# MA2 2020-2021

## Contents



## 1. Overview of the MA2

Courses are organized in thematic programs, and there will be three such programs, described in detail below. The calendar is roughly as follows.

- Refresher courses. August 31st–mid September 2020
- Basic courses. September–December 2020
- Advanced courses. January–March 2021
- Research internship. April–July 2021

## 2. Number theory and arithmetic geometry

This program consists of six courses. The first three courses will take place during the first semester and cover the fundamentals of algebraic geometry and of the theory of modular forms. The other three courses will take place during the second semester and cover a broader range of techniques of number theory and arithmetic geometry.

**Modular and automorphic forms (Gabriel Dospinescu, 24h).** This course is an introduction to the theory of automorphic forms and representations of real reductive groups. These objects play a fundamental role in modern arithmetic, via the Langlands program. Their study is a mixture of analysis, representation theory and arithmetic, and the goal of the course is to explain several points of view on these objects, as well as to prove some classical results concerning them (for instance some deep finiteness theorems due to Harish-Chandra, or the results of Cartan and Mostow on the structure of real reductive groups). Depending on time, we will discuss a subset of the following set of themes:

- classical modular forms (examples, arithmetic applications, $L$-functions).
- unitary representations of real reductive groups.
- the decomposition of $L^2(\Gamma\backslash G(\mathbf{R}))$ for an arithmetic subgroup $\Gamma$ of a reductive group over $\mathbf{Q}$.
- automorphic forms and adèles.


[1] D. Bump *Automorphic forms and representations*, Cambridge University Press 1997
[2] S. Gelbart *Automorphic forms on adele groups*, Annals of Maths Studies, Vol. 83
[3] H. Jacquet, R. Langlands *Automorphic forms on* $\mathrm{GL}_2$, Lecture notes in maths, 1970

**The fundamental group in arithmetic geometry (Philippe Gille, 24h).** There is a deep analogy between Galois theory and the theory of covers of topological spaces. The guiding principle of the course is to pursue that analogy toward Grothendieck's theory of the fundamental group in algebraic geometry, in connection with Sophie Morel's course. The course will start with the modern viewpoint on Galois theory: étale/Galois algebras, profinite groups, ... Étale finite morphisms and *fibre* functors are the main ingredients of the theory of the fundamental group in algebraic geometry. We shall discuss various examples (e.g. discrete valuation rings, local rings, Dedekind rings, ... ).

[1] J.-B. Bost, F. Loeser, M. Raynaud *Courbes semi-stables et groupe fondamental en géometrie algébrique*, Progress in Math. 187 (2000), Birkhäuser.
[2] A. Grothendieck *SGA 1.*
[3] S. Lang, J.P. Serre *Sur les revêtements non ramifiés des variétés algébriques*, Amer. J. Math. **79** (1957), 319-330.
[4] T. Szamuely *Galois groups and fundamental groups*, Cambridge University Press.
[5] The Stacks Project.

**Modern algebraic geometry (Sophie Morel, 24h).** This course is an introduction to the theory of schemes, aiming to cover the contents of Chapters II and III of [1]. More precisely, here is a list of topics that will hopefully be covered (the order of the topics may change):

- Definition of schemes.
- Global properties of schemes (noetherian, irreducible, reduced, projective schemes).
- Dimension.
- Global properties of morphisms of schemes (morphisms of finite type, separated, proper, projective morphisms).
- Local properties of schemes and of morphisms of schemes (normal and regular schemes, flat, smooth and Ètale morphisms).
- Zariski's main theorem.
- Quasi-coherent sheaves.
- Kähler differentials.
- Cohomology of quasi-coherent sheaves, flat base change, Serre duality.

[1] R. Hartshorne *Algebraic geometry*, Graduate Texts in Mathematics, No. 52 (1977).

**The stack of vector bundles on a curve (Vincent Pilloni, 24h).** Moduli spaces play an important role in arithmetic geometry: moduli spaces of curves, abelian varieties, Shimura varieties, Jacobians, stack of vector bundles on a curve, stack of Shtukas. In particular, the conjectural Langlands correspondence is often formulated with the help of these spaces.

The first part of the course will be an introduction to the language of algebraic stacks. This theory relies on the theory of schemes that will be studied during the first semester. This language is well adapted to the study of many moduli problems that are not representable by schemes.

