

La simulation numérique 1/2

Introduction sur les modèles

Les modèles permettent de décrire une approximation de la réalité

Quand on souhaite comprendre comment fonctionne le monde qui nous entoure, comment il fonctionne, et que l'on veut essayer de prévoir des choses, on cherche à définir des lois qui semblent être celles suivies pour expliquer ce que les expériences montrent.

Ces lois fonctionnent dans un domaine défini qui va correspondre à des échelles d'observation, à des intervalles de températures, etc.

L'ensemble constitué par ces lois et le domaine dans lequel elles s'appliquent forment le modèle que l'on va utiliser pour décrire un phénomène, pour le comprendre et peut-être même prévoir son évolution.

En fonction des connaissances que l'on acquière au cours du temps, ces lois et ces modèles peuvent être rendues plus précis, ils peuvent être étendus à d'autres phénomènes ou même réfutés. L'évolution de la compréhension de la gravité, de l'attraction, etc en est un bon exemple.

Afin de pouvoir mettre en place un modèle la démarche scientifique va être la suivante :

- Avoir une question dont on recherche la réponse (comportement d'une aile, évolution du temps, des galaxies, etc)
- Limiter le champ de travail afin de pouvoir faire des calculs en un temps fini et de ne pas être parasité par des données non pertinentes
- Qu'il représente le phénomène que l'on souhaite observer
- Valider le modèle avec les données connues pour s'assurer que l'on est sur quelque chose de valable
- Qu'il soit réutilisable et reproductible

L'intérêt des modèles ¹ est ainsi de pouvoir permettre différentes choses :

- La prévision des phénomènes afin de prévoir comment ils vont se comporter dans le futur ou soumis à des contraintes spécifiques
- La prévision des interactions entre différents phénomènes où il est nécessaire de savoir comment ceux-ci interagissent
- Valider que les modèles représentent bien ce qu'ils sont censés modéliser

¹http://fr.wikipedia.org/wiki/Mod%C3%A8le_math%C3%A9matique

Comme l'explique Wikipedia ^{2 3} : “On appelle modèle un élément, analogique ou numérique, dont le comportement vis-à-vis d'un phénomène est similaire à celui de l'élément à étudier. Les sorties sont les éléments que l'on veut étudier. Les entrées, paramètres et contraintes sont les éléments dont la variation influe sur le comportement du modèle ; on appelle entrée ceux qui sont commandés par l'expérimentateur, paramètres ceux que l'opérateur choisit de fixer et contraintes ceux qui dépendent d'éléments extérieurs. On appelle simulation l'ensemble constitué par un modèle, les ordres d'entrée, les paramètres et contraintes, et les résultats obtenus. Comme indiqué plus haut les maquettes, prototypes, etc. peuvent être considérés comme des modèles analogiques et les essais, tests, manœuvres, etc. comme des simulations analogiques. Les équations sont des simulations numériques. Aujourd'hui ce terme s'applique essentiellement aux modèles et simulations réalisés sur ordinateur. Dans certains cas on peut réaliser des simulations hybrides, analogiques - numériques, qui intègrent divers éléments dont certains seulement sont représentés par des équations.”

Dans bon nombre de cas, on peut arriver à reproduire les phénomènes en laboratoire afin de les étudier, comprendre si nos modèles sont bons, comment ils se comportent, etc. On va donc parler ici de simulation des phénomènes. On peut reproduire énormément de choses en laboratoire en passant par la chute des pommes pour la gravité, les trous noirs avec du sable, et jusqu'aux comportements des particules les plus élémentaires avec le LHC. Les outils, les méthodes ou mêmes les tailles de labo vont changer en fonction des phénomènes que l'on souhaite étudier.

Mais il y a quand même des choses qui parfois sont trop complexes à reproduire comme les conditions extrêmes de vol pour un avion, des réactions nucléaires ou l'univers lui-même.

La simulation numérique comme un outil pour pallier aux problèmes d'expérimentation

Pour résoudre ce problème, l'informatique a été utilisée pour réaliser des simulations dites numériques. Comme ceci est expliqué sur la page dédiée à la simulation de phénomènes sur wikipédia :

Le principe de base est celui de la représentation du phénomène par une équation. Grâce à une puissance de calcul toujours croissante et à l'augmentation du volume de données stockables il est possible de découper un phénomène complexe en milliers, voire en millions, de phénomènes simples et donc de calculer les résultats sur le phénomène complexe. Exemple : on sait, en aérodynamique, représenter par une équation les forces (portance, traînée) qui résultent de l'action d'un courant d'air sur une plaque plane. On ne sait

²http://fr.wikipedia.org/wiki/Simulation_de_ph%C3%A9nom%C3%A8nes

³http://fr.wikipedia.org/wiki/Mod%C3%A8le_math%C3%A9matique

pas représenter par une équation ces mêmes forces lorsque l'action est exercée sur une surface complexe telle que l'aile d'un avion. La simulation numérique permet de découper l'aile en plusieurs millions de petits éléments qu'on considère comme étant des plaques planes. On peut alors calculer les forces qui s'exercent sur chacune d'entre elle et les combiner pour calculer les forces sur l'aile complète.

Problématique des modèles et des simulations

Attention, je ne prétends pas que la simulation est la réponse ultime et qu'elle permet de tout prendre en compte et tout comprendre. Tous les phénomènes physiques ne pourront pas forcément être pris en compte dans la simulation (les passagers qui bougent dans un avion, des phénomènes locaux pour la météo, etc), et les modèles utilisés ... sont des modèles. Ils intègrent donc des hypothèses qui ne seront pas forcément vraie IRL ou des simplifications qui permettent à la simulation d'être pertinente en temps de calcul ou d'espace mémoire à utiliser.

La simulation doit être validée par des expériences physiques afin de s'assurer que le modèle utilisé pour représenter le phénomène étudié est valable et que la simulation numérique basée sur ce modèle est pertinente (les techniques de discrétisation et de calcul doivent aussi être validées). Et pour cela l'expérimentation est indispensable.

Autre point : il est important de noter que l'on ne pourra rien apprendre sur ce qui n'est pas inclus dans notre modèle. On ne peut pas faire de découverte. Il ne faut pas faire croire que l'on peut se servir des modèles pour découvrir des phénomènes, au contraire, pour des modèles complexes, on peut en arriver à des résultats qui peuvent être totalement inattendus : non pas que ce soient des découvertes, mais des conséquences d'interactions non prévues à priori :

- Dans des modèles multi-échelle (comme pour la simulation un système biologique au niveau macro-cellulaire, micro-cellulaire, biochimique, génétique, etc)
- Ou des modèles multi-physique (comme pour la simulation de mécanique des fluides en conditions extrêmes : air/eau/glace résistance des matériaux, thermodynamique, etc).

