



Mémoire de
Magistère de Mathématiques

Lucile Morelle

Août 2013

1 Cursus au sein du Magistère

Les mathématiques ont toujours été une passion pour moi, d'où mon vœux de les enseigner. Lorsque je suis rentrée en classes préparatoires aux grandes écoles au lycée Henri Wallon de Valenciennes, j'ai choisi spontanément le parcours MPSI/MP et c'est en deuxième année que j'ai pris connaissance de l'existence du Magistère de Mathématiques d'Orsay, une formation d'excellence, plus riche qu'une formation universitaire habituelle. J'ai su tout de suite que choisir de suivre cette voie me permettra de réussir le concours de l'agrégation de Mathématiques, dans les meilleures conditions. Je vais revenir avec vous sur les grandes lignes de mon expérience au sein du Magistère.

Je suis arrivée à Orsay en septembre 2009, après mes deux années en classes préparatoires de Mathématiques spéciales. Mon année de L3 MFA s'est bien déroulée, dans le cadre exceptionnel du campus d'Orsay. Au premier semestre, j'ai suivi le tronc commun très intéressant qui regroupait les enseignements d'*Algèbre*, d'*Intégrale de Lebesgue* et de *Topologie et Calcul Différentiel*. Pour compléter cette formation, j'ai choisi les options : *Programmation, Algorithmique et Complexité* et *Mathématiques et Biologie*. Cette dernière option m'a permise de découvrir une partie des applications des mathématiques au sciences du vivant. Au second semestre, j'ai suivi des cours également très enrichissants tels que l'*Analyse de Fourier*, les *Equations Différentielles*, les *Fonctions Holomorphes* et les *Probabilités*. Pour compléter cette formation, se sont ajoutés le cours d'*Algèbre Effective et le Calcul Formel*, le cours spécifique de magistère, qui apportait des *Compléments de topologie et de théorie de la mesure*, ainsi que le *Projet*, un travail que j'ai eu la chance d'effectuer avec le mathématicien Daniel Perrin. Ce projet, aboutissant à une démonstration de la loi de réciprocity quadratique, m'a permis en plus de découvrir la Théorie de Galois. Pour clôturer cette année de licence et cette première année de magistère, j'ai choisi de faire mon apprentissage « hors murs » dans un Laboratoire d'Analyses Médicales à Valenciennes, sous la direction du biologiste Guillaume Obert. Pendant ce stage de trois semaines, j'ai pu découvrir une partie du rôle des mathématiques dans le domaine de la médecine comme l'analyse, le contrôle ou l'ajustement des données.

En septembre 2010, débute mon année de M1 MFA. Des choix importants concernant les matières proposées ont été indispensables. Mon but étant d'obtenir l'agrégation l'année suivante, j'ai décidé de suivre les cours d'*Analyse et Mathématiques Générales*, les cours de *Probabilités* et de *Statistiques* ainsi que le *MAO Probabilités-Statistiques*. En plus de ces enseignements, j'ai apprécié les deux cours spécifiques de magistère, à savoir une

Introduction à la Théorie Spectrale et à l'Analyse Harmonique ainsi qu'une *Introduction aux Systèmes Dynamiques*. Enfin, j'ai effectué un *Travail Encadré de Recherche*, avec le chercheur Francis Choucroun, dont le sujet était la classification des formes quadratiques sur l'ensemble \mathbb{Q} des nombres rationnels. Cette année bien terminée me fournit alors un bagage très solide pour entamer la préparation décisive qui est à suivre.

En septembre 2011, j'ai donc interrompu le magistère pour passer le M2 Formation des Professeurs Agrégés de mathématiques et me présenter au concours de l'agrégation externe, option « Probabilités et Statistiques ». Cette préparation a été très enrichissante, elle m'a permis de consolider les connaissances qui m'étaient acquises et de m'essayer à l'enseignement dans le cadre d'un stage pédagogique au lycée Parc-de-Vilgénis de Massy, en classe de Terminale S, encadré par Delphine Cheymol.

Finalement, j'ai réussi le concours de l'agrégation ! J'ai donc eu la possibilité de prendre un poste de fonctionnaire stagiaire pour la rentrée 2012, mais j'ai choisi le report de stage. Pourquoi ?

C'est le magistère qui m'a encouragée à continuer en M2 Recherche. Les trois années écoulées à Orsay m'ont motivée à découvrir ce qu'est vraiment le monde de la recherche en mathématiques. Au début, je pensais faire la spécialité « Probabilités et Statistiques », celle qui me semblait correspondre le mieux à mon parcours. Finalement, une nouvelle spécialité a été proposée : « Mathématiques pour les Sciences du Vivant ». Cette formation à l'interface de la biologie, de l'écologie et de la médecine, m'a tout de suite attirée et c'est pourquoi je l'ai choisie.

En septembre 2012, je suis donc revenue à Orsay pour suivre les enseignements du M2 MathSV, très complets dans les différentes branches des mathématiques au contact des sciences du vivant. Au premier semestre, j'ai suivi le cours accéléré introduisant les *Concepts fondamentaux de la biologie et de l'Ecologie*, très instructif, ainsi que les cours d'*Optimisation et Simulation Numérique*, de *Modélisation Déterministe*, d'*Apprentissage Statistique en grande dimension* et de *Processus Stochastiques*. En plus de cela, j'ai assisté au *Groupe de Travail*, un séminaire hebdomadaire de sensibilisation à la modélisation en sciences du vivant. Ce sont ces exposés qui ont déterminés le choix de mon parcours pour le second semestre : « Ecologie et modèles d'évolution ». Pour ce dernier semestre, j'ai sélectionné les quatre cours qui me plaisaient le plus (mes choix ont été validés par le responsable du parcours) : *Modèles Statistiques pour la Génomique*, *Statistiques Spatiales pour l'Environnement*, *Arbres aléatoires et modèles d'évolution* et enfin *Modélisation Stochastique des populations structurées*. En parallèle des cours, j'ai

eu la chance de pouvoir effectuer un *Projet* à l'interface entre mathématiques et sciences du vivant. Pour cette première expérience concrète, j'ai choisi un sujet en liaison avec l'écologie : « la dynamique de la dépression de consanguinité », qui était encadré par Emmanuelle Porcher (MNHM), Vincent Bansaye (Ecole Polytechnique) et Amandine Véber (Ecole Polytechnique). Cette collaboration a été très intéressante et enrichissante. A ce stade, ma décision était prise, j'avais choisi de ne pas poursuivre en thèse. Cependant, j'ai quand même pris plaisir à effectuer un stage de recherche de quatre mois. J'ai décidé de travailler avec Amandine Véber et Hélène Morlon, dans l'équipe *Modélisation pour l'évolution du Vivant*, au Centre de Mathématiques Appliquées de l'Ecole Polytechnique. Le sujet de mon mémoire consistait à donner des prédictions théoriques pour la similarité entre deux communautés écologiques séparées par une distance géographique donnée. Mon expérience au sein de cette équipe et du laboratoire fut excellente !

En septembre 2013, je vais avoir un poste de professeur de Mathématiques au lycée Fustel de Coulanges de Massy. Le Magistère de Mathématiques d'Orsay m'a permis d'atteindre mes objectifs et m'a poussée à les dépasser en effectuant un M2 Recherche. Pendant mon cursus, j'ai eu la chance de bénéficier d'une formation exceptionnelle, de rencontrer des Professeurs, des Enseignants-Chercheurs, des Etudiants... qui partagent ma passion, avec lesquels j'ai eu des échanges formidables. Je suis très heureuse et fière d'avoir suivi cette voie !

2 Rapports réalisés au cours du Magistère

:- Projet en L3, sous la direction de Daniel Perrin :

LOI DE RECIPROCITE QUADRATIQUE.

:- Stage de Magistère, sous la direction de Guillaume Obert :

INTERVENTION DES MATHÉMATIQUES AU SEIN D'UN LABORATOIRE D'ANALYSES MÉDICALES.

:- Travail Encadré de Recherche en M1, sous la direction de Francis Chou-
croun :

NOMBRES P-ADIQUES ET FORMES QUADRATIQUES.

:- Stage Pédagogique en M2 FPA, sous la direction de Delphine Cheymol,
au lycée Parc-de-Vilgenis de Massy, avec une classe de Terminale S.

:- Projet de Recherche en M2 MathSV, sous la direction de Emmanuelle
Porcher, Vincent Bansaye et Amandine Véber :

DYNAMIQUE DE LA DEPRESSION DE CONSANGUINITE.

:- Stage de Recherche en M2 MathSV, sous la direction de Amandine
Véber et Hélène Morlon, au Centre de Mathématiques Appliquées de l'Ecole
Polytechnique :

PREDICTIONS THEORIQUES POUR LA SIMILARITE ENTRE DEUX COMMUNAUTES ECOLOGIQUES SEPARÉES PAR UNE DISTANCE GEOGRAPHIQUE DONNÉE.

3 Introduction à un domaine de recherche :

Modélisation de β -diversité sous théorie « neutre »

3.1 Introduction

La β -diversité est une mesure de la biodiversité qui consiste à comparer la diversité des espèces entre écosystèmes. Ceci suppose de comparer le taux de variation en composition d'espèces parmi des communautés.

Ici, nous allons nous intéresser au renouvellement des espèces en forêt tropicale.

Au cours des 20 dernières années, la compréhension des processus maintenant la diversité des arbres dans une région tropicale donnée a bien augmenté grâce à l'établissement de grandes parcelles où la floristique a été répertoriée.

Nous avons moins d'informations sur les facteurs qui gouvernent la β -diversité, comme en témoigne la divergence en composition des espèces d'arbre entre parcelles de forêt à différents endroits. En effet, le renouvellement des espèces parmi les arbres tropicaux n'est qu'en voie d'être élucidé.

Pour mesurer la β -diversité, nous allons considérer la relation distance-décroissance, un des patrons les plus décrits et étudiés en écologie, mais dont il existe très peu de prédictions théoriques. Cette relation dit que la similarité en composition des communautés écologiques décroît avec la distance géographique qui les sépare. Il est alors intéressant d'obtenir des informations pour cette relation, pour différents modèles de la biodiversité.

Pour débiter, nous avons utilisé une extension de la *théorie neutre de la biodiversité* des arbres qui a été développée par Hubbell (2001). Cette *théorie neutre de la biodiversité* inspirée de la biogéographie insulaire, remet complètement en cause l'idée d'une répartition selon les niches en considérant les espèces comme équivalentes. C'est une théorie mécaniste basée sur les mécanismes de base de dynamique des espèces (démographie, spéciation, extinction). La répartition des espèces est expliquée par la stochasticité démographique de la dispersion et le phénomène d'échantillonnage. Nous pouvons distinguer dans cette théorie deux échelles, la métacommunauté et la population locale.

Nous avons donc considéré les espèces comme équivalentes, dans le but de montrer comment la spéciation et la limitation de dispersion des graines peut contribuer au renouvellement des espèces. Ensuite, nous pourrions par

exemple considérer le cas où les espèces ont différentes capacités de dispersion...

Ici, nous n'allons considérer que des modèles « neutres », à savoir des modèles qui ne prennent pas en compte les différences d'adaptation des espèces à l'environnement. Les espèces seront considérées équivalentes du point de vue de leur démographie et de leur faculté à se diversifier pour former de nouvelles espèces.

Il existe un problème équivalent en *génétique des populations*, une discipline qui a été initiée dans les années 1920 à 1940 par Ronald Fisher, J.B.S. Haldane et Sewall Wright. La génétique des populations est une application des principes fondamentaux de la génétique mendélienne à l'échelle des populations. Il s'agit de l'étude de la distribution et des changements de la fréquence des versions d'un gène (allèles) dans les populations d'êtres vivants, sous l'influence des « pressions évolutives » (sélection naturelle, dérive génétique, mutations, et migration).

Si on considère une espèce et un gène d'intérêt tel que les individus peuvent porter des allèles différents de ce gène, alors lorsque la population a une structure spatiale, les fréquences alléliques deviennent une fonction de la coordonnée géographique.

Pour décrire les corrélations entre les diversités locales en fonction de la distance qui les sépare, des outils existent déjà. Si la population est représentée par un nuage de points dans \mathbb{R}^2 qui évolue au cours du temps, grâce à des événements de reproduction, alors nous pouvons décrire explicitement les arbres généalogiques reliant un échantillon d'individus. Dans ce cas, nous utilisons la relation de dualité que nous avons entre le modèle de reproduction et ses généalogies (*théorie de la coalescence*).

La *théorie de la coalescence* est un modèle rétrospectif de génétique des populations. Son objectif est de suivre l'évolution de tous les allèles d'un gène donné de tous les individus d'une population, jusqu'à une seule copie ancestrale, appelée ancêtre commun le plus récent. Les relations d'hérédité entre les allèles sont représentées sous la forme d'un arbre similaire à un arbre phylogénétique. Cet arbre est aussi appelé coalescent, et la compréhension des propriétés statistiques du coalescent sous différentes hypothèses forme la base de la théorie de la coalescence. Le coalescent utilise des modèles de dérive génétique, en remontant le temps pour reconstruire la généalogie des ancêtres. La théorie mathématique de la coalescence a été élaborée indépendamment par plusieurs chercheurs au début des années 1980, mais la formulation définitive est attribuée à John Kingman (1982).

En supposant la taille de la population finie, si nous ne nous intéressons plus aux allèles mais seulement aux relations de parenté entre les individus, le coalescent peut alors être interprété comme la généalogie d'un échantillon d'individus dans le modèle.

L'approche coalescente est rétrospective : au lieu de travailler en avant dans le temps, nous travaillons en arrière dans le temps, en retraçant les lignées des individus de l'échantillon jusqu'à ce qu'elles mutent ou coalescent. Lorsque nous regardons en arrière dans le temps, les lignées fusionnent, ce qui signifie qu'en avant dans le temps, la coalescence est un cas où une lignée se divise en deux branches. Deux individus au temps t ont le même allèle si leurs ancêtres au temps 0 ont le même allèle et qu'il n'y a pas eu de mutation entre-temps. En utilisant cette dualité, nous sommes alors en mesure d'exprimer les corrélations spatiales entre les diversités locales en termes de probabilité de rencontre de deux marches aléatoires (les lignées ancestrales) avant qu'apparaisse une mutation.

3.2 Un modèle neutre de β -diversité

D'après l'article de Jérôme Chave et Egbert G. Leigh Jr., paru en 2002 dans la revue « Theoretical Population Biology », *A spatially Explicit Neutral Model of β -Diversity in Tropical Forests*.

Pour prédire le renouvellement des espèces en forêt tropicale, nous considérons deux parcelles séparées par une distance r et nous nous intéressons à la probabilité que deux arbres pris au hasard dans chacune des parcelles soient de la même espèce : $F(r)$, que nous appelons *fonction de similarité*.

Remarque. Nous nous plaçons ici en dimension deux, mais des résultats similaires peuvent être trouvés en dimension un ou en dimension supérieure.

Dans un premier temps, nous nous intéressons à $F(x, y, t)$ avec $x = (x_1, x_2) \in \mathbb{R}^2$, $y = (y_1, y_2) \in \mathbb{R}^2$ et $t \in \mathbb{R}$, la probabilité qu'un arbre choisi au hasard au temps t dans le voisinage $dS_x = dx_1 dx_2$ incluant le site x soit de la même espèce qu'un arbre choisi au hasard dans un voisinage $dS_y = dy_1 dy_2$ incluant le site y , en supposant les arbres distribués suivant un processus ponctuel de Poisson d'intensité ρ , le temps de vie moyen de chaque arbre égal à une unité de temps (indépendamment de son espèce), le noyau de dispersion des graines $P : (x, y) \in \mathbb{R}^2 \mapsto P(x, y) \in \mathbb{R}$ identique pour chaque individu et la même probabilité de spéciation $\nu \in [0; 1]$ pour chaque nouvel arbre.

En trouvant la relation qui lie $F(x, y, t)$ à $F(x, y, t + dt)$ et en passant à la limite continue, nous obtenons le résultat suivant :

Proposition 1. *Nous avons l'équation principale :*

$$\frac{\partial F(x, y, t)}{\partial t} = -2F(x, y, t) + 2(1 - \nu)Q(x, y, t),$$

où $Q(x, y, t)$ représente la probabilité que si x meurt, le parent du descendant qui le remplace soit de la même espèce que y . Elle est donnée par :

$$Q(x, y, t) = \int_{u \in \mathbb{R}^2} F(u, y, t) P(u, x) dS_u + \frac{1 - F(y, y, t)}{\rho} P(y, x).$$

En supposant que la probabilité que deux arbres soient conspécifiques ne dépende que de leurs positions relatives, nous avons :

$$\begin{cases} F(x, y, t) &= F(x - y, t) &= F(z, t) \\ P(x, y) &= P(x - y) &= P(z) \end{cases}$$

où $z = (z_1, z_2) \in \mathbb{R}^2$ et nous obtenons :

Proposition 2. *À l'équilibre,*

$$F(z) = (1 - \nu) \int_{u \in \mathbb{R}^2} P(u) F(z - u, z - u) du + \frac{1 - F(0)}{\rho} P(z).$$

En supposant que $F(z_1, z_2)$ et $P(z_1, z_2)$ sont des fonctions uniquement de $r = \sqrt{z_1^2 + z_2^2}$, nous pouvons dériver une équation formelle : l'équation de Malécot.

En supposant le noyau de dispersion de variance finie $2\sigma^2$ et $r \gg \sigma$, nous arrivons au théorème suivant :

Théorème 1. *Si $r \gg \sigma$ et $P(r)$ décroît avec r plus vite qu'une exponentielle négative, on a :*

$$F(z_1, z_2) \sim \frac{1 - \nu}{2\nu} \sigma^2 \Delta F(z_1, z_2).$$

où Δ est l'opérateur Laplacien.

En passant en coordonnées polaires, nous obtenons l'équation de Bessel :

$$\frac{\partial^2 F(r)}{\partial r^2} + \frac{1}{r} \frac{\partial F(r)}{\partial r} = \frac{2\nu}{\sigma^2} F(r),$$

et nous savons que la seule solution positive vérifiant $\lim_{r \rightarrow \infty} F(r) = 0$ est :

$$F(r) = \frac{1}{c_0} K_0 \left(\frac{r\sqrt{2\nu}}{\sigma} \right),$$

où c_0 est une constante et K_0 est la fonction de Bessel modifiée de seconde espèce d'ordre 0.

Remarque. Cette équation de F est vérifiée pour $r \gg \sigma$, mais ne peut pas être valable pour r trop petit car $K_0(r)$ diverge logarithmiquement quand r approche zéro. Dans la suite, tous les arbres d'une même parcelle seront donc considérés comme étant à une certaine distance $\kappa \in \mathbb{R}^+$ les uns des autres (cette distance est appelée « échelle locale »). La probabilité précédemment notée $F(0)$ correspond alors à $F(\kappa)$.

Nous avons les équivalents suivants pour la fonction de Bessel modifiée de seconde espèce d'ordre 0 (Abramowitz M. et Stegun I. A., *Handbook of Mathematical Functions.*) :

$$K_0(r) \sim \sqrt{\frac{\pi}{2r}} \exp(-r) \text{ si } r \text{ grand } (r > \kappa, r > \sigma)$$

$$K_0(r) \sim -\log(r) \text{ si } r \text{ petit } (r < \kappa)$$

Nous pouvons alors écrire des équivalents pour $F(r)$.

Théorème 2. *Pour r grand :*

$$F(r) \simeq \frac{1}{c_0} \left(\frac{\pi\sigma}{2r\sqrt{2\nu}} \right)^{1/2} \exp \left(-\frac{r\sqrt{2\nu}}{\sigma} \right),$$

et pour r petit :

$$F(r) \simeq -\frac{1}{c_0} \log \left(\frac{r\sqrt{2\nu}}{\sigma} \right).$$

Cette fonction de similarité peut être vue différemment. Dans l'article de Barton et al. (2013), la formule de Wright-Malécot est utilisée pour approcher la fonction de similarité, alors vue comme fonction génératrice des temps de coalescence de deux lignées. En plus, une statistique permettant de l'estimer est proposée. Cette statistique est définie en fonction des fréquences alléliques (fréquences des différentes espèces), d'où l'idée d'un nouveau modèle que nous allons présenter, permettant d'obtenir d'autres (voire meilleures) prédictions, à partir de ce modèle...

3.3 Modèle neutre proposé

On considère deux parcelles disjointes A et B séparées l'une de l'autre par une distance r . La distribution spatiale des arbres est un processus ponctuel de Poisson d'intensité ρ . Le sort d'un arbre est indépendant de son espèce et la dispersion des graines est limitée à partir des parents.

Soit N_A le nombre de points dans A . La variable aléatoire N_A suit une loi de Poisson de paramètre $\rho|A|$, où $|A|$ représente l'aire de A .

Si A et B sont deux parcelles disjointes alors N_A et N_B sont des variables aléatoires indépendantes.

Soit $p_i(A)$ la probabilité qu'un arbre de la parcelle A soit d'espèce $i \in \mathbb{N}$,

$$p_i(A) = \frac{1}{N_A} \sum_{x \in A} \mathbb{1}_{x \text{ d'espèce } i}.$$

Cette expression contient une variable aléatoire, N_A . Pour faciliter les calculs, nous allons donc remplacer N_A par son espérance et finalement nous posons :

$$p_i(A) = \frac{1}{\rho|A|} \sum_{x \in A} \mathbb{1}_{x \text{ d'espèce } i}.$$

Remarque. La probabilité $p_i(A)$ peut valoir 0.

Nous allons nous intéresser à l'indice suivant :

Définition 1. On définit l'indice de dissimilarité :

$$\mathcal{F}(A, B) = \sum_i (p_i(A) - p_i(B))^2.$$

Cet indice va nous permettre d'estimer la similarité en composition d'espèces des deux parcelles A et B séparées par une distance donnée.

En effet, lorsque nous développons, nous voyons apparaître l'hétérozygotie entre A et B :

$$\mathcal{F}(A, B) = 2H_{AB} - H_A - H_B.$$

Nous allons supposer que $\mathcal{F}(A, B)$ ne dépend que de la distance entre les parcelles A et B , que les graines se dispersent à une distance limitée et que les arbres sont distribués de façon homogène, avec densité ρ .

Nous pouvons calculer l'espérance de $\mathcal{F}(A, B)$.

Proposition 3. *L'espérance de $\mathcal{F}(A, B)$ est donnée par :*

$$\mathbb{E}_r[\mathcal{F}(A, B)] = 2(F(\kappa) - F(r)) \underbrace{\frac{|A| + |B|}{\rho|A||B|}}_{:=C_1},$$

où κ détermine l'échelle « locale » et r la distance entre les parcelles A et B .

Le premier terme dépend de la distance qui sépare les parcelles A et B et le second terme dépend des aires de ces parcelles A et B ainsi que de l'intensité ρ selon laquelle les arbres sont répartis dans la forêt. Le seconde terme correspond à la dissimilarité locale.

Si la distance entre les parcelles augmente alors $F(r)$ diminue et donc $\mathbb{E}_r[\mathcal{F}(A, B)]$ augmente, donc la similarité diminue (plus les parcelles sont éloignées, moins les compositions en espèce des parcelles A et B sont similaires). Si ρ augmente ou les aires des parcelles augmentent alors $\mathbb{E}_r[\mathcal{F}(A, B)]$ diminue, donc la similarité augmente (plus il y a d'arbres ou plus les parcelles sont grandes, plus les espèces sont représentées).

En s'attaquant au calcul de la variance, nous remarquons la nécessité d'introduire une nouvelle fonction de similarité dite par paires, que nous notons $F^{(4)}$. Il s'agit de la probabilité que quatre arbres pris au hasard forment deux couples d'arbres conspécifiques. En supposant que cette probabilité ne dépend que des positions relatives des différents arbres, la fonction $F^{(4)}$ peut être définie sur \mathbb{R}^6 à valeurs dans $[0, 1]$. Ainsi, nous sommes en mesure de déterminer une expression de la variance en fonction de F , $F^{(4)}$ et r .

Avec ces résultats concernant la loi de notre indice de dissimilarité considéré ici, nous décidons de regarder s'il est possible d'obtenir des estimations pour F , ainsi que pour les paramètres qui régissent l'équation de Malécot, à savoir σ et ν .

À partir des résultats concernant l'espérance, si nous sommes en possession d'un jeu de données, nous sommes alors en mesure d'estimer $\mathcal{F}(A, B)$ et nous notons $\hat{\mathcal{F}}(A, B)$ cette estimation.

Pour commencer, nous pouvons alors estimer $F(0)$ par :

$$\hat{F}_0 = \frac{1}{C_1} \lim_{r \rightarrow 0} \mathbb{E}_r[\hat{\mathcal{F}}(A, B)]$$

et en posant $\hat{c} = (2 + C_1)\hat{F}_0$, nous pouvons alors estimer $F(r)$ par :

$$\boxed{\hat{F}(r) = \frac{\hat{c} - \mathbb{E}_r[\hat{\mathcal{F}}(A, B)]}{2}}$$

Maintenant, nous allons nous intéresser aux paramètres σ et ν , inconnus.