The second part of the course will be dedicated to a detailed study of the stack of vector bundles on an algebraic curve. We will talk about the following subjects: Hilbert schemes, Harder-Narasimhan filtrations, algebraicity of the stack, smoothness, dimension, affine grassmanian, uniformization...

If time permits, in the last part of the course we will talk about: geometric class field theory, introduction to the geometric Langlands program, Shtukas.

[1] V.G. Drinfeld et C Simpson *B-structures on G bundles and local triviality*, Math Res Letter, 1995.
[2] A. Beauville et Y. Laszlo *Un lemme de descente*, Comptes Rendus Acad Sci Paris, 1995.
[3] A. Grothendieck *Techniques de construction et théorèmes d'existence en géométrie algébrique IV: les schémas de Hilbert.* Séminaire Bourbaki 221, 1960/61.
[4] G. Laumon et L. Moret-Bailly *Champs algébriques*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics, Springer, 2000.
[5] G. Laumon *Travaux de Frenkel, Gaitsgory et Vilonen sur la correspondance de Drinfeld-Langlands*, Séminaire Bourbaki: volume 2001/2002, exposés 894-908, Astérisque, no. 290 (2003), Exposé no. 906, p. 267-284.
[6] C. Sorger *Lectures on moduli of principal G-bundles over algebraic curves.* In School on Algebraic Geometry, (Trieste, 1999), volume 1 of ICTP Lect. Notes, pages 1-57. Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2000.
[7] X. Zhu *An introduction to affine Grassmannians and the geometric Satake equivalence*, IAS/Park City Mathematics Series.
[8] The Stacks Project.

**An introduction to Drinfeld modules (Tuan Ngo Dac and Federico Pellarin, 24h).** Drinfeld modules were invented by Drinfeld (1974) in his proof of the Langlands correspondence for $GL(2)$ over function fields in some special cases. The goal of this course is to introduce these objects and to explain their applications to explicit class field theory and to the Langlands correspondence over function fields.

In the first part we focus on explicit class field theory for global fields of positive characteristic. Let $F$ be such a global function field and let $F^{ab}$ be its maximal abelian extension. The work of David Hayes allows to construct a continuous homomorphism $\rho$ from the Galois group of $F^{ab}/F$ to the idèle class group of $F$. In fact, by using class field theory it is possible to go further, by showing that $\rho$ is an isomorphism of topological groups, the inverse of which is the Artin map. The proof uses normalised Drinfeld modules of rank 1 and their torsion points in a way which is analogous to that of Kronecker-Weber theorem, where the torsion of the multiplicative group allows to generate $\mathbf{Q}^{ab}/\mathbf{Q}$.

In the second part we study Drinfeld modular varieties which classify Drinfeld modules of fixed rank. They play the role of Shimura varieties. We recall the construction and the basic properties of these varieties. Then we review the theory of Drinfeld modules in finite characteristic. As an application, we obtain a formula counting the number of fixed points under the action of the Frobenius and a Hecke operator. Finally we relate this number to the Arthur-Selberg trace formula for some test functions.

[1] D. Goss *Basic Structures of Function Field Arithmetic*, volume 35 of *Ergebnisse der Mathematik und ihrer Grenzgebeite.* Springer-Verlag, Berlin, 1996.
[2] G. Laumon *Cohomology of Drinfeld modular varieties. Part I*, volume 41 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1996. Geometry, counting of points and local harmonic analysis.
[3] G. Laumon *Cohomology of Drinfeld modular varieties. Part II*, volume 56 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997. Automorphic forms, trace formulas and Langlands correspondence, With an appendix by Jean-Loup Waldspurger.

**The Arthur-Selberg trace formula and the local Jacquet-Langlands correspondence (Olivier Taïbi, 24h).** The Arthur-Selberg trace formula is a fruitful tool for the Langlands programme. It gives a "geometric" expression for the trace of a trace function acting on a space of automorphic forms, in terms of orbital integrals. In the lectures we will first study smooth representations of $\mathrm{GL}_2(\mathbf{Q}_p)$ and its harmonic analysis. Then we will prove (a special case of) the trace formula and use it to prove a case of the local Jacquet-Langlands correspondence, between representations of $\mathrm{GL}_2(\mathbf{Q}_p)$ and of $D^\times$ where $D$ is a quaternion algebra over $\mathbf{Q}_p$. This kind of result is motivated by conjectures of Langlands, linking automorphic representations and Galois representations, and useful to prove cases of these conjectures (from "automorphic" to "Galois").

