

Processus de Hawkes discret.

Laure Coutin *

Les processus stochastiques sont très utilisés pour modéliser des signaux tels que le nombre de personnes dans une file d'attente, la consommation d'électricité en France et bien d'autres phénomènes. Un processus stochastique est une famille de variables aléatoires $((X_t)_{t \in T})$ où $T = \mathbf{N}$ ou $T = I$ un intervalle de \mathbf{R} . Parmi les plus utilisés on peut citer les marches aléatoires, le processus de Poisson qui modélise l'évolution du nombre de personnes dans une file d'attente. Ces processus ont la propriété que leur valeur à l'instant $t+1$ ne dépend que de la valeur à l'instant t et d'un bruit indépendant. Lors du choix d'une option à l'aide d'un vote à main levée, le choix des individus dépend en partie du choix de tous (ou beaucoup) individus ayant déjà votés et pas uniquement le dernier. Le processus de Hawkes, introduit en 1974 par G. Hawkes et D. Oakes dans [1] permet de modéliser cette situation de dépendance au passé de façon simple. Une version à temps discret existe aussi. Elle repose sur des variables aléatoires de Bernoulli conditionnelles dans [2]. Ces processus font l'objet de nombreux développements récents.

L'objet de ce stage est d'étendre les travaux de Y. Seol, [2], et de H. Wang, [3] au cas de variables aléatoires de Poisson, c'est à dire de proposer :

(i) Une construction simple des processus de Hawkes à temps discret en utilisant des variables aléatoires de Poisson conditionnelles et d'étudier leurs propriétés loi fortes, théorème de la limite centrale, (pouvant être abordé dès le niveau L3).

(ii) De montrer la convergence (en un sens à préciser), après renormalisation, vers un processus à temps continu en s'appuyant sur les processus de Poisson et mesures aléatoires de Poisson à partir d'un niveau M1.

Ce stage pourra se dérouler de mai à mi juillet.

References

- [1] A. G. Hawkes and D. Oakes, A cluster process representation of a self-exciting process, *Journal of Applied Probability*, Volume 11, Number 3, September 1974 , pp. 493-503
- [2] Y. Seol, Limit theorems for discrete Hawkes processes, *Stat. and Proba. Letters* 99 (2015), 223-229
- [3] H. Wang, Limit theorems for a discrete-time marked Hawkes process, arXiv:2003.02810

*IMT, University of Toulouse, France, laure.coutin@math.univ-toulouse.fr

- [4] P. Bremaud, Point Processes and Queues, Martingale Dynamics; Springer-Verlag New York Inc., 1981

Master internship and PhD thesis offer

Additive constrained Gaussian processes in high dimension

General context. This master internship and PhD thesis are part of the ANR JCJC “GAP” project, a 5-year research project funded by the French national research agency. GAP focuses on both theoretical and practical aspects related to Gaussian processes (GPs), a class of stochastic processes widely used in machine learning and computer experiments. The guiding principle of GAP is to push the boundaries of the current mathematical models and algorithmic procedures based on GPs, with a strong emphasis on reliability. The research proposal seeks to exploit additive structures in constrained GPs for which constraints are verified everywhere. This reliability property is key for real-world applications concerned by risk assessment, especially in criticality safety.

The framework of constrained GPs. Adding physical information in models can significantly reduce the size of the training set which is necessary to reach a certain level of accuracy. We consider here the information given by inequality constraints, such as monotonicity. This research considers a class of GP models proposed by [3], for which the inequality constraints are verified everywhere in the space. This GP model has been investigated under different angles: simulation [6], inference properties [1], noisy observations [5] and Bayesian inference [8]. However, in its original form, the model is constructed by tensorization of 1-dimensional models, and thus structurally limited to low dimensional problems (typically less than 5 dimensions). To scale to higher dimensions, a promising algorithm called MaxMod has been introduced, which sequentially finds the more active input variables [2]. Recently, MaxMod has been extended to hundreds of dimensions under additivity assumptions of the GP model [7].

Research directions. Constrained additive GPs are of great interest to address high dimensional problems with inequality information. However, the additivity assumption used in [7] excludes the presence of interactions of input variables, which may be very restrictive in real-world applications. To conciliate these two objectives, we propose to consider block-additive GPs of the form,

$$\sum_{I \in \mathcal{I}} Z_I(x_I). \quad (1)$$

which are additive with respect to *groups* of input variables. In (1), \mathcal{I} contains groups of input variables. While considering interactions in additive model is not new (see, e.g., [4] in machine learning, or [9] in computer experiments), there still remain difficult questions to solve, especially when inequality constraints are considered. To cite a few: how to select the groups, namely the subsets $I \in \mathcal{I}$? How to adapt the MaxMod algorithm? How to deal with non-overlapping groups? How to obtain theoretical guarantees?

Expected research outputs. The output is methodological and will lead to publications in international journals in applied mathematics and/or conferences in machine learning. Furthermore, the methodology may also be applied on real-world applications, such as the environmental risk studies considered in [2, 7]. We expect to significantly improve the efficiency of the current models and algorithms by using both the wider class of block-additive GPs and the physical information brought by inequality constraints.

Supervising team and organizational details. The master internship is funded for about 6 months and will start in Spring 2023. The PhD is funded for 3 years and will start in Fall 2023. For a student already holding a master degree, the PhD training can start as soon as possible. The student will be supervised by a leading team on constrained Gaussian processes:

- François Bachoc, Ass. Prof. at *Institut de Mathématiques de Toulouse (IMT)*,
- Andrés F. López-Lopera, Ass. Prof. at *Université Polytechnique Hauts-de-France*,
- Olivier Roustant, Prof. at *INSA Toulouse & IMT*.

The student will be located at IMT, France. An international mobility during the PhD is envisaged, as a collaboration with other researchers of the ANR JCJC GAP project.

How to apply? Applications will be considered until the position is filled. The candidates should have master-level skills in mathematics / statistics / machine learning. Please send a CV (either in English or French), application letter and grade transcripts (bachelor and master level) to

- François Bachoc (francois.bachoc@math.univ-toulouse.fr),
- Andrés F. López-Lopera (andres.lopezlopera@uphf.fr), and
- Olivier Roustant (roustant@insa-toulouse.fr).

References related to the PhD subject.

- [1] F. Bachoc, A. Lagnoux, and A. F. López-Lopera. Maximum likelihood estimation for Gaussian processes under inequality constraints. *Electronic Journal of Statistics*, 13(2):2921–2969, 2019.
- [2] F. Bachoc, A. F. López-Lopera, and O. Roustant. Sequential construction and dimension reduction of Gaussian processes under inequality constraints. *SIAM Journal on Mathematics of Data Science*, 4(2):772–800, 2022.
- [3] X. Bay, L. Grammont, and H. Maatouk. Generalization of the Kimeldorf-Wahba correspondence for constrained interpolation. *Electronic journal of statistics*, 10(1):1580–1595, 2016.
- [4] D. Duvenaud, H. Nickisch, and C. Rasmussen. Additive Gaussian processes. In *NeurIPS*, 2011.
- [5] A. F. López-Lopera, F. Bachoc, N. Durrande, J. Rohmer, D. Idier, and O. Roustant. Approximating Gaussian process emulators with linear inequality constraints and noisy observations via MC and MCMC. In *Monte Carlo and Quasi-Monte Carlo Methods*, 2020.
- [6] A. F. López-Lopera, F. Bachoc, N. Durrande, and O. Roustant. Finite-dimensional Gaussian approximation with linear inequality constraints. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1224–1255, 2018.
- [7] A. F. López-Lopera, F. Bachoc, and O. Roustant. High-dimensional additive Gaussian processes under monotonicity constraints. *Accepted for publication in NeurIPS 2022* <https://arxiv.org/abs/2205.08528>, 2022.
- [8] A. F. Lopez-lopera, S. John, and N. Durrande. Gaussian process modulated Cox processes under linear inequality constraints. In *AISTATS*, 2019.
- [9] T. Muehlenstaedt, O. Roustant, L. Carraro, and S. Kuhnt. Data-driven Kriging models based on FANOVA-decomposition. *Statistics & Computing*, 22(3):723–738, 2012.

Description of the internship project “Chen-Stein”

Title: “Distribution of the number of occurrences of segments
with a level higher than a threshold”

-

Simona GRUSEA¹, Sabine MERCIER², Toulouse Mathematics Institute,
¹ grusea@insa-toulouse.fr, INSA Toulouse,
² mercier@univ-tlse2.fr, Toulouse 2 University Jean Jaurès

The internship will be paid.
Duration : 4 or 6 months.
Expected beginning: during spring 2023.

Context and goal

We are interested in highlighting in a sequence, atypical periods or regions corresponding to breakpoints or interferences. Numerous results exist in the literature for highlighting atypical regions or breakpoints, but they generally focus on the most atypical region (in the case of biological sequence analysis for example), or on the quickest detection of an eventual breakpoint (in the case of industrial data or signal transmission). This project combines both a local and a global approach, in order to take into account not only one region, but several regions at the same time. Hansen in 2006 established a Poisson approximation for the distribution of the number of pairs of segments of two sequences with a similarity level higher than a given threshold. His work is based on the Chen-Stein approximation. We propose here to adapt this work to the case of the number of occurrences of segments in one sequence with a characteristic level higher than a given threshold. We are interested to detect situations where each segment taken alone could be of a lower atypical level, but the segments taken together create an atypical event by their multiple occurrences. These segments are considered to be possibly of unknown and different lengths.

Such tools can be applied in many diverse areas such as : biological sequences, industrial data, quality control, signal transmission as terrestrial activity, medical or public health surveillance, risk management and many others...

Description of the project

Let $(A_i)_{1 \leq i \leq n}$ be a sequence of observations and let f be a score function quantifying a given characteristic of interest that depends on the context of study: hydrophobicity scale for amino acids, observation likelihood under a given hypothesis for signal transmission, shift in number of standard deviations for industrial data for example.

Dembo *et al.* (1992) proposed, with $(A_i)_{1 \leq i \leq n}$ independent and identically distributed variables or a Markov chain, a Poisson approximation for the distribution of the number of regions of interest of a common and fixed length. Nonetheless, results in the context of different and unknown lengths have not been established yet.