Dans le cas où r est petit, en utilisant l'équivalent de $F(r)$ (théorème 2), nous avons :

$$\mathbb{E}_r[\mathcal{F}(A, B)] \simeq c + \frac{2}{c_0} \log\left(\frac{r\sqrt{2\nu}}{\sigma}\right)$$

où $c = (2 + C_1)F(0)$.

Ainsi :

$$\frac{c_0}{2} (\mathbb{E}_r[\mathcal{F}(A, B)] - c) = \log(r) + \frac{\log(2)}{2} + \frac{\log(\nu/\sigma^2)}{2}$$

Nous pouvons alors écrire :

$$\boxed{\frac{\nu}{\sigma^2} = \exp [c_0 (E_r[\mathcal{F}(A, B)] - c) - 2 \log(r) - \log(2)]}$$

Dans le cas où r est grand, en utilisant l'équivalent de $F(r)$ (théorème 2), nous avons :

$$\mathbb{E}_r[\mathcal{F}(A, B)] \simeq c - \frac{2}{c_0} \left(\frac{\pi\sigma}{2r\sqrt{2\nu}}\right)^{1/2} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right)$$

Ainsi :

$$\left(\frac{\sqrt{2\pi}}{r}\right)^{1/2} c_0 (\mathbb{E}_r[\mathcal{F}(A, B)] - c) = \left(\frac{\sqrt{\nu}}{\sigma}\right)^{-1/2} \exp\left(-\sqrt{2}r \left(\frac{\sqrt{\nu}}{\sigma}\right)\right)$$

Nous pourrions alors obtenir des informations sur $\sqrt{\nu}/\sigma$, en résolvant l'équation :

$$x^{-1/2} \exp(-A_r x) - B_r = 0$$

où $A_r = \sqrt{2}r$ et $B_r = (\sqrt{2\pi}/r)^{1/2} c_0 (\mathbb{E}_r[\mathcal{F}(A, B)] - c)$.

Remarque. Il semble difficile d'obtenir des informations séparées en ce qui concerne les paramètres σ et ν , mais si nous avons des données, nous serons quand même en mesure d'estimer la quantité ν/σ^2 .

3.4 Conclusion

Des simulations nous ont permis de valider notre modèle neutre utilisant l'indice de dissimilarité que nous avons défini. Si nous avons accès à des données (la distribution spatiale, son intensité et les fréquences alléliques), grâce à ce modèle neutre de β -diversité, nous serions en mesure de proposer des estimations pour $F(0)$ et $F(r)$. Nous ne serions malheureusement pas capable de fournir des estimateurs pour le taux de flux de gènes σ^2 et le taux de mutation ν . Nous ne pourrions avoir que des informations sur le rapport ν/σ^2 , ce qui peut quand même être intéressant !

Pour aller plus loin, nous pourrions considérer le cas où les espèces ont différentes capacités de dispersion et voir en quoi ceci affecte la similarité entre deux communautés...

3.5 Références principales

Barton N., Etheridge A., Kelleher J. et Véber A., 2013, *Inference in two dimensions : allele frequencies versus lengths of shared sequence blocks*, Theor. Pop. Biol.

Chave J. et Leigh E. G., 2002, *A spatially Explicit Neutral Model of β -Diversity in Tropical Forests*, Theor. Pop. Biol. 62, 153-168.

Fisher R. A., 1930, *The genetical theory of natural selection*, Oxford Clarendon Press.

Haldane J. B. S., 1932, *The Causes of Evolution*, London, Longman Green.

Hubbell S. P., 2001, *A Unified Neutral Theory of Biodiversity and Biogeography*, Monographs in Pop. Biol., Princeton University Press.

Kingman J. F. C., 1982, *The coalescent*, Stochastic Proc. Appl. 13, 235-248.

Wright S., 1931, *Evolution in Mendelian populations*, Genetics 16, 97-159.

LA LOI DE RÉCIPROCITÉ QUADRATIQUE

Lucile Morelle

31 mai 2010

INTRODUCTION

Nous allons étudier les carrés du corps fini \mathbb{F}_q et notamment prouver la loi de réciprocité quadratique.

La recherche des carrés dans les corps finis est utile dans nombre de questions d'algèbre. Par exemple le quotient k^*/k^{*2} intervient dans l'étude du groupe linéaire, des formes quadratiques, etc., où k est un corps (commutatif, de caractéristique quelconque).

La loi de réciprocité quadratique est également essentielle dans de nombreuses questions d'arithmétique, par exemple la recherche des entiers de la forme $x^2 + dy^2$.

Expliquons le problème de la recherche de conditions nécessaires pour $d = 5$.

Soit e un entier strictement positif (si e vaut 0, on a de suite le résultat).

- Si e est de la forme $x^2 + 5y^2$ avec x et y dans \mathbb{N} , cela implique que e est un carré modulo 5, donc e est congru à 0 ou ± 1 modulo 5.

- Si il existe p premier qui divise e , on a, modulo p , $x^2 + 5y^2 \equiv 0$, alors $x^2 \equiv -5y^2$. Deux cas se présentent.

Si p divise x et y alors il existe a et b dans \mathbb{N} tels que $a^2 + 5b^2 = n/p^2$, on est alors ramené au nombre n/p^2 .

Sinon, le nombre -5 est un carré modulo p . On a $(\frac{-5}{p}) = (\frac{-1}{p})(\frac{5}{p}) = (\frac{-1}{p})(\frac{p}{5})$.

- Si p est congru à 1 modulo 4, alors on obtient que p est un carré modulo 5.

- Si p n'est pas congru à 1 modulo 4, alors on obtient que p n'est pas un carré modulo 5.

- Si il n'existe pas p premier qui divise e , alors e est premier. Dans ce cas, -5 est un carré modulo e . De plus, on sait que e est obligatoirement un carré modulo 5, on en déduit que e est aussi congru à 1 modulo 4.

Conjecture. Soit e un entier de la forme $x^2 + 5y^2$, avec x et y dans \mathbb{N} , on obtient les conditions nécessaires suivantes sur e :

1) Si e est premier, on a :

$$e = 5 \quad \text{ou} \quad e \equiv 1 [4] \quad \text{et} \quad e \equiv \pm 1 [5].$$

2) Si e est divisible par p premier tel qu'on ait :

$$\text{soit } p = 5 \quad \text{ou} \quad p \equiv 1 [4] \quad \text{et} \quad p \equiv \pm 1 [5], \quad (1)$$

$$\text{soit } p = 2 \quad \text{ou} \quad p \equiv -1 [4] \quad \text{et} \quad p \equiv \pm 2 [5]. \quad (2)$$

3) Si e est divisible par q^2 , où q est premier, ne vérifiant pas les conditions précédemment données sur p .

On constate que les cas 1. et 2.1. se rejoignent, on peut dire qu'il s'agit des bons nombres premiers.

En ce qui concerne le cas 2.2., on remarque que si on considère deux nombres premiers impairs distincts vérifiant ces conditions, leur produit vérifie les conditions du cas 1. (il est congru à 1 modulo 4 et à ± 1 modulo 5), on peut donc dire qu'il s'agit des nombres premiers moyens, car ils doivent aller deux par deux. Maintenant, dans le cas 3., il s'agit des autres nombres premiers tels que leur carré doit diviser e , on peut dire qu'il s'agit des mauvais nombres premiers.

Soit P l'ensemble des nombres premiers. On peut alors considérer trois sous-ensembles de nombres premiers qui correspondent aux trois cas évoqués ci-dessus :

$$P_1 = \{5\} \cup \{p \in P \mid p \equiv 1 [4] \text{ et } p \equiv \pm 1 [5]\}$$

$$P_2 = \{2\} \cup \{p \in P \mid p \equiv -1 [4] \text{ et } p \equiv \pm 2 [5]\}$$

$$P_3 = P \setminus (P_1 \cup P_2)$$

Soit e un entier dont la décomposition en nombres premiers est $\prod_{p \in P} p^{v_p(e)}$.

Finalement, e est de la forme $x^2 + 5y^2$ avec x et y dans \mathbb{N} si et seulement si on a les deux conditions suivantes :

- 1) $\forall p \in P_3, v_p(e)$ est pair,
- 2) $\sum_{p \in P_2} v_p(e)$ est pair.

Chapitre 1

Les carrés de \mathbb{F}_q

On pose $\mathbb{F}_q^2 = \{x \in \mathbb{F}_q \mid \exists y \in \mathbb{F}_q ; x = y^2\}$ et $\mathbb{F}_q^{*2} = \mathbb{F}_q^2 \cap \mathbb{F}_q^*$.

1.1 Différents cas suivant la caractéristique

1.1.1 Cas de la caractéristique égale à 2

On rappelle le résultat suivant :

Proposition 1.1.1. *Soit K un corps de caractéristique p strictement supérieure à 0.*

L'application $F : K \rightarrow K$ définie par $F(x) = x^p$ est un homomorphisme de corps appelé homomorphisme de Frobenius.

Si K est fini, c'est un automorphisme.

Si $K = \mathbb{F}_p$, c'est l'identité.

On déduit de la proposition précédente que, dans un corps fini de caractéristique 2, tout élément est un carré. On écarte définitivement ce cas.

1.1.2 Cas de la caractéristique supérieure à 2

Soit \mathbb{F}_q un corps fini de cardinal $q = p^n$ (où p est un nombre premier impair).

Théorème 1.1.2. *Les carrés de \mathbb{F}_q^* forment un sous-groupe d'indice 2 de \mathbb{F}_q^* .*

Ce sous-groupe est le noyau de l'homomorphisme $x \rightarrow x^{\frac{(q-1)}{2}}$, à valeurs dans $\{\pm 1\}$.

En d'autres termes, on a les deux suites exactes fondamentales suivantes :

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathbb{F}_q^* \xrightarrow{u} \mathbb{F}_q^{*2} \longrightarrow 1 \quad (1.1)$$

$$1 \longrightarrow \mathbb{F}_q^{*2} \longrightarrow \mathbb{F}_q^* \xrightarrow{v} \{\pm 1\} \longrightarrow 1 \quad (1.2)$$

avec $u(x) = x^2$ et $v(x) = x^{\frac{(q-1)}{2}}$.

Preuve du théorème :

- Soit $x \in \mathbb{F}_q^*$. On a $x^{q-1} = 1$, donc $(x^{\frac{q-1}{2}})^2 = 1$, donc $x = \pm 1$. Cela montre que v est à valeurs dans $\{\pm 1\}$.
- Si x est dans \mathbb{F}_q^{*2} , on a $x = y^2$, avec $y \in \mathbb{F}_q^*$, donc $x^{\frac{q-1}{2}} = y^{q-1} = 1$. Cela montre que $\mathbb{F}_q^{*2} \subset \text{Kerv}$. Comme l'équation $x^{\frac{q-1}{2}} = 1$ admet au plus $\frac{q-1}{2}$ racines, on a $\mathbb{F}_q^{*2} = \text{Kerv}$.
- Il en résulte que $|\text{Im}v| = (q-1)/|\text{Kerv}| = 2$, de sorte que v est surjective. \square

Remarques. 1) La première suite donne le nombre de carrés non nuls : $\frac{(q-1)}{2}$. En effet, on déduit de cette suite exacte, $|\mathbb{F}_q^{*2}| = \frac{q-1}{2}$.
 2) La seconde suite donne un critère pour que x , différent de 0, soit un carré, à savoir :

$$x^{\frac{(q-1)}{2}} = 1.$$

En effet, soit $X = \{x \in \mathbb{F}_q | x^{\frac{q-1}{2}} = 1\}$, on a $|X| \leq \frac{q-1}{2}$. (Un polynôme de degré $\frac{q-1}{2}$ a au plus $\frac{q-1}{2}$ racines). D'autre part, si x appartient à \mathbb{F}_q^{*2} , alors on a $x = y^2$, donc $x^{\frac{q-1}{2}} = y^{q-1}$ qui vaut 1 et donc \mathbb{F}_q^{*2} est inclus dans X . Il en résulte que X est égal à \mathbb{F}_q^{*2} .

Exemple. Le nombre -1 est un carré dans \mathbb{F}_q^* si et seulement si on a $q \equiv 1[4]$.

En effet, dire que -1 est un carré dans \mathbb{F}_q^{*2} équivaut à dire que $(-1)^{\frac{q-1}{2}}$ est égal à 1, donc que $\frac{q-1}{2}$ est pair, ou encore qu'on a $q \equiv 1[4]$.

Dans le même ordre d'idées on a :

Proposition 1.1.3. *Le nombre 2 est un carré de \mathbb{F}_q si et seulement si on a $q \equiv \pm 1[8]$.*

Preuve : Si ζ désigne une racine primitive huitième de l'unité dans une extension de \mathbb{F}_q , alors l'élément $y = \zeta^1 + \zeta^{-1}$ vérifie $y^2 = 2$.

En effet, comme ζ est une racine primitive huitième de l'unité, on a $\zeta^4 = -1$, donc $\zeta^2 + \zeta^{-2} = 0$. On en déduit $y^2 = \underbrace{\zeta^2 + \zeta^{-2}}_{=0} + 2 \underbrace{\zeta \zeta^{-1}}_{=1}$.

Par ailleurs, par Frobenius, on a $y^q = \zeta^q + \zeta^{-q}$

- Si $q \equiv \pm 1[8]$, on a $y^q = \zeta + \zeta^{-1}$, alors on obtient $y^q = y$ d'où $2^{\frac{q-1}{2}} = y^{q-1} = 1$.
- Si $q \equiv \pm 5[8]$, on a $y^q = \zeta^5 + \zeta^{-5}$, donc $y^q = -(\zeta + \zeta^{-1})$ car $\zeta^4 = -1$, alors on obtient $y^q = -y$, d'où $2^{\frac{q-1}{2}} = (y)^{q-1} = -1$.

D'après le critère donné par la seconde suite exacte, et comme q est impair, on en déduit que le nombre 2 est un carré de \mathbb{F}_q si et seulement si $q \equiv \pm 1[8]$. \square

Remarque. À partir de maintenant, on suppose que q est un nombre premier. On va désormais le noter p .

Définition 1.1.4. *Soit p un nombre premier différent de 2 et soit $x \in \mathbb{F}_p^*$. Le symbole de Legendre de x , noté $(\frac{x}{p})$, est défini par la formule suivante :*

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}}.$$

On rappelle une propriété importante du symbole de Legendre :

Proposition 1.1.5. *Le symbole de Legendre est multiplicatif, on a :*

$$\left(\frac{x}{p}\right) \left(\frac{y}{p}\right) = \left(\frac{xy}{p}\right).$$

Remarque. La seconde suite exacte nous indique que le symbole de Legendre est un homomorphisme en x et que x est un carré si et seulement si $\left(\frac{x}{p}\right) = 1$.

On a vu comment calculer le symbole de Legendre si $x = -1$ et $x = 2$, et on a rappelé sa multiplicativité.

Il reste à calculer $\left(\frac{n}{p}\right)$ lorsque n est un nombre premier impair. Ceci est l'objet de la loi de réciprocité quadratique.

1.2 La loi de réciprocité quadratique

Théorème 1.2.1. *Soient n et p deux nombres premiers impairs distincts. On a la formule :*

$$\left(\frac{n}{p}\right) = (-1)^{\frac{(n-1)(p-1)}{4}} \left(\frac{p}{n}\right).$$

Traduction du signe $(-)$ en termes de congruences modulo 4 :

On écrit : $n\mathbf{R}p$ si n est un carré (mod p),
c'est-à-dire si n est "reste quadratique" (mod p),
 $n\mathbf{N}p$ sinon.

On a donc : $n\mathbf{R}p \Leftrightarrow p\mathbf{R}n$ si p ou $n \equiv 1 \pmod{4}$,

$n\mathbf{R}p \Leftrightarrow p\mathbf{N}n$ si p et $n \equiv -1 \pmod{4}$.

Remarque. On peut utiliser cette formule pour calculer, par réductions successives, les symboles de Legendre.

Exemple numérique : Grâce à la loi de réciprocité quadratique, on peut par exemple faire le calcul suivant :

$$\begin{aligned} \left(\frac{29}{43}\right) &= \left(\frac{43}{29}\right) && (\text{car } 29 \equiv 1[4]) \\ &= \left(\frac{14}{29}\right) && (\text{car } 43 - 29 = 14) \\ &= \left(\frac{2}{29}\right) \left(\frac{7}{29}\right) && (\text{par multiplicativité}) \\ &= -\left(\frac{7}{29}\right) && (\text{car } 29 \equiv -3[8]) \\ &= -\left(\frac{29}{7}\right) && (\text{car } 29 \equiv 1[4]) \\ &= -\left(\frac{1}{7}\right) && (\text{car } 29 \equiv 1[7]) \\ &= -1. \end{aligned}$$

Le nombre 29 n'est donc pas un carré. Comme $43 \equiv -1[4]$, alors $-29 = 14$ en est un.

Chapitre 2

Démonstration de la loi de réciprocité.

2.1 Rappels sur la théorie de Galois.

- Groupe de Galois.

Définition 2.1.1. - Un *automorphisme de corps* K est un homomorphisme bijectif de K dans lui-même.

- Un *automorphisme d'une extension* L/K est un automorphisme de corps de L qui fixe les éléments de K (dont la restriction à K est l'identité).

Définition 2.1.2. Si $L \supset K$ est une extension de corps (on la note également L/K), le **groupe de Galois** de L sur K , $G = \text{Gal}(L/K)$, est le groupe des automorphismes de corps de L qui laissent fixe K .

Lemme 2.1.3. Soit σ un automorphisme de corps de K . L'ensemble $\{x \in K/\sigma(x) = x\}$ est un sous-corps de K . On le note $\text{Inv}(\sigma)$.

Preuve :

On a $\sigma(1) = 1$, donc $1 \in \text{Inv}(\sigma)$.

Soit $(a, b) \in (\text{Inv}(\sigma))^2$. On a $\sigma(a) = a$ et $\sigma(b) = b$, ce qui implique $\sigma(a + b) = \sigma(a) + \sigma(b) = a + b$ et $\sigma(ab) = \sigma(a)\sigma(b) = ab$. Par conséquent, $a + b$ et ab appartiennent à $\text{Inv}(\sigma)$.

Soit $x \in \text{Inv}(\sigma)$ avec $x \neq 0$. Alors $\sigma(x) = x$.

Or $\sigma(x^{-1}) = (\sigma(x))^{-1}$. Donc $\sigma(x^{-1}) = x^{-1}$, c'est-à-dire $x^{-1} \in \text{Inv}(\sigma)$. \square

Proposition 2.1.4. Soit K un corps, P son sous-corps premier. On a $\text{Aut}(K) = \text{Gal}(K/P)$.

Preuve :

Évidemment, on a $\text{Gal}(K/P) \subseteq \text{Aut}(K)$.

Soit $\sigma \in \text{Aut}(K)$. $\text{Inv}(\sigma)$ est un sous-corps de K donc, puisque P est le plus petit sous-corps de K , on a $P \subseteq \text{Inv}(\sigma)$, c'est-à-dire : $(\forall x \in P, \sigma(x) = x)$. Ainsi σ appartient à $\text{Gal}(K/P)$. \square

Définition 2.1.5. Soient K un corps, L une extension de K .

L'ensemble $F = \{x \in L \mid \forall g \in \text{Gal}(L/K), g(x) = x\}$ est un sous-corps de L qui contient K et est appelé le **corps fixe** de $\text{Gal}(L/K)$.

Proposition 2.1.6. Avec les notations précédentes, avec $K \subset F \subset L$, on a $Gal(L/K) = Gal(L/F)$.

Preuve :

Si l'automorphisme σ de L laisse fixe tous les éléments de F , alors puisque $K \subseteq F$, σ laisse fixe tous les éléments de K . Ainsi $Gal(L/K) \supseteq Gal(L/F)$.

Si $g \in Gal(L/K)$, alors par définition même de F , l'automorphisme g de L laisse fixe tous les éléments de F , donc $g \in Gal(L/F)$. Ainsi $Gal(L/K) \subseteq Gal(L/F)$.
□

Proposition 2.1.7. Soit H un sous-groupe d'un groupe de Galois $Gal(F/K)$, alors l'ensemble L^H des éléments fixés par tous les éléments de H est un sous-corps de L qui contient K . Il s'agit du **corps fixe de H** .

• **Ordre du groupe de Galois.**

Théorème 2.1.8. (de l'ordre du groupe de Galois). Soient K un corps, L une extension de degré fini de K . Alors on a l'inégalité : $|Gal(L/K)| \leq [L : K]$.

Preuve :

Raisonnons par l'absurde : notant $n = [L : K]$, supposons $|Gal(L/K)| > n$.

Alors il existe $(n+1)$ K -automorphismes distincts $\sigma_1, \dots, \sigma_{n+1}$ de L . Soit (x_1, \dots, x_n) une base du K -espace vectoriel L . Considérons la matrice $M = (\sigma_j(x_i)) \in M_{n,n+1}(L)$. Les vecteurs colonnes C_1, \dots, C_{n+1} de M sont $n+1$ vecteurs du L -espace vectoriel L^n de dimension n et sont donc liés :

$$\exists(\lambda_1, \dots, \lambda_{n+1}) \in L^{n+1} \setminus \{0\} \text{ tels que } \sum_{j=1}^{n+1} \lambda_j C_j = 0,$$

c'est à dire,

$$\exists(\lambda_1, \dots, \lambda_{n+1}) \in L^{n+1} \setminus \{0\} \text{ tels que } \forall i \in [1; n], \sum_{j=1}^{n+1} \lambda_j \sigma_j(x_i) = 0.$$

L'application K -linéaire $\sum_{j=1}^{n+1} \lambda_j \sigma_j$ de L dans lui-même prend la valeur 0 sur chacun des vecteurs de la famille (x_1, \dots, x_n) qui est une base du K -espace vectoriel L . Elle est donc nulle. Ainsi $(\lambda_1, \dots, \lambda_{n+1}) \in L^{n+1} \setminus \{0\}$, et $\sum_{j=1}^{n+1} \lambda_j \sigma_j = 0$.

Or $\sigma_1, \dots, \sigma_{n+1}$ sont $(n+1)$ homomorphismes distincts du groupe multiplicatif L^* dans le groupe multiplicatif L^* ; donc ils forment une famille libre sur L , en admettant le théorème d'Artin.

D'où la contradiction. □

Définition 2.1.9. Soit K un corps. On appelle **extension galoisienne finie** de K toute extension L de K de degré fini vérifiant $|Gal(L/K)| = [L : K]$.

Exemples. 1) Tout corps est une extension galoisienne finie de lui-même (puisque $Gal(K/K) = \{id_K\}$ et $[K : K] = 1$).

2) L'extension $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ n'est pas galoisienne.

Théorème 2.1.10. Soient K un corps, L une extension de degré fini de K , F le corps fixe de $Gal(L/K)$.

1) On a l'égalité $|Gal(L/K)| = |Gal(L/F)| = [L : F]$.

2) L'extension L/K est galoisienne si et seulement si $K = F$.

Nous allons admettre ce résultat ici.

• **Normalité. Séparabilité.**

Définition 2.1.11. Soient K un corps et L une extension algébrique de K . On dit que L est une **extension normale** de K si, et seulement si, il existe P dans $K[X]$ tel que $L = D_K(P)$.

Il revient au même de dire que chaque fois qu'un polynôme $P(X)$ irréductible de $K[X]$ a une racine dans L , $P(X)$ a toutes ses racines dans L ($P(X)$ est scindé sur L).

Définitions 2.1.12. - Soient K un corps et L une extension de K . Un élément α de L algébrique sur K est dit **séparable** sur K s'il est racine simple de son polynôme minimal.

- Une extension algébrique L d'un corps K est dite **séparable** si tout élément de L est séparable sur K .

- Un polynôme irréductible de $K[X]$ n'ayant aucune racine multiple dans une extension de K est dit **séparable**.

Remarques. 1) L'extension finie L/K est automatiquement séparable si K est fini ou de caractéristique 0.

2) L'extension finie L/K est automatiquement normale sur un corps K fini. Si on a $L = D_K(P)$ avec P séparable, on pose $Gal_K(P) = Gal(L/K)$.

Dans le cas galoisien, on dispose du théorème de Galois suivant :

Théorème 2.1.13. Soit L une extension galoisienne de dimension finie sur K et G son groupe de Galois.

Soit H un sous-groupe de G et L^H le corps fixe de H . Alors :

- On a L qui est une extension galoisienne de L^H et H , le groupe de Galois de l'extension $L^H \subset L$.

- On a deux bijections décroissantes réciproques de l'ensemble des sous-groupes de G sur l'ensemble \mathcal{E} des extensions intermédiaires $K \subset M \subset L$:

• $\phi : \mathcal{G} \longrightarrow \mathcal{E}$ qui à H associe L^H ,

• $\psi : \mathcal{E} \longrightarrow \mathcal{G}$ qui à M associe $Gal(L/M)$.

- L'extension L^H de K est galoisienne si et seulement si H est un sous-groupe distingué de G . Alors le groupe de Galois de L^H sur K est isomorphe au groupe quotient G/H .

Ce théorème établit une bijection entre sous-groupes de G et sous-extensions de L/K .

