = \frac{1}{v^2} \frac{1}{I_n(\infty)} \int_0^1 y\, g(\frac{y}{v}, y) e^{nh(\frac{y}{v},y)} dy. \tag{5.6}$$

Les relations précédentes montrent que la méthode de Laplace est bien adaptée à la recherche des parties principales de G et de g, lorsque n est grand. Par ailleurs, G et g ont exactement le même dénominateur qui est un coefficient de normalisation. Dans un premier temps, on déterminera la partie principale de la fonction de répartition, puis celle de la densité. Les résultats obtenus précisent l'étude en moyenne effectuée dans le Chapitre 3. Ils feront apparaître des phénomènes de seuils dans les réseaux aléatoires. Nous utiliserons ces résultats dans le Chapitre 6 pour évaluer les comportements en moyenne des algorithmes de réduction.

5.1 Comportement asymptotique de la variable l_a/l_b

La Proposition 8 du Chapitre 4 montre que les comportements asymptotiques des intégrales de la forme $I_n(v)$ et $I_n(\infty)$ qui interviennent dans l'expression (5.4) de la fonction de répartition G, dépendent fortement du maximum de la fonction h sur le domaine d'intégration. La fonction h définie en (5.2) atteint son maximum absolu sur $\Delta(\infty) = [0,1]^2$ en $(\sqrt{1-\alpha}, \sqrt{1-\beta})$. Il apparaît alors trois comportements asymptotiques différents pour la fonction de répartition G (et deux comportements différents pour la densité g), selon la localisation de ce point par rapport au domaine $\Delta(v)$.

1. Le point $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ n'appartient pas à $\Delta(v)$. On montre dans ce cas que la fonction h a un maximum strict et global sur $\Delta(v)$ et que ce maximum est atteint en un point du bord $y = vx$ du polygône $\Delta(v)$, soit en $(\xi, v\xi)$. Comme $\Delta(v)$ est par définition inclus dans $\Delta(\infty)$, ce maximum vérifie clairement $h(\xi, v\xi) < h(\sqrt{1-\alpha}, \sqrt{1-\beta})$. Lorsque toutes les hypothèses d'application de la Proposition 8 sont satisfaites, la fonction de répartition tend vers 0, exponentiellement avec n.
   De plus, grâce aux Propositions 6 et 7, il en est de même pour la densité g. Plus précisément, la quantité $D_{\alpha,\beta,v} := \exp\big(h(\xi, v\xi) - h(\sqrt{1-\alpha}, \sqrt{1-\beta})\big)$, vérifie $D_{\alpha,\beta,v} < 1$ et nos résultats du Chapitre 4 montrent qu'il existe $p_{(i,j)}$, tels que

$$\left\{ \begin{array}{lcl} G_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \;\; n^{p_{(i,j)}}\right] \\ g_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \;\; n^{p_{(i,j)}+1}\right] \end{array} \right.$$

2. Le point $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ appartient à $\Delta(v)$, mais n'appartient pas à la frontière $y = vx$ de $\Delta(v)$. La fonction h admet alors le même maximum sur le domaine $\Delta(v)$ et sur $\Delta(\infty)$. De plus, le point $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ a le même voisinage en tant qu'élément de $\Delta(v)$ et en tant qu'élément de $\Delta(\infty)$. Lorsque les hypothèses de la Proposition 8

du Chapitre précédent sont remplies, le numérateur et le dénominateur de la fonction de répartition ont alors le même équivalent et G tend vers 1. De plus comme

$$G_{n,a,b}(v) = 1 - G_{n,b,a}(1/v),$$

la vitesse de convergence est exponentielle avec n. Par ailleurs, le maximum de h sur $\Delta(\infty)$ est encore strictement supérieur au maximum de h sur le segment $y = vx$. De nouveau, grâce aux Propositions 6 et 7, la densité g tend vers 0, exponentiellement avec n. En résumé, nos résulats du Chapitre précédent montrent qu'il existe D et p tels que

$$\left\{ \begin{array}{lcl} 1 - G_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \; n^{p(i,j)}\right] \\ g_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \; n^{p(i,j)+1}\right] \end{array} \right. .$$

3. Le point $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ appartient à la frontière $y = vx$ de Δ. Dans ce cas, la fonction h admet le même maximum sur le domaine $\Delta(v)$ et sur $\Delta(\infty)$. Mais le point $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ n'a pas le même voisinage comme élément de $\Delta(v)$ et comme élément de $\Delta(\infty)$. Lorsque n tend vers l'infini, la fonction de répartition tend alors en général vers une constante qui dépend de v et qui est dans l'intervalle $]0,1[$. Quant à la densité g, elle tend en général vers l'infini.

Dans la suite, les résultats précédents sont récapitulés sous forme de trois propositions. Les preuves détaillées sont fournies en annexe.

5.1.1 Cas où la fonction de répartition tend vers zéro

Proposition 11. *Soient $(b_1, b_2, ..., b_n)$, n vecteurs pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. On considère deux indices affines $a := \alpha n + i$ et $b := \beta n + j$. Alors, pour tout réel positif v tel que*

$$v < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}, \tag{5.7}$$

la fonction de répartition $G_{n,a,b}$ et la densité $g_{n,a,b}$ des variables aléatoires $l_{b(n)}/l_{a(n)}$ admettent les équivalents suivants lorsque n tend vers l'infini:

$$\begin{cases} G_{n,a,b}(v) &= \Theta\left[D^n_{(\alpha,\beta,v)}\ n^{p(i,j)}\right] \\ g_{n,a,b}(v) &= \Theta\left[D^n_{(\alpha,\beta,v)}\ n^{p(i,j)+1}\right] \end{cases}$$

Les deux Tableaux 5.1 et 5.2 résument les valeurs de $D_{(\alpha,\beta,v)}$ et de $p_{(i,j)}$, d'abord lorsque $v \leq 1$, puis lorsque $v > 1$. La quantité $D_{(\alpha,\beta,v)}$ est toujours strictement inférieure à 1. En tout point v vérifiant (5.7), la fonction de répartition et la densité tendent toutes deux exponentiellement vers 0, avec la dimension n de l'espace ambiant.

	rapport étudié	$D_{(\alpha,\beta,v)}$	$p_{(i,j)}$
$(\alpha,\beta) \in]0,1[^2$	$\frac{l_{\beta n+j}}{l_{\alpha n+i}}$	$\frac{e^{h(\xi_{(\alpha,\beta,v)}, v\xi_{(\alpha,\beta,v)})}}{e^{h(\sqrt{1-\alpha},\sqrt{1-\beta})}}$	$-\frac{1}{2}$
$\alpha \in]0,1[\quad \beta = 0$	$\frac{l_j}{l_{\alpha n+i}}$	$\frac{v}{\sqrt{2}} \frac{(1-\frac{\alpha}{2})^{1-\frac{\alpha}{2}}}{\left(\frac{1-\alpha}{2}\right)^{\frac{1-\alpha}{2}}}$	$\frac{j-1}{2}$
$\alpha = 1 \quad \beta = 0$	$\frac{l_j}{l_{n+i}}$	$\frac{v}{2}$	$\frac{-i+j-1}{2}$
$\alpha = 0 \quad \beta = 0$	$\frac{l_j}{l_i}$	v	$\frac{j-1}{2}$
$\alpha = 0 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_i}$	$\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}} \left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}$	$-\frac{1}{2}$
$\alpha = 1 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_{n+i}}$	$\frac{e^{h(\xi_{(\beta,v)}, v\xi_{(\beta,v)})}}{\beta^{\frac{\beta}{2}}\ (1-\beta)^{\frac{1-\beta}{2}}}$	$\frac{-i-1}{2}$

TAB. 5.1 – *Les valeurs de $D_{(\alpha,\beta,v)}$ et de $p_{(i,j)}$ lorsque $v < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$ et $v \leq 1$*

	rapport étudié	v	$D_{(\alpha,\beta,v)}$	$p_{(i,j)}$
$(\alpha,\beta)\in]0,1[^2$	$\frac{l_{\beta n+j}}{l_{\alpha n+i}}$	$v\geq 1$	$\frac{e^{h(\xi_{(\alpha,\beta,v)},v\xi_{(\alpha,\beta,v)})}}{e^{h(\sqrt{1-\alpha},\sqrt{1-\beta})}}$	$-\frac{1}{2}$
$\alpha\in]0,1[\quad\beta=0$	$\frac{l_j}{l_{\alpha n+i}}$	$\frac{1}{v}<\sqrt{1-\frac{\alpha}{2}}$	$\left(\frac{1-\frac{1}{v^2}}{\alpha}\right)^{\frac{\alpha}{2}}\left(\frac{\frac{1}{v^2}}{1-\alpha}\right)^{\frac{1-\alpha}{2}}$	$-\frac{1}{2}$
$\alpha\in]0,1[\quad\beta=0$	$\frac{l_j}{l_{\alpha n+i}}$	$\frac{1}{v}=\sqrt{1-\frac{\alpha}{2}}$	$\left(\frac{1-\frac{1}{v^2}}{\alpha}\right)^{\frac{\alpha}{2}}\left(\frac{\frac{1}{v^2}}{1-\alpha}\right)^{\frac{1-\alpha}{2}}$	$\frac{j-1}{4}$
$\alpha\in]0,1[\quad\beta=0$	$\frac{l_j}{l_{\alpha n+i}}$	$\frac{1}{v}>\sqrt{1-\frac{\alpha}{2}}$	$\frac{v}{\sqrt{2}}\frac{(1-\frac{\alpha}{2})^{1-\frac{\alpha}{2}}}{\left(\frac{1-\alpha}{2}\right)^{\frac{1-\alpha}{2}}}$	$\frac{j-1}{2}$
$\alpha=1\quad\beta=0$	$\frac{l_j}{l_{n+i}}$	$\frac{1}{v}<\sqrt{\frac{1}{2}}$	$\sqrt{1-\frac{1}{v^2}}$	$\frac{-i-1}{2}$
$\alpha=1\quad\beta=0$	$\frac{l_j}{l_{n+i}}$	$\frac{1}{v}=\sqrt{\frac{1}{2}}$	$\sqrt{1-\frac{1}{v^2}}$	$\frac{-2i+j-1}{4}$
$\alpha=1\quad\beta=0$	$\frac{l_j}{l_{n+i}}$	$\frac{1}{v}>\sqrt{\frac{1}{2}}$	$\frac{v}{2}$	$\frac{-i+j-1}{2}$
$\alpha=1\quad\beta\in]0,1[$	$\frac{l_{\beta n+j}}{l_{n+i}}$	$v\geq 1$	$\frac{e^{h(\xi_{(\beta,v)},v\xi_{(\beta,v)})}}{\beta^{\frac{\beta}{2}}\,(1-\beta)^{\frac{1-\beta}{2}}}$	$\frac{-i-1}{2}$

TAB. 5.2 – *Les valeurs de $D_{(\alpha,\beta,v)}$ et de $p_{(i,j)}$ lorsque $v<\frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$ et $v>1$*

5.1.2 Cas où la fonction de répartition tend vers une constante dépendant de v

Proposition 12. *Soient $(b_1, b_2, ..., b_n)$, n vecteurs pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. On considère deux indices affines $a := \alpha n + i$ et $b := \beta n + j$. Alors, pour tout réel positif v tel que*

$$v = \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}},$$

le Tableau 5.3 donne les équivalents de la fonction de répartition $G_{n,a,b}$ et la densité $g_{n,a,b}$ des variables aléatoires $l_{b(n)}/l_{a(n)}$ en v, lorsque n tend vers l'infini. Dans ce cas, la fonction de répartition tend vers une constante de l'intervalle $]0,1[$.

	rapport étudié	$G_{n,\alpha,\beta}(v)$	$g_{n,\alpha,\beta}(v)$
$(\alpha,\beta) \in]0,1[^2 \quad v = \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$	$\frac{l_{\beta n+j}}{l_{\alpha n+i}}$	$\sim \frac{1}{2}$	$\Theta(\sqrt{n})$
$\alpha = 0 \quad \beta \in]0,1[\quad v = \sqrt{1-\beta}$	$\frac{l_{\beta n+j}}{l_i}$	A_v	$\Theta(\sqrt{n})$
$\alpha \in]0,1[\quad \beta = 0 \quad v = \frac{1}{\sqrt{1-\alpha}}$	$\frac{l_j}{l_{\alpha n+i}}$	$A_{\frac{1}{v}}$	$\Theta(\sqrt{n})$
$\alpha = 0 \quad \beta = 0 \quad v = 1$	$\frac{l_j}{l_i}$	C	$\Theta(n)$
$\alpha = 1 \quad \beta = 1 \quad v \leq 1$	$\frac{l_{n+j}}{l_{n+i}}$	$\sim \frac{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}, \frac{v^2}{1+v^2}\right)}{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}$	$\sim \frac{4\, v^{-j}\, (1+v^2)^{\frac{i+j-2}{2}}}{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}$
$\alpha = 1 \quad \beta = 1 \quad v \geq 1$	$\frac{l_{n+j}}{l_{n+i}}$	$\sim \frac{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}, \frac{v^2}{1+v^2}\right)}{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}$	$\sim \frac{4\left(\frac{1}{v}\right)^{-j} (1+\frac{1}{v^2})^{\frac{i+j-2}{2}}}{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}$

TAB. 5.3 – *Le comportement asymptotique de G et de g en $v = \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$*

Les valeurs précises des constantes A et C ne sont pas calculées, puisqu'elles ne servent pas dans la suite. Par ailleurs, il est important de noter que hormis le cas où a et b sont tous deux des indices de la fin, i.e. $\alpha = \beta = 1$, la Proposition précédente donne les comportements asymptotiques de G et g, uniquement en un point isolé $v = \sqrt{1-\beta}/\sqrt{1-\alpha}$. Le cas $\alpha = \beta = 1$ est donc un cas dégénéré, comme le soulignera le Théorème 9.

5.1.3 Cas où la fonction de répartition tend vers 1

Proposition 13. *Soient $(b_1, b_2, ..., b_n)$, n vecteurs pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. On considère deux indices affines $a := \alpha n + i$ et $b := \beta n + j$. Alors, pour tout réel positif v tel que*

$$v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}, \tag{5.8}$$

la fonction de répartition $G_{n,a,b}$ et la densité $g_{n,a,b}$ des variables aléatoires $l_{b(n)}/l_{a(n)}$ admettent les équivalents suivants lorsque n tend vers l'infini:

$$\begin{cases} 1 - G_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \ n^{p(i,j)}\right] \\ g_{n,a,b}(v) & = & \Theta\left[D^n_{(\alpha,\beta,v)} \ n^{p(i,j)+1}\right] \end{cases}$$

Les deux Tableaux 5.4 et 5.5 résument les valeurs de $D_{(\alpha,\beta,v)}$ et de $p_{(i,j)}$, d'abord lorsque $v > 1$, puis lorsque $v \leq 1$. La quantité $D_{(\alpha,\beta,v)}$ est toujours strictement inférieure à 1. En tout point v vérifiant (5.8), la fonction de répartition tend vers 1, la densité tend vers 0 et les vitesses de convergence sont exponentielles en la dimension n de l'espace ambiant.

	rapport étudié	$D_{(\alpha,\beta,v)}$	$p_{(i,j)}$
$(\alpha,\beta) \in]0,1[^2$	$\frac{l_{\beta n+j}}{l_{\alpha n+i}}$	$\frac{e^{h(\xi_{(\beta,\alpha,\frac{1}{v})}, v\xi_{(\beta,\alpha,\frac{1}{v})})}}{e^{h(\sqrt{1-\beta},\sqrt{1-\alpha})}}$	$-\frac{1}{2}$
$\alpha = 0 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_i}$	$\frac{1}{\sqrt{2}v} \frac{(1-\frac{\beta}{2})^{1-\frac{\beta}{2}}}{\left(\frac{1-\beta}{2}\right)^{\frac{1-\beta}{2}}}$	$\frac{i-1}{2}$
$\alpha = 0 \quad \beta = 1$	$\frac{l_{n+j}}{l_i}$	$\frac{1}{2v}$	$\frac{i-j-1}{2}$
$\alpha = 0 \quad \beta = 0$	$\frac{l_j}{l_i}$	$\frac{1}{v}$	$\frac{i-1}{2}$
$\alpha \in]0,1[\quad \beta = 0$	$\frac{l_j}{l_{\alpha n+i}}$	$\left(\frac{1-\frac{1}{v^2}}{\alpha}\right)^{\frac{\alpha}{2}} \left(\frac{\frac{1}{v^2}}{1-\alpha}\right)^{\frac{1-\alpha}{2}}$	$-\frac{1}{2}$
$\alpha \in]0,1[\quad \beta = 0$	$\frac{l_{n+j}}{l_{\alpha n+i}}$	$\frac{e^{h(\xi_{(\alpha,\frac{1}{v})}, \frac{1}{v}\xi_{(\alpha,\frac{1}{v})})}}{\alpha^{\frac{\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}}}$	$\frac{-j-1}{2}$

TAB. 5.4 – *Les valeurs de $D_{(\alpha,\beta,v)}$ et de $p_{(i,j)}$ lorsque $v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$ et $v > 1$*

	rapport étudié	v	$D_{(\alpha,\beta,v)}$	$p_{(i,j)}$
$(\alpha,\beta)\ \in]0,1[^2$	$\frac{l_{\beta n+j}}{l_{\alpha n+i}}$	$v \leq 1$	$\frac{e^{h(\xi_{(\beta,\alpha,\frac{1}{v})},\frac{1}{v}\xi_{(\beta,\alpha,\frac{1}{v})})}}{e^{h(\sqrt{1-\beta},\sqrt{1-\alpha})}}$	$-\frac{1}{2}$
$\alpha = 0 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_i}$	$v < \sqrt{1-\frac{\beta}{2}}$	$\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}} \left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}$	$-\frac{1}{2}$
$\alpha = 0 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_i}$	$v = \sqrt{1-\frac{\beta}{2}}$	$\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}} \left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}$	$\frac{i-1}{4}$
$\alpha = 0 \quad \beta \in]0,1[$	$\frac{l_{\beta n+j}}{l_i}$	$v > \sqrt{1-\frac{\beta}{2}}$	$\frac{1}{\sqrt{2}v}\frac{\left(1-\frac{\beta}{2}\right)^{\left(1-\frac{\beta}{2}\right)}}{\left(\frac{1-\beta}{2}\right)^{\left(\frac{1-\beta}{2}\right)}}$	$\frac{i-1}{2}$
$\alpha = 0 \quad \beta = 1$	$\frac{l_{n+j}}{l_i}$	$v < \sqrt{\frac{1}{2}}$	$\sqrt{1-v^2}$	$\frac{-j-1}{2}$
$\alpha = 0 \quad \beta = 1$	$\frac{l_{n+j}}{l_i}$	$v = \sqrt{\frac{1}{2}}$	$\sqrt{1-\frac{1}{v^2}}$	$\frac{i-2j-1}{4}$
$\alpha = 0 \quad \beta = 1$	$\frac{l_{n+j}}{l_i}$	$v > \sqrt{\frac{1}{2}}$	$\frac{1}{2v}$	$\frac{i-j-1}{2}$
$\alpha \in]0,1[\quad \beta = 1$	$\frac{l_{n+j}}{l_{\alpha n+i}}$	$v \leq 1$	$\frac{e^{h(\xi_{(\alpha,\frac{1}{v})},\frac{1}{v}\xi_{(\alpha,\frac{1}{v})})}}{\alpha^{\frac{\alpha}{2}}\,(1-\alpha)^{\frac{1-\alpha}{2}}}$	$\frac{-j-1}{2}$

TAB. 5.5 – *Les valeurs de* $D_{(\alpha,\beta,v)}$ *et de* $p_{(i,j)}$ *lorsque* $v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$ *et* $v \leq 1$

5.2 Conclusion

Le théorème suivant résume les 3 Propositions précédentes.

Théorème 9. *Soient les vecteurs $b_1, b_2, \ldots, b_n$ choisis indépendamment et uniformément dans la boule unité $\mathcal{B}_n$ de $\mathbb{R}^n$ et l_a la norme du a-ème orthogonalisé.*

(i) Lorsque $a := \alpha n + i, b := \beta n + j$ sont deux indices affines du début, i.e. α et β sont des constantes réelles de l'intervalle $[0, 1[$ et i, j deux constantes entières, la fonction de répartition $G(v)$ du rapport l_b/l_a suit asymptotiquement avec n une loi $0-1$ *et la convergence est exponentielle en n. Le saut se produit pour $v = (\sqrt{1-\beta}/\sqrt{1-\alpha})$.*

(ii) Lorsque $a := n - i, b := n - j$ sont des indices de la fin, i.e. $\alpha = \beta = 1$ et i, j sont deux entiers positifs, la fonction de répartition $G(v)$ du rapport l_b/l_a ne suit plus une loi $0-1$*. Lorsque n tend vers l'infini, la limite de $G(v)$ est une fonction* continue *de i, j et v. Plus précisément, $G(v)$ tend vers une constante de l'intervalle* $]0, 1[$*, donnée par la relation suivante*

$$\Pr\{\frac{l_{n-j}}{l_{n-i}} < v\} \to \frac{B\left(\frac{i+1}{2}, \frac{j+1}{2}, \frac{v^2}{1+v^2}\right)}{B\left(\frac{i+1}{2}, \frac{j+1}{2}\right)}.$$

Remarque 9. Quand un des deux indices a, b est un indice de début et l'autre est un indice de la fin, on est toujours dans le (i) du Théorème précédent. Cependant la fonction de répartition de l_b/l_a ne suit plus vraiment une loi $0-1$, puisque le suat se produit soit en $v = 0$ soit en $v = \infty$. Par exemple, lorsque $a := n - i$ est un indice de la fin et b un indice du début, le saut se produit à l'infini. Autrement dit, dans ce cas, quelle que soit la valeur de v (indépendant de n), la fonction de répartition $G(v)$ tend vers 0 avec n.

Nous avons caractérisé de façon complète le comportement asymptotique des variables aléatoires l_b/l_a. En particulier en comparant les résultats de ce Chapitre avec ceux obtenus dans le Chapitre 3 par une étude asymptotique en moyenne, on constate que l'étude en distribution complète l'étude en moyenne: ce dernier restait trop vague sur les indices de la fin. Par ailleurs, dans tous les autres cas, les vitesses de convergence sont précisées: toutes ces vitesses de convergence sont exponentielles avec la dimension n de l'espace ambiant.

Le théorème précédent exprime un *double* phénomène de seuil qui a un intérêt en soi. D'une part, pour une paire d'indices du début, il énonce clairement un premier seuil (ou loi $0-1$), quant au rapport des orthogonalisés. Ce premier seuil montre notamment la convergence avec n de la concentration autour de la moyenne. Le deuxième seuil consiste en la différence du comportement asymptotique lorsque l'on passe d'une paire d'indices du début à une paire d'indices de la fin. Le théorème montre que l'on passe d'une loi $0-1$ à un phénomène continu. Nous verrons qu'alors, le moyenne n'est plus très significative. Notons pour finir, qu'afin de localiser plus précisément ce deuxième seuil, il faudrait sans doute une échelle plus fine que l'échelle affine.

Dans le chapitre suivant, nous allons utiliser les résultats précis de ce chapitre pour analyser les comportements des différents algorithmes de réduction présentés dans le Chapitre 2.

Les deux Sections suivantes fournissent les preuves détaillées des équivalents donnés pour la fonction de répartition et la densité.

5.3 Les preuves pour la fonction de répartition

Toutes les preuves sont basées sur la méthode de Laplace pour les intégrales et en particulier, les généralisations que nous en avons données pour les intégrales doubles dans le Chapitre 4 (Propositions 8 et 9). On remarque que les hypothèses $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ communes à ces deux Propositions sont vérifiées de la même façon pour tous les cas considérés. Les deux dernières hypothèses distinguent ces cas. Il est en particulier important de noter à chaque fois, la localisation du maximum sur le domaine d'intégration (intérieur, bord, coin), la valeur de la fonction h en son maximum ainsi que les valeurs des quatre constantes $\lambda, \mu, \rho, \sigma$, introduites par l'hypothèse $\mathcal{H}_4$.

5.3.1 Cas où $\mathbf{v} < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$

Les maxima de la fonction h sur $\Delta(v)$ et sur $\Delta(\infty)$ sont alors toujours atteints en des points différents.

• **Cas général:** $(\alpha, \beta) \in]\mathbf{0}, \mathbf{1}[$

Pour ce qui est du dénominateur, le maximum de la fonction h définie en (5.2) sur le domaine $[0,1]^2$ est bien de nature elliptique et il se trouve à l'intérieur du domaine d'intégration $[0,1]^2$. De plus la fonction g, définie en (5.1) est continue et non nulle en $(\sqrt{1-\alpha}, \sqrt{1-\beta})$. Toutes les hypothèses de la Proposition 8 sont réunies, avec en particulier

$$(\lambda, \mu, \rho, \sigma) = (0, 0, 2, 2).$$

Quand n tend vers l'infini, le dénominateur $\mathcal{D}$ vérifie

$$\mathcal{D} = \Theta\left(\left(\alpha^{\frac{\alpha}{2}}\, \beta^{\frac{\beta}{2}}\, (1-\alpha)^{\frac{1-\alpha}{2}}\, (1-\beta)^{\frac{1-\beta}{2}}\right)^n \left(\frac{1}{n}\right)\right). \tag{5.9}$$

On montre aisément que la maximum de la fonction h, sur le domaine $\Delta(v)$ se situe sur le bord $y = vx$. Pour évaluer la partie principale du numérateur $\mathcal{N}$, il faut situer ce maximum. Or,

$$\frac{\partial h(v, vx)}{\partial x} = \frac{2v^2x^4 + (-2v^2 + \beta + \alpha v^2 - 2)x^2 + (2 - \alpha - \beta)}{(1-x^2)(1-v^2x^2)x}.$$

Une simple étude des zéros de la dérivée précédente montre que $h(x, vx)$ atteint son maximum sur $[0,1]$ en ξ, avec $(\xi, v\xi) \in]0,1[^2$ (le maximum n'est pas en un coin) et ξ^2 vérifie

$$\xi^2 = \frac{1}{4v^2}\left((2-\alpha)v^2 + (2-\beta) - \sqrt{(2-\alpha)^2v^4 + (2-\beta)^2 + 2(2\alpha + 2\beta + \alpha\beta - 4)v^2}\right). \tag{5.10}$$

Le point $(\xi, v\xi)$ n'est pas un point critique pour h. Par ailleurs, g est continue et non nulle en $(\xi, v\xi)$. Toutes les hypothèses de la Proposition 8 sont réunies avec

$$(\lambda, \mu, \rho, \sigma) = (0, 0, 2, 1),$$

et lorsque n tend vers l'infini, le numérateur $\mathcal{N}$, de G vérifie

$$\mathcal{N} = \Theta\left(\left((1-\xi^2)^{\frac{\alpha}{2}}\, (1-v^2\xi^2)^{\frac{\beta}{2}}\, \xi^{\,(2-\alpha-\beta)} v^{(1-\beta)}\right)^n \left(\frac{1}{n}\right)^{\frac{3}{2}}\right)$$

Finalement, lorsque n est grand, la fonction de répartition $G_{n,(\alpha,\beta)}(v)$ a comme partie principale

$$A(a,b,v)\,(D_{\alpha,\beta,v})^n\ \frac{1}{\sqrt{n}}$$

avec $D_{\alpha,\beta,v} = \dfrac{e^{h(\xi,v\xi)}}{e^{h(\sqrt{1-\alpha},\sqrt{1-\beta})}} < 1$, car $\Delta \subset [0,1]^2$.

Ci-dessous, tous les cas particuliers sont traités successivement.

• **Premier cas particulier:** $\boldsymbol{\beta = 0}$ et $\mathbf{v < 1/\sqrt{1-\alpha}}$

a - $\boldsymbol{\beta = 0}$ et $\mathbf{0 < \alpha < 1}$.

Pour le dénominateur $\mathcal{D}$, le maximum est atteint en $(\sqrt{1-\alpha},1)$, soit sur un bord de $[0,1]^2$. Par le Lemme 6, il existe un voisinage de $(\sqrt{1-\alpha},1)$, inclus dans $[0,1]^2$, tel que pour tout (x,y) lui appartenant,

$$\begin{cases} h(x,y) & \sim & -A(x-\sqrt{1-\alpha})^2 - D(y-1) - B(y-1)^2 \\ g(x,y) & \sim & C(y-1)^{\frac{j-1}{2}} \end{cases}$$

(Ici, c'est particulièrement simple car $v=0$, $Y(x,y)=y$ et $X(x,y)=x$.)

La Proposition 8 avec

$$(\lambda,\mu,\rho,\sigma) = (0,(j-1)/2,2,1),$$

entraîne alors

$$\mathcal{D} = \Theta\left(\left[e^{h(\sqrt{1-\alpha},1)}\right]^n \left(\frac{1}{n}\right)^{\frac{j+2}{2}}\right) \quad \text{pour} n \to \infty. \tag{5.11}$$

Pour le numérateur $\mathcal{N}$ de G, le maximum de la fonction h sur le domaine d'intégration se situe toujours sur la droite $y = vx$ et l'étude du signe de

$$\frac{\partial h(x,vx)}{\partial x} = \frac{-2x^2+2-\alpha}{x(1-x^2)}$$

montre que ce maximum est précisément atteint en $(\sqrt{1-\frac{\alpha}{2}}, v\sqrt{1-\frac{\alpha}{2}})$.

La constante α étant ici non nulle, on a toujours $\sqrt{1-\alpha/2} < 1$. Trois cas sont à considérer:

(i) Pour $v<1$, ou pour $v>1$ et $\sqrt{1-\frac{\alpha}{2}} < \frac{1}{v}$, la Proposition 8 s'applique une fois de plus. Il existe d'après le Lemme 6 un voisinage de $(\sqrt{1-\alpha/2},1)$, inclus dans Δ et des constantes A, B, C et D, ainsi que deux formes linéaires X et $Y = -vx+y$ tels que pour tout (x,y) appartenant à ce voisinage, $h(x,y) \sim -AX^2 - DY - BY^2$. Par ailleurs, la fonction g est continue et non nulle en $(\sqrt{1-\alpha/2},1)$.

En résumé le maximum $(\sqrt{1-\alpha/2},1)$ est sur un bord de Δ, et la Proposition 8 avec

$$(\lambda,\mu,\rho,\sigma) = (0,0,2,1),$$

entraîne que lorsque n tend vers l'infini,

$$\mathcal{N} = \Theta\left(\left[e^{h(\sqrt{1-\frac{\alpha}{2}},v\sqrt{1-\frac{\alpha}{2}})}\right]^n \left(\frac{1}{n}\right)^{\frac{3}{2}}\right).$$

Et après simplification,

$$G = \Theta\left(\left[\frac{v}{\sqrt{2}}\,\frac{(1-\frac{\alpha}{2})^{(1-\frac{\alpha}{2})}}{\left(\frac{1-\alpha}{2}\right)^{\left(\frac{1-\alpha}{2}\right)}}\right]^n n^{\frac{j-1}{2}}\right).$$

(ii) Lorsque $v > 1$ et $\sqrt{1-\alpha/2} > \frac{1}{v}$, le maximum de la fonction h sur Δ est atteint en $(\frac{1}{v},1)$, c'est à dire sur un coin du domaine d'intégration. Par ailleurs, la dérivée en ce point ne s'annule pas selon les directions des bords de Δ auxquels il appartient. Il existe des réels A, B et C et un voisinage de $(\sqrt{1-\alpha/2},1)$, inclus dans Δ, tels que pour tout (x,y) lui appartenant,

$$\begin{cases} h(x,y) & \sim & -AX - BY \\ g(x,y) & \sim & C\,Y^{\frac{j-1}{2}} \end{cases}$$

La Proposition 8 avec

$$(\lambda,\mu,\rho,\sigma) = (0,(j-1)/2,1,1),$$

donne alors la partie principale du numérateur:

$$\mathcal{N} = \Theta\left(\left[e^{h(\frac{1}{v},1)}\right]^n \left(\frac{1}{n}\right)^{\frac{j+3}{2}}\right).$$

Et après simplification,

$$G = \Theta\left(\left[\left(\frac{1-\frac{1}{v^2}}{\alpha}\right)^{\frac{\alpha}{2}}\left(\frac{\frac{1}{v^2}}{1-\alpha}\right)^{\frac{1-\alpha}{2}}\right]^n \left(\frac{1}{\sqrt{n}}\right)\right).$$

(iii) En revanche, lorsque $v \geq 1$ et $\sqrt{1-\frac{\alpha}{2}} = \frac{1}{v}$, les raisonnements précédents ne permettent plus de donner une partie principale pour G: le maximum de h est sur un coin de Δ mais la fonction h ne se sépare plus en fonction des bords. On pourra toutefois tirer des conclusions similaires sur la limite de G en encadrant correctement la fonction h.

b - $\boldsymbol{\beta = 0}$ $\quad\boldsymbol{\alpha = 1}$

Ici, le maximum de la fonction h sur $[0,1]^2$ est atteint en $(0,1)$, soit encore sur un coin de ce domaine. La dérivée en $(0,1)$ s'annule selon la direction d'un des deux bords ($x=0$). Pour le dénominateur, on remarque qu'au voisinage de $(0,1)$ et à l'intérieur de $[0,1]^2$,

$$\begin{cases} h(x,y) & \sim & -A(y-1)^2 - Dx + Bx^2 \\ g(x,y) & \sim & C\,x^{-i}\,(1-y)^{\frac{j-1}{2}} \quad \text{où seules les valeurs négatives de } i \text{ ont un sens.} \end{cases}$$

Par ailleurs, $h(0,1)$ est nul. Donc, en vertu de la Proposition 8 avec

$$(\lambda,\mu,\rho,\sigma) = (-i,(j-1)/2,1,2),$$

le dénominateur $\mathcal{D}$ a comme partie principale

$$\mathcal{D} = \Theta\left(\left(\frac{1}{n}\right)^{\frac{j-i+2}{2}}\right), \quad \text{quand } n \to \infty. \tag{5.12}$$

Pour le numérateur, tout se passe comme précédemment, et on note qu'ici la dérivée de $x \mapsto h(x, vx)$ comme fonction de x s'annule en $1/\sqrt{2}$.

(i) Lorsque $v < 1$ ou lorsque $v \geq 1$ et $1/\sqrt{2} < 1/v$, le numérateur $\mathcal{N}$ vérifie

$$\mathcal{N} = \Theta\left(\left[e^{h(\frac{1}{\sqrt{2}}, \frac{v}{\sqrt{2}})}\right]^n \left(\frac{1}{n}\right)^{\frac{3}{2}}\right),$$

et après simplification, $\qquad G = \Theta\left(\left[\frac{v}{2}\right]^n n^{\frac{j-i-1}{2}}\right)$, quand $n \to \infty$.

(ii) Lorsque $v \geq 1$ et $1/\sqrt{2} > 1/v$, et quand n tend vers l'infini,

$$\mathcal{N} = \Theta\left(\left[e^{h(\frac{1}{v}, 1)}\right]^n \left(\frac{1}{n}\right)^{\frac{j+3}{2}}\right), \text{ et } \quad G = \Theta\left(\left[\sqrt{1 - \frac{1}{v^2}}\right]^n n^{\frac{i+1}{2}}\right)$$

Là encore, si $1/\sqrt{2} = 1/v$, on se contente d'un résultat sur la limite de G, en encadrant correctement la fonction h.

c - $\beta = 0 \quad \alpha = 0$.

Dans ce cas, les constantes qui interviennent dans les équivalents ont une expression simple. C'est pourquoi à titre d'exemple le traitement de ce cas est détaillé davantage et les constantes sont données. En remplaçant α et β par 0 dans la relation (5.2), l'expression de la fonction h se simplifie: $\quad h(x, y) = \ln(xy)$.