Let \mathbb{W} be the Lindley process, also called the CUSUM process, associated to the score sequence $(f(A_i))_{1 \leq i \leq n}$ and defined as follows: $W_0 := 0$ and $W_{i+1} := \max(0, W_i + f(A_{i+1}))$ for $0 \leq i \leq n$. The process \mathbb{W} defines successive non negative excursions, i.e. regions of the sequence with a potentially high level of the studied characteristic. Let $T_0 := 0$

and for $k \geq 0$ $T_{k+1} := \inf\{i \geq T_k : W_i = 0\}$ be the successive stopping times corresponding to the starting points of the successive excursions. Let us denote for $k \geq 1$, $Q^k = \max_{T_{k-1} \leq i < T_k} W_i$ the height of the k -th excursion. We propose to establish an approximation for the law of the number, $C_n(t)$, of regions in the sequence with a property level higher than a given threshold t :

$$C_n(t) = \text{card}\{k : Q^k > t\} .$$

We will model the sequence of observations $(A_i)_{1 \leq i \leq n}$ by a Markov chain, in order to take into account an eventual dependence between the successive observations. Hansen (2006) performed a comparison of two Markovian sequences and considered the number of pairs of segments with a similarity level higher than a given threshold. We also plan to use the Chen-Stein method, by adapting it to the context of a single sequence analysis. The Poisson approximation obtained by the Chen-Stein method is controlled in total variation distance by four quantities

$$d_{TV}(\mathcal{L}(C_n(t)), \mathcal{P}(\lambda_n)) \leq \beta_1(n) + \beta_2(n) + \beta_3(n) + \beta_4(n)$$

where $\mathcal{L}(C_n(t))$ is the distribution of $C_n(t)$, $\mathcal{P}(\lambda_n)$ is a Poisson distribution of parameter λ_n and the $(\beta_i(n))_{1 \leq i \leq 4}$ are bounds taking into account different elements of the model. The Chen-Stein method stands on an appropriate choice of the neighborhood of strong dependence between the $(X_i)_i$. The bound $\beta_1(n)$ is related to the choice of the Poisson parameter λ_n ; the other bounds are related to the chosen dependence neighborhood.

Using mathematical tools from the theory of Markov chains, large deviations and extreme values, we will start to study the bound behavior, by adapting the definitions of the dependence neighborhoods, with the goal of controlling them and showing their convergence to zero when the sequence size goes to infinity.

Key words: Extreme values, Large deviations, random walk, break-point detection, on-line detection, atypical region detection.

Required skills: Markov chain theory, stochastic process, random walk, R software and/or C++ language or Python ; Rigorous, clear communication ; B2-C1 English level.

Schedule of the internship project: a first step consists in reading the key references, principally Hansen 2006, in order to assimilate the method developed. The student will then adapt the definitions and methods to the context of a single sequence analysis, and start to study the bounds of the Chen-Stein approximation. A computational and quantitative study could be developed in order to illustrate the approximation on simulated or real data.

The work could be continued during a doctoral position.

References

1. A. Dembo and S. Karlin. Poisson approximation for r-scan processes. *Annals of Applied Probability*, 2(2): 329-357 (1992).
2. L. H. Y. Chen. Poisson Approximation for dependent trials. *Annals of Probability*, 3(3) :534-545 (1975).
3. N.R. Hansen. Local Alignment in Markov Chains. *Annals of Applied Probability*, 16(3):1262-1296 (2006).

Martingales additives du mouvement brownien branchant

Michel Pain

michel.pain@math.univ-toulouse.fr

Stage de niveau M2 (adaptable au niveau M1)

Durée de 4 à 6 mois entre janvier et septembre 2023

Prolongement en thèse possible

6 novembre 2022

Le mouvement brownien branchant (BBM) est un processus de branchement dans lequel une population de particules se reproduit et se déplace aléatoirement sur la droite réelle. Il est défini comme suit : on commence avec une unique particule en 0 à l'instant 0, puis chaque particule se déplace selon un mouvement brownien pendant une durée de vie de loi exponentielle, avant de donner naissance à un nombre aléatoire d'enfants choisi selon une loi de reproduction fixée, ces enfants apparaissant à la position de leur parent et continuant de même indépendamment des autres particules. Voir la Figure 1 pour une illustration. Notons $\mathcal{N}(t)$ l'ensemble des particules en vie à l'instant t , et $X_u(t)$ la position de la particule u à l'instant t . Il y a de nombreuses motivations à l'étude du mouvement brownien branchant, provenant de la biologie (modèles de populations), de l'informatique (algorithmique), de la physique statistique (verres de spin) ainsi que d'autres modèles probabilistes et du lien avec les équations de réaction-diffusion.

Les *martingales additives* sont définies, pour $\beta \in \mathbb{R}$, par

$$W_t(\beta) = e^{-c(\beta)t} \sum_{u \in \mathcal{N}(t)} e^{\beta X_u(t)}, \quad t \geq 0,$$

où la constante $c(\beta)$ peut être choisie explicitement de sorte que $(W_t(\beta))_{t \geq 0}$ soit en effet une martingale. Comme elle est positive, elle converge presque sûrement vers une limite $W_\infty(\beta)$. Cette quantité joue un rôle important pour décrire le BBM. En effet, en temps long, le BBM grandit comme un cône (voir Figure 1) et la valeur de $W_\infty(\beta)$ décrit la taille de la population dans la direction de pente β .

L'objectif de ce stage est tout d'abord de faire une revue de la littérature et de présenter de manière unifiée les résultats suivants :

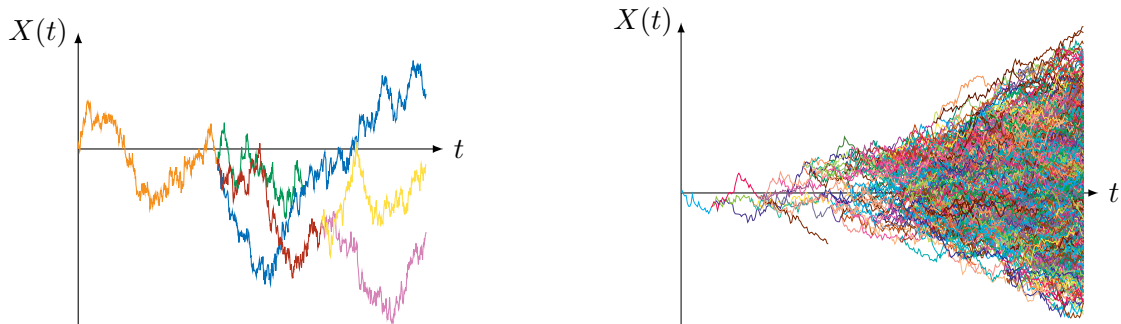


FIGURE 1 – Réalisation d'un mouvement brownien branchant, en temps court et en temps plus long.

- Il existe un paramètre $\beta_c > 0$ explicite tel que

$$\begin{cases} W_\infty(\beta) > 0 \text{ p.s.} & \text{if } |\beta| < \beta_c, \\ W_\infty(\beta) = 0 \text{ p.s.} & \text{if } |\beta| \geq \beta_c. \end{cases}$$

Ce résultat est démontré de différentes manières dans [9, 2, 3, 7]. On se concentre dans la suite sur le cas sous-critique $|\beta| < \beta_c$.

- Pour $|\beta| < \beta_c$, $W_\infty(\beta)$ décrit en effet la taille du BBM dans la direction de pente β . Cela a d'abord été formalisé dans [1].
- Pour $|\beta| < \beta_c$, la queue de distribution de $W_\infty(\beta)$ satisfait $\mathbb{P}(W_\infty(\beta) > x) \sim c(\beta)x^{-\alpha(\beta)}$ quand $x \rightarrow \infty$, où $c(\beta)$ est implicite et $\alpha(\beta)$ est explicite, voir [8]. En particulier, $\alpha(\beta) \rightarrow 1$ quand $\beta \rightarrow \beta_c$, c'est-à-dire que les variables $W_\infty(\beta)$ ont des queues de plus en plus lourdes.
- Les fluctuations de $W_t(\beta)$ autour de sa limite $W_\infty(\beta)$ sont gaussiennes pour β suffisamment petit, puis de loi α -stable avec $\alpha \in (1, 2)$, voir [4, 5, 6].

Pour certains de ces résultats, des méthodes plus récentes permettent une approche plus simple que dans les papiers mentionnés ci-dessus. D'autres résultats ont été montrés uniquement pour la marche aléatoire branchante, la version en temps discret du BBM, il s'agira alors d'adapter les démonstrations au cas du BBM (qui est généralement plus facile).

Selon la durée du stage, différentes extensions ou applications des résultats ci-dessus pourront fournir des directions de recherche. Par exemple, les fluctuations de quantités plus générales décrivant le BBM dans la direction β pourront être obtenues. Le lien avec l'équation F-KPP (une EDP de réaction-diffusion) pourra également être étudié. Une autre direction serait une application de ses méthodes à une question provenant des verres de spins : si l'on tire deux particules à l'instant t proportionnellement à leurs poids dans $W_t(\beta)$, quelle est la probabilité que leur plus proche ancêtre commun soit mort après l'instant at pour $a \in (0, 1)$.

Références

- [1] J. D. Biggins. Growth rates in the branching random walk. *Z. Wahrsch. Verw. Gebiete*, 48(1) :17–34, 1979.
- [2] J. D. Biggins. Uniform Convergence of Martingales in the Branching Random Walk. *Ann. Probab.*, 20(1) :137–151, 1992.
- [3] S. C. Harris. Travelling-waves for the FKPP equation via probabilistic arguments. *Proc. Roy. Soc. Edinburgh Sect. A*, 129(3) :503–517, 1999.
- [4] A. Iksanov and Z. Kabluchko. A central limit theorem and a law of the iterated logarithm for the Biggins martingale of the supercritical branching random walk. *J. Appl. Probab.*, 53(4) :1178–1192, 2016.
- [5] A. Iksanov, K. Kolesko, and M. Meiners. Stable-like fluctuations of Biggins' martingales. *Stochastic Process. Appl.*, 2018.
- [6] A. Iksanov, K. Kolesko, and M. Meiners. Fluctuations of Biggins' martingales at complex parameters. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(4) :2445–2479, 2020.
- [7] A. E. Kyprianou. Travelling wave solutions to the K-P-P equation : Alternatives to Simon Harris' probabilistic analysis. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(1) :53–72, 2004.
- [8] Q. Liu. On generalized multiplicative cascades. *Stochastic Process. Appl.*, 86(2) :263–286, 2000.
- [9] J. Neveu. Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987 (Princeton, NJ, 1987)*, volume 15 of *Progr. Probab. Statist.*, pages 223–242. Birkhäuser Boston, Boston, MA, 1988.