[1] H. Jacquet et R. Langlands *Automorphic forms on* $\mathrm{GL}_2$, Lecture Notes in Mathematics, Vol. 114, 1970.
[2] P. Deligne, D. Kazhdan et M.-F. Vignéras *Représentations des algèbres centrales simples p-adiques*, dans Representations of reductive groups over a local field, Travaux en Cours, 1984.
[3] H. Jacquet, R. P. Langlands *Automorphic forms on* GL(2), Lecture Notes in Mathematics, Vol. 114, Springer, 1970.
[4] P. Deligne, D. Kazhdan, D. and M.-F. Vignéras *Représentations des algèbres centrales simples p-adiques*, Representations of reductive groups over a local field, Travaux en Cours, 33–117, Hermann, Paris, 1984.

## 3. Partial differential equations and applications

This program aims to prepare students for research in the field of theoretical and numerical analysis of problems involving partial differential equations (PDEs). It has three components:

(1) Refresher courses (see §5) in the first two weeks aimed at ensuring a common knowledge base for students from various mathematical backgrounds. These courses are optional but **very strongly advised**.
(2) Three 24h basic courses which offer a broad introduction to the analysis techniques of a large class of partial differential equations.
(3) Four 18h advanced courses on subjects closely related to current research: the analysis of equations with stochastic components, numerical methods for approximation of PDEs, the optimal transport theory for parabolic equations, and the kinetic theory of gases.

The advanced courses will welcome the attendance of PhD students and colleagues.

**Evolution equations (Emmanuel Grenier, 24h).** The aim of these lectures is to study some of the following evolution equations:

(1) reaction-diffusion
(2) hyperbolic systems of conservation laws
(3) parabolic equations
(4) simple kinetic equations
(5) Euler and Navier-Stokes equations.

We will investigate existence, uniqueness, smoothness and qualitative properties of the corresponding solutions.

**Calculus of variations and elliptic equations (Filippo Santambrogio, 24h).** The course will be mainly devoted to the study of the minimizers of integral functionals, their existence, their regularity, and their characterization in terms of solutions of some partial differential equations, but regularity results for the equations themselves will also be presented for their own interest. The course will be roughly structured into 8 classes as follows:

(1) **Introduction and 1D** examples of 1D variational problems (geodesics, brachistochrone, economical growth models) and their applications, tools for existence, Euler-Lagrange equation (both in 1D and in higher dimension).
(2) **Convexity and semicontinuity** conditions to ensure the semicontinuity for the weak Sobolev convergence of integral functionals and applications to existence results. Notions of convex analysis (Fenchel-Legendre transforms, subdifferentials...).
(3) **Convex duality and regularity** duality for some "simple" variational problems and application of convex duality to some $H^1$ regularity result.
(4) $L^p$ **estimates for the Poisson equation**. Proof by interpolation of the result $\Delta u = f$, $f \in L^p \Rightarrow u \in W^{2,p}$.
(5) **Hölder regularity with smooth coefficients**. Morrey-Campanato spaces and applications to the result $\nabla \cdot (a(x)\nabla u) = \nabla \cdot F$, $a, F \in C^{k,\alpha} \Rightarrow u \in C^{k+1,\alpha}$.
(6) **Hölder regularity with bounded coefficients**. Proof by Moser's iterations of the De Giorgi regularity result $\nabla \cdot (a(x)\nabla u) = 0$, $a$ bounded and uniformly elliptic but not smooth $\Rightarrow u \in C^{0,\alpha}$ and applications to the solution of the 19th Hilbert problem.
(7) $\Gamma-$**convergence and examples**. The general theory of the $\Gamma-$convergence for the limits of variational problems and some example, in particular the optimal quantization of measures (aka optimal location problem).
(8) $BV$ **functions, perimeters, and the Modica-Mortola functional**. Few words about the space BV and its role in defining sets of finite perimeter. Proof of the $\Gamma-$convergence of the functionals $\int \varepsilon|\nabla u|^2 + \varepsilon^{-1}W(u)$ towards the perimeter functional.

A detailed bibliography and a list of exercises will be provided. The knowledge of some functional analysis (in particular, compactness for weak-$*$ convergence and Sobolev spaces) and some measure theory is the main prerequisite for the course.