Pour l'étude du dénominateur, le maximum de la fonction h sur $[0, 1]^2$ est atteint en $(1, 1)$ (un *coin* du domaine) et au voisinage de $(1, 1)$:

$$\begin{cases} h(x, y) &= (x-1) + (y-1) + o\left(\sqrt{(x-1)^2 + (y-1)^2}\right) \\ g(x, y) &\sim 2^{\frac{i+j}{2}-1} (1-x)^{\frac{i-1}{2}} (1-y)^{\frac{j-1}{2}} \end{cases}.$$

La Proposition 8 avec

$$(\lambda, \mu, \rho, \sigma) = ((i-1)/2, (j-1)/2, 1, 1),$$

donne alors la partie principale du dénominateur,

$$2^{\frac{i+j}{2}-1} \Gamma\left(\frac{i+1}{2}\right) \Gamma\left(\frac{j+1}{2}\right) \left(\frac{1}{n}\right)^{\frac{i+1}{2}} \left(\frac{1}{n}\right)^{\frac{j+1}{2}}. \tag{5.13}$$

Pour ce qui est du numérateur, pour $v < 1$, la fonction h atteint son maximum de nouveau en un coin de Δ, soit en $(1, v)$. Elle est de classe $\mathcal{C}^1$ en $(1, v)$ et au voisinage de ce point, elle vérifie

$$h(x, y) = \ln(v) + 2(x-1) + \frac{1}{v}(-vx + y) + o\left(\sqrt{(x-1)^2 + (y-v)^2}\right).$$

Par ailleurs, $\qquad g(x, y) \sim_{(1,v)} \dfrac{(2(1-x))^{\frac{i-1}{2}} (1-v^2)^{\frac{j-1}{2}}}{v^j}.$

Toutes les hypothèses de la Proposition 8 sont ici vérifiées. En particulier,

$$\left\{\begin{array}{lcl} X(x,y) & = & x \\ Y(x,y) & = & -vx+y \end{array}\right. \qquad \text{donc} \qquad \left|\frac{\partial(x,y)}{\partial(X,Y)}\right| = 1.$$

et

$$\left\{\begin{array}{ll} A=2 & B=1/v \\ \lambda=(i-1)/2 & \mu=0 \\ \rho=1 & \sigma=1 \end{array}\right.$$

Le numérateur de $G_{n,i}(v)$ a donc pour partie principale

$$\frac{2^{\frac{i-1}{2}}\,(1-v^2)^{\frac{j-1}{2}}}{v^j}\Gamma\left(\frac{i+1}{2}\right)\left(\frac{1}{2n}\right)^{\frac{i+1}{2}}\left(\frac{v}{n}\right)\,v^n.$$

Finalement, lorsque n est grand, pour tout $i \neq j$ fixés et pour tout $v<1$:

$$G_{n,(a,b)}(v) \sim \frac{\left(1-v^2\right)^{\frac{j-1}{2}}}{2^{\frac{i+j}{2}}\;v^{j-1}\;\Gamma\left(\frac{j+1}{2}\right)}\;v^n\,n^{\frac{j-1}{2}}.$$

• **Deuxième cas particulier:** $\beta \neq \mathbf{0}$ et $\mathbf{v} < \sqrt{\mathbf{1}-\beta}$

a - $\beta \neq \mathbf{0}$ et $\alpha = \mathbf{0}$.

Les mêmes raisonnements que les précédents permettent ici de conclure. (Pour le numérateur de G, le maximum est sur un bord et pour le dénominateur, le maximum est sur un coin) et les résultats obtenus sont similaires. Pour $\sqrt{1-\beta} > v$ et n suffisamment grand,

$$\Pr\{\frac{l_{\beta n+j}}{l_i} < v\} = \Theta\left(\left[\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}}\left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}\right]^n\frac{1}{\sqrt{n}}\right)$$

b - $\beta \neq \mathbf{0}$ et $\alpha = \mathbf{1}$.

Le maximum de la fonction h sur $[0,1]^2$ est ici atteint en $(0,\sqrt{1-\beta})$. Ce maximum est de nature elliptique et il se trouve sur un bord du domaine d'intégration. De plus, au voisinage de $(0,1)$ et à l'intérieur de $[0,1]^2$,

$$\left\{\begin{array}{lcl} h(x,y) & \sim & -AX^2-BY^2 \\ g(x,y) & \sim & C\,X^{-i} \quad \text{où pour avoir un sens, } i \text{ est négatif.} \end{array}\right.$$

Donc, en vertu de la Proposition 8 avec

$$(\lambda,\mu,\rho,\sigma) = (-i,0,2,2),$$

le dénominateur $\mathcal{D}$ vérifie

$$\mathcal{D} = \Theta\left(\left[\beta^{\frac{\beta}{2}}\,(1-\beta)^{\frac{1-\beta}{2}}\right]^n\left(\frac{1}{n}\right)^{\frac{-i+2}{2}}\right), \quad \text{quand } n\to\infty. \tag{5.14}$$

Par ailleurs, la fonction h atteint son maximum sur Δ en $(\xi, v\xi)$, avec $(\xi,v\xi)\in]0,1[^2$. Donc le maximum est sur un bord du domaine. La fonction g est continue et non nulle en $(\xi,v\xi)$. Cette fois, la Proposition 8 avec,

$$(\lambda,\mu,\rho,\sigma) = (0,0,1,2),$$

montre que le numérateur $\mathcal{N}$ vérifie

$$\mathcal{N} = \Theta\left(\left[e^{h(\xi,v\xi)}\right]^n \left(\frac{1}{n}\right)^{\frac{3}{2}} \right), \quad \text{quand } n \to \infty.$$

Finalement, quand n est assez grand, $$G = \Theta\left(\left[\frac{e^{h(\xi,v\xi)}}{\beta^{\frac{\beta}{2}}(1-\beta)^{\frac{1-\beta}{2}}} \right]^n n^{\frac{-i-1}{2}} \right).$$

En conclusion nous avons exhibé la partie principale de la fonction de répartition dans presque tous les cas où $v < \sqrt{(1-\beta)/(1-\alpha)}$. Pour des raisons de simplification d'écriture, la plupart des constantes qui interviennent dans ces équivalents ont été omises.

On en déduit en particulier que si $v < \dfrac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$, alors $\lim_{n\to\infty} G_{n,a,b}(v) = 0.$

5.3.2 Cas où $\mathbf{v} = \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$

Les maxima de la fonction h sur $\Delta(v)$ et sur $[0,1]^2$ sont atteints au même point mais le voisinage de ce point est différent selon le domaine considéré.
Le maximum de h est atteint sur la frontière $y = vx$ de $\Delta(v)$. Ce maximum $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ est bien de nature elliptique pour la fonction h et lorsque $(\alpha, \beta) \in]0,1[^2$ la fonction g est continue et non nulle en le maximum. On en déduit l'existence d'un voisinage $\mathcal{V}$ de $(\sqrt{1-\alpha}, \sqrt{1-\beta})$ et de trois réels A, B et C, strictement positifs tels que pour tout $(x,y) \in \mathcal{V}$:

$$\begin{cases} h(x,y) &= -AX^2 - BY^2 + o(x^2+y^2) \\ g(x,y) &\sim C \end{cases}$$

où $Y = -vx + y$ représente la direction du bord et X est une autre forme linéaire (exhibée dans la preuve du Lemme 6).
De nouveau, la méthode de Laplace permet de conclure. Ici, pour conclure on a besoin d'un équivalent précis. Pour le numérateur $\mathcal{N}$, le maximum est sur un bord du domaine $\Delta(v)$, et pour le dénominateur $\mathcal{D}$, le maximum est à l'intérieur du domaine $\Delta(\infty)$. Dans les deux cas, la Proposition 8 avec

$$(\lambda, \mu, \rho, \sigma) = (0, 0, 2, 2),$$

donne les équivalents souhaités. Lorsque n tend vers l'infini,

$$\mathcal{N} \sim \frac{\pi C}{2\sqrt{AB}} \frac{1}{n} \quad \text{et} \quad \mathcal{D} \sim \frac{\pi C}{\sqrt{AB}} \frac{1}{n}.$$

En conclusion, si $v = \dfrac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$ et $(\alpha,\beta) \in]0,1[^2$, alors $\lim_{n\to\infty} G_{n,\alpha,\beta} = \dfrac{1}{2}.$

Il existe ici, trois cas où pour le numérateur de G, le maximum de la fonction h sur Δ se trouve à l'intersection de deux bords de ce domaine, mais au voisinage du maximum, la fonction h ne se « *sépare* » pas selon les directions de ces deux bords. Il s'agit des trois cas suivants:

- $\alpha = \beta = 1$

- $\alpha = 0$ et $\sqrt{1-\beta} = v$ (avec $v \leq 1$)
- $\beta = 0$ et $\sqrt{1-\alpha} = 1/v$ (avec $v \geq 1$)

Il est toutefois possible de donner des équivalents pour ces cas particuliers en utilisant la Proposition 9. C'est ce qui est fait ci-dessous pour $\alpha = \beta = 1$.
La fonction h atteint son maximum absolu sur $[0,1]^2$ en $(0,0)$, qui appartient également à Δ et au voisinage de l'origine,

$$\begin{cases} h(x,y) & \sim & -\frac{1}{2}(x^2+y^2) \\ g(x,y) & \sim & x^{-i}y^{-j} \end{cases}.$$

Toutes les hypothèses de la Proposition 9 sont satisfaites et lorsque n tend vers l'infini, le numérateur $\mathcal{N}$ et le dénominateur $\mathcal{D}$ de la fonction de répartition ont respectivement comme parties principales

$$\mathcal{D} \sim \frac{1}{4}\Gamma\left(\frac{-i-j+2}{2}\right)\left(\frac{2}{n}\right)^{\frac{-i-j+2}{2}} B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right),$$

$$\text{et} \quad \mathcal{N} \sim \frac{1}{4}\Gamma\left(\frac{-i-j+2}{2}\right)\left(\frac{2}{n}\right)^{\frac{-i-j+2}{2}} B\left(\frac{-i+1}{2}, \frac{-j+1}{2}, \frac{v^2}{1+v^2}\right).$$

Finalement, après simplification,

$$G(v) \sim \frac{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}, \frac{v^2}{1+v^2}\right)}{B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}.$$

5.3.3 Cas où $\mathbf{v} > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$

Les maxima de la fonction h sur Δ et sur $[0,1]^2$ sont atteints au même point et ce point a le même voisinage dans les deux domaines. Plus précisément, le maximum de la fonction h sur le domaine $[0,1]^2$ est atteint en un point de l'intérieur de Δ. Donc, Le numérateur et le dénominateur ont exactement le même équivalent, lorsque n tend vers l'infini (celui du dénominateur dans les calculs du premier cas) et en général, $G_{n,(a,b)}(v)$ tend vers 1 avec n. Pour avoir plus de précision, il suffit de remarquer que

$$\Pr\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} < v\} \;=\; \Pr\{\frac{l_{\alpha n+i}}{l_{\beta n+j}} > \frac{1}{v}\} \;=\; 1 \;-\; \Pr\{\frac{l_{\alpha n+i}}{l_{\beta n+j}} < \frac{1}{v}\},$$

avec $1/v < \sqrt{1-\alpha}/\sqrt{1-\beta}$. D'où un équivalent pour $G_{n,(a,b)}(v)$, lorsque n devient grand et $v > \sqrt{1-\beta}/\sqrt{1-\alpha}$:

$$G_{n,(a,b)}(v) \;\sim\; 1 \;-\; G_{n,(b,a)}(1/v).$$

$$\text{En particulier, lorsque } v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}, \qquad \lim_{n\to\infty} G_{n,\alpha,\beta} \;=\; 1.$$

Nous venons donc de mettre évidence un phénomène de seuil:

Pour tous réels α, β de $[0,1]^2$, et tout couple d'entiers i, j , on rappelle que les indices affines a et b sont déterminés par

$$a := \alpha n + i \quad \text{et} \quad b := \beta n + j.$$

Pour toute constante réelle positive v, lorsque n tend vers l'infini,

$$\begin{cases} \lim_{n\to\infty} G_{n,(a,b)}(v) & = 0 \quad \text{pour } v < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}} \\ \lim_{n\to\infty} G_{n,(\alpha,\beta)}(v) & = 1 \quad \text{pour } v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}} \end{cases}$$

Lorsque $v = \dfrac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$, $\lim_{n\to\infty} G_{n,(\alpha,\beta)}(v) = A_v$ et la constante A_v vérifie toujours $0 < A_v < 1$.

En particulier, lorsque $(\alpha,\beta) \in]0,1[^2$, $A_v = 1/2$.

5.4 Les preuves pour la densité

On cherche d'abord la partie principale de la densité, lorsque $v \leq 1$. Le cas où $v > 1$ s'en déduira aisément ensuite[2].

5.4.1 Partie principale de $g_{n,(a,b)}(v)$ avec $v \leq 1$

La relation (5.5) et l'étude asymptotique de la fonction de répartition permettent facilement de trouver la partie principale de la densité lorsque $v \leq 1$. D'une part, la densité a exactement le même dénominateur que la fonction de répartition comme le montrent les écritures (5.4) et (5.5) et les parties principales de ces dénominateurs sont données dans la Section précédente. D'autre part, pour le numérateur, c'est la méthode de Laplace en une dimension (Propositions 6 et 7) qui sera utilisée. On note ici

$$\begin{cases} h_1 & : \quad x \mapsto h(x, vx) \\ g_1 & : \quad x \mapsto xg(x, vx) \end{cases}, \quad \text{où } h \text{ et } g \text{ sont définies par les relations (5.1) et (5.2).}$$

Pour tous les cas où $\sqrt{1-\beta} > v\sqrt{1-\alpha}$, la localisation du maximum de h_1 sur l'intervalle $[0,1]$ est également explicitement donnée dans la section précédente, puisque ce maximum est le même que celui de la fonction h sur le domaine $\Delta(v)$.

2. Plutôt que de chercher les équivalents des numérateurs de la densité comme on le fait par la suite, on pourrait établir le rapport précis entre l'équivalent du numérateur de la densité et l'équivalent du numérateur de la fonction de répartition déjà calculée dans la Section précédente. On montrerait ainsi que l'équivalent de la densité se déduit toujours de celle de la fonction de répartition comme annoncé par les Propositions 11 et 13. Cette deuxième façon pour donner les équivalents de la densité reste technique, car le lien entre les équivalents de la densité et la fonction de répartition est dû à la forme très particulière des fonctions g et h définies par (5.1) et (5.2). Une preuve rigoureuse par cette deuxième façon exige aussi une étude détaillée des différents cas et n'apporte pas d'éclaircissement supplémentaire.

1. $(\alpha,\beta)\in]0,1[^2$

La partie principale du dénominateur est donnée en (5.9). Pour le numérateur, la fonction h_1 atteint son maximum sur $[0,1]^2$ en ξ avec $\xi\in]0,1[$ et la définition de ξ est donnée dans la section précédente en (5.10). De plus, la fonction h_1 est de classe $\mathcal{C}^2$ en ξ avec $h_1'(\xi)=0$ et $h_1''(\xi)<0$. Par ailleurs, la fonction g_1 est continue et non nulle en ξ. Dans ces conditions, la Proposition 7 donne la partie principale du numérateur. En résumé, lorsque n tend vers l'infini, le numérateur $\mathcal{N}$ et le dénominateur $\mathcal{D}$ de $g_{n,(a,b)}(v)$ vérifient

$$\mathcal{N}=\Theta\left(e^{nh(\xi,v\xi)}\,\frac{1}{\sqrt{n}}\right)\qquad\qquad \mathcal{D}=\Theta\left(e^{nh(\sqrt{1-\alpha},\sqrt{1-\beta})}\,\left(\frac{1}{n}\right)\right)$$

Finalement,
$$g_{n,(a,b)}(v)=\Theta\left(\left[\frac{e^{nh(\xi,v\xi)}}{e^{nh(\sqrt{1-\alpha},\sqrt{1-\beta})}}\right]^n\,\sqrt{n}\right)$$

2. $\beta=0$ **et** $\alpha\in]0,1[$

La partie principale du dénominateur $\mathcal{D}$ est donnée en (5.11). Pour le numérateur, la fonction h_1 atteint son maximum sur $[0,1]^2$ en $\sqrt{1-\alpha/2}$, avec encore $h_1'(\sqrt{1-\alpha/2})=0$, $h_1''(\sqrt{1-\alpha/2})<0$ et la fonction g_1 continue et non nulle en $\sqrt{1-\alpha/2}$. Dans ces conditions, la Proposition 7 donne la partie principale du numérateur $\mathcal{N}$. Pour n assez grand,

$$\mathcal{N}=\Theta\left(e^{nh(\sqrt{1-\frac{\alpha}{2}},v\sqrt{1-\frac{\alpha}{2}})}\,\frac{1}{\sqrt{n}}\right)\qquad\qquad \mathcal{D}=\Theta\left(e^{nh(\sqrt{1-\alpha},1)}\,\left(\frac{1}{n}\right)^{\frac{j+2}{2}}\right),$$

et après simplification,

$$g_{n,(a,b)}(v)=\Theta\left(\left[\frac{v}{\sqrt{2}}\,\frac{(1-\frac{\alpha}{2})^{(1-\frac{\alpha}{2})}}{\left(\frac{1-\alpha}{2}\right)^{\left(\frac{1-\alpha}{2}\right)}}\right]^n\,n^{\frac{j+1}{2}}\right).$$

3. $\beta=0$ **et** $\alpha=1$

La partie principale du dénominateur est donnée en (5.12). Pour le numérateur, il suffit de remplacer α par 1, dans l'équivalent trouvé du cas précédent. Donc, lorsque n tend vers l'infini,

$$\mathcal{N}=\Theta\left(e^{nh(\frac{1}{\sqrt{2}},\frac{v}{\sqrt{2}})}\,\frac{1}{\sqrt{n}}\right)\qquad \text{et}\qquad \mathcal{D}=\Theta\left(\left(\frac{1}{n}\right)^{\frac{-i+j+2}{2}}\right).$$

D'où
$$g_{n,(a,b)}(v)=_{n\to\infty}\Theta\left(\left[\frac{v}{2}\right]^n\,n^{\frac{-i+j+1}{2}}\right).$$

4. $\beta=0$ **et** $\alpha=0$

La partie principale du dénominateur est donnée en (5.13). Pour le numérateur, la fonction h_1 atteint ici son maximum en 1 et au voisinage de 1,

$$\left\{\begin{array}{rcl} h_1(x) & = & \ln(v)+2(x-1)+o\,(x-1)\\ g_1(x) & \sim & 2^{\frac{i-1}{2}}\,(1-v^2)^{\frac{j-1}{2}}\,v^{-j}\,(1-x)^{\frac{i-1}{2}}\end{array}\right. .$$

Toutes les hypothèses de la Proposition 6 sont remplies et lorsque n tend vers l'infini,

$$\mathcal{N} \sim \frac{2^{\frac{i-1}{2}}\,(1-v^2)^{\frac{j-1}{2}}}{v^j}\Gamma\left(\frac{i+1}{2}\right)\left(\frac{1}{2n}\right)^{\frac{i+1}{2}}\,v^n.$$

Pour finir après simplification,

$$g_{n,(a,b)}(v) \;\sim\; \frac{\left(1-v^2\right)^{\frac{j-1}{2}}}{2^{\frac{i+j}{2}}\;v^j\;\Gamma\left(\frac{j+1}{2}\right)}\;v^n\,n^{\frac{j+1}{2}}.$$

5. $\alpha = 0$ **et** $\beta \in]0,1[$

Comme lors de l'étude de G, le dénominateur $\mathcal{D}$, de $g_{n,(a,b)}(v)$ vérifie la relation suivante lorsque n tend vers l'infini.

$$\mathcal{D} = \Theta\left(e^{nh(1,\sqrt{1-\beta})}\,\left(\frac{1}{n}\right)^{\frac{i+1}{2}}\right).$$

Par ailleurs, $\quad h'_1(x) = \dfrac{2-\beta-2v^2x^2}{(1-v^2x^2)x}$

La variation de h_1 pour $x \in [0,1]$ dépend donc de la localisation de $\dfrac{1}{v}\sqrt{1-\beta/2}$ par rapport à 1.

Si $\sqrt{1-\beta/2} > v$, alors le maximum de h_1 est de nouveau en 1 et exactement comme dans le cas précédent, pour n assez grand,

$$\mathcal{N} = \Theta\left(e^{nh(1,v)}\,\left(\frac{1}{n}\right)^{\frac{i+1}{2}}\right), \quad \text{et} \quad g_{n,(a,b)}(v) = \Theta\left(\left[\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}}\,\left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}\right]^n\,\sqrt{n}\right).$$

Si $\sqrt{1-\beta/2} = v$, alors le maximum de h_1 est toujours en 1, mais cette fois, $h'_1(1) = 0$ et $h''_1(1) < 0$. On applique encore la Proposition 6 et lorsque n tend vers l'infini,

$$\mathcal{N} = \Theta\left(e^{nh(1,v)}\,\left(\frac{1}{n}\right)^{\frac{i+1}{4}}\right), \quad \text{et} \quad g_{n,(a,b)}(v) = \Theta\left(\left[\left(\frac{1-v^2}{\beta}\right)^{\frac{\beta}{2}}\,\left(\frac{v^2}{1-\beta}\right)^{\frac{1-\beta}{2}}\right]^n\,n^{\frac{i+3}{4}}\right).$$

Si $\sqrt{1-\beta/2} < v$, alors la fonction h_1 atteint son maximum en $(1/v)\sqrt{1-\beta/2}$. Comme dans les cas 1 et 2, en le maximum, la dérivée h'_1 s'annule et la dérivée seconde est strictement négative. Par ailleurs, la fonction g_1 est continue et non nulle en $(1/v)\sqrt{1-\beta/2}$. La Proposition 7 entraîne alors, pour n assez grand,

$$\mathcal{N} = \Theta\left(e^{nh(\frac{\sqrt{1-\frac{\beta}{2}}}{v},\sqrt{1-\frac{\beta}{2}})}\,\left(\frac{1}{\sqrt{n}}\right)\right) \quad , \text{et} \quad g_{n,(a,b)}(v) = \Theta\left(\left[\frac{1}{\sqrt{2}\,v}\,\frac{(1-\frac{\beta}{2})^{(1-\frac{\beta}{2})}}{\left(\frac{1-\beta}{2}\right)^{\left(\frac{1-\beta}{2}\right)}}\right]^n\,n^{\frac{i+1}{2}}\right).$$

6. $\alpha = 0$ **et** $\beta = 1$

La partie principale du dénominateur $\mathcal{D}$ se déduit de celle du dénominateur dans le cas 3, en échangeant les fonctions a et b:

$$\mathcal{D} = \Theta\left(\left(\frac{1}{n}\right)^{\frac{i-j+2}{2}}\right), \quad \text{quand} \quad n \to \infty.$$

Pour le numérateur, il suffit de remplacer dans les équivalents trouvés du cas précédent (cas 5) β par 1.
De nouveau, trois situations se présentent et lorsque n tend vers l'infini,

(*i*) Si $1/\sqrt{2} > v$, alors $g_{n,(a,b)}(v) = \Theta\left([\sqrt{1-v^2}]^n \, n^{\frac{-j+1}{2}}\right)$.

(*ii*) Si $1/\sqrt{2} = v$, alors $g_{n,(a,b)}(v) = \Theta\left([\sqrt{1-v^2}]^n \, n^{\frac{i-2j+3}{4}}\right)$.

(*iii*) Si $1/\sqrt{2} < v$, alors $g_{n,(a,b)}(v) = \Theta\left(\left[\frac{1}{2v}\right]^n \, n^{\frac{i-j+1}{2}}\right)$.

7. $\alpha = 1$ **et** $\beta \in]0,1[$

La partie principale du dénominateur $\mathcal{D}$, de $g_{n,(a,b)}(v)$ est donnée en (5.14). Pour ce qui est du numérateur, la fonction h_1 atteint son maximum en un point $\xi \in]0,1[$, avec $h_1'(\xi) = 0$ et $h_1''(\xi) < 0$. Par ailleurs, la fonction g_1 est continue et non nulle en ξ. La Proposition 7 donne alors la partie principale du numérateur $\mathcal{N}$ de $g_{n,(a,b)}(v)$. Finalement, quand n tend vers l'infini,

$$\mathcal{N} = \Theta\left(e^{nh(\xi,v\xi)} \frac{1}{\sqrt{n}}\right), \quad \text{et} \quad g_{n,(a,b)}(v) = \Theta\left(\left[\frac{e^{h(\xi,v\xi)}}{\beta^{\frac{\beta}{2}} (1-\beta)^{\frac{1-\beta}{2}}}\right]^n n^{\frac{-i+1}{2}}\right).$$

8. $\alpha \in]0,1[$ **et** $\beta = 1$

En échangeant les fonctions a et b dans la relation (5.14), on a un équivalent pour le dénominateur $\mathcal{D}$ de G:

$$\mathcal{D} = \Theta\left(\left[\alpha^{\frac{\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}}\right]^n \left(\frac{1}{n}\right)^{\frac{-j+2}{2}}\right).$$

La fonction h_1 atteint son maximum en un point $\xi' \in]0,1[$. Tout se passe comme dans le cas précédent et lorsque n tend vers l'infini, le numérateur $\mathcal{N}$ de $g_{n,(a,b)}(v)$ vérifie

$$\mathcal{N} = \Theta\left(e^{nh(\xi',v\xi')} \frac{1}{\sqrt{n}}\right), \quad \text{d'où} \quad g_{n,(a,b)}(v) = \Theta\left(\left[\frac{e^{h(\xi',v\xi')}}{\alpha^{\frac{\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}}}\right]^n n^{\frac{-j+1}{2}}\right).$$

9. $\alpha = 1$ **et** $\beta = 1$

Pour le dénominateur, le maximum de la fonction h sur $[0,1]^2$ se situe en $(0,0)$, soit sur un coin du domaine et il est de nature elliptique. De plus, au voisinage de l'origine et à l'intérieur du domaine Δ,

$$\left\{ \begin{array}{lcl} h(x,y) & \sim & -\frac{1}{2}(x^2+y^2) \\ g(x,y) & \sim & x^{-i} y^{-j} \end{array} \right. .$$

Alors, grâce à la Proposition 8 avec $(\lambda, \mu, \rho, \sigma) = (-i, -j, 2, 2)$, lorsque n est grand, le dénominateur $\mathcal{D}$, de $g_{n,(a,b)}(v)$ vérifie

$$\mathcal{D} \sim \frac{1}{4}\,\Gamma\left(\frac{-i+1}{2}\right)\,\Gamma\left(\frac{-j+1}{2}\right)\,\left(\frac{2}{n}\right)^{\frac{-i-j+2}{2}}.$$

Pour le numérateur, le maximum de la fonction h_1 est au point 0. De plus, au voisinage de 0 et pour $x > 0$,

$$\left\{\begin{array}{lll} h_1(x) & \sim & -\frac{1+v^2}{2}(x^2) \\ g_1(x) & \sim & v^{-j}x^{-i-j+1} \end{array}\right. .$$

La Proposition 6 donne alors la partie principale du numérateur $\mathcal{N}$ de $g_{n,(a,b)}(v)$:

$$\mathcal{N} \sim v^{-j}\left(\frac{1}{1+v^2}\right)^{\frac{-i-j+2}{2}}\,\Gamma\left(\frac{-i-j+2}{2}\right)\,\left(\frac{2}{n}\right)^{\frac{-i-j+2}{2}}.$$

Pour finir,
$$g_{n,(a,b)}(v) \sim_{n\to\infty} \frac{4\,v^{-j}}{(1+v^2)^{\frac{-i-j+2}{2}} B\left(\frac{-i+1}{2}, \frac{-j+1}{2}\right)}$$

5.4.2 Partie principale de $g_{n,(a,b)}(v)$ avec $\mathbf{v} > 1$

De la remarque

$$\Pr\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} < v\} \;=\; \Pr\{\frac{l_{\alpha n+i}}{l_{\beta n+j}} > \frac{1}{v}\} \;=\; 1 \;-\; \Pr\{\frac{l_{\alpha n+i}}{l_{\beta n+j}} < \frac{1}{v}\},$$

on déduit successivement:

$$\frac{\partial}{\partial v}\,\Pr\left\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} < v\right\} \;=\; \frac{1}{v^2}\,\frac{\partial}{\partial(\frac{1}{v})}\,\Pr\left\{\frac{l_{\alpha n+i}}{l_{\beta n+j}} < \frac{1}{v}\right\}$$

et
$$g_{n,(a,b)}(v)\ \text{(avec } v > 1) \quad = \quad \frac{1}{v^2}\,g_{n,(b,a)}\left(\frac{1}{v}\right)\ \text{(avec } \tfrac{1}{v} < 1)\ .$$

Nous pouvons préciser le phénomène de seuil du paragraphe précédent:

Pour tous α, β, i et j et v fixés, on rappelle que les deux indices affines a et b sont définis par

$$a := \alpha n + i \quad \text{et} \quad b := \beta n + j.$$

Nous avons montré en particulier que

$$\left\{\begin{array}{ll} \lim\limits_{n\to\infty} g_{n,(a,b)}(v) & = 0 \quad \text{pour } v < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}} \\ \lim\limits_{n\to\infty} g_{n,(a,b)}(v) & = 0 \quad \text{pour } v > \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}} \end{array}\right.$$

Lorsque $v = \dfrac{\sqrt{1-\beta}}{\sqrt{1-\alpha}}$, $\quad g_{n,(\alpha,\beta)}(v) \sim_{n\to\infty} C(v,i,j)\,n^{B(i,j)}$, avec: $0 < C(v,i,j) < 1$.
En particulier, quand $\alpha = \beta = 1$, $\quad B(i,j) = 0$.

Chapitre 6

Analyse en moyenne des algorithmes de réduction

Dans ce chapitre, nous allons appliquer les études asymptotiques précédentes, pour classer les différents algorithmes de réduction, selon leurs comportements sur les réseaux aléatoires. Nous parlerons de *réduction lente* et de *réduction rapide*. En particulier, nous montrons que certains des algorithmes que nous avons introduits sont en moyenne plus rapides que l'algorithme LLL original, tout en fournissant des bases de sortie de qualité similaire.
Dans toute la suite, par "quasi-sûrement", on entend que la probabilité tend vers 1, de façon exponentielle avec la dimension n de l'espace ambiant. La hiérarchisation des algorithmes de réduction selon leurs complexités moyennes se fera en trois étapes.

Tout d'abord, on se préoccupe des conditions de réduction pour un indice affine isolé, indépendamment des autres indices. Cette étude précise et complète les résultats issus de la première étude en moyenne du Chapitre 3 (Proposition 5). Pour les indices du début, nous précisons toutes les vitesses de convergence des probabilités asymptotiques de réduction. Pour les indices de la fin, nous exhibons les probabilités–limites de réduction (elles n'étaient pas connues jusqu'alors). Cette première étape confirme que seulement pour les indices de la fin, les différentes conditions de réduction ne sont éventuellement pas satisfaites. Pour les indices du début, presque sûrement, tous les algorithmes de réduction vérifient juste que les conditions de réduction sont remplies. (En dehors de cette vérification, ils rendent les bases commençantes propres). Autrement dit, quasi-sûrement, les phases non–triviales *des algorithmes de réduction* se produisent pour les indices de la fin et c'est en considérant ces indices qu'on donne un premier classement des différents processus de réduction.

Puis, on considère les indices, non plus isolés, mais groupés et on étudie les différentes catégories de bases aléatoires définies dans le Chapitre 3.

1. Les bases aléatoires triviales (la dimension est une constante devant celle de la dimension de l'espace ambiant).
Le Chapitre 3 montrait déjà le résultat suivant qu'on précise ici: pourvu que la dimension n de l'espace ambiant soit assez grand, tous les algorithmes de réduction sont extrêmement rapides puisqu'on montre que quasi-sûrement, ils vérifient juste que la condition de réduction (peu importe laquelle parmi celles de la Définition 1) est satisfaite pour tous les indices du réseau et ils rendent la base propre. La précision apportée ici concerne la vitesse de

convergence. Si n désigne la dimension de l'espace ambiant, on savait déjà que la vitesse de convergence était polynomiale (Corollaire 5). On montre ici qu'elle est exponentielle en n. Ajoutons que grâce aux études en moyenne et en distribution des rapports l_b/l_a des Chapitres 3 et 5, ce résultat était attendu: d'une part le Théorème 7 montre que les bases triviales sont en moyenne orthogonales et d'autre part, l'étude en distribution montre que quand a et b sont des indices du début, la moyenne de la variable aléatoire l_b/l_a est une excellente approximation de ce rapport, lorsque n tend vers l'infini.

2. Les bases aléatoires incomplètes, c'est à dire de cardinal $\omega n + i$ avec $\omega < 1$. Le résultat précédent reste vrai pour les nouvelles réductions introduites.

En revanche, pour les réductions aux sens de s-Siegel et t-Lovász, tout dépend du poids ω de la base. Plus exactement, pour un paramètre réel $0 < s < 1$, on montre qu'il existe un poids $\omega_0(s) < 1$ pour lequel toute base aléatoire incomplète de dimension $p := \omega n + i$ avec $\omega < \omega_0(s)$ soit quasi-sûrement réduite, après $p - 1$ itérations. Mais lorsque $\omega_0(s) < \omega < 1$, le problème reste ouvert.

3. Les bases aléatoires complètes (et en particulier pleines), c'est à dire de cardinal $n - i$. Le classement introduit par la comparaison des probabilités de réduction pour un indice isolé se confirme. En particulier, lorsque la dimension de l'espace ambiant devient élevée, toute base aléatoire (même pleine) sera réduite au sens de nouvelles réductions après seulement $n - 1$ itérations. Pour les réductions “classiques” (LLL), nous donnons plusieurs majorations des probabilités–limites de réduction des bases aléatoires complètes. Autrement dit, avec une probabilité importante, l' algorithme LLL exige un nombre d'itération strictement supérieur à $n - 1$.

6.1 Quelques majorants pour la fonction de répartition

Dans le chapitre précédent, nous avons étudié très précisément la fonction de répartition et la distribution asymptotiques des rapports l_b/l_a lorsque a et b sont deux indices affines différents. Cette étude précise est riche en information et sera amplement utilisée dans la suite.

Cependant, dans l'étude précédente le rapport l_b/l_a est toujours comparé à une constante v. Par ailleurs, les résultats obtenus ne sont valables que pour les indices affines et ce de manière "non-uniforme" . Or, dans ce chapitre, nous allons étudier les probabilités de réduction des réseaux aléatoires, pour les réductions précédemment introduites où souvent les rapports l_b/l_a sont comparés avec une quantité qui dépend de n. Bien sûr, l'étude du chapitre précédent conduit en corollaire à ce type d'information. Mais les renseignements restent parfois trop peu précis pour les besoins de ce chapitre.

Les deux lemmes qui suivent donnent des majorations de la fonction de répartition, lorsque dans l'expression de la probabilité

$$\Pr\{l_b/l_a < v\}$$

la quantité v est une variable *totalement quelconque.* Elle peut en particulier dépendre exponentiellement de la dimension n de l'espace ambiant. Les résultats des lemmes ne sont alors que des majorations (le Chapitre 5 fournissait des équivalents).

Le Lemme 7 donne deux majorations grossières mais uniformes, donc valable pour toute paire d'indices (mêmes non affines) (a, b), tels que $1 \leq a < b \leq n$. Les majorations (6.1)

et (6.2) seront utilisées dans les Sections 6.3 et 6.4, lorsqu'on considère les probabilités de réduction d'une base aléatoire (globalement). Comme l'échelle affine ne constitue pas une partition de l'ensemble $\{1, \ldots, n\}$, on ne peut pas se contenter des résultats valables uniquement pour a et b des indices affines.
Dans le Lemme 8, a et b restent deux indices affines, mais v peut dépendre exponentiellement de n. Quand on considère la probabilité de réalisation d'une condition de réduction pour un indice affine isolé, excepté pour la réduction de s-Siegel, v dépend toujours de n de façon exponentielle. Les majorations de ce Lemme (6.6), (6.7) , (6.8) servent à évaluer les vitesses de convergence pour les probabilités de réalisation de toutes les conditions de réduction autres que celle de s-Siegel, pour un indice affine isolé. Les résultats seront alors moins précis que pour la condition de s-Siegel, puisque pour cette dernière on utilise, non pas des majorations, mais des équivalents issus de l'étude du Chapitre 5.