Encore un point sur les modèles : les modélisation choisies pour représenter des phénomènes sont choisies la plupart de temps pour étudier certains éléments spécifiques. Quand on veut étudier le comportement d'une aile d'avion au sein d'un écoulement d'air, on peut proposer des modèles différents si l'on veut étudier :

- son comportement avec un écoulement sub-sonique
- son comportement avec un écoulement super-sonique

- son comportement dans des conditions extrêmes (air + basse température + eau)

Pour chacune de ces simulations, on va avoir des équations qui seront modélisées de manière différente car les composantes qui seront importantes seront différentes, voire il sera nécessaire d'ajouter des éléments à la simulation comme la prise en compte de la température, où la prise en compte des interactions air/eau, etc.

Je sais que part le passé, il y a eu tout un échange sur les modèles, et notamment leur limites (un HS d'ailleurs), est ce que vous pourriez donner votre point de vue sur leur usage ?

Le tableau dressé sur les modèles ou la simulation jusqu'ici est un peu noir, mais il ne faut pas trop s'affoler. J'ai appuyé sur ces points pour vraiment expliquer la réalité des choses. Avec les simulations numériques on arrive à faire des choses vraiment extraordinaires ! En fait on peut rapidement comprendre l'intérêt de la simulation numérique ^{4 5} si on prend quelques exemples d'expérimentation et de tests (de la simulation analogique) et c'est ce que l'on va faire dans la suite du podcast.

Mécanique des fluides, aéronautique et automobile

Dans ces domaines l'expérimentation et les tests coûtent excessivement cher. On a grosso modo deux types d'expériences que l'on peut réaliser :

- On peut faire des tests en soufflerie
- On peut faire des tests en conditions réelles

Expériences et tests en soufflerie

Il s'agit ici de tester l'écoulement de fluides autour d'éléments dont on veut comprendre le comportement. On se sert la plupart du temps de fumée pour "voir" les écoulements. Les éléments testés peuvent être instrumentés (les objets sont équipés de capteurs qui vont permettre de connaître différentes grandeurs physiques comme la température, la pression, etc) pour comprendre les températures ou les contraintes auxquels ils sont soumis.

La réalisation de tests en soufflerie nécessite de disposer de soufflerie suffisamment grande pour arriver à tester les systèmes en grandeur nature : voiture, éléments d'avion. Mais aussi de faire un très grand nombre d'essai pour arriver à tester les différentes conditions que l'on souhaite observer. Il est même

⁴http://fr.wikipedia.org/wiki/Simulation_de_ph%C3%A9nom%C3%A8nes

⁵http://fr.wikipedia.org/wiki/Simulation_informatique

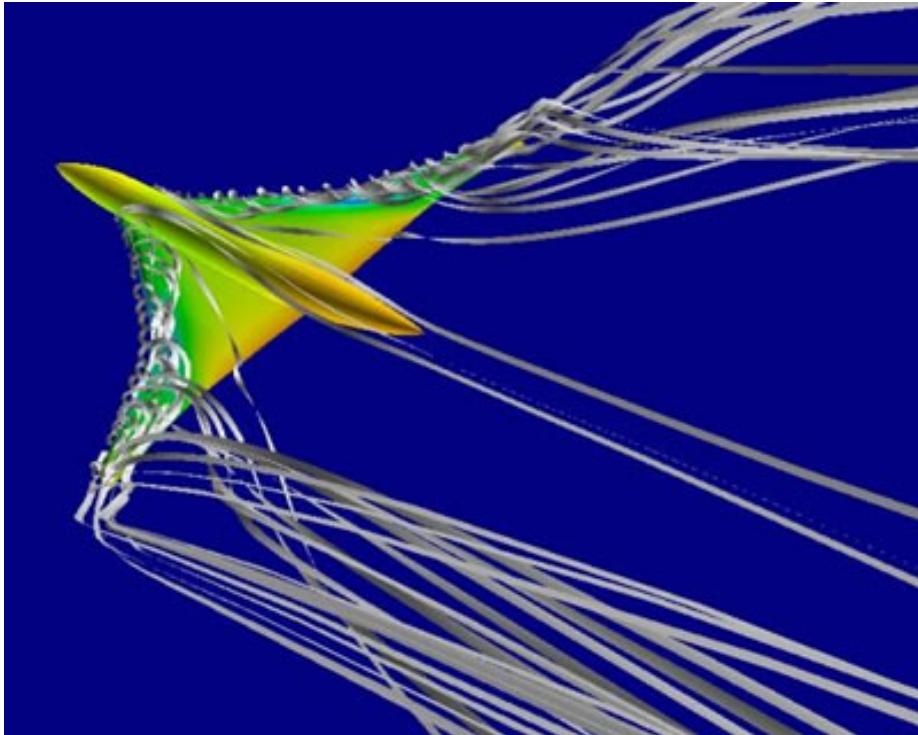


Figure 1: Simulation numérique de l'aérodynamique de l'hypersustentation de l'avion de transport supersonique Epistle (DAAP - Gérard Carrier) <http://hathor.onera.fr/images-science/simulation-numerique/sst-epistle.php>

compliqué voire impossible de tester des avions en soufflerie : soit ils sont trop volumineux et alors on ne teste que des composants (les ailes par exemple), soit cela se fait avec des réductions.

Seules certaines composantes des écoulements vont pouvoir être testé, on ne peut en effet qu'envoyer des filets de fumée afin de pouvoir voir comment ils se comportent.



Figure 2: Vu de face d'un BMW en soufflerie <http://www.usinenouvelle.com/article/bmw-joue-a-fond-la-carte-ecolo.N70485>

On voit donc que cela peut devenir compliqué de tester des conditions extrêmes de vol comme par exemple (températures basses, écoulement d'air, gouttelettes d'eau pour simuler des nuages, etc).

En grandeur réelle

Ceci peut par exemple se voir pour le test de crash de voiture. Ici des conditions spécifiques vont être testées : impact de collision de voitures à différentes vitesses, avec différents angles ou points d'impact, etc. On voit rapidement que cela va coûter cher car il est nécessaire d'utiliser des véhicules ... et de les faire percuter des objets de manière plus ou moins violente.

Pour ce qui est de l'aéronautique, c'est encore plus complexe. Déjà il est nécessaire de construire des avions en taille réelle pour qu'ils puissent être pilotés et cela coûte extrêmement cher! Je ne prétends pas que cela ne soit pas fait,

mais dans une mesure qui ne permet pas de pouvoir s'en servir comme source principale d'information concernant le comportement des avions construits. En effet, et si le design possédait une faille majeur qui, une fois en vol provoquerait un crash ?

Il apparaît donc indispensable de pouvoir simuler des avions pour comprendre comment cela se passe. Une fois que tout est validé par la simulation, il peut ensuite être temps de faire des tests de vérification d'éléments de fuselage et enfin des tests en vol pour valider que ce qui était attendu pour l'avion au global. Et là, je ne parle même pas des tests de crash d'avion comme cela pourrait être le cas avec voiture comme je l'ai expliqué plus haut.