Description of the project “Chen-Stein”

Title: “Distribution of the number of occurrences of segments with a level higher than a threshold”

-

Simona GRUSEA¹, Sabine MERCIER², Toulouse Mathematics Institute,

¹ grusea@insa-toulouse.fr, INSA Toulouse,

² mercier@univ-tlse2.fr, Toulouse 2 University Jean Jaurès

Context and goal

We are interested in highlighting in a sequence, atypical periods or regions corresponding to breakpoints or interferences. Numerous results exist in the literature for highlighting atypical regions or breakpoints, but they generally focus on the most atypical region (in the case of biological sequence analysis for example), or on the quickest detection of an eventual break-point (in the case of industrial data or signal transmission). This project combines both a local and a global approach, in order to take into account not only one region, but several regions at the same time. We propose to establish the probability of observing several segments of a sequence with a level of a given property higher than a threshold: each region taken alone being possibly of a lower atypical level, but forming together an atypical event by their multiple occurrences. These segments are considered to be possibly of different lengths and also of unknown lengths. We expect to obtain a better detection sensitivity.

Such tools can be applied in many diverse areas such as : biological sequences, industrial data, quality control, transmission signal as terrestrial activity, medical or public health surveillance, risk management and many others...

Description of the project

Dembo *et al.* (1992) proposed a Poisson approximation for the distribution of the number of regions of interest of a common and fixed length. Nonetheless, results in the context of different and unknown lengths have not been established yet.

Let $(A_i)_{1 \leq i \leq n}$ be a sequence of observations and let f be a score function quantifying a given characteristic of interest which depends on the context of study: hydrophobicity scale for amino acids, observation likelihood under a given hypothesis for transmission signal, shift in number of standard deviations for industrial data for example.

Let \mathbb{W} be the Lindley process associated to the score sequence $(f(A_i))_{1 \leq i \leq n}$ and defined as follows: $W_0 := 0$ and $W_{i+1} := \max(0, W_i + f(A_{i+1}))$ for $i \geq 0$. The process \mathbb{W} defines successive non negative excursions, i.e. regions of the sequence with a potentially high level of the given property. Let $T_0 := 0$ and for $k \geq 0$ $T_{k+1} := \inf\{i \geq T_k : W_i = 0\}$ be the successive stopping times corresponding to the starting points of the successive excursions. Denote $Q^k = \max_{T_{k-1} \leq i < T_k} W_i$ for $k \geq 1$ the height of the k -th excursion. We propose to establish an approximation for the law of the number $C_n(t)$ of regions in the sequence with a property level higher than a given threshold : $C_n(t) = \text{card}\{k : Q^k > t\}$.

We will model the sequence of observations $(A_i)_{1 \leq i \leq n}$ by a Markov chain, in order to take into account an eventual dependence between the successive observations. The idea is to adapt the work of Hansen (2007), who performed a comparison of two Markovian sequences and considered the number of pairs of segments with a similarity level higher

than a given threshold. We plan to use the Chen-Stein method, by adapting it to the context of a single sequence analysis. The Poisson approximation obtained by the Chen-Stein method is controlled in total variation distance by four quantities

$$d_{TV}(\mathcal{L}(C_n(t)), \mathcal{P}(\lambda_n)) \leq \beta_1(n) + \beta_2(n) + \beta_3(n) + \beta_4(n)$$

where $\mathcal{L}(C_n(t))$ is the distribution of $C_n(t)$, $\mathcal{P}(\lambda_n)$ is a Poisson distribution of parameter λ_n and the $(\beta_i(n))_{1 \leq i \leq 4}$ are bounds taking into account different elements of the model (the choice of the Poisson parameter for $\beta_1(n)$, the dependence for $\beta_2(n)$ for example).

Using mathematical tools from the theory of Markov chains, large deviations and extreme values, the goal is to study the four bounds, and to prove their convergence to zero when the sequence size n goes to infinity.

A first step consists in reading the key references, in order to assimilate the tools and the methods developed. In parallel, the research student will get familiar to the application context. He will then adapt the previous tools and methods to the context of a single sequence analysis. A computational and quantitative study of the achieved work could be developed in order to illustrate the quality of the approximation.

Key words: Extreme values, Large deviations, random walk, break-point detection, on-line detection, atypical region detection.

Required skills: Markov chain theory, stochastic process, random walk, R software and/or C++ language or Python ; Rigorous, clear communication ; B2-C1 English level.

References

- A. Dembo and S. Karlin. Poisson approximation for r-scan processes. *Annals of Applied Probability*, 2(2): 329-357 (1992).
- L. H. Y. Chen. Poisson Approximation for dependent trials. *Annals of Probability*, 3(3) :534-545 (1975).
- S. Grusea and S. Mercier. Improvements on the distribution of maximal segmental scores in a Markovian sequence. *Journal in Applied Probability* 57.1 (March 2020). (<https://hal.archives-ouvertes.fr/hal-01726031>).
- N.R. Hansen. Local Alignment in Markov Chains. *Annals of Applied Probability*, 16(3):1262-1296 (2006).
- S. Mercier. Distribution du score local pour la mise en évidence de segments atypiques au sein de séquences, Université Toulouse III, 3 déc. 2018. Habilitation à diriger les recherches (<https://hal.archives-ouvertes.fr/tel-01944432>).

Compacité par compensation et application à l'étude des modèles nonlinéaires

Stage de Master M1 ou M2

Contact : Radu IGNAT
Institut de Mathématiques de Toulouse
& Institut Universitaire de France
(email : Radu.Ignat@math.univ-toulouse.fr)

La compacité par compensation est un des outils les plus utilisés à la résolution des modèles nonlinéaires de la physique. La situation la plus simple consiste à analyser la limite du produit $u_n \cdot v_n$ pour deux suites des champs de vecteurs u_n et v_n qui convergent faiblement dans l'espace L^2 vers les limites u et v lorsque $n \rightarrow \infty$. Si aucune de ces deux suites ne converge pas fortement dans L^2 , il suffit d'un bon contrôle sur la divergence et le rotationnel de ces champs afin de *compenser* le manque de convergence forte et conclure que $u \cdot v$ est la limite faible du produit $u_n \cdot v_n$ (ce qu'on appelle le lemme div-curl de Murat-Tartar). Le premier but du projet est d'étudier cette méthode de compacité, ainsi que les phénomènes de perte de compacité causés en général par les fortes oscillations des champs de vecteurs.

Le deuxième but du projet est d'illustrer l'utilisation de la compacité par compensation pour des modèles issus de l'élasticité nonlinéaire, l'homogénéisation et les lois de conservation scalaires. Dans cette analyse, la notion de mesure de Young sera fondamentale afin d'analyser le comportement limite des suites $f(u_n)$ pour une nonlinéarité f . En particulier, cette problématique sera étudiée pour le modèle suivant de transition de phase : on considère la fonctionnelle d'énergie

$$E_\varepsilon(u) = \int_\Omega \left\{ \varepsilon |\nabla u|^2 + \frac{1}{\varepsilon} (1 - |u|^2)^2 \right\} dx,$$

définie pour des champs de gradients deux-dimensionnels $u = \nabla \phi : \Omega \rightarrow \mathbb{R}^2$ où $\varepsilon > 0$ est un petit paramètre. Pour une suite $\varepsilon_n \rightarrow 0$, le but est de montrer que les suites de champs de gradients $\{u_n = \nabla \phi_n\}$ d'énergie uniformément bornée (i.e., $E_{\varepsilon_n}(u_n) \leq C$) sont relativement compactes dans L^2 lorsque $n \rightarrow \infty$; de plus, les configurations limites $u = \nabla \phi$ satisfont l'équation eikonale $|\nabla \phi| = 1$ dans Ω . La preuve de ce résultat repose sur une jolie combinaison du lemme div-curl, les mesures de Young et la notion d'entropie pour l'équation eikonale (en lien avec les lois de conservation scalaires).

- Les outils requis pour ce projet sont les cours d'intégration de Lebesgue, le cours de calcul différentiel et le cours d'Analyse fonctionnelle.
- Ce stage de Master pourrait être suivi par une thèse de doctorat.

Références

- [1] L. Ambrosio, C. De Lellis, C. Mantegazza, *Line energies for gradient vector fields in the plane*, Calc. Var. Partial Differential Equations **9** (1999), 327-255.
- [2] A. De Simone, R.V. Kohn, S. Müller, F. Otto, *A compactness result in the gradient theory of phase transitions*, Proc. Roy. Soc. Edinburgh Sect. A **131** (2001), 833-844.
- [3] L.C. Evans, *Weak convergence methods for nonlinear partial differential equations*, CBMS Regional Conference Series in Mathematics, 74. AMS, Providence, 1990.

- [4] R. Ignat, A. Monteil, *A DeGiorgi type conjecture for minimal solutions to a nonlinear Stokes equation*, Comm. Pure Appl. Math. **73** (2020), 771-854.
- [5] R. Ignat, C. De Lellis, *A regularizing property of the 2D-eikonal equation*, Comm. Partial Differential Equations **40** (2015), 1543-1557.
- [6] L. Tartar, *Compensated compactness and applications to partial differential equations*, Heriot-Watt Symposium, Vol. IV, Pitman, Boston, 1979.

Unicité et minimalité des solutions vortex dans des modèles variationnels

Stage de Master M1 ou M2

Contact : Radu IGNAT
Institut de Mathématiques de Toulouse
& Institut Universitaire de France
(email : Radu.Ignat@math.univ-toulouse.fr)

Les symétries sont au cœur de la description mathématique de la nature. Les lois physiques sont invariantes par certains changements de référentiels, ce qui impose certaines exigences de symétrie aux équations décrivant ces lois. Il y a des nombreuses questions ouvertes dans la théorie de Ginzburg-Landau pour la supraconductivité (sur la minimalité des solutions symétriques correspondant à la singularité de vortex ou filament de vortex, ou à la lattice triangulaire d'Abrikosov), dans les films ferromagnétiques sur la symétrie des parois magnétiques ou skyrmions, ou bien, dans les cristaux liquides sur la structure des défauts point, ligne ou interface.