Lemme 7. *Soient p vecteurs $b_1, \ldots, b_p$ pris uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$ $(1 \leq p \leq n)$ et soit l_a la norme du a-ème vecteur orthogonalisé. Alors, pour toute paire d'indices (mêmes non affines) (a, b), tels que $1 \leq a < b \leq p$, et pour tout réel positif v, la fonction de répartition $G_{n,a,b}(v)$ du rapport l_b/l_a vérifie les majorations suivantes.*

$$G_{n,a,b}(v) \ < \ 2 \, \frac{n^{\frac{n}{2}}}{(b-1)^{\frac{b-1}{2}} (n-b+1)^{\frac{n-b+1}{2}}} \, v^{n-b+1}. \tag{6.1}$$

$$G_{n,a,b}(v) \ < \ 2 \, \frac{n^n \, (2n-a-b)^{\frac{2n-a-b}{2}}}{(b-1)^{\frac{b-1}{2}} \, (n-b+1)^{\frac{n-b+1}{2}} \, (2n-b+1)^{\frac{2n-b+1}{2}}} \, (n-a-1)^{\frac{n-a-1}{2}} \, v^{n-b+1}. \tag{6.2}$$

Preuve. Comme a est différent de b, le Théorème 4 [DV94] montre que les variables l_a et l_b sont indépendantes et donne la fonction de répartition du rapport l_b/l_a,

$$\Pr\left\{\frac{l_b}{l_a} < v\right\} \ = \ \frac{\int_{x=0}^{1} \int_{y=0}^{vx} (1-x^2)^{\frac{a-1}{2}} x^{n-a} (1-y^2)^{\frac{b-1}{2}} y^{n-b} \, dx dy}{\int_{x=0}^{1} (1-x^2)^{\frac{a-1}{2}} x^{n-a} \, dx \ \int_{y=0}^{1} (1-y^2)^{\frac{b-1}{2}} y^{n-b} \, dy}. \tag{6.3}$$

Dans le numérateur, la quantité $(1-y^2)$ est toujours trivialement majorée par 1.
1. La fonction de répartition de l_b/l_a est donc majorée par une expression ne faisant intervenir que des intégrales Beta. Les intégrales Beta s'expriment ensuite à l'aide de la fonction Gamma (3.2) et après simplification,

$$\Pr\left\{\frac{l_b}{l_a} < v\right\} < \frac{v^{n-b+1} \Gamma\left(n - \frac{a+b}{2} + 1\right) \Gamma^2\left(\frac{n}{2} + 1\right)}{\Gamma\left(\frac{b+1}{2}\right) \Gamma\left(\frac{n-b+3}{2}\right) \Gamma\left(\frac{n-a+1}{2}\right) \Gamma\left(n - \frac{b-1}{2} + 1\right)}. \tag{6.4}$$

2. De plus, en notant préalablement que l'intégrale $B(p,q)$ (3.1) est une fonction décroissante séparément de p et de q, et en exprimant ensuite seulement les intégrales Beta à l'aide de la fonction Γ, on obtient la majoration suivante, qui est plus grossière mais plus simple à manipuler.

$$\Pr\left\{\frac{l_b}{l_a} < v\right\} \ < \ \frac{v^{n-b+1} \Gamma\left(\frac{n+2}{2}\right)}{\Gamma\left(\frac{b+1}{2}\right) \Gamma\left(\frac{n-b+2}{2}\right)} \tag{6.5}$$

Puis, grâce à l'approximation classique de Stirling sur la fonction Gamma, [WW27],

$$x^x e^{-x}\sqrt{2\pi x} \leq \Gamma(x+1) \leq x^x e^{-x}\sqrt{2\pi x} e^{1/12x},$$

la relation (6.4) entraîne

$$\Pr\left\{\frac{l_b}{l_a} < v\right\} \;<\; \frac{e^{\frac{1}{18n}}}{\sqrt{\pi}}\sqrt{\frac{n}{(b-1)(n-b+1)}}\sqrt{\frac{n}{(2n-b+1)}}\sqrt{\frac{2n-a-b}{n-a-1}}M(a,b,n)v^{n-b+1},$$

$$\text{avec} \quad M(a,b,n) := \frac{n^n(2n-a-b)^{\frac{2n-a-b}{2}}}{(b-1)^{\frac{b-1}{2}}(n-b+1)^{\frac{n-b+1}{2}}(n-a-1)^{\frac{n-a-1}{2}}(2n-b+1)^{\frac{2n-b+1}{2}}}.$$

De même en utilisant la formule de Stirling et la relation (6.5), on obtient

$$\Pr\left\{\frac{l_b}{l_a} < v\right\} \;<\; \frac{e^{1/6n}}{\sqrt{\pi}}\sqrt{\frac{n}{(b-1)(n-b+1)}}\;\frac{n^{\frac{n}{2}}}{(b-1)^{\frac{j-1}{2}}(n-b+1)^{\frac{n-b+1}{2}}}\,v^{n-b+1}.$$

Pour obtenir les majorations (6.1) et (6.2) annoncées par le Lemme, il suffit de remarquer que pour $1 \leq a < b \leq p \leq n$,

$$\begin{aligned}(b-1)(n-b+1) &\geq n-1,\\ 2n-b+1 &\geq n+1,\\ 2(n-a-1) &\geq 2n-a-b.\end{aligned}$$

$\square$

Lemme 8. *Soient n vecteurs $b_1,\ldots,b_n$ pris uniformément et indépendamment dans la boule unité $\mathbb{R}^n$ et soit l_i la norme du i-ème vecteur orthogonalisé. Alors, pour toute paire d'indices affines (a,b), tels que $0 \leq \alpha \leq \beta \leq 1$, et pour tout réel positif v, la fonction de répartition $G_{n,a,b}(v)$ du rapport l_b/l_a est majorée, par une des relations suivantes selon les valeurs des pentes de α,β.*
Si $a := \alpha n + i, b := \beta n + j$ sont des indices de début, i.e. $(\alpha,\beta) \in [0,1[^2$, alors

$$G_{n,a,b}(v) = O\left(\left(\frac{(2-\alpha-\beta)^{\frac{2-\alpha-\beta}{2}}}{\beta^{\frac{\beta}{2}}(1-\beta)^{\frac{1-\beta}{2}}(2-\beta)^{\frac{2-\beta}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}}\right)^n v^{(1-\beta)n}\frac{1}{\sqrt{n}}\right). \tag{6.6}$$

Si $a := \alpha n + i$ est un indice de début, i.e. $\alpha < 1$, et $b := n - j$ est un indice de la fin, i.e. $\beta = 1$, alors

$$G_{n,a,b}(v) < \frac{v^{j+1}n^{\frac{j+1}{2}}}{(1-\alpha)^{\frac{j-1}{2}}2^{\frac{j-1}{2}}\Gamma\left(\frac{j-1}{2}\right)}. \tag{6.7}$$

Si $a := n - i, b := n - j$ sont deux indices de la fin, i.e. $(\alpha,\beta) = (1,1)$, alors

$$G_{n,a,b}(v) < \frac{v^{j+1}}{2B\left(\frac{i+1}{2},\frac{j+1}{2}\right)}. \tag{6.8}$$

Preuve. En utilisant encore la définition de la fonction de répartition comme rappelée en (6.3) et de nouveau en remarquant que dans la numérateur, la quantité $(1-y^2)$ est toujours trivialement majorée par 1, on déduit une majoration de la fonction de répartition de l_b/l_a,

$$G_{n,a,b}(v) < \frac{v^{n-b+1}}{n-b+1} \frac{\overbrace{\int_0^1 (1-x^2)^{\frac{a-1}{2}} x^{2n-a-b+1}\, dx}^{I_1}}{\underbrace{\int_0^1 (1-x^2)^{\frac{a-1}{2}} x^{n-a}\, dx}_{I_2} \underbrace{\int_0^1 (1-x^2)^{\frac{b-1}{2}} x^{n-b}\, dx}_{I_3}}. \tag{6.9}$$

La méthode de Laplace pour les intégrales simples (4) donne facilement un équivalent pour les intégrales de la forme I_1, I_2 ou I_3 (Les deux dernières sont les mêmes). En remplaçant dans I_2, a par son expression $a := \alpha n + i$, I_2 se met sous la forme

$$\int_0^1 g(x) e^{nh(x)}\, dx, \tag{6.10}$$

$$\text{avec} \quad \begin{array}{rcl} g(x) & = & (1-x^2)^{\frac{i-1}{2}}\, x^{-i} \\ h(x) & = & (\alpha/2)\ln(1-x^2) + (1-\alpha)\ln(x). \end{array}$$

Lorsque $\alpha \neq 0$ et $\alpha \neq 1$, le maximum de la fonction h est atteint en $\sqrt{1-\alpha}$ soit à l'intérieur de l'intervalle d'intégration et on a respectivement,

$$h(\sqrt{1-\alpha}) = \ln\left(\alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}\right), \quad h''(\sqrt{1-\alpha}) = -2/\alpha, \quad g(\sqrt{1-\alpha}) = \alpha^{\frac{i-1}{2}}(1-\alpha)^{-\frac{i}{2}}.$$

La Proposition 7 donne alors la partie principale de I_2, lorsque n tend vers l'infini,

$$I_2 \sim \sqrt{\pi}\left(\frac{\alpha}{1-\alpha}\right)^{\frac{i}{2}} \frac{1}{\sqrt{n}} \left(\alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}\right)^n. \tag{6.11}$$

Si maintenant $\alpha = 0$ ou $\alpha = 1$, le maximum de la fonction h est atteint en 1 respectivement (en 0), soit sur le bord de l'intervalle d'intégration et c'est la Proposition 6 qui fournit un équivalent pour I_2,

$$I_2 \sim \frac{1}{2}\Gamma\left(\frac{|i|+1}{2}\right)\left(\frac{2}{n}\right)^{\frac{|i|+1}{2}}. \tag{6.12}$$

L'intégrale I_1 est également de la forme (6.10). En remplaçant a et b par leurs expressions, I_1 s'écrit comme en (6.10) avec

$$\begin{array}{rcl} g(x) & = & (1-x^2)^{\frac{i-1}{2}}\, x^{-i-j+1} \\ h(x) & = & (\alpha/2)\ln(1-x^2) + (2-\alpha-\beta)\ln(x). \end{array}$$

De nouveau si α, β sont différents de 1, le maximum de la fonction h est atteint en $\xi := \sqrt{1-\alpha/(2-\beta)}$, soit à l'intérieur du domaine d'intégration et on a respectivement

$$h(\xi) = \ln\left(\frac{\alpha^{\frac{\alpha}{2}}(2-\alpha-\beta)^{\frac{2-\alpha-\beta}{2}}}{(2-\beta)^{\frac{2-\beta}{2}}}\right), \quad h'(\xi) = 0, \quad h''(\xi) = -\frac{2(2-\beta)^2}{\alpha},$$

$$\text{et} \quad g(\xi) = \alpha^{\frac{i}{2}}(2-\beta)^{\frac{j-2}{2}}(2-\alpha-\beta)^{-\frac{i+j-1}{2}}.$$

La Proposition 7 donne alors la partie principale de I_1:

$$I_1 \sim \sqrt{\pi}\frac{\alpha^{\frac{i}{2}}(2-\beta)^{\frac{j-2}{2}}}{(2-\alpha-\beta)^{\frac{i+j-1}{2}}}\,\frac{1}{\sqrt{n}}\,\left(\frac{\alpha^{\frac{\alpha}{2}}(2-\alpha-\beta)^{\frac{2-\alpha-\beta}{2}}}{(2-\beta)^{\frac{2-\beta}{2}}}\right)^n. \tag{6.13}$$

Lorsque $(\alpha, \beta) \in]0,1[^2$, la première majoration (6.6) du Lemme découle alors des relations (6.13),(6.11) et (6.9).

Si $\alpha = 0$ et $\beta > 0$, le maximum de la fonction h est en 1, soit encore sur un bord de l'intervalle d'intégration. La Proposition 6 montre alors que

$$I_1 \sim \frac{2^{\frac{i-1}{2}}}{(2-\beta)^{\frac{i+1}{2}}}\Gamma\left(\frac{i+1}{2}\right)\left(\frac{1}{n}\right)^{\frac{i+1}{2}}. \tag{6.14}$$

La dernière relation (6.14), ainsi que les relations (6.11), (6.12) et (6.9) conduisent alors à la première majoration donnée par le Lemme (6.6) , lorsque $\alpha = 0$ et $\beta > 0$.

Si $\alpha < 1$ et $\beta = 1$, la fonction h est alors exactement la même que pour l'intégrale I_2 et toujours d'après la Proposition 7,

$$I_1 \sim \sqrt{\pi}\frac{\alpha^{\frac{i}{2}}}{(1-\alpha)^{\frac{i+j-1}{2}}}\,\frac{1}{\sqrt{n}}\,\left(\alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}\right)^n. \tag{6.15}$$

Les relations (6.15), (6.12), (6.9) entraînent alors la deuxième majoration du Lemme (6.7). Enfin si $(\alpha, \beta) = (1,1)$, le maximum de la fonction h est en 0, soit encore sur un bord de l'intervalle d'intégration. La Proposition 6 montre alors que,

$$I_1 \sim \frac{1}{2}\,\Gamma\left(\frac{i+j+2}{2}\right)\left(\frac{2}{n}\right)^{\frac{i+j+2}{2}}. \tag{6.16}$$

Les relations (6.16), (6.12), (6.9) entraînent alors la troisième majoration du Lemme, (6.8). $\square$

6.2 Probabilité d'une condition de réduction pour un indice affine isolé

Soit une base aléatoire pleine. Les Propositions suivantes utilisent l'étude en distribution du Chapitre précédent et en particulier, le Théorème 9 pour montrer que pour tout indice affine m du début, toute condition de réduction $\mathcal{C}_m$ est quasi-sûrement satisfaite, lorsque n est assez grand. Pour les vitesses de convergence (pour une condition distincte de celle de s-Siegel), on utilisera souvent les majorations de la Section précédente. Nous montrons ainsi que les difficultés éventuelles des algorithmes de réduction ne se produisent que pour les indices de la fin. Puis, en considérant ces indices, on sépare les algorithmes définis dans le Chapitre 2 en deux classes. Cette classification sera reprise et confirmée par les sections suivantes.

Définition 4.

1. *Les algorithmes de réduction rapide*: ce sont les algorithmes de réduction au sens de (s,γ)-Gram, au sens (s,k)-Schmidt et au sens de (s,k,γ)-Gram–Schmidt avec $k := \omega n + i$ un indice du début.
2. *Les algorithmes de réduction lente*: ce sont les algorithmes de réduction au sens de s-Siegel, de t-Lovász, de (s,k)-Schmidt et (s,k,γ)-Gram–Schmidt avec $k := n - j$ un indice de la fin.

Pour les conditions de réduction relatives aux algorithmes de réduction rapide, l'assertion (i) du Théorème 9 montre que ces conditions sont quasi-sûrement satisfaites, même pour un indice m de la fin $(m := n - i)$.

En revanche, pour les conditions de réduction relatives aux algorithmes de réduction lente, l'assertion (ii) du Théorème 9 montre qu'avec une probabilité non–négligeable, ces conditions ne sont pas satisfaites pour un indice $m := n - i$ de la fin.

Remarque. Comme pour la proposition 5 du chapitre 3, pour la condition de t-Lovász, on doit supposer en plus que le coefficient sub–diagonal $m_{n-i+1,n-i}$ de la matrice de passage exprimant la base aléatoire b dans la base orthogonalisée b^*, vérifie: $|m_{n-i+1,n-i}| < 1/2$. D'un point de vue algorithmique, ce n'est pas vraiment une hypothèse supplémentaire car l'algorithme de $\mathcal{C}$–réduction effectue toujours le nécessaire pour assurer $|m_{n-i+1,n-i}| < 1/2$ avant de faire le test de la condition $\mathcal{C}_{n-i}$.

6.2.1 Les indices du début

Nous allons d'abord considérer un indice du début et montrer que toutes les conditions de réduction sont quasi-sûrement satisfaites lorsque la dimension n de l'espace ambiant tend vers l'infini.

Proposition 14. *Soit b une base aléatoire pleine et $m := \theta n + i$ un indice affine du début. Lorsque $\mathcal{C}$ désigne une des cinq conditions de la Définition 1, alors la condition $\mathcal{C}_m$ est quasi-sûrement satisfaite quand la dimension n de l'espace ambiant est assez élevée.*
Les vitesses de convergence sont toujours exponentielles en n mais varient selon la condition considérée et selon la localisation de l'indice, c'est à dire la pente θ.

Remarque. Les vitesses de convergence sont illustrées par les Figures 6.1, 6.2 et 6.3.

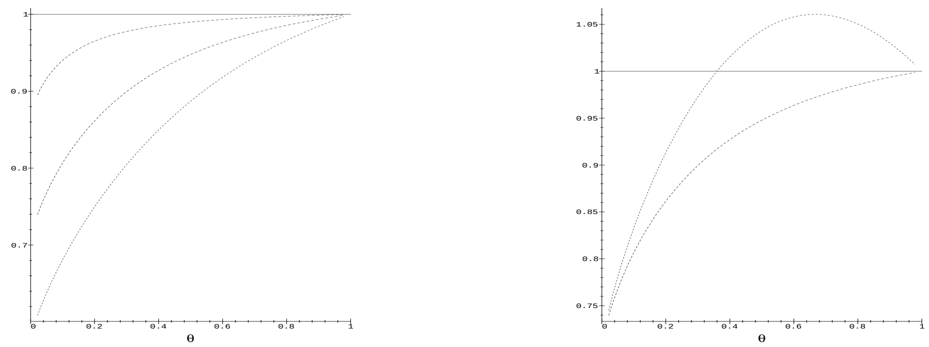

FIG. 6.1 – *La condition de s-Siegel pour un indice du début.*

Pour un indice m du début, la probabilité que la condition de s-Siegel soit satisfaite tend vers 1 en $\Theta_{\theta,i}\left((D_{\theta,s})^n/\sqrt{n}\right)$. La figure de gauche représente $D_{\theta,s}$ en fonction de θ pour trois valeurs de s^2 $(1/3, 1/2, 3/4)$. La figure de droite compare pour $s^2 = 1/2$, la valeur de D obtenue par la méthode de Laplace en dimension deux, à celle obtenue par majoration par le Lemme 8. Notons enfin, que les courbes tracées ne sont valables que pour $\theta < 1$.

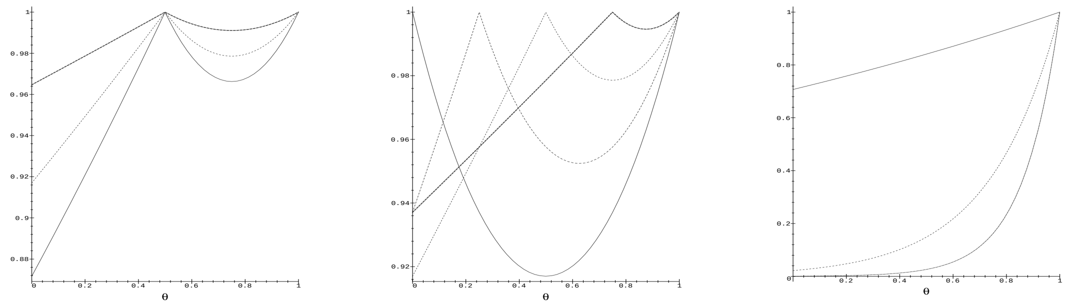

FIG. 6.2 – *La condition de (s,k)-Schmidt pour un indice du début.*

Pour un indice m de pente θ du début, la probabilité que la condition de (s,k)-Schmidt, (avec $k := \omega n + j$ également un indice du début), soit satisfaite tend vers 1 en $\mathrm{O}\left(s^{|\theta-\omega|(1-\max(\theta,\omega))n^2}\right)$. Les courbes de gauche et du milieu représentent $s^{|\theta-\omega|(1-\max(\theta,\omega))}$ en fonction de θ. À la différence des courbes de la Figure 6.1, ici les résultats sur les convergences sont obtenus par majoration, et non par équivalence. Dans la Figure de gauche, ω est fixé à $1/2$: les trois courbes correspondent aux trois valeurs $(1/3, 1/2, 3/4)$ pour s^2. Dans la Figure de droite, s est fixé à sa valeur habituelle $\sqrt{1/2}$ et les quatre courbes représentent différentes valeurs pour k (ω vaut $0, 1/4, 1/2, 3/4$). Les courbes ci-dessus sont valables pour $\theta \neq \omega$. Si $\theta = \omega$, la vitesse de convergence de la condition de (s,k)-Schmidt tend vers celle de la condition de s-Siegel d'indice k: l'ordre de grandeur passe de $s^{\mathrm{O}(n^2)}$ à $s^{\mathrm{O}(n)}$. Enfin, les courbes de droite illustrent les vitesses de convergence des conditions de $(s, n-i)$-Schmidt, pour les indices du début. La probabilité que cette condition soit satisfaite tend vers 1 en $\mathrm{O}\left(s^{(i+1)(1-\theta)n}\right)$. Dans les courbes de droite, s est toujours fixé à $1/\sqrt{2}$. Elles représentent $s^{(i+1)(1-\theta)}$, en fonction de $\theta < 1$, lorsque i est fixée respectivement aux valeurs $0, 10, 20$.

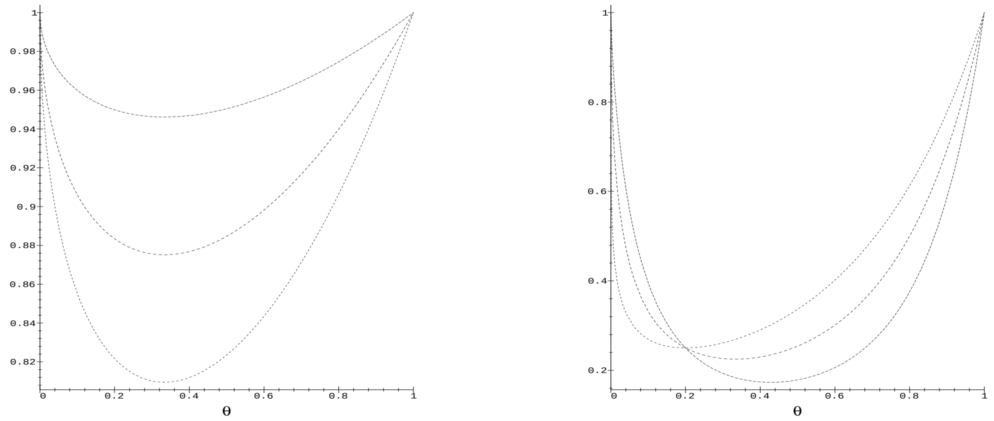

FIG. 6.3 – *La condition de (s,γ)-Gram pour un indice du début.*

Pour un indice m du début, la probabilité que la condition de (s,γ)-Gram soit satisfaite tend vers 1 en $\mathrm{O}\left(s^{\theta^\gamma(1-\theta)n^{1+\gamma}}\right)$. Ici encore, toutes les courbes des vitesses de convergence sont obtenues par majoration, et non par équivalence. Les courbes de gauche représentent $s^{\theta^\gamma(1-\theta)}$. Le paramètre γ est fixé à 1/2: les trois courbes correspondent aux trois valeurs $(1/3, 1/2, 3/4)$ pour s. Les courbes de la figure de droite représentent $s^{\theta^\gamma(1-\theta)n^{1+\gamma}}$, où s est fixé à $\sqrt{1/2}$, n est fixé à 5 et γ prend les valeurs $1/4, 1/2, 3/4$. Là encore, les courbes tracées ne sont valables que pour $\theta < 1$.

Preuve. On remarque d'abord que lorsque $\mathcal{C}$ désigne une des conditions de réduction précédemment définies et m un indice,

$$\Pr\{\mathcal{C}_m \text{ n'est pas satisfaite }\} \tag{6.17}$$

est toujours directement exprimée à l'aide de la fonction de répartition du rapport des normes de deux orthogonalisés. Dans tous les cas, les résultats du chapitre précédent sur les fonctions de répartition asymptotiques ou les deux Lemmes de la première Section permettent de majorer cette probabilité.

Dans la suite de cette preuve, $m := \theta n + i$ est un indice du début.

1. On suppose d'abord que $\mathcal{C}$ est la condition de s-Siegel.
Comme m est un indice du début, $m+1$ l'est également. L'assertion (i) du Théorème 9 montre alors que la fonction de répartition $G_{n,m+1,m}(s)$ du rapport l_{m+1}/l_m, converge (avec une vitesse exponentielle en n) vers une loi $0-1$ et ici le saut arrive pour $s = 1$, car $\alpha = \beta = \theta$. Ainsi, si $s < 1$, la fonction de répartition $G_{n,m+1,m}(s)$ converge vers 0, toujours avec une vitesse exponentielle en n. Les vitesses de convergence précises sont fournies par les Propositions récapitulatives du chapitre 5 et la Figure 6.1 illustre la variation de cette probabilité asymptotique lorsque s est fixé et lorsque $0 \le \theta < 1$ varie.

2. La condition $\mathcal{C}$ désigne maintenant celle de (s,k)-Schmidt avec $k := \omega n + j$ un indice du début, i.e. $0 \leq \theta < 1$:
On pose $\beta := \max(\theta, \omega)$, $\alpha := \min(\theta, \omega)$, et en changeant éventuellement les noms de i et de j, sans perte de généralité, on peut supposer que $j \geq i$.
Pour tout $s < 1$, et pour tout $\alpha < \beta < 1$, il existe n_0 et v_0 tels que,

$$\text{pour } n > n_0 \qquad s^{(\beta-\alpha)n+j-i} < v_0 < \frac{\sqrt{1-\beta}}{\sqrt{1-\alpha}},$$

et par la suite, pour $n > n_0$,

$$\Pr\left\{\frac{l_{\max(k,m)}}{l_{\min(k,m)}} < s^{|k-m|}\right\} < \Pr\left\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} < s^{(\beta-\alpha)n+j-i}\right\} < \Pr\left\{\frac{l_{\beta n+j}}{l_{\alpha n+i}} < v_0\right\}.$$

L'assertion (i) du Théorème 9 montre alors que la probabilité que la condition de (s,k)-Schmidt d'indice m n'est pas satisfaite, converge exponentiellement vers 0 avec n.

3. La condition $\mathcal{C}$ désigne celle de (s,k)-Schmidt avec $k := n - j$ un indice de la fin.
Cette fois, c'est le Lemme 8 qui permet de majorer la probabilité (6.17). En effet, en prenant $a := \omega n + i$, $b := n - j$ et $v := s^{b-a}$, la relation (6.7) entraîne la majoration souhaitée,

$$\Pr\{\mathcal{C}_m \text{ n'est pas satisfaite }\} = \mathrm{O}\left(s^{(j-1)(1-\omega)n}\right).$$

Une estimation des vitesses de convergence pour les conditions de (s,k)-Schmidt (toujours exponentielles en n) est fournie par le Lemme 8 et illustrée par les courbes de la Figure 6.2.

4. La condition de (s,k,γ)-Gram–Schmidt pour un indice de début m est quasi-sûrement satisfaite, car pour $\gamma < 1$, c'est une conséquence de la condition de s-Siegel (Lemme 1). Plus précisément, par définition, la condition de (s,k,γ)-Gram–Schmidt équivaut aux deux conditions simultanées de (s,γ)-Gram et de (s,k)-Schmidt. La partie principale de la vitesse de convergence (toujours exponentielle en n) provient de la condition de (s,γ)-Gram dont une estimation est là encore fournie par la relation (6.6) du Lemme 8 et illustrée par la Figure 6.3, pour différentes valeurs de γ et de s. □

6.2.2 Les algorithmes de réduction rapide et les indices de la fin

Nous allons montrer que les conditions correspondant aux algorithmes de réduction rapide sont quasi-sûrement satisfaites, même pour les indices de la fin, lorsque la dimension n de l'espace ambiant tend vers l'infini.

Proposition 15. *Soit b une base aléatoire de cardinal n et $m := n - i$ un indice affine de la fin. Si $\mathcal{C}$ désigne une condition de réduction relative à un algorithme de réduction rapide (condition de (s,γ)-Gram, de (s,k)-Schmidt ou de (s,k,γ)-Gram–Schmidt avec $k := \omega n + i$ un indice du début), alors $\mathcal{C}_{n-i}$ est quasi-sûrement satisfaite, lorsque n tend vers l'infini.*

La variation des vitesses de convergence est illustrée par la figure 6.4.

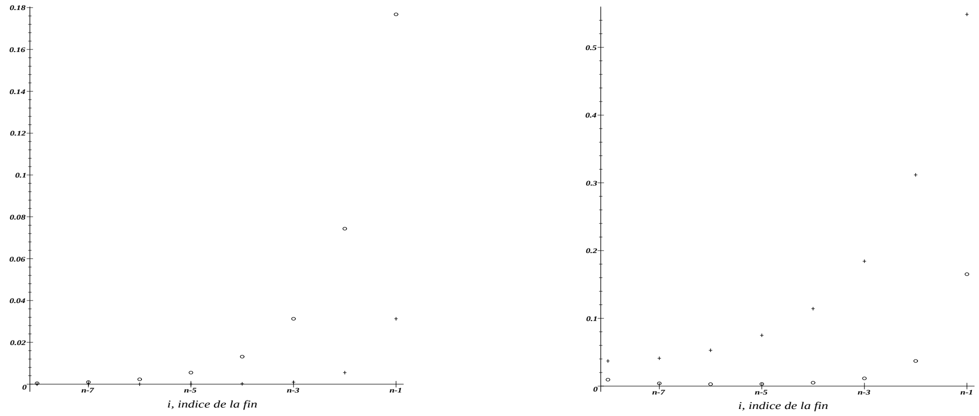

FIG. 6.4 – *Les conditions de (s,k)-Schmidt et (s,γ)-Gram pour les indices de la fin.*

Pour un indice $m := n-i$ de la fin, la probabilité que la condition de (s,k)-Schmidt (avec $k := \omega n+j$, un indice du début) est satisfaite tend vers 1 en $\mathrm{O}\left(s^{(i+1)(1-\omega)n}\right)$. Les courbes de gauche représentent $s^{(i+1)(1-\omega)n}$, lorsque n est fixé à 5 et ω est fixé aux valeurs respectives 0 et $1/2$.
Pour un indice $m := n-i$ de la fin, la probabilité que la condition de (s,γ)-Gram est satisfaite tend vers 1 en $\mathrm{O}\left(s^{i(n-i)^\gamma}\right)$. Les courbes de droite représentent $s^{i(n-i)^\gamma}$, n est fixé à 10 et γ prend les valeurs $1/4, 3/4$. Dans les deux cas, l'abscisse correspond au paramètre $n-i$ pour i variant de 8 à 0 et le pramètre s est fixé à sa valeur habituelle $1/\sqrt{2}$.

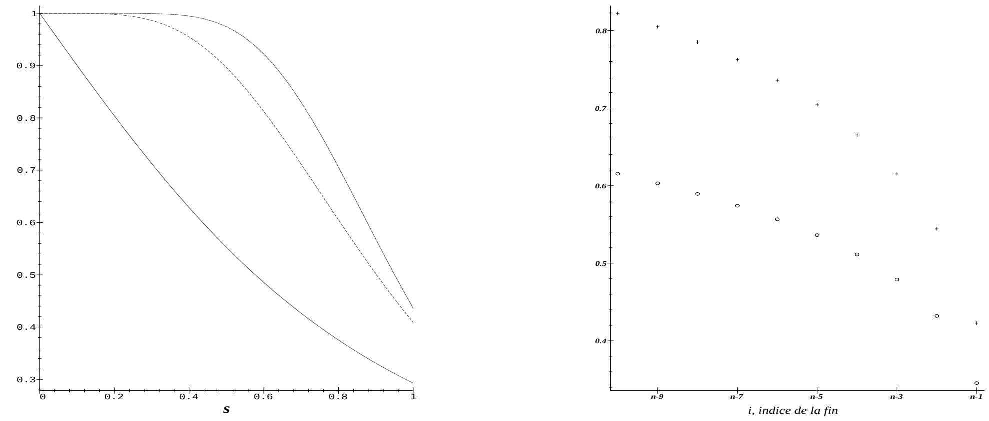

FIG. 6.5 – *Les probabilités–limites de la condition de s-Siegel pour les indices de la fin*

À gauche, les courbes représentent la probabilité–limite de la réalisation de la dernière (respectivement $(n-5)$-ème, $(n-10)$-ème) condition de s-Siegel en fonction de s: plus s esr grand, plus la réalisation est improbable. À droite, pour la valeur habiruelle $1/\sqrt{2}$ et pour $\sqrt{3}/\sqrt{4}$ pour s, les courbes représentent les probabilités–limites de réalisation des dix dernières conditions de s-Siegel: la réalisation est de plus en plus improbable quand l'indice s'approche de la fin.

Preuve de la Proposition 15.
1. La condition $\mathcal{C}$ désigne celle de (s,k)-Schmidt avec k un indice du début.
Comme k est un indice du début, c'est déjà montré par le Lemme précédent (point 4 de la preuve), car la probabilité que $\mathcal{C}_m$ est satisfaite est la même que la probabilité que la condition de (s,m)-Schmidt est satisfaite pour l'indice k. Les vitesses de convergence sont illustrées par la Figure 6.4 (courbes de droites).
2. La condition $\mathcal{C}$ désigne celle de (s,γ)-Gram. En choisissant $a := n-i$, $b := n-i-1$ et $v := s^{(n-i)^\gamma}$ dans la relation (6.8), on obtient

$$Pr\{\mathcal{C}_{n-i} \text{ n'est pas satisfaite } \} = \mathrm{O}\left(s^{i(n-i)^\gamma}\right)$$

Les variations de cette probabilité en fonction de l'indice $n-i$ et du paramètre $0 \le \gamma < 1$ sont illustrées par les Figures 6.4 (courbes de droites).
3. Si $\mathcal{C}$ désigne la condition de (s,k,γ)-Gram–Schmidt avec k un indice du début, C_{n-i} est quasi–sûrement satisfaite comme conséquence des points 1 et 2 et de nouveau pour $\gamma > 0$, la partie principale de la vitesse de convergence , exponentielle en n, provient de la probabilité que la condition de (s,γ)-Gram n'est pas satisfaite. □

6.2.3 Les algorithmes de réduction lente et les indices de la fin

Nous allons montrer que les conditions correspondant aux algorithmes de réduction lente ne sont pas satisfaites pour les indices de la fin, avec des probabilités–limites non–négligeables, lorsque la dimension n de l'espace ambiant tend vers l'infini.

Proposition 16. *Soit b une base aléatoire de cardinal n et $m := n-i$ un indice affine de la fin. Si $\mathcal{C}$ désigne une condition de réduction relative à un algorithme de réduction lente, alors*

$$\Pr\{\mathcal{C}_{n-i}(s) \textit{n'est pas satisfaite}\}$$

tend asymptotiquement vers une constante de l'intervalle $]0,1[$. *Plus précisément, si $\mathcal{C}$ désigne la condition de (s,k)-Schmidt ou de (s,k,γ)-Gram–Schmidt avec $k := n-j$ un indice de la fin, alors,*

$$\Pr\{\mathcal{C}_{n-i}(s) \textit{n'est pas satisfaite}\} \to \frac{B\left(\frac{\min(i,j)+1}{2}, \frac{\max(i,j)+1}{2}, \frac{s^{2|i-j|}}{1+s^{2|i-j|}}\right)}{B\left(\frac{j+1}{2}, \frac{i+1}{2}\right)}. \qquad (6.18)$$

Si $\mathcal{C}$ désigne la condition de s-Siegel ou de de t-Lovász, alors

$$\Pr\{\mathcal{C}_{n-i}(s) \textit{n'est pas satisfaite}\} \to \frac{B\left(\frac{i}{2}, \frac{i+1}{2}, \frac{s^2}{1+s^2}\right)}{B\left(\frac{i}{2}, \frac{i+1}{2}\right)}. \qquad (6.19)$$

Les probabilités limites des conditions de la fin au sens de s-Siegel sont illustrées par la Figure 6.5.

Preuve. C'est est un corollaire de l'assertion (ii) du Théorème 9.

Pour montrer (6.18), il suffit de remarquer que par définition,

$$\Pr\{\mathcal{C}_{n-i}(s)\text{n'est pas satisfaite}\} = \Pr\{\frac{l_{n-\min(i,j)}}{l_{n-\max(i,j)}} < s^{|i-j|}\}.$$

Le résultat découle alors de l'assertion (ii) du Théorème 9, en notant bien que la quantité $s^{|i-j|}$ ne dépend pas de n.
Pour montrer (6.19), il suffit par exemple de choisir $j := i - 1$ dans (6.18). □

Les probabilités asymptotiques de réduction de la condition de s-Siegel pour un indice de la fin $n - i$, sont d'une part décroissantes avec s (plus on exige une bonne qualité pour la base de sortie, plus les probabilités asymptotiques de réduction sont faibles) et d'autre part croissantes avec i. (La probabilité de réduction pour un indice isolé est d'autant plus faible que l'indice s'approche de n, la dimension de l'espace ambiant). Les courbes de la Figure 6.5 illustrent la variation de ces probabilités–limites. À titre d'exemples, quelques valeurs numériques sont données par le Tableau 6.1.

Indice (de la fin)	$s^2 = 1$	$s^2 = 3/4$	$s^2 = 1/2$	$s^2 = 1/3$
$(n-1)$	0.293	0.345	0.423	0.500
$(n-2)$	0.354	0.432	0.544	0.649
$(n-3)$	0.381	0.479	0.615	0.734
$(n-10)$	0.436	0.615	0.822	0.937
$(n-20)$	0.455	0.702	0.922	0.989
$(n-50)$	0.472	0.827	0.991	0.999

TAB. 6.1 – *Probabilités–limite de réalisation de la condition de s-Siegel pour quelques indices de la fin*

6.2.4 Conclusion

Cette Section permet de retrouver tout d'abord la Proposition 5 ainsi que les Corollaires 3, 4 et 5 du Chapitre 3 en les précisant puisque nous montrons ici que toutes les vitesses de convergence sont exponentielles en la dimension de l'espace ambiant (les méthodes du Chapitre 3 prouvaient qu'elles étaient polynomiales).

L'étude menée dans cette Section montre surtout que les conditions de réduction correspondant aux algorithmes de réduction lente (les réductions classiques) ne sont pas satisfaites pour les indices de la fin (avec des probabilités non–négligeables). Nous avons montré que les algorithmes de réduction rapide arrivent à résoudre cette difficulté, puisque les conditions de réduction même pour les indices de la fin restent quasi-sûrement satisfaites. L'étude en moyenne du Chapitre 3 ne fournissait pas d'information sur les indices de la fin. On rappelle (Chapitre 2) que les bases de sortie des algorithmes de (s, γ)-Gram réduction et (s, k, γ)-Gram–Schmidt réduction (avec k un indice du début) gardent une qualité similaire aux bases de sortie de l'algorithme LLL original. Nous avons donné des majorations pour tous les défauts de longueur et pour le défaut d'orthogonalité de ces bases.