Domaines où l'expérience est impossible

La météorologie

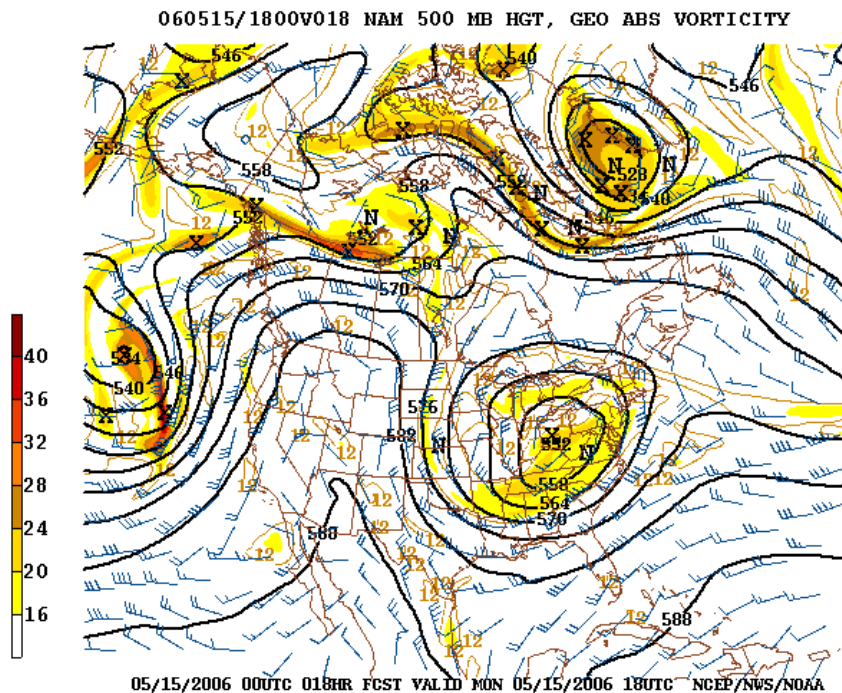


Figure 3: Prévision météorologique réalisée par modèle de prévision américain NAM http://commons.wikimedia.org/wiki/File:NAM_500_MB.PNG

Prenons un autre domaine où la simulation est indispensable : la météorologie ⁶

⁶http://fr.wikipedia.org/wiki/Pr%C3%A9vision_m%C3%A9t%C3%A9orologique#Pr.C3.A9vision_num.C3.A9rique

⁷ . Ici la reproductibilité n'est même plus possible. L'idée est plutôt de pouvoir prévoir le temps. Réaliser des tests à échelle réduite ne me semble pas possible, ni même pertinent. Il est indispensable de pouvoir simuler l'atmosphère, son interaction avec les sols, avec l'océan, l'impact de la nuit et du jour, etc, etc. Quand on parle de météorologie, certaines personnes tiquent un peu quand on dit que c'est une science. Il est vrai que l'on observe souvent nos chers présentateurs météo annoncer de la pluie en Bretagne à longueur de temps et on entend régulièrement certains quimperois expliquer à corps et à cris que non, "il faisait grand beau ce jour là, et qu'ils en ont ras la crêpe". Ceci vient d'au moins deux phénomènes :

- **La discrétisation et le maillage choisi.** Pour réaliser des simulations sur un ordinateur (un système digital et pas analogique qui fonctionne donc par étape discrète), il faut discrétiser son domaine d'étude (le découper en morceaux) et produire un maillage (une grille grosso modo) dont les intersections seront les points auxquels nous chercherons les valeurs de la température (pour faire simple). Sauf que vouloir simuler tout un pays, voire plus pour prendre en compte la mer ou des choses comme les montagnes, etc. donc de grands territoires, c'est compliqué. Afin de pouvoir aller vite, la distance entre deux points du maillage peut-être de plusieurs kilomètres. Le problème c'est que des "micro-climats" ou des orages par exemple qui vont avoir de forts impacts à de plus grandes échelles. Pour résoudre cela on peut "paramétriser" les mailles, c'est-à-dire ajouter des bouts de modèles à ces échelles qui vont prendre en compte certains phénomènes. Cela reste néanmoins approximatif et pas forcément très efficace.

Je ne sais pas si tout le monde a suivi jusqu'ici, est ce qu'il y a des questions sur ces problématiques? Je peux vous donner quelques exemples de maillage pour que vous puissiez voir ce que cela peut donner pour des modèles utilisés "dans la vraie vie".

Pour donner un ordre d'idée de maillages utilisés, voici ce que Wikipédia donne comme info ⁸ : "Météo-France utilise actuellement pas moins de trois modèles numériques, telles trois boîtes imbriquées les unes dans les autres pour émettre ses bulletins. L'IFS (Integrated Forecasting System), calculé par le Centre européen de prévision météorologique à moyen terme à Reading en Grande-Bretagne, donne une prédiction pour l'Europe avec une maille mondiale de 25 km de côté. Arpège (Action de recherche petite échelle /grande échelle) de Météo France utilise le résultat et refait le calcul avec une maille de 25 km sur la France pour une

⁷http://fr.wikipedia.org/wiki/Pr%C3%A9vision_num%C3%A9rique_du_temps

⁸http://fr.wikipedia.org/wiki/Pr%C3%A9vision_m%C3%A9t%C3%A9orologique#Pr.C3.A9vision_num.C3.A9rique

échéance de un à trois jours, mais moins précise ailleurs. Aladin (Aire limitée et adaptation dynamique) reprend les résultats pour l'Europe de l'Ouest avec une précision de 10 km. Depuis 2008, le modèle AROME (Application de la recherche à l'opérationnel à mésoéchelle) calcule sur une maille de 2,5 km (cent fois plus précise en surface que le modèle Arpège et près de dix fois plus qu'Aladin). Le calcul intègre en continu et réajuste les prévisions à partir des informations des stations météorologiques, navires, bouées, avions, radar, satellites... et intègre aux données de base : les vents, les précipitations et l'humidité de l'air fournies par les satellites GPS. Le modèle évalue aussi la fiabilité de la prédiction".