**Approximation by PDEs (Julien Vovelle, 24h).** In this course, we will see how to understand and describe the large scale limit of various discrete evolution systems (random and deterministic) with the help of partial differential equations. This will be the occasion to use and discover some standard tools from the theory of PDEs, of numerical analysis, and of statistical physics.

Contents:

(1) Martingale in continuous time
(2) Discrete conservation laws, systems of interacting particles and their asymptotic description by PDEs
(3) Interacting particles systems; independent random walk; model of random interface

(4) Discrete conservation laws (Finite Volume method); parabolic equations; hyperbolic equations; the case of the linear transport equation: optimal convergence estimate by two different methods (deterministic/probabilistic).

**Stochastic PDEs and their asymptotic behaviour (Alexandre Boritchev, 18h).** This course is on the interface between PDEs and probability theory. It is in some way "complementary" with respect to the basic course given by Julien Vovelle, being concerned with a more specific topic of PDEs with random noise. Nevertheless, the notions of martingales in continuous time, Markov processes, invariant measures, seen in the course of Julien Vovelle, will be used. The main technicalities come from the complex interplay between the regularity of the solution to the deterministic equation and that of the noise. We also give some key notions on long-term behaviour of solutions and on the role of the stationary measure. The latter is a probability measure on a functional space which is the random counterpart of a stationary solution for a deterministic equation.

This is a course which introduces many new concepts: therefore we will follow a plan which underlines the differences and similarities between the finite-dimensional situation (stochastic ODEs) and the infinite-dimensional one (stochastic PDEs).

We use heavily the material from refresher courses, especially the ones on "Basic tools of functional analysis" (convolution, Sobolev spaces. . . ) and "Stochastic tools" (Brownian motion and regularity of its trajectories).

(1) Introduction:
  - Wiener process: a reminder. Regularity of the trajectories.
  - Construction and properties of the Itô integral.

(2) SDEs:
  - What is a stochastic differential equation (SDE)? Two points of view: ODEs with random coefficients and through the Itô integral.
  - Two Markov semigroups: on spaces of probability measures and of continuous bounded functions.
  - Invariant measure and convergence to the equilibrium.
  - Example: the Langevin equation.

(3) An introduction to SPDEs:
  - Itô integral in Hilbert spaces.
  - What is a stochastic partial differential equation (SPDE)? Two points of view: PDEs with random coefficients and through the Itô integral.
  - Two Markov semigroups: on spaces of probability measures and of continuous bounded functionals.
  - Invariant measure and convergence to the equilibrium.
  - Examples: the stochastic heat and Burgers equations.
  - More involved examples (if time allows): the stochastic 2D incompressible Navier-Stokes equation.

**Optimal transport theory and links with parabolic equations (Ivan Gentil, 18h).** Goal of the course: the goal of this course is to introduce the Wasserstein distance between two probability measures. This distance has many recent developments. For instance, this is a natural distance to show the asymptotic behavior of evolutional PDE, asymptotic behavior of stochastic process, to show concentration inequalities, and also to prove that the heat equation is the gradient flow of the Boltzmann entropy with respect to the Wasserstein metric.

We will see in this course analytic properties of the Wasserstein distance and also links between parabolic PDE and optimal transport. Parabolic PDE can be seen in this context as the law of diffusion Markov processes, solutions of SDE driven by a Brownian motion. Table of contents:

(1) Properties of the Wasserstein distance between two probabilities measures.
   (a) Optimal transportation, Monge and Monge-Kantorovich problems.
   (b) Properties of the Wasserstein space, in particular its geodesic property.
   (c) Duality of Kantorovich.
   (d) Brenier's Theorem and proof of the optimal Sobolev inequality by using the optimal transport.
(2) Links between parabolic PDE and Otto calculus.
   (a) Notion of operator's curvature.
   (b) Von Renesse-Sturm's Theorem: equivalence between curvature of an operator and the contraction in Wasserstein distance.
   (c) Heat equation as a gradient flow of the entropy with respect to the Wasserstein metric, introduction of the Otto Calculus.

[1] L. Ambrosio, N. Gigli, and G. Savaré *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2005.
[2] D. Bakry, I. Gentil et M. Ledoux *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2014.
[3] C. Villani *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
[4] C. Villani *Optimal transport*, Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer-Verlag, Berlin, 2009.