Après avoir étudié *localement*, les probabilités asymptotiques de réalisation des différentes conditions, nous allons dans les sections qui suivent considérer les probabilités de réalisation de ces conditions *de façon globale* sur des bases aléatoires de poids différents.

6.3 Réductions des bases aléatoires incomplètes

On rappelle que lorsque la base d'entrée est entière et la norme choisie pour exprimer sa taille est la norme euclidienne, alors le nombre maximum d'itérations des algorithmes de réductions est majoré par $p^2 \log M$ (Théorème 3).
Pour les bases d'entrée de taille inférieure à une constante M et de cardinal $p := \omega n$ avec $\omega < 1$, (c'est à dire des base incomplètes), la borne précédente est en $\mathrm{O}(n^2)$.
Nous avons montré dans la Chapitre 3 que la qualité des bases aléatoires se dégrade, à mesure que le rapport entre la dimension de la base et la dimension de l'espace ambiant, c'est à dire son poids, s'approche de 1.
Nous allons montrer dans cette section que le poids de la base (et donc la dimension de l'espace ambiant) intervient de manière beaucoup plus *essentielle* dans le comportement en moyenne des algorithmes de réduction. En effet, le Théorème suivant montre que pour un paramètre réel $s < 1$ fixé, il existe un poids $\omega_0(s)$, pour lequel les bases aléatoires incomplètes de poids inférieurs à $\omega_0(s)$ se réduisent en un nombre d'itérations en $\mathrm{O}(n)$. Mieux, pourvu que n soit assez grand, pour une base aléatoire incomplète de dimension $p < \omega_0(s)n$ (du modèle probabiliste choisi), ce nombre est asymptotiquement égal à $p - 1$.

De plus, nous exhibons en (6.20) une borne inférieure pour $\omega_0(s)$, valable pour toutes les réductions précédemment définies. Pour un paramètre $0 < s < 1$ fixé, le poids limite $\omega_0(s)$ est le plus grand ω vérifiant la relation (6.20).
Pour les algorithmes de réduction rapide, nous montrons dans la Section suivante que $\omega_0(s) = 1$. Autrement dit pour ceux-la, le résultat est valable pour toutes les bases incomplètes, et pour tout $s < 1$.

Théorème 10. *Pour tout réel positif s vérifiant*

$$s < (1 - \frac{\omega}{2})\,(\omega(2-\omega))^{\frac{\omega}{1-\omega}}\,, \tag{6.20}$$

l'algorithme de réduction de s-Siegel réduit une base aléatoire incomplète de dimension $p := \omega n + j < n$ en un nombre d'itérations K_S qui est quasi-sûrement égal à $p - 1$:

$$Pr\{K_S = p-1\} \geq 1 - 7\omega n \left(\frac{(2s)^{(1-\omega)}}{\omega^{\frac{\omega}{2}}\,(2-\omega)^{\frac{(2-\omega)}{2}}} \right)^n . \tag{6.21}$$

Preuve. Soit $p := \omega n$ avec $\omega < 1$ et b une base aléatoire de cardinal p.
La probabilité qu'il y ait durant une exécution de l'algorithme de réduction au moins un test négatif, est trivialement majorée par

$$\sum_{i=1}^{p} \Pr\left\{\frac{l_{i+1}}{l_i} < s\right\} := \sum_{i=1}^{p} G_{n,i,i+1}(s).$$

Or le lemme 7 donne une majoration de la fonction de répartition $G_{n,i,i+1}(s)$ du rapport l_{i+1}/l_i. En choisissant $a := i$ et $b := i+1$ dans (6.2), on obtient la majoration suivante valable pour tout $i \in \{1, \ldots, n-1\}$:

$$\Pr\left\{\frac{l_{i+1}}{l_i} < v\right\} < 7\frac{n^n}{(2i)^{\frac{i}{2}}\,(n-i/2)^{n-\frac{i}{2}}}\, s^{n-i}.$$

Puis la variation de la fonction $i \mapsto (2i)^{\frac{i}{2}}(n-i/2)^{n-i/2}\,(1/s)^{n-i}$ pour $i \in [0,p]$ montre que, pourvu que $s^2 < 1/(2\omega) - 1/4$, le minimum de la fonction est atteint en $i := p$ (on rappelle que $p \sim \omega n$). Ainsi, tous les termes de la somme sont majorés par le dernier terme et la somme entière vérifie

$$\sum_{i=1}^{p} \Pr\left\{\frac{l_{i+1}}{l_i} < s\right\} < 7\omega n \left(\frac{(2s)^{(1-\omega)}}{\omega^{\frac{\omega}{2}}\,(2-\omega)^{\frac{(2-\omega)}{2}}}\right)^n.$$

La base d'entrée $(b_1, \ldots, b_p)$ est réduite après $p-1$ itérations si et seulement si tous les tests sont positifs. Donc, si $s^2 < 1/(2\omega) - 1/4$, le nombre K_S d'itérations de l'algorithme de s-Siegel vérifie bien la relation (6.21). Si s vérifie la relation (6.20), alors s vérifie également $s^2 < 1/(2\omega) - 1/4$. La majoration précédente est alors légitime et de plus la borne précédente converge vers 0, avec une vitesse exponentielle en n. □

Comme les autres réductions sont plus faibles que la réduction de s-Siegel d'après le Lemme 1, on obtient le résultat suivant.

Corollaire 6. *Pour tout réel positif s vérifiant la relation* (6.20)*,*

$$s < (1 - \frac{\omega}{2})\,(\omega(2-\omega))^{\frac{\omega}{1-\omega}},$$

le nombre K d'itérations de tous les algorithmes de réduction pour réduire une base aléatoire incomplète de dimension $p := \omega n + j < n$ est quasi-sûrement $p - 1$.

On montre dans la prochaine Section, que pour les algorithmes de réduction rapide, le poids limite $\omega_0(s)$ est toujours égal à 1. Plus généralement, on montrera le résultat pour les bases complètes et comme conséquence, toute base incomplète de cardinal p est donc quasi-sûrement réduite, après $p-1$ itérations.
Je pense que pour tous les algorithmes de réductions, y compris les algorithmes de réduction lente et en particulier l'algorithme de réduction au sens de s-Siegel, le poids limite $\omega_0(s)$ est égal à 1. En effet, dans la preuve précédente, on commence par majorer la probabilité de non–satisfaisabilité de la condition de réduction pour un indice i, par une expression $P(n,i)$ (en utilisant le Lemme 7). Puis, on remarque que pour $i \leq p$, avec p suffisamment petit

devant la dimension n de l'espace ambiant, la suite des majorants $P(n,i)$ est croissante avec i. Enfin, on majore la probabilité de non–réduction du réseau par

$$pP(n,p).$$

Si dans l'expression précédente, on remplace $P(n,p)$ par l'équivalent de la la probabilité de non–satisfaisabilité de la condition de s-Siegel, pour le dernier indice $p-1$, (équivalent obtenu par la méthode de Laplace en dimension deux), alors le résultat est prouvé pour tout indice p du début, i.e. pour tout $s<1$, $\omega_0(s)=1$.

Pour conclure, il suffit donc de montrer que la suite des probabilités de non–satisfaisabilité de la condition de réduction pour un indice i est croissante avec i, au moins pour $n > n_0(p)$. (La preuve précédente le montre non pas pour la suite des probabilités, mais pour la suite des majorants $P(n,i)$ des probabilités, et cela uniquement pour $i < p_0$.)

Les courbes de la figure 6.1 semblent justifier cette dernière conjecture. En particulier, les courbes de droite de cette Figure montre bien l'apparition d'une "fausse–limite" $\omega_0(s)$, due à une majoration grossière. Notons enfin que que cette croissance n'est pas prouvée, malgré l'allure des équivalents, puisque la méthode de Laplace (et donc les courbes des équivalents) n'est valable que pour les indices affines.

Conjecture 1. *Pour tout $s<1$ et pour tout $\omega<1$, il existe $N_0(s,\omega)$, tel que pour $n > N_0(s)$, et pour tout $1 \le i < \omega n$, la suite des probabilités $G_{n,i,i+1}(s)$ est une suite croissante de l'indice i.*

Corollaire 7. *Sous la conjecture 1, le nombre d'itérations de l'algorithme de réduction au sens de Siegel (et donc aussi l'algorithme LLL), lorsqu'il opère sur une base aléatoire incomplète de cardinal p est quasi-sûrement $p-1$, quand la dimension n de l'espace ambiant tend vers l'infini.*

En conclusion, l'intérêt de cette section est moins de classer les algorithmes de réductions (comme le font les Sections 6.2 et 6.4) que de montrer, comme l'a fait le Chapitre 3 l'importance du paramètre poids de la base ou encore le rapport

$$\frac{\text{Nombre de vecteurs de la base}}{\text{dimension de l'espace ambiant}}$$

dans le comportement des algorithmes de réduction sur les réseaux aléatoires.

6.4 Réductions des bases aléatoires complètes

Dans cette Section, nous confirmons le classement des algorithmes de réductions, tel qu'il est fait dans la section 6.2, selon le comportement des algorithmes sur les réseaux aléatoires.

Pour les algorithmes de réduction rapide, la Proposition 14 montre que pour tout indice affine $i \in \{1,\dots,n\}$, la condition de réduction $\mathcal{C}_i$ est quasi-sûrement satisfaite. Les Propositions 17 et 19 et le Corollaire 8 précisent que les conditions de réduction sont quasi-sûrement satisfaites, *toutes ensembles*, pour les algorithmes de réductions aux sens respectifs de (s,k)-Schmidt, (s,γ)-Gram et (s,k,γ)-Gram–Schmidt, avec k un indice du début. Plus précisément la pente de l'indice k doit être inférieure à une pente limite fonction du paramètre d'approximation s, mais cette restriction peut être levée sous la conjoncture 1. Ainsi, durant l'exécution d'un algorithme de réduction rapide sur une base aléatoire complète (et en particulier pleine), tous les tests des conditions $\mathcal{C}_i$ sont quasi-sûrement positifs et le nombre d'itérations de l'algorithme de réduction est quasi-sûrement $n-1$: *Sur les bases aléatoires du modèle uniforme, quasi-sûrement, les algorithmes de réduction rapide vérifient juste que les conditions de réduction sont bien satisfaites et rendent la base propre.*

Pour les algorithmes de réduction lente, la situation est très différente: Si une base est réduite, alors la condition de réduction est satisfaite en particulier pour le dernier indice. Pour une base aléatoire complète de cardinal $n-i+1$, la relation (6.19) exhibe la probabilité asymptotique que la dernière condition de réduction relative à un algorithme de réduction lente n'est pas satisfaite et donne ainsi un minorant pour la probabilité asymptotique de non–réduction d'une base aléatoire complète. Ainsi, durant l'exécution d'un algorithme de réduction lente sur une base aléatoire complète, avec une probabilité non–négligeable, au moins une fois, le dernier test est négatif. Le nombre d'itérations est alors strictement supérieure à $n-i$ avec la même probabilité. Le théorème suivant résume la situation.

Théorème 11. *On suppose que les algorithmes de réductions opèrent sur des bases aléatoires pleines, i.e. de dimension n. Soit k un indice du tout début, i.e. k est une constante entière. On note par*
K_1, le nombre d'itérations de l'algorithme de s-Siegel, de t-Lovász réduction, de (s,k)-Schmidt, de (s,γ,k)-Gram–Schmidt réduction, avec k un indice de la fin,
K_2, le nombre d'itérations de l'algorithme de (s,γ)-Gram, ou de (s,γ,k)-Gram–Schmidt réduction, avec k un indice du tout début, i.e. k une constante entière,
K_3, le nombre d'itérations de l'algorithme de (s,k)-Schmidt réduction, avec k un indice du début.
Alors, pour tout $0<\varepsilon<1$, il existe $N(\varepsilon,k,\gamma)$ tel que pour tout $n>N$,

$$\Pr\{K_1 > n-1\} \geq 1/\sqrt{1+(1/s)^2} \tag{6.22}$$

$$\Pr\{K_2 = n-1\} \geq 1 - s^{(1-\varepsilon)n^\gamma} \tag{6.23}$$

$$\Pr\{K_3 = n-1\} \geq 1 - s^{(1-\varepsilon)n} \tag{6.24}$$

Remarque 10. Par souci de simplicité, nous avons choisi d'énoncer les définitions de K_2 et de K_3 avec k une constante. En fait, les propositions qui suivent montrent que k peut être plus généralement un indice du début, i.e. $k := \omega n + i$ avec une condition sur la pente ω (6.27) qui fait intervenir le paramètre s d'approximation. Cette condition qui est de même nature que celle sur le poids d'une base incomplète dans le Théorème 10 devrait pouvoir disparaître sous la Conjecture 1.

Schéma de la preuve. Le théorème précédent est fondé sur plusieurs propositions, correspondant aux différents algorithmes de réduction.

Les Propositions 17 et 19 montrent les minorations (6.24) et (6.23) respectivement pour les algorithmes de réduction au sens de (s,k)-Schmidt, avec k un indice du début, et au sens de (s,γ)-Gram. Le Corollaire 8 en déduit la minoration (6.23) également pour l'algorithme de (s,γ,k)-Gram–Schmidt réduction. Voici le schéma de la preuve des Propositions 17 et 19.

Tout d'abord, on rappelle que le nombre d'itérations de l'algorithme de $\mathcal{C}$-réduction est $n-1$ si et seulement si les tests des conditions $\mathcal{C}_i$ sont positifs pour tous les $i \in \{1, \ldots n-1\}$.
La probabilité qu'il existe un indice $i \in \{1, \ldots n-1\}$, pour lequel la condition de s-Schmidt (respectivement la condition de (s, γ)-Gram) n'est pas satisfaite est majorée par une somme et on doit montrer qu'elle converge vers 0 avec une vitesse exponentielle en n. Pour évaluer le comportement asymptotique de cette somme (qui diffère selon la réduction considérée), nous utilisons l'idée de la méthode de Laplace (Chapitre 4, Section 3). Les preuves se font en trois étapes.

1. Nous cherchons l'indice i, pour lequel le i-ème terme de la somme est maximale. Comme l'expression exacte du terme général est trop compliqué pour le permettre, on commence par majorer tous les termes de la somme grâce au Lemme 7 par des expressions plus simples à manipuler.
2. En étudiant la variation du nouveau terme général obtenu, on décompose la somme en plusieurs "morceaux". La restriction de l'indice de sommation i à un intervalle adéquat plus petit permet alors de montrer que chaque *sous-somme* tend vers 0, avec une vitesse de convergence exponentielle en n.
3. On regroupe les majorants obtenus pour chaque *sous-somme* et on montre que la somme globale tend vers 0, avec une vitesse de convergence exponentielle en n.

Enfin la Proposition 9 montre la minoration (6.22), pour le nombre d'itérations des algorithmes de réduction lente.

□

6.4.1 L'algorithme de $(s, 1)$-Schmidt réduction

Dans ce paragraphe, nous montrons la relation (6.24) du théorème précédent. Le prochain paragraphe donne une généralisation pour tous algorithmes de (s, k)-Schmidt réduction, avec $k := \omega n + r$ un indice de début tel que $\omega < \omega_1(s)$.

Proposition 17. *Le nombre K_3 d'itérations de l'algorithme de réduction au sens de $(s, 1)$-Schmidt, lorsqu'il opère sur une base aléatoire pleine de cardinal n et lorsque k est un indice du début est quasi-sûrement $n-1$, quand n tend vers l'infini. Plus précisément, K_3 vérifie la relation* (6.24).

Preuve. La probabilité qu'il existe un indice i pour lequel, la condition de de $(s, 1)$-Schmidt n'est pas satisfaite est trivialement majorée par la somme

$$\sum_{i=1}^{n} \Pr\left\{\frac{l_{i+1}}{l_1} < s^i\right\},$$

$$\text{et par suite,} \quad \Pr\{K_S = n-1\} \geq 1 - \sum_{i=1}^{n} \Pr\left\{\frac{l_{i+1}}{l_1} < s^i\right\}.$$

En utilisant la majoration (6.1) du Lemme 7, on obtient

$$\sum_{i=1}^{n} \Pr\left\{\frac{l_{i+1}}{l_1} < s^i\right\} \leq 2 \underbrace{\sum_{i=1}^{n-1} \frac{n^{\frac{n}{2}}}{i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}}\, s^{i(n-i)}}_{P_1}.$$

Il reste à montrer que la somme P_1 converge vers 0, avec une vitesse exponentielle. Nous allons majorer P_1 dans l'esprit de la méthode de Laplace. (Chapitre 4, Section 3)

Un majorant pour P_1.

Ici, on divise P_1 en trois parties. Plus précisément, pour $s < 1$ fixé, pour tout $0 < \alpha < 1/2$ et lorsque $\alpha n > 2$, P_1 est décomposé en en trois morceaux de la façon suivante. ($\lfloor x \rfloor$ est la partie entière de x):

$$P_1 := \underbrace{\sum_{i=2}^{\lfloor \alpha n \rfloor} (\ldots\ldots\ldots)}_{S_1} \quad + \underbrace{\sum_{i=\lfloor \alpha n \rfloor+1}^{\lfloor (1-\alpha)n \rfloor} (\ldots\ldots\ldots)}_{S_2} \quad + \underbrace{\sum_{i=\lfloor (1-\alpha)n \rfloor+1}^{n} (\ldots\ldots\ldots)}_{S_3}$$

Nous allons montrer que

$$S_1 \text{ et } S_3 \text{ sont en } \mathrm{O}\left(n \, \left(s\, \alpha^{-\frac{\alpha}{2}} (1-\alpha)^{-\frac{1-\alpha}{2}}\right)^{-n}\right), \tag{6.25}$$

$$\text{et} \qquad S_2 = \mathrm{O}\left(n \;\, 2^n \;\, s^{\alpha(1-\alpha)n^2}\right). \tag{6.26}$$

Or, la fonction $\left(\alpha \mapsto \alpha^{\frac{\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}}\right)$ est continue, positive et décroissante sur l'intervalle $]0, 1/2]$ et elle tend vers 1, lorsque $\alpha \to 0$. Donc, pour tout ε, $0 < \varepsilon < 1$, il existe α_0 tel que

$$\text{pour tout } \alpha, \quad 0 < \alpha < \alpha_0, \qquad \alpha^{\frac{\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}} > s^{\varepsilon}, \qquad \text{donc} \qquad s\alpha^{-\frac{\alpha}{2}} (1-\alpha)^{-\frac{1-\alpha}{2}} < s^{(1-\varepsilon)}.$$

Ainsi, pour tout $\varepsilon \in]0, 1[$, on peut choisir $\alpha_0(s)$ tel que $P_1 \;=\; \mathrm{O}\left(n\, s^{(1-\varepsilon)n}\right)$.

Il reste à montrer les relations (6.25) et (6.26).

(i) On considère d'abord la somme S_1. Comme $2 \leq i \leq \lfloor \alpha n \rfloor$, dans chaque terme de S_1, l'exposant $i(n-i)$ de s est minoré par $n-1$ et le dénominateur $i^{\frac{i}{2}} (n-i)^{\frac{n-i}{2}}$ est minoré par $(\alpha n)^{(\alpha n)/2} ((1-\alpha)n)^{(1-\alpha)n/2}$. Enfin, en remarquant que la somme contient moins de n termes la majoration (6.25) est obtenue pour S_1.

(ii) Pour la somme S_3, les mêmes considérations permettent de conclure.

(iii) On considère maintenant la somme S_2. Dans chaque terme, l'exposant de s est minoré par $\alpha n(1-\alpha)n$ et le dénominateur est toujours minoré par $2^{n/2}$ et là encore, la somme contient moins de n termes.

□

6.4.2 Les algorithmes de (s, k)-Schmidt réduction

On rappelle qu'une base de cardinal n est (s, k)-Schmidt réduite si elle est propre et si les normes des orthogonalisés vérifient (2.10):

$$\text{pour } i \geq k, \quad l_i / l_k \geq s^{i-k},$$

$$\text{pour } i \leq k, \quad l_k/l_i \geq s^{k-i}.$$

Dans la proposition suivante, la relation (6.27) introduit un nouveau seuil $\omega_1(s)$ (pour un s fixé, le plus grand ω qui vérifie (6.27). Pour un paramètre s fixé, $\omega_1(s)$ est inférieur à $\omega_0(s)$ introduit par le Théorème 10. Le nouveau seuil $\omega_1(s)$ pourrait être facilement remplacée par $\omega_0(s)$, avec une étude plus fine.

Proposition 18. *Soit un paramètre $s < 1$ et un indice de début $k := \omega n + r$, tels que*

$$s < (1-\omega)^{1/2}\, \omega^{\frac{\omega}{2(1-\omega)}}. \tag{6.27}$$

Le nombre K_3' d'itérations de l'algorithme de réduction au sens de (s,k)-Schmidt lorsqu'il opère sur une base aléatoire pleine de cardinal n, est quasi-sûrement $n-1$, quand n tend vers l'infini. Plus précisément, K_3' vérifie la relation suivante. Pout tout $\varepsilon > 0$, il existe $N_0(\varepsilon, k)$ tel que pour tout $n > N_0(\varepsilon, k)$,

$$\Pr\left\{K_3' = n-1\right\} > 1 - \left(\frac{s^{(1-\omega)}}{\omega^{\frac{\omega}{2}}\,(1-\omega)^{\frac{(1-\omega)}{2}}}\right)^{(1-\varepsilon)n}. \tag{6.28}$$

Preuve. Pour montrer que l'algorithme de réduction s'arrête quasi-sûrement au bout de $n-1$ itérations, il suffit donc de montrer que la somme suivante converge vers 0 lorsque n tend vers l'infini, avec une vitesse exponentielle.

$$\underbrace{\sum_{i=1}^{k-1} \Pr\left\{\frac{l_k}{l_i} < v^{k-i}\right\}}_{S'} + \underbrace{\sum_{i=k}^{n-1} \Pr\left\{\frac{l_{i+1}}{l_k} < v^{i+1-k}\right\}}_{S}. \tag{6.29}$$

La somme S' correspond à un majorant de la probabilité de non–réduction au sens de (s,k)-Schmidt, d'une base aléatoire incomplète de dimension k. La Section précédente (Théorème 10) montre que lorsque $k := \omega n + r$ avec $\omega < \omega_0(s)$ (6.20), la somme S' tend exponentiellement vers 0 avec n et $\omega_1(s) < \omega_0(s)$. Plus précisément, en utilisant la majoration plus grossière (6.1), au lieu de (6.2), comme dans la preuve du Théorème 10, on obtient

$$\sum_{i=1}^{k-1} \Pr\left\{\frac{l_k}{l_i} < v^{k-i}\right\} < 2\omega n \left(\frac{s^{(1-\omega)}}{\omega^{\frac{\omega}{2}}\,(1-\omega)^{\frac{(1-\omega)}{2}}}\right)^{n}.$$

Donc, S' tend exponentiellement vers 0 pourvu que $\omega < \omega_1(s)$ défini par la relation (6.27).

Dans l'étude de la deuxième somme de la partie droite de l'égalité (6.29), désignée par S, nous allons appliquer des raisonnements similaires que dans le paragraphe précédent.
En l'occurrence, on suppose que $s < 1$ et $k := \omega n + r$ avec $\omega < \omega_0$ sont fixés.
Supposons d'abord que $\omega < 1/2$. On choisit alors α et β dans les intervalles suivants

$$\omega < \alpha < \frac{1}{2} \quad \text{et} \quad \frac{1+\omega}{2} < \beta < 1,$$

et on sépare la somme S en trois parties comme suit: ($\lfloor x \rfloor$ est la partie entière de x)

$$s := \underbrace{\sum_{i=\lfloor \omega n + r\rfloor}^{\lfloor \alpha n\rfloor} (\ldots\ldots\ldots)}_{S_1} \quad + \underbrace{\sum_{i=\lfloor \alpha n\rfloor+1}^{\lfloor \beta n\rfloor} (\ldots\ldots\ldots)}_{S_2} \quad + \underbrace{\sum_{i=\lfloor \beta n+1\rfloor}^{n-1} (\ldots\ldots\ldots)}_{S_3}$$

Si $\omega \geq 1/2$, on partagerait S en deux parties seulement en choisissant uniquement β. On aurait donc à traiter juste les sommes S_2 et S_3 et tous les arguments qui suivent restent valables. (Ce cas n'a pas d'intérêt pratique. En effet, ω et s sont reliés par la relation (6.27) et si $\omega > 1/2$, alors s serait beaucoup trop petit pour que la réduction soit réellement intéressante. On rappelle q'une valeur habituellement choisie pour s est $1/\sqrt{2}$.)

Nous allons montrer respectivement que:

$$S_1 = \mathrm{O}\left(n \left(s^{1-\omega}\, \alpha^{-\frac{\alpha}{2}}(1-\alpha)^{-\frac{1-\alpha}{2}}\right)^n\right), \tag{6.30}$$

$$S_3 = \mathrm{O}\left(n \left(s^{1-\omega}\, \beta^{-\frac{\beta}{2}}(1-\beta)^{-\frac{1-\beta}{2}}\right)^n\right), \tag{6.31}$$

$$S_2 = \mathrm{O}\left(n\ 2^n\ s^{An^2}\right),\ \text{avec } A := \min{(\alpha-\omega)(1-\alpha), (\beta-\omega)(1-\beta)}. \tag{6.32}$$

Et comme la fonction $\left(\alpha \mapsto \alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}\right)$ est continue, positive, symétrique par rapport à $\alpha = 1/2$ et décroissante sur l'intervalle $]0, 1/2]$, pour tout ε, $0 < \varepsilon < 1$,

$$S = \mathrm{O}\left(\left(\frac{s^{(1-\omega)}}{\omega^{\frac{\omega}{2}}\ (1-\omega)^{\frac{(1-\omega)}{2}}}\right)^{(1-\varepsilon)n}\right).$$

Il reste à montrer les relations (6.30),(6.31) et (6.32).

(*i*) On considère d'abord la somme S_1. Comme $\omega n + r \leq i \leq \lfloor \alpha n \rfloor$, dans chaque terme de S_1, l'exposant $(i-k)(n-i)$ de s est minoré par $n - \omega n - r$ et le dénominateur $i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}$ est minoré par $(\alpha n)^{(\alpha n)/2}((1-\alpha)n)^{(1-\alpha)n/2}$. Enfin, en remarquant que la somme contient moins de n termes la majoration (6.30) est obtenue pour S_1.

(*ii*) Pour la somme S_3, les mêmes considérations permettent de conclure. En effet, comme $(1+\omega)/2 < \beta$ le maximum absolu de l'exposant de s, c'est à dire $(i-k)(n-i)$, sur l'intervalle $\{k, \ldots n-1\}$ est à l'extérieur de l'intervalle à sommer. L'exposant est toujours majoré par $n-1-k$, soit encore $n - \omega n - r + 1$. Cette fois, le dénominateur $i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}$ est minoré par $(\beta n)^{(\beta n)/2}((1-\beta)n)^{(1-\beta)n/2}$.

(*iii*) On considère enfin la somme S_2. Dans chaque terme, l'exposant de s est minoré par le plus petit des deux exposants $(\beta-\omega)n(1-\beta)n$ et $(\alpha-\omega)n(1-\alpha)n$. Le dénominateur est toujours minoré par $2^{n/2}$ et là encore, la somme contient moins de n termes.

□

6.4.3 L'algorithme de (s, γ)-Gram réduction

Dans ce paragraphe, nous montrons la relation (6.23) du théorème précédent.

Proposition 19. *Le nombre K_2 d'itérations de l'algorithme de réduction au sens de (s,γ)-Gram, lorsqu'il opère sur une base aléatoire pleine de cardinal n, est quasi-sûrement $n-1$, quand n tend vers l'infini. Plus précisément, pour tout $0<\varepsilon<1$, il existe $N(\varepsilon,\gamma)$ tel que pour tout $n>N$,*

$$Pr\{K_2 = n-1\} \geq 1 - s^{(1-\varepsilon)n^{\gamma}}.$$

Preuve. La probabilité qu'il existe un indice i pour lequel, la condition de de (s,γ)-Gram n'est pas satisfaite est trivialement majorée par

$$\sum_{i=1}^{n-1} \Pr\left\{\frac{l_{i+1}}{l_i} < s^{i^{\gamma}}\right\}.$$

$$\text{Donc,} \qquad \Pr\{K_S = n-1\} \geq 1 - P_1(n,n,s^i).$$

La relation (6.1) du Lemme 7 montre que

$$\sum_{i=1}^{n-1} \Pr\left\{\frac{l_{i+1}}{l_i} < s^{i^{\gamma}}\right\} \leq \underbrace{2\sum_{i=1}^{n-1} \frac{n^{\frac{n}{2}}}{i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}}\, s^{i^{\gamma}(n-i)}}_{P_2}. \tag{6.33}$$

Il nous reste à montrer que la somme P_2 converge vers 0, avec une vitesse exponentielle. Nous allons majorer la somme P_2 comme dans les preuves précédentes dans l'esprit de la méthode de Laplace. Mais cette fois, nous aurons besoin d'une subdivision plus fine de l'intervalle $[1,n]$.

Un majorant pour P_2.

Ici, nous allons diviser la somme P_2 en $k+4$ parties comme suit. Pour tout couple de réels fixés α et ϕ tels que $0<\alpha<1/2$ et $0<\phi<1$, on définit la suite croissante u_j, pour $j \in \{0,\ldots,k+4\}$, par les relations suivantes:

$$\begin{cases}
u_0 & := & 1; \\
u_1 & := & \alpha n; \\
u_2 & := & n - n^{(1+(k)\phi)\gamma}; \\
u_3 & := & n - n^{(1+(k-1)\phi)\gamma}; \\
\vdots & \vdots & \vdots \\
u_j & := & n - n^{\delta_j} \quad \text{pour } 2 \leq j \leq k+3 \quad \text{où } \delta_j := [1+(k-j+2)\phi]\,\gamma; \\
\vdots & \vdots & \vdots \\
u_{k+1} & := & n - n^{(1+\phi)\gamma}; \\
u_{k+2} & := & n - n^{\gamma}; \\
u_{k+3} & := & n - n^{(1-\phi)\gamma}; \\
u_{k+4} & := & n-1;
\end{cases}$$

La suite $u_0, u_2, \ldots, u_{k+4}$ forme une subdivision de l'intervalle $[1, n-1]$, pourvu que $(1+k\phi)\gamma < 1$. En choisissant l'entier k et le réel $0<\phi<1$ tels que

$$k\phi < \frac{1}{\gamma} - 1,$$

la subdivision est possible. On décompose alors P_2 de la façon suivante.

$$P_2 := \sum_{j=1}^{k+4} S_j \quad \text{où } S_j := \sum_{i \in [u_{j-1}, u_j]} \frac{n^{\frac{n}{2}}}{i^{\frac{i}{2}}(n-i)^{\frac{n-i}{2}}} \, s^{i^\gamma (n-i)}$$

(i) Pour la somme S_1, le majorant est obtenu exactement comme dans la preuve pour l'algorithme de $(s,1)$-Schmidt réduction (Il s'agissait alors de S_1 dans le majorant de P_1):

$$S_1 = \mathrm{O}\left(n \left(\frac{s}{\alpha^{\frac{\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}}\right)^n\right).$$

Là encore, on peut toujours choisir $\alpha < 1/2$ tel que la somme S_1 converge exponentiellement vers 0 avec n.

(ii) Pour chacune des sommes S_j, avec $j \in \{2 \dots k+4\}$, l'indice i varie dans l'intervalle $[u_{j-1}, u_j]$.

Le minimum de l'exposant $i^\gamma(n-i)$ de s est toujours atteint en $i = u_j$, car la fonction $(i \mapsto i^\gamma(n-i))$ est décroissante pour $i > \gamma n/(1+\gamma)$.

Par ailleurs, la fonction $\left(i \mapsto i^i(n-i)^{n-i}\right)$ est décroissante pour $i \in [0, n/2]$ et croissante pour $i \in [n/2, n]$. Donc, la quantité $(1/2)^n$ minore tous les dénominateurs et en particulier, ceux des termes de S_2. Pour chaque autre somme S_j, le minimum du dénominateur est atteint en $i = u_{j-1}$, ce qui fournit pour chaque somme autre que S_2, un minorant plus fin pour le dénominateur du terme général.

Ces remarques ainsi que des développements asymptotiques élémentaires permettent alors de montrer que

$$\begin{aligned}
S_2 &= \mathrm{O}\left(n\, 2^n\, s^{n^{(2+k\phi)\gamma}}\right) \\
S_j &= \mathrm{O}\left(n^{n^{\delta_{j-1}}}\; s^{n^{\mu_j}}\right) \quad \text{pour } j \in \{3, \dots, k+3\} \\
\text{où } \mu_j &= [2 + (k-j+2)\phi]\,\gamma \\
S_{k+4} &= \mathrm{O}\left(n^{n^{(1-\phi)\gamma}}\; s^{n^\gamma}\right)
\end{aligned}$$

Ensuite, on constate que pour $j \in \{3, \dots, k+3\}$,

$$\mu_j - \delta_{j-1} = 1 - \phi > 0 \quad \text{car } \phi < 1.$$

Donc, pour tout $0 < \varepsilon < 1$, il existe $N(\varepsilon, \gamma, s)$ tel que pour tout $j \in \{3, \dots, k+3\}$ et pour tout $n > N(\varepsilon, \gamma, s)$

$$S_j < s^{(1-\varepsilon)n^{(2+(k-j+2)\phi)\gamma}} < s^{(1-\varepsilon)n^\gamma}.$$

Il en est de même pour la somme S_{k+4}, car $1 - \phi < 1$.

Ainsi, les sommes $S_1, S_3, \dots, S_{k+4}$ convergent toutes vers 0, de façon exponentielle lorsque n tend vers l'infini.

Pour prouver que la somme P_2 toute entière converge vers 0 de façon exponentielle avec n, il suffit de montrer qu'il existe toujours un entier k et un réel ϕ tel que,

1. La somme S_2 converge vers 0 avec n, i.e. $(2 + k\phi)\gamma > 1$.

2. La subdivision précédente (en $k+4$ morceaux) de l'intervalle $[1, n-1]$ est effectivement possible, i.e. $(1 + k\phi)\gamma < 1$.

Ces deux conditions sont toutes deux remplies, pourvu que k et ϕ soient choisis tels que

$$\frac{1}{\gamma} - 2 < k\phi < \frac{1}{\gamma} - 1, \quad \text{ce qui est toujours possible.}$$

□

6.4.4 L'algorithme de (s, k, γ)-Gram–Schmidt réduction, avec k un indice du début

On rappelle qu'une base est (s, k, γ)-Gram–Schmidt réduite si elle est réduite au sens de (s, γ)-Gram et au sens de (s, k)-Schmidt. On déduit donc des propositions précédentes,

Corollaire 8. *Soit un paramètre $s < 1$ et $k := \omega n + r$ un indice de début tel que $\omega < \omega_1(s)$, où $\omega_1(s)$ est défini par la relation* (6.27)*. Le nombre K_3'' d'itérations de l'algorithme de réduction au sens de (s, k, γ)-Schmidt lorsqu'il opère sur une base aléatoire pleine de cardinal n, est quasi-sûrement $n-1$, quand n tend vers l'infini. Plus précisément, K_3'' vérifie la relation* (6.23)*.*

6.4.5 L'algorithme de s-Siegel, t-Lovász, (s, k)-Schmidt ou (s, k)-Gram–Schmidt réduction, avec k un indice de la fin

Les résultats du Lemme 16 permettent de majorer la probabilité de réduction d'une base aléatoire pleine au sens des réductions lentes. On rappelle que si une base est réduite au sens de t-Lovász, alors elle est réduite au sens de $s-$Siegel, avec

$$\frac{1}{s^2} = \frac{1}{t^2} - \frac{1}{4}.$$

De même, si une base est réduite au sens de (s, k)-Schmidt avec $k := n - j$ un indice de la fin, alors elle vérifie la condition de s-Siegel d'indice $k - 1$
et si une base est réduite au sens de $s-$Siegel, alors la condition de s-Siegel est satisfaite en particulier pour le dernier indice. On en déduit le corollaire suivant ou la deuxième majoration est bien sûr moins fine.