- **La couverture en données du domaine considéré et l'erreur sur les mesures de ces données.** Il apparaît clair, que les zones géographiques terrestres sont plus couvertes et permettent d'avoir plus de données, donc une plus grande précision sur les conditions initiales du domaine. Par contre, pour les océans c'est plus compliqué. Et même avec les informations que les balises en mer peuvent fournir (sur la température ou les courants marins), on se retrouve avec beaucoup moins de données. Pour palier à cela, on va mitiger un peu ces observations avec des résultats précédant pour essayer de se rapprocher au mieux de la réalité. Un point important à savoir quand on fait de la simulation temporelle comme pour les prévisions météo, c'est que l'on fait évoluer le système à partir de ses conditions initiales et donc plus celles-ci sont précises, plus le résultat final sera précis. Et inversement. Ajouté à cela les erreurs numériques des modèles (des approximations de la réalité) et celles des ordinateurs, on se retrouve avec un système qui peut totalement diverger et ceci vers des solutions qui pour des simulations indépendantes avec des écarts minimes au niveau des données d'entrée peuvent être totalement différentes ! Cet aspect de "chaos" (faudrait faire un dossier sur le sujet, c'est juste un truc de fou) a notamment été mis en exergue par Poincaré pour la simulation à N corps du système solaire en 1890 (Il y a une jolie histoire sur ce qui a amené Poincaré à découvrir cela :) et ensuite par Edward Lorenz ⁹, (c'est pas le même que pour les invariants de Lorentz en relativité) pour la météo en 1963. C'est d'ailleurs lui qui est à l'origine de la fameuse métaphore du papillon et les simplifications qu'il avait du faire pour tenter des simulations plus simple de météo ont abouti à la découverte des "attracteurs étranges" qui font qu'en fait, pour presque toutes les conditions initiales possibles, on se retrouve attiré par les zéros des équations différentielles qu'il considérait^{10 11}.

En 1890, Poincaré publie le fameux mémoire intitulé : "Sur le problème des trois corps et les équations de la dynamique", qui lui vau-

⁹http://fr.wikipedia.org/wiki/Edward_Norton_Lorenz

¹⁰http://fr.wikipedia.org/wiki/Syst%C3%A8me_dynamique_de_Lorenz

¹¹http://fr.wikipedia.org/wiki/Th%C3%A9orie_du_chaos

dra le prix du roi Oscar, roi de Norvège et de Suède et passionné de mathématiques. L'histoire est célèbre : le mémoire lauréat comportait une erreur détectée par le jeune mathématicien Phragmén alors qu'il prépare le manuscrit pour l'imprimeur. Cette erreur obligera Poincaré à procéder à de profonds remaniements dans son mémoire, et aussi à rembourser les frais d'impression du premier mémoire, une somme supérieure de quelque mille couronnes au prix qu'il avait reçu. Mais cette erreur fut féconde, car en lieu et place de la stabilité du Système solaire, Poincaré découvrit le chaos potentiel caché dans les équations de la dynamique.

Il s'agit ici d'un des domaines dans lequel la théorie du Chaos peut dire des choses, je ne sais pas si vous avez des exemples d'autres applications de cette théorie ?

Pour compenser ces sources d'erreurs, une nouvelle approche a été mise en place : les prévisions d'ensemble. Il est ici question de faire varier de les conditions initiales dans le domaine d'incertitude des mesures et des modèles disponibles afin de pouvoir, comme expliqué sur Wikipédia : "permettre de quantifier la prédictibilité de l'atmosphère et d'offrir une marge d'erreur statistique sur la prévision."

Pour infos, trois chercheurs, dont un très connus : Charney, Fjortoft et von Neumann réussirent la première prévision numérique du temps par ordinateur ¹² . Les premiers programmes de prévisions numériques opérationnelles furent instaurés au début des années 1960.

La cosmologie

Prenons encore un autre domaine, et je pense que j'aurais bien illustré la situation : la cosmologie. Ici le but est de comprendre l'évolution des galaxies, la mécanique céleste, etc. Un bon exemple est de comprendre la forme des galaxies et notamment ce qu'il se passe quand deux galaxies se rencontrent, comme Andromède et la Voie Lactée vont le faire. A part la simulation, la seule solution que j'entrevois est d'attendre 4,5 milliards d'années et de prendre des photos pour voir ce que cela va donner ... Problème encore plus complexe en lien avec la cosmologie : Quelle est la répartition de masse dans l'univers et comment celle-ci a influencé son évolution jusqu'à nos jours ? La grande simulation Deus [deus] qui a été mise en place par JM Alimi et son équipe au sein de l'Observatoire de Paris a, pour comprendre, tenté de simuler tout cela de manière massive sur le super ordinateur nommé CURIE qui se situe au sein du TGCC (Le Très Grand Centre de Calcul) à Bruyères le Châtel et qui fait partie du GENCI (le Grand Équipement National de Calcul Intensif), la structure qui fédère les moyens informatiques, dédiés à la recherche scientifique, les plus importants de France (à