**Numerical approximation methods for fluid mechanics (Khaled Saleh, 18h).** A first objective of this course is to present recent methods for the numerical approximation of fluid mechanics models. The considered numerical schemes are used in industrial codes. They are designed so as to mimic the main properties satisfied by the exact solutions: positivity of the density, mass and momentum conservation, energy estimates, etc. A second objective of the course is to establish the convergence of the numerical schemes by following the lines of the theory of existence of weak solutions for the considered PDE models. For this purpose, famous functional analysis results must be adapted to discrete functional spaces. Contents of the course:

(1) PDE models in fluid mechanics: compressible Navier-Stokes and Euler systems. Incompressible models.
(2) Study of the incompressible Stokes model. Well-posedness for weak solutions. Analysis of a finite element numerical scheme: a priori estimates, compactness is assumed at this level, convergence of the approximate solutions towards the exact weak solution.
(3) Discrete functional analysis. We prove continuous/compact embedding results for discrete functional spaces arising from the numerical discretization. These are discrete counterparts to the famous Sobolev continuous embedding and Rellich's compact embedding theorems.

(4) Numerical analysis of a finite volume - finite element numerical scheme for the compressible Navier-Stokes equations: positivity and energy estimates, convergence of the numerical method (compactness, passing to the limit in non-linear terms with only weak convergence) for the stationary model.
(5) Non stationary models. Compactness in time: Aubin-Simon theorem and its discrete counterpart. Study of a numerical scheme for the non stationary Stokes model.

**Kinetic theory (Laure Saint-Raymond, 18h).** In this course, we will introduce basic tools for the mathematical study of kinetic equations. We will then present two important classes of kinetic equations (mean field equations and collisional equations) and study their main features.

(1) Mathematical tools for the analysis of kinetic transport equations
  1.1. A priori estimates in Lebesgue spaces
  1.2. Averaging lemma
  1.3. Dispersion and control of concentrations
  1.4. Boundary conditions and trace estimates
(2) Mean field models
  2.1. Some classical models
  2.2. Weak solutions for the Vlasov-Poisson model
  2.3. Propagation of moments and regularity
  2.4. A uniqueness criterion
  2.5. The mean field approximation: convergence of the empirical measure
  2.6. The mean field approximation: propagation of chaos
(3) Collisional models
  3.1. The Boltzmann equation and its BGK approximation
  3.2. Weak solutions for the BGK equation
  3.3. Propagation of moments and uniqueness
  3.4. On the trend to global equilibrium
  3.5. Renormalization techniques

## 4. Probability and statistics

This program is made of two closely intertwined lecture series in probability (P) and statistics (S), that will provide a strong and interdisciplinary training in the modern mathematics of random phenomena.

Via their choice of lectures (3 out of 4 in the first semester and 4 out of 6 in the second), students may choose to give a stronger focus on the probabilistic or statistical aspects of this program. It is also possible to follow lectures from other programs, especially in PDEs. It is strongly advised to follow the corresponding refresher courses.

**P1: Statistical physics (Christophe Garban, 24h).** In statistical physics, one is interested in physical models made of a large number of microscopic elements which interact together in a simple fashion. The goal is then to understand how come such simple microscopic mechanisms can generate interesting (and surprising!) macroscopic phenomena such as phase transitions or symmetry breaking. This program has lead to the development of an important branch of probability theory and the aim of this course is to give a panorama of the field together with tools and techniques that are used in statistical mechanics. We shall focus on three fundamental models: percolation, Ising model and O(n) spherical spin model. Program of the course:

(1) Percolation • Definition, phase transition • FKG inequality, $p_c = 1/2$ • Exponential decay in the sub-critical regime • Russo-Seynour-Welsh theorem for critical percolation
(2) Ising model • Definition, correlation inequalities • Infinite volume limit, Free energy, phase transition • Low temperature and Peierls argument • Uniqueness at high temperature
(3) Phase transition KT (Kosterlitz-Thouless) Nobel prize in phyics 2016 • Spin models with continuous symmetry ($\sigma_x \in S^d, d \geq 1$) • No symmetry breaking in dimension 2 (Mermin-Wagner theorem) • Gaussian Free Field • Vortices and Coulomb gas • A glimpse of Frölich-Spencer Theorem on the KT transition for the XY model $[\sigma = (\sigma_x)_{x \in Z^2} \in (S^1)^{Z^2}]$

[1] W. Werner, Percolation et modèle d'Ising, Soc. Math. France, 2009.
[2] Y. Velenik, Introduction aux champs aléatoires markoviens et gibbsiens, http://www.unige.ch/math/folks/velenik/Cours/2006-2007/Gibbs/gibbs.pdf.