Corollaire 9. *Soit b une base de cardinal n dont les vecteurs sont choisis de façon aléatoire, uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Alors, lorsque n est assez grand,*

$$\textit{pour tout } s \geq 1, \quad \Pr\{ \textit{ la base } b \textit{ est } s\textit{−Siegel réduite}\} < 1 - \frac{1}{\sqrt{1 + (1/s)^2}},$$

$$\textit{et pour tout } 1 \leq t \leq 2, \quad \Pr\{\textit{la base } b \textit{ est } t\textit{−Lovász réduite}\} < 1 - \sqrt{\frac{4 - t^2}{3 + t^2}}$$

Par ailleurs, les majorations données précédemment ne sont clairement pas optimales. Les variables aléatoires l_i étant indépendantes [DV94], il en est de même, pour des rapports l_i/l_j et $l_{i'}/l_{j'}$ lorsque $\{i, j\}$ et $\{i', j'\}$ ont une intersection vide. Autrement dit, lorsque $\mathcal{C}$

désigne la condition de s-Siegel, les événements "$\mathcal{C}_{n-i}$ est satisfaite", pour des j différents et de mêmes parités sont indépendants. Ainsi pour tout entier positif r, si $\Pr_\infty$ désigne la probabilité asymptotique quand n tend vers l'infini,

$$\Pr_\infty\{\text{ la base } (b) \text{ est } s\text{−Siegel réduite}\} < \prod_{\substack{j\geq 0,\\ j \text{ pair ou impair}}}^{r} \left(1 - \frac{B\left(\frac{j+1}{2}, \frac{j+2}{2}, \frac{1}{1+s^2}\right)}{B\left(\frac{j+1}{2}, \frac{j+2}{2}\right)}\right) \tag{6.34}$$

L'inégalité précédente permet de majorer numériquement la probabilité de s-Siegel–réduction des bases aléatoires, de plus en plus finement. Par exemple, en choisissant $r := 100$ et en considérant les j pairs, le Tableau 6.4.5 fournit quelques valeurs numériques qui majorent les probabilités asymptotiques de réductions des bases aléatoires, pour quelques différentes valeurs de s.

	$s^2 = 1$	$s^2 = 3/4$	$s^2 = 1/2$	$s^2 = 1/3$
Pr {La base aléatoire b est s−Siegel réduite }	$0.7\,10^{-17}$	$0.29\,10^{-5}$	0.034	0.209

TAB. 6.2 – *Majorants pour les probabilités asymptotiques de réductions des bases aléatoires*

Remarque 11. Pour $s = 1$, on ne peut même pas montrer que l'algorithme finit dans le pire des cas. C'est normal puisqu'il n'existe pas forcément une base 1-Siegel réduite dans un réseau. En revanche la valeur $s^2 = 4/3$ de la réduction de Siegel correspond au paramètre $t = 1$ de l'algorithme LLL. Dans ce cas, on sait que l'algorithme finit, mais la borne donnée par le théorème 3 est exponentielle en la taille de la base d'entrée et en la dimension du réseau. La question du nombre maximum d'itérations de l'algorithme LLL, lorsque le paramètre $t = 1$ est une question ouverte. Nous y apportons une réponse dans le chapitre suivant: lorsque la dimension du réseau est fixé, l'algorithme reste polynomial en la taille de la base d'entrée.

Pour tout r fixé, la probabilité que les r dernières conditions de s-Siegel soient satisfaites a bien une limite lorsque n tend vers l'infini. Pour $r = 1$, c'est une fonction simple de s donnée par le Corollaire 9.
Pour r quelconque fixé, la méthode de Laplace avec des intégrales multiples (à $r+1$ variables), nous permet là encore de conclure et avoir une expression précise de la limite. Mais le calcul, ainsi que l'expression de la limite, sont de plus en plus complexes.

La Proposition 10 permet de trouver cette limite lorsque $r = 2$.

Proposition 20. *Soit b une base aléatoire pleine, i.e. les n vecteurs sont choisis de façon aléatoire, uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Alors, la probabilité* *asymptotique* $\Pr_\infty^{(2)}(s)$, *que les deux dernières conditions de s-Siegel soient satisfaites vérifie*

$$\Pr_\infty^{(2)}(s) \;=\; \frac{16}{\pi} \iint_{\substack{\arctan s \le \theta \le \pi/2 \\ 0 \le \phi \le \pi/2 - \arctan(s\,\sin\theta)}} \cos^2\theta\ \sin\theta\ \sin^4\phi\ d\theta d\phi. \tag{6.35}$$

Preuve. Par définition, la probabilité $\Pr_{(n)}^{(2)}(s)$, que les deux dernières conditions de s-Siegel soient satisfaites est

$$\Pr_{(n)}^{(2)}(s) \;=\; \Pr\{\frac{l_n}{l_{n-1}} \ge s \quad \text{et} \quad \frac{l_{n-1}}{l_{n-2}} \ge s\}$$

De nouveau, grâce au théorème 4,

$$\Pr_{(n)}^{(2)}(s) := \frac{\int_{\Omega(s)} (1-x^2)^{\frac{n-3}{2}} x^2 (1-y^2)^{\frac{n-2}{2}} y (1-z^2)^{\frac{n-1}{2}} dxdydz}{\int_{\Omega(0)} (1-x^2)^{\frac{n-3}{2}} x^2 (1-y^2)^{\frac{n-2}{2}} y (1-z^2)^{\frac{n-1}{2}} dxdydz},$$

où $$\Omega(s) \;=\; \big\{(x,y,z) \in [0,1]^3 \;:\quad y \ge sx \quad \text{et} \quad z \ge sy\big\}.$$

Les numérateur et dénominateur de l'expression précédente sont tous deux de la forme (4.9),

$$\int_{\Omega(s)} g(x,y,z) e^{nh(x,y,z)} dxdydz,$$

avec $$g(x,y,z) = \frac{x^2 y}{(1-x^2)^{\frac{3}{2}}(1-y^2)(1-z^2)^{\frac{1}{2}}},$$

et $$h(x,y,z) = \frac{1}{2}\ln(1-x^2) + \frac{1}{2}\ln(1-y^2) + \frac{1}{2}\ln(1-z^2).$$

La fonction h atteint son maximum absolu sur $[0,1]^3$ en $(0,0,0)$ (qui est aussi un point de $\Omega(s)$, pour tout s) et au voisinage de l'origine, les donctions h et g vérifient

$$\begin{cases} h(x,y,z) & \sim & -\frac{1}{2}(x^2+y^2+z^2) \\ g(x,y,z) & \sim & x^2 y \end{cases}$$

Toutes les hypothèses de la Proposition 10 sont réunies, avec

$$(A, C, \lambda, \mu, \nu) = (1/2, 1, 2, 1, 0),$$

pour le numérateur et le dénominateur.

La Proposition 10 (une généralisation de la méthode de Laplace pour les intégrales triples) donne alors les parties principales du numérateur et du dénominateur, lorsque n tend vers l'infini:

$$\Gamma(\frac{5}{2}) \;\left(\frac{2}{n}\right)^{\frac{5}{2}} \iint_{\substack{\arctan s \le \theta \le \pi/2 \\ 0 \le \phi \le \pi/2 - \arctan(s\,\sin\theta)}} \cos^2\theta\ \sin\theta\ \sin^4\phi\, d\theta\, d\phi,$$

avec $s = 0$ pour le dénominateur, d'où la conclusion de la Proposition.

$\square$

La généralisation de la méthode précédente pour r fixé montre que la suite de fonctions $\Pr_{(n)}^{(r)}$ converge simplement à r et s fixés, vers une suite de fonctions $\Pr_\infty^{(r)}(s)$.

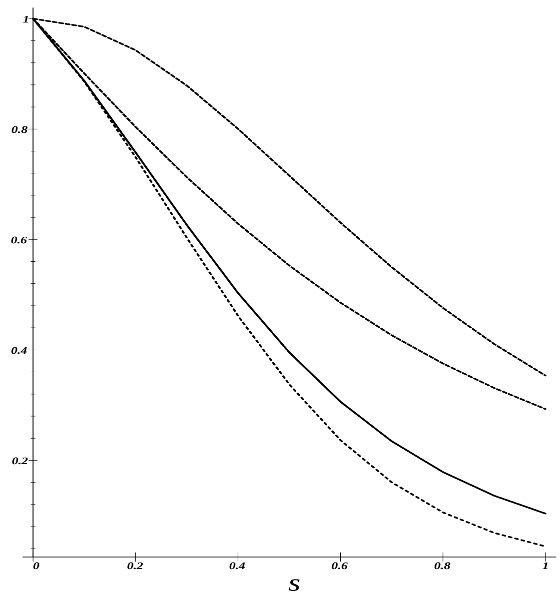

FIG. 6.6 – *Compraison entre les majorants de la probabilité de réduction au sens de s-Siegel d'une base aléatoire pleine.*

Les deux courbes supérieures correspondent aux probabilités-limites de réduction de la dernière, respectivement avant–dernière, condition. La courbe la plus basse correspond à la probabilité de réduction des deux dernières conditions *ensembles* (elle est obtenue grâce à la Proposition 20). La deuxième courbe en partant du bas illustre le simple produit des probabilités de réduction des deux deux dernières conditions. La figure montre bien que les événements correspondants aux conditions de réduction pour les deux derniers indices ne sont pas indépendants et que chacun des deux événements diminue la probabilité de réalisation de l'autre. Il semblerait donc que dans le majorant de la relation (6.34), la condition "j paire" ou "j impaire" soit superflue.

La suite (indicée par r) des $\Pr_\infty^{(r)}(s)$ est donc bien définie à chaque rang r. C'est une suite trivialement décroissante et minorée par 0.

Conjecture 2. *Soit b une base aléatoire pleine, i.e. dont les n vecteurs sont choisis uniformément et indépendamment dans la boule unité de $\mathbb{R}^n$. Pour tout $s \leq 1$, la probabilité que la base b est $s-$Siegel réduite a une limite lorsque n tend vers l'infini.*

Sous cette conjecture, le Corollaire 9 montre que cette limite est strictement inférieure à $1 - 1/\sqrt{1 + (1/s)^2}$ et la Proposition précédente (relation 6.35) donne une borne plus précise de sa valeur.

Pour vérifier la conjecture, il faudrait soit pouvoir généraliser la méthode de Laplace à des espaces de suites à valeurs dans $[0, 1]$, soit pouvoir donner le deuxième terme des asymptotiques des variables aléatoires $l_{n-j-i+1}/l_{n-i}$.

6.5 Conclusion

Pour une réduction donnée, si on distingue la condition de propreté de la deuxième condition de réduction, nous avons évalué dans ce Chapitre les probabilités asymptotiques de réductions des bases aléatoires tant pour les réductions classiques que pour les réductions que nous avons introduites dans cette thèse. Il n'existait jusque là aucune estimation de la probabilité de réduction des réseaux aléatoires au sens de Siegel. Ainsi, nous avons classé ces différents algorithmes en considérant leur comportement sur les réseaux aléatoires de notre modèle. Une des motivations premières de notre travail consiste à classer ces algorithmes. En effet, puisque les conditions exigées pour les nouvelles réductions introduites sont plus faibles que les réductions classiques de s-Siegel ou t-Lovász, il est naturel de s'attendre à ce que les algorithmes de réduction soient plus rapides. Cependant, l'étude du Chapitre 2, dans le pire des cas, donnait la même borne pour tous ces algorithmes.

Les bornes asymptotiques que nous avons établi pour la complexité de ces algorithmes dépendent bien sûr du modèle aléatoire choisi. Mais l'expérience semble montrer (Chapitre 8) que l'ordre relatif entre ces bornes est respecté dans d'autres modèles et que *les algorithmes de réduction rapide* soient toujours plus efficaces en moyenne.

Chapitre 7

Complexité de l'algorithme LLL pour le paramètre 1

7.1 Introduction

L'algorithme LLL est un algorithme itératif qui à partir d'une base quelconque d'un réseau euclidien, construit une base réduite au sens de t-Lovasz. Le paramètre d'approximation t doit être choisi dans l'intervalle $[1, \sqrt{3}]$. Ce chapitre traite de la complexité dans le pire des cas de l'algorithme LLL lorsque $t = 1$. Plus exactement, on majorera le nombre maximum d'itérations effectuées par cet algorithme.
Les résultats qui vont suivre sont valables pour des réseaux entiers (et peuvent facilement être généralisés à des réseaux rationnels). Un réseau entier $\mathcal{L}$ est donné par une base $(b_1, b_2, ..., b_n)$. Si M désigne la longueur (euclidienne) maximale des vecteurs de cette base

$$M = \max_{i \in \{1,2,..,n\}} |b_i|,$$

la taille de l'entrée de l'algorithme LLL est alors en $(n, \log M)$.
Les inventeurs de l'algorithme LLL ont montré [LLL82] que l'algorithme LLL est polynomial en la taille de l'entrée $(n, \log M)$, pour tout $t \neq 1$. Leur preuve échoue si $t = 1$ et le problème de la complexité de LLL$(t = 1)$ reste ouvert.
Kannan a presenté en 84 [Kan83] un algorithme qui utilise l'algorithme LLL (avec $t > 1$) et qui construit lorsque la dimension n est fixée, en temps polynomial en $\log M$ une base réduite au sens de Korkine-Zolotareff. La réduction de Korkine-Zolotareff étant plus fine que celle de $t-$Lovász même avec $t = 1$, il semblait naturel que l'algorithme LLL$(t = 1)$ soit également polynomial en $\log M$ lorsque la dimension n est fixée. C'est l'objet de ce chapitre que de le montrer. Plus précisémment, nous montrons que le nombre maximum d'itérations de LLL$(t = 1)$ est majoré par un polynôme en $\log M$ (chaque itération ne comporte que des opérations très simples et dont les temps d'exécution sont linéaires en $\log M$).
Plusieurs raisons motivent ce travail:

- C'est lorsque $t = 1$ que l'algorithme LLL peut être considéré comme une généralisation de l'algorithme bien connu de Gauss en dimension 2.
- L'algorithme LLL(t) est un algorithme d'approximation et $t = 1$ permet d'obtenir la meilleure approximation possible. Les critères habituels qui permettent de mesurer la

qualité d'une base b sont son défaut d'orthogonalité $\rho(b)$ et ses défauts de longueur $\mu_i(b)$. Pour la base de sortie de l'algorithme LLL(t),

$$\rho(b) \leq s^{\frac{n(n-1)}{2}} \quad \text{et} \quad \mu_i \leq s^{n-1} \quad \text{avec} \quad s = \sqrt{\frac{4t^2}{4-t^2}}.$$

Plus t est proche de 1, plus les majorations ci-dessus sont fines et la base de sortie est garantie d'une meilleure qualité.

- Même si la réponse apportée ici n'est que très partielle par rapport à la question ouverte de la complexité de LLL($t=1$), à notre connaissance il n'existait jusqu'alors aucune preuve montrant que l'algorithme LLL($t=1$) etait polynomial, ne serait-ce qu'en dimension 3. (En dimension 2, la complexité polynomiale a été établie par Lagarias [Lag82] et la meilleure borne donnée par Vallée [Val91].)

Dans la suite, la première partie donne une formulation de l'algorithme LLL et introduit quelques notations utiles. Nous y définissons en particulier ce qu'est *une situation de $t-$quasi-réduction* durant une exécution de l'algorithme LLL. La troisième partie rappelle les résultats de la dimension 2. Une situation de $t-$quasi-réduction entraîne alors la fin de l'exécution de l'algorithme en 2 itérations (Lemme 9). Dans la quatrième partie, nous établirons par une étude combinatoire qu'une situation de $t-$quasi-réduction entraîne la fin de l'exécution de l'algorithme en au plus 9 itérations (le Théorème de cette partie). Nous en déduirons une borne explicite linéaire en $\log M$ pour le nombre maximum d'itérations de l'algorithme LLL($t=1$), lorsque la dimension est 3. Quelques preuves techniques de l'étude combinatoire sont détaillées dans la cinquième partie. La dernière partie généralise les raisonnements précédents au cas de la dimension n fixée.

7.2 Description générale de l'algorithme LLL

Soit $\mathcal{L}$ un réseau donné par une base $(b_1, ..., b_n)$ de $\mathbb{R}^n$. On note $(b_1^*, ..., b_n^*)$, l'orthogonalisée de Schmidt de $(b_1, ..., b_n)$ et $m = (m_{ij})$ la matrice exprimant $(b_1, ..., b_n)$ dans $(b_1^*, ..., b_n^*)$. En notant $(u)_{\perp H}$ le projeté du vecteur u orthogonalement au sous-espace vectoriel H et $\langle u_1, u_2, ..., u_r\rangle$ l'espace vectoriel engendré par la famille de vecteurs $(u_1, u_2, ..., u_r)$, nous avons:

$$\left\{\begin{array}{lcl} b_1^* & = & b_1 \\ b_2^* & = & b_{2\perp\langle b_1\rangle} \\ \vdots & & \\ b_i^* & = & b_{i\perp\langle b_1,\cdots,b_{i-1}\rangle} \\ \vdots & & \\ b_n^* & = & b_{n\perp\langle b_1,\cdots,b_{n-1}\rangle} \end{array}\right. \quad \text{et} \quad m = \begin{array}{c} \\ b_1 \\ b_2 \\ \vdots \\ b_i \\ b_{i+1} \\ \vdots \\ b_n \end{array} \begin{array}{c} \begin{array}{ccccccc} b_1^* & b_2^* & \cdots & b_i^* & b_{i+1}^* & \cdots & b_n^* \end{array} \\ \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ m_{21} & 1 & \ddots & & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & & & \vdots \\ \cdots & \cdots & \cdots & 1 & 0 & & \vdots \\ \cdots & \cdots & \cdots & m_{i+1,i} & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & 0 \\ m_{n1} & m_{n2} & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix} \end{array}. \tag{7.1}$$

Le procédé même d'orthogonalisation de Gramm-Schmidt montre que la matrice m est triangulaire avec une diagonale formée de 1. Cette orthogonalisation s'effectue en temps polynomial par rapport à la taille $(n, \log M)$ de l'entrée.

La base $(b_1, ..., b_n)$ est dite *propre* si dans la matrice m de passage de la base (b) à sa base orthogonalisée (b^*),

$$-\frac{1}{2} < m_{ij} \leq \frac{1}{2} \quad \text{pour tout } i > j. \tag{7.2}$$

Il existe un algorithme très simple qui étant donné une base quelconque du réseau la rend propre en un temps polynomial en $(n, \log M)$. Cette algorithme est rappelé dans le Chapitre 2.

L'algorithme LLL(1) est une généralisation possible à toute dimension de sa version en dimension 2, appelé l'algorithme de Gauss, et dont tous les résultats de complexité sont maintenant connus.

Définition 5. La base $(b_1, ..., b_n)$ est réduite ou LLL($t = 1$)-réduite si elle vérifie les deux conditions suivantes:

(i) $(b_1, ..., b_n)$ est propre.

(ii) $\forall i \in \{1, 2, \cdots, n-1\} \qquad |b_i^*|^2 \ \leq \ {m_{i+1,i}}^2 |b_i^*|^2 + |b_{i+1}^*|^2$.

Si l'on note, pour tout i appartenant à $\{1, 2, \cdots, n-1\}$:

$$\begin{cases} u_i &= b_i^* \\ v_i &= m_{i+1,i} b_i^* + b_{i+1}^* \end{cases} \quad \text{et la } \textit{boîte } B_i \text{ le couple } (u_i, v_i),$$

par définition, la base $(b_1, ..., b_n)$ est réduite si elle est propre et les boîtes B_i sont réduites, pour tout i appartenant à $\{1, 2, \cdots, n-1\}$.

Définition 6. Soit t un réel vérifiant $t > 1$. On dira que la base $(b_1, \cdots, b_n)$ est *t-quasi-réduite* lorsqu'elle vérifie les conditions suivantes:

(i) la base $(b_1, \cdots, b_{n-1})$ est propre.

(ii) Pour $1 \leq i \leq n-2$, toutes les boîtes B_i sont réduites.

(iii) La dernière boîte B_{n-1} n'est pas réduite mais

$$|m_{n,n-1}| < \frac{1}{2} \quad \text{et} \quad \frac{1}{t^2} |b_{n-1}^*|^2 \ \leq \ {m_{n,n-1}}^2 |b_{n-1}^*|^2 + |b_n^*|^2.$$

(La dernière boîte est $t-$Gauss réduite ou simplement $t-$réduite.)

Autement dit, si la base $(b_1, \cdots, b_{n-1})$ est réduite et la base $(b_1, \cdots, b_n)$ est non-réduite alors pour tout $t > 1$ tel que la dernière boîte est $t-$réduite, la base $(b_1, \cdots, b_n)$ est $t-$quasi-réduite.

Voici un énoncé possible de l'algorithme LLL(1):

Algorithme LLL

Donnée : une base b d'un réseau $\mathcal{L}$ de dimension n de $\mathbb{R}^p$.
Résultat : une base b de $\mathcal{L}$ réduite ou $LLL(1)$ réduite

$i := 1$;
Tant que $i < n$ **répéter**
- Calculer la matrice m.
- **Translater** b_{i+1} parallèlement à b_i pour le rendre le plus court possible:
  Calculer r_i, l'entier le plus proche de $m_{i+1,i}$;
  Faire $b_{i+1} := b_{i+1} - r_i b_i$;
- **Tester si** $|b_i^*|^2 \ \leq \ m_{i+1,i}{}^2 |b_i^*|^2 + |b_{i+1}^*|^2$
  Si OUI :
  Rendre $(b_1, ..., b_{i+1})$ propre.
  $i := i + 1$;
  Si NON :
  Échanger b_i et b_{i+1};
  Si $i \neq 1$ alors $i := i - 1$;

Dans toute la suite, une itération de l'agorithme LLL est une itération de la boucle *tant que* dans l'énoncé précédent. Chaque itération comporte exactement un test de réduction ou non-réduction d'une boîte B_i $(1 \leq i \leq n-1)$. Le nombre d'itérations de l'algorithme est donc exactement le nombre de tests de réductions des boîtes. Les itérations écoulées entre deux passages successifs de i à la valeur $n-1$ forment *une phase* de l'algorithme.
Par ailleurs, pendant une exécution de l'algorithme LLL, l'indice i de l'énoncé ci-dessus est à un instant donné, le plus grand entier k, tel que $(b_1, ..., b_k)$ soit réduite. À l'issue de chaque test, sa valeur évolue d'une des trois façons suivantes.

1. l'indice i peut être incrémenté de 1 (test positif)
2. l'indice i peut être décrémenté de 1 (test négatif)
3. l'indice i peut garder son ancienne valeur, qui est alors nécessairement 1. (i valait déjà 1 avant le test et le test est négatif).

Ainsi, la variation de l'indice i pendant une exécution de l'algorithme LLL peut être représentée par une marche aléatoire (figure 7.1).

Au terme de chaque phase non finale de l'algorithme, l'indice i a la valeur $n-1$ (juste avant le test) et il existe $t > 1$ tel que la base $(b_1, \ldots, b_n)$ est $t-$quasi-réduite. On dit alors qu'on est en *situation de t-quasi-réduction*. Ces situations sont indiquées sur la figure 7.1.
Par définition, lors d'une situation de quasi-réduction, le test concerne la réduction de la dernière boîte B_{n-1} et le résultat de ce test est négatif.
Nous séparons les situations de t-quasi-réduction en deux types selon qu'entre deux passages successifs à la dernière boîte, on passe plusieurs fois de suite dans la première boîte. Plus précisément, une situation de $t-$quasi-réduction sera dite
de type I, si pendant la phase qui suit, une des 3 conditions équivalentes suivantes est remplie:
(i) l'algorithme ne fait jamais l'échange entre b_1 et b_2.

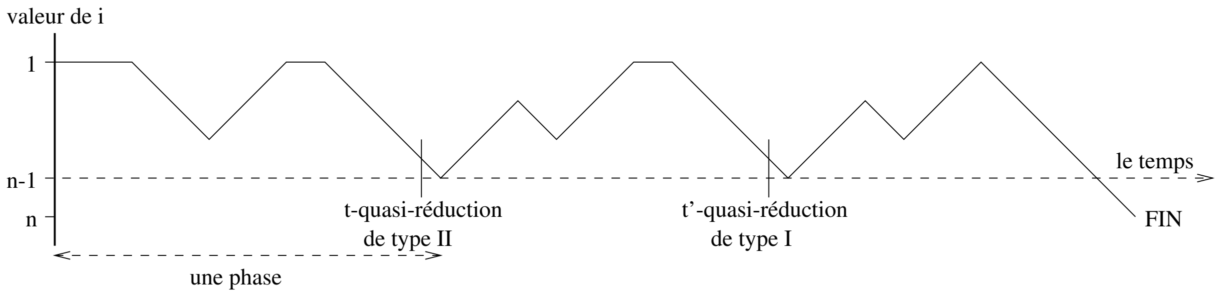

FIG. 7.1 – *Un exemple de la variation de l'indice i représentée sous-forme de marche aléatoire en dimension* 4

(*ii*) il n'existe aucun palier (figure 7.1).

(*iii*) il n'y a aucun test négatif, dans la première boîte B_1.

de type II, si pendant la phase qui suit, une des 3 conditions équivalentes suivantes est remplie:

(*ii*) l'algorithme fait au moins une fois l'échange entre b_1 et b_2.

(*ii*) il existe au moins un palier (figure 7.1).

(*iii*) il y a au moins un test négatif, dans la première boîte.

On notera en abrégé:

$t-$qr.	pour	$t-$quasi-réduction
$t-$qr.I	pour	$t-$quasi-réduction de type I
$t-$qr.II	pour	$t-$quasi-réduction de type II
FIN	pour	fin de l'algorithme

7.3 Les résultats connus en dimension 2: l'algorithme de Gauss

En dimension 2, une phase de l'algorithme est tout simplement une itération de l'algorithme (une itération de la boucle *tant que*) et le premier test positif signifie la fin de l'algorithme. Pour compter le nombre maximum d'itérations, il suffit donc de compter le nombre de tests négatifs. Nous allons montrer qu'en un nombre polynomial en $\log_t M$ d'itérations, la base (b_1, b_2) sera au moins $t-$quasi-réduite ($t-$quasi-réduite ou réduite). En outre, pour tout t, $1 < t < \sqrt{3}$, une situation de $t-$quasi-réduction entraîne la fin de l'algorithme en 2 itérations. (la première est l'itération durant laquelle la situation de quasi-réduction se produit)

On pose $$d_1 \;=\; |b_1|^2 \;=\; |b_1^*|^2.$$

À t fixé, avant d'arriver à une situation de $t-$quasi-réduction, lors de chaque test négatif, nous avons $|b_2|^2 < (1/t^2)|b_1|^2$ et b_1 et b_2 seront échangés, ce qui signifie encore que la quantité d_1 décroît d'au moins $1/t^2$. Après k_- itérations et sachant que le réseau est entier,

$$1 \;\leq\; d_1 \;<\; \frac{1}{(t^2)^{k_-}} M^2$$

Donc le nombre k_- d'itérations pour arriver à un situation de $t-$quasi-réduction est au plus $\log_t M$.

Le lemme suivant permet de conclure dans le cas de la dimension 2.

Lemme 9. *Soit t un réel vérifiant $1 < t < \sqrt{3}$. En dimension 2, une situation de $t-$quasi-réduction entraîne la fin de l'exécution de l'algorithme en 2 itérations de tests négatif, puis positif.*

$$t - qr. \xrightarrow{\quad 1<t<\sqrt{3} \quad} \textit{FIN}, \qquad \textit{en 2 itérations de tests } (-,+).$$

Preuve. Lors d'une situation de $t-$quasi-réduction, la base (b_1, b_2) vérifie

$$(b_1, b_2) \text{ est propre } (|m_{21}| < \frac{1}{2}) \quad \text{et} \quad \frac{1}{t^2}|b_1|^2 \leq |b_2|^2 < |b_1|^2. \tag{7.3}$$

Avant l'échange, $m = \begin{array}{c c} & \begin{array}{cc} b_1^* & b_2^* \end{array} \\ \begin{array}{c} b_1 \\ b_2 \end{array} & \begin{pmatrix} 1 & 0 \\ m_{21} & 1 \end{pmatrix} \end{array}.$

Après l'échange, $\hat{m} = \begin{array}{c c} & \begin{array}{cc} \hat{b_1^*} & \hat{b_2^*} \end{array} \\ \begin{array}{c} b_2 \\ b_1 \end{array} & \begin{pmatrix} 1 & 0 \\ \hat{m}_{21} & 1 \end{pmatrix} \end{array}$, avec $\hat{m}_{21} = \dfrac{(b_1, b_2)}{|b_2|^2} = \dfrac{(b_1, b_2)}{|b_1|^2}\dfrac{|b_1|^2}{|b_2|^2} = m_{21}\dfrac{|b_1|^2}{|b_2|^2}$

Les relations (7.3) donnent $\quad |\hat{m}_{21}| < \dfrac{t^2}{2} < \dfrac{3}{2}, \quad$ pour $1 < t < \sqrt{3}$.

Soit encore, $$\lfloor \hat{m}_{21} \rceil \in \{-1, 0, 1\}.$$

Donc le vecteur b_1 sera transformé en $b_1 + \varepsilon b_2$, avec $\varepsilon \in \{-1, 0, 1\}$.

Or $$\left\{ \begin{array}{lcrcrcr} |b_1 + \varepsilon b_2|^2 & = & (1 + \varepsilon m_{21})^2 & |b_1^*|^2 & + & \varepsilon^2 & |b_2^*|^2 \\ |b_2|^2 & = & (m_{21})^2 & |b_1^*|^2 & + & & |b_2^*|^2 \end{array} \right.$$

Et $\quad \forall \varepsilon \in \{-1, 0, 1\}, \quad$ on a bien: $|b_1 + \varepsilon b_2|^2 \geq |b_2^*|^2$.

(Si $\varepsilon = 0$, c'est une conséquence de la $t-$qr. Sinon, $\varepsilon^2 = 1$ et comme $|m_{21}| \leq \frac{1}{2}$, on a bien $(1 + \varepsilon m_{21})^2 \geq m_{21}^2$.)
Le test suivant la $t-$quasi-réduction est donc bien positif. □

Le nombre maximum d'itérations de l'algorithme de Gauss est donc majoré par $\log_{\sqrt{3}} M + 2$. Cette borne n'est pas optimale, mais c'est ce raisonnement que nous allons généraliser à la dimension 3, dans le prochain paragraphe.

7.4 Cas de la dimension 3

On procède comme précédemment:
On montre d'abord qu'en un nombre polynomial d'itérations (en $\log_t M$), la base d'entrée (b_1, b_2, b_3) sera $t-$quasi-réduite. Puis, nous verrons (le théorème de ce paragraphe) que pour $1 < t < \sqrt{6/5}$, une base $t-$quasi-réduite sera réduite en au plus 9 itérations. Les raisonnements que nous utilisons donnent par ailleurs une caractérisation précise de la qualité d'une base réduite en dimension 3.

7.4.1 D'une base quelconque à la quasi-réduction

En partant d'une base quelconque, l'algorithme LLL produit une base quasi-réduite en un nombre polynomial d'itérations. Plus exactement,

Proposition 21. *Soit t un réel vérifiant $1 < t < \sqrt{3}$. Le nombre d'itérations effectuées par l'algorithme LLL, pour arriver à la première $t-$quasi-réduction est majoré par $\log_{\sqrt{3}} M + 6 \log_t M + 2$.*

Preuve. On note: $$\begin{cases} d_1 = |b_1^*|^2 = |b_1|^2 \\ d_2 = |b_1^*|^2 |b_2^*|^2 \end{cases}$$

Il est clair que $d_1 < M^2$ et $d_2 < M^4$.
On va compter le nombre maximum k, de tests pour arriver à la $t-$quasi-réduction (chaque itération comporte 1 unique test).
On note respectivement

- k_1^- , le nombre de tests négatifs avec $i = 1$, avant la $t-$quasi-réduction,
- k_2^- , le nombre de tests négatifs avec $i = 2$, avant la $t-$quasi-réduction.(on ne décomptera pas le dernier test négatif), et
- k^+ , le nombre de tests positifs avant la $t-$quasi-réduction (nécessairement avec $i = 1$).

$$\text{Nous avons alors} \quad k = k_1^- + k_2^- + k^+. \tag{7.4}$$

En dimension 3, une phase de l'algorithme est la suite d'itérations qui s'exécutent entre deux passages successifs de i à la valeur 2. Chaque phase comporte ici exactement un test positif. Donc, en notant p le nombre de phases avant la $t-$quasi-réduction,

$$k^+ = p = k_2^- + 1. \tag{7.5}$$

Pendant l'exécution d'une phase tous les tests négatifs avec $i = 1$ entraînent une décroissance de la quantité d_1 de $1/3$ sauf éventuellement la dernière (cf. le lemme 9: un réseau (b_1, b_2), $\sqrt{3}-$qr. sera réduite en deux itérations de test $-, +$).
Si $\hat{k}_1^-$ est le nombre de test négatifs qui font décroître la quantité d_1, comme le réseau est entier et cette quantité ne croît jamais, elle vérifie

$$1 \ \le \ d_1 \ < \ \frac{1}{(\sqrt{3}^2)^{\hat{k}_1^-}} M^2. \quad \text{D'où:}$$

$$k_1^- \ \le \ \log_{\sqrt{3}} M \ + \ p. \tag{7.6}$$

Par ailleurs les tests négatifs ou positifs avec $i = 1$, ne font pas varier la quantité d_2, qui est le déterminant de Gramm du réseau (b_1, b_2).
Les tests négatifs avec $i = 2$ font décroître $|b_2^*|^2$ donc d_2 de $1/t^2$. Donc pour un réseau entier, au terme de k_2^- tests négatifs, la quantité d_2 vérifie

$$1 \ \le \ d_2 \ < \ \frac{1}{(t^2)^{k_2^-}} M^4. \quad \text{D'où:}$$

$$k_2^- \ \leq \ 2\log_t M \tag{7.7}$$

Finalement, d'après les équations (7.4), (7.5), (7.6) et (7.7), le nombre total k d'itérations avant la $t-$quasi-réduction est majoré:

$$k \leq \log_{\sqrt{3}} M + 6\log_t M + 2.$$ □

En particulier pour $1 < t^2 \leq 6/5$, $\quad k \leq \log_{\sqrt{3}} M + 6\log_{\sqrt{\frac{6}{5}}} M + 2.$

7.4.2 De la quasi-réduction à la réduction

Rappelons que pour tout $t > 1$, (b_1, b_2, b_3) est $t-$quasi-réduite lorsqu'elle vérifie les 2 conditions suivantes.

(i) La boîte B_1 est 1-Lovász-réduite (ou encore Gauss-réduite).

(ii) La boîte B_2 est t-Lovász-réduite mais non 1-Lovász-réduite.

Théorème 12. *Soit t un réel vérifiant $1 < t < \sqrt{6/5}$. En dimension 3, une situation de $t-$quasi-réduction entraîne la fin de l'exécution de l'algorithme en au plus 9 itérations.*

$$t - qr. \xrightarrow[\textit{au plus 9 itérations.}]{1<t<\sqrt{6/5}} \textit{FIN.}$$

Preuve. Elle est basée sur le Lemme 10 et la Proposition 22. Si la $t-$quasi-réduction est de type I, le Lemme 10 montre que l'algorithme termine son exécution en 3 itérations. La Proposition 22 prouve par ailleurs qu'une $t-$quasi-réduction de type II entraîne en au plus 6 itérations soit la fin de l'exécution de l'algorithme, soit une nouvelle situation de $t-$quasi-réduction nécessairement de type I et dans ce dernier cas, de nouveau le Lemme 10 permet de conclure.
Toutes les possibilités sont résumées dans le diagramme suivant où au dessus de chaque flèche, sont indiqués les résultats des tests successifs effectués par l'algorithme entre deux situations considérées (le test positif final est omis). À gauche, sont décomptés les nombres totaux d'itérations (ou de tests) après la quasi-réduction, dans chaque cas.

$$\begin{array}{c} \text{non}(t\text{−qr.}) \\ (1 < t^2 < 6/5) \end{array} \quad \Longrightarrow \quad \left\{ \begin{array}{lllllllr} \text{FIN} & & & & & & & (1) \\ t\text{−qr.I} & \stackrel{-+}{\longrightarrow} & \text{FIN} & & & & & (3) \\ t\text{−qr.II} & \stackrel{--+}{\longrightarrow} & \text{FIN} & & & & & (4) \\ t\text{−qr.II} & \stackrel{--+}{\longrightarrow} & t'\text{−qr.I} & \stackrel{-+}{\longrightarrow} & \text{FIN} & & & (6) \\ t\text{−qr.II} & \stackrel{--+}{\longrightarrow} & t'\text{−qr.II} & \stackrel{--+}{\longrightarrow} & \text{FIN} & & & (7) \\ t\text{−qr.II} & \stackrel{--+}{\longrightarrow} & t'\text{−qr.II} & \stackrel{--+}{\longrightarrow} & t''\text{−qr.I} & \stackrel{-+}{\longrightarrow} & \text{FIN} & (9) \end{array} \right.$$

□

Remarque: Une étude combinatoire plus fine du cas

$$t\text{−qr.II} \longrightarrow t'\text{−qr.II}$$

permettrait probablement de remplacer dans la Proposition 22, $1 < t^2 < 6/5$ par $1 < t^2 < 12/7$, mais $12/7$ nous paraît le meilleur majorant possible pour cette modélisation.