Le projet se concentrera sur de la minimalité de la solution vortex dans le modèle de Ginzburg-Landau. Pour les champs de vecteurs $u : B^N \rightarrow \mathbb{R}^N$ définis sur la boule unité en dimension $N \geq 2$, on considère la fonctionnelle de Ginzburg-Landau :

$$E_\varepsilon(u) = \int_{B^N} \left(|\nabla u|^2 + \frac{1}{\varepsilon^2} (1 - |u|^2)^2 \right) dx$$

avec la donnée au bord $u(x) = x$ pour $x \in \partial B^N = \mathbb{S}^{N-1}$ correspondant au vortex de degré topologique un. Il est bien connu l'existence d'une unique solution symétrique de la forme $u_\varepsilon(x) = f_\varepsilon(|x|) \frac{x}{|x|}$ pour $x \in B^N$. La conjecture est de montrer que u_ε est l'unique minimiseur de E_ε pour tout $\varepsilon > 0$. Si la question est facile pour $\varepsilon > \varepsilon_0$ grand (engendrant la stricte convexité de E_ε), la question est délicate pour ε petit. La littérature étant très riche sur ce sujet, je rappelle seulement les résultats suivants : Pacard-Rivière [7] ont montré la conjecture en dimension $N = 2$ pour $\varepsilon < \varepsilon_1$ petit, mais la question reste ouverte pour $\varepsilon \in [\varepsilon_1, \varepsilon_0]$. Dans [3], la conjecture a été montrée en dimension $N \geq 7$ et pour tout $\varepsilon > 0$ par un argument assez simple inspiré des techniques de décomposition de Hardy. Très récemment, il a été prouvé dans [5] que u_ε est un minimiseur *local* de E_ε pour tout $N \geq 2$ et tout $\varepsilon > 0$. Le but est maintenant d'étendre les techniques dans [5] afin de montrer d'abord la conjecture pour $\varepsilon \leq \varepsilon_N$ petit en toute dimension $2 \leq N \leq 6$ et ensuite analyser le "trou" $\varepsilon \in [\varepsilon_N, \varepsilon_0]$.

Même type de question se pose pour les champs d'aimantation $m = (m_1, \dots, m_{N+1}) : B^N \rightarrow \mathbb{S}^N$ minimisant l'énergie

$$F_\varepsilon(m) = \int_{B^N} \left(|\nabla m|^2 + \frac{1}{\varepsilon^2} m_{N+1}^2 \right) dx$$

sous la condition au bord $m(x) = (x, 0)$ pour $x \in \partial B^N = \mathbb{S}^{N-1}$. Plus précisément, montrer que les deux solutions de vortex magnétique de la forme $m_\varepsilon = (g_\varepsilon(|x|) \frac{x}{|x|}, \pm \sqrt{1 - g_\varepsilon^2(|x|)})$ sont les uniques minimiseurs globaux de F_ε pour tout $\varepsilon > 0$ en dimension $2 \leq N \leq 6$. En effet, en dimension $N \geq 7$, il est connu que la solution harmonique $\bar{m}(x) = (\frac{x}{|x|}, 0)$ à valeurs \mathbb{S}^N est l'unique minimiseur de F_ε pour tout ε ; un pas en avant a été fait dans [5], où il est prouvé que m_ε est un minimiseur *local* de F_ε pour tout $2 \leq N \leq 6$ et tout $\varepsilon > 0$.

- Les outils requis pour ce projet sont les cours d'intégration de Lebesgue, le cours de calcul différentiel et le cours d'Analyse fonctionnelle.

- Ce stage de Master pourrait être suivi par une thèse de doctorat.

Références

- [1] F. Bethuel, H. Brezis, F. Hélein, *Ginzburg-landau vortices*, Birkhäuser, Boston, 1994.
- [2] F. B. Hang, F. H. Lin, *Static theory for planar ferromagnets and antiferromagnets*, Acta Math. Sin. (Engl. Ser.) **17** (2001), 541-580.
- [3] R. Ignat, L. Nguyen, V. Slastikov, A. Zarnescu, *Uniqueness of degree-one Ginzburg-Landau vortex in the unit ball in dimensions $N \geq 7$* , C. R. Math. Acad. Sci. Paris **356** (2018), 922-926.
- [4] R. Ignat, L. Nguyen, V. Slastikov, A. Zarnescu, *On the uniqueness of minimisers of Ginzburg-Landau functionals*, Ann. Sci. Éc. Norm. Supér. **53** (2020), 589-613.
- [5] R. Ignat, L. Nguyen, *Local minimality of \mathbb{R}^N -valued and \mathbb{S}^N -valued Ginzburg-Landau vortex solutions in the unit ball B^N* , preprint.
- [6] P. Mironescu, *Les minimiseurs locaux pour l'équation de Ginzburg-Landau sont à symétrie radiale*, C. R. Acad. Sci. Paris Sér. I Math. **323** (1996), 593-598.
- [7] F. Pacard, T. Rivière, *Linear and nonlinear aspects of vortices. The Ginzburg-Landau model*, Birkhäuser, Boston, 2000.

Analysis of singularities in variational models

Master M2 project

Advisor : Radu IGNAT
Institut de Mathématiques de Toulouse
& Institut Universitaire de France
(email: Radu.Ignat@math.univ-toulouse.fr)

The topic of this Master project (M2) lies in the area of nonlinear partial differential equations (PDE) and calculus of variations. The aim is to introduce the student to the challenging research field of singular phenomena arising in variational principles. The Landau-Lifshitz-Gilbert equation (from micromagnetics) and the Landau-de Gennes equation (from liquid crystals) are two paradigms for the models we address. Each problem involves topological information, as order parameters tend (by a penalty) to take values into a nontrivial manifold (for instance the unit sphere or the projective space). The models are also multiscale and can be analyzed for large or small parameter regimes by means of asymptotic analysis. The formation of singularities in certain limits is due either to topological constraints or to the failure of some critical functional embeddings, and there is substantial overlap in some of the analytic difficulties. Therefore, the connecting thread is given by the analysis of singular structure of vector fields satisfying certain constraints (their target is a certain manifold, they can be divergence-free or curl-free).

There are three possible directions:

- Pattern formation in micromagnetics.
- Analysis of defects in liquid crystals.
- Singularities of vector fields with constraints.

The Master (M2) thesis can be continued with a PhD thesis.

The student is expected to have attended a course in Functional Analysis, Theory of distributions, Elliptic PDEs.

References

- [1] A. DeSimone, R. V. Kohn, S. Müller and F. Otto, *Recent analytical developments in micromagnetics*, in: Giorgio Bertotti, Isaak Mayergoyz (Eds.), *The science of Hysteresis*, vol. 2, Elsevier, Academic Press, 2005, 269-381 (Chap. 4).
- [2] R. Ignat, *Singularities of divergence-free vector fields with values into \mathbb{S}^1 or \mathbb{S}^2 . Applications to micromagnetics*, *Confluentes Mathematici* **4** (2012), 1-80.
- [3] R. Ignat, L. Nguyen, V. Slastikov, A. Zarnescu, *Symmetry and multiplicity of solutions in a two-dimensional Landau-de Gennes model for liquid crystals*, *Arch. Ration. Mech. Anal.* **237** (2020), 1421-1473.
- [4] W. E, *Nonlinear continuum theory of smectic-A liquid crystals*, *Arch. Rational Mech. Anal.* **137** (1997), 159-175.

Stage L3-M1 : Sur la stabilité des ondes périodiques de KdV

Contact : Pascal NOBLE : pascal.noble@math.univ-toulouse.fr

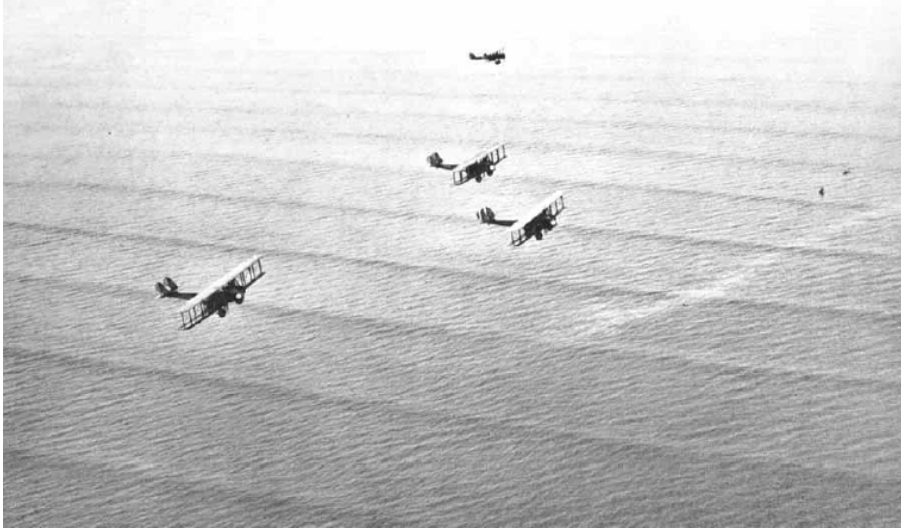


FIGURE 1 – Ondes périodiques dans l’océan

L’équation de Korteweg de Vries est un modèle simplifié de propagation de vagues dans des écoulements de “faible” profondeur. Si on note u l’élévation de la surface du fluide par rapport à un état de référence, l’équation s’écrit

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) + \frac{\partial^3 u}{\partial x^3} = 0.$$

Cette équation contient les deux ingrédients principaux permettant de décrire correctement la propagation des vagues. D’une part, elle est non linéaire : la vitesse de propagation dépend de l’amplitude de la vague. D’autre part, elle est “dispersive” : la vitesse de propagation dépend de la longueur d’onde de la vague.