**P2: Stochastic calculus (Grégory Miermont, 24h).** This lecture series will present some of the most important tools allowing to build and study continuous-time stochastic processes, the central example of which is of course Brownian motion. To this end, we will have to introduce and study semimartingales, a rich class of processes for which one can develop a differential and integral calculus, and set and solve certain type of stochastic differential equations.

Just as for the familiar ordinary differential equations (or PDEs), the motivation to study such stochastic differential equations comes from the goal of understanding the global behavior of random processes by equations describing their infinitesimal behavior. But since we are dealing with random processes, these equations contain a random "noise", which informally is an infinitesimal increment of Brownian motion. The main problem of their study comes from the fact that Brownian motion (and therefore the other processes of interest) have too rough trajectories (nowhere differentiable, for instance) for the usual differential and integral calculus to make sense.

In front of this obstacle, we will develop a notion of stochastic integral, due to Itô. It will give rise to a particular integral calculus, in which Itô's formula acts as an integration by parts (or a fundamental theorem of analysis) of a new kind. This integral calculus will allow us to study the stochastic differential equations for continuous semimartingales, and will shed a new light on these processes, for instance via Lévy's characterization of Browian motion, or the Dubins-Schwarz theorem according to which continuous martingales are appropriate time-changes of Brownian motion. Contents:

(1) Generalities on continuous-time processes
(2) Continuous-time martingales, regularization. Local martingales, semimartingales. Bracket of a continuous semimartingale.
(3) Stochastic integral with respect to a continuous semimartingale.
(4) Itô's formula and applications. The theorems of Lévy, Dubins-Schwarz, Girsanov. Burkholder-Davis-Gundy inequalities.
(5) Stochastic differential equations. The Lipschitz case.
(6) (Time allowing) Continuous-time Markov processes. Generators. Diffusions.

[1] Karatzas-Shreve: Brownian motion and stochastic calculus

[2] Le Gall: Brownian motion and stochastic calculus
[3] Mörters-Peres: Brownian motion
[4] Revuz-Yor: Continuous martingales and Brownian motion
[5] Varadhan: Stochastic processes

**S1: Concentration of measure in probability and high-dimensional statistical learning (Guillaume Aubrun, Aurélien Garivier, Rémi Gribonval, 24h+8h).** This course will introduce the notion of concentration of measure and highlight its applications, notably in high dimensional data processing and machine learning. The course will start from deviations inequalities for averages of independent variables, and illustrate their interest for the analysis of random graphs and random projections for dimension reduction. It will then be shown how other high-dimensional random functions concentrate, and what guarantees this concentration yields for randomized algorithms and machine learning procedures to learn from large training collections. This course will be based on a sample of the classical textbook “Concentration Inequalities” by Boucheron, Massart, Lugosi, and on “High-Dimensional Probability” by Roman Vershynin. Applications to machine learning will rely on “Understanding Machine Learning”, by Shalev-Schwartz and Ben-David.

**S2: Non-parametrics (Irène Ganaz, Clément Marteau, Franck Picard, 24h).** In this course we will focus on recent developments in non parametric statistics, with a special focus on non-parametric model selection and high dimensional statistics. Variable selection through the LASSO (Least Absolute Shrinkage and Selection Operator) has revolutionized high dimensional statistics thanks to the use of a L1-constrained optimization problem that ensures powerful statistical properties. These connections between non-convex optimization and Statistics has been very fruitful from both the applied and theoretical aspects of Machine Learning. A non-negligible part of this course will focus on the theoretical properties (model selection, convergence ...) of penalized estimators, with the use of oracle inequalities as a building block. These penalized methods will be put into perspective with kernel-based estimation and regression, another popular non-parametric strategy that requires fine theoretical and computational calibration (bandwith choice). Some attention may also be payed to related topics such as wavelet transform, multiple testing issues or signal processing on graphs.