Lemme 10. *Soit t un réel vérifiant $1 < t < \sqrt{3}$. En dimension 3, une situation de $t-$quasi-réduction de type I entraîne la fin de l'exécution de l'algorithme en 3 itérations de tests respectifs $(-,+,+)$.*

$$t - qr.I \xrightarrow[\text{3 itérations de tests respectifs } (-,+,+)]{1<t<\sqrt{3}} FIN.$$

Preuve. Comme la base (b_1, b_2, b_3) n'est pas réduite, les vecteurs b_2 et b_3 seront échangés, puis l'algorithme translatera b_3 parallèlement à b_1, afin de le rendre le plus court possible. Soit b_3' le vecteur obtenu. Comme la quasi-réduction est de type I, le vecteur b_1 ne sera pas échangé à ce stade, (ou encore $|b_3'| > |b_1|$) et la base (b_1, b_3') est Gauss-réduite. En résumé,

$$m = \begin{array}{c} \\ b_1 \\ b_2 \\ b_3 \end{array} \begin{array}{c} \begin{array}{ccc} b_1^* & b_2^* & b_3^* \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \end{array} \xrightarrow[\text{de tests: (-,+)}]{\text{en 2 itérations}} \hat{m} = \begin{array}{c} \\ b_1 \\ b_3' \\ b_2 \end{array} \begin{array}{c} \begin{array}{ccc} \hat{b}_1^* & \hat{b}_2^* & \hat{b}_3^* \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ \hat{m}_{21} & 1 & 0 \\ \hat{m}_{31} & \hat{m}_{32} & 1 \end{pmatrix} \end{array}.$$

Puisque (b_1, b_2, b_3) est $t-$qr, il en est de même en particulier pour la base propre $((b_2)_{\perp b_1}, (b_3')_{\perp b_1})$. D'après le Lemme 9 (algorithme de Gauss), la base $((b_3')_{\perp b_1}, (b_2)_{\perp b_1})$ sera réduite en une itération de test positif[1]. □

Proposition 22. *Soit t un réel vérifiant $1 < t < \sqrt{6/5}$. En dimension 3, une situation de $t-$quasi-réduction de type II entraîne la fin de l'exécution de l'algorithme ou une situation de $t'-$qr.I avec $1 < t' < \sqrt{3}$, en au plus 6 itérations.*

$$t - qr.II \xrightarrow[\text{au plus 6 itérations}]{1<t<\sqrt{6/5}} \left\{ \begin{array}{c} FIN \\ ou \\ t' - qr.I \quad avec\ 1 < t' < \sqrt{3} \end{array} \right. .$$

Preuve. Si pendant l'exécution de l'algorithme LLL, se produit une situation de $t-$qr.II pour une base (b_1, b_2, b_3), alors les vecteurs b_2 et b_3 seront échangés, puis l'algorithme translatera b_3 parallèlement à b_1, afin de le rendre le plus court possible. Soit b_3' le vecteur ainsi obtenu. Comme la quasi-réduction est de type II, le vecteur b_3' vérifie $|b_3'| < |b_1|$ et donc b_1 et b_3' seront échangés. En résumé,

$$m = \begin{array}{c} \\ b_1 \\ b_2 \\ b_3 \end{array} \begin{array}{c} \begin{array}{ccc} b_1^* & b_2^* & b_3^* \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \end{array} \xrightarrow[\text{de tests: (-,-)}]{\text{en 2 itérations}} \hat{m} = \begin{array}{c} \\ b_3' \\ b_1 \\ b_2 \end{array} \begin{array}{c} \begin{array}{ccc} \hat{b}_1^* & \hat{b}_2^* & \hat{b}_3^* \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ \hat{m}_{21} & 1 & 0 \\ \hat{m}_{31} & \hat{m}_{32} & 1 \end{pmatrix} \end{array}. \tag{7.8}$$

La preuve de cette proposition repose alors sur le Lemme 13 et le Corollaire 10. D'après le Lemme 13, une $t-$quasi-réduction de type II peut éventuellement déboucher sur une

1. Le Lemme 9 n'exige pas que la base $t-$Gauss réduite soit entière.

nouvelle, mais le Corollaire 10 interdit plus de 2 quasi-réductions de type II, ce qui prouve la Proposition.
Toutes les possibilités sont décrites dans le diagramme suivant où chaque flèche schématise 3 itérations de l'algorithme LLL.

$$t-\text{qr.II} \xrightarrow[\text{(Lemme 13)}]{1 < t^2 < 6/5} \begin{cases} \text{FIN} \\[4pt] \begin{array}{l} t'-\text{qr.I} \\ 1 < t'^2 \leq \frac{12}{7} < 3 \end{array} \\[10pt] \begin{array}{l} t'-\text{qr.II} \\ 1 < t'^2 \leq \frac{12}{7} \end{array} \xrightarrow[\begin{pmatrix}\text{Lemme 13 et} \\ \text{Corollaire 10}\end{pmatrix}]{\qquad\qquad} \begin{array}{l} t''-\text{qr.I} \\ 1 < t''^2 \leq 3 \end{array} \end{cases}$$

□

Lemme 11. *Soit t un réel vérifiant $1 < t < (3/2)$. Si la base (b_1, b_2, b_3) est $t-$quasi-réduite et propre, alors parmi les vecteurs du réseau se trouvant en dehors du plan $\langle b_1, b_2 \rangle$, il existe au signe près, au plus un vecteur du réseau de norme strictement inférieure à $|b_3|$.*

Ce résultat qui reste vrai a fortiori lorsque la base (b_1, b_2, b_3) est réduite, permet par ailleurs de caractériser précisément la qualité d'une base réduite au sens de 1-Lovász en dimension 3. Une démonstration détaillée est donnée dans la Section 7.5.

Corollaire 10. *Soit t un réel vérifiant $1 < t < (3/2)$. Durant une exécution de l'algorithme LLL(1), il se produit au plus 2 situations de $t-$qr.II.*

Preuve. Supposons que durant l'exécution de l'algorithme LLL(1), il se produit une situation $t-$quasi-réduction de type II avec $1 < t < (3/2)$. D'aprés le Lemme précédent, si $(b_1, b_2, b'_3) := \text{Propre}(b_1, b_2, b_3)$, alors au signe près, le réseau contient en dehors du plan $\langle b_1, b_2 \rangle$ au plus un vecteur de norme strictement inférieure à b'_3.
Comme la $t-$quasi-réduction est de type II, $|b'_3| < |b_1|$ et l'algorithme échangera nécessairement $|b_1|$ et $|b'_3|$.
Mais (b_1, b_2) étant Gauss-réduite, b_1 est le plus court vecteur du réseau appartenant au plan $\langle b_1, b_2 \rangle$ et $|b'_3| < |b_1|$ implique qu'il n'existe dans le plan $\langle b_1, b_2 \rangle$ aucun vecteur du réseau plus court que b'_3. Par ailleurs, comme en dehors du plan $\langle b_1, b_2 \rangle$, il existe au plus un vecteur plus court que b'_3, on déduit que dans le réseau tout entier, il existe, au signe près, au plus un vecteur plus court que b'_3. Donc, dans la suite de l'exécution de l'algorithme, b'_3 pourra être échangé au plus une fois. Cela signifie qu'au plus une nouvelle situation de $t-$quasi-réduction de type II pourrait se présenter. □

Le Lemme suivant montre que les $\hat{m}_{i,j}$ (cf. (7.8)) sont bornés. Plus précisément,

Lemme 12. *Lors d'une situation de $t-$quasi-réduction de type II, après 2 itérations de tests négatifs, les coefficients de la nouvelle matrice $\hat{m}$ de passage vérifient:*

(a)	*Pour* $1 < t < (3/2)\sqrt{3}$,	$\lfloor \hat{m}_{21} \rceil \in \{-1, 0, 1\}$
(b)	*Pour* $1 < t < 2$,	$\lfloor \hat{m}_{32} \rceil \in \{-1, 0, 1\}$
(c)	*Pour* $1 < t^2 < \sqrt{10} - 1$,	$\lfloor \hat{m}_{31} \rceil \in \{-1, 0, 1\}$

La preuve détaillée est donnée dans la Section 7.5.

Lemme 13. *Soit t un réel vérifiant $1 < t < 2$. En dimension 3, une situation de $t-$quasi-réduction de type II entraîne la fin de l'exécution de l'algorithme ou une situation de $t'-$qr. avec,*

$$\frac{1}{t'^2} > \frac{1}{t^2} - \frac{1}{4},$$

en au plus 3 itérations de tests respectifs $(-,-,+)$.

$$t - qr.II \quad \xrightarrow[\text{au plus 3 itérations de tests respectifs } (-,-,+).]{1<t<2} \quad \begin{cases} \mathit{FIN} \\ \mathit{ou} \\ t' - qr.\ \mathit{avec} \quad \frac{1}{t'^2} > \frac{1}{t^2} - \frac{1}{4} \end{cases}.$$

La preuve de ce dernier fait est basée sur l'assertion (a) du lemme précédent et la définition d'une situation de $t-$qr.II. Elle est donnée dans la Section 7.5.

7.5 Les preuves

7.5.1 Preuve du Lemme 11

Soit u un vecteur du réseau. Il s'écrit dans la base (b) sous la forme $u = x\,b_1 + y\,b_2 + z\,b_3$ avec $(x, y, z) \in \mathbb{Z}^3$.
La matrice triangulaire $(m_{i,j})$ de passage de (b) à (b^*) est rappelée en (7.8). Le vecteur u s'exprime dans la base orthogonale (b^*) et sa norme vérifie

$$|u|^2 = (x + y\,m_{21} + z\,m_{31})^2\,|b_1^*|^2 + (y + z\,m_{32})^2\,|b_2^*|^2 + |b_3^*|^2. \tag{7.9}$$

Nous allons montrer successivement:

1. Si $|z| > 1$ ou $|y| > 1$ ou $|x| > 1$, alors $|u| > |b_3|$.
2. Si $z = 1$, alors il existe au plus un couple $(x, y) \in \{0, 1, -1\}^2 \backslash \{(0, 0)\}$ tel que $|u| < |b_3|$

1. Lorsque (b_1, b_2, b_3) est $t-$qr. et propre, nous avons:

$$\frac{3}{4} \le 1 - m_{21}^2 \le \frac{|b_2^*|^2}{|b_1^*|^2} \quad (7.10.1) \qquad \text{et} \qquad \frac{4-t^2}{4t^2} \le \frac{1}{t^2} - m_{32}^2 \le \frac{|b_3^*|^2}{|b_2^*|^2} \quad (7.10.2) \tag{7.10}$$

Par ailleurs, les vecteurs b_2 et b_3 s'expriment dans la base orthogonale (b^*), comme en (7.8), et grâce à la propreté leurs normes vérifient

$$|b_2|^2 \ \le\ |b_2^*|^2 + \frac{1}{4}|b_1^*|^2 \quad \text{et} \quad |b_3|^2 \ \le\ |b_3^*|^2 + \frac{1}{4}\left(|b_1^*|^2 + |b_2^*|^2\right).$$

Les relations (7.10) donnent alors

$$\frac{|b_2^*|}{|b_2|} > \frac{\sqrt{3}}{2} \quad (7.11.1) \qquad \text{et} \qquad \frac{|b_3^*|}{|b_3|} > \sqrt{\frac{12-3t^2}{12+4t^2}} \quad (7.11.2). \tag{7.11}$$

(a) La relation (7.9) entraîne $\quad |u| \ \ge\ |z|\,|b_3^*|$.
Donc si $|z| \ge 2$, pour $1 < t \le \dfrac{3}{2}$ et d'après (7.11.2), on a bien $|u| > |b_3|$.

(b) Si $|z| = 1$, la relation (7.9) montre que $|u|^2 \geq (y \pm m_{32})^2 \, |b_2^*|^2 + |b_3^*|^2$.
Pour $|y| \geq 2$, nous avons:

$$|u|^2 \quad \geq \quad \frac{9}{4} |b_2^*|^2 + |b_3^*|^2 \quad \geq \quad 2\,|b_2^*|^2 + \frac{1}{4} |b_2^*|^2 + |b_3^*|^2$$

Finalement comme (b_1, b_2) est réduite (7.10.1) et b_3 est propre,

$$|u|^2 \quad \geq \quad 2\frac{3}{4} |b_1^*|^2 + \frac{1}{4} |b_2^*|^2 + |b_3^*|^2 \quad > \quad |b_3|^2$$

Notons par ailleurs que si $|z| = 1$ et $|y| \geq 2$, alors $|u| > |b_2|$.
En effet, pour $y \in \mathbb{Z}^*$ et grâce à la propreté: $|u| \geq |y \pm m_{32}| \, |b_2^*| \geq (|y| - 1/2) \, |b_2^*|$

D'où si $|y| \geq 2$, en utilisant (7.11.1) $|u| \geq \dfrac{3}{2} \dfrac{\sqrt{3}}{2} |b_2| > |b_2|$

(c) Si $|z| = 1$, $|y| \leq 1$ et $|x| \geq 2$, d'après la relation(7.9) et grâce encore à la propreté:

$$|u|^2 \; > \; |b_1^*|^2 + (m_{32})^2 \, |b_2^*|^2 + |b_3^*|^2 \; > \; |b_3|^2$$

Notons également que dans ce cas, $|u| > |b_1|$.

2. Soit u un vecteur du plan $z = 1$ de la forme $u = b_3 + x\, b_1 + y\, b_2$ avec $(x, y) \in \{-1, 0, 1\}^2 \backslash \{(0, 0)\}$.
Nous allons montrer qu'il existe au plus un vecteur u, vérifiant $|u| < |b_3|$. Ce vecteur dépend des signes de m_{21}, m_{31} et m_{32}. Le tableau de la figure 7.1 résume la situation. Par la suite, on va étudier le rapport $|u|^2/|b_3|^2$.

$$\frac{|u|^2}{|b_3|^2} = 1 + x^2 \frac{|b_1|^2}{|b_3|^2} + y^2 \frac{|b_2|^2}{|b_1|^2} \frac{|b_1|^2}{|b_3|^2} + 2xy \frac{(b_1, b_2)}{|b_1|^2} \frac{|b_1|^2}{|b_3|^2} + 2x \frac{(b_1, b_3)}{|b_1|^2} \frac{|b_1|^2}{|b_3|^2} + 2y \frac{(b_2, b_3)}{|b_1|^2} \frac{|b_1|^2}{|b_3|^2}$$

Or les vecteurs b_2 et b_3 s'expriment dans la base orthogonale (b^*) comme en (7.8) et

$$\frac{(b_2, b_3)}{|b_1|^2} \;=\; m_{21} m_{31} + m_{32} \frac{|b_2^*|^2}{|b_1|^2}.$$

On pose alors: $$\frac{|u|^2}{|b_3|^2} = 1 + A(m_{21}, m_{31}, m_{32}, x, y) \frac{|b_1|^2}{|b_3|^2},$$

avec $A(m_{21}, m_{31}, m_{32}, x, y) := x^2 + 2m_{21}xy + 2m_{31}x + 2m_{21}m_{31}y + y^2 \dfrac{|b_2|^2}{|b_1|^2} + 2m_{32} \dfrac{|b_2^*|^2}{|b_1|^2} y$.

Il ne reste plus qu'à montrer que dans l'ensemble $\{-1, 0, 1\}^2 \backslash \{(0, 0)\}$, il existe au plus un couple (x, y) tel que $A(m_{21}, m_{31}, m_{32}, x, y) < 0$.
On remarque d'abord que:

$$A(m_{21}, m_{31}, m_{32}, -x, y) \;=\; A(-m_{21}, -m_{31}, m_{32}, x, y) \tag{7.12}$$
$$A(m_{21}, m_{31}, m_{32}, x, -y) \;=\; A(-m_{21}, m_{31}, -m_{32}, x, y) \tag{7.13}$$

m_{21}	m_{31}	m_{32}	u
$+$	$+$	$+$	$b_3 - b_2$
$+$	$+$	$-$	$b_3 - b_1 + b_2$
$+$	$-$	$+$	$b_3 + b_1 - b_2$
$+$	$-$	$-$	$b_3 + b_2$
$-$	$+$	$+$	$b_3 - b_2 - b_1$
$-$	$+$	$-$	$b_3 + b_2$
$-$	$-$	$+$	$b_3 - b_2$
$-$	$-$	$-$	$b_3 + b_2 + b_1$

TAB. 7.1 – *L'unique vecteur u suscesptible d'être plus court que le vecteur b_3 en fonction des signes de m_{21}, m_{31} et m_{32}.*

Lorsqu'il n'y a pas d'ambiguïté, $A(m_{21}, m_{31}, m_{32}, x, y)$ est noté simplement $A(x, y)$. Tous les couples (x, y) de $\{-1, 0, 1\}^2 \backslash \{(0, 0)\}$ sont envisagés ci-dessous.

Si $y = 0$, aucun vecteur u du plan $z = 1$ n'est plus court que b_3.
En effet, $A(x, 0) = x^2 + 2m_{31}x$.
Etant donné que $|x| = 1$ et que la base est propre: $\quad A(x, 0) \geq 0$.

Le vecteur $b_3 + b_1 + b_2$ est plus court que b_3 si $A(1, 1) < 0$.

Or, $A(1,1) = \overbrace{\frac{|b_2|^2}{|b_1|^2} + 2m_{32}\,\frac{|b_2^*|^2}{|b_1|^2}}^{\geq 0} + 1 + 2m_{21} + 2m_{31} + 2m_{21}m_{31}$

1. La fonction f définie par $f(m_{21}, m_{31}) := 1 + 2m_{21} + 2m_{31} + 2m_{21}m_{31}$,

$$\text{a une dérivée} \quad \frac{\partial f}{\partial m_{31}} = 2(1 + m_{21}) > 0.$$

$$\text{Donc} \quad f(m_{21}, -\frac{1}{2}) = m_{21} < f(m_{21}, m_{31}).$$

   En particulier, l'inégalité $m_{21} \geq 0$ entraîne $A(1, 1) > 0$. et la fonction f étant symétrique, il en est de même si $m_{31} \geq 0$.
2. Supposons maintenant que $m_{21} < 0$ et $m_{31} < 0$.

$$\text{Si} \quad m_{32} \geq 0, \qquad \text{alors} \quad \frac{|b_2|^2}{|b_1|^2} + 2m_{32}\,\frac{|b_2^*|^2}{|b_1|^2} \geq 1,$$

$$\text{et} \quad A(1,1) \geq \overbrace{2 + 2m_{21} + 2m_{31}}^{\geq 0} + 2m_{31}m_{21} \; > \; 0.$$

On déduit de ce qui précède:

$$A(1,1) < 0 \longrightarrow m_{21} < 0 \text{ et } m_{31} < 0 \text{ et } m_{32} < 0. \tag{7.14}$$

La condition nécessaire (7.14) et les relations (7.13) et (7.13) montrent alors

$$A(1,-1) < 0 \longrightarrow m_{21} > 0 \text{ et } m_{31} < 0 \text{ et } m_{32} > 0. \tag{7.15}$$

$$A(-1,1) < 0 \longrightarrow m_{21} > 0 \text{ et } m_{31} > 0 \text{ et } m_{32} < 0. \tag{7.16}$$

$$A(-1,-1) < 0 \longrightarrow m_{21} < 0 \text{ et } m_{31} > 0 \text{ et } m_{32} > 0. \tag{7.17}$$

Le vecteur $b_3 + b_2$ est plus court que b_3 si $A(0,1) < 0$.
Or, $A(0,1) = \dfrac{|b_2|^2}{|b_1|^2} + 2m_{32}\dfrac{|b_2^*|^2}{|b_1|^2} + 2m_{21}m_{31}$.

$$\text{Mais si} \quad m_{32} \geq 0, \quad \text{alors } \frac{|b_2|^2}{|b_1|^2} + 2m_{32}\frac{|b_2^*|^2}{|b_1|^2} \geq 1.$$

Et donc $\quad \forall (m_{21}, m_{31}) \in [-\frac{1}{2}, \frac{1}{2}]^2, \quad A(0,1) \geq 0.$

$$\text{Par ailleurs,} \quad \forall m_{32} \in [-\frac{1}{2}, \frac{1}{2}], \quad \frac{|b_2|^2}{|b_1|^2} + 2m_{32}\frac{|b_2^*|^2}{|b_1|^2} \geq 0.$$

Finalement l'inégalité $A(0,1) < 0$ implique $m_{21}m_{31} < 0$. On en déduit:

$$A(0,1) < 0 \quad \longrightarrow \left\{ \begin{array}{ll} m_{21} > 0\,,\ m_{31} < 0\,,\ m_{32} < 0 & (7.18.1) \\ \text{ou} & \\ m_{21} < 0\,,\ m_{31} > 0\,,\ m_{32} < 0 & (7.18.2) \end{array} \right. \tag{7.18}$$

La condition nécessaire (7.18) et la relation (7.13) montrent

$$A(0,-1) < 0 \quad \longrightarrow \left\{ \begin{array}{ll} m_{21} > 0\,,\ m_{31} > 0\,,\ m_{32} > 0 & (7.19.1) \\ \text{ou} & \\ m_{21} < 0\,,\ m_{31} < 0\,,\ m_{32} > 0 & (7.19.2) \end{array} \right. \tag{7.19}$$

Pour terminer la preuve, on remarque que les conditions (7.14), (7.15), (7.16), (7.17), (7.18.1), (7.18.2), (7.19.1), (7.19.2) s'excluent deux à deux. Par exemple, (b_1, b_2, b_3) est $t-$quasi-réduite et propre et les m_{ij} sont tous strictement positifs, seul le vecteur $b_3 - b_2$ est susceptible d'être plus court que b_3.

$\square$

7.5.2 Preuve du Lemme 12 (Encadrements des m_{ij})

Rappel: Si une base (b_1, b_2, b_3) est $t-$qr.II, b_2 et b_3 seront échangés, puis l'algorithme translatera b_3 parallèlement à b_1, afin de le rendre le plus court possible. Plus précisément, avec $m'_{31} := m_{31} - \lfloor \hat{m}_{31} \rceil$, le vecteur b'_3 ainsi obtenu vérifie $b'_3 := m'_{31} b^*_1 + m_{32} b^*_2 + b^*_3$. Comme la quasi-réduction est de type II, $|b'_3| < |b_1|$ et donc les vecteurs b_1 et b'_3 seront échangés. En résumé,

$$m = \begin{array}{c} \\ b_1 \\ b_2 \\ b_3 \end{array} \!\!\begin{array}{c} \begin{array}{ccc} b^*_1 & b^*_2 & b^*_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \end{array} \xrightarrow[\text{de tests: (-,-)}]{\text{en 2 itérations}} \hat{m} = \begin{array}{c} \\ b'_3 \\ b_1 \\ b_2 \end{array} \!\!\begin{array}{c} \begin{array}{ccc} \hat{b}^*_1 & \hat{b}^*_2 & \hat{b}^*_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ \hat{m}_{21} & 1 & 0 \\ \hat{m}_{31} & \hat{m}_{32} & 1 \end{pmatrix} \end{array}$$

Encadrement de $\hat{m}_{21}$

Par définition,

$$\hat{m}_{21} = \frac{(b_1, b'_3)}{|b'_3|^2} = \frac{m'_{31} |b^*_1|^2}{|b'_3|^2}. \tag{7.20}$$

Or, le vecteur b'_3 s'exprime dans la base (b^*) et sa norme vérifie

$$|b'_3|^2 = {m'_{31}}^2 |b^*_1|^2 + {m_{32}}^2 |b^*_2|^2 + |b^*_3|^2.$$

Et comme (b_1, b_2, b_3) est $t-$qr., les relations (7.10.2) et (7.10.1) montrent l'inégalité

$$|b'_3|^2 > \frac{1}{t^2} |b^*_2|^2 > \frac{3}{4} \frac{1}{t^2} |b^*_1|^2. \tag{7.21}$$

La propreté et les relations (7.20) et (7.21) permettent alors de conclure:

$$|\hat{m}_{21}| \leq \frac{2}{3}\, t^2. \tag{7.22}$$

Donc pour $1 < t < \dfrac{3}{2}$, on a bien $|\hat{m}_{21}| < \dfrac{3}{2}$ et $\lfloor \hat{m}_{21} \rceil \in \{-1, 0, 1\}$.

Une étude plus précise montrerait que le résultat reste valable pour $1 < t < \frac{3}{2}\sqrt{3}$.

Encadrement de $\hat{m}_{31}$

Par définition, $\quad \hat{m}_{31} = \dfrac{(b_2, b'_3)}{|b'_3|^2} = \dfrac{m_{21} m'_{31} |b^*_1|^2 + m_{32} |b^*_2|^2}{|b'_3|^2}.$

D'après la définition de $\hat{m}_{21}$ (7.20) et la relation (7.21),

$$|\hat{m}_{31}| < |m_{21}||\hat{m}_{21}| + |m_{32}| t^2.$$

Et grâce à la propreté et la majoration de $|\hat{m}_{21}|$ en (7.22), $\quad |\hat{m}_{31}| < \dfrac{5}{6}\, t^2.$

Finalement lorsque $1 < t^2 < \dfrac{9}{5}$, on a $|\hat{m}_{31}| < \dfrac{3}{2}$ et $\lfloor \hat{m}_{31} \rceil \in \{-1, 0, 1\}$.

Une étude plus précise montre que le résultat reste valable pour $1 < t^2 < \sqrt{10} - 1$.

Encadrement de $\hat{m}_{32}$

Par définition, $\hat{m}_{32} = \dfrac{(b_2, (b_1)_{\perp b'_3})}{\left|(b_1)_{\perp b'_3}\right|^2} = \dfrac{(b_2, b_1 - \hat{m}_{21} b'_3)}{\left|(b_1)_{\perp b'_3}\right|^2}.$

En exprimant les vecteurs b_2 et b'_3 dans la base orthogonale (b^*) comme dans le rappel du début de la preuve, on obtient

$$\hat{m}_{32} = \frac{m_{21}(1 - \hat{m}_{21} m'_{31})|b_1|^2 - \hat{m}_{21} m_{32} |b_2^*|^2}{|(b_1)_{\perp b'_3}|^2}. \tag{7.23}$$

Or, $|(b_1)_{\perp b'_3}|\,|b'_3| = |(b'_3)_{\perp b_1}|\,|b_1|$ et $(b'_3)_{\perp b_1} = (b_3)_{\perp b_1}$,. D'où,

$$|(b_1)_{\perp b'_3}|^2 = \frac{\left(m_{32}^2 |b_2^*|^2 + |b_3^*|^2\right)|b_1|^2}{|b'_3|^2}. \tag{7.24}$$

Les relations (7.20), (7.23) et (7.24) montrent après simplification

$$\hat{m}_{32} = m_{21} - m'_{31} m_{32} \frac{|b_2^*|^2}{m_{32}^2 |b_2^*|^2 + |b_3^*|^2}.$$

Pour finir, comme (b_1, b_2, b_3) est $t-$quasi-réduite, $\hat{m}_{32} \leq \dfrac{1}{2} + \dfrac{t^2}{4}.$

Pour $1 < t < 2$, on a $|\hat{m}_{32}| \leq \dfrac{3}{2}$ et $\lfloor \hat{m}_{32} \rceil \in \{-1, 0, 1\}$. □

7.5.3 Preuve du Lemme 13

Nous partons ici de la base (b'_3, b_1, b_2) (obtenue nécessairement en 2 itérations de tests négatifs à partir de la base (b_1, b_2, b_3), supposée $t-$qr.II) et nous nous intéresserons à l'action de l'algorithme sur cette base.
Le vecteur b_1 sera échangé par $b'_1 := b_1 - \lfloor \hat{m}_{21} \rceil b'_3$. Le test qui s'en suit est positif. (En effet (b'_3, b_1) était propre et la propreté de (b_1, b'_3) n'exige au pire qu'une translation entière de b_1 selon $\pm b'_3$, car d'après le Lemme 12, $\lfloor \hat{m}_{21} \rceil \in \{-1, 0, 1\}$.)
En conséquence, le deuxième orthogonalisé $\hat{b}_2^* = (b_1)_{\perp b'_3}$ ne sera pas modifié et sa norme est donnée par la relation (7.24).
Comme (b_1, b_2, b_3) est $t-$qr.II, en particulier la deuxième boîte $(b_2^*, m_{32} b_2^* + b_3^*)$ est non réduite. D'où:

$$|(b'_1)_{\perp b'_3}|^2 \leq \frac{|b_2^*|^2 |b_1|^2}{|b'_3|^2}. \tag{7.25}$$

A ce stade, le vecteur b_2 sera translaté parallèlement à b'_1 (afin de rendre propre la base $\left((b'_1)_{\perp b'_3}, (b_2)_{\perp b'_3}\right)$). Appelons b'_2 le vecteur obtenu après la translation.

Le test qui suivra concerne le rapport $\dfrac{|(b'_2)_{\perp b'_3}|^2}{|(b'_1)_{\perp b'_3}|^2}$, que nous allons minorer dans la suite.

Nous avons: $b'_2 = b_2 - \lfloor \hat{m}_{32} \rceil b'_1 = b_2 - \lfloor \hat{m}_{32} \rceil b_1 + \lfloor \hat{m}_{32} \rceil \lfloor \hat{m}_{21} \rceil b'_3$
Le vecteur b'_2 est donc de la forme $b_2 - nb_1 + mb'_3$.

$$\text{D'où}: \quad (b'_2)_{\perp b'_3} = (b_2 - nb_1)_{\perp b'_3} = (b_2 - nb_1) - \frac{(b_2 - nb_1, b'_3)}{|b'_3|^2} b'_3 \qquad (n \text{ désigne } \lfloor \hat{m}_{32} \rceil)$$

$$\text{et}: \quad \left|(b'_2)_{\perp b'_3}\right|^2 = |b_2 - nb_1|^2 - \frac{(b_2 - nb_1, b'_3)^2}{|b'_3|^2}$$

Les vecteurs b_1, b_2 et b'_3 s'expriment dans la base orthogonale (b^*), et après simplification,

$$\left|(b_2 - nb_1)_{\perp b'_3}\right|^2 = \frac{|b^*_2|^2 |b^*_1|^2}{|b'_3|^2} \left(\frac{|b^*_3|^2}{|b_1|^2} + (m_{21} - n)^2 \frac{m_{32}^2 |b^*_2|^2 + |b^*_3|^2}{|b^*_2|^2} + {m'}^2_{31} - 2(m_{21} - n) m'_{31} m_{32} \right).$$

Etant donnée la majoration obtenue pour $|(b'_1)_{\perp b'_3}|^2$ en (7.25),

$$\frac{\left|(b_2 - nb_1)_{\perp b'_3}\right|^2}{|(b'_1)_{\perp b'_3}|^2} \geq \frac{|b^*_3|^2}{|b_1|^2} + (m_{21} - n)^2 \frac{m_{32}^2 |b^*_2|^2 + |b^*_3|^2}{|b^*_2|^2} + {m'}^2_{31} - 2(m_{21} - n) m'_{31} m_{32}.$$

Mais la base (b_1, b_2, b_3) est $t-$quasi-réduite. D'après les relation(7.10),

$$\frac{\left|(b_2 - nb_1)_{\perp b'_3}\right|^2}{|(b'_1)_{\perp b'_3}|^2} \geq (1 - m_{21}^2) \left(\frac{1}{t^2} - m_{32}^2 \right) + \frac{(m_{21} - n)^2}{t^2} + {m'}^2_{31} - 2(m_{21} - n) m'_{31} m_{32}$$

Après simplification et comme la base (b_1, b_2, b_3) est propre,

$$\frac{\left|(b_2 - nb_1)_{\perp b'_3}\right|^2}{|(b'_1)_{\perp b'_3}|^2} \geq (1+n^2-2n\, m_{21}) \left(\frac{1}{t^2} - m_{32}^2 \right) + \big(m_{32}(m_{21} - n) - m'_{31}\big)^2 \geq (1+n^2-|n|) \left(\frac{1}{t^2} - \frac{1}{4} \right)$$

Dans la relation précédente, le minorant est clairement une fonction croissante de n. Aussi,

$$\frac{\left|(b_2 - nb_1)_{\perp b'_3}\right|^2}{|(b'_1)_{\perp b'_3}|^2} \geq \frac{1}{t^2} - \frac{1}{4} \qquad \text{valable pour tout } n \in \mathbb{Z}.$$

Finalement, pour $1 < t < 2$, l'algorithme débouche bien au pire sur une situation de $t'-$quasi-réduction avec:

$$\frac{1}{t'^2} = \frac{1}{t^2} - \frac{1}{4}$$

□

7.6 Cas de la dimension n fixée

Les raisonnements des dimensions 2 et 3 se généralisent à une dimension n fixée. Ils permettent de montrer que là encore le nombre d'itérations de l'algorithme LLL(1) est un

polynôme en $\log M$. Mais les bornes obtenues par cette modélisation se dégradent de façon exponentielle avec la dimension.
La difficulté principale est la suivante. Il existe à priori de nombreuses bases réduites au sens de 1-Lovász dans un réseau donné. En particulier, une estimation grossière du nombre de vecteurs du réseau "candidats" à être le premier vecteur d'une base réduite (Lemme 14) laisse supposer que ce nombre croît exponentiellement avec la dimension et le raisonnement qui suit est basé sur ce nombre.
Dans la suite, le Lemme 14 joue le même rôle que le Lemme 11 pour la dimension 3. Ce Lemme est une version plus courte d'un Lemme utilisé par Kannan [Kan83] pour mettre en évidence un algorithme qui construit une base réduite au sens de Korkine-Zolotarev avec une complexité en $O(n^n)$. La borne polynomiale en $\log M$ pour le nombre d'itérations de l'algorithme LLL(1) à dimension n fixée s'en déduit par réccurence sur n.

Lemme 14. *Soient t un réel vérifiant $t > 1$ et $(b) := (b_1, \ldots, b_n)$ une base $t-$quasi-réduite et propre d'un réseau $\mathcal{L}$. Le nombre de vecteurs du réseau $\mathcal{L}$, de normes strictement inférieures à $|b_1|$, est majoré par une constante qui ne dépend que de la dimension n et du paramètre t. Elle est notée $\alpha(n,t)$.*

Preuve. Tout vecteur $v = \sum_{i=1}^n v_i b_i$ du réseau s'exprime dans la base orthogonale (b^*) comme en (7.1) et si il est plus court que b_1, il vérifie:

$$|v|^2 = \sum_{i=1}^{n} \left(\sum_{j=i}^{n} v_j m_{j,i} \right)^2 |b_i^*|^2 < |b_1|^2.$$

La recherche d'un vecteur plus court que $|b_1|$ revient donc à chercher un vecteur entier $(v_1, \ldots, v_n)$ vérifiant

$$0 \le |v_n| < \frac{|b_1|}{|b_n^*|}$$

et

$$\left(\sum_{j=i}^{n} v_j m_{j,i} \right)^2 |b_i^*|^2 < |b_1|^2 - \sum_{k=i+1}^{n} \left(\sum_{j=k}^{n} v_j m_{j,i} \right)^2 |b_k^*|^2 \quad \text{pour } i = 1, \ldots, n-1.$$

Comme $|\sum_{j=i}^n v_j m_{j,i}| < |b_1|/|b_i^*|$, le nombre de valeurs possibles pour v_i lorsque $v_{i+1}, \ldots, v_n$ sont fixés est au plus $\lfloor 2|b_1|/|b_i^*| \rfloor + 1$. Donc, au plus

$$\prod_{i=2}^{n} \left(\lfloor 2\frac{|b_1|}{|b_i^*|} \rfloor + 1 \right)$$

vecteurs $(v_1, \ldots, v_n)$ sont susceptibles d'être plus courts que b_1. Mais la base $(b_1, \ldots, b_n)$ est $t-$quasi-réduite. La Définition (6) entraîne alors

$$\frac{|b_i^*|^2}{|b_{i+1}^*|^2} \le \frac{4}{3} \quad \text{pour } i = 1, \ldots, n-2$$

et

$$\frac{|b_{n-1}^*|^2}{|b_n^*|^2} \le \frac{4t^2}{4-t^2}.$$

Ces dernières relations entraînent après simplification, que le nombre de vecteurs $(v_1, \ldots, v_n)$ susceptibles d'être plus courts que b_1 est au plus

$$\left(\frac{6t}{\sqrt{4-t^2}}\right)^n \left(\frac{2}{\sqrt{3}}\right)^{\frac{(n-1)(n-2)}{2}}.$$

□

Remarque. Il est légitime de penser que la majoration donnée par le Lemme précédent sur $\alpha(n,t)$ pourrait être considérablement amélioriée. Une approche (inspirée du cas de la dimension 3) serait de montrer que les signes des $m_{i,j}$ déterminent le(s) vecteur(s) du réseau de norme strictement inférieure à $|b_1|$, lorsque b_1 est le premier vecteur d'une base t-quasi–réduite et propre du réseau.