¹²http://fr.wikipedia.org/wiki/Pr%C3%A9vision_num%C3%A9rique_du_temps

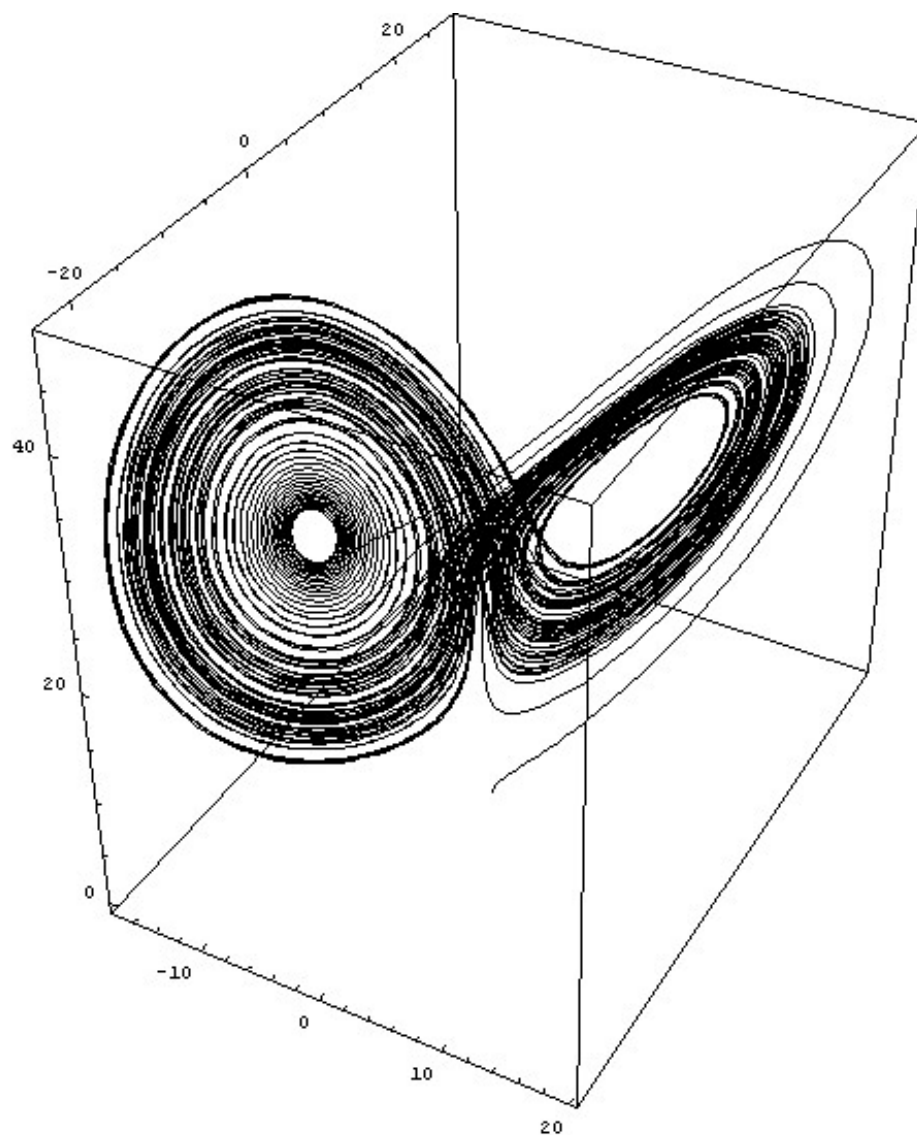


Figure 4: Attracteur étrange de Lorenz (1963) <http://commons.wikimedia.org/wiki/File:Lorenz.jpg>

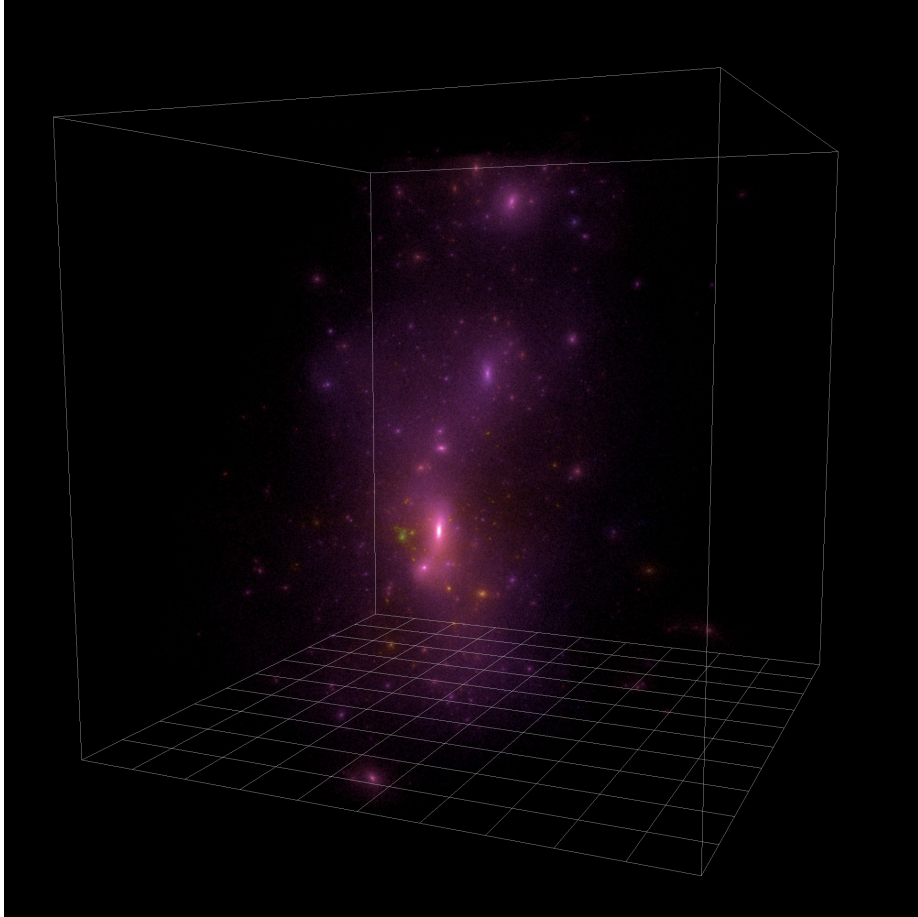


Figure 5: Halos pour une simulation 162 Mpc/h pour une cosmologie de type Λ CDM $w=5$ <http://www.deus-consortium.org/gallery/images/>

savoir que la machine CURIE dispose de 92000 coeurs et que sa performance crête est de 2 millions de milliards d'opérations à la seconde (Pour comparer la machine la plus puissante qui est en chine et se nomme Tianhe-2 possède un peu plus de 3 millions de coeurs et possède une performance de plus 33 millions de milliards d'opérations à la seconde). Pour informations, l'idée était de simuler tout l'univers depuis le Big Bang jusqu'à maintenant avec 3 scénarios de matière noire ^{13 14} :

- Λ CDM : où l'on considère la relativité générale avec la constante cosmologique (équation d'état cosmologique ¹⁵ constante, pour simplifier le ratio nommé w entre la pression et la densité d'un fluide considéré parfait qui remplirait l'univers vaut -1) et de la matière noire froide (La matière noire froide ou **CDM**, de l'anglais *Cold dark matter*, est une théorie améliorée à partir de celle du Big bang. Elle repose sur le postulat supplémentaire que la majeure partie de la matière de l'Univers serait constituée d'éléments non observables par le rayonnement électromagnétique - noire - et dont les constituants se déplacent lentement - froide - ¹⁶)
- modèle à base de quintessence ¹⁷: où l'on considère la relativité générale avec une énergie noire plus générique (qui évolue au cours du temps et qui correspond à un champ scalaire avec une équation d'état qui est dynamique, pour simplifier w varie au cours du temps) et de la matière noire froide
- modèle d'énergie fantôme ¹⁸: qui est un dérivé du modèle à base de quintessence où l'équation d'état est satisfaite avec une valeur de w inférieure à -1. Ces modèles pousse pour des univers cycliques ou dit de "Big Rip".

Cela représentait la simulation de l'évolution de plus de 550 milliards de particules en utilisant 76000 coeurs de processeurs et plus de 300 Téra-octets (37500 DVDs) de mémoire vive pendant l'équivalent de 50 millions d'heures de calcul (presque 28 jours grâce aux 76k coeurs)! Au final ils ont généré 1,5 Péta-octets de données utiles (187500 DVDs de données) qui serviront à ALMA, VLT, au futur télescope spatial Euclid, mais aussi au Radio télescope Quare Kilometer Array qui sera opérationnel en 2020.