2.2 Galois sur un corps fini.

Posons $q = p^n$. On a alors $\mathbb{F}_q = D_{\mathbb{F}_p}(X^q - X)$, avec $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Proposition 2.2.1. Le groupe de Galois $Gal(\mathbb{F}_q/\mathbb{F}_p)$ est cyclique d'ordre n engendré par l'homomorphisme de Frobenius F défini par $F(x) = x^p$.

Preuve :

• Tout élément σ de $Aut(\mathbb{F}_q)$ vérifie $\sigma(1_{\mathbb{F}_q}) = 1_{\mathbb{F}_q}$, donc laisse fixe le sous-corps premier \mathbb{F}_p de \mathbb{F}_q . L'automorphisme σ est donc un \mathbb{F}_p -automorphisme

de \mathbb{F}_q . Ainsi on a $Aut(\mathbb{F}_q) = Gal(\mathbb{F}_q/\mathbb{F}_p)$. Or \mathbb{F}_q est un espace vectoriel sur \mathbb{F}_p et $[\mathbb{F}_q : \mathbb{F}_p] = n$. Le théorème (de l'ordre du groupe de Galois) donne : $|Gal(\mathbb{F}_q/\mathbb{F}_p)| \leq [\mathbb{F}_q : \mathbb{F}_p]$. Ainsi on a l'inégalité $|Aut(\mathbb{F}_q)| \leq n$.

- Le Frobenius F défini par $F(x) = x^p$ est un automorphisme de corps de \mathbb{F}_q .
- On montre aisément par récurrence que pour tout k dans \mathbb{N} , pour tout x dans \mathbb{F}_q , on a $F^k(x) = x^{p^k}$. L'ordre de F est $\omega = \min\{k \in \mathbb{N}^* | \forall x \in \mathbb{F}_q, F^k(x) = x\}$. L'ensemble \mathbb{F}_q^* est un groupe d'ordre $p^n - 1$, donc pour tout x dans \mathbb{F}_q^* , on a $x^{p^n - 1} = 1_{\mathbb{F}_q}$, donc pour tout x dans \mathbb{F}_q , on a $x^{p^n} = x$, soit $F^n = id_{\mathbb{F}_q}$. Ainsi ω divise n . Or l'égalité $F^\omega = id_{\mathbb{F}_q}$ est équivalente à dire que pour tout x appartenant à \mathbb{F}_q , on a $x^{p^\omega} = x$. Le polynôme $X^{p^\omega} - X$ à coefficients dans le corps commutatif \mathbb{F}_q et de degré p^ω a donc $card(\mathbb{F}_q)$ racines, soit p^n racines. On a donc $p^\omega \geq p^n$, soit $\omega \geq n$. Ainsi on en déduit que $\omega = n$. \square

2.3 Galois pour une extension cyclotomique.

On rappelle que pour un polynôme cyclotomique Φ_n , de degré $\varphi(n)$, on a les résultats suivants :

- Si n est premier, $\varphi(n) = n - 1$.
- Soit K un corps. Si n est premier avec la caractéristique de K , alors on a $K_n = D_K(X^n - 1) = D_K(\Phi_n)$.

Proposition 2.3.1. *On a un homomorphisme injectif*

$$Gal_K(\Phi_n) = Gal(K_n/K) \longrightarrow (\mathbb{Z}/n\mathbb{Z})^*.$$

Preuve :

Soit ζ une racine n -ième de l'unité. Soit σ appartenant à $Gal(K_n/K)$. On a $\zeta^n = 1$ et $\zeta^m \neq 1$ pour $0 < m < n$, d'où la même relation pour $\sigma(\zeta)$. Cela montre que $\sigma(\zeta)$ est une racine primitive n -ième de l'unité. Elle s'écrit donc $\sigma(\zeta) = \zeta^{a_\sigma}$ avec a_σ un entier premier à n et ζ^{a_σ} ne dépend que de la classe $\overline{a_\sigma}$ modulo n . On a ainsi l'application suivante :

$$\begin{array}{ccc} \phi : Gal(K/K_n) & \longrightarrow & (\mathbb{Z}/n\mathbb{Z})^* \\ \sigma & \longmapsto & a_\sigma \end{array}$$

On vérifie que c'est un homomorphisme de groupes :

$$\sigma_{\overline{a}} \circ \sigma_{\overline{b}}(\zeta) = (\zeta^a)^b = \zeta^{ab} = \sigma_{\overline{ab}}(\zeta), \text{ donc } \sigma_{\overline{a}} \circ \sigma_{\overline{b}} = \sigma_{\overline{ab}}.$$

De plus, ϕ est injectif. En effet, si $a_\sigma = 1$, on a $\sigma(\zeta) = \zeta$ et comme $\sigma|_K = Id_K$, $\sigma = Id$. Donc $Gal_K(\Phi_n) = Gal(K_n/K) \longrightarrow (\mathbb{Z}/n\mathbb{Z})^*$ est un homomorphisme injectif. \square

Proposition 2.3.2. *Avec les notations précédentes, si on décompose Φ_n en produit de polynômes irréductibles sur K : $\Phi_n = P_1 \dots P_r$, les P_i ont tous le même degré d et on a $d = [K_n : K] = |Gal(\Phi_n)|$. On a $\varphi(n) = rd$.*

Preuve :

On a $K \subset K_n = D_K(X^n - 1) = K(\zeta)$, où ζ est une racine primitive n -ième de l'unité. Soit $\Phi_n(X)$ le polynôme cyclotomique qui appartient à $K[X]$ et est de degré $\varphi(n)$. On décompose Φ_n en produit de polynômes irréductibles sur K : $\Phi_n = P_1 \dots P_r$, avec $r \in \mathbb{N}$. Soit ζ_i une racine de P_i . C'est une racine de Φ_n ,

donc une racine primitive n -ième de 1. On a donc $K(\zeta_i) = K_n$. Par ailleurs, on a $K(\zeta_i) \simeq K[X]/P_i$, d'où $d^\circ P_i = [K_n : K]$. Ainsi on a donc que les P_i sont des polynômes irréductibles de même degré $d = [K_n : K] = |\text{Gal}(\Phi_n)|$, car il s'agit d'une extension galoisienne finie. Et on a immédiatement $\varphi(n) = rd$. \square

On a le critère suivant pour que Φ_n soit irréductible :

Corollaire 2.3.3. *Soient K un corps et Φ_n le polynôme cyclotomique de degré $\varphi(n)$. Le polynôme Φ_n est irréductible sur K si et seulement si il n'a pas de racine dans les extensions L de K telles que $[L : K] \leq \frac{n-1}{2}$.*

Preuve :

Si Φ_n est irréductible et si x appartenant à L est une racine de Φ_n , $K[x]$ est un corps de rupture de Φ_n , donc de degré $\varphi(n)$, donc on a $[L : K] \geq n - 1$. Réciproquement, si Φ_n n'est pas irréductible, $\Phi_n = PQ$ et $d^\circ P$ ou $d^\circ Q$ est inférieur ou égal à $(n-1)/2$. Si on a $d^\circ P \leq \frac{n-1}{2}$ et si P' est un facteur irréductible de P , Φ_n aura une racine dans un corps de rupture de P' , de degré inférieur ou égal à $(n-1)/2$. \square

Remarque. Avec les notations précédentes, le polynôme Φ_n est irréductible si, et seulement si, on a $[K_n : K] = |\text{Gal}(\Phi_n)| = \varphi(n)$.

2.4 Galois et permutations.

Voici une propriété importante du groupe de Galois :

Proposition 2.4.1. *Soit L/K une extension galoisienne et soient $x, y \in L$. Les conditions suivantes sont équivalentes :*

- 1) *Il existe $\sigma \in \text{Gal}(L/K)$ tel que $y = \sigma(x)$,*
 - 2) *x et y ont même polynôme minimal sur K .*
- On dit alors que x et y sont conjugués sur K .*

Preuve :

- Montrons que 1) implique 2) :

Soient x et y dans L , on suppose qu'il existe $\sigma \in \text{Gal}(L/K)$ tel que l'on ait $y = \sigma(x)$. Soient P_1 et P_2 les polynômes minimaux respectifs de x et y sur K . On suppose que P_1 et P_2 sont des polynômes irréductibles, distincts, sur K . On a $P_2(y) = 0$, or $y = \sigma(x)$ donc on obtient $P_2(y) = P_2(\sigma(x)) = 0$. Comme on se place sur K , et que σ appartient à $\text{Gal}(L/K)$, on a $\sigma|_K = \text{id}_K$, alors on a $P_2(y) = P_2(\sigma(x)) = P_2(x) = 0$, et donc P_1 divise P_2 , car P_1 est le polynôme minimal de x sur K . De la même façon, on montre que P_2 divise P_1 , car P_2 est le polynôme minimal de y sur K . Les éléments x et y ont le même polynôme minimal sur K .

- Montrons que 2) implique 1) :

Soient x et y dans L , qui ont le même polynôme minimal P sur K . Soient $K(x)$ et $K(y)$ des extensions algébriques. Tout élément a appartenant à $K(x)$ a une unique expression sous la forme :

$$a = a_0 + a_0x + \cdots + a_nx^n \quad (a_1, \dots, a_n \in K)$$

où $n = d^\circ P - 1$. On définit l'application $\phi : K(x) \rightarrow K(y)$ par

$$\phi(a) = a_0 + a_1y + \cdots + a_ny^n.$$

Vérifions que ϕ est un isomorphisme : on a immédiatement

$$\phi(a + b) = \phi(a) + \phi(b).$$

Montrons à présent que $\phi(ab) = \phi(a)\phi(b)$, pour tout a et b dans $K(x)$. On prend $a = f(x)$, $b = g(x)$, $ab = h(x)$; où f, g, h sont des polynômes sur K de degrés inférieurs à $d^\circ P$. Alors

$$f(x)g(x) - h(x) = ab - ab = 0.$$

Or comme le polynôme minimal de x divise tout polynôme s'annulant en x , alors P divise $fg - h$, et donc il existe un polynôme Q sur K tel que $fg = PQ + h$. On a $d^\circ h < d^\circ P$, ce qui implique que h est le reste de la division euclidienne de fg par P . Avec le même raisonnement, on peut avoir $f(y)g(y) = h(y)$, d'où

$$\phi(ab) = h(y) = f(y)g(y) = \phi(a)\phi(b),$$

donc ϕ est un isomorphisme. De plus, comme ϕ est l'identité sur K , les deux extensions $K(x)$ et $K(y)$ sont isomorphes. On a clairement $\phi(x) = y$. Avec x et y dans L , et comme on a $L = D_K(P) = D_{K(x)}(P) = D_{K(y)}(P)$, grâce à l'unicité du corps de décomposition, on peut prolonger ϕ en $\sigma \in \text{Gal}(L/K)$ tel que $y = \sigma(x)$. \square

Remarque. Soient K un corps et $P \in K[X]$ un polynôme de degré m . Posons $L = D_K(P)$ et supposons que P a m racines distinctes x_1, \dots, x_m dans L (on dit alors que P est séparable et l'extension L/K est galoisienne). Alors, on voit aussitôt que le groupe de Galois $G = \text{Gal}(P) := \text{Gal}(L/K)$ permute les x_i et on peut l'identifier à un sous-groupe du groupe symétrique \mathfrak{S}_m . Précisément, si $P = P_1 \cdots P_r$ avec les P_i irréductibles sur K , la proposition précédente montre que les orbites dans l'action de G sur les x_i sont exactement les paquets de racines d'un même P_j .

Corollaire 2.4.2. Soit $P \in K[X]$ un polynôme séparable de degré m , soient x_1, \dots, x_m ses racines et soit $P = P_1 \cdots P_r$ sa décomposition en produit d'irréductibles, avec $d^\circ P_i = d_i$. Supposons $\text{Gal}(P)$ cyclique engendré par σ . Alors, comme permutation des x_i , l'élément σ est produit de r cycles disjoints d'ordre d_i .

Preuve :

Notons dans un premier temps que dire qu'une permutation σ agissant sur un ensemble X est produit de r cycles d'ordre d_i signifie que le groupe cyclique S engendré par σ admet exactement r orbites de cardinal d_i sur X .

Soient X l'ensemble des racines de P et $S = \langle \sigma \rangle$.

- Si x est une racine de P_i , on a $P_i(x) = 0$ et comme P est à coefficients dans K , $P(\sigma(x)) = 0$, de sorte que l'orbite de x est incluse dans l'ensemble des racines de P_i .

- Inversement, si x et y sont deux racines de P_i , on considère les corps $K[x]$ et $K[y]$ inclus dans L . Ce sont deux corps de rupture de P_i sur K . Il existe donc un K -isomorphisme de $K[x]$ sur $K[y]$ qui envoie x sur y . De plus, comme L est le corps de décomposition de P sur K , donc sur $K[x]$ et $K[y]$, l'isomorphisme s'étend en un K -isomorphisme de L sur lui-même. On a donc trouvé un élément u dans G tel que $u(x) = y$, mais comme σ engendre le groupe de Galois, il existe i tel que $\sigma^i(x) = y$ et x et y sont dans la même orbite sous S .

Le polynôme P étant séparable, décomposable en r polynômes irréductibles $P_1 \cdots P_r$, il existe donc exactement r orbites sous S , qui correspondent aux paquets disjoints des racines de P_i . Finalement, l'élément σ vu comme permutation des x_i , est produit de r cycles disjoints d'ordre d_i . \square

2.5 Discriminant.

On a vu que le groupe de Galois d'un polynôme séparable P de degré m s'injecte dans \mathfrak{S}_m . Pour savoir s'il s'injecte dans \mathfrak{A}_m (les permutations paires), on introduit le discriminant $\Delta(P)$:

Définition 2.5.1. Soient x_1, \dots, x_m les racines de P dans $L = D_K(P)$. On pose

$$\delta(P) = \prod_{i < j} x_i - x_j \text{ et } \Delta(P) = \delta(P)^2.$$

Le nombre $\Delta(P)$ est dans K^* et s'appelle le discriminant de P .

On a bien $\Delta(P)$ est dans K^* car le discriminant de P est invariant par le groupe de Galois.

Proposition 2.5.2. On a $\text{Gal}(P) \subset \mathfrak{A}_m$ si et seulement si $\Delta(P)$ est un carré de K , i.e. si $\delta(P)$ est dans K .

Preuve :

- On prend $G = \text{Gal}(P)$. L'ensemble G est un sous-groupe de \mathfrak{S}_m . Si $\Delta(P)$ est un carré de K , alors δ appartient à K , donc δ est fixe par G . De plus, les permutations impaires changent δ en $-\delta$ et comme la caractéristique de K est différente de 2, on a $\delta \neq -\delta$. Par conséquent, toutes les permutations dans G sont paires, et donc on a $G \subset \mathfrak{A}_m$.

- Réciproquement, si $G \subset \mathfrak{A}_m$, alors δ appartient au corps fixe de G qui est égal à K . On obtient que $\Delta(P)$ est un carré de K . \square

Remarque. La proposition suivante donne des moyens de calculer le discriminant. On l'applique ensuite dans le cas cyclotomique.

Proposition 2.5.3. Soit P' le polynôme dérivé de P et soient y_1, y_2, \dots, y_{m-1} ses racines dans une extension de K . On a les formules :

1)

$$\Delta(P) = (-1)^{n(n-1)/2} \prod_{i=1}^n P'(x_i),$$

2)

$$\Delta(P) = (-1)^{n(n-1)/2} n^n \prod_{j=1}^{n-1} P(y_j).$$

Preuve :

1) Si x_1, x_2, \dots, x_n désignent les racines de P , on a

$$P(x) = \prod_{k=1}^n (x - x_k) \text{ donc, } P'(x) = \sum_{k=1}^n \prod_{j \neq k} (x - x_j).$$

On en déduit pour tout k dans $[1, n]$,

$$P'(x_k) = \prod_{j \neq k} (x_k - x_j).$$

Par conséquent, on a

$$(-1)^{n(n-1)/2} \prod_{k=1}^n P'(x_k) = (-1)^{n(n-1)/2} \prod_{j \neq k} (x_k - x_j).$$

On a donc aussi

$$\Delta(P) = (-1)^{n(n-1)/2} \prod_{k=1}^n P'(x_k).$$

2) Si y_1, y_2, \dots, y_{n-1} désignent les racines de P' , on a

$$P'(x_j) = n \prod_{k=1}^{n-1} (x_j - y_k).$$

En utilisant la première formule, on trouve

$$\Delta(P) = (-1)^{n(n-1)/2} n^n \underbrace{\prod_{k=1}^{n-1} \prod_{j=1}^n (x_j - y_k)}_{=(-1)^n P(y_k)}.$$

Comme on a $\prod_{k=1}^{n-1} (-1)^n = ((-1)^n)^{n-1} = (-1)^{n(n-1)} = 1$, on a donc aussi

$$\Delta(P) = (-1)^{n(n-1)/2} n^n \prod_{k=1}^{n-1} P(y_k).$$

□

Proposition 2.5.4. *Soit n un nombre premier. Le discriminant de Φ_n vaut*

$$\Delta(\Phi_n) = (-1)^{\frac{n(n-1)}{2}} n^{n-2}.$$

Preuve :

Soit n un nombre premier, et $P(X) = \Phi_n(X) = (X^n - 1)/(X - 1)$ le polynôme cyclotomique d'indice n . Notons $\zeta = e^{\frac{2i\pi}{n}}$. Les racines de $P(X)$ dans \mathbb{C} sont les ζ^k , avec k appartenant à $[1, n - 1]$.

On a $P(X) = (X^n - 1)/(X - 1)$, alors on en déduit

$$P'(X) = \frac{nX^{n-1}(X - 1) - (X^n - 1)}{(X - 1)^2}.$$

Donc pour tout k dans $[1, n - 1]$, on a

$$P'(\zeta^k) = \frac{n\zeta^{k(n-1)}}{\zeta^k - 1} = \frac{n}{\zeta^k(\zeta^k - 1)},$$

d'où :

$$\prod_{k=1}^{n-1} P'(\zeta^k) = \frac{n^{n-1}}{\prod_{k=1}^{n-1} \zeta^k \cdot \prod_{k=1}^{n-1} (\zeta^k - 1)}.$$

Or les relations entre coefficients et racines d'un polynôme montrent que l'on a $\prod_{k=1}^{n-1} \zeta^k = (-1)^{n-1}$, et puisque les $\zeta^k - 1$, pour k dans $[1, n-1]$, sont les racines dans \mathbb{C} du polynôme $P(X+1)$, on a

$$\prod_{k=1}^{n-1} (\zeta^k - 1) = (-1)^{n-1} n.$$

On en déduit donc

$$\prod_{k=1}^{n-1} P'(\zeta^k) = n^{n-2},$$

d'où

$$\Delta(P) = (-1)^{(n-1)(n-2)/2} \prod_{k=1}^{n-1} P'(\zeta^k) = (-1)^{(n-1)(n-2)/2} n^{n-2}.$$

Remarquons que l'on a $(n-1)(n-2) = n(n-1) + 2(1-n)$, ce qui donne $(-1)^{\frac{(n-1)(n-2)}{2}} = (-1)^{\frac{n(n-1)}{2}} \cdot (-1)^{(1-n)} = (-1)^{\frac{n(n-1)}{2}}$, car n est un nombre premier impair. Finalement on a :

$$\Delta(\phi_n) = (-1)^{\frac{n(n-1)}{2}} n^{n-2}.$$

□

Remarque 2.5.5. Comme le nombre n est impair, on voit que $\Delta(\Phi_n)$ est un carré dans K si et seulement si $(-1)^{n(n-1)/2} n$ est un carré dans K .

Corollaire 2.5.6. *Soient p et n deux nombres premiers impairs distincts. On a $\text{Gal}_{\mathbb{F}_p}(\Phi_n) \subset \mathfrak{A}_{n-1}$ si et seulement si $(-1)^{\frac{n(n-1)}{2}} n$ est un carré de \mathbb{F}_p .*

Preuve :

Le polynôme Φ_n est de degré $n-1$. En vertu de la proposition 2.5.2., $\text{Gal}(\Phi_n)$ est contenu dans \mathfrak{A}_{n-1} si et seulement si $\Delta(\Phi_n)$ est un carré, ou encore si et seulement si $(-1)^{\frac{n(n-1)}{2}} n$ est un carré dans \mathbb{F}_p , d'après la remarque précédente.

□

2.6 Cyclotomie et corps finis.

Proposition 2.6.1. *Soit p premier et n un entier tel que p ne divise pas n . On décompose Φ_n en produit de r polynômes irréductibles sur \mathbb{F}_p , tous de degré d . Alors le nombre d est l'ordre de p dans $(\mathbb{Z}/n\mathbb{Z})^*$.*

Preuve :

On considère l'extension $\mathbb{F}_p \subset \mathbb{F}_q = D_{\mathbb{F}_p}(\Phi_n)$.

D'après la proposition 2.3.2., on a $d = [\mathbb{F}_q : \mathbb{F}_p]$, donc $q = p^d$ et le groupe de Galois, $\text{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ est cyclique d'ordre d , engendré par le Frobenius F défini

par $F(x) = x^p$.

D'après la proposition 2.3.1., on a un homomorphisme injectif

$$\theta : Gal_{\mathbb{F}_p}(\Phi_n) \longrightarrow (\mathbb{Z}/n\mathbb{Z})^*.$$

En particulier, on a $\theta(F) = p$. Comme θ est injectif, l'ordre de F dans $Gal(\mathbb{F}_q/\mathbb{F}_p)$ est égal à l'ordre de p dans $(\mathbb{Z}/n\mathbb{Z})^*$. Finalement, le nombre d est l'ordre de p dans $(\mathbb{Z}/n\mathbb{Z})^*$. \square

Le point suivant est l'une des clés de la démonstration de la loi de réciprocité :

Corollaire 2.6.2. *On suppose p et n premiers impairs distincts. Alors, avec les notations précédentes, p est un carré de $\mathbb{F}_n = \mathbb{Z}/n\mathbb{Z}$ si et seulement si r est pair.*

Preuve :

Soit p et n premiers impairs distincts. On a $rd = \varphi(n) = n - 1$.

- Si p est un carré de \mathbb{F}_n alors $p^{\frac{n-1}{2}} = 1$, or p est d'ordre d , donc d divise $(n-1)/2$, soit d divise $rd/2$. Il existe alors k dans \mathbb{Z} tel qu'on a $kd = rd/2$, on a alors $r = 2k$ et r est pair.
- Si r est pair alors il existe alors k dans \mathbb{Z} tel qu'on a $n-1 = 2kd$. De plus, p est d'ordre d , donc on a $p^{kd} = 1$, ce qui implique $p^{\frac{n-1}{2}} = 1$ et p est un carré de \mathbb{F}_n . \square

Remarque. Nous allons énoncer quelques conséquences de la proposition 2.6.1., bien que cela nous éloigne un peu de notre sujet.

Corollaire 2.6.3. *On a les équivalences :*

- 1) *Le polynôme Φ_n est scindé sur \mathbb{F}_p .*
- 2) *Le polynôme Φ_n a une racine dans \mathbb{F}_p .*
- 3) *On a $p \equiv 1 \pmod{n}$.*

Preuve :

- L'équivalence entre 1) et 2) résulte du fait qu'il s'agit ici de racines de l'unité et qu'alors, dès qu'on a une racine primitive, on les a toutes.
- Si p est congru à 1 modulo n , alors p ne divise pas n . On peut appliquer la proposition 2.6.1. et donc décomposer le polynôme Φ_n en produit de r polynômes irréductibles sur \mathbb{F}_p , tous de degré d , de telle sorte que d soit l'ordre de p dans $(\mathbb{Z}/n\mathbb{Z})^*$. Dans ce cas, dire que p est congru à 1 modulo n est équivalent à dire que d est égal à 1, donc que le polynôme Φ_n est scindé sur \mathbb{F}_p . Finalement, on a l'équivalence entre 3) et 1). \square

Corollaire 2.6.4. *Soit n supérieur ou égal à 2, il existe une infinité de nombres premiers congrus à 1 (mod n).*

Preuve :

Soit k un entier strictement positif et soit p un facteur premier de $\Phi_n(k!)$ (on prend k assez grand pour avoir $\Phi_n(k!)$ différent de ± 1). Le corollaire précédent montre que p est congru à 1 modulo n . De plus, on voit que p est nécessairement plus grand que k (sinon il diviserait $k!$ et, comme on a $\Phi_n(0) = 1$, on aurait une contradiction). Finalement, il existe une infinité de nombres premiers congrus à 1 modulo n . \square

Corollaire 2.6.5. *On a les équivalences :*

- 1) *Le groupe $(\mathbb{Z}/n\mathbb{Z})^*$ est un groupe cyclique.*
- 2) *On a $n = 2, 4, p^\alpha$ ou $2p^\alpha$ avec p premier impair.*
- 3) *Il existe p premier tel que le polynôme Φ_n est irréductible sur \mathbb{F}_p .*

Preuve :

- L'équivalence entre 1) et 2) résulte de la description de $(\mathbb{Z}/n\mathbb{Z})^*$.
- D'après le corollaire 2.3.3., si il existe p premier tel que le polynôme Φ_n est réductible sur \mathbb{F}_p , alors il existe un entier m avec $m \leq \varphi(n)/2$ tel que Φ_n ait une racine dans \mathbb{F}_{p^m} , c'est-à-dire tel que n divise $p^m - 1$, puisque $\mathbb{F}_{p^m}^*$ est cyclique d'ordre $p^m - 1$. Finalement, dire que Φ_n est réductible sur \mathbb{F}_p est équivalent à dire que p est d'ordre inférieur à $\varphi(n)$, dans $(\mathbb{Z}/n\mathbb{Z})^*$. Si $(\mathbb{Z}/n\mathbb{Z})^*$ n'est pas cyclique, cette condition est toujours réalisée, ce qui prouve que 3) implique 1). Réciproquement, si $(\mathbb{Z}/n\mathbb{Z})^*$ est cyclique et si on a un entier a tel que \bar{a} engendre $(\mathbb{Z}/n\mathbb{Z})^*$, en vertu du théorème de progression arithmétique, on peut supposer l'égalité entre a et p premier. Mais alors, comme p est d'ordre $\varphi(n)$ dans $(\mathbb{Z}/n\mathbb{Z})^*$, Φ_n est irréductible sur \mathbb{F}_p . On a donc prouvé l'équivalence entre 1) et 3). \square

2.7 Frobenius et décomposition en cycles.

Le résultat suivant est une conséquence du corollaire 2.4.2. :

Corollaire 2.7.1. *Soit p un nombre premier impair et n un entier tel que p ne divise pas n . Posons $\mathbb{F}_q = D_{\mathbb{F}_p}(\Phi_n)$ et soit F l'homomorphisme de Frobenius de \mathbb{F}_q . Alors, si d est le degré des facteurs irréductibles de Φ_n sur \mathbb{F}_p , F , vu comme élément de $\mathfrak{S}_{\varphi(n)}$ est un produit de r cycles disjoints d'ordre d .*

Preuve :

Pour commencer, on rappelle que l'on a $rd = \varphi(n)$.