Cette équation possède des solutions de type ondes progressives ($u(t, x) = U(\xi)$ où $\xi = x - ct$) qui sont soit des ondes périodiques (il existe $X > 0$ tel que $U(\xi + X) = U(\xi)$ pour tout $\xi \in \mathbb{R}$) soit des ondes solitaires ($\lim_{\xi \rightarrow \pm\infty} U(\xi) = \bar{u}$ où \bar{u} est une constante). L’objectif de ce stage est d’étudier la stabilité spectrale des ondes périodiques de (KdV). Le problème mathématique se formule ainsi : étant donné U une onde périodique de (KdV) de vitesse c et de période X , trouver $\lambda \in \mathbb{C}$, $\xi \in \mathbb{R}$ et une fonction v définie sur $[0, X]$ telle que

$$v'''(\xi) + ((U - c)v)' = \lambda v, \quad v(\xi + X) = e^{i\xi X} v(\xi).$$

La description complète des couples λ, ξ est désormais connue et repose sur des outils d’analyse fonctionnelle (spectre d’un opérateur compact) et de théorie des équations différentielles (théorie de Floquet).



Jean-François Coulombel & Grégory Faye
Institut de Mathématiques de Toulouse
CNRS, UMR 5219
Université Paul Sabatier
Toulouse

Proposition de stage de M1

Titre : Comportement en temps long de schémas numériques dispersifs

Résumé. Le but du stage est d'aborder l'étude de schémas aux différences finies pour l'équation de transport, et notamment leur comportement en temps grand. Ces schémas mènent à l'étude d'opérateurs de convolution de la forme :

$$L_a \quad : \quad (u_j)_{j \in \mathbb{Z}} \longmapsto \left(\sum_{\ell \in \mathbb{Z}} a_\ell u_{j-\ell} \right)_{j \in \mathbb{Z}} .$$

Le comportement en temps grand du schéma revient à étudier les puissances L_a^n pour n grand. Quand les nombres a_j sont positifs de somme 1, l'asymptotique est régie par le théorème centrale limite (comme illustré sur la gauche de la Figure 1). On s'intéressera ici aux cas où le schéma numérique développe des oscillations, comme pour les équations aux dérivées partielles dispersives (comme illustré sur la droite de la Figure 1). On tâchera notamment d'améliorer le théorème de la limite locale fourni par E. Randles et L. Saloff-Coste en tirant parti de la dissipation inhérente au schéma numérique. L'implémentation sur machine (qui est simple) permettra d'illustrer les résultats obtenus et/ou de les prédire.

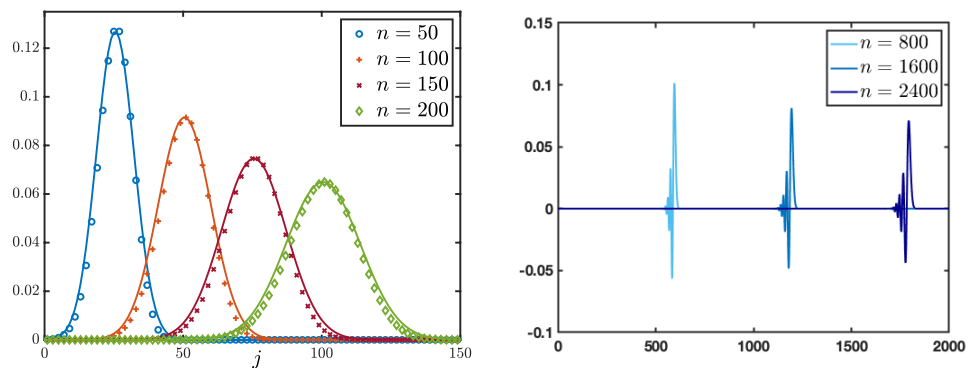


FIGURE 1 – Sur la gauche : le comportement gaussien de type “théorème centrale limite”. Sur la droite : le comportement oscillant propre aux schémas dispersifs.

Prérequis : pour aborder ce stage, il est souhaitable de maîtriser les bases de la théorie de l’intégration (espaces L^p , convolution...), de l’analyse de Fourier (séries et transformée de Fourier, égalité de Parseval-Bessel, théorème de Plancherel...) ainsi que de l’analyse complexe.

Contacts : jean-francois.coulombel@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~jcoulomb/>

gregory.faye@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~gfaye/>

Lieu : le stage s’effectuera à l’institut de mathématiques de Toulouse

Durée : entre 2 à 4 mois d’Avril à Juillet 2023

Gratification : une gratification de stage est possible.

Références : voici quelques références que l’on pourra consulter au cours du mémoire.

- G. W. Hedstrom, *The near stability of the Lax-Wendroff method*, Numerische Mathematik, 1965.
- P. Diaconis et L. Saloff-Coste, *Convolution powers of complex functions on \mathbb{Z}* , Mathematische Nachrichten, 2014.
- E. Randles et L. Saloff-Coste, *On the convolution powers of complex functions on \mathbb{Z}* , Journal of Fourier Analysis and Applications, 2015.
- L. Cœuret, *Local limit theorem for complex valued sequences*, preprint, 2022.



Grégory Faye & Jean-Michel Roquejoffre
Institut de Mathématiques de Toulouse
CNRS, UMR 5219
Université Paul Sabatier
Toulouse

Proposition de stage de M1

Titre : Phénomène de propagation pour les équations de réaction-diffusion

Résumé. L'objectif du stage est de se familiariser avec certaines propriétés qualitatives d'une classe d'équations aux dérivées partielles (EDPs) de type parabolique dites de réaction-diffusion. On s'intéressera plus particulièrement à l'équation de Fisher-KPP, pour Kolmogorov, Petrovsky et Piscunov, qui se trouve à l'interface entre les probabilités et les EDPs. Elle intervient notamment dans de très nombreux problèmes de modélisation. L'étude de cette équation, initiée au début des années 1930, a eu une influence majeure dans de très nombreux domaines des mathématiques et a permis d'ouvrir des pistes d'investigations qui se sont avérées très fructueuses. Néanmoins, il reste encore des problèmes ouverts, et notamment la compréhension du comportement en temps grand des solutions pour des données initiales générales. Le stage permettra de se confronter à une question pas encore abordée et d'apporter sa pierre à l'édifice.

Prérequis : notions d'EDO et d'EDPs, bases d'analyse fonctionnelle.

Contact : gregory.faye@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~gfaye/>

jean-michel.roquejoffre@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~roque/>

Lieu : le stage s'effectuera à l'institut de mathématiques de Toulouse

Durée : entre 2 à 4 mois d'Avril à Juillet 2023

Gratification : une gratification de stage est possible.

Descriptif détaillé du projet

Ce projet s'intéresse à l'étude de l'équation de Fisher-KPP qui prend la forme

$$\begin{cases} \partial_t u = \partial_x^2 u + u(1 - u), & x \in \mathbb{R}, \quad t > 0, \\ u(t = 0, x) = u_0(x), & x \in \mathbb{R}, \end{cases} \quad (1)$$

où l'inconnue $u(t, x) \in \mathbb{R}$ représente typiquement une densité de population à l'instant $t > 0$ et à la position d'espace $x \in \mathbb{R}$, et $u_0(x)$ est la densité à l'instant initiale. Le terme de réaction $f(u) = u(1 - u)$ traduit des mécanismes de reproduction et de mort. Il est bien connu que pour cette classe d'équations, les solutions particulières, dites en ondes progressives, c'est à dire de la forme $u(t, x) = \varphi_c(x - ct)$ pour un certain profil φ_c et une vitesse $c \in \mathbb{R}$, constituent une brique de base pour la compréhension du comportement en temps grand de (1). Plus précisément le couple (φ_c, c) est cherché comme solution du problème suivant

$$\begin{cases} \varphi_c'' + c\varphi_c' + f(\varphi_c) = 0, & x \in \mathbb{R}, \\ \varphi_c(-\infty) = 1, \quad \varphi_c(+\infty) = 0, & \text{et } 0 < \varphi_c < 1, \end{cases} \quad (2)$$

et pour tout $c \geq 2$, il existe un unique (à translation près) couple solution de (2), voir [1].

L'objectif principal du stage est de comprendre le comportement asymptotique lorsque $t \rightarrow +\infty$ de la solution $u(t, x)$ en fonction de la condition initiale u_0 lorsque cette dernière vérifie

$$\varphi_c(x + m) \leq u_0(x) \leq \varphi_c(x - m), \quad x \in \mathbb{R}, \quad (3)$$

pour un certain $m \in \mathbb{R}$ fixé. Cette question a été résolue dans le cas des vitesses sur-critiques [2], c'est à dire pour $c > 2$. Il s'agira donc dans un premier temps de comprendre la méthode de preuve de [2], et dans un second temps d'explorer des pistes pour attaquer le problème critique, c'est à dire pour $c = 2$.

Références

- [1] D.G. Aronson and H.F. Weinberger. *Multidimensional nonlinear diffusion arising in population genetics*. Advances in Mathematics, vol 30, 1978, pp. 33–76.
- [2] M. Bages, P. Martinez, and J.M-Roquejoffre. *How travelling waves attract the solutions of KPP-type equations*. Transactions of the AMS, vol. 364, no 10, 2012, pp. 5415–5468.

Stage M1/M2

Autour des applications harmoniques / On harmonic maps

[Niveau et durée adaptables, dans la période mars-juillet 2023, possibilité de continuer en thèse]

Xavier Lamy, Institut de Mathématiques de Toulouse, xlamy@math.univ-toulouse.fr

L'application singulière $u_*: B_1 \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $x \mapsto x/|x|$, est à valeurs dans la sphère unité \mathbb{S}^2 , et son énergie $\int_{B_1} |\nabla u|^2 dx$ est minimale parmi toutes les applications $u: B_1 \rightarrow \mathbb{S}^2$ telles que $u = u_*$ sur ∂B_1 [1]. Cette propriété de minimalité explique qu'on observe ce type de singularité dans la nature (cristaux liquides, supraconducteurs, micromagnétisme). On parle d'applications harmoniques à valeurs dans \mathbb{S}^2 (ou dans une autre variété) – par analogie avec les fonctions harmoniques réelles qui sont aussi d'énergie minimale

L'objectif principal sera de comprendre, en lisant les deux premiers chapitres de [3], quels types de singularités une application harmonique peut avoir. On pourra ensuite s'intéresser à des énergies plus générales qui soulèvent de nombreuses questions encore ouvertes, avec l'article [2] comme point de départ.