**P3 : Branching random walks (Xinxin Chen, 18h).** Branching random walks (BRWs) generalise both the concept of random walks and that of branching processes. Branching Brownian motion (BBM) is a simple example of BRWs. The study of BBM and BRW, not only has its own interests, but also leads to the understanding of other models belonging in the BBM-universality class, such as the 2-dimensional Gaussian free field, the 2-dimensional cover times, the characteristic polynomials of random unitary matrices, etc.

We mainly give an elementary introduction to branching Brownian motion and branching random walk in the real line and describe the spinal decomposition which can be used to obtain the convergence in law of the minimum.

Contents: • Branching Brownian motion and F-KPP equation • Branching random walks and martingales • Spinal decomposition and change of measures • Applications of spinal decomposition: Seneta-Heyde norming; weak convergence of minimum • Other models in BBM-universality class

[1] Bovier, A. Gaussian Processes on Trees: From spin glasses to branching Brownian motion, Cambridge Univ. Press, 2016.

[2] Shi, Z. Branching Random Walks, Ecole d'Eté de Probas de Saint-Flour XLII-2012.

**P4 : Random Graphs (Dieter Mitsche, 18h).** In the last years, complex networks have become central elements in many areas (telecommu- nication networks, internet, neural networks, social networks, propagation of infectious diseases, propagation of rumors, ....). It is a booming area, and it is crucial to develop mathematical models to represent these networks.

A network is often modelled by a random graph, and this course proposes the study of different random graph models, in particular the Erdös-Rényi random graph model, the configuration model and random geometric graphs. A special focus of this course will be given on the threshold of the giant component in different graph models. Contents:

(1) Erdös-Rényi model • Introduction, subgraph count • Local weak convergence • Phase transition, appearance of a giant component • Hamiltonicity
(2) Configuration model • Differential equation method • Emergence of a giant component 3. Random geometric graphs • Euclidean random geometric graphs - emergence of the giant component • Introduction to random hyperbolic graphs

[1] N. Alon, J. Spencer, The probabilistic method, 3rd ed., John Wiley & Sons, 2008.

[2] C. Bordenave, Lecture notes on random graphs and combinatorial optimization, https://www.math.univ-toulouse.fr/~bordenave/coursRG.pdf

[3] A. Frieze, M. Karonski, Introduction to random graphs, CUP, 2015.

[4] M. Penrose, Random geometric graphs, Oxford Univ. Press, 2003.

[5] R. van der Hofstad, Random graphs and complex networks,
Vol. 1, Cambridge Series in Statistical and Probabilistic Mathematics, 2017.
Volume 2, https://www.win.tue.nl/~rhofstad/ NotesRGCN.html
https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf

**P5 : Determinantal processes (Adrien Kassel, 18h).** A determinantal process on a nice topological measured space $S$ is a random discrete collection of points such that the correlation functions – loosely speaking, the density of probability of seeing a finite subcollection of points at given locations – exist, and may be written in the form

$$\rho_m(x_1, \dots, x_m) = \det[(K(x_i, x_j))_{1 \le i,j \le m}] \,,$$

where $K : S^2 \to \mathbb{C}$ is a two-point function, also called a kernel.

There are many examples of stochastic models which give rise to interesting determinantal processes, many of which find their origin in mathematical physics. The corresponding kernel can sometimes be computed rather explicitly, which enables the study of fine properties of the model. We may broadly distinguish two classes of processes according to the topology of $S$: discrete ones (e.g. $S = \mathbb{Z}^d$) and continuous ones (e.g. $S = \mathbb{R}^d$). Examples of discrete processes include random spanning forests on finite and infinite graphs; examples of continuous processes include eigenvalues of certain random matrices of finite or infinite size.

The goal of this course will be to present a theory of determinantal processes, namely to present what is common to these examples beyond their particularities. For that matter, we will focus on the better understood case where $K$ is self-dual, namely when the symmetry $K(x, y) = \overline{K(y, x)}$ holds.

We will start with the case where $S$ is finite, for which a very complete understanding is available. The kernel $K$ is then a Hermitian matrix, and the process is completely described in terms of linear algebra in $\mathbb{C}^S$, and its Euclidean geometry. This allows to explain the link to theoretical physics in quite a transparent way. This part of the theory may easily be extended to the case where $S$ is countable, where now $\ell^2(S)$ is the relevant geometry. To move on to the case of uncountable $S$ requires extra caution, and the dictionary between kernels and Euclidean geometry now needs to be enhanced to the setup of bounded integral operators on the Hilbert space $L^2(S)$.