La proposition 2.2.1 donne que $D_{\mathbb{F}_p}(\Phi_n)$ est engendré par le Frobenius F , défini par $F(x) = x^p$. En regardant F comme un élément de $\mathfrak{S}_{\varphi(n)}$, si d est le degré des facteurs irréductibles de Φ_n sur \mathbb{F}_p , alors d'après le corollaire 2.4.2., on obtient que F est un produit de r cycles disjoints d'ordre d . \square

On suppose à présent que n est un nombre premier (impair), de sorte que $\varphi(n) = n - 1 = rd$ est pair. On a alors :

Corollaire 2.7.2. *Le groupe de Galois $Gal(\Phi_n)$ sur \mathbb{F}_p est contenu dans le groupe alterné \mathfrak{A}_{n-1} si et seulement si $r = (n - 1)/d$ est pair, c'est-à-dire encore si p est un carré de \mathbb{F}_n .*

Preuve :

- Soit F appartenant à $Gal_{\mathbb{F}_p}(\Phi_n)$, que l'on suppose inclus dans \mathfrak{A}_{n-1} . Si d est impair, comme $\varphi(n) = rd$ est pair, alors le nombre r est pair. Si d est pair, F se décompose en r cycles disjoints de cardinal pair, qui sont donc impairs, comme F est pair, alors on obtient que r est pair.
- Soit F appartenant à $Gal_{\mathbb{F}_p}(\Phi_n)$, on suppose que r est pair. La permutation F est composée d'un nombre pair de cycles disjoints d'ordre d , alors la permutation F est nécessairement paire et on a $Gal_{\mathbb{F}_p}(\Phi_n)$ qui est contenu dans \mathfrak{A}_{n-1} . \square

2.8 Conclusion.

On considère deux nombres premiers impairs distincts p et n . On regarde le polynôme cyclotomique Φ_n sur le corps \mathbb{F}_p . On va montrer la loi de réciprocité en traduisant de deux façons différentes la condition $Gal(\Phi_n) \subset \mathfrak{A}_{n-1}$.

D'après le corollaire 2.5.5., on a $Gal(\Phi_n) \subset \mathfrak{A}_{n-1}$ si et seulement si $(-1)^{n(n-1)/2}n$ est un carré de \mathbb{F}_p .

D'après le corollaire 2.7.2., on a $Gal(\Phi_n) \subset \mathfrak{A}_{n-1}$ si et seulement si p est un carré de \mathbb{F}_n .

On en déduit donc que p est un carré de \mathbb{F}_n si et seulement si $(-1)^{n(n-1)/2}n$ est un carré de \mathbb{F}_p . On a donc :

$$\begin{aligned}
 p^{\frac{n-1}{2}} &= ((-1)^{\frac{n(n-1)}{2}} n)^{\frac{p-1}{2}} \\
 \iff \underbrace{p^{\frac{n-1}{2}}}_{= \left(\frac{p}{n}\right)} &= (-1)^{\frac{n(n-1)(p-1)}{4}} \underbrace{n^{\frac{p-1}{2}}}_{= \left(\frac{n}{p}\right)} \\
 \iff \left(\frac{p}{n}\right) &= \left(\frac{n}{p}\right) (-1)^{\frac{(n-1)(p-1)}{4}} n \\
 \iff \left(\frac{p}{n}\right) &= \left(\frac{n}{p}\right) (-1)^{\frac{(n-1)(p-1)}{4}} \quad \text{car } n \text{ est premier impair.}
 \end{aligned}$$

Ainsi on a donc démontré la loi de réciprocité quadratique. \square

ANNEXE

Rappels sur la caractéristique d'un corps

On rappelle d'abord que si K et K' sont des corps finis avec $K \subset K'$, $|K'|$ est une puissance de $|K|$.

Définition. Soit K un corps (quelconque). On appelle **sous-corps premier** de K le plus petit sous-corps de K .

Description du sous-corps premier :

Soit $\varphi : \mathbb{Z} \rightarrow K$ l'homomorphisme d'anneaux défini par $\varphi(n) = n.1 = 1 + \dots + 1$ (n fois). Le noyau de φ est alors un idéal de \mathbb{Z} , donc on a $\text{Ker}\varphi = p\mathbb{Z}$ et comme $\mathbb{Z}/p\mathbb{Z} \simeq \text{Im}\varphi$, qui est inclus dans K , est intègre, $p\mathbb{Z}$ est un idéal premier. Donc p est nul ou un nombre premier.

Définition. Le nombre p , générateur de $\text{Ker}\varphi$ est appelé la **caractéristique** du corps K . La caractéristique d'un corps est donc 0, ou un nombre premier. On la note $\text{car}(K)$.

Remarques. 1) Si $\text{car}(K) = 0$, alors $\varphi(\mathbb{Z})$ est inclus dans K , donc comme on a $\varphi(\mathbb{Z}) \simeq \mathbb{Z}$, K est infini.

De plus, K contient le corps des fractions de \mathbb{Z} , i.e \mathbb{Q} qui est le sous-corps premier de K .

2) Si K est fini, alors on a $\text{car}(K) = p > 0$. Le sous-corps premier de K est $\mathbb{Z}/p\mathbb{Z}$. On le note aussi \mathbb{F}_p .

3) Si K est fini et si on a $\text{car}(K) = p > 0$, la remarque initiale montre que l'on a $q = |K| = p^n$.

Rappels sur les extensions de corps commutatifs

• Degré d'une extension. Éléments algébriques.

Définition. Soient K, L des corps avec $K \subset L$. On dit que L est une **extension** de K .

Exemples. Nous avons les inclusions suivantes : $\mathbb{R} \subset \mathbb{C}$, $\mathbb{R} \subset \mathbb{R}(T)$, $\mathbb{Q} \subset \mathbb{Q}(i)$...

Remarques. 1) Si $K \subset L$, L est un K -espace vectoriel.

2) Si $\dim_K L$ est finie, on pose $[L : K] = \dim_K L$ et $[L : K]$ s'appelle le degré de L sur K .

3) Si K et L sont des corps finis, on a $|L| = |K|^n$ avec $n = [L : K]$.

Théorème. (de la base télescopique). Soient $K \subset L \subset M$ des corps, $(e_i)_{i \in I}$ une base de L sur K , $(f_j)_{j \in J}$ une base de M sur L . Alors $(e_i f_j)_{(i,j) \in I \times J}$ est une base de M sur K .

Corollaire. (multiplicativité du degré). On a l'égalité suivante :

$$[M : K] = [M : L][L : K].$$

Preuve :

1) La famille $e_i f_j$ est libre sur K :

Si $\sum_{i,j} \lambda_{ij} e_i f_j = 0$, $\lambda_{ij} \in K$ on a $\sum_j f_j (\sum_i \lambda_{ij} e_i = 0)$, donc puisque f_j est une base de M sur L , on a pour tout j , $\sum_i \lambda_{ij} e_i = 0$; et puisque e_i est une base de L sur K , $\lambda_{ij} = 0$ pour tous i, j .

2) Elle engendre M : Soit $x \in M$, $x = \sum_j \mu_j f_j$, $\mu_j \in L$, et $\mu_j = \sum_i \lambda_{ij} e_i$ d'où $x = \sum_{i,j} \lambda_{ij} e_i f_j$ avec $\lambda_{ij} \in K$. \square

Définition. 1) Soit $K \subset L$ une extension et $A \subset L$. On dit que A **engendre** L sur K et on écrit alors $L = K(A)$ si L est le plus petit sous-corps de L contenant K et A .

Si A est fini, $A = \{\alpha_1, \dots, \alpha_n\}$, on note $L = K(\alpha_1, \dots, \alpha_n)$.

2) L'extension $K \subset L$ est dite **monogène** s'il existe $\alpha \in L$ tel que $L = K(\alpha)$.

Remarques. 1) Si K est inclus dans L et si α appartient à L , on note $K[\alpha]$ le sous-anneau de L engendré par K et α . On a $K[\alpha] \subset K(\alpha)$.

Attention, $K[\alpha]$ n'est pas en général isomorphe à l'anneau des polynômes $K[T]$, ni $K(\alpha)$ à $K(T)$.

2) On peut décrire $K[\alpha]$ et $K(\alpha)$:

Si $x \in K[\alpha]$, $x = P(\alpha)$ avec $P \in K[T]$, i.e. $x = a_n \alpha^n + \dots + a_0$, $a_i \in K$.

Si $x \in K(\alpha)$, $x = P(\alpha)/Q(\alpha)$ avec $Q(\alpha) \neq 0$, $P, Q \in K[T]$.

La différence entre $K[\alpha]$ et $K[T]$ (resp. $K(\alpha)$ et $K(T)$) vient du fait que l'on peut avoir $Q(\alpha) = 0$ avec $Q \in K[T]$, $Q \neq 0$.

De façon précise, il y a deux types d'éléments de L relativement à K :

Définition. Soit K qui est inclus dans L et α qui appartient à L . Soit $\varphi : K[T] \rightarrow L$ l'homomorphisme défini par $\varphi|_K = id_K$ et $\varphi(T) = \alpha$.

Si φ est injectif, on dit que α est **transcendant** sur K .

Sinon, on dit que α est **algébrique** sur K .

Ceci signifie qu'il existe un polynôme $P(T)$, non nul, tel que $P(\alpha) = 0$.

Exemples. 1) On peut montrer que e et π sont transcendants sur \mathbb{Q} (mais pas sur \mathbb{R}). Dans $K(T)$, T est transcendant sur K .

2) Les éléments : $\sqrt{2}$, i , $\sqrt[3]{2}$, \dots sont algébriques sur \mathbb{Q} .

Structure de $K(\alpha)$:

1) Si α est transcendant, on a $K[\alpha] \simeq K[T]$ et $K(\alpha) \simeq K(T)$ (donc $K(\alpha)$ est différent de $K[\alpha]$).

2) Si α est algébrique, soit $I = Ker \varphi$. C'est un idéal principal, $I = (P)$, $P \neq 0$ et on peut prendre P unitaire. Le polynôme P s'appelle alors le polynôme minimal de α sur K .

On a un isomorphisme : $\bar{\varphi} = K[T]/P \rightarrow K[\alpha]$

Comme $K[\alpha] \subset L$, $K[\alpha]$ est intègre, donc (P) premier, donc P irréductible dans $K[T]$ principal, donc (P) maximal. Donc $K[\alpha]$ est un corps et $K[\alpha] = K(\alpha)$.

On peut résumer les propriétés d'un élément algébrique dans le théorème suivant :

Théorème. Soit $K \subset L$ et $\alpha \in L$. Les propriétés suivantes sont équivalentes :

- 1) α est algébrique sur K , de polynôme minimal P .
- 2) $K[\alpha] = K(\alpha)$.
- 3) $\dim_K K[\alpha] < +\infty$; précisément $\dim_K K[\alpha] = d^{\circ}(P)$.

Preuve :

On a $1 \Rightarrow 2$, mais aussi $2 \Rightarrow 1$ (car $K(T) \neq K[T]$).

$3 \Rightarrow 1$ est clair également, car si α est transcendant $K[\alpha] \simeq K[T]$ est de dimension infinie sur K .

Enfin $1 \Rightarrow 3$ résulte de l'isomorphisme $K[T]/(P) \longrightarrow K[\alpha]$, si P est de degré n , $(1, \alpha, \dots, \alpha^{n-1})$ est une base de $K[\alpha]/K$. \square

Définition. 1) Une extension $K \subset L$ est dite **finie** si $[L : K] < +\infty$.

2) Une extension $K \subset L$ est dite **algébrique** si pour tout α dans L , α est algébrique sur K .

Théorème. Soit $K \subset L$ et $M = \{x \in L \mid x \text{ est algébrique sur } K\}$.

Alors M est un sous-corps de L .

Preuve :

Soient $\alpha, \alpha' \in M$; Considérons le sous-anneau $K[\alpha, \alpha']$ engendré par α et α' . On a $K[\alpha, \alpha'] = K[\alpha]K[\alpha']$ et donc comme α' est algébrique sur K donc a fortiori sur $K[\alpha]$, le théorème 2 montre que $K[\alpha]$ et $K[\alpha, \alpha']$ sont des corps. De plus le second théorème et la multiplicativité des degrés donnent $[K[\alpha, \alpha'] : K] < +\infty$. Mais alors comme $K[\alpha + \alpha']$ et $K[\alpha\alpha']$ sont inclus dans $K[\alpha, \alpha']$, ils sont de dimension finie sur K et $\alpha + \alpha'$ et $\alpha\alpha'$ sont donc algébriques d'où $\alpha + \alpha', \alpha\alpha' \in M$. \square

Exemple. Soit $A = \{\alpha \in \mathbb{C} \mid \alpha \text{ algébrique sur } \mathbb{Q}\}$. L'ensemble A est un corps, algébrique sur \mathbb{Q} , mais l'extension $\mathbb{Q} \subset A$ n'est pas finie.

Définition. Un corps K est dit **algébriquement clos** s'il vérifie l'une quelconque des propriétés suivantes :

- 1) Tout polynôme $P \in K[X]$ de degré ≥ 1 admet une racine dans K .
- 2) Tout polynôme $P \in K[X]$ est produit de polynômes de degré un.
- 3) Les éléments irréductibles de $K[X]$ sont les $X - a$, $a \in K$.
- 4) Si $K \subset L$ est algébrique, on a $L = K$.

Exemples. 1) L'ensemble \mathbb{C} est algébriquement clos.

2) L'ensemble A défini ci-dessus est lui aussi algébriquement clos. C'est même la clôture algébrique de \mathbb{Q} [voir définition plus loin].

• Corps de rupture. Corps de décomposition.

Remarque. Nous n'allons pas rappeler la démonstration des théorèmes qui vont suivre.

Définition. Soit K un corps, $P \in K[X]$ irréductible. Une extension $L \supset K$ est un **corps de rupture** de P si L est une extension monogène $L = K(\alpha)$ avec $P(\alpha) = 0$.

Théorème. Soit $P \in K[X]$, irréductible. Il existe un corps de rupture pour P , unique à isomorphisme près.

Remarque. Si L est un corps de rupture de P , P n'est pas, en général, entièrement factorisé sur L : si on a $k = \mathbb{Q}$ et $P(X) = X^3 - 2$, alors on a $L = \mathbb{Q}(\sqrt[3]{2})$ et on peut prendre la racine cubique réelle, de sorte que L soit contenu dans \mathbb{R} , mais alors, les autres racines de P ne sont pas dans L (ce sont $j\sqrt[3]{2}$ et $j^2\sqrt[3]{2}$). Ceci nous amène à la définition suivante :

Définition. Soit $P \in K[X]$ (irréductible ou non), de degré n . Un **corps de décomposition** de P sur K est une extension L de K telle que :

- 1) Dans $L[X]$, P est produit de facteurs de degré 1 (ou encore, P a toutes ses racines dans L).
- 2) L est minimal pour cette propriété (ou encore, les racines de P engendrent L).

Théorème. Pour tout P appartenant à $K[X]$, il existe un corps de décomposition de P , unique à isomorphisme près.

Notation. Le corps de décomposition de P sur K est noté $D_K(P)$.

Exemples. Soit $K = \mathbb{Q}$, si on a $P(X) = X^3 - 2$, alors on a $D_K(P) = \mathbb{Q}(\sqrt[3]{2}; j)$
 Soit $K = \mathbb{Q}$, si on a $P(X) = X^4 - 2$, alors on a $D_K(P) = \mathbb{Q}(\sqrt[4]{2}; i) \dots$

Un prolongement naturel de la notion de corps de décomposition est celle de clôture algébrique :

Définition. Une extension \overline{K} de K est appelée une **clôture algébrique** de K si :

- 1) L'ensemble \overline{K} est algébriquement clos.
- 2) L'ensemble \overline{K} est algébrique sur K .

Exemples. On a les égalités suivantes : $\mathbb{C} = \overline{\mathbb{R}}$, $A = \overline{\mathbb{Q}}$.

Rappels sur les polynômes cyclotomiques

• Racines de l'unité. Racines primitives.

Soit K un corps, $n \in \mathbb{N}^*$, on considère le polynôme $P_n(X) = X^n - 1$.

Remarque. La dérivée de P_n est $P'_n(X) = nX^{n-1}$.

Si la caractéristique p de K ne divise pas n , la seule racine de P'_n est 0, qui n'annule pas P_n . Donc P_n n'a que des racines simples.

Si p divise n , on a $n = mp$ et $X^n - 1 = (X^m - 1)^p$ (Frobenius) donc P_n a des racines multiples dans un corps de décomposition.

Dans la suite, on supposera toujours n premier à la caractéristique de K .

L'ensemble des racines n -ièmes de l'unité dans K sera noté $\mu_n(K)$:

$$\mu_n(K) = \{\zeta \in K / \zeta^n = 1\}.$$

C'est un sous-groupe de K^* , de cardinal $\leq n$, donc cyclique.

Notons $K_n = D_K(P_n)$, le corps de décomposition de P_n sur K . Alors, $|\mu_n(K_n)| = n$ et $\mu_n(K_n) \simeq \mathbb{Z}/n\mathbb{Z}$. De plus, comme $\mu_n(K) \subset \mu_n(K_n)$, $\mu_n(K) \simeq \mathbb{Z}/d\mathbb{Z}$ pour un d , diviseur de n .

On étudie à présent le groupe $\mu_n(K_n)$.

Définition. Une *racine n -ième primitive de l'unité* est un élément ζ de K_n tel que $\zeta^n = 1$ et $\zeta^d \neq 1$ pour $d < n$. Autrement dit, ζ est un générateur de $\mu_n(K_n)$.

Il y a donc $\varphi(n)$ racines primitives de l'unité. Leur ensemble sera noté $\mu_n^*(K_n)$.

Définition. Le n -ième polynôme cyclotomique $\phi_{n,k} \in K_n[X]$ est donné par la formule :

$$\phi_{n,k}(X) = \prod_{\zeta \in \mu_n^*(K_n)} (X - \zeta).$$

Remarque. Le polynôme $\phi_{n,k}$ est un polynôme unitaire, de degré $\varphi(n)$

• **Etude de $\phi_{n,K}$:**

Proposition. On a la formule (en omettant l'indice de K) : $X^n - 1 = \prod_{d|n} \varphi_d(X)$.

Preuve :

Cela résulte de l'égalité : $\mu_n(K_n) = \bigcup_{d|n} \mu_d^*(K_n)$ (union disjointe) qui exprime que si $\zeta^n = 1$, l'ordre de ζ est un diviseur de n . \square

Remarque. En comparant les degrés, on retrouve la formule : $n = \sum_{d|n} \varphi(d)$.

Exemple. On a : $\phi_1(X) = X - 1$; $X^2 - 1 = \phi_1(X)\phi_2(X) = (X - 1)(X + 1)$. Cette formule permet le calcul des ϕ_n par récurrence :

$$\phi_n(X) = \frac{X^n - 1}{\prod_{d|n, d \neq n} \phi_d(X)}.$$

Proposition. 1) On a $\phi_{n,\mathbb{Q}}(X) \in \mathbb{Z}[X]$.

2) Soit K un corps quelconque et $\sigma : \mathbb{Z} \rightarrow K$ l'homomorphisme canonique. On a alors : $\phi_{n,k}(X) = \sigma(\phi_{n,\mathbb{Q}}(X))$.

En particulier ϕ_{n,\mathbb{F}_p} s'obtient à partir de $\phi_{n,\mathbb{Q}}$ par réduction modulo p .

Preuve :

1) On raisonne par récurrence sur n . On a $\phi_1(X) = X - 1 \in \mathbb{Z}[X]$.

Supposons la propriété établie pour $d < n$. Soit $F(X) = \prod_{d|n, d \neq n} \phi_d(X)$, on a $F(X) \in \mathbb{Z}[X]$ et F unitaire. On effectue alors la division euclidienne de $X^n - 1$ par $F(X)$ dans $\mathbb{Z}[X]$:

$$X^n - 1 = F(X)P(X) + R(X); P, R \in \mathbb{Z}[X] \text{ et } d^e R < d^n F.$$

Mais comme on a $X^n - 1 = \phi_n(X)F(X)$ dans $\mathbb{Q}[X]$, alors on a $F(X)(\phi_n(X) - P(X)) = R(X)$ et, pour une raison de degré, on a $\phi_n = P \in \mathbb{Z}[X]$.

2) On raisonne encore par récurrence, le cas $n = 1$ étant trivial.

Dans $\mathbb{Z}[X]$, on a : $X^n - 1 = \prod_{d|n, d \neq n} \phi_{d,\mathbb{Q}}(X) = \phi_{n,\mathbb{Q}}(X)F(X)$.

Comme σ est un homomorphisme, on a, dans $K_n[X]$ (avec $K_n = D_K(X^n - 1)$) : $X^n - 1 = \sigma(X^n - 1) = \sigma(\phi_{n,\mathbb{Q}})\sigma(F)$.

Mais $\sigma(F) = \prod_{d|n, d \neq n} \sigma(\phi_{d,\mathbb{Q}}) = \prod_{d|n, d \neq n} \phi_{d,K}$,

et comme $X^n - 1 = \prod_{d|n} \phi_{d,K}$, il en résulte, puisque $K[X]$ est intègre,

$\phi_{n,K} = \sigma(\phi_{n,\mathbb{Q}})$. \square

BIBLIOGRAPHIE

SERRE J.P., Cours d'arithmétique, Hermann.

STEWART I., Galois theory, Chapman et Hall.

PERRIN D., Cours d'algèbre, Ellipse.

GOZARD I., Théorie de Galois, Ellipse.

Table des matières

1	Les carrés de \mathbb{F}_q	3
1.1	Différents cas suivant la caractéristique	3
1.1.1	Cas de la caractéristique égale à 2	3
1.1.2	Cas de la caractéristique supérieure à 2	3
1.2	La loi de réciprocité quadratique	5
2	Démonstration de la loi de réciprocité.	6
2.1	Rappels sur la théorie de Galois.	6
2.2	Galois sur un corps fini.	8
2.3	Galois pour une extension cyclotomique.	9
2.4	Galois et permutations.	10
2.5	Discriminant.	12
2.6	Cyclotomie et corps finis.	14
2.7	Frobenius et décomposition en cycles.	16
2.8	Conclusion.	17



Rapport de stage

dans un **Laboratoire d'Analyses Médicales Biogroup**

à Valenciennes

du 5 au 24 juillet 2010.

Lucile MORELLE
Année 2009-2010

Tuteur : Mr Guillaume OBERT,
Biologiste.