¹³http://www.asa2012.phys.unsw.edu.au/HWWS/Harley%20Wood%20Winter%20School%202012/speaker%20notes/Davis_notes.pdf

¹⁴<http://www.deus-consortium.org/a-propos/dark-energy-universe-simulation-full-universe-run/cosmological-models-2/>

¹⁵[http://en.wikipedia.org/wiki/Equation_of_state_\(cosmology\)](http://en.wikipedia.org/wiki/Equation_of_state_(cosmology))

¹⁶http://fr.wikipedia.org/wiki/Mati%C3%A8re_noire_froide

¹⁷[http://en.wikipedia.org/wiki/Quintessence_\(physics\)](http://en.wikipedia.org/wiki/Quintessence_(physics))

¹⁸http://en.wikipedia.org/wiki/Phantom_energy

D'autres domaines où les limitations ne sont pas forcément physiques

On pourrait illustrer cela avec encore d'autres domaines comme la génétique où l'on ne peut pas tester toutes les maladies avec tous les remèdes possibles pour d'évidentes raisons éthiques ... Ou bien la physique nucléaire où soit il est nécessaire d'avoir des instruments très complexes à disposition pour faire des expériences (LHC, Iter, etc), soit il faut pouvoir faire des tirs de bombes nucléaires pour observer ce qu'il se passe (mais ce n'est plus autorisé dans nos démocraties, et puis de manière générale ce n'est pas top). Un autre domaine où il est compliqué aussi de faire des expériences en grandeur réelle est celle de la sélection naturelle et/ou de l'évolution biologique, le podcast de Marco sur l'expérience de Lenski montre bien que c'est complexe et que la simulation peut avoir un vrai apport. David pourra sûrement nous en parler plus en détail!

David, est ce que tu as des précisions sur les aspects de simulation de l'évolution biologique et de la sélection naturelle ?

Physique nucléaire et méthodes de Monté Carlo : première simulation numérique

Avant de parler des méthodes que je connais mieux pour la simulation numérique, je voudrais aborder le cas de la première simulation numérique qui a été réalisé en 1953 sous l'impulsion de John Van Neumann dans le domaine de la physique nucléaire ^{19 20 21}. Il souhaitait pouvoir, dans le cadre du projet Manhattan simuler le comportement des neutrons lors des réactions en chaîne dans des armes explosives basées sur la fission. Il est apparu qu'une bonne méthode pour arriver à faire cela est celle de Monté Carlo. Comme le précise Wikipédia, "Le nom de ces méthodes, qui fait allusion aux jeux de hasard pratiqués à Monte-Carlo, a été inventé en 1947 par Nicholas Metropolis, et publié pour la première fois en 1949 dans un article coécrit avec Stanislaw Ulam". S. Ulam que l'on retrouve notamment dans l'expérience de John Van Neumann qui s'appelle d'ailleurs l'expérience de Fermi-Pasta-Ulam (Fermi était le chercheur, Pasta l'informaticien et Ulam le mathématicien).

La méthode proposée (qui est la base de toute bonne méthode de Monte Carlo) consiste dans le découpage d'un processus complexe en un grand nombre de petites étapes, où, à chaque pas plusieurs choix sont possibles avec pour chacun d'entre eux une probabilité définie de se réaliser. Chaque run complet va fournir un résultat et en effectuant un grand nombre on sera en capacité de connaître la fonction de probabilité du processus complexe global en incluant l'incertitude inhérente due au hasard.

¹⁹http://fr.wikipedia.org/wiki/Exp%C3%A9rience_de_Fermi-Pasta-Ulam

²⁰<http://132.187.98.10:8080/encyclopedia/en/monteCarlo.pdf>

²¹http://fr.wikipedia.org/wiki/M%C3%A9thode_de_Monte-Carlo

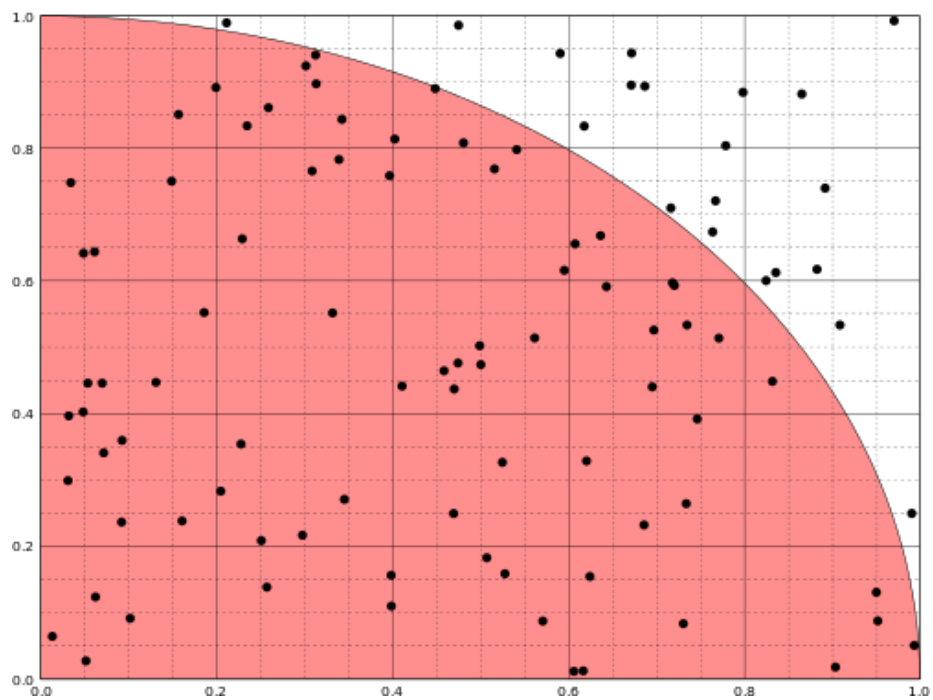


Figure 6: Représentation du calcul de la valeur de pi par rapport au nombre de points aléatoires étant contenus dans un quart de cercle, l'ensemble des possible étant un carré de coté R. L'exemple est pris avec 100 points <http://commons.wikimedia.org/wiki/File:Montecarlo-valeur-pi.svg>

John Von Neumann a immédiatement reconnu les avantages de cette méthode et notamment dans son application à l'exploration du comportement des chaînes de réaction à base de neutre dans les armes basées sur la fission, pour la prédiction de leurs explosions, mais aussi pour les problèmes de diffusion et de multiplication des neutrons.

Dans la simulation que faisait John Von Neumann il y avait plusieurs étapes qui peuvent être représentées sous forme de diagramme :

- La première était la génération d'un neutron dont la position et la vitesse était fournies de manière aléatoire. Les lois de probabilité correspondantes sont déterminées afin de refléter les conditions initiales réputées connues.
- La seconde étape consistait en un déplacement libre du neutron sur une certaine distance définie par une distribution de probabilité spécifique qui va dépendre par les propriétés du matériaux et les sorties de la première étape.
- Trois cas peuvent ensuite apparaître pour l'étape suivante :
 - Soit le neutron entre en collision au sein du même matériaux
 - Soit le neutron traverse la frontière afin de réaliser un nouveau déplacement libre dans ce nouveau matériau.
 - Soit la branche se termine car le neutron a été absorbé par le matériau sans générer de nouveau neutron ou est sorti complètement de l'objet.