The singular map $u_*: B_1 \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $x \mapsto x/|x|$ takes values in \mathbb{S}^2 , and its energy $\int_{B_1} |\nabla u|^2 dx$ is minimal among all maps $u: B_1 \rightarrow \mathbb{S}^2$ such that $u = u_*$ on ∂B_1 [1]. This minimizing property explains why this type of singularity can be observed in nature (liquid crystals, superconductors, micromagnetism). One speaks about harmonic maps with values into \mathbb{S}^2 (or another manifold) – by analogy with real harmonic functions which are also energy-minimizing.

The main goal will be to understand, by reading the first two chapters of [3], which types of singularities harmonic maps can have. A possible further direction will be to investigate more general energies for which several questions are still open, with the article [2] as a starting point.

Références

- [1] LIN, F.-H. Une remarque sur l'application $x/|x|$. (a remark on the map $x/|x|$). *C. R. Acad. Sci., Paris, Sér. I 305* (1987), 529–531.
- [2] LUCKHAUS, S. Partial Hölder continuity for minima of certain energies among maps into a Riemannian manifold. *Indiana Univ. Math. J.* 37, 2 (1988), 349–368.
- [3] SIMON, L. *Theorems on regularity and singularity of energy minimizing maps. Based on lecture notes by Norbert Hungerbühler.* Basel : Birkhäuser, 1996.

Stage M1/M2

Autour du 19e problème de Hilbert / On Hilbert's 19th problem

[Niveau et durée adaptables, dans la période mars-juillet 2023, possibilité de continuer en thèse]

Xavier Lamy, Institut de Mathématiques de Toulouse, xlamy@math.univ-toulouse.fr

Le 19e problème de Hilbert soulève la question de la régularité des solutions $u: \mathbb{R}^n \rightarrow \mathbb{R}$ de l'équation aux dérivées partielles

$$\nabla \cdot H(Du) = \sum_{j=1}^n \partial_j [H_j(Du)] = 0,$$

où $H: \mathbb{R}^n \rightarrow \mathbb{R}^n$ est de la forme $H = \nabla F$, pour une certaine fonction régulière $F: \mathbb{R}^n \rightarrow \mathbb{R}$, strictement convexe. Si les valeurs propres de $DH = D^2F$ restent dans un segment $[\lambda, \Lambda] \subset]0, \infty[$, cette question a été résolue par Morrey en dimension $n = 2$ dans les années 1930, et par de Giorgi, Nash et Moser en dimension $n \geq 3$ dans les années 1950. Le cas dégénéré où la plus petite valeur propre de $DH = D^2F$ peut s'annuler sur $K \subset \mathbb{R}^n$, est plus délicat et soulève encore des questions ouvertes.

L'objectif premier sera de comprendre l'introduction à ces questions présentée dans [2]. On pourra ensuite l'approfondir par la lecture de [1, 3], puis s'intéresser à l'extension de ce type de résultats à des champs H qui ne sont pas de la forme $H = \nabla F$.

Hilbert's 19th problem raises the question of regularity of solutions $u: \mathbb{R}^n \rightarrow \mathbb{R}$ to the partial differential equation

$$\nabla \cdot H(Du) = \sum_{j=1}^n \partial_j [H_j(Du)] = 0,$$

where $H: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is of the form $H = \nabla F$, for a smooth, strictly convex function $F: \mathbb{R}^n \rightarrow \mathbb{R}$. If the eigenvalues of $DH = D^2F$ stay in a segment $[\lambda, \Lambda] \subset (0, \infty)$, this was solved by Morrey in dimension $n = 2$ in the 1930's, and by de Giorgi, Nash and Moser in dimension $n \geq 3$ in the 1950's. The degenerate case, where the smallest eigenvalue of $DH = D^2F$ vanishes on $K \subset \mathbb{R}^n$, is more complicated and still raises open questions.

The main goal will be to understand the introduction to these questions provided in [2]. This may then lead to a more in-depth study of [1, 3], and eventually to the question of extending such results to fields H which are not of the form $H = \nabla F$.

Références

- [1] DE SILVA, D., AND SAVIN, O. Minimizers of convex functionals arising in random surfaces. *Duke Math. J.* 151, 3 (2010), 487–532.
- [2] MOONEY, C. Hilbert's 19th problem revisited. *Boll. Unione Mat. Ital.* 15, 4 (2022), 483–501.
- [3] SANTAMBROGIO, F., AND VESPRI, V. Continuity in two dimensions for a very degenerate elliptic equation. *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods* 73, 12 (2010), 3832–3841.

Research Internship :

Travelling waves in dispersive models in fluid mechanics :
existence, stability and numerical simulation

Contact : Pascal NOBLE : pascal.noble@math.univ-toulouse.fr



We consider the evolution in time of thin film flows in the presence of dispersive effects. The first model we will consider is the viscous shallow water system modified by a surface tension term :

$$(1) \quad \begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial h u}{\partial x} &= 0, \\ \frac{\partial h u}{\partial t} + \frac{\partial}{\partial x} (h u^2 + p_0(h)) &= \kappa h \frac{\partial^3 h}{\partial x^3} + \nu \partial_x (h \partial_x u), \quad \forall t > 0, \quad \forall x \in I \subset \mathbb{R}, \end{aligned}$$

where h is the fluid height, u is the fluid velocity and $p_0(h) = h^2/2F^2$ is a pressure term whereas κ is a surface tension coefficient and ν a viscosity parameter. If $I = \mathbb{R}$, one can prove that the associated Cauchy problem is well posed for an initial datum in a suitable Sobolev space.

The second model under consideration is the Serre-Green-Naghdi system used in oceanography and hydraulic engineering to model, e.g., the propagation of tidal waves. The system is also a modification of the shallow water equations :

$$(2) \quad \begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial h u}{\partial x} &= 0, \\ \frac{\partial h u}{\partial t} + \frac{\partial}{\partial x} \left(h u^2 + p_0(h) + \frac{h^2}{3} \ddot{h} \right) &= \nu \partial_x (h \partial_x u), \quad \forall t > 0, \quad \forall x \in I \subset \mathbb{R}, \end{aligned}$$

with $\ddot{h} = (\partial_t + u\partial_x)^2 h$. If $I = \mathbb{R}$, one can prove that the associated Cauchy problem is well posed for an initial datum in a suitable Sobolev space.

The main purpose of this internship is to study “simple” nonlinear waves of these models like solitary waves (a propagating wave relating two identical steady states at $\pm\infty$), shocks/heteroclinic waves (namely hydraulic jumps) or periodic travelling waves. These simple waves are usually the building blocks of more complex behaviors.

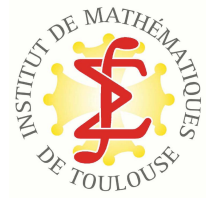
The internship will consist in

1. Proving the existence of travelling waves (solitary waves, periodic waves, fronts depending on the situation) : a first approach consists in studying all small amplitude waves with tools coming from dynamical system (centre manifold reduction, normal forms, bifurcation theory).
2. Studying the spectral stability of these waves by using Evans functions : this consists in showing that eigenvalues are zeros of an analytic function which is not explicit and finding, eventually numerically, those zeros.
3. Carrying out direct numerical simulations to validate the stability results : the question of boundary conditions is then a central question.

This internship comes with a gratification. There is a possibility to continue with a PhD thesis.

References

- B. Barker, M.A. Johnson, P. Noble, L.M. Rodrigues, K. Zumbrun, *Stability of viscous St. Venant roll-waves : from onset to infinite-Froude number limit*, Journal of Nonlinear Science, Vol. 27 (2017), no. 1, p. 285-342
- C. Besse, P. Noble, D. Sanchez, “Discrete transparent boundary conditions for the mixed KDV-BBM equation”, J. Comp. Phys 345 (2017), pp. 484-509.
- P. Noble, J.-P. Vila : “Stability theory for difference approximations of some dispersive shallow water equations and applications to capillary fluids”, SIAM J Numerical Analysis 52 (2014) no 6, 2770-2791.
- D. Bresch, F. Couderc, P. Noble, J.-P. Vila “A generalization of the quantum Bohm identity : Hyperbolic CFL condition for Euler-Korteweg equations”, CRAS Mathématiques 2015.



Grégory Faye & Jean-Michel Roquejoffre

Institut de Mathématiques de Toulouse

CNRS, UMR 5219

Université Paul Sabatier

Toulouse

M2 Internship Proposal

Title : Acceleration phenomena for reaction-diffusion equations set on domains with a line of fast diffusion

Internship summary. The overall objective of this internship is to investigate the qualitative properties of a class of systems of parabolic partial differential equations which couple a reaction-diffusion equation on a half-plane (or a band) with another reaction-diffusion equation on its boundary via a non homogeneous Robin boundary condition. Such types of systems naturally appear in the modeling of population dynamics, for example in ecology [1] or epidemiology [2]. More specifically, we will focus on the case where the diffusion on the boundary is much larger than the diffusion in the half-space and investigate some acceleration phenomena induced by this line of fast diffusion.

Prerequisites : some basic PDE courses.

Contact : gregory.faye@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~gfaye/>

jean-michel.roquejoffre@math.univ-toulouse.fr

<https://www.math.univ-toulouse.fr/~roque/>

Location : the internship will take place at the Institut de Mathématiques de Toulouse.

Dates : 5 months from March to July 2023

NB : an internship stipend is possible and this internship may lead to a PhD thesis.