Examples we will present, at least superficially, include: uniform spanning forests of infinite lattices; zeros of the Gaussian analytic function on the unit disc; eigenvalues of a random Hermitian matrix distributed according to the Gaussian unitary ensemble.

[1] A. Borodin. Determinantal point processes. *Oxford handbook of random matrix theory*, pp 231–249, Oxford Univ. Press, 2011.

[2] R. Lyons. Determinantal probability measures. *Publ. Math. Inst. Hautes Études Sci.*, (98):167–212, 2003.

[3] A. Soshnikov. Determinantal random point fields. *Uspekhi Mat. Nauk*, 55(5(335)):107–160, 2000.

**S3: Mathematical foundations of deep neural networks (Rémi Gribonval, Aurélien Garivier, 18h).** This course will provide a detailed overview of the mathematical foundations of modern learning techniques based on deep neural networks. Starting with the universal approximation property of neural networks, the course will then show why depth improves the capacity of networks to provide accurate function approximations for a given computational budget. Tools to address the optimization problems appearing when training networks on large collections will then be covered, and their convergence properties will be reviewed. Finally, statistical results on the generalization guarantees of deep neural networks will be described, both in the classical underfitting scenario and in the overfitting scenario leading to the so-called “double descent” phenomenon.

**S4: Inverse problems and high dimension (Yohann de Castro (course at Centrale Lyon), complements by Rémi Gribonval, 18h).** Sparsity and convexity are ubiquitous notions in Machine Learning and Statistics. In this course, we study the mathematical foundations of some powerful methods based on convex relaxation: L1-regularisation techniques in Statistics and Signal Processing; Nuclear Norm minimization in Matrix Completion; K-means and Graph Clustering. These approaches turned to be Semi-Definite representable (SDP) and hence tractable in practice. The theoretical part of the course will focus on the guarantees of these algorithms under the sparsity assumption. The practical part of this cours will present the standard SDP solvers of these learning problems.

Keywords: L1-regularisation; Matrix Completion; K-Means; Graph Clustering; Semi-Definite Programming;

**S5: Advanced machine learning theory (Laurent Jacob, Joseph Salmon, 18h).** Choosing an appropriate data representation is a key element in modern machine learning. While most existing learning algorithms manipulate vectors, important data types including webpages, sequences or graphs do not admit a straightforward vectorial representation. Other data types admit a natural vectorial encoding, but the resulting representation is not necessarily appropriate for learning (this is the case for images).

This course will introduce positive definite kernels, a powerful mathematical framework to create, analyze and manipulate data representation. Kernels are functions measuring the similarity between pairs of objects. We will show how they implicitly define a mapping of the data to a Hilbert space, and how this fact can be used to manipulate large or even infinite sets of descriptors. We will specifically discuss how these kernels can be used in supervised and unsupervised learning algorithms, and provide examples of kernels defined on sequences and graphs. Finally, we will consider how this framework can shed some light on convolutional neural networks.

Keywords: positive definite kernel, RKHS, kernel methods, sequences, graphs, libsvm, large scale learning, deep kernel machines.

## 5. Refresher Courses

**Basic tools of functional analysis (Simon Masnou, 16h).**

(1) Duality: Hahn-Banach theorem, weak and weak-* topologies, Lebesgue spaces;
(2) Distributions: weak derivatives, convolution, fundamental solutions of differential operators;
(3) Fourier transform;
(4) Sobolev spaces: embeddings, extension and traces, compactness;
(5) Weak solutions of PDEs;
(6) Spectral analysis in Hilbert spaces.

**Stochastic tools (Grégory Miermont, 10h).**

(1) Discrete time martingales: stopping theorems and convergence. Extensions for continuous time martingales.
(2) Construction of Brownian motion. Regularity of trajectories.
(3) Some properties of Brownian trajectories. Connection with the heat equation.

**Starting with PDEs (Petru Mironescu, 16h).**

(1) Geometric aspects of integration theory. Area, co-area, Jacobians.
(2) The Bochner integral.
(3) Comparison principles for first and second order partial differential equations. A priori estimates.
(4) Perron's method. A potential theoretic point of view of smoothness.
(5) The energy method. Uniqueness and domain of influence.
(6) Semi-group methods in the study of evolution equations. Existence, smoothness and additional smoothness.