Pour chacune de ces étapes, des distributions de probabilités sont utilisées pour chacun des cas identifiés et il est donc nécessaire de disposer d'une bonne source de nombres aléatoires.

Importance de la qualité du hasard

Ce que l'on remarque et qui a déjà été abordé dans de précédents podcasts (notamment ceux de Robin sur le hasard), c'est qu'il est important d'avoir une bonne source de nombres aléatoires.

Il est tout à fait possible d'utiliser des générateurs de nombres aléatoires qui sont basés sur des phénomènes assez imprévisibles pour être considérés aléatoires (dés, roulette, etc). Le problème est que le temps d'obtention des nombres peut-être lent, voire trop lent pour les expériences à réaliser pour lesquels il faut parfois un grand nombre ... de nombres aléatoires ²².

Il existe aussi des générateurs de nombres aléatoires reposant sur des phénomènes physiques (radioactivité, bruits thermiques, mécanique quantique, etc).

²²http://fr.wikipedia.org/wiki/G%C3%A9n%C3%A9rateur_de_nombres_al%C3%A9atoires

Il existe enfin des algorithmes qui peuvent être utilisés pour générer des nombres aléatoires. Nous sommes tous d'accord que des algorithmes déterministes, exécutés sur des machines déterministes (les ordinateurs), ne risquent pas de donner des nombres vraiment aléatoires. C'est d'ailleurs pour cela que l'on parle plutôt de nombre pseudo-aléatoires²³. L'intérêt de ces générateurs de nombres pseudo-aléatoires (ou Pseudo-Random Number Generator, PRNG), est qu'ils permettent de fournir des nombres qui présentent certaines propriétés du hasard, comme par exemple l'indépendance des nombres entre eux ou le fait qu'il est difficile de repérer des groupes de nombres suivant certaines règles. L'autre intérêt est tout simple : ils peuvent être implémentés et peuvent générer des nombres très rapidement sans avoir de matériel spécifique (comme pour les générateurs physiques). Il existe d'ailleurs plusieurs ensembles de suite de tests qui ont été mises en place pour valider le hasard produit par ces PRNG :

- La DieHard test suite de George Marsaglia (<http://stat.fsu.edu/pub/diehard/>) comme le film avec notre Bruce Willis préféré.
- La TestU01 de Pierre L'Ecuyer et Richard Simard (<http://www.iro.umontreal.ca/~simardr/testu01/tu01.html>)

Ces tests vont par exemple tester la répartition dans un cube des nombres générés quand on en prend trois consécutifs pour faire une coordonnée ou encore le test dit de “parking lot” en plaçant des voitures sur une parking de carré 1 (chaque voiture est une sphère). Si il y a des crashes (une voiture veut se garer sur une place déjà prise), on recommence. Au bout d'un certain nombre de tentatives, on regarde le nombre de voitures garées, et on doit s'attendre à obtenir une probabilité qui suit une loi normale.

Parmi les failles qu'il faut éviter chez ces PRNG, on peut notamment citer :

- La variation de la qualité du hasard des nombres générés en fonction de la graine fournie
- Valeurs successives qui ne sont pas indépendantes
- Une période trop petite
- Certains bits qui ne sont pas assez aléatoire dans les sorties (en informatique toutes les données sont représentées sous forme de 0 et de 1 et assemblés en un certain nombre d'octets (ensemble de 8 bits). Les entiers sont par exemple codés sur 32 bits, ou 4 octets)

Parmi les PRNG les plus connus, Robin avait déjà parlé de celui de John Van Neumann (Robin avait notamment abordé ses failles avec l'histoire des zéros).

Robin tu pourrais juste nous redonner l'explication du fonctionnement de ce générateur s'il te plaît ?

Avec Nico ils avaient abordé le cas des générateurs congruentiels linéaires²⁴ (à

²³http://fr.wikipedia.org/wiki/G%C3%A9n%C3%A9rateur_de_nombres_pseudo-al%C3%A9atoires

²⁴http://fr.wikipedia.org/wiki/G%C3%A9n%C3%A9rateur_congruentiel_lin%C3%A9aire

base de module) qui sont des suites numériques dites arithmético-géométriques et dont on prend le modulo d'une certaine valeur (la valeur à l'étape suivante est la somme du produit de la valeur à l'étape courante par un coefficient (a) plus un nombre (b) et l'on prend le module pour une certaine valeur (m)).

Ces générateurs ont été introduits par Lehmer en 1948 sous une forme réduite (b était nul). Ici la période est au maximum m (dans les implémentations 32 bits et bien c'est 2^{32}). Un des générateurs de ce type qui a été le plus utilisé se nomme RANDU ²⁵. Il fut introduit par IBM dans les années 60. Pour lui, les paramètres a et m valent respectivement 65539 et 2^{31} . Ce générateur a beaucoup été utilisé, jusqu'à ce que l'on se rende compte qu'il générait des nombres, qui utilisés pour représenter des points en dimension supérieur à 2, n'était pas placé au hasard dans les cubes ou plus considérés. Pour les plus ferrus de mathématiques, RANDU échoue violemment sur le test spectral, notamment à partir de la dimension 3.

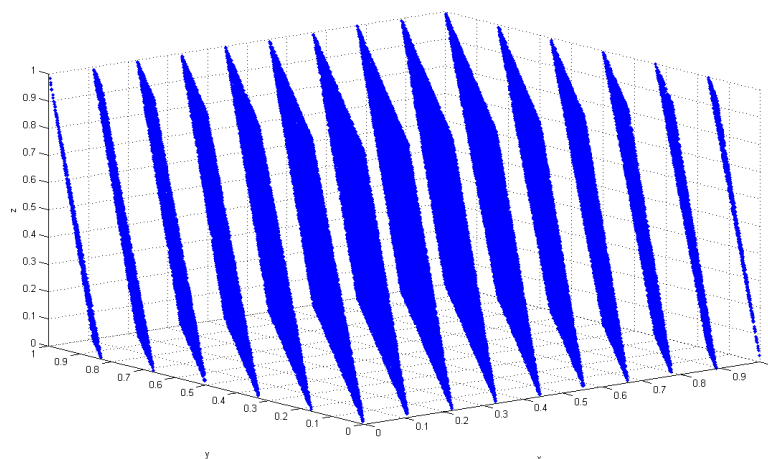


Figure 7: Représentation créée avec MATLAB en générant 100002 valeurs utilisant RANDU et en positionnant dans un graphe 100000 triplets (x,y,z) consécutifs <http://en.wikipedia.org/wiki/File:Randu.png>

Pour finir sur les PRNG, l'un des plus connus pour être de bonne qualité (sauf pour la crypto) est celui inventé par Makoto Matsumoto et Takuji Nishimura en 1997 nommé Mersenne Twister ²⁶. Il possède déjà une très grande période : $2^{19937} - 1$, il est uniformément distribué sur 623 dimensions (pour des nombres sur 32 bits), il est très rapide, aléatoire quelque soit le poids du bit considéré et pas les tests de Die Hard. Je ne vais pas forcément rentrer dans les détails de son implémentation, bien plus complexe que les LCG, mais l'idée est de faire

²⁵<http://fr.wikipedia.org/wiki/RANDU>

²⁶http://fr.wikipedia.org/wiki/Mersenne_Twister

des déplacements de bits, et des opérations logiques de type OU exclusif sur les éléments d'une matrice qui contient des bits de nombres aléatoires et qui est réutilisée, un peu comme avec les LCG pour l'itération suivante.