Table des matières

REMERCIEMENTS.....	2
INTRODUCTION.....	2
I- Présentation du laboratoire d'analyses médicales Biogroup de Valenciennes.....	3
1. Historique des laboratoires d'analyses médicales Biogroup.....	3
2. Organisation du laboratoire d'analyses biologiques Biogroup de Valenciennes.....	4
II- Découverte du plateau technique.....	5
1. Le travail des techniciens.....	5
1.1.Le service des envois.....	6
1.2.Le service hématologie et coagulation.....	6
1.3.Le service biochimie et immunologie.....	9
1.4.Le groupe 3.....	10
1.5.Le service bactériologie.....	12
2. Le travail des biologistes.....	15
2.1.Validation.....	15
2.2.Contrôle de qualité.....	15
3. Le secrétariat, les coursiers, le déballage et le courrier.....	15
3.1.Le secrétariat.....	15
3.2.Les coursiers.....	15
3.3.Le déballage.....	15
3.4.Le courrier.....	16
III- L'intervention des mathématiques au sein du laboratoire d'analyses médicales.....	17
1. Les contrôles de qualité.....	17
1.1.Qu'est ce que le contrôle de qualité?.....	18
1.2.CQI et CQE : quels points communs et quelles spécificités?.....	18
1.2.1.Le CQI.....	18
1.2.2.Le CQE.....	19
1.3.L'exploitation statistique.....	21
1.3.1.L'exploitation pour le CQE.....	21
1.3.2.L'exploitation pour le CQI.....	23
1.4.L'analyse des contrôles de qualité.....	25
2. Les modèles mathématiques.....	26
2.1.Les méthodes de traitement de données de l'analyseur c.....	26
2.1.1.Méthodes de traitement de données photométriques.....	26
2.1.2.Méthodes de traitement de données potentiométriques.....	27
2.2.Les méthodes de traitement de données de l'analyseur i.....	27
2.2.1.Méthode en point par point.....	27
2.2.2.Méthode par régression linéaire.....	28
2.2.3.Méthode 4PLC.....	28
2.2.4.Méthode des dosages par seuil.....	28
2.3.Les méthodes d'ajustement de l'analyseur i.....	28
CONCLUSION.....	30
ANNEXE.....	31
I- Exemples de résultats d'analyses sur le CAPILLARYS.....	31
II- Exemple de rapport mensuel de CQE exploité, pour un des réactifs de l'ARCHITECT.....	37
III- Exemple d'analyse d'une représentation de Levey-Jennings, pour un des réactifs de l'ARCHITECT.....	38

REMERCIEMENTS

Je tiens à remercier Madame WILLAERT de m'avoir permis d'effectuer mon stage dans ce laboratoire d'analyses médicales de Valenciennes.

Un grand merci à mon tuteur, le biologiste Guillaume OBERT, celui-ci m'a guidée et suivie tout au long de mon stage.

Merci à toute l'équipe technique avec laquelle j'ai partagé cette expérience professionnelle, qui m'a rapidement intégrée et beaucoup appris.

Je remercie également les secrétaires, les coursiers... pour leur accueil.

INTRODUCTION

J'ai choisi d'effectuer mon stage de trois semaines dans un laboratoire d'analyses biologiques et médicales, car ce milieu scientifique m'intrigue. Je trouve en effet intéressant, de savoir comment les analyses de notre prélèvement se déroulent, entre ce dernier et les résultats que l'on récupère. De plus, je trouve aussi intéressant de savoir quel rôle peuvent jouer les mathématiques dans ce domaine.

D'abord, j'ai pu me familiariser avec le plateau technique et m'intégrer au personnel du laboratoire, observer leurs modes opératoires...

Ensuite, pour mieux comprendre les analyses des prélèvements et savoir où les mathématiques interviennent, j'ai consulté et étudié les manuels techniques des différentes machines.

Finalement, j'ai pu m'occuper des contrôles de qualité, indispensables au bon fonctionnement des analyses au sein du laboratoire. Autrement dit, apprendre en quoi ils consistent et aider à les analyser.

Dans ce rapport, je vais dans un premier temps présenter le laboratoire d'analyses médicales dans lequel j'ai effectué mon stage. Dans un second temps, je vais retracer ma découverte du plateau technique, tout en présentant les machines qui effectuent les analyses, dans les différents services du laboratoire. Et enfin, je vais mettre en évidence l'intervention des mathématiques dans un laboratoire d'analyses médicales, en évoquant les contrôles de qualité et le fonctionnement des machines.

I- Présentation du laboratoire d'analyses médicales Biogroup de Valenciennes :

1. Historique des laboratoires d'analyses médicales Biogroup.

Création en 1949 d'un laboratoire par Monsieur Michel DEMONCHY situé 158 Avenue de Liège à Valenciennes.

Installation le 1er Janvier 1959 d'un laboratoire au 17 rue St Géry à Valenciennes par le biologiste Pierre HOURIEZ

Le docteur Jean LEDUC, successeur du Dr DELAUNAY, est alors chirurgien à l'hôpital de Le Quesnoy. Ainsi se créait un axe

Valenciennes – Le Quesnoy- Berlaimont qui est en grande partie responsable du développement du laboratoire de l'Avesnois.

Concrétisation de l'axe par Madame HOURIEZ qui créa un laboratoire autonome dans les locaux de la clinique de Berlaimont.

Association du laboratoire par Monsieur HOURIEZ à Haumont, puis celle du Parc à Saint-Saulve.

Arrivée de Jacques MOMAL le 1er Janvier 1971 comme assistant, puis associé en Janvier 1972.

Création de la Société Civile qui regroupe trois associés égaux (Mr HOURIEZ, Mme HOURIEZ- DANGRE et Mr MOMAL).

Arrivée de Madame Monique BRAUGE, Monsieur Philippe HIVES et Monsieur Didier BACLE comme assistants au sein de l'équipe.

Prise de participation par Monsieur BACLE dans la société qui deviendra par la suite Société Civile Professionnelle alors intitulée SCP HOURIEZ-DANGRE, MOMAL et BACLE.

Création en 1985 d'un laboratoire situé au 11 rue Anatole France à Anzin par Messieurs Marc DEMAEGHT et Daniel GADEYNE, anciens biologistes à l'hôtel Dieu.

Fusion en Avril 1990 des 2 laboratoires (Anzin et Valenciennes) à cause d'une importante baisse des tarifs de la biologie imposée par le Ministère de la Santé. Les 2 laboratoires fusionnés bénéficient d'une restructuration. Emménagement en Juin 1991

des locaux sis au 8 place verte après le transfert des Ets DRECQ en zone industrielle, ce qui permettra d'accroître l'activité. Il devient alors SCP MOMAL BACLE DEMAEGHT et GADEYNE.

En Mai 1997, augmentation du capital, ce qui permet à la société de devenir par la suite une SELARL, Société d'Exercice Libérale à Responsabilité Limitée.

Ouverture le 1er Octobre 1998 d'un nouveau laboratoire à Aulnoye-Aymeries par le biologiste Mr BACLE.

Ouverture en 1999 d'un nouveau laboratoire au sein de la clinique du Val de Sambre par le biologiste Mr GADEYNE.

Fusion avec le laboratoire BERNARDIN de Saint-Saulve qui travaille avec la clinique du Hainaut. C'est un apport de 12 salariés au sein de SELARL.

Création d'une Délégation Unique du Personnel, d'un CHSCT et mise en place de la participation aux bénéfices.

Arrivée de Mr Frédéric LECOMPTE en Novembre 2002 comme assistant puis comme associé en 2004.

Arrivée de Monsieur Guillaume OBERT en Juin 2004 comme assistant puis comme associé en 2006.

Départ en retraite de Mr Jacques MOMAL en fin d'année 2004.

Arrivée de Mr Nicolas CHATELAIN en Novembre 2005 comme assistant puis comme associé en 2007.

La SELARL BBDGLOC devient BIOGROUP Selarl en Janvier 2008.

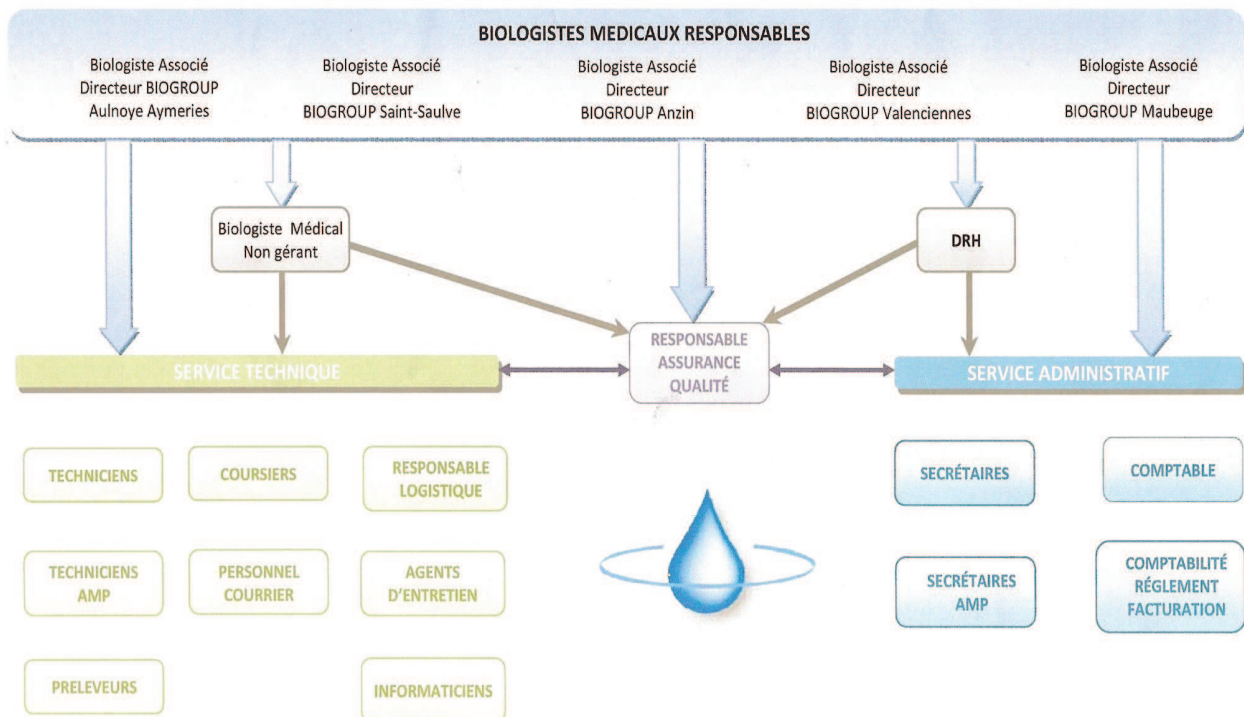
Transfert du laboratoire BERNARDIN de Saint-Saulve en face de la Clinique du Parc dans un laboratoire tout neuf en Février 2008.

Finalement, Biogroup est un **réseau de laboratoires d'analyses médicales**. Il regroupe des laboratoires situés dans le Nord de la France, sélectionnés pour la qualité de leur pratique, la compétence de leurs équipes et la fiabilité de leurs services. C'est pourquoi j'ai choisi d'effectuer mon stage dans le laboratoire d'analyses médicales de Valenciennes qui fait parti de ce réseau.

2. Organisation des laboratoires d'analyses biologiques Biogroup.







ORGANIGRAMME STRUCTUREL SELARL BIOGROUP



II- Découverte du plateau technique :

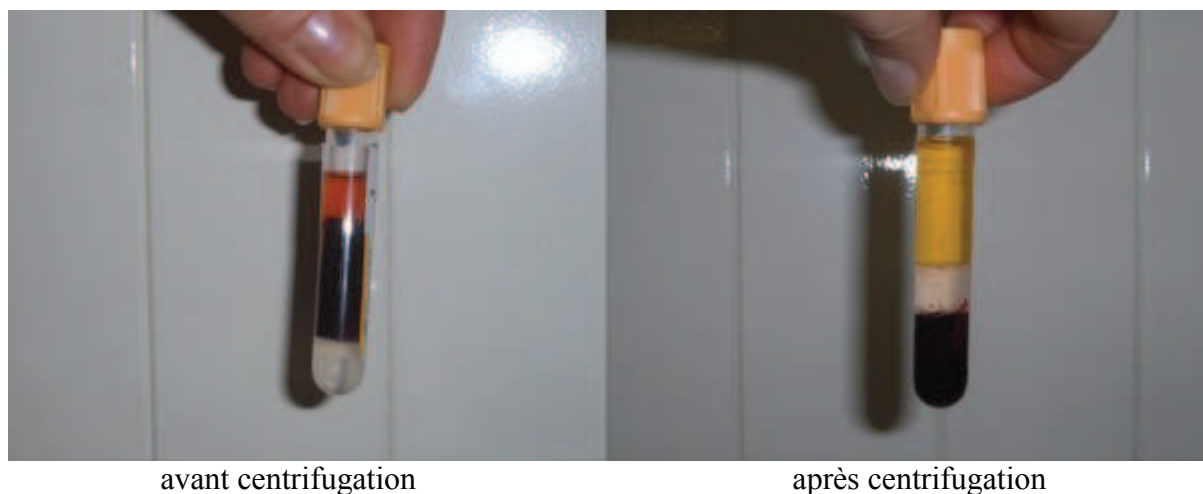
1. Le travail des techniciens.

Pour faciliter le travail des techniciens, les préleveurs bénéficient d'un code couleur concernant les tubes, suivant l'analyse qui doit être faite :

TUBE SEC	TUBE CITRATE	TUBE EDTA	TUBE FLUORURE
			
<p>BIOCHIMIE sauf ammoniémie, PTH, lactates</p> <p>IMMUNOLOGIE</p> <p>HORMONOLOGIE</p> <p>MARQUEURS TUMORAUX</p> <p>ALLERGIE</p> <p>MEDICAMENTS sauf la ciclosporine.</p> <p>TOXIQUES</p>	<p>COAGULATION :</p> <p>TP, INR, TCA, Fibrine, D Dimères, Activité anti Xa, Antithrombine III, ACC, facteurs de la coagulation.</p> <p>TS in vitro (prélever un tube supplémentaire)</p> <p>REPLIR ENTIEREMENT LES TUBES DE 5 ml</p>	<p>HEMATOLOGIE</p> <p>GROUPE SANGUIN RA!</p> <p>BIEN IDENTIFIE LE TUBE AVEC NOM, PRENOM ET DATE DE NAISSANCE</p> <p>AMMONIEMIE</p> <p>CICLOSPORINE</p> <p>HEMOGLOBINE</p> <p>GLYCOSYLEE</p> <p>VS</p>	<p>GLYCEMIE : Glycémie à jeun, Post – prandiale.</p> <p>LACTATES</p>
<p>ORDRE DE REMPLISSAGE DES TUBES →</p>			
<p>Certaines analyses spécialisées nécessitant le prélèvement de tubes particuliers (tubes héparinés, tubes aluminium,...) Disponibles au laboratoire sur simple demande</p>			

Suivant l'analyse à effectuer, le tube doit ou non être centrifugé. La centrifugation sépare le sérum des éléments figurés (globules rouges, globules blancs, plaquettes).

La figure ci-dessous met en évidence la différence entre un prélèvement avant centrifugation et un prélèvement après centrifugation.



- Le tube sec est automatiquement centrifugé. Il est ensuite analysé dans le service « biochimie et immunologie » ou le groupe 3.
- Le tube citrate est analysé dans le service « hématologie et coagulation ». S'il s'agit de sang total, il n'est pas centrifugé et l'analyse à effectuer peut alors être : le temps de saignement ou la numération des plaquettes (hématologie). En revanche, s'il s'agit de plasma, il doit passer par la centrifugation avant d'aller en coagulation.
- Le tube EDTA est analysé en « hématologie ». S'il s'agit de sang total, il n'est pas centrifugé. S'il s'agit de plasma EDTA ou de plasma EDTA et globules rouges, le tube est passé à la centrifugeuse.
- Le tube fluorure (plasma fluoré) est analysé dans le service « biochimie et immunologie ».

Si l'analyse prescrite ne peut pas être effectuée dans les services du laboratoire, le prélèvement est analysé dans un autre laboratoire du réseau Biogroup, si l'un d'entre eux est en mesure d'effectuer l'analyse. C'est un coursier qui se charge alors du transfert. Sinon il est dirigé dans le service des « envois ».

1.1. Le service des « envois »

Le technicien qui se charge des envois est celui qui va se charger également des prélèvements au laboratoire.

Les « envois » correspondent aux prélèvements dont les analyses ne peuvent pas être effectuées sur place ou dans un des laboratoires Biogroup. Ils sont emballés dans un paquet spécial avec tous les renseignements nécessaires, conservés à la température adaptée (congélateur, température ambiante, température du corps...) et en fin de journée, ils sont envoyés dans un laboratoire d'analyses médicales à Paris.

1.2. Le service « hématologie et coagulation »

Les différentes machines du service « hématologie et coagulation » sont les suivantes :

L'AEROSPRAY hématologie : c'est un automate de coloration des lames et de centrifugation.



Le ALIFAX-TEST1 : c'est un analyseur automatique pour la détermination de la vitesse de sédimentation (en mm/h) dont le principe analytique est la photométrie capillaire.



Le COULTER LH780 : c'est un système qui transporte, mélange, aspire et traite automatiquement les échantillons. Il utilise une méthode qui effectue le comptage et l'analyse volumétrique des cellules en détectant et en mesurant les variations de résistance électrique observées lorsqu'une particule (telle qu'une cellule) se trouvant dans un liquide conducteur passe à travers un micro-orifice. Il réalise la numération globulaire complète, la formule sanguine et la numération des réticulocytes (pourcentage et valeur absolue).



L'ADAMS : c'est un analyseur automatique d'hémoglobine glyquée qui mesure l'HbA1c, un indicateur de la glycémie moyenne du patient au cours du dernier mois (ou des deux derniers). Il permet de contrôler la glycémie et apporte des données utiles sur l'efficacité des traitements destinés à contrôler la progression du diabète. Il fournit également des données nécessaires au diagnostic et au contrôle des affections, telles que le diabète, qui résultent des habitudes quotidiennes. Lorsqu'il est configuré à cet effet, l'analyseur permet également de mesurer l'HbA2 pour le dépistage de la bêta-thalassémie.



Le DIAGNOSTICA STAGO : c'est un dispositif médical de diagnostic in vitro composé d'un instrument de laboratoire et d'un logiciel destiné à être utilisé en combinaison avec des consommables et des produits réactifs. Ce système est conçu pour réaliser des tests in vitro destinés au diagnostic et à la surveillance des pathologies liées à l'hémostase. Il permet de réaliser des tests de chronométrie (mesure d'un temps de coagulation), de colorimétrie ou d'immunologie sur des échantillons de plasma.



Le rôle des techniciens est de placer les tubes (code-barres apparent) sur les portoirs correspondants à la machine adaptée à l'analyse demandée. Ensuite, de disposer le ou les portoirs à leur place attribuée dans la machine et enfin de démarrer les analyses.

Une fois les analyses effectuées par la machine, s'il n'y a pas de contestation possible, les résultats sont directement validés et envoyés au programme central du laboratoire par la machine, sinon les résultats sont envoyés au logiciel relié à la machine et le technicien les analyse. Il choisit alors soit de valider, de refaire l'analyse en cas de doute ou de vérifier si l'échantillon est suffisant, s'il n'y a pas de caillots...

En ce qui concerne les analyses effectuées sur le COULTER, si le technicien a un doute, il valide le résultat et fait une lame, qu'il colore à l'aide de l'AEROSPRAY, et que le biologiste regardera au microscope.

Lorsque le technicien a fini de valider, il range les tubes.

1.3. Le service « biochimie et immunologie »

Les machines du service « biochimie et immunologie » sont les suivantes :

L'AxSYM : c'est un système qui utilise trois technologies différentes pour effectuer les dosages immunologiques :

- le dosage immunoenzymologique microparticulaire (MEIA),
- le dosage immunologique par polarisation de fluorescence (FPIA),
- l'atténuation de l'intensité de fluorescence (REA).

La concentration de substance à doser présente dans les échantillons des patients est déterminée par comparaison des résultats d'analyse avec la courbe de calibration, sachant que chaque dosage doit être calibré avant de pouvoir procéder à l'analyse des échantillons.



L'ARCHITECT ci8200 : c'est un automate qui utilise les méthodes photométriques, potentiométriques ou CMIA (dosage immunologique microparticulaire par chimiluminescence), pour mesurer les concentration en substance à doser dans les échantillons.

Il est composé d'un *analyseur c* de biochimie et d'un *analyseur i* d'immunologie (qui détermine la présence d'antigènes, d'anticorps ou d'analytes dans les échantillons analysés).



Le service « biochimie et immunologie » étant entièrement automatisé, le travail du technicien est de déboucher les tubes, les mettre sur les portoirs et mettre les portoirs dans la machine. La machine effectue les analyses demandées (fournies grâce au code-barres) et renvoie les résultats au logiciel associé, puis au programme central. Le technicien vérifie les résultats, les valide puis contacte, en cas d'urgence, le médecin ou le patient lui-même, afin de lui fournir les résultats.

Ensuite, le technicien récupère les tubes, les range et les garde en chambre froide, le temps réglementaire.

1.4. Le groupe 3

Les machines du groupe 3 sont les suivantes :

Le CAPILLARYS : c'est un instrument multiparamétrique automatique qui assure l'analyse des protéines sériques sur 8 capillaires en parallèle.



Cet appareil d'électrophorèse capillaire multiparamétrique, offre un large choix d'analyses : protéines, immunotyping (technique correspondant à l'immunofixation), HR (haute résolution), CDT (carbohydre déficient transferrin), hémoglobine. Complètement robotisé et pourvu d'un passeur d'échantillons à alimentation continue, il ne nécessite aucune manipulation entre le chargement des échantillons et l'obtention des résultats. (Les tubes sont identifiés au moyen d'un système de lecture code-barres.)

Remarque : pour des exemples de résultats d'analyses sur le CAPILLARYS, voir l'annexe.

Le système HYDRASYS : c'est un instrument multiparamétrique semi-automatique qui assure le traitement des HYDRAGEL (gels d'agarose) et permet la réalisation de tout type d'électrophorèse. Ce système fait appel à l'isoélectrofocalisation pour la migration électrophorétique.

Il a de nombreuses applications : protéines sériques et urinaires, immunofixations (serum-IF, Bence Jones, CSF), hémoglobines, lipoprotéines/ISO-enzymes...

Cet appareil demande un peu de manipulation de la part du technicien : c'est lui qui doit déposer 2ml de l'échantillon à analyser dans les petites cuves prévues à cet effet, de même pour les différents réactifs...



Les tests effectués dans le groupe 3 sont les suivants :

Waler Rose et Monogen.

Waler Rose : test semi-quantitatif permettant le titrage des facteurs rhumatoïdes de classe IgM dans le sérum humain par une réaction d'agglutination en microplaque.

Monogen : test rapide au latex pour la détection de facteurs rhumatoïdes de classe IgM dans le sérum humain.

Les différentes techniques peuvent être divisées en deux groupes selon l'origine des IgG.

Les IgG animales => réaction de Waler Rose.

Les IgG humaines => réaction dite au latex.

Les réactions de Waler Rose sont plus spécifiques alors que les méthodes au latex sont plus sensibles.

syphagen TPHA et RPR reditest.

syphagen TPHA : test d'hémagglutination indirecte pour la détection des anticorps spécifiques *anti-Treponema pallidum* dans le sérum ou le plasma humain.

RPR reditest : test rapide pour la détection qualitative et quantitative de syphilis dans du sérum ou du plasma.

SureStep.

SureStep : test de dépistage rapide pour la détection qualitative drogues et de leurs métabolites dans l'urine humaine.

SureStep *Mult-Drogues* : test de dépistage rapide pour la détection qualitative et simultanée d'un panel de toxiques et de leurs principaux métabolites dans l'urine humaine.

En ce qui concerne ces tests, le technicien doit suivre les modes opératoires indiqués dans les différentes notices.

1.5. Le service « bactériologie »

Les machines en bactériologie sont les suivantes :

L'AEROSPRAY microbiologie : c'est un automate de coloration de GRAM.

L'AEROSPRAY hématologie : c'est un automate de coloration et centrifugation.

Le VITEK2 : c'est un automate d'identification microbienne et de détection des résistances bactériennes, qui interprète les résultats des antibiogrammes. Il est capable d'identifier plus de 330 espèces microbiennes grâce à une base de données optimisée et de nombreuses cartes d'identification.



Le BacT/ALERT3D : c'est un système de détection microbienne automatisé, pour les hémocultures, les cultures de liquides stériles, et/ou de mycobactéries. Il est composé d'un module de contrôle et d'un module d'incubation. (Son système de détection colorimétrique basé sur des algorithmes sophistiqués permet la détection précoce des microorganismes.)



L'IRIS IQ200 : c'est un automate d'analyse par cytométrie en flux de sédiments urinaires. L'analyse s'effectue directement sur l'échantillon, sans centrifugation.



La procédure suivie en bactériologie est la suivante :

La recherche d'infection urinaire ou Examen CytoBactériologique Urinaire (ECBU) s'effectue après recueil du second jet urinaire sur petit pot stérile et transmission rapide au laboratoire (à défaut conservation à 4°C).

Si l'analyse ne peut pas être effectuée sur place ou dans un des laboratoires Biogroup, le technicien fait un envoi dans un laboratoire extérieur à Paris.

Sinon, pour révéler la présence de germes dans un prélèvement, celui-ci doit être ensemencé sous hotte à flux laminaire (protection du personnel) sur des milieux de culture adaptés à une croissance optimale des germes fragiles, il s'agit de géloses.

L'étude de la cytologie urinaire à la recherche de signes d'infection ou d'hématurie (passage de sang dans les urines) est automatisée sur IQElite (IRIS) avec archivage photographique des éléments décomptés. Une fois cette étude effectuée, suivant si le résultat est positif ou négatif, le prélèvement est ensemencé sur la gélose correspondante (respectivement CPS3 ou BCP).

