

Unification des méthodes hiérarchiques pour le traitement systèmes linéaires

Emmanuel Agullo (Emmanuel.Agullo@inria.fr), Olivier Coulaud (Olivier.Coulaud@inria.fr)

Stage M2/3A Printemps 2024

1 Contexte

La méthode multipolaire rapide (fast multipole method, FMM) [1] est une technique mathématique hiérarchique développée pour accélérer le calcul des forces de longue portée dans le problème à N -corps réduisant la complexité de quadratique à linéaire. Elle a été reconnue comme l'un des dix plus importants algorithmes du XXe siècle par la Society for Industrial and Applied Mathematics (SIAM). D'un point de vue algébrique, elle peut être vue comme produit matrice-vecteur (MV) où la matrice est découpée hiérarchiquement en sous-matrices. Contrairement à un calcul complètement algébrique, les sous-matrices sont traitées analytiquement, via l'expansion de la fonction de Green du système par une expansion multipolaire, le groupement des sources voisines permettant de les traiter comme une source unique.

Au sein de l'équipe-projet Inria Concace, une version haute performance (HPC) de la méthode FMM a été développée dans la bibliothèque ScalFMM, qui supporte une parallélisation classique MPI+OpenMP ainsi qu'une abstraction à bases de tâches pures au-dessus du moteur d'exécution StarPU [4,5].

2 Objectif

L'objectif du stage est de revisiter la FMM pour l'écrire explicitement comme un produit matrice-vecteur (MV) accéléré. L'intérêt est de bénéficier des travaux récents [2] sur la parallélisation du MV et d'aboutir à un code générique permettant de faire ou bien un MV classique (type GEMM) ou un MV accéléré (type FMM) en factorisant un maximum de code (et de concepts mathématiques). Les machines visées étant des machines hétérogènes distribuées. L'écriture moderne d'un MV parallèle requiert de supporter plusieurs schémas de communication (dits A-, B- et C- stationnaires, notamment) [2]. De même, l'écriture moderne d'une FMM requiert de supporter différents algorithmes de compressions hiérarchiques (analytiques type FMM voire algébriques de type H -matrice ou H^2 -matrice). Une tâche importante de du stage sera donc d'abstraire clairement les schémas de parallélisation (sur la base de [2]) ainsi que les opérations de compression. Les produits matrice-vecteur locaux compressés devront appeler les opérateurs de la FMM au lieu des produits classiques.

Le développement des opérateurs FMM efficaces à la fois sur CPU et GPU est important. Ces travaux se feront au sein du framework de composabilité `compose` développé au sein de l'équipe-projet `concace` et dans lequel ScalFMM sera refondu.

3 Lieu du stage

Le stage aura lieu au sein de l'équipe-projet `concace` de l'Inria Bordeaux.

4 Poursuite en thèse

Le stage pourra être poursuivi en thèse dans l'équipe-projet `concace` à l'Inria Bordeaux.

5 Références

- [1] Greengard, L., & Rokhlin, V. (1997). A new version of the Fast Multipole Method for the Laplace equation in three dimensions. *Acta Numerica*, 6, 229–269.
- [2] Emmanuel Agullo, Alfredo Buttari, Abdou Guermouche, Julien Herrmann, Antoine Jegou. Task-Based Parallel Programming for Scalable Algorithms: application to Matrix Multiplication. Research Report RR-9461, Inria Bordeaux - Sud-Ouest. 2022, pp.29. (hal-03588491v2) <https://hal.inria.fr/hal-03588491/>
- [3] Börm, Steffen, Lars Grasedyck, et Wolfgang Hackbusch. « Hierarchical Matrices ». Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig, 2003.
- [4] Agullo, E., Bramas, B., Coulaud, O., Darve, E., Messner, M., & Takahashi, T. (2014). Task-based FMM for multicore architectures. *SIAM Journal on Scientific Computing*, 36(1), C66-C93
- [5] Agullo, E., Bramas, B., Coulaud, O., Darve, E., Messner, M., & Takahashi, T. (2016). Task-based FMM for heterogeneous architectures. *Concurrency and Computation: Practice and Experience*, 28(9), 2608-2629.