Une des dernières problématiques avec l'usage de PRNG est dans le cas de parallélisation de code. Quand un code est déployé sur des serveurs de calcul, une des problématiques qui apparaît est de pouvoir générer pour chaque bout de code une série de nombres aléatoires qui sont indépendants et identiquement distribués et que notamment, on ne se retrouve pas à utiliser les mêmes nombres pseudo-aléatoires sur chaque noeud de calcul.

Conclusion

J'espère qu'en vous présentant tout ces cas où l'expérimentation n'est plus vraiment possible (et ceci pour diverses raisons) et où l'informatique a beaucoup apporté, j'ai pu vous montrer qu'il est donc devenu indispensable de simuler grâce à l'informatique.

Je pense que je vais m'arrêter là sur la simulation pour ce podcast. On a pu voir plusieurs choses :

- Les raisons du besoin de simulation,
- L'attention qu'il faut porter aux problématiques des modèles et des simulations,
- Les problématiques rencontrées avec différents champs de recherche et notamment la première simulation numérique qui était probabiliste.

Dans la suite de ce podcast, on ira plus sur les simulations pour des équations déterministes. On verra trois grandes méthodes utilisées pour passer des modèles à l'ordinateur, et ensuite les ramifications en terme de mathématiques avec les méthodes utilisées pour résoudre les problèmes posés et ce que l'informatique a fait pour les résoudre dans des temps corrects.

Note

Matière noire, énergie noire et la composition de l'univers

Si l'on souhaite s'intéresser à ces notions, il est possible de se référer aux sources suivantes :

- Wikipédia ²⁷ et ²⁸
- Le livre de Trinh Xuan Thuan : Désir d'infini ²⁹
- Ou encore le très bon dossier de Mathieu, l'épisode 31 qui traitait de la question ³⁰

En ce qui concerne la matière noire il y a grosso modo deux grands types de matière noire :

- La matière noire chaude : composée de particules très rapides dont la vitesse est proche de celle de la lumière
- La matière noire froide : composée de particules plus massives et plus lentes

Comme expliqué sur Wikipédia et très bien aussi dans le dossier de Mathieu :

- Si l'Univers était dominé par de la matière noire chaude, la très grande vitesse des particules la constituant empêcherait dans un premier temps la formation d'une structure plus petite que le superamas de galaxies qui ensuite se fragmente en amas de galaxies, puis en galaxies, etc. C'est le scénario dit "du haut vers le bas", puisque les plus grosses structures se forment d'abord, pour ensuite se diviser. Le meilleur candidat pour constituer la matière noire chaude est le neutrino.
- En revanche, si la matière noire froide dominait l'Univers, les particules vont parcourir une distance plus petite et donc effacer les fluctuations de densité sur des étendues plus petites que dans le cas de matière noire chaude. La matière ordinaire va alors se regrouper pour former d'abord des galaxies (à partir de nuages de gaz), qui elles-mêmes se regrouperont en amas, puis superamas. C'est le scénario dit "du bas vers le haut". Les candidats à la matière noire froide sont les WIMP et les MACHO."

Actuellement, c'est le modèle de matière noire froide qui semble l'emporter. En effet, les galaxies sont en équilibre dynamique, ce qui laisse penser qu'elles se sont créées avant les amas — dont tous ne semblent pas encore stables — à qui il faut plus de temps pour atteindre cet équilibre. Cependant, les théories introduisent aujourd'hui un peu de matière noire chaude. Celle-ci est nécessaire pour expliquer la formation des amas ; la matière froide seule ne pouvant la permettre en si peu de temps.

²⁷http://fr.wikipedia.org/wiki/Mati%C3%A8re_noire

²⁸http://fr.wikipedia.org/wiki/%C3%89nergie_sombre

²⁹<http://www.fayard.fr/desir-dinfini-9782213635118>

³⁰<http://www.podcastscience.fm/dossiers/2011/04/07/dossier-la-matiere-noire/>

Énergie fantôme

Comme expliqué sur Wikipédia ³¹ :

En cosmologie, l'énergie fantôme désigne une forme hypothétique d'énergie dont la densité aurait la particularité surprenante d'augmenter lors de l'expansion de l'Univers. Cette énergie est un candidat potentiel à l'énergie noire. Cette forme d'énergie noire implique la violation d'un principe physique connu sous le nom de condition faible sur l'énergie ³² (valable pour les fluides parfaits). Cet argument est l'un des principaux écueils théoriques à l'existence de l'énergie fantôme. Cependant, la condition faible sur l'énergie n'est valable que pour les fluides parfaits. L'énergie sombre pourrait donc ne pas enfreindre ce principe, à condition de postuler une interaction avec d'autres composants cosmologiques, comme, par exemple, la matière sombre. L'énergie fantôme, si elle existe, serait responsable d'un emballement de l'expansion de l'Univers, qui causerait un éloignement arbitrairement grand des différents objets célestes les uns des autres en un temps fini, puis une dislocation de ceux-ci : c'est le modèle cosmologique du Big Rip ³³. Dans ce scénario, la densité de l'univers atteint une valeur infinie en un temps fini, terminant son existence par une singularité gravitationnelle où toute structure, des amas de galaxies aux atomes, sont détruits.

Ce modèle diffère du Big Crunch, dans lequel l'expansion s'arrête pour laisser place à une phase de contraction. Dans ce dernier cas, les objets astrophysiques sont détruits car ils entrent en collision les uns avec les autres (ils sont "écrasés", d'où Big "Crunch"), alors que dans le cas du Big Rip, ils sont étirés par une expansion de plus en plus violente, jusqu'à être disloqués (ils sont « déchirés », d'où Big "Rip"). L'hypothèse du scénario du Big Rip est pour l'heure marginalement compatible avec les données observationnelles et reste donc relativement spéculative.

³¹http://fr.wikipedia.org/wiki/%C3%89nergie_fant%C3%B4me

³²http://fr.wikipedia.org/wiki/Condition_sur_l%27%C3%A9nergie

³³http://fr.wikipedia.org/wiki/Big_Rip