Une température de 37°C est nécessaire à la multiplication des germes microscopiques jusqu'à l'obtention de "colonies" bactériennes visibles à l'œil nu. La lecture des boîtes s'effectue 24h et 48h après l'ensemencement.



Après ensemencement et croissance 24h à 48h à 37°C, des colonies bactériennes sont visibles. L'observation des germes se fait au microscope par un technicien après coloration spéciale, coloration automatisée. (utilisation de l'AEROSPRAY)

Tout l'art du technicien spécialisé en bactériologie est de différencier la flore commensale non pathogène de la flore responsable de l'infection pour ensuite identifier le germe si un rôle pathogène est suspecté.



Pour les germes les plus courants, le laboratoire utilise des techniques automatisées (VITEK2). Pour identifier les germes délicats et étudier leur sensibilité aux antibiotiques, étape nécessaire à une bonne prise en charge thérapeutique, des techniques manuelles et standardisées sont utilisées.



Finalement, le service bactériologie et le groupe 3 sont les seuls services où il reste encore un peu de technique manuelle. Maintenant, pour des raisons d'efficacité et de rapidité, le monde de l'analyse biologique et médicale est de plus en plus automatisé.

2. Le travail des biologistes

2.1. Validation

Après que le technicien ait validé ses résultats d'analyses, c'est au biologiste de les regarder à son tour et de les valider. Ce n'est qu'une fois validé par le biologiste que le résultat d'analyse est pris en compte dans le programme central du laboratoire et peut alors être édité.

2.2. Contrôle de qualité

Tous les jours, les techniciens effectuent des contrôles sur leurs machines, il s'agit du contrôle de qualité interne. Et une fois par mois environ, les techniciens font un contrôle ponctuel de leurs machines, les résultats de ce contrôle sont envoyés à un organisme d'étude et sont comparés aux résultats des autres laboratoires, il s'agit du contrôle de qualité externe. Le biologiste doit analyser les résultats de ces contrôles de qualité, pour permettre de garantir la qualité et la validité des résultats que fournissent les machines du laboratoire.

3. Le secrétariat, les coursiers, le déballage et le courrier

3.1. Le secrétariat

La secrétaire médicale se charge de la prise en charge administrative. Elle enregistre les informations concernant le patient (nom (et nom de jeune fille), prénom, adresse, date de naissance, médecin traitant, numéro de sécurité sociale) dans le programme central lors de sa première visite au laboratoire. Ensuite, suivant ce qui est prescrit sur l'ordonnance, elle enregistre les prélèvements et donc les analyses à effectuer. Enfin, elle se charge d'imprimer des étiquettes à code-barres, pour coller sur les différents prélèvements, et une spéciale pour l'ordonnance, afin d'avoir une traçabilité. Les prélèvements sont ensuite dirigés vers le plateau technique, dans le service qui convient.

3.2. Les coursiers

Les coursiers se chargent d'aller récupérer les prélèvements dans les différentes pharmacies (lorsque que le prélèvement a été effectué à domicile), cliniques (lorsque le patient est hospitalisé), etc. puis ils les ramènent au laboratoire, où ils sont vérifiés avec l'ordonnance et rangés sur des portoirs mère, il s'agit du déballage primaire, effectué par les techniciens.

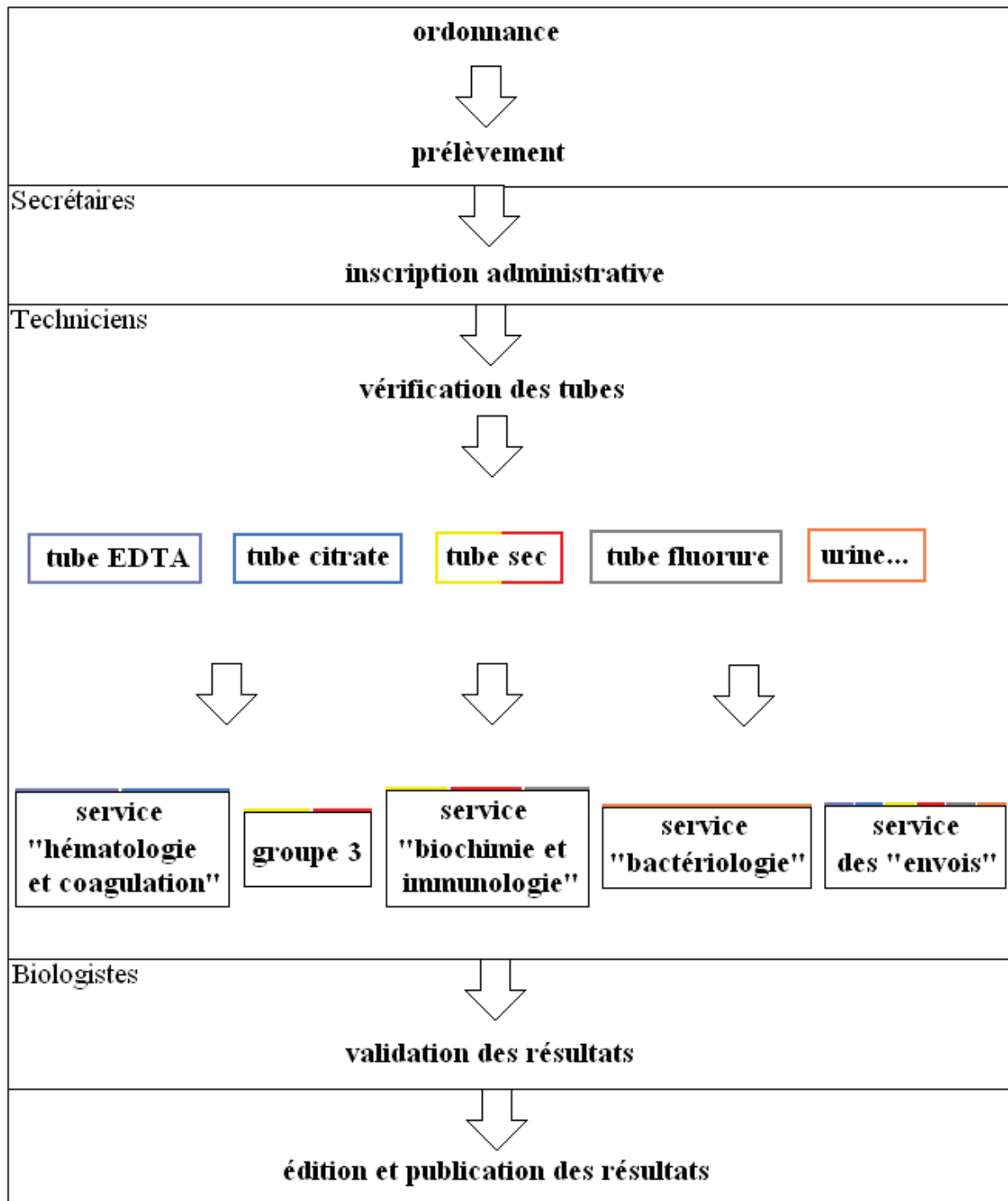
3.3. Le déballage

Après le déballage primaire, les secrétaires enregistrent les patients et les prélèvements effectués, puis sortent les étiquettes avec le numéro du portoir mère (avant analyse), celui du portoir fille (après analyse), l'analyse prescrite et le code-barres. Les techniciens étiquettent les tubes et les trient selon l'analyse à effectuer, il s'agit du déballage secondaire.

3.4. Le courrier

C'est une fois que les résultats sont validés par le biologiste que la personne chargée du courrier peut éditer les résultats d'analyse des patients et est donc en mesure de les envoyer au patient lui-même ou au médecin... Les feuilles de résultats sont ensuite pliées et enveloppées par une machine. Finalement, les enveloppes passent les unes après les autres dans la machine à affranchir, puis un coursier se charge d'aller à la poste.

Finalement, on peut résumer le parcours, de l'ordonnance jusqu'aux résultats d'analyse, sur le schéma suivant :



III- L'intervention des mathématiques au sein du laboratoire d'analyses médicales

Au sein d'un laboratoire d'analyses médicales, les mathématiques ont leur rôle à jouer. Les mathématiques sont en effet présentes dans la conception des machines, dans leur fonctionnement, puis également dans les contrôles de qualité. C'est pourquoi je me suis intéressée aux différents contrôles de qualité ainsi qu'aux modèles mathématiques utilisés par les machines.

1. Les contrôles de qualité

Le contrôle de qualité tel que nous le connaissons aujourd'hui, résulte de l'évolution de plusieurs facteurs sur les huit dernières décennies :

- les pratiques des laboratoires ont progressivement évolué,
- le savoir-faire industriel est devenu de plus en plus performant,
- le tout, dans un contexte réglementaire toujours plus exigeant.

La naissance du contrôle de qualité dans le milieu industriel...

Le contrôle de qualité naît dans les années 30, de la conjonction de travaux de statisticiens et du besoin des industriels de fournir au client des produits constants. En 1931, W. Shewhart élabore le concept de contrôle statistique, applicable à la production de série, et plus généralement, à tout processus répétitif.

L'arrivée du contrôle de qualité en biologie...

Intégrés plus tardivement au sein des laboratoires d'analyses de biologie médicale, les premiers contrôles s'appuyant sur des résultats chiffrés apparaissent en biochimie vers 1950 : S. Levey et E.R. Jennings proposent alors des graphiques de contrôle pour le laboratoire d'analyses médicales. L'intérêt et l'expertise dans ce domaine s'intensifient dès les années 70. En ce qui concerne l'hémostase, le contrôle de qualité se développe réellement dans les années 80.

Tout au long de ces décennies, la matière première utilisée dans le cadre du contrôle de qualité a également évolué : en hémostase par exemple, les laboratoires sont passés progressivement de pool de plasmas de patients dits « normaux », à l'utilisation de plasmas de contrôle industriels.

Le contrôle de qualité et la réglementation...

De nos jours, la mise en place du contrôle de qualité dans les laboratoires se fait dans un cadre réglementaire, qui lui aussi, a largement évolué ces dernières années. En France, le point de départ est la loi du 11 juillet 1975, texte fondateur de la biologie médicale. Dans ce dernier, l'article 761-3 introduit la notion de contrôle de qualité.

Depuis, de nouveaux textes réglementaires sont apparus, notamment le « Guide de Bonne Exécution des Analyses de Biologie Médicale » (GBEA) dont l'application est obligatoire par tous les laboratoires d'analyses de biologie médicale publics et privés. Ce guide, initialement paru au Journal Officiel du 4 décembre 1994, en est déjà à sa deuxième version (arrêté du 26 novembre 1999). Sa vocation est d'instituer la mise en place d'un système d'assurance qualité fondé sur des procédures et des modes opératoires écrits, concernant l'ensemble des étapes, allant du prélèvement de l'échantillon biologique au rendu des résultats.

Dans le cadre d'une démarche volontaire d'accréditation, des référentiels normatifs encadrent également la mise en place du contrôle de qualité :

- la norme NF EN ISO/CEI 17025 définit les prescriptions générales concernant la compétence des laboratoires d'étalonnage, d'essais et d'analyses ;
- la norme ISO/EN 15189 définit les exigences particulières concernant la qualité et la compétence, pour les laboratoires d'analyses de biologie médicale.

Le contrôle de qualité et les institutions...

Enfin, on ne saurait parler de contrôle de qualité sans évoquer les associations régionales qui ont largement contribué à l'essor du contrôle de qualité. De même, le Contrôle de Qualité National (CQN) a joué un rôle moteur dans ce domaine. Ce dernier prévu dans la loi du 11 juillet 1975 a été mis en place par Alain Leblanc (1931-2004), avec l'aide des Professeurs M. Bailly (biochimie), A. Goguel (hématologie dont hémostase), M. Véron et J-C. Petithory (microbiologie). Alain Leblanc met alors sur pied un contrôle de qualité fondé sur le volontariat, regroupant au départ quelque 1500 laboratoires. En 1979, les premiers CQN virent le jour.

En 1993, le Laboratoire National de Santé (LNS) est remplacé par l'Agence du Médicament, puis en 1999, par l'Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSAPS). Ces évolutions ont profondément changé l'organisation du CQN : auparavant géré par six associations (biochimie, virologie, hématologie, parasitologie, hormonologie, gazométrie sanguine), le CQN a d'abord été centralisé au sein de l'Agence du Médicament, puis au sein de l'AFSSAPS.

Le contrôle de qualité qui évolue depuis le siècle dernier, fait désormais partie du quotidien des laboratoires d'analyses de biologie médicale. Cependant, des facteurs tels que l'intensification de la démarche qualité, la modernisation des méthodes statistiques et l'intégration du management de la qualité au sein des analyseurs dessinent progressivement le contrôle de qualité de demain...

1.1. Qu'est ce que le contrôle de qualité?

Au sein du laboratoire d'analyses de biologie médicale, le contrôle de qualité est l'ensemble des dispositions qui permettent au biologiste d'évaluer et de contrôler en continu le processus analytique produisant les résultats rendus aux patients.

Dans la pratique, le contrôle de qualité se définit par la réalisation du « Contrôle de Qualité Interne » (CQI) et par la participation à un programme d'évaluation externe de la qualité (EEQ), appelé plus communément « Contrôle de Qualité Externe » (CQE).

1.2. CQI et CQE : quels points communs et quelles spécificités?

1.2.1. Le CQI

Quel est le rôle du CQI ?

Le rôle du CQI est de permettre au biologiste de valider en permanence la fiabilité du couple « réactif-instrument ». jour après jour, la gestion et l'analyse des résultats obtenus via le CQI, permettent non seulement d'évaluer les performances des méthodes d'analyse, mais surtout de valider la fiabilité des résultats rendus aux patients.

Pourquoi utiliser un CQI ?

En matière de réglementation, le Guide de Bonne Exécution des Analyses de Biologie Médicale (GBEA II) et les normes en vigueur, imposent l'utilisation du CQI ; ceci explique une mise en œuvre bien standardisée.

Dans la pratique, les laboratoires d'analyses de biologie médicale utilisent quotidiennement et/ou à intervalles réguliers des plasmas de contrôle de qualité, appelés plus communément « Contrôle de Qualité Interne » (CQI) :

- à l'issue d'une calibration afin de la valider
- en début, au milieu et/ou fin d'une série de patients, afin de valider les résultats en temps réel.

Dans les deux cas, les CQI sont mesurés dans les mêmes conditions que les plasmas des patients.

L'exploitation statistique de ces résultats permet d'évaluer la justesse et la fidélité des méthodes d'analyses :

- justesse : les résultats du CQI sont confrontés aux valeurs attendues (fourchettes et valeurs cibles préétablies par le fournisseur ou estimées par le biologiste sur une période probatoire). L'écart observé peut être quantifié par le biais.
- Fidélité : la dispersion des résultats obtenus sur le CQI est quantifiée par le coefficient de variation (CV). Son interprétation se fait en fonction des spécifications préalablement établies et de la pertinence clinique du paramètre.

L'utilisation régulière du CQI permet de :

- valider les résultats patients en temps réel
- détecter les erreurs et de mettre en place des actions correctives immédiates
- prévenir les erreurs, en observant d'éventuelles dérives.

1.2.1. Le CQE

Qu'est-ce que le CQE ?

Le CQE est un outil permettant une évaluation externe de la qualité (EEQ). Les performances individuelles d'un laboratoire sont évaluées via la comparaison inter laboratoire des résultats obtenus sur un même spécimen.

Il existe deux types de CQE :

- les programmes permanents : ils exploitent et confrontent les résultats issus du CQI des différents laboratoires participants. Par ailleurs, la confrontation de ces valeurs se fait toujours sur les mêmes zones de mesures que celles du CQI (il s'agit du même et seul produit). Cela implique que le participant connaît au préalable les valeurs attendues.
- Les programmes ponctuels : ils exploitent et confrontent les résultats obtenus par les laboratoires participants, sur un même échantillon, en aveugle. Les avantages de ce type de programme sont nombreux : une totale objectivité dans le rendu des résultats, la multiplicité des niveaux de concentration, complémentaires des niveaux couverts par le CQI qui permettent de couvrir toutes les zones de mesure et une mise en œuvre simple pour les laboratoires participants.

Pourquoi utiliser un CQE ?

Participer à un CQE, c'est s'inscrire dans une démarche qualité : déjà fortement conseillée dans le cadre GBEA II, l'utilisation d'un CQE devient obligatoire dans le cadre de l'accréditation COFRAC (référentiels ISO 17025 et ISO 15189).

Remarque : utiliser un CQE permet de répondre pleinement aux exigences réglementaires et normatives.

Qu'apporte un CQE ?

Avant tout, le CQE ne se substitue pas au CQI dont il est complémentaire. Ainsi, les règles de gestion du CQI restent inchangées.

Associé au CQI, le CQE est un outil de mesure indispensable qui permet de confirmer les performances des méthodes utilisées :

- grâce au CQE, la justesse et la fidélité des techniques utilisées sont évaluées :
justesse : les résultats sont comparés à ceux des autres participants. L'écart observé entre ces résultats peut être quantifié par le biais
fidélité : la reproductibilité inter laboratoire des méthodes utilisées est évaluée sur l'ensemble des participants (CV inter laboratoire)
- l'intérêt d'un CQE est d'offrir une analyse pertinente : la comparaison au groupe des pairs. Le laboratoire peut ainsi être situé au sein d'un groupe de laboratoires travaillant dans des conditions similaires (couple « instrument-réactif » identique).

Remarque : l'utilisation des deux types de contrôles, CQI et CQE, permet d'assurer la qualité de la prestation d'analyse en continu.

Le suivi de l'ensemble de ces indicateurs fait partie intégrante du système de maîtrise mis en place dans le laboratoire : il constitue une aide précieuse dans les décisions techniques et médicales à prendre et contribue ainsi à la fiabilité des résultats des patients.

Remarque : en utilisant un CQE régulièrement, le laboratoire gagne en confiance et en sérénité.

Le tableau ci-dessous résume ce qu'est le contrôle de qualité :

	CQI	CQE
JUSTESSE	Comparaison des résultats du laboratoire (obtenus en individuel) aux valeurs attendues (valeur cible ou fourchette). Indicateur possible : biais	Comparaison des résultats du CQE à la moyenne du groupe des pairs. Indicateur possible : biais
FIDELITE	Analyse de la dispersion des résultats du laboratoire obtenus jour après jour. Indicateur : CV inter laboratoire	Analyse de la dispersion des résultats obtenus sur un ensemble de laboratoires participants. Indicateur : CV inter laboratoire

1.3. Une exploitation statistique.

Les CQ sont effectués sur trois niveaux de contrôle (de concentration) : bas, moyen et haut.

1.3.1. L'exploitation pour le CQE.

Il existe différents niveaux d'exploitation en ce qui concerne le CQE. Pour chaque paramètre, les résultats sont analysés suivant différents degrés de comparaison, allant du plus global au plus fin :

- 1^{er} degré, global : résultat comparé à l'ensemble des résultats des laboratoires participants.
- 2^{ème} degré, méthodologie : résultat comparé à ceux des laboratoires participants utilisant la même méthodologie.
- 3^{ème} degré, technique : résultat comparé à ceux des laboratoires participants utilisant le même réactif.
- 4^{ème} degré, groupe des pairs : résultat comparé à ceux des laboratoires participants utilisant le même couple « réactif-instrument ».

Les résultats sont exploités pour chaque « degré de comparaison » uniquement si le nombre de résultats est statistiquement relevant, soit 12 minimum.

Remarque : grâce à ce gradient d'exploitation, le laboratoire dispose d'une analyse exhaustive, fine et personnalisée de ses performances.

Remarque : l'exploitation statistique des résultats s'effectue sous forme de rapports et chaque rapport est spécifique d'un instrument.

Lors de l'exploitation statistique, les indicateurs utilisés sont les suivants :

- Le nombre total de laboratoires en comparaison (n) : le nombre minimum retenu est 12.
- La moyenne (M.labo) : la moyenne d'une série de résultats d'un laboratoire. Cet indicateur permet d'évaluer la précision des mesures.
- La cible théorique (M.groupe) : la moyenne du groupe de comparaison, des résultats des n laboratoires du groupe.
- L'écart type (Et) : il quantifie la dispersion des résultats d'un laboratoire.
- Le coefficient de variation (Cv) : le rapport entre Et et M, exprimé en pourcentage. Il quantifie, comme l'écart type, la dispersion des résultats.
 $Cv \text{ (en \%)} = Et / M * 100.$

Il existe deux coefficients de variation. Le Cv de répétabilité et le Cv de reproductibilité.

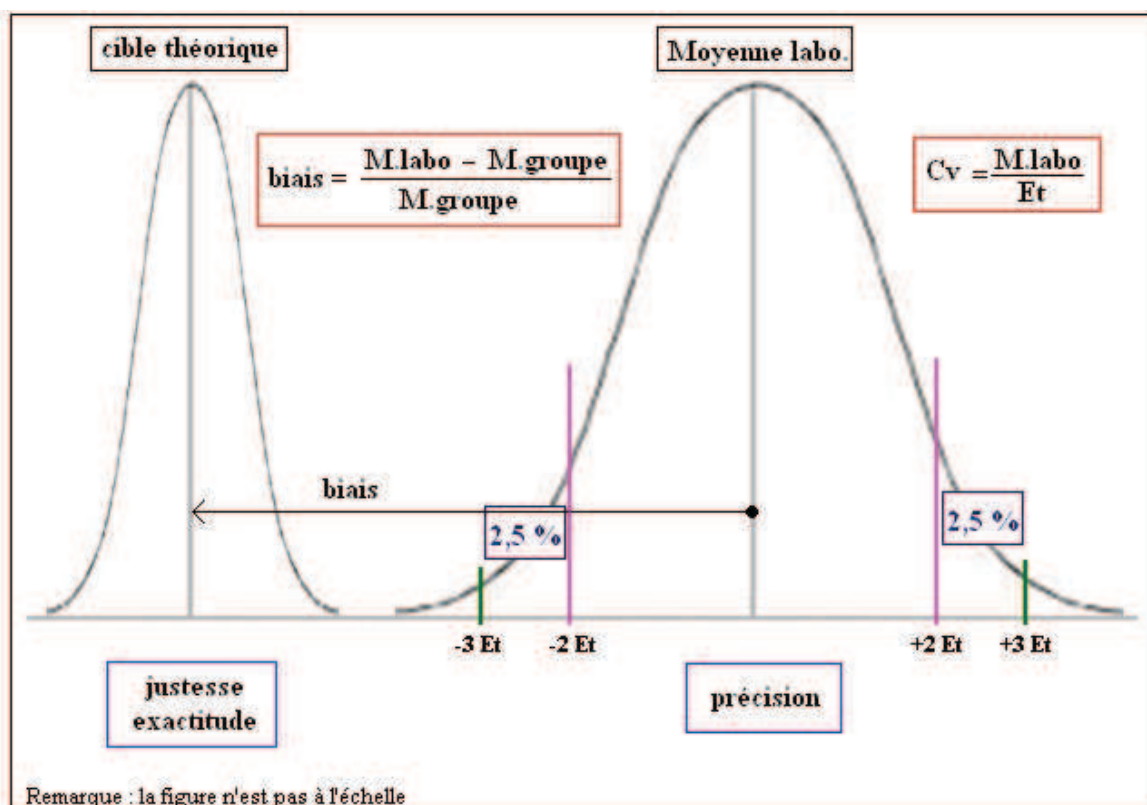
La condition de répétabilité est une condition de **mesurage** dans un ensemble de conditions qui comprennent la même **procédure opératoire**, les mêmes opérateurs, le même **système de mesure**, les mêmes conditions de fonctionnement et le même lieu, ainsi que des mesurages répétés sur le même objet ou des objets similaires pendant une courte période de temps.

La condition de reproductibilité est une condition de **mesurage** dans un ensemble de conditions qui comprennent des lieux, des opérateurs et des **systèmes de mesure** différents, ainsi que des mesurages répétés sur le même objet ou des objets similaires. Les différents systèmes de mesure peuvent utiliser des **procédures opératoires** différentes.

On a l'inégalité suivante : $Cv.repeta < Cv.reprod < Et$.

- Le biais : il quantifie l'écart entre le résultat du laboratoire et la valeur cible théorique (moyenne du groupe de comparaison). Cet indicateur permet d'évaluer la justesse et l'exactitude des mesures.
Biais (en %) = $(M.labo - M.groupe) / M.groupe * 100$.
- Le score : il exprime le nombre d'écart type qui sépare le résultat du laboratoire de celui de la cible théorique.
Score (en %) = $(M.labo - M.groupe) / Et * 100$.
Comme le biais, il permet d'évaluer la justesse de la mesure : une valeur proche de zéro traduit une absence de biais par rapport à la valeur de comparaison. A l'inverse, plus la valeur absolue de cet indicateur est élevée, plus elle traduit un problème de justesse.
 - $|score| < 2$: acceptable.
 - $2 < |score| < 3$: signe de surveillance.
 - $|score| > 3$: signe d'action.

On peut récapituler ces indicateurs sur la figure suivante :



La démarche suivie pour l'exploitation statistique du CQE est la suivante :

- le technicien effectue les mesures de contrôle, en suivant des conditions de répétabilité ou de reproductibilité ;
- le technicien remplit le rapport (mensuel ou semestriel) spécifique à la machine, qu'il envoie ensuite au service d'assurance de la qualité externe (EQAS ici) ;
- EQAS va gérer l'exploitation statistique du contrôle de qualité externe du laboratoire dans un groupe de comparaison, puis renvoie les résultats du contrôle au laboratoire.

Remarque : pour un exemple de rapport mensuel de CQE exploité, pour un réactif de l'ARCHITECT, voir l'annexe.

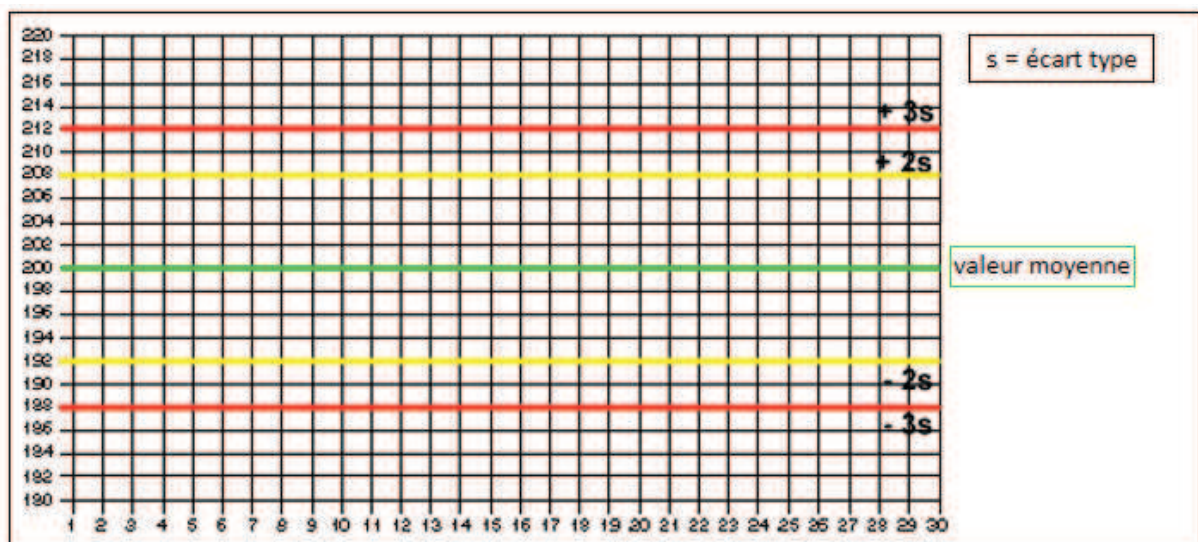
1.3.2. L'exploitation pour le CQI.

Le CQI fait appel au tableau de Levey-Jennings et l'évaluation des résultats s'effectue grâce aux principales règles de Westgard.

- La représentation de Levey-Jennings :

Si la représentation graphique sous forme de distribution permet de visualiser les différentes mesures, sous forme condensée et représentative, elle ne donne aucun renseignement historique. Pour évaluer une série par rapport au reste des mesures, il est plus facile d'avoir une représentation qui tienne compte de l'historique des mesures. On représente alors chaque point en fonction du jour auquel la mesure a été faite. C'est la représentation de Levey-Jennings. Pour faciliter l'interprétation d'un tel graphe, il est courant de représenter les points en fonction de l'écart à la moyenne par rapport à la déviation standard.

Le tableau de Levey-Jennings se présente comme suit (les valeurs dépendent du réactif contrôlé) :



- Les règles de Westgard :

L'utilisation de règles permet pour chaque nouveau point de décider si ce point peut être considéré comme acceptable ou non. Ces règles donnent des moyens objectifs de valider techniquement une série.

Selon une habitude répandue, on abrège ces règles sous la forme A_L , où A représente le nombre de mesures prises en compte et L représente la limite utilisée.

Par exemple : 1_{2s} indique une mesure excédant $2s$ (2 déviations standard ou 2 écarts type).

Il est possible de combiner plusieurs règles pour améliorer l'interprétation. On les sépare alors par une barre de fraction.

Par exemple : $1_{3s}/2_{2s}$ indique que l'on applique les deux règles 1_{3s} et 2_{2s} .

Les principales règles de Westgard sont représentées et décrites dans le tableau suivant :

<i>règle</i>	<i>représentation</i>	<i>description</i>
1_{2s}	<p>Le graphique montre une série de 10 mesures. L'axe vertical est gradué de -3s à +3s, avec une ligne horizontale au niveau de la moyenne (Mean). Les limites de $\pm 2s$ sont indiquées par des lignes jaunes en pointillés. Une seule mesure à l'indice 2 est représentée par un point jaune au-dessus de la limite de +2s, avec une bulle d'appel qui indique '1_{2s} rule violation'.</p>	<p>Seuil d'avertissement, la règle 1_{2s} est violée lorsqu'une seule valeur de contrôle est en dehors des limites de $\pm 2s$. Cette règle est en général considérée comme avertissement et non pas comme critère de rejet d'une série.</p>
1_{3s}	<p>Le graphique montre une série de 10 mesures. L'axe vertical est gradué de -3s à +3s, avec une ligne horizontale au niveau de la moyenne (Mean). Les limites de $\pm 3s$ sont indiquées par des lignes rouges pleines. Une mesure à l'indice 1 est représentée par un point rouge au-dessous de la limite de -3s, avec une bulle d'appel qui indique '1_{3s} rule violation'.</p>	<p>Cette règle est la plus commune, elle stipule que la série de mesure doit être rejetée lorsqu'une mesure exède la moyenne de plus ou de moins de 3 écart-types.</p>
2_{2s}	<p>Le graphique montre une série de 10 mesures. L'axe vertical est gradué de -3s à +3s, avec une ligne horizontale au niveau de la moyenne (Mean). Les limites de $\pm 2s$ sont indiquées par des lignes jaunes en pointillés. Deux mesures consécutives à l'indice 2 et 3 sont représentées par des points jaunes au-dessus de la limite de +2s, avec une bulle d'appel qui indique '2_{2s} rule violation'.</p>	<p>Seuil d'avertissement, la règle 2_{2s} est violée lorsque deux résultats consécutifs et du même côté de la cible, sont supérieurs à $\pm 2s$. Cette règle peut être utilisée avec un contrôle ou avec deux. Si on utilise deux contrôles, la comparaison est faite sur une même série. Avec un contrôle, la comparaison est faite dans la même série.</p>
R_{4s}	<p>Le graphique montre une série de 10 mesures. L'axe vertical est gradué de -3s à +3s, avec une ligne horizontale au niveau de la moyenne (Mean). Les limites de $\pm 3s$ sont indiquées par des lignes rouges pleines. Deux mesures consécutives à l'indice 2 et 3 sont représentées par des points jaunes. L'écart entre ces deux points est supérieur à 4s, ce qui est illustré par une ellipse rouge englobant les deux points et une bulle d'appel qui indique 'R_{4s} rule violation'.</p>	<p>Seuil d'alarme, la règle R_{4s} est violée si l'écart entre les deux contrôles est supérieure à $4s$. Cette règle détecte les erreurs aléatoires trop importantes et implique le rejet de la série. Uniquement avec deux contrôles.</p>

<i>règle</i>	<i>représentation</i>	<i>description</i>
4_{1s}		<p>Avec un contrôle : seuil d'alarme, la règle 4_{1s} violation si quatre valeurs (de quatre séries) sont en dessus ou en dessous de 1s. Avec deux contrôles, la comparaison se fait sur deux séries. Cette règle détecte les erreurs systématiques même de faible importance.</p>
10_{moy}		<p>Seuil d'alarme, la règle 10_{moy} est violée lorsque 10 valeurs qui se suivent se trouvent sur le même côté de la moyenne. Cette règle détecte les erreurs systématiques même de très faible importance. Mais, elle nécessite une validation précise de la valeur cible.</p>

Une fois que le biologiste a rentré les seuils d'avertissement (plus ou moins 2 et 3s), qui peuvent varier d'un lot de réactifs à un autre, pour la représentation de Levey-Jennings des différents réactifs, sur un logiciel adapté (ici, Unity Real Time), il sélectionne les règles de Westgard qu'il souhaite actives. Ensuite, c'est le logiciel, relié à l'automate qui travaille : il enregistre les valeurs des contrôles (quotidiens, mensuels...) et effectue les représentations de Levey-Jennings pour tous les réactifs.

1.4. L'analyse des contrôles de qualité.

– Une analyse quotidienne.

Tout les jours, c'est le technicien qui vérifie si les contrôles quotidiens sont corrects (les contrôles qui ont violés une ou plusieurs des règles de Westgard en vigueur ont déjà été rejetés par le logiciel). Si le technicien découvre une anomalie, il en parle au biologiste, qui cherche alors à trouver une solution au problème.

– Une analyse mensuelle.

Avec les résultats du CQE fournis par l'EQAS, le biologiste relève les biais (en %) pour les différents réactifs. Il analyse ensuite ces résultats : il regarde si le biais est correct en se référant (par exemple) aux tables de RICOS et si ce biais alterne bien entre des valeurs positives et des valeurs négatives.

Si les résultats ne sont pas bons, le biologiste cherche alors à trouver l'origine du problème afin de le résoudre. Pour cela le biologiste va s'aider les représentations de Levey-Jennings, réalisées à partir de tous les contrôles.

Les principales sources d'erreur sont : une erreur de transcription des résultats, une mauvaise compréhension des règles et des diagrammes, une mauvaise utilisation des contrôles ou encore un mauvais étalonnage de l'appareil.

En regardant quelles règles de Westgard ont été violées, le biologiste peut avoir une idée sur le type de l'erreur :

- Si les règles de Westgard 2_{2s} , 4_{1s} et 10_{moy} ont été violées, le biologiste doit trouver et corriger la source de l'erreur systématique (étalonnage...).
- Si les règles de Westgard 1_{3s} et R_{4s} ont été violées, le biologiste doit trouver et corriger la source de l'imprécision (pipetage...).
- Si la règle de Westgard 1_{2s} a été violée, le biologiste doit continuer la série en surveillant les valeurs suivantes.

Remarque : pour un exemple d'analyse d'une représentation de Levey-Jennings, pour un réactif de l'ARCHITECT, voir l'annexe.

Finalement, pour mener à bien les contrôles de qualité, il est nécessaire de connaître les différences et complémentarités entre les deux contrôles de qualité interne (CQI) et externe (CQE). Il faut être capable de mettre en place un recueil de résultats de CQI pour chaque analyte : calculer l'écart type ou déviation standard, créer un tableau de Levey-Jennings et utiliser ce tableau. Il faut être en mesure d'évaluer les résultats d'un analyte à l'aide des règles de Westgard et d'identifier puis corriger les principales sources d'erreur.

2. Les modèles mathématiques

Selon les automates, les méthodes de traitement des données ou de calibration utilisent différents modèles mathématiques pour analyser un ensemble de données défini et calculer des résultats. La complexité de ces modèles mathématiques varie et peut aller d'une simple soustraction à des équations polynomiales.

Pour exemple, je me suis intéressée plus en détail aux méthodes de traitement de données des analyseurs c et i de l'ARCHITECT, ainsi qu'aux méthodes d'ajustement de l'analyseur i.

2.1. Les méthodes de traitement de données de l'analyseur c.

L'ARCHITECT c utilise une technique photométrique et potentiométrique pour mesurer les concentrations en substances à analyser dans les échantillons. Les méthodes de traitement des données ou de calibration sont propres à chaque type de technique.

2.1.1. Méthodes de traitement de données photométriques.

Les méthodes de traitement des données ou de calibration analysent les données de lecture photométrique pour calculer les résultats. Les méthodes de traitement des données photométrique utilisées sont les suivantes :

- la méthode d'absorbance, qui repose sur la comparaison entre l'absorbance de l'échantillon et l'absorbance de l'eau pour calculer les résultats ;
- la méthode du facteur, qui utilise un blanc réactif et un facteur de calibration fixe pour calculer les résultats ;

- la méthode linéaire, qui utilise un blanc réactif et entre un et six calibrateurs pour générer une courbe de calibration en point par point. La pente est calculée pour chaque segment de la courbe entre les concentrations des calibrateurs ;
- la méthode logit-4 et la , qui utilise un blanc réactif et trois à six calibrateurs pour générer une courbe de calibration ;
- la méthode spline, qui utilise un blanc réactif et trois à six calibrateurs pour générer une courbe de calibration et l'axe de concentration de la courbe de calibration est divisé en plusieurs parties, dont les subdivisions correspondent aux concentrations des calibrateurs. Chaque partie de la courbe fait l'objet d'une interpolation à l'aide d'une expression polynomiale de manière à ce que les parties se joignant soient reliées sans écart ;
- la méthode d'utilisation du facteur et du blanc réactif, qui utilise le facteur et le blanc réactif d'une courbe de calibration générée pour un autre dosage (dosage de référence) pour calculer les résultats.

2.1.2. Méthodes de traitement de données potentiométriques.

La méthode de traitement des données potentiométriques analyse les données de potentiel électrique pour calculer les résultats des électrolytes. La méthode potentiométrique comporte trois volets :

- la mesure de la force électromotrice, qui utilise la différence (différence de potentiel) entre la force électromotrice d'un échantillon patient et la force électromotrice de la solution de référence ICT, mesurée immédiatement après l'échantillon, pour déterminer la concentration en électrolytes ;
- le calcul de la pente, qui utilise la différence (différence de potentiel) entre les forces électromotrices de deux calibrateurs et la force électromotrice de la solution de référence ICT et les compare aux concentrations des calibrateurs pour générer une courbe de calibration. Deux graphes sont alors générés, le graphe de la calibration faisant apparaître la relation entre la concentration en électrolyte ionique et la force électromotrice, puis la courbe de réponse de l'électrode pendant la calibration ;
- la mesure de l'échantillon, qui utilise les données suivantes pour déterminer la concentration (ICT) en électrolytes dans un échantillon patient : la différence entre les forces électromotrices de l'échantillon patient et la solution de référence ICT, la différence entre les forces électromotrices du calibrateur de niveau bas et la solution de référence ICT, puis la pente de l'électrode ICT concernée.

2.2. Les méthodes de traitement de données de l'analyseur i.

L'ARCHITECT utilise la technique CMIA (dosage immunologique microparticulaire par chimiluminescence) pour mesurer les concentrations en substances à doser dans les échantillons. Les méthodes de traitement des données ou de calibration sont propres à ce type de technique.

2.2.1. Méthode en point par point.

La méthode en point par point utilise la valeur moyenne URL (unité relative de lumière) obtenue pour chaque calibrateur et la compare à la concentration du calibrateur pour générer une courbe de calibration. Une ligne droite est utilisée pour relier chaque point et la pente est calculée pour chaque portion de ligne. La concentration d'un échantillon patient est calculée à partir de la portion de la droite qui comprend la valeur URL de l'échantillon.

2.2.2. Méthode par régression linéaire.

La méthode par régression linéaire utilise le rapport linéaire existant entre la valeur URL et la concentration en substance à doser dans l'échantillon comme suit pour générer une courbe de calibration : $\text{valeur URL} = \text{ordonnée à l'origine} + \text{pente} * \text{concentration}$.

Dans ce cas, l'ordonnée à l'origine est la valeur URL à une concentration zéro. La pente indique l'augmentation de la concentration par rapport à l'augmentation de la valeur URL.

2.2.3. Méthode 4PLC

Les méthodes 4PLC (ajustement de la courbe logistique à quatre paramètres ou calibration logistique à quatre paramètres) utilisent la différence entre la concentration prévue et la concentration mesurée du calibrateur (ou le signal) pour générer une courbe de calibration.

Il existe deux types de méthode d'ajustement de la courbe logistique à 4 paramètres :

- la méthode 4PLC avec minimisation résiduelle de la pondération x, qui utilise la différence entre la concentration théorique et la concentration réelle du calibrateur pour générer une courbe de calibration ;
- la méthode 4PLC avec minimisation résiduelle de la pondération y, qui utilise la différence entre l'URL théorique et l'URL réel du calibrateur pour générer une courbe de calibration.

2.2.4. Méthode des dosages par seuil.

La méthode des dosages par seuil utilise une calibration en un ou deux points pour déterminer le point (valeur seuil) qui distingue les échantillons réactifs des échantillons non réactifs. Les formules valeur seuil peuvent utiliser l'addition, la soustraction et la multiplication et il est possible d'ajouter des constantes. Une fois la valeur seuil déterminée, la valeur URL de l'échantillon patient est divisée par la valeur seuil pour déterminer le ratio valeur seuil (indice).

Pour d'avantage de détail sur ces méthodes, on peut regarder :

- la méthode qualitative en un point, qui permet l'utilisation d'un calibrateur dans la formule indice ;
- la méthode qualitative en deux points, qui permet l'utilisation de deux calibrateurs dans la formule indice.

2.3. Les méthodes d'ajustement de l'analyseur i.

Les méthodes d'ajustement permettent d'analyser deux niveaux de calibrateurs au lieu de six pour calibrer un dosage. Les deux niveaux de calibrateurs (d'ajustement) servent à adapter la courbe de calibration principale mémorisée dans le code-barres du réactif aux valeurs URL mesurées sur l'analyseur.

L'analyseur i permet les méthodes suivantes d'ajustement de la calibration :

- la technique des ratios, qui compare la valeur URL des calibrateurs aux valeurs correspondantes des calibrateurs à partir de la courbe de calibration principale. Elle est utilisée si le bout inférieur d'une courbe de calibration a des caractéristiques différentes du reste de la courbe, alors que le tracé général de la courbe est proportionnel et linéaire ;

- la méthode de transformation linéaire, qui présuppose un rapport linéaire entre les valeurs URL du calibrateur générées par le système et les informations relatives à la calibration principale enregistrées dans le code-barres réactif ;
- la méthode d'ajustement des paramètres, qui utilise les valeurs URL mémorisées dans le code-barres du réactif pour les calibrateurs A à F pour déterminer les paramètres 4PLC de la calibration principale. Les calibrateurs 1 et 2 sont ensuite utilisés pour générer une courbe de calibration spécifique à l'appareil utilisant ces paramètres ;
- la méthode d'ajustement de la forme de la courbe, qui utilise les valeurs URL mémorisées dans le code-barres du réactif pour les calibrateurs A à F pour déterminer les paramètres 4PLC de la calibration principale. Les calibrateurs 1 et 2 sont ensuite utilisés pour générer une courbe de calibration spécifique à l'appareil tenant compte de la forme de cette courbe.

Cette étude donne un aperçu des modèles mathématiques qui peuvent être utilisés dans les méthodes de traitement des données ou de calibration, pour analyser un ensemble de données défini et calculer des résultats.

Notons que ceci n'est valide que pour l'ARCHITECT. En effet, comme les méthodes de traitement sont spécifiques à chaque type de technique et que les techniques sont propres à chaque appareil, suivant les automates les modèles mathématiques utilisés pour l'analyse et le calcul peuvent varier et être plus ou moins complexes.

CONCLUSION

Ce stage de trois semaines dans ce laboratoire d'analyses médicales Biogroup de Valenciennes a été vraiment enrichissant, d'une part pour ma culture personnelle, et d'autre part pour ma culture professionnelle, scientifique et mathématique.

En effet, la vie au sein d'un laboratoire d'analyses biologiques et médicales m'était inconnue et m'a donné un aperçu du travail en entreprise, dans un milieu scientifique.

J'ai pu remarquer que les mathématiques ont un rôle important en ce qui concerne l'analyse médicale. Elles interviennent effectivement dans un domaine incontournable de l'analyse qui fait appel à des connaissances statistiques : le contrôle de qualité.

De plus, le monde de l'analyse médicale, est aujourd'hui presque entièrement automatisé, et cela, grâce aux progrès scientifiques, qui permettent la conception de machines toujours plus performantes, pour un gain d'efficacité, de rapidité, de précision, de fidélité...

Dans ce laboratoire par exemple, un certain nombre de machines vont être retirées pour laisser place à une nouvelle machine plus performante, équipée d'un rail, permettant plus de rapidité et de fonctionnalité, pour les analyses.

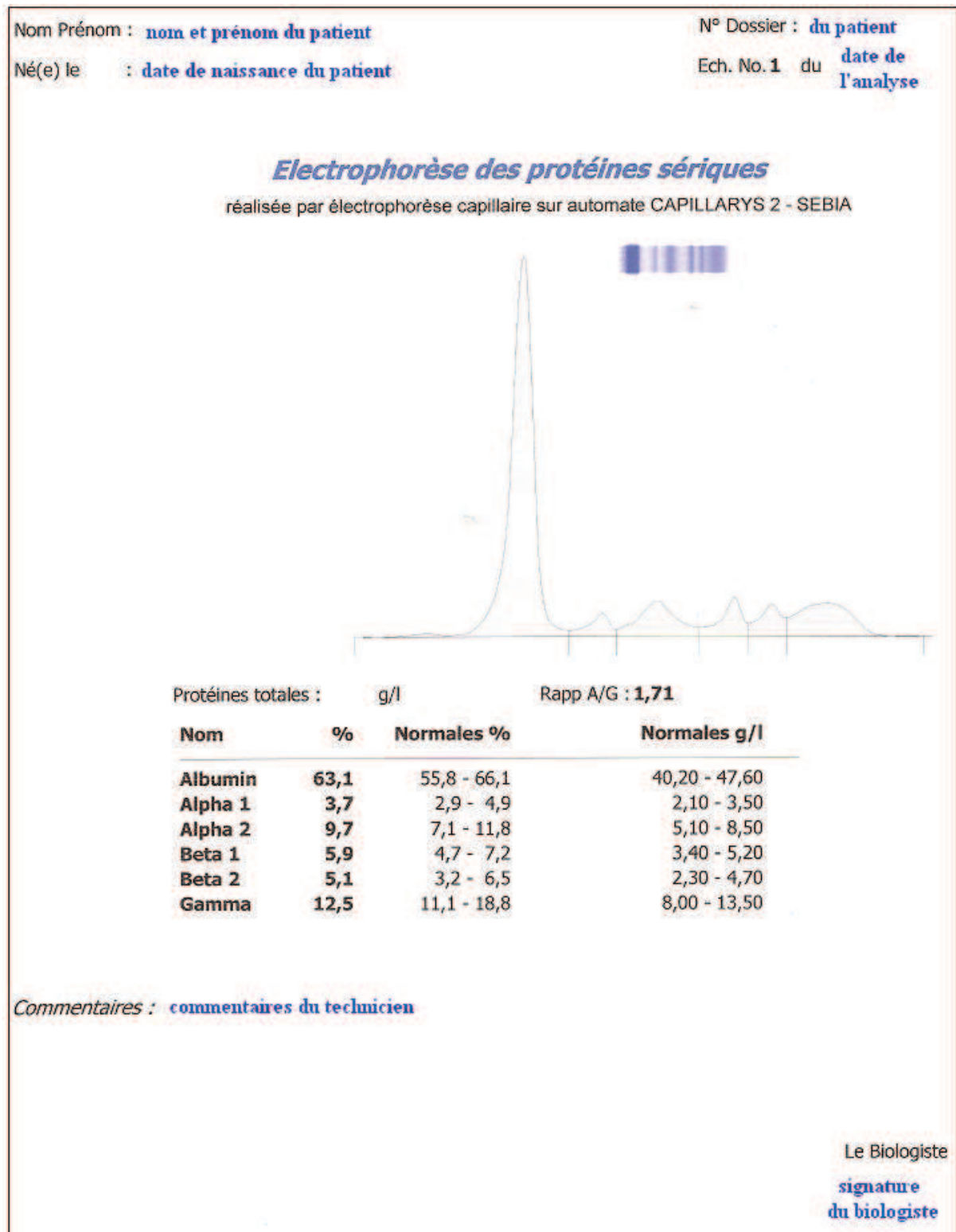
L'étude des manuels techniques des machines mis à ma disposition, m'a permis de constater la complexité de la conception de ces machines et de mettre en évidence l'intervention non négligeable de la physique et des mathématiques dans leur fonctionnement.

Finalement, les mathématiques sont bel et bien présentes dans les laboratoires d'analyses médicales et y jouent malgré tout un rôle important, que ce soit dans le cadre des contrôles de qualité ou dans le fonctionnement des machines de plus en plus optimisées.

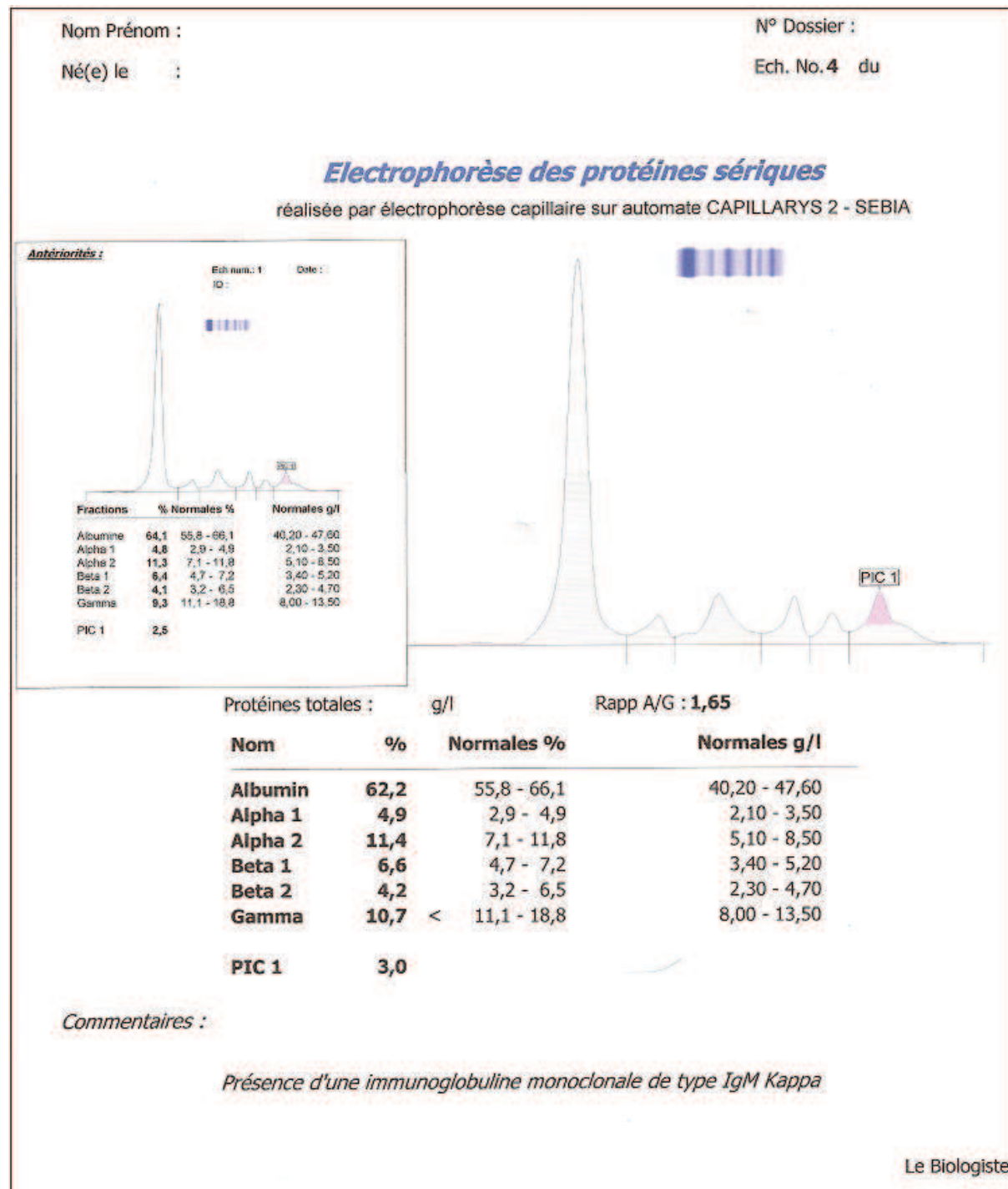
ANNEXE

1. Exemples de résultats d'analyses sur le CAPILLARYS

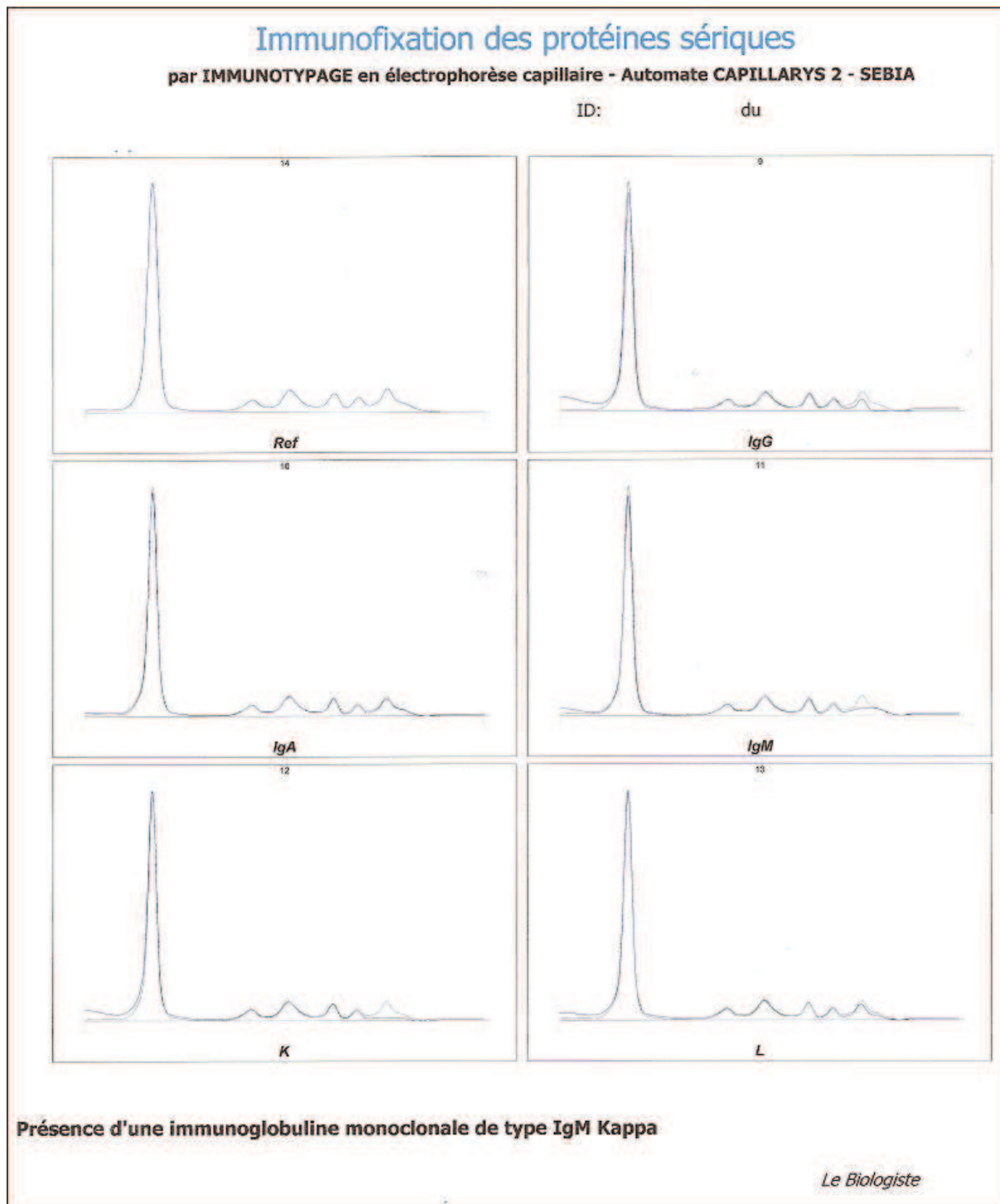
Un résultat négatif se présente comme suit :



Si l'échantillon est positif, le résultat peut être le suivant :



Le technicien a pu commenté ce résultat, en analysant l'immunofixation des protéines sériques, c'est à dire les graphiques qui suivent.



Avec le CAPILLARYS, le dosage de la CDT peut être effectué.

Le kit CAPILLARYS CDT permet la séparation en milieu alcalin des isoformes de la transferrine du sérum humain par électrophorèse capillaire dans le système automatique CAPILLARYS. Les formes faiblement sialylées représentent les formes de la transferrine appelées CDT, marqueurs biochimiques de la consommation chronique d'alcool.

Les isoformes de la transferrine sont séparées en cinq fractions selon leur degré de sialylation : asialotransferrine (forme non sialylée), di-sialotransferrine, tri-sialotransferrine, tétra-sialotransferrine et penta-sialotransferrine. La quantité de 0- et de di-sialo transferrine est donnée en pourcentage de la transferrine totale.

Sur les figures suivantes sont représentés les graphiques de plusieurs résultats possibles au test CDT :

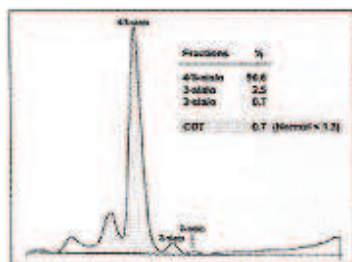


Figure 1 : Contrôle normal.

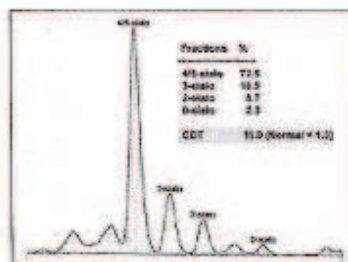


Figure 2 : Contrôle pathologique.

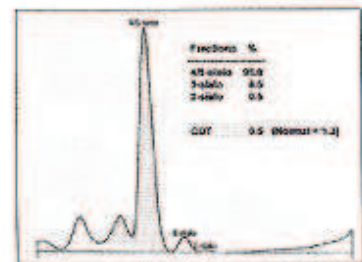


Figure 3 : Échantillon Normal.

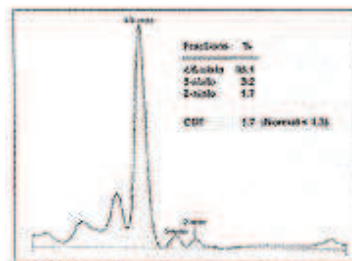


Figure 4 : Échantillon positif avec une CDT élevée.

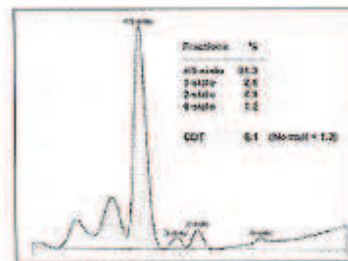


Figure 5 (a) : Profil d'un buveur excessif.

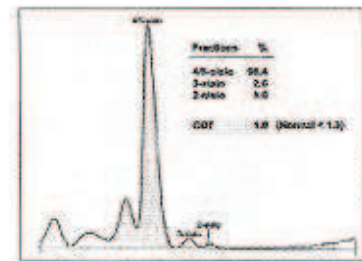


Figure 5 (b) : Profil de patient avec un mois d'abstinence.

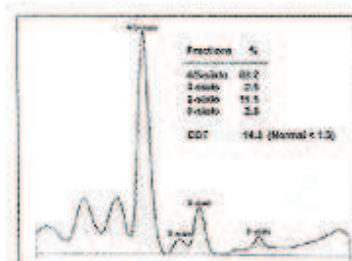


Figure 6 : Échantillon positif avec une forte CDT due à une maladie du foie.

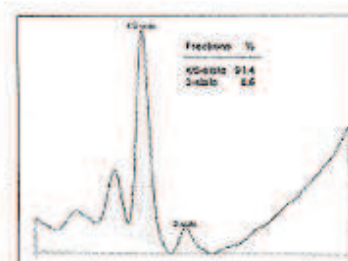


Figure 7 : Profil d'un patient chronique.

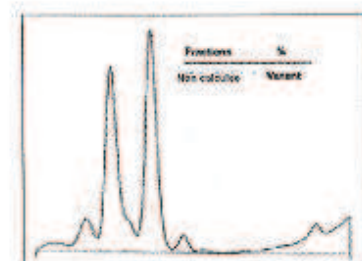
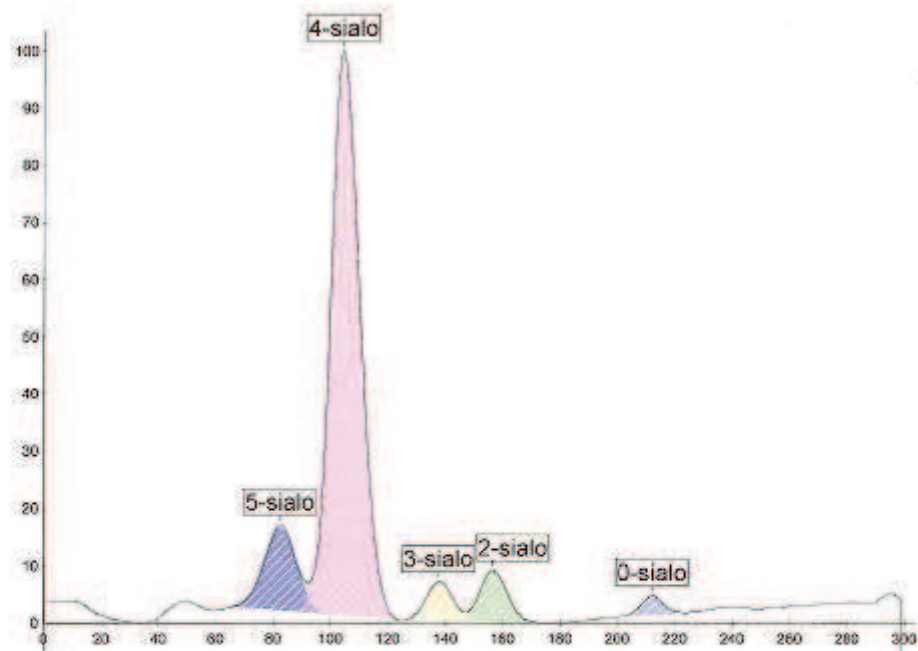


Figure 8 : Variant génétique de la transferrine.

Exemples d'un résultat positif puis d'un résultat négatif au test CDT :

Electrophorèse des isoformes de la Transferrine

par électrophorèse capillaire sur automate Capillarys 2 - SEBIA

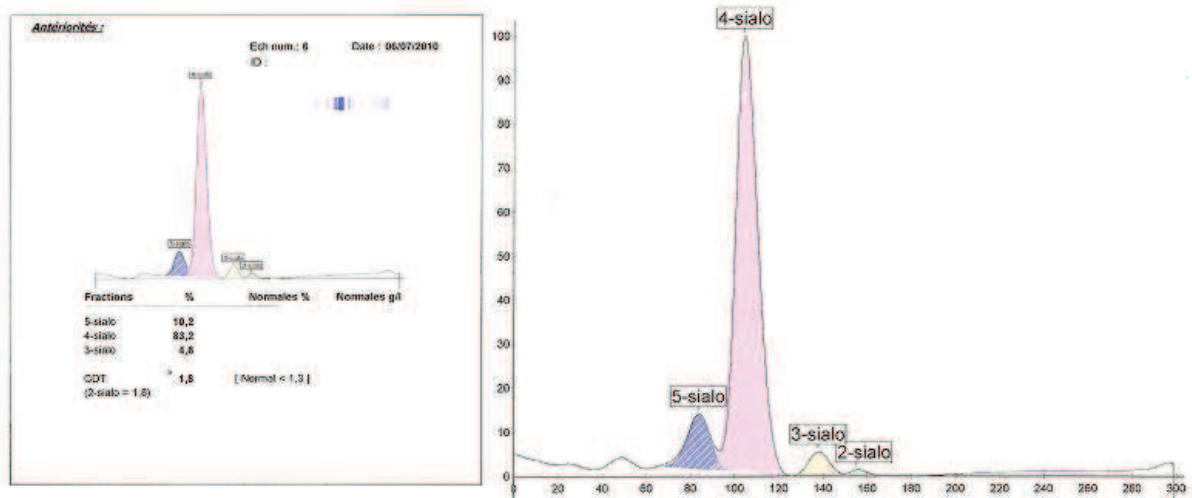


Nom	%
5-sialo	12,1
4-sialo	75,0
3-sialo	4,6

CDT **8,3** [Normal < 1,3]
(2-sialo = 6,4) + (0-sialo = 1,9)

Electrophorèse des isoformes de la Transferrine

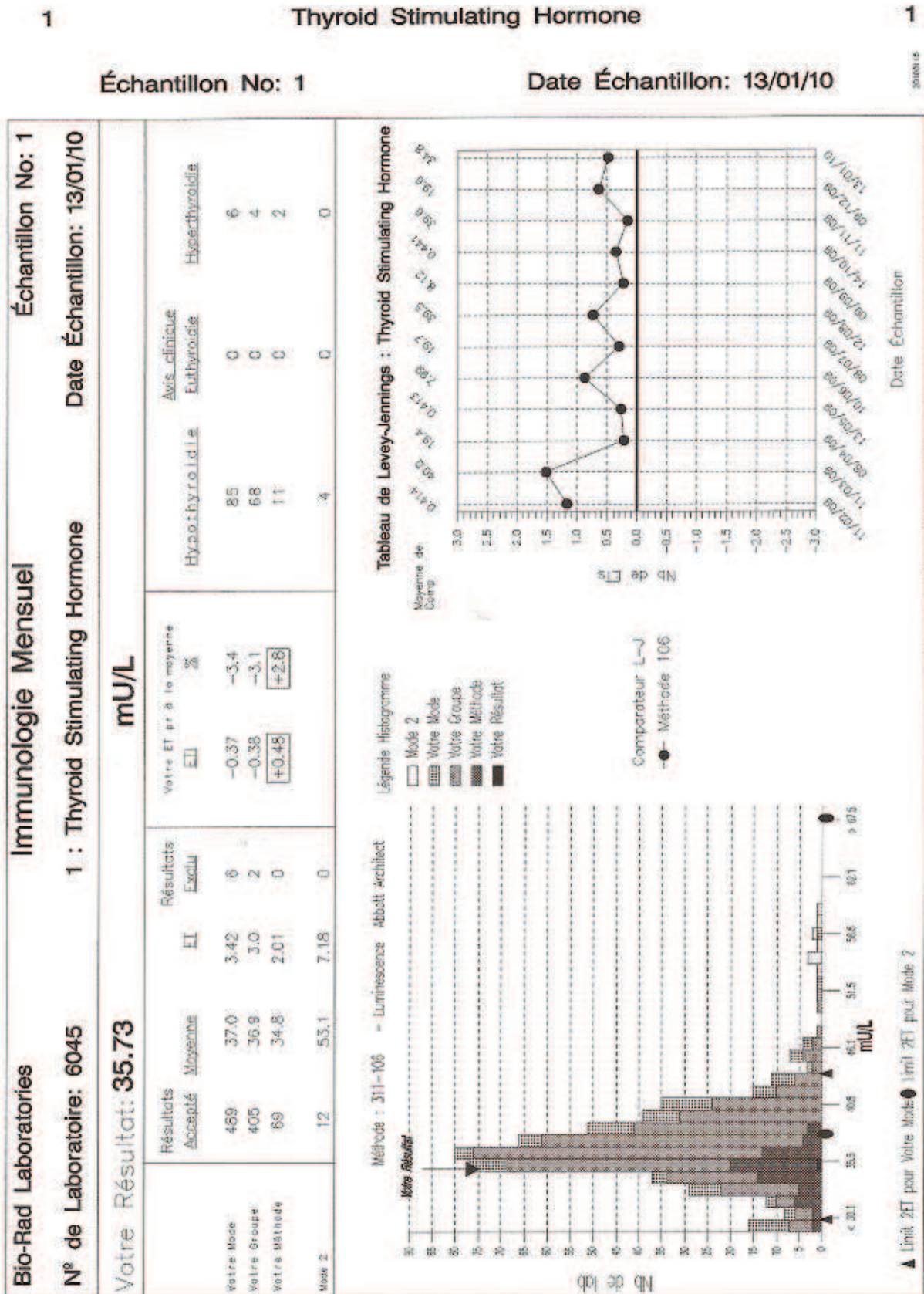
par électrophorèse capillaire sur automate Capillarys 2 - SEBIA



Nom	%
5-sialo	10,8
4-sialo	84,1
3-sialo	4,3

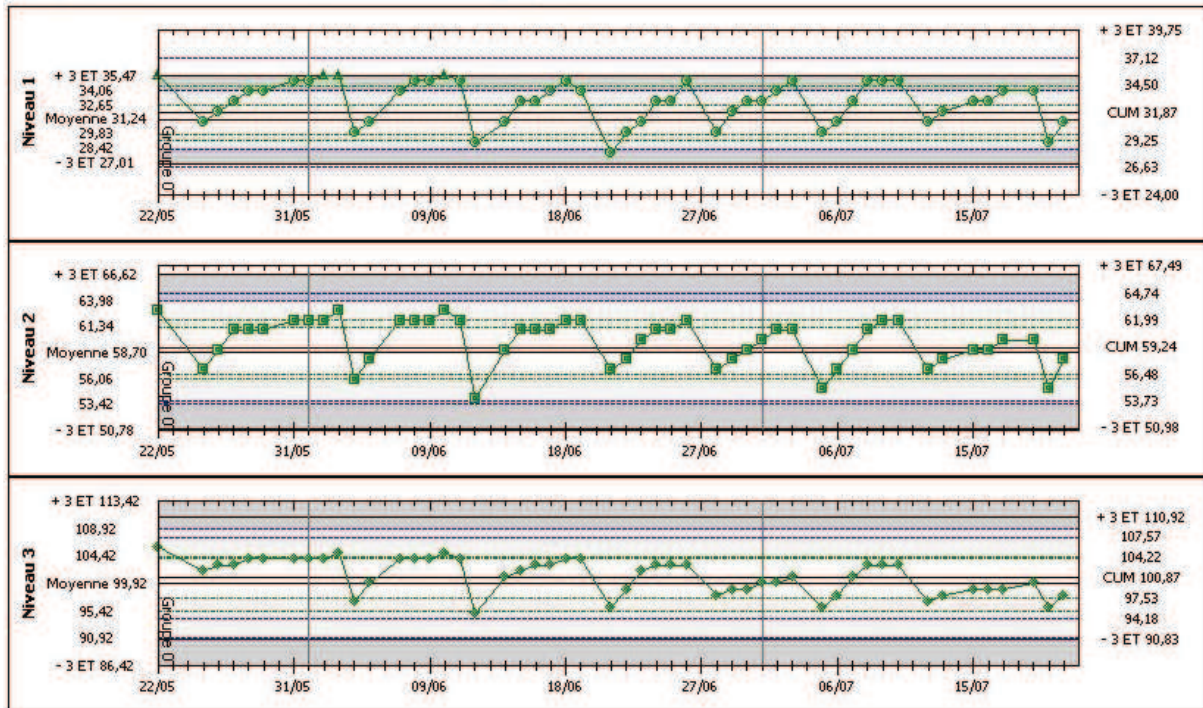
CDT 0,8 [Normal < 1,3]
(2-sialo = 0,8)

II- Exemple de rapport mensuel de CQE exploité, pour un des réactifs de l'ARCHITECT



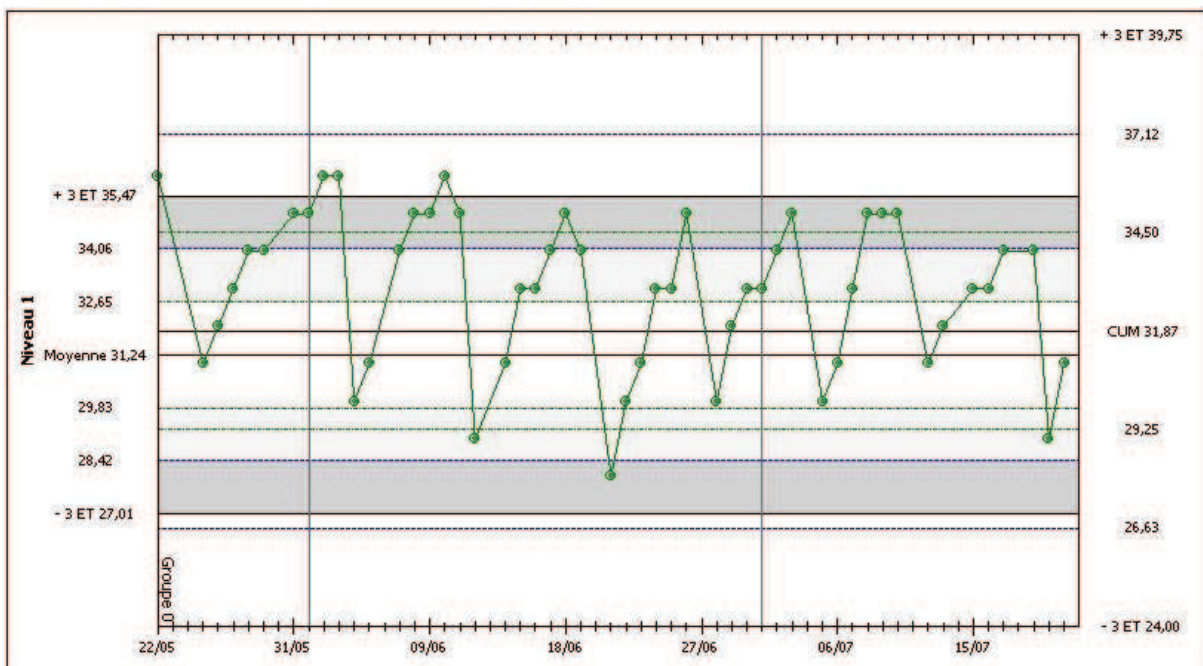
III- Exemple d'analyse d'une représentation de Levey-Jennings, pour un des réactifs de l'ARCHITECT