Théorème 13. *Lorsque la dimension n est fixée, le nombre maximum d'itérations de l'algorithme LLL(1) est polynomial en la taille $\log M$, de la base d'entrée.*

Preuve. Elle se fait par récurrence sur n. La propriété est vraie en dimension 2 et 3 comme le montrent les parties précédentes.
Supposons que la propriété est vraie au rang $n-1$. Soit t un réel vérifiant $t > 1$. Les phases non finales qui se produisent pendant une exécution de l'algorithme peuvent être classées en trois groupes:

1. Les phases non finales qui produisent une base $(b_1, \ldots, b_n)$ non $t-$quasi-réduite. Si d désigne le produit $\prod_1^{n-1} |b_i^*|^2$, une telle phase se termine par un test négatif dans la dernière boîte entraînant une décroissance de la quantité d d'un facteur $1/t^2$. La quantité d et le nombre maximum k de ces phases vérifient alors respectivement
$$1 \leq d \leq \frac{1}{(t^2)^k} M^{2(n-1)} \quad \text{et} \quad k \leq (n-1)\log_t M.$$
2. Les phases qui finissent sur des situations $t-$qr.II. Mais d'après le Lemme 14, le nombre total de situation $t-$qr.II, se produisant pendant une exécution de l'algorithme est une constante qui ne dépend que de la dimension n et du paramètre t. $(\alpha(n,t)+1)$.
3. Les phases qui finissent sur des situations $t-$qr.I. Au plus, un nombre polynomial de phases de ce type peuvent se succéder, car tant que les phases débouchent sur des $t-$qr.I successives, l'algorithme travaille sur la base
$$\left((b_2)_{\perp\langle b_1\rangle}, \ldots, (b_n)_{\perp\langle b_1\rangle}\right)$$
qui a $n-1$ vecteurs (et qui est d'ailleurs également $t-$qr.I) et l'hypothèse de récurrence entraîne que le nombre de phases pour réduire une base en dimension $n-1$ est majoré par un polynôme en $\log M$. On le note $\mathcal{P}_{n-1}(\log M)$.

Donc le nombre maximum de phases durant une exécution de l'algorithme LLL(1) en dimension n est bien majoré par un polynôme $\mathcal{P}_n$ en $\log M$ avec

$$\mathcal{P}_n(\log M) = ((n-1)\log_t M + \alpha(n,t) + 1)\,\mathcal{P}_{n-1}(\log M).$$

Les deux remarques suivantes permettent alors de conclure.

1. Toutes les bases intermédiaires obtenues par l'algorithme sont de tailles polynomiales par rapport à la taille de la base initiale.

2. Une phase de l'algorithme consiste en la réduction de la base $(b_1, \ldots, b_{n-1})$, qui se fait d'après l'hypothèse de récurrence en un nombre ploynomial d'itérations, suivie d'une translation du vecteur b_n parallèlement au vecteur b_{n-1}. Le nombre d'itération d'une phase est donc polynomial en $\log M$.

□

Chapitre 8

Étude expérimentale comparative

Dans ce chapitre nous décrivons une série d'expérimentations en relation avec les résultats énoncés dans cette thèse. Ce chapitre décrit deux types de résultats expérimentaux:

- Les premiers confirment les études que nous avons menées dans le modèle uniforme.
- Les seconds concernent le modèle "sac–à–dos" et montrent que le comportement relatif des algorithmes dans ce modèle est proche de ce qu'il est dans le modèle uniforme, ce que l'on ne savait pas a priori.

L'étude expérimentale sur le modèle uniforme confirme de façon très nette les résultats de l'étude comparative du Chapitre 6. Elle montre d'ailleurs que le régime asymptotique s'installe très rapidement: en dimension 20 déjà, les nouveaux algorithmes finissent presque toujours après le nombre minimal d'itérations. Mais ces expériences n'apportent pas seulement la simple confirmation des résultats des chapitres précédents: ils mettent l'accent sur certaines défaillances de l'algorithme LLL classique. L'étude des bases aléatoires de ce modèle (Chapitre 3) montrait qu'elles ont en moyenne de très bonnes propriétés euclidiennes. Nous montrons dans le Chapitre 6 qu'une base pleine de ce modèle une fois rendue propre est quasi-sûrement réduite au sens des nouvelles réductions que nous avons introduites et qu'avec une probabilité non–négligeable, elle n'est pas réduite au sens de Siegel et Lovász. On pouvait a priori envisager plusieurs éventualités:

1. Bien que prouvé strictement supérieur, le nombre moyen d'itérations de l'algorithme LLL pouvait rester très proche du nombre minimal. (on rappelle que, pour un réseau plein de dimension n, ce nombre minimal est $n-1$.)
2. La qualité de la base de sortie de l'algorithme LLL pouvait être en moyenne bien supérieure à celle des nouveaux algorithmes de réduction.

Les expérimentations montrent qu'aucune de ces deux éventualités n'est réalisée: sur les bases aléatoires pleines du modèle uniforme, *l'algorithme LLL classique tourne longtemps pour rien*. Plus précisément, le nombre d'itérations est en moyenne effectivement bien supérieur au nombre minimal pour une qualité de sortie du même ordre que la base initiale rendue propre. Parfois, cette qualité de sortie est même moins bonne. C'est donc clairement une défaillance de l'algorithme LLL qui est ici exhibée expérimentalement. En revanche, les nouveaux algorithmes introduits s'arrêtent après un nombre minimal d'itérations (nécessaires pour vérifier la bonne qualité de la base initiale).

Les expérimentations du second type sont effectuées dans le modèle sac–à–dos. Nous commençons par une description succincte des réseaux de ce modèles et montrons que ce modèle correspond aux réseaux intervenant dans la plupart des applications pratiques (approximations diophantiennes, factorisation de polynômes à coefficients entiers cryptanalyse des systèmes d'Okamoto et de Merkle–Hellmann). Puis on fournit une étude comparative expérimentale du nombre d'itérations et de la qualité de la base de sortie des algorithmes de réduction sur les réseaux aléatoires de ce modèle. Ces expériences montrent que tout comme dans le modèle uniforme, les nouveaux algorithmes sont nettement plus rapides. Par ailleurs, la qualité des bases de sortie des nouveaux algorithmes est certes légèrement moins bonne, mais reste comparable à celle de l'algorithme LLL. De nouveau, on peut envisager deux possibilités:

1. C'est la légère supériorité des bases de sortie de l'algorithme LLL qui justifie le nombre d'itérations toujours nettement plus élevé de l'algorithme LLL par rapport aux nouveaux algorithmes . Dans ce cas, on peut alors considérer ces expérimentations comme un pas vers une quantification *fine* des problèmes algorithmiques dans les réseaux.
2. Tout comme dans le modèle uniforme le nombre élevé d'itérations de l'algorithme LLL témoigne d'une certaine défaillance de l'algorithme. Notre étude peut alors être vue comme le point de départ pour la mise au point d'une stratégie algorithmique utilisant toute la diversité des outils de réduction (LLL, ses version en flottant ou en parallèle, ainsi que les nouveaux algorithmes introduits dans cette thèse)

L'étude expérimentale menée dans ce chapitre donne également d'autres indications en aparté intéressantes tant pour une étude théorique que pour la mise au point de stratégies algorithmiques. Pour finir, je précise que ce chapitre ne fait que suggérer de telles stratégies. Une stratégie précise dépend sans aucun doute de l'application précise envisagée et ceci n'est pas le propos de cette thèse.

8.1 Description générale des expériences

Les expériences ont été effectuées sur station Sun UltraSPARC à 168 MHz et à l'aide du logiciel MapleV.4.

On rappelle que tous les algorithmes de réduction qui font l'objet de cette thèse sont des algorithmes d'approximation et comportent donc des paramètres d'approximation. On rappelle également que le paramètre t de l'algorithme LLL classique et le paramètre s des nouveaux algorithmes de réductions sont reliés par la relation

$$s^2 = \frac{1}{t^2} - \frac{1}{4}, \quad \text{avec } t > 1 \text{ ou encore } s < \sqrt{3/4}. \tag{8.1}$$

La valeur $4/3$ pour t^2 qui correspond à $1/2$ pour s^2 est la valeur habituellement choisie. Le tableau suivant décrit le choix des paramètres pour les algorithmes de réduction dans toutes les expériences qui vont suivre.

t-Lovász	$t^2 = 4/3$		
(s,k)-Schmidt	$s^2 = 1/2$	$k = 1$	
(s,γ)-Gram	$s^2 = 1/2$		$\gamma = 3/4$
(s,k,γ)-Gram–Schmidt	$s^2 = 1/2$	$k = 1$	$\gamma = 3/4$
$(s,1)$-Schmidt optimale	$s^2 = 3/4$	$k = 1$	

La dernière ligne du tableau précédent correspond à un choix un peu différent des autres. Le choix de la valeur $s^2 = 3/4$ correspond en effet à celle de la valeur 1 pour le paramètre t. Le Théorème 3 du Chapitre 2 ne permet pas de montrer que le nombre d'itérations de l'algorithme est polynomial dans le pire des cas. Nous avons montré dans le Chapitre 7 que la complexité de l'algorithme LLL($t = 1$) reste polynomial en la taille de la base lorsque la dimension de la base est fixée. Les expériences qui suivent semblent montrer que l'algorithme optimal de Schmidt–réduction (avec le choix optimal de la valeur $3/4$ pour s^2) reste en moyenne plus rapide que l'algorithme LLL usuel (avec $t^2 = 4/3$). Il est clair que pour la mise au point d'une stratégie algorithmique précise, outre le choix de ou des algorithmes de réduction, le bon choix des paramètres d'approximation est sans doute très important.
Lors de chaque expérience, nous avons mesuré d'une part le nombre d'itérations de chaque algorithme et d'autre part la qualité de la base de sortie. Nos deux critères d'appréciation de la qualité sont les suivants:

1. La racine n-ème du défaut d'orthogonalité
2. Le rapport entre les normes du vecteur de sortie le plus court d'un algorithme donné et la norme du vecteur de sortie le plus court de l'algorithme LLL.

Les résultats des expériences sont donnés sous forme de tableaux et, lors de l'étude du modèle sac–à–dos, des courbes illustrent ces résultats. Les résultats donnés sont toujours des valeurs moyennes. Pour les expériences sur le modèle uniforme, chaque moyenne est calculée à partir de 10 expériences. Lors des expériences sur le modèle sac–à–dos chaque moyenne est calculée en général à partir de 20 expériences. (pour les réseaux de dimension strictement supérieures à 60 et de densité strictement inférieures à 1 les moyennes sont données à partir de 5 expériences seulement)

8.2 Le modèle uniforme

8.2.1 Génération des bases aléatoires

Pour générer les bases aléatoires du modèle uniforme, il faut pouvoir tirer uniformément et indépendamment des vecteurs à coordonnées entières à l'intérieur d'une boule de rayon M. La méthode *naïve* est la méthode de rejet: on choisit un vecteur uniformément dans l'hyper–cube centré sur l'origine et dont les arêtes sont de longueur $2M$ (le plus petit hyper-cube

contenant la sphère), et on garde le vecteur si il est à l'intérieur de la sphère (on le rejette, sinon). Mais comme le rapport entre le volume $\omega_n(M) = M^n \pi^{n/2}/\Gamma(1+n/2)$ de l'hyper–sphère et le volume $v_n(M) = (2M)^n$ de l'hyper–cube tend exponentiellement vers 0 avec la dimension, cette méthode naïve n'est pas raisonnable même en petite dimension ($n \geq 5$).
Pour générer un vecteur uniformément dans la boule, on utilise la méthode suggérée par Knuth [Knu81]: si l'on choisit n variables aléatoires indépendantes $x_1, \ldots, x_n$ suivant une loi normale de moyenne 0 et de variance 1, alors le point

$$(x_1/r, \ldots, x_n/r), \qquad \text{avec} \qquad r := \sqrt{x_1^2 + \ldots x_n^2}$$

est régulièrement réparti *sur* la sphère unité. On montre en effet que la fonction de répartition du point $(x_1, \ldots, x_n)$ ne dépend que de r. Pour en déduire ensuite de façon uniforme, un point *à l'intérieur* de la sphère, on utilise une idée due à Brent: si $(y_1, \ldots, y_n)$ est uniformément réparti sur la sphère et U est une variable aléatoire de loi uniforme sur $[0, 1]$, alors le point $U^{1/n}(y_1, \ldots, y_n)$ est uniformément distribué à l'intérieur de la sphère.

8.2.2 Résultats des expériences

Le tableau 8.1 regroupe tous les moyennes pour les nombres d'itérations et la qualité des bases de sortie des différents algorithmes de réduction. Les résultats sont en fonction de la dimension n du réseau. Nous avons considéré les cas où la taille de la base initiale est respectivement en $\mathrm{O}(2^{n/2}), \mathrm{O}(2^n), \mathrm{O}(2^{3n/2})$. Par ailleurs la première ligne du tableau du milieu concerne la base aléatoire initiale (entrée des algorithmes). Cette ligne confirme les résultats du Chapitre 3 concernant la bonne qualité des bases du modèle uniforme. Le tableau 8.1 montre clairement la rapidité des nouveaux algorithmes comparés à l'algorithme LLL. On remarque en particulier que la qualité de sortie est sensiblement la même dans tous les cas que celle de la base d'entrée (qui est déjà de bonne qualité), mais les nouveaux algorithmes détectent cette qualité satisfaisante beaucoup plus rapidement.

Les expériences effectuées montrent par ailleurs[1] que le vecteur le plus court de la base de sortie n'est pas souvent en première position, alors que la majoration théorique la plus fine sur les normes des vecteurs de sortie fait intervenir le premier vecteur de la base. On rappelle que quand s et t sont reliés par la relation (8.1), la norme du i-ème vecteur d'une base t-Lovász réduite vérifie la majoration

$$|b_i| < (1/s)^{n-1}\Lambda_i, \qquad \text{où } \Lambda_i \text{ est le } i\text{-ème minimum du réseau.}$$

Ce phénomène n'est pas très surprenant, car les réseaux aléatoires sont assez “réguliers” (les minima successifs sont proches les uns des autres).

Pour terminer, il faut noter que le rapport précis entre les nombres d'itérations des différents algorithmes de réduction n'est pas le même que le rapport entre leur temps d'exécution. En effet, lorsque les entiers manipulés sont “grands”, l'orthogonalisation de Gram–Schmidt qui est la phase d'initialisation commune de tous ces algorithmes de réduction est coûteuse. Si on compare les temps d'exécution, en dehors de la phase d'initialisation, on retrouve l'extrême rapidité des nouveaux algorithmes par rapport à l'algorithme classique: les nouveaux algorithmes sont plus rapides encore que ne le montre le rapport de leur nombre d'itérations à celui de l'algorithme LLL.

1. Cette constatation n'est pas à chercher dans le tableau 8.1.

Nombre d'itérations

dim	10	20	30	40	50	10	20	30	40	10	20	30
taille	$O(2^{n/2})$					$O(2^n)$				$O(2^{3n/2})$		
LLL	62.8	141.9	135.8	178.2	177.7	59.8	132.7	149.8	191.4	78	129	150
Sch	9.2	19	29	39	49	9.8	19	29	39	10.8	19	29
Gr	9.4	19	29	39	49	9.9	19	29	39	11.2	19	29
Gr–Sch	9.4	19	29	39	49	10.2	19	29	39	13.3	19	29
Sch opt	19.8	19.6	29.4	39	49	22.9	19.8	29	39	32.9	19.9	29.3

Racine n-ème du défaut d'orthogonalité

dim	10	20	30	40	50	10	20	30	40	10	20	30
taille	$O(2^{n/2})$					$O(2^n)$				$O(2^{3n/2})$		
aléa	1.65	1.71	1.63	1.67	1.66	1.69	1.67	1.64	1.68	1.84	1.64	1.68
LLL	1.20	1.54	1.71	1.82	1.87	1.19	1.52	1.73	1.86	1.58	1.64	1.75
Sch	1.42	1.69	1.71	1.77	1.81	1.44	1.66	1.72	1.82	1.57	1.64	1.75
Gr	1.42	1.69	1.71	1.77	1.81	1.43	1.66	1.72	1.82	1.57	1.64	1.75
Gr–Sch	1.42	1.69	1.71	1.77	1.81	1.43	1.66	1.72	1.82	1.53	1.64	1.75
Sch opt	1.33	1.69	1.71	1.77	1.81	1.27	1.66	1.72	1.82	1.36	1.63	1.75

Comparaison des normes des vecteurs les plus courts par rapport à LLL

dim	10	20	30	40	50	10	20	30	40	10	20	30
taille	$O(2^{n/2})$					$O(2^n)$				$O(2^{3n/2})$		
Sch	1.04	1.07	1.00	1.00	1.00	1.04	1.08	1.02	0.99	1.08	1.07	1.02
Gr	1.04	1.07	1.00	1.00	1.00	1.06	1.08	1.02	0.99	1.04	1.07	1.02
Gr–Sch	1.00	1.07	1.00	1.00	1.00	1.06	1.08	1.02	0.99	1.04	1.07	1.02
Sch opt	1.04	1.07	1.00	1.00	1.00	1.01	1.08	1.02	0.99	1.04	1.06	1.02

TAB. 8.1 – *Comparaison entre le nombre d'itérations et la qualité de la base de sortie sur les réseaux aléatoires du* modèle uniforme.

8.3 Le modèle sac–à–dos

8.3.1 Introduction

Nous allons définir ici les réseaux du *modèle sac–à–dos*. Ces réseaux interviennent naturellement dans de nombreuses applications. Dans cette introduction quelques exemples d'applications sont détaillés, puis nous décrivons le cadre précis des expériences.

Approximations diophantiennes simultanées [Lag82]

Soit un ensemble arbitraire de nombre réels $\alpha_1, \ldots \alpha_n$. Dans de nombreuses applications, il est intéressant d'en trouver des approximations rationnelles $p_1/q, \ldots, p_n/q$ toutes de même dénominateur (entier) q. Il s'agit là du problème *d'approximation diophantienne simultanée*. Plus précisément, étant donnés des réels positifs $\alpha_1, \ldots \alpha_n, \varepsilon$ et un entier Q, on voudrait trouver lorsqu'ils existent des entiers $p_1, \ldots, p_n, q$ tels que $0 < q < Q$ et

$$\text{pour tout } i \in \{1, \ldots, n\}, \quad |\alpha_i - \frac{p_i}{q}| \leq \frac{\varepsilon}{q}.$$

Le changement de variable $p'_i = -p_i$ dans les inégalités $|q\alpha_i - p_i|$ permet de formuler le problème précédent sous forme matricielle,

$$-\varepsilon \leq Ax \leq \epsilon,$$

où x est le vecteur formé des entiers recherchés $p'_1, \ldots, p'_n, q$ et

$$A = \begin{pmatrix} 1 & 0 & 0 & . & . & . & \alpha_1 \\ 0 & 1 & 0 & . & . & . & \alpha_2 \\ 0 & 0 & 1 & . & . & . & \alpha_3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & 1 & \alpha_n \\ 0 & . & . & . & . & 0 & \varepsilon/Q \end{pmatrix}.$$

Le théorème fondamental de Dirichlet donne une condition suffisante pour l'existence des entiers $p'_1, \ldots, p'_n, q$: si les réels positifs ε et Q vérifient $Q \geq \varepsilon^{-n}$ alors le problème a des solutions. (Cette condition découle directement du premier théorème de Minkowski.)
Pour trouver effectivement les entiers $p'_1, \ldots, p'_n, q$, on remarque que lorsque x est un vecteur de $\mathbb{Z}^{n+1}$, le vecteur Ax est un élément du réseau L engendré par les vecteurs colonnes de la matrice A. Le problème revient alors à trouver un vecteur non nul du réseau L dont la norme infinie est au plus ε. Comme pour tout vecteur x, le norme infinie de x est inférieure à sa norme euclidienne, il suffit de trouver un vecteur non nul du réseau dont la norme euclidienne est au plus ε. En utilisant les algorithmes de réduction, on peut trouver un vecteur assez court pour la norme euclidienne. En particulier, on montre dans le Chapitre 2 (Théorème 1) que la norme euclidienne du premier vecteur b_1 d'une base réduite au sens de s-Siegel (ou t-Lovász), $(s, 1)$-Schmidt, $(s, 1, \gamma)$-Gram–Schmidt vérifie $|b_1| \leq (1/s)^n \Lambda_1(L)$, où $\Lambda_1(L)$ est le premier minimum du réseau (la norme d'un vecteur le plus court).

Par ailleurs, d'après le théorème de Minkowski,

$$\Lambda_1(L) \leq \sqrt{n+1}d(L)^{1/(n+1)}, \quad \text{où} \quad d(L) \ \text{désigne le détérminant du réseau} \ L.$$

Or le déterminant du réseau L est égal à ε/Q. Ainsi, pour que le premier vecteur b_1 des bases réduites citées soit une solution du problème, il suffit que

$$\left(\frac{1}{s}\right)^n \sqrt{n+1}\left(\frac{\varepsilon}{Q}\right)^{\frac{1}{n+1}} \leq \varepsilon.$$

On montre que pour n assez grand, il suffit que $Q \geq \left(\frac{1}{s}\right)^{n^2/2} \varepsilon^{-n}$.

D'autres applications pratiques telles que la factorisation de polynômes à coefficients entiers [LLL82, Sch84] ou la recherche d'une relation linéaire "courte" entre p vecteurs de $\mathbb{Z}^n$ peuvent être également formulées sous forme de recherche de vecteur court dans un réseau de type semblable à celle formé par les vecteurs colonnes de la matrice A.

Une première application cryptographique: cassage des cryptosystèmes d'Okamoto[VGT88]

Il s'agit de la recherche des racines n-ème modulo ℓ. Soient deux entiers ℓ et $n \geq 2$. Étant donné deux points x_0 et y_0, un voisinage I de x_0, un voisinage J de y_0 qui contiennent respectivement un point x et une point y vérifiant $x^\ell \equiv y[\ell]$, on veut trouver x et y.
Plus précisément, on recherche trois entiers u_1, u_2, v assez petits, solutions de l'équation

$$(u_1 x_0 + u_2)^\ell \equiv y_0 + v \quad [\ell].$$

En développant l'équation précédente avec la formule du binôme et en posant

$$w_i = u_2^i u_1^{n-i}, \ \text{pour} \ 0 \leq i \leq n-1 \quad \text{et} \quad w_n = y_0 + v - u_2^n,$$

on obtient:

$$\sum_{i=0}^{n-1} \binom{n}{i} x_0^{n-i} w_i - w_n \equiv 0 \quad [\ell]$$

Les points $w = (w_0, \ldots, w_n)$ de $\mathbb{Z}^{n+1}$ vérifiant la relation précédente appartiennent à un réseau de dimension $n+1$, généré par les vecteurs lignes de la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & . & . & . & x_0^n \\ 0 & 1 & 0 & . & . & . & \binom{n}{1}x_0^{n-1} \\ 0 & 0 & 1 & . & . & . & \binom{n}{2}x_0^{n-2} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & 1 & \binom{n}{n-1}x_0 \\ 0 & . & . & . & . & 0 & \ell \end{pmatrix}.$$

Le problème revient alors à rechercher –en un sens à préciser– un point du réseau L assez proche du point $(0, \ldots, 0, y_0)$. Lorsque l'on dispose d'une base du réseau réduite au sens de t-Lovász, il existe deux algorithmes dus à Babai [Bab86] qui permettent de trouver un point du réseau assez proche d'un point donné de $\mathbb{Z}^{n+1}$. On peut utiliser les même heuristiques en partant d'une base réduite au sens de (s, γ)-Gram ou (s, k, γ)-Gram–Schmidt.

Une deuxième application cryptographique: cassage des cryptosystèmes de Merkle–Hellmann [LO83, SE91]

Voici un énoncé du célèbre problème du sac–à–dos:
Soient n entiers positifs a_i –les paquets – et un entier M –le sac –, trouver un n–uplet $(x_1, \ldots, x_n)$ de $\{0,1\}^n$ solution de l'équation

$$\sum_{i=1}^{n} a_i x_i = M. \tag{8.2}$$

Autrement dit, quels paquets doit–on mettre pour remplir exactement le sac, lorsque cela est possible? Les cryptosystèmes à clé publique basés sur la difficulté du problème précédent ont été attrayants en particulier par la grande simplicité du chiffrement. Dans ces cryptosystèmes, les a_i forment la clé publique et un texte clair $(e_1, \ldots, e_n)$, vecteur de $\{0,1\}^n$, est très simplement crypté, en utilisant la relation (8.2). L'entier M est alors le message crypté. La connaissance d'une certaine information sur la suite des a_i (clé secrète) permet de retrouver le texte clair. On définit la densité d du sac ou encore la densité des poids a_i par

$$d(a) = \frac{n}{\log(\max_i a_i)}.$$

L'injectivité nécessaire pour la fonction de chiffrement fait que seules les instances du problème sac–à–dos de densité inférieure à 1 ont un intérêt cryptographique. Par ailleurs, les instances les plus difficiles sont celles de densité proche de 1.
Il est clair que toute solution de l'équation (8.2) est un vecteur du réseau L engendré par les lignes de la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & . & . & . & -a_1 \\ 0 & 1 & 0 & . & . & . & -a_2 \\ 0 & 0 & 1 & . & . & . & -a_3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & 1 & -a_n \\ 0 & . & . & . & . & 0 & M \end{pmatrix}.$$

On montre que les réseaux engendrés par les vecteurs ligne de la matrice A, avec les a_i vérifiant $1 \leq a_i \leq B$ représentent bien les problèmes de sac–à–dos de densité $n/\log_2 B$, aussitôt que pour $n \to \infty$, il existe un réel positif c tel que $\log_2 B \sim cn$.
Le défaut d'orthogonalité des bases formées des vecteurs ligne de la matrice A se calcule aisément et, pour des bases de densité d et de cardinalité n, le défaut d'orthogonalité $\rho(b)$ vérifie la majoration

$$\rho(b) = \mathrm{O}(2^{\frac{n^2}{d}}). \tag{8.3}$$

De même en adaptant la preuve du Théorème 3 du Chapitre 2, on montre que le nombre maximum d'itérations K des algorithmes de réduction pour réduire une base d'un réseau sac–à–dos régulier, c'est à dire dont les minima sont "à peu près" de même taille, est majoré en fonction de la dimension n et de la densité d du réseau [2]:

$$K(n,d) = \mathrm{O}\left(\frac{n^2}{d\log t}\right).$$

8.3.2 Le modèle sac–à–dos choisi pour les expériences

Le modèle que nous avons choisi pour les expériences est celui utilisé par Schnorr et Euchner [SE91] pour résoudre de façon pratique une grande majorité des instances des problèmes sac–à–dos. On montre que pour presque toutes les instances du problème sac–à–dos de densité inférieure à 0.9408, la solution de (8.2) est un plus court vecteur du réseau engendré par les vecteurs ligne de la matrice suivante (réseau de dimension $n+1$ dont les éléments sont des vecteurs de $\mathbb{Z}^{n+2}$)

$$\begin{pmatrix} 2 & 0 & 0 & . & . & . & na_1 & 0 \\ 0 & 2 & 0 & . & . & . & na_2 & 0 \\ 0 & 0 & 2 & . & . & . & na_3 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & 2 & na_n & 0 \\ 1 & 1 & . & . & . & 1 & nM & 1 \end{pmatrix}.$$

Nous avons considéré des bases aléatoires de densités respectives $2, 1, 2/3, 1/2$ (Seules les trois dernières densités ont un réel intérêt cryptographique). Dans la suite, nous allons étudier expérimentalement les performances des différents algorithmes de réduction sur des réseaux de ce modèle. Nous allons en particulier voir que les nouveaux algorithmes sont là encore plus rapides et que les bases de sortie restent de qualité comparable à celle de l'algorithme LLL.

8.3.3 Étude comparative à densité “fixée”

Les Tableaux 8.2, 8.3, 8.4, 8.5 présentent des résultats en moyenne sur des réseaux sac–à–dos de densités fixées (respectivement à 2, 1, 2/3, 1/2) en fonction de la dimension du réseau. À chaque fois, le tableau du haut représente le nombre moyen d'itérations pour les différents algorithmes de réduction. Le tableau du milieu représente la valeur moyenne des racines n-ème du défaut d'orthogonalité de la base aléatoire initiale (première ligne) et des bases de sortie des différents algorithmes. Le tableau du bas représente le rapport entre les normes du vecteur de sortie le plus court des nouveaux algorithmes de réduction par celle du vecteur de sortie le plus court de l'algorithme LLL. Les Figures 8.1, 8.2, 8.3, 8.4 illustrent les résultats rassemblés dans ces tableaux.
Là encore, la remarque faite pour les expériences du modèle uniforme concernant la longueur de la phase d'initialisation reste valable: la différence entre les nombre d'itérations pour les algorithmes de réduction n'est pas reflétée exactement dans les temps d'exécution en particulier, à cause d'une identique phase d'initialisation coûteuse pour tous les algorithmes. Toutefois, ici, la matrice exprimant la base de départ dans la base canonique est assez “creuse”. Les entiers qui interviennent dans cette phase d'initialisation sont moins grands, et l'orthogonalisation de Schmidt est ainsi plus rapide que dans le cas du modèle uniforme.

2. La densité du réseau est la densité des poids a_i associés.

Densité 2

Nombre d'itérations

dim	10	20	30	40	50	60	80	100	110
LLL	97.9	211	340	490	657	859	1297	2209	2322
Sch	19.1	39	67.5	98	130	168	238	345	379
Gr	26	55.4	84	110	134	157	225	295	323
Gr–Sch	25.9	57	89	113	150	187	262	357	391
Sch opt	26.1	58	104	160	217	289	429	625	613

Racine n-ème du défaut d'orthogonalité

dim	10	20	30	40	50	60	80	100	110
aléa	49	$2.8e3$	$1.3e5$	$5.7e6$	$2.1e8$	$8.1e9$	$1.1e13$	$1.4e16$	$5.2e17$
LLL	1.2	1.5	1.9	2.2	2.6	3	3.9	4.3	4.3
Sch	2.5	4.0	5.7	8.4	12	17	32	47	56
Gr	1.5	2.2	3.7	6.1	10	19	36	73	97
Gr–Sch	1.5	2.2	3.4	5.9	7.9	12	21	37	50
Sch opt	2.3	2.6	2.9	3.3	5.1	5.6	7.9	9.9	15

Comparaison des normes des vecteurs les plus courts par rapport à LLL

dim	10	20	30	40	50	60	80	100	110
Sch	1	1.2	1.5	1.9	2	2.6	3.8	5	5.7
Gr	1	1.1	1.4	1.7	2	3.4	4.5	11	14
Gr–Sch	1	1.1	1.2	1.7	1.9	2.3	2.7	5.9	5.2
Sch opt	1	1.1	1.1	1.2	1.3	1.4	1.8	1.9	1.9

TAB. 8.2 – *Comparaison entre le nombre d'itérations et la qualité de la base de sortie des différents algorithmes de réduction sur les réseaux sac–à–dos de densité* 2.

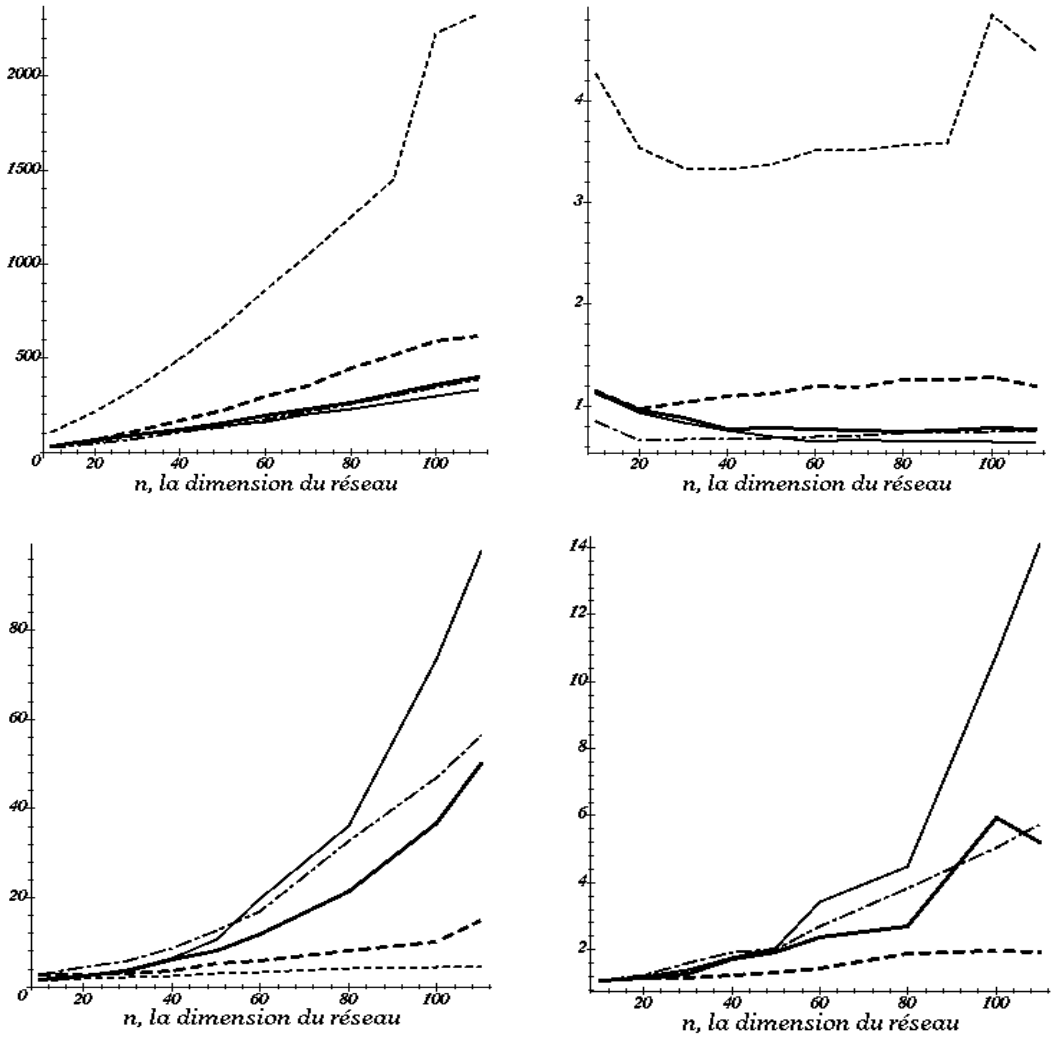

FIG. 8.1 – *Étude comparative sur des réseaux sac–à–dos de densité* 2*. En haut, la figure de gauche représente la variation du nombre moyen d'itérations pour les différents algorithmes de réduction en fonction de la dimension. La figure de droite représente le rapport entre le nombre moyen d'itérations et la fonction* $n \log n$ *en fonction de la dimension* n*. En bas, la figure de gauche représente la variation de la racine* n*-ème du défaut d'orthogonalité. La figure de droite représente les rapports entre la norme du vecteur de sortie le plus court des nouveaux algorithmes et celle du vecteur de sortie le plus court de LLL.*

LLL – – –   *Schm opt* – – –   *Gr-Sch* ——   *Schm* – · – · –   *Gr* ——

Densité 1										
Nombre d'itérations										
dim	10	20	30	40	50	60	80	90	100	110
LLL	140	449	782	1131	1775	1763	2163	2266	2363	2600
Sch	29.9	75.1	137	208	279	340	462	483	500	596
Gr	35	72.5	129	179	231	286	761	932	1058	1329
Gr–Sch	36	87	148	218	286	372	813	951	1115	1363
Sch opt	62	144	254	393	542	614	737	758	779	810
Racine n-ème du défaut d'orthogonalité										
dim	10	20	30	40	50	60	80	90	100	110
aléa	$1.1e3$	$2e6$	0.29	$4.3e12$	$5e15$	$6.6e18$	$8.9e24$	$1e28$	$1.2e31$	$1.4e34$
LLL	1.2	1.6	2.1	2.8	3	3.3	3.8	3.8	4.2	4.1
Sch	2.3	5.9	10	18	31	54	64	48	38	35
Gr	2	5.9	12	26	54	97	38	27	32	33
Gr–Sch	2	4.3	8.6	16	28	31	19	19	20	19
Sch opt	1.5	2.6	4	6.3	8.3	7.3	9	8.3	10	9.9
Comparaison des normes des vecteurs les plus courts par rapport à LLL										
dim	10	20	30	40	50	60	80	90	100	110
Sch	1.4	2.4	2.7	4.2	6.5	5.4	5.7	6.4	5.4	3.8
Gr	1.5	2.7	3.3	6.4	12	15	5.7	4	4.7	5.9
Gr–Sch	1.4	2	2.6	3.9	5.9	5.4	2.6	3.8	3.5	3.4
Sch opt	1.1	1.4	1.6	1.6	1.9	1.7	1.8	1.7	1.7	2.1

TAB. 8.3 – *Comparaison entre le nombre d'itérations et la qualité de la base de sortie des différents algorithmes de réduction sur les réseaux sac–à–dos de densité* 1.