Extended description of the project

It is well known and documented that transportation networks can enhance biological invasions and epidemics. This project aims at studying a class of partial differential equations which precisely models such effects. More specifically, we shall work within the framework of reaction-diffusion equations and investigate during the internship the effects of a line of fast diffusion on the boundary of the domain on the spreading properties of the overall system. The model under study reads as follows

$$\left\{ \begin{array}{l} \partial_t u = d\Delta u + f(u), \quad x \in \mathbb{R}, \quad y > 0, \quad t > 0, \\ \partial_t v = D\partial_x^2 v + \nu u(t, x, 0) - \mu v, \quad x \in \mathbb{R}, \quad t > 0, \\ -d\partial_y u(t, x, 0) = \mu v - \nu u(t, x, 0), \quad x \in \mathbb{R}, \quad t > 0, \end{array} \right. \quad (1)$$

and can be represented schematically by the following diagram

$$\begin{array}{r} \partial_t u = d\Delta u + f(u) \\ \hspace{15em} x \in \mathbb{R} \\ \hspace{15em} y > 0 \\ -d\partial_y u(t, x, 0) = \mu v - \nu u(t, x, 0) \\ \hline \partial_t v = D\partial_x^2 v + \nu u(t, x, 0) - \mu v \\ \hspace{15em} x \in \mathbb{R} \end{array}$$

for some strictly positive parameters $D, d, \nu, \mu > 0$ and $\Delta = \partial_x^2 + \partial_y^2$. In (1), u can be interpreted as the density of a given species that can move in the half-plane $\Omega = \{x \in \mathbb{R}, y > 0\}$, and may leave the half-space Ω to the real line \mathbb{R} and become v . The term $f(u)$ traduces some nonlinear effects such as birth, death and competition among the individuals of the species. Note that the diffusion coefficients in the half-space and on the boundary \mathbb{R} are allowed to be different, and actually we will be mainly interested in the regime where $D \gg d$. The non homogeneous Robin boundary condition is really natural in the sense that without nonlinear effects (that is with $f(u) = 0$), it is the only natural boundary condition which ensures the preservation of the total mass. Let us finally note that many generalizations (and sophistications) of the above PDE system (1) have been considered in the literature with different reaction terms, nonlocal forms of diffusion or different types of geometries to name a few.

The first step of the internship will consist in a review of the existing literature on the subject, starting with the seminal work [1] and the follow up papers. As previously emphasized, the key objective will be to understand the long time dynamics of (1) in the regime where $D \gg d$. This is a very relevant question both from the modeling and mathematical point of view since one expects nontrivial transitions in the dynamics typically of slow/fast nature (see [3] in the case of an ignition nonlinearity). Depending on the student interests and strengths the next steps of the internship will be *à la carte* and can take various colorations ranging from the exploration of new models (predator-prey, deadly field, epidemic spreading, etc...), to an in depth numerical analysis of the

problem or to the more specific theoretical study of acceleration phenomena in the regime of large diffusion along the boundary.

Références

- [1] H. Berestycki, J.-M. Roquejoffre and L. Rossi. *The influence of a line of fast diffusion in Fisher-KPP propagation*. J. Math. Biol. 66 (2013), pp. 743-766.
- [2] H. Berestycki, J.-M. Roquejoffre and L. Rossi. *Propagation of epidemics along lines with fast diffusion*. Bulletin of Mathematical Biology 83.1 (2021), pp. 1-34.
- [3] L. Dietrich and J.-M. Roquejoffre. *Front propagation directed by a line of fast diffusion : large diffusion and large time asymptotics*. Journal de l'École Polytechnique, Mathématiques, 4 (2017), pp. 141-176.

Stage M2 :

Conditions aux limites transparentes pour les systèmes hyperboliques avec relaxation : applications à la génération d'ondes pour l'océanographie.

Contact :

1. Institut de Mathématiques de Toulouse : Pascal NOBLE
(pascal.noble@math.univ-toulouse.fr),
2. Service Hydrographique et Océanographique de la Marine : Remy Baraille,
(remy.baraille@shom.fr)

Les simulations numériques de phénomènes de propagation de vagues requiert de travailler sur des domaines bornés. Pour réaliser des simulations réalistes, la question des conditions à imposer aux bords du domaine de calcul se pose alors et doit remplir deux objectifs :

1. laisser passer sans réflexion les ondes de propagation sortantes,
2. introduire un forçage pertinent pour les ondes de propagation entrantes.

Deux grandes classes de méthodes coexistent dans la simulation de phénomènes de propagation d'ondes. La première consiste à modifier les équations dans une zone artificielle supplémentaire entourant le domaine d'intérêt dans laquelle les ondes sortantes se propagent tout en étant amorties. C'est la méthode des couches parfaitement raccordées ("perfectly matched layer" ou PML en anglais). Une modification de la méthode permet aussi de traiter les ondes entrantes : cela revient à mettre en place une méthode de relaxation dans la couche artificielle. L'intérêt de la méthode est qu'elle est en général très simple à implémenter. Cependant, selon la nature du phénomène de propagation (dispersion, anisotropie), le système modifié peut être instable et donc générer des instabilités numériques (voir [1] et les références associées). Une deuxième méthode consiste à obtenir des conditions aux limites adaptées directement sur le bord du domaine qui absorbent ou laissent passer parfaitement les ondes sortantes : ce sont des conditions aux limites transparentes ou absorbantes. Ce sont des conditions en général non locales de type Dirichlet-Neumann (relations algébriques entre les dérivées de la fonction sur le bord du domaine). Ces méthodes ont de bonnes propriétés de stabilité [2]. Elles sont cependant un peu plus complexes à implémenter sur les problèmes multi-dimensionnels (gestion des coins dans un domaine rectangulaire) et assez coûteuses (stockage des données aux bords). Un des enjeux pour en améliorer la performance consiste donc à en produire des versions simplifiées.

L'objectif de ce stage est d'examiner le problème des conditions aux limites transparentes ou absorbantes ainsi que des conditions de forçage pour des systèmes hyperboliques de lois de conservation avec relaxation. Ces systèmes sont de la forme :

$$\frac{\partial U}{\partial t} + \operatorname{div}(F(U)) = \frac{1}{\varepsilon}G(U), \forall t > 0, \quad \forall x \in \mathbb{R}^d,$$

où $U(t, x) \in \mathbb{R}^n$ pour tout $(t, x) \in \mathbb{R}_*^+ \times \mathbb{R}^d$ et $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ sont des fonctions de classe \mathcal{C}^1 . Ce formalisme a été introduit récemment pour approcher des modèles de propagation de vagues. Un des systèmes les plus connus et très utilisé est le système de Serre-Green-Naghdi qui s'écrit, en une dimension d'espace :

$$(1) \quad \begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial h u}{\partial x} &= 0, \\ \frac{\partial h u}{\partial t} + \frac{\partial}{\partial x} \left(h u^2 + p_0(h) + \frac{h^2 \ddot{h}}{3} \right) &= 0, \quad \forall t > 0, \quad \forall x \in I \subset \mathbb{R}, \end{aligned}$$

avec $\ddot{h} = (\partial_t + u \partial_x)^2 h$. La gestion des conditions aux limites pour ce type de modèle est assez délicate [3, 2], notamment à cause de la présence de dérivées d'ordre 3. Récemment, un modèle hyperbolique avec relaxation a été proposé pour approcher le modèle de Serre-Green-Naghdi [5]. Une version uni-dimensionnelle (sans topographie) de ce système s'écrit :

$$(2) \quad \begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial}{\partial x} (hu) &= 0, & \frac{\partial(hu)}{\partial t} + \frac{\partial}{\partial x} \left(hu^2 + \frac{gh^2}{2} + hp \right) &= 0, \\ \frac{\partial(hw)}{\partial t} + \frac{\partial}{\partial x} (huw) &= \frac{3}{2}p, & \frac{\partial(hp)}{\partial t} + \frac{\partial}{\partial x} (hup) &= -\frac{1}{\varepsilon^2} \left(2w + h \frac{\partial u}{\partial x} \right). \end{aligned}$$

L'objectif de ce stage consistera :

1. Obtenir, analyser et implémenter des conditions aux limites transparentes pour des modèles simples de systèmes hyperboliques avec terme source (système de Jin-Xin, système de Saint Venant). On s'inspirera de la méthodologie mise en place dans [4] pour l'équation des ondes amortie.
2. Obtenir formellement des conditions pour le système (2) et implémenter des cas simplifiés.

Le stage se déroulera à l'INSA de Toulouse. Il est financé par le SHOM avec une possibilité de poursuite en thèse également financée par le SHOM.

Références

- [1] Besse, C., Gavriluk, S., Kazakova, M. & Noble, P. (2022). Perfectly Matched Layers Methods for Mixed Hyperbolic-Dispersive Equations. *Water Waves*, 1-31.
- [2] Kazakova, M., & Noble, P. (2020). Discrete Transparent Boundary Conditions for the Linearized Green-Naghdi System of Equations. *SIAM Journal on Numerical Analysis*, 58(1), 657-683.
- [3] Lannes, D., & Weynans, L. (2020). Generating boundary conditions for a Boussinesq system, *Nonlinearity*, 33(12), 6868.
- [4] Nguyen, T.H.T. (2020) Numerical Approximation of boundary conditions and stiff source terms in hyperbolic equations, *Thèse Université Rennes 1*.
- [5] Richard, G.L. (2021). An extension of the Boussinesq-type models to weakly compressible flows, *European Journal of Mechanics-B/Fluids* 89, 217-240.

**MASTER 2 INTERNSHIP IN APPLIED MATHEMATICS:
DISCOVERING INFLUENTIAL VARIABLES WITH POINCARÉ
INEQUALITIES**

ALDÉRIC JOULIN AND OLIVIER ROUSTANT

Many physical phenomena and industrial systems are now investigated by costly computer codes (also called simulators). For the analysis of such computer models, the Global Sensitivity Analysis (GSA) methodology is an important tool that allows to rank the relative importance of each input of the system on the output variable, cf. [2]. Referring to a probabilistic modeling of the type $Y = f(X)$, where $X = (X_1, \dots, X_d)$ corresponds to the input random vector with independent inputs X_i , the random variable Y is the scalar output and the function f stands for the computer code, the inputs responsible for such uncertainties are determined through some relevant quantities called sensitivity indices.

Among the sensitivity indices appearing in the literature, variance-based indices (or Sobol indices) are often preferred by practitioners. However, their estimation requires numerous computations. When the derivatives of f are available (because either directly provided by the computer code, or cheap-to-compute), other sensitivity indices, called DGSM (Derivative-based Global Sensitivity Measure), can be used as a proxy. As observed in [3] and further studied in detail in [4], Sobol indices and DGSM are connected by a Poincaré-type inequality. In other words, when it is satisfied, such a functional inequality provides an upper bound of the variance-based index by using the derivative-based one, cheaper to compute. As a result, this allows performing a low-cost screening of unessential variables.

To apply these approaches, finding the Poincaré constant is of crucial importance. Since in theory it is quite hard to find it explicitly, some numerical methods are required. For that purpose, the idea is to investigate the spectral interpretation of the Poincaré constant, which is nothing but the inverse spectral gap of a convenient second-order differential operator without constant term. More precisely, if the distribution of the inputs defined on a (compact) connected domain $\Omega \subset \mathbb{R}^k$ (with k the dimension of each input X_i) is encoded by some probability measure μ with Lebesgue density proportional to $\exp(-V)$ on Ω , the spectral gap corresponds to the first non-zero eigenvalue λ of the spectral problem

$$\Delta f - \langle \nabla V, \nabla f \rangle = -\lambda f,$$

endowed with Neumann conditions $\langle \nabla f, n \rangle = 0$ at the boundary $\partial\Omega$ (where n is the outer normal vector on $\partial\Omega$). When the inputs are one-dimensional, *i.e.*, $k = 1$, the authors in [4] take advantage of this particular setting for which the theory of Poincaré inequalities on intervals is by now well understood, and test their theoretical results (both

Key words and phrases. Global Sensitivity Analysis, Poincaré inequalities, Eigenvalue problems, Finite elements.

explicit, semi-explicit expressions and bounds) through numerical simulations obtained by discretization according to the finite elements method.

At this stage, the identified tasks are the following:

- To discover the existing works at the interface between GSA and Poincaré inequalities, at least in the one-dimensional case in the spirit of [4].
- To discover the theory around functional inequalities and in particular Poincaré inequalities in higher dimension, according to the reference book [1].
- To understand how the finite elements method might be implemented to obtain an approximation of the optimal Poincaré constants in the one and two dimensional case, as initiated in [4] and [5].

As such this 4-month internship, which will start in March or April and for which we are able to provide a financial support, can be seen as a project at the interface between applied statistics and probability, functional analysis and numerical analysis. Such an internship might lead to a PhD Thesis prepared at the Institut de Mathématiques de Toulouse in the case of a potential financial support.

REFERENCES

- [1] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Grundlehren Math. Wiss., 348, Springer, Cham, 2014.
- [2] B. Iooss and P. Lemaître, A review on global sensitivity analysis methods. In *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, Springer, 2015.
- [3] M. Lamboni, B. Iooss, A.-L. Popelin, and F. Gamboa. Derivative-based global sensitivity measures: general links with Sobol indices and numerical tests. *Math. Comput. Simulat.*, 87:45-54, 2013.
- [4] O. Roustant, F. Barthe and B. Iooss. Poincaré inequalities on intervals - application to sensitivity analysis. *Electronic Journal of Statistics*, 11:3081-3119, 2017.
- [5] C. Steiner. *Sur l'utilisation des relations d'entrelacement dans l'étude des générateurs de Markov auto-adjoints. Application aux inégalités spectrales et fonctionnelles et à l'analyse de sensibilité*. Thèse de doctorat, Institut de Mathématiques de Toulouse, 2022.

(A. Joulin) UMR CNRS 5219, INSTITUT DE MATHÉMATIQUES DE TOULOUSE, UNIVERSITÉ DE TOULOUSE, FRANCE

Email address: <mailto:ajoulin@insa-toulouse.fr>

URL: <http://perso.math.univ-toulouse.fr/joulin/>

(O. Roustant) UMR CNRS 5219, INSTITUT DE MATHÉMATIQUES DE TOULOUSE, UNIVERSITÉ DE TOULOUSE, FRANCE

Email address: <mailto:roustant@insa-toulouse.fr>

URL: <https://olivier-roustant.fr/>

Stage de Master 2 Recherche :
Modèles Théoriques et Numériques de Circulation Océanique

Contacts :

Institut de Mathématiques de Toulouse : P. Noble (pascal.noble@math.univ-toulouse.fr)

Service Hydrographique et Océanographique de la Marine : R. Baraille (remy.baraille@shom.fr)

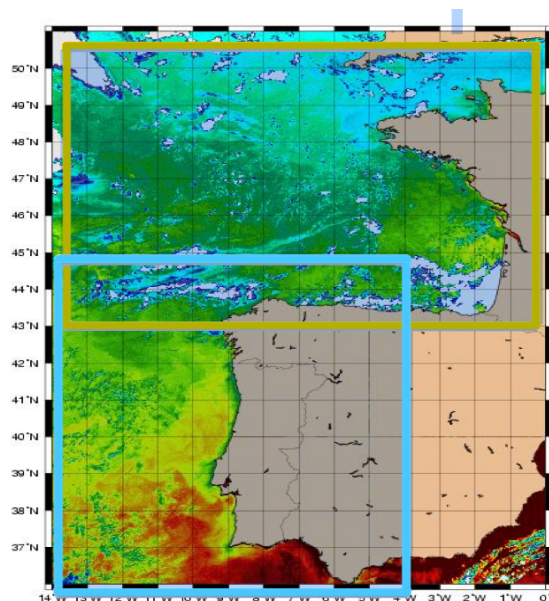


FIGURE 1 – Carte des températures au large du Golfe de Gascogne et de la Péninsule Ibérique. Simulation numérique à l'aide de HYCOM (source : Service Hydrographique et Océanographique de la Marine)

Contexte : HYCOM (HYdrodynamic Coordinates Ocean Modelling) est un modèle numérique de circulation océanique dont les objectifs sont

1. de fournir des conditions aux limites fiables pour des modèles régionaux et côtiers plus précis de courants marins pour décrire par exemple l'évolution des courants et hauteurs d'eau près des côtes et des estuaires.
2. de fournir une description précise de la circulation à l'échelle régionale grâce à la prise en compte de forçages réalistes et d'observations (in situ et satellitales) introduites par assimilation de données.

Ce modèle est utilisé par le Service Hydrographique et Océanographique de la Marine (SHOM) pour la prédiction des courants océaniques dans le Golfe de Gascogne et au large de

la Péninsule Ibérique. HYCOM est composé de différents modules numériques permettant de prendre en compte un grand nombre de phénomènes : son couplage avec un modèle atmosphérique et une résolution à l'échelle du globe permet de prévoir des variations à l'échelle climatique, celui avec un modèle de vagues et une résolution métrique décrit la physique à très courte échéance à l'échelle littorale. La prise en compte d'une telle variété de phénomènes physiques débouche sur des questions mathématiques, numériques et informatiques qui représentent de vrais défis.

Une des difficultés rencontrées est le couplage des conditions aux limites fournies par HYCOM avec le modèle numérique régional utilisé par le SHOM. C'est un modèle à 32 couches de fluides avec d'une part, un modèle d'équations aux dérivées partielles (EDP) portant sur la hauteur totale de l'océan (pondérée par le poids relatif de chaque couche) et la vitesse moyenne permettant de décrire les ondes de gravité externes et les ondes de Rossby et un système d'EDP pour chaque couche qui traduit simplement le transport de masse et de quantité de mouvement dans chaque couche.

L'objectif du stage est de tester et valider différentes améliorations des schémas numériques mis en oeuvre dans HYCOM. Les améliorations porteront d'abord sur une modélisation numérique plus efficace des ondes de gravité externe via un schéma numérique "faible nombre de Mach" (utilisé en général pour modéliser des fluides faiblement compressibles). Une autre piste d'amélioration consiste à intégrer des conditions aux limites adaptées. Les expériences numériques de validation consisteront à simuler la dérive d'un tourbillon sous l'effet de la variation de la force de Coriolis avec la latitude et à représenter les propagations d'ondes internes de gravité.

Les stagiaires auront accès aux machines de calculs et aux codes développés au SHOM. Les sujets proposés sont avant tout une base pour démarrer le travail : les objectifs et thèmes abordés pourront évoluer au gré des goûts du stagiaire. Le travail fourni pendant le stage pourra aisément se poursuivre et déboucher sur une thèse (bourse de thèse SHOM-IMT).

Proposition de Stage

Etude du lien entre prise de psychotropes et issue de grossesse

Tuteurs de stage : C. Chouquet, S. Mercier, IMT
Contact : cecile.chouquet@math.univ-toulouse.fr

Contexte

Ce stage est réalisé dans le cadre d'une collaboration avec l'équipe de pharmaco-épidémiologie de l'INSERM (UMR1027). La cohorte EFEMERIS (Evaluation chez la femme enceinte des médicaments et de leurs risques – <http://www.efemeris.fr/>) est la première base de données française sur l'évaluation du risque médicamenteux pendant la grossesse. Elle contient à ce jour les données de plus 137600 femmes (prescription de médicaments et informations sur leur issue de grossesse). De nombreux projets sont réalisés sur ces données par l'unité 1027 de l'INSERM (C.Damase-Michel, C.Hurault-Delarue).

Données

Pour chaque femme enregistrée dans la cohorte EFEMERIS, on dispose :

- de la mesure d'exposition hebdomadaire par les DDD (defined-daily dose) au cours de la grossesse sous forme de trajectoires d'exposition pour les femmes exposées,
- du type d'issue de grossesse (naissance aboutie et enfant sans problème néonatal, grossesse non aboutie ou aboutie mais avec malformations néonatales),
- des caractéristiques de la femme (âge, nombre de grossesses précédentes, antécédents de fausses-couches, ...),
- des caractéristiques de l'enfant (sexe, allaitement, malformation congénitale, ...).

Objectifs

L'objectif principal du stage est d'étudier les trajectoires d'exposition aux psychotropes pendant la grossesse et leur éventuelle influence sur les types d'issues de grossesse.

Dans un premier temps, on mettra en œuvre une classification des trajectoires pour construire des clusters d'exposition (Méthode Kmeans sur données longitudinales, package R KML).

Dans un deuxième temps, ces clusters seront ensuite pris en compte, avec les caractéristiques de la mère et de l'enfant, pour modéliser le type d'issues de grossesse telles que les pathologies néonatales. Des régressions logistiques seront mises en place, dans le contexte particulier des événements rares (correction de Firth).

En fonction de l'avancée du travail, d'autres analyses classiques pourront être effectuées telles que la description des individus atypiques, ou des individus mal classés. Des méthodes innovantes pourront être envisagées (comme les données augmentées par exemple).

Un programme sous R commenté devra être réalisé pour automatiser le processus d'analyses.

Profil de l'étudiant

- Etudiant en filière Statistique/informatique, niveau M1/M2.
- Compétences en modélisation statistique avec un bon niveau en programmation R.
- Savoir-être : Rigoureux, méthodique, savoir identifier et communiquer sur les problèmes rencontrés au cours du travail

Lieu : Institut de Mathématiques de Toulouse et/ou INSERM – Durée : 3-4 mois (début souhaité avril 2023) – Rémunération au tarif stage