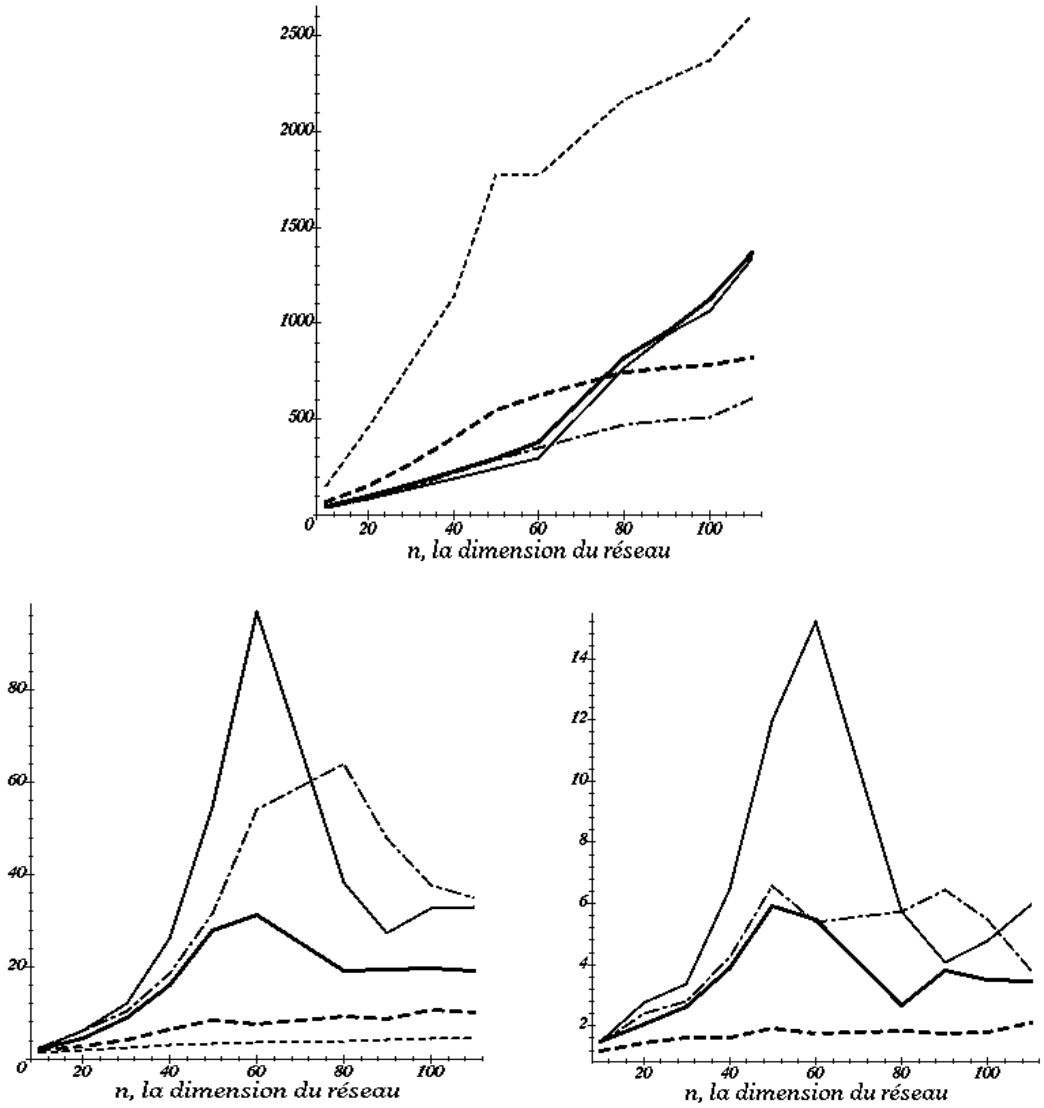

FIG. 8.2 – *Étude comparative sur des réseaux sac–à–dos de densité* 1*. En haut, la figure représente la variation du nombre moyen d'itérations pour les différents algorithmes de réduction en fonction de la dimension. En bas, la figure de gauche représente la variation de la racine n-ème du défaut d'orthogonalité. La figure de droite représente les rapports entre la norme du vecteur de sortie le plus court des nouveaux algorithmes et celle du vecteur de sortie le plus court de LLL.*

LLL – – – *Schm opt* – – – *Gr-Sch* —— *Schm* – · – · – *Gr* ——

Densité 2/3								
Nombre d'itérations								
dim	10	20	30	40	50	80	90	100
LLL	178	699	1306	1596	1674	2546	2569	2735
Sch	45	119	219	318	356	576	640	672
Gr	51.5	118	186	271	613	1332	1465	1587
Gr–Sch	54	132	233	359	648	1332	1482	1631
Sch opt	96	237	451	585	620	906	914	929
Racine n-ème du défaut d'orthogonalité								
dim	10	20	30	40	50	80	90	100
aléa	$2.3e4$	$1.5e9$	$7.3e13$	$3.2e18$	$1.2e23$	$6.8e36$	$2.6e41$	$9.6e45$
LLL	1.2	1.5	2.3	2.5	2.9	3.5	4.1	4.1
Sch	2.7	8.3	18	28	27	45	39	49
Gr	2.3	7.8	29	63	24	25	26	36
Gr–Sch	2.1	6.2	14	21	15	19	20	16
Sch opt	1.5	2.9	5	5.9	6.2	7.7	9.6	10
Comparaison des normes des vecteurs les plus courts par rapport à LLL								
dim	10	20	30	40	50	80	90	100
Sch	2.7	5	6.7	5.8	5.7	5.3	5.3	5.8
Gram	2.5	5.8	10	16	6	4.1	4.3	5.7
Gr–Sch	2.1	4.7	5.8	5.3	4.1	2.8	3.6	3.4
Sch 3/4	1.1	2.2	2.2	1.8	1.7	1.8	1.9	2

TAB. 8.4 – *Comparaison entre le nombre d'itérations et la qualité de la base de sortie des différents algorithmes de réduction sur les réseaux sac–à–dos de densité* 2/3.

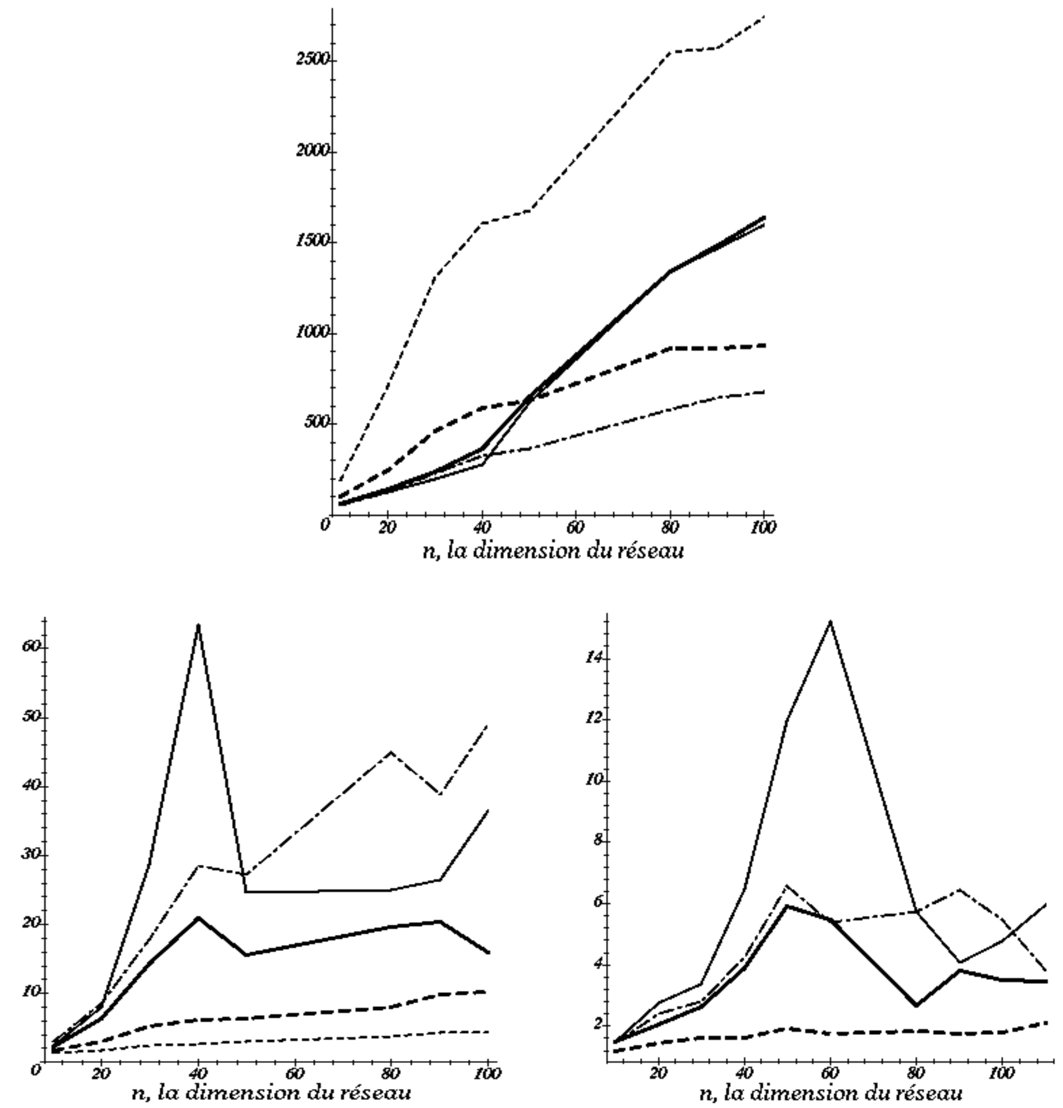

FIG. 8.3 – *Étude comparative sur des réseaux sac–à–dos de densité* 2/3. *En haut, la figure représente la variation du nombre moyen d'itérations pour les différents algorithmes de réduction en fonction de la dimension. En bas, la figure de gauche représente la variation de la racine* n*-ème du défaut d'orthogonalité. La figure de droite représente les rapports entre la norme normes du vecteur de sortie le plus court des nouveaux algorithmes et celle du vecteur de sortie le plus court de LLL.*

LLL – – –   *Schm opt* — — —   *Gr-Sch* ——   *Schm* – · – · –   *Gr* ——

Densité 1/2					
Nombre d'itérations					
dim	10	20	30	40	50
LLL	207	941	1451	1756	1843
Sch	64	173	308	402	461
Gr	63	151	252	652	794
Gr–Sch	69	186	361	713	828
Scmidt opt	128	370	565	710	733
Racine n-ème du défaut d'orthogonalité					
dim	10	20	30	40	50
aléa	$5.8e5$	$1.1e12$	$1.7e18$	$2.3e24$	$2.9e30$
LLL	1.2	1.5	1.9	2.4	2.8
Sch	2.8	9.3	25	32	26
Gr	2.7	13	48	26	20
Gr–Sch	2.4	8.1	16	11	15
Sch opt	1.4	3	4.3	5	5.9
Rapport des normes des plus courts vecteurs					
dim	10	20	30	40	50
Sch	3.4	7.8	9.6	7	5.7
Gr	3.6	13	23	8.3	5.4
Gr–Sch	2.8	8.1	8.6	4	3.8
Sch opt	1	3.1	2.5	1.9	1.7

TAB. 8.5 – *Comparaison entre le nombre d'itérations et la qualité de la base de sortie des différents algorithmes de réduction sur les réseaux sac–à–dos de densité* 1/2.

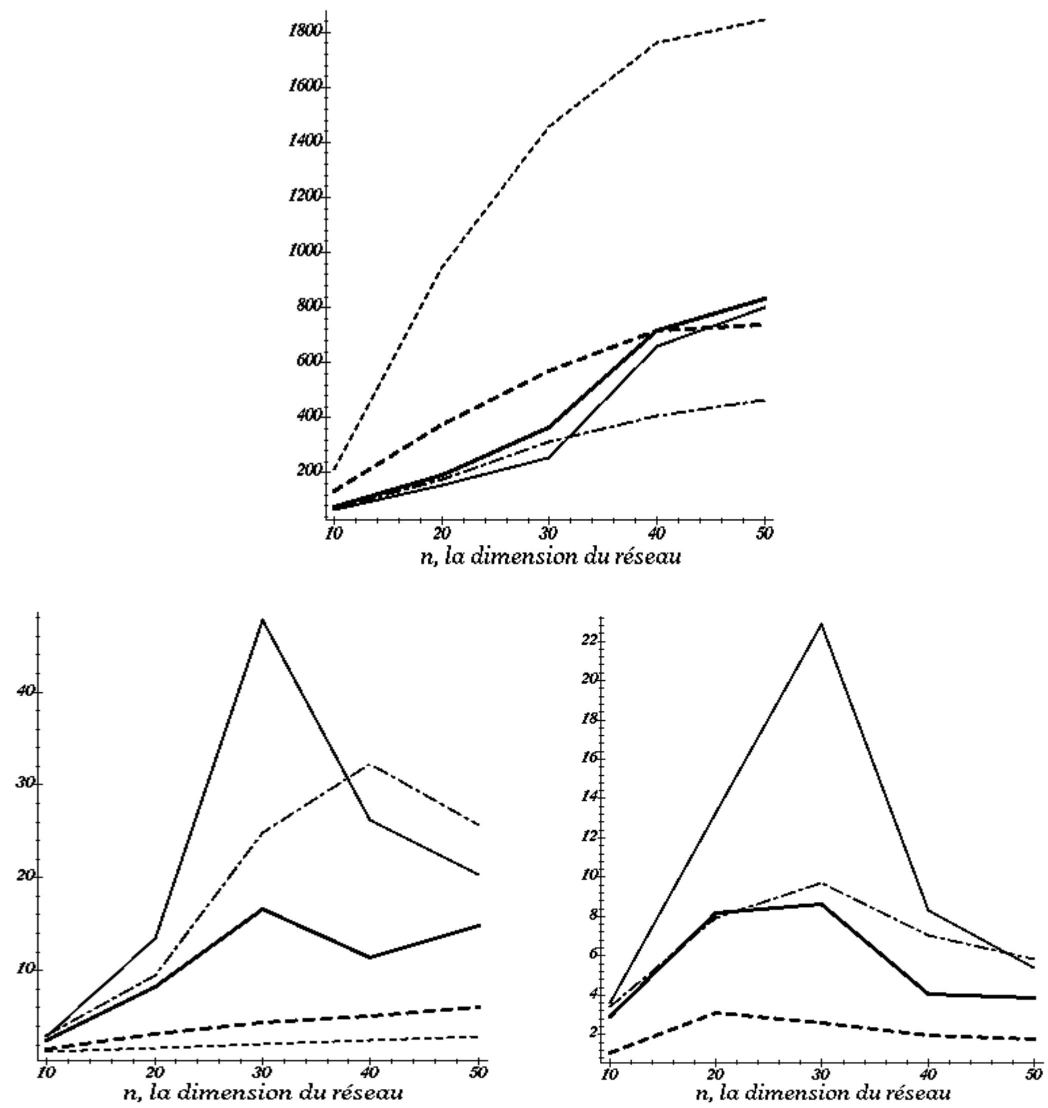

FIG. 8.4 – *Étude comparative sur des réseaux sac–à–dos de densité* 1/2. *En haut, la figure représente la variation du nombre moyen d'itérations pour les différents algorithmes de réduction en fonction de la dimension. En bas, la figure de gauche représente la variation de la racine n-ème du défaut d'orthogonalité. La figure de droite représente les rapports entre la norme du vecteur de sortie le plus court des nouveaux algorithmes et celle du vecteur de sortie le plus court de LLL.*

LLL – – – *Schm opt* – – – *Gr-Sch* —— *Schm* – · – · – *Gr* ——

Les tableaux précédents montrent avant tout que les nouveaux algorithmes sont en moyenne plus rapides même sur les réseaux sac–à–dos (plusieurs densités différentes ont été envisagées). Autrement dit, le comportement relatif des algorithmes de réduction sur les réseaux du modèle uniforme est ici conservé.
Par ailleurs les expériences semblent montrer (moyennant d'autres essais et aux incertitudes de l'expérience près) que le nombre d'itérations sur ce modèle pour tous les différents algorithmes de réduction est “sub-quadratique”. On rappelle que le nombre maximum d'itérations est borné en $O(n^2)$.
Par ailleurs, les performances des différents algorithmes varient en fonction de la densité et de la dimension du réseau. En particulier, pour les réseaux de densité inférieurs ou égale à 1 (ceux qui sont intéressant en cryptographie), les algorithmes de Schmidt sont ceux qui se comportent le mieux en moyenne. On peut noter les performances en moyenne de l'algorithme de Schmidt optimal qui, en un nombre moyen d'itérations de l'ordre du tiers de l'algorithme LLL, fournit un vecteur qui est, en moyenne, au plus 2 fois plus long que le vecteur le plus court construit par LLL.

Les résultats expérimentaux précédents peuvent être un point de départ pour la mise au point de stratégies algorithmiques, dépendant de l'application.

- Choix d'un algorithme de réduction en fonction de la qualité exigée par l'application visée, de la dimension et de la densité.
- Enchaînement de plusieurs algorithmes de réduction différents.
- Combinaison avec les versions de LLL en flottant [Sch88] ou en parallèle [HT93, Jou93, Vil92].

8.3.4 Influence de la taille de la base, de la densité du réseau

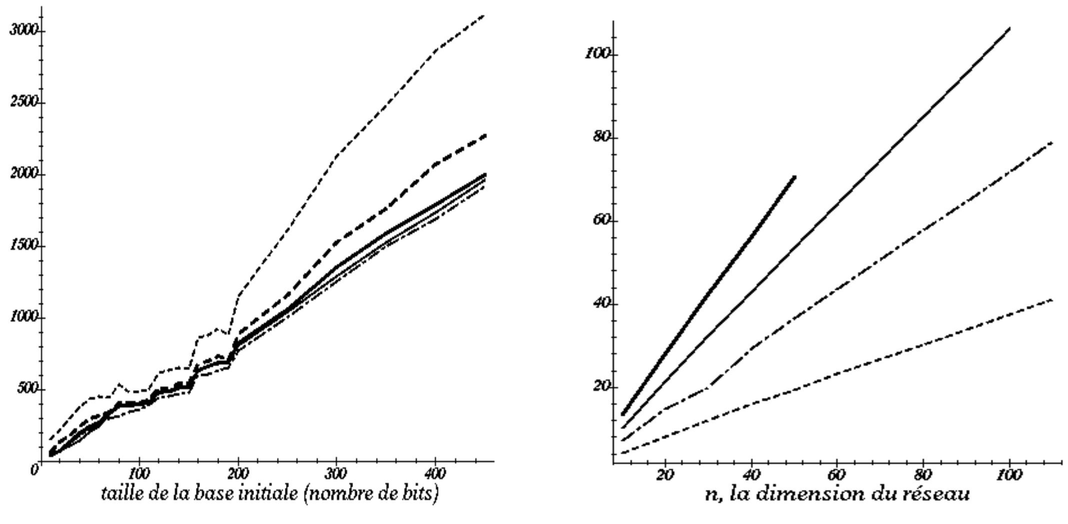

FIG. 8.5 – *La figure de droite représente les nombres moyens d'itérations des différents algorithmes de réduction en fonction de la taille de la base à dimension fixée (*10*). La taille de la base varie de* 2^{10} *à* 2^{500}*. Lorsque* n *désigne la dimension du réseau aléatoire, la figure de gauche représente* $(1/n)\log\rho$ *en fonction de* n*, pour des réseaux aléatoires de densités* 1/2, 1, 3/2, 2.

Un premier phénomène assez clairement exhibé par ces expériences est une évaluation du défaut d'orthogonalité des bases aléatoires du modèle sac–à–dos en fonction de la densité et de la dimension du réseau. Le résultat est en accord avec la majoration (8.3). La Figure 8.5 (courbes de droites) semble montrer en effet que le défaut d'orthogonalité moyen $\mathrm{E}[\rho]$ des réseaux aléatoires du modèle sac–à–dos est de la forme:

$$\mathrm{E}[\rho] = \Theta(e^{\frac{n^2}{d}}).$$

Par ailleurs, comme le suggérait l'ensemble des expériences dans le modèle sac–à–dos, ici le nombre moyen d'itérations semble dépendre aussi de la taille de la base. (On sait que dans le modèle uniforme, ce nombre est asymptotiquement indépendant de la taille [DV94].)

Chapitre 9

Conclusion générale

Le propos général de cette thèse est d'éclaircir le fonctionnement des différents processus de réduction des réseaux euclidiens et de suggérer de nouvelles stratégies algorithmiques en fonction des applications visées. Dans ce but, nous avons introduit de nouvelles réductions. Puis, nous avons mené une étude comparative des différents algorithmes, en considérant leur nombre d'itérations et la qualité de leur base de sortie.

9.1 Analyse dans le pire des cas

9.1.1 Principaux résultats

L'analyse dans le pire des cas du Chapitre 2 donne la même borne pour la complexité de tous ces algorithmes. Cette borne est valable pour toutes les valeurs des paramètres d'approximation, exceptées les valeurs optimales. Dans ce cadre, cette première analyse montre que tous les algorithmes de réduction fonctionnent en temps polynomial par rapport à la taille et à la dimension de la base d'entrée. Le Chapitre 2 décrit également la qualité des différentes bases de sortie dans le pire des cas. Cette étude est complétée dans le Chapitre 7, par celle de la complexité de l'algorithme LLL, pour la valeur optimale du paramètre d'approximation ($t = 1$). Il s'agissait là d'un problème ouvert auquel nous avons apporté une réponse partielle: l'algorithme LLL reste polynomial en la taille de la base d'entrée à dimension fixée. En dimension 3, je donne une borne explicite, *linéaire* en la taille de l'entrée, ainsi qu'une caractérisation précise de la base de sortie.

9.1.2 Problèmes ouverts

Le problème, important à mon sens, qui reste ouvert est la mise en évidence pour chaque dimension n, des bases qui exigent le plus d'étapes pour être réduites. Vallée l'a déjà fait en dimension 2 [Val91] et Kaib et Schnorr [KS96] ont généralisé son analyse à d'autres normes que la norme euclidienne[1]. L'algorithme LLL peut être considéré en particulier comme un système de réécriture dans le groupe des matrices unimodulaires de rang n. En dimension 2, les règles qui régissent cette réécriture peuvent être assez facilement mises en évidence. Liées bien sûr à la représentation de $GL_2(\mathbb{Z})$, ces règles ne sont pour autant pas triviales.

1. En dimension 2, l'algorithme LLL coïncide avec l'algorithme bien connu de Gauss.

Elles permettent de retrouver la borne de [Val91], en dimension 2. Plus généralement, la connaissance de ces règles conduit, au moins intuitivement, à reconnaître, pour une taille fixée de matrice, celles qui admettent les décompositions les plus longues et par suite, les types de bases qui demandent le plus d'étapes, pour être réduites. La mise en évidence de ces règles lorsque la dimension dépasse 2 semble beaucoup plus difficile. A mon avis, considérer les algorithmes de réduction comme des systèmes de réécriture dans $GL_n(\mathbb{Z})$ pourrait sans doute apporter une compréhension différente de leur fonctionnement. Cette approche permettrait de relier le problème de la réduction des réseaux à la théorie de la présentation des groupes discrets tels que $GL_n(\mathbb{Z})$ [CM72, MKS66]. Je pense qu'elle apportera des éléments de réponse à la question de la dépendance exacte du nombre maximum d'itérations en la dimension du réseau.

9.2 Analyse en moyenne

9.2.1 Principaux résultats

Afin de mieux comprendre ces algorithmes et aussi de les classer, là où l'analyse dans le pire des cas ne le permettait pas, nous avons opté pour une étude probabiliste, dans un modèle simple et naturel, le modèle uniforme. Ce modèle est a priori, plutôt théorique: notre étude du Chapitre 3 montre que ces bases sont en moyenne, déjà de bonne qualité. Mais l'étude expérimentale du Chapitre 8 semble montrer que nos résultats sur les comportements relatifs des différents algorithmes de réduction restent valables dans d'autres modèles, provenant cette fois des applications réelles de la réduction. Je rappelle les principaux résultats de l'étude probabiliste.

– Les bases aléatoires du modèle uniforme.

Ces bases ont été déjà considérées par Daudé et Vallée [DV94]. En utilisant leur résultat concernant la norme des orthogonalisés (rappelé en Théorème 4), j'obtiens des bornes sur la moyenne de tous les défauts de longueur et du défaut d'orthogonalité de ces bases. Je montre en particulier que la qualité moyenne d'une base aléatoire dépend de façon sensible d'un paramètre, appelé *poids*, qui est le rapport entre la dimension de la base et celle de l'espace ambiant. Par exemple, une base de cardinal constant est en moyenne presque orthogonale, pourvu que ses vecteurs soient choisis dans un espace ambiant dont la dimension tend vers l'infini.

– Introduction et analyse comparative de nouveaux algorithmes de réduction.

J'ai introduit trois nouveaux algorithmes de réduction. Je montre (Chapitre 2) que ces algorithmes, dont les conditions de sortie sont plus faibles que celle de l'algorithme LLL, construisent toutefois des bases de qualité similaire et suffisante pour la plupart des applications. Une large partie de mon travail a eu comme but de "séparer" les comportements de ces différents algorithmes de réduction et de montrer que les nouveaux algorithmes sont plus rapides. C'est une analyse en moyenne précise qui a rendu possible le classement de ces algorithmes: je montre que le nombre d'itérations des algorithmes de Schmidt, de Gram, de Gram–Schmidt (avec des choix appropriés des paramètres) est asymptotiquement le minimum possible. En même temps, je montre qu'avec une probabilité–limite non–négligeable, l'algorithme classique exige plus que ce nombre minimal d'itérations.

9.2.2 Problèmes ouverts

Voici quelques perspectives pour la suite de notre étude probabiliste. Les 3 premières questions évoquées concernent toujours le modèle uniforme.

1. Pour comparer les différentes réductions, j'ai été amené à évaluer les probabilités asymptotiques de réduction des bases aléatoires [2]. En particulier, le Corollaire 9 et la Proposition 20 donnent des majorants pour la probabilité asymptotique de réduction d'une base aléatoire pleine, au sens de Siegel ou de Lovász. Une première question ouverte est de savoir si lorsque la dimension tend vers l'infini, la probabilité que la base aléatoire soit réduite au sens de s-Siegel, a bien une limite (Conjecture 2). L'existence de cette probabilité–limite et sa dépendance en fonction du paramètre s d'approximation, sont pour moi des questions théoriques intéressantes.
2. J'ai montré que si la base d'entrée est de poids strictement inférieur à un poids–seuil –qui dépend du paramètre d'approximation–, alors le nombre d'itérations de l'algorithme LLL est le minimum possible (ou encore la base d'entrée, rendue propre, est déjà réduite). Le Théorème 10 donne un minorant pour ce poids–seuil. La connaissance plus précise de ce poids–seuil est une autre question intéressante (cf. Conjecture 1). Plus généralement, étant donnée une base aléatoire pleine, quelle proportion précise de la fin de la base n'est pas déjà réduite?
3. La seule borne connue sur le nombre moyen d'itérations de l'algorithme LLL, due à Daudé et Vallée [DV94], est en $\mathrm{O}(n^2 \log n)$. D'une part, j'ai montré qu'une partie (sans doute importante) d'une base aléatoire pleine est déjà réduite (mais la totalité de la base ne l'est pas). D'autre part, la réduction de Siegel coïncide avec celle de (s, γ)-Gram, en choisissant $\gamma = 0$. Il serait vraiment étonnant que le nombre moyen d'itérations passe de $n - 1$ $(\gamma > 0)$ à $n^2 \log n$ $(\gamma = 0)$. C'est pourquoi, je pense que $n^2 \log n$ n'est pas une borne optimale pour le nombre d'itérations de l'algorithme LLL sur les réseaux aléatoires du modèle uniforme. Les expériences du Chapitre 8 semblent montrer que le nombre moyen d'itérations est “sub–quadratique”. Intuitivement, on peut comparer l'algorithme LLL à “une sorte de tri” et espérer alors des bornes en $\mathrm{O}(n \log n)$.
4. Analyse en moyenne dans d'autres modèles.
   Dans le Chapitre 8, nous avons considéré expérimentalement, le modèle des réseaux sac–à–dos. En effet, dans un grand nombre d'applications, les bases de départ sont de cette forme. Sur de tels réseaux, “difficiles”, les bornes (optimales) établies dans le modèle uniforme ne sont plus valides. En revanche, l'expérience montre que l'ordre relatif parmi ces bornes est maintenu et que les nouveaux algorithmes restent plus rapides que l'algorithme LLL en moyenne. De plus, les bases de sortie conservent une qualité satisfaisante. L'analyse en moyenne des algorithmes de réduction dans le modèle sac–à–dos est un sujet intéressant (et sans doute difficile).
5. Cette thèse est aussi un point de départ pour l'élaboration de stratégie algorithmique, en fonction de l'application de la réduction. Pour être efficaces, ces stratégies pourront combiner les améliorations déjà existantes apportées à l'algorithme LLL, ainsi que les nouvelles réductions que nous avons introduites.

---

2. Ces probabilités concernent les bases aléatoires, rendues propres préalablement.

Bibliographie

[ABD+95] F. Avnaim, J.-D. Boissonnat, O. Devillers, F. Preparata, and M. Yvinec. Evaluation of a new method to compute signs of determinants. In *Eleventh Annual ACM Symposium on Computational Geometry*, pages C16–C17, 1995. Full paper to appear in *Algorithmica*, 1997.

[ABSS97] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The Hardness of Approximate Optima in Lattices, Codes and Systems of Linear Equations. *Journal of Computer and System Sciences*, 54:317–331, 1997.

[Ajt96] M. Ajtai. Generating hard instances of lattice problems. In *Proc. 28-th Symposium On Theory of Computing*, pages 99–108, 1996.

[Ajt97] M. Ajtai. The shortest vector problem in L_2 is NP-hard for randomized reduction. *Electronic Colloquim on Computationnal Complexity*, 1997. available via: http://www.eccc.uni-trier.de/eccc/.

[Akh99] A. Akhavi. Threshold phenomena in random lattices and efficient reduction algorithms. Preprint, 1999.

[AL97] S. Arora and C. Lund. *Hardness of approximations*, chapter X. PWS Publishing, 1997. in D. S. Hochbaum (ed.).

[Bab86] L. Babai. On Lovász' lattice reduction and the nearest lattice point problem. *Combinatorica*, 6(1):1–13, 1986.

[BO78] C.M. Bender and S.A. Orzag. *Advanced Mathematical Methods for Scientsits and Engineers.* MacGraw-Hill, New York, 1978.

[BS98] J. Blömer and J. P. Seifert. On the Complexity of Computing Short Linearly Independant Vectors and Short Bases in a Lattice. Preprint, 1998.

[Cas71] J. W. S. Cassels. *An introduction to the Geometry of Numbers.* Springer-Verlag, 1971.

[CM72] H. S. M. Coxeter and W. O. J. Moser. *Generators and relations for discrete groups.* Springer–Verlag, 1972.

[Dau93] H. Daudé. *Des fractions continues à la réduction des réseaux: analyse en moyenne.* PhD thesis, Université de Caen, 1993.

[Dav] H. A. David. *Order Statistics.*

[DeB81] N. G. DeBruijn. *Asymptotic methods in Analysis.* Dover, New York, 1981.

[DFV97] H. Daudé, Ph. Flajolet, and B. Vallée. An average–case analysis of the Gaussian algorithm for lattice reduction. *Combinatorics, Probability and Computing*, 123:397–433, 1997.

[DV94] H. Daudé and B. Vallée. An upper bound on the average number of iterations of the LLL algorithm. *Theoretical Computer Science*, 123:95–115, 1994.

[FHK+] A. Frieze, J. Hastad, R. Kannan, J. C. Lagarias, and A. Shamir. Reconstructing truncated integer variables satisfying linear congruences. *S.I.A.M. Journal of Computing.*

[GG98] O. Goldreich and S. Goldwasser. On the limits of Non-Approximability of Lattice Problems. In *Proc. 30-th Symposium On Theory of Computing*, pages 1–9, 1998.

[HJ85] B. Helfrich-Just. Algorithms to construct Minkowski and Hermite reduced bases. *Theoretical Computer Science*, 41:125–139, 1985.

[HT93] C. Heckler and L. Thiele. A parallel lattice basis reduction for mesh–connected processor arrays and parallel complexity. In *Proc SPDP'93, Dallas USA*, December 1–4 1993.

[Jou93] A. Joux. A fast parallel lattice reduction algorithm. In *Proc. Second Gauss Symposium (Munich)*, 1993.

[Kan83] R. Kannan. Improved algorithm for integer programming and related lattice problems. In *15th Ann. ACM Symp. on Theory of Computing*, pages 193–206, 1983.

[Kan87] R. Kannan. Algorithmic geometry of numbers. *Ann. Rev. Comput. Sci.*, 2:231–267, 1987.

[Knu81] Donald E. Knuth. *The Art of Computer Programming*, volume 2: Seminumerical Algorithms. Addison-Wesley, 2nd edition, 1981.

[KS96] M. Kaib and C. P. Schnorr. The generalized Gauss reduction algorithm. *Journal of Algorithms*, 21:565–578, 1996.

[KZ73] A. Korkine and G. Zolotareff. Sur les formes quadratiques. *Math. Annalen*, 6:366–389, 1873.

[Lag80] J. C. Lagarias. Worst–case complexity bounds for algorithms in the theory of integral quadratic forms. *J. Algorithms*, 1:142–186, 1980.

[Lag82] J. C. Lagarias. The computationnal complexity of simultaneous diphantine approximation problems. In *23rd IEEE Symposium On Foudations of Computer Science*, pages 32–39, 1982.

[Lag83] J. C. Lagarias. Solving low-density subset problems. *IEEE*, 1983.

[Len83] H.W. Lenstra. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8:538–548, 1983.

[LG87] C. Lekkerkerker and P.M. Grüber. *Geometry of numbers.* Noth Holland, 1987.

[LLL82] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:513–534, 1982.

[LLS86] J. C. Lagarias, H. W. Lenstra, and C. P. Schnorr. Korkine-Zolotarev bases and successive minima of a lattice and its reciprocal lattice. Technical report, Mathematical Science Resaerch Institute, Berkeley, 1986.

[LO83] J. C. Lagarias and A. M. Odlyzko. Solving low–density subset sum peoblems. In *24th IEEE Symposium FOCS*, pages 1–10, 1983.

[Mic98] D. Micciancio. The Shortest Vector in a lattice is hard to approximate to within some constant. In *to appear in Proc. 39-th Symposium On Foundations of Computer Science*, 1998.

[Min91] H. Minkowski. Über die positiven quadratischen Formen und über kettenbruchähnliche Algorithmen. *Crelles Journal fur die reine und angewandte mathematik*, 107:278–297, 1891.

[Min10] H. Minkowski. *Geometrie der Zahlen.* Teubner, Leipzig, 1910.

[MKS66] W. Magnus, A. Karrass, and D. Solitar. *Combinatorial group theory: presentation of groups in terms of generators and relations.* Interscience Publishers, 1966.

[New72] M. Newman. *Integral matrices.* Academic Press, 1972.

[Sch84] A. Schönhage. Factorization of univariate integer polynomials by diophantine approximation and an improved basis reduction algorithm. In *Lect. Notes Comput. Sci.*, volume 172, pages 436–447, 1984.

[Sch87] C. P. Schnorr. A hierarchy of polynomial time lattice basis reduction algorithm. *Theoretical Computer Science*, 53:201–224, 1987.

[Sch88] C. P. Schnorr. A more efficient algorithm for lattice basis reduction. *J. Algorithms*, 9:47–62, 1988.

[Sch89] A. Schrijver. *Theory of linear and integer programming.* Wiley-Interscience Publication, 1989.

[Sch94] C. P. Schnorr. Block reduced lattice bases and successive minima. *Combinatorics, Probability & Computing*, 1994.

[Sch95] C. P. Schnorr. Attacking the Chor-Rivest cryptosystem by improved lattice reduction. In *Eurocrypt*, 1995.

[SE91] C. P. Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. In *Proceedings of the FCT'91 (Altenhof, Germany), LNCS 529*, pages 68–85. Springer, 1991.

[SF96] R. Sedgewick and Ph. Flajolet. *An Introduction to the Analysis of Algorithms.* Addison-Wesley Publishing Company, 1996.

[Sie89] C. L. Siegel. *Lectures on the Geometry of Numbers.* Springer-Verlag, 1989.

[Val86] B. Vallée. *Une approche géométrique aux réseaux euclidiens en petite dimension.* PhD thesis, Université de Caen, 1986.

[Val89a] B. Vallée. Algorithmique en géométrie des nombres. Applications à la cryptographie et à la factorisation des entiers. Master's thesis, Université de Caen, 1989.

[Val89b] B. Vallée. Un problème central en géométrie algorithmique des nombres: la réduction des réseaux. Autour de l'algoritme LLL. *Informatique Théorique et Applications*, 3:345–376, 1989. English translation by E.Kranakis, CWI Quaterly 1990-3. Amsterdam.

[Val91] B. Vallée. Gauss' algorithm revisited. *Journal of Algorithms*, 12:556–572, 1991.

[vEB81] P. van Emde Boas. Another NP-complete problem and the complexity of finding short vectors in a lattice. *Rep. 81-04 Math. Inst. Univ. Amsterdam*, 1981.

[VGT88] B. Vallée, M. Girault, and Ph. Toffin. How to braek Okamoto's cryptosystem by reducing lattice bases. In *Proceedings of Eurocrypt*, 1988.

[Vil92] G. Villard. Prallel lattice basis reduction. In *International Symposium on Symbolic and Algebric Computation, Berkeley USA*, pages 269–277. ACM Press, 1992.

[Won] Wong. *Asymptotic Approximations of Integrals*, chapter IX.5. Academic Press Inc.

[WW27] E.T. Whittaker and G.N. Watson. *A course of Modern Analysis*. Cambridge University Press, Cambridge (England), 4th edition, 1927. reprinted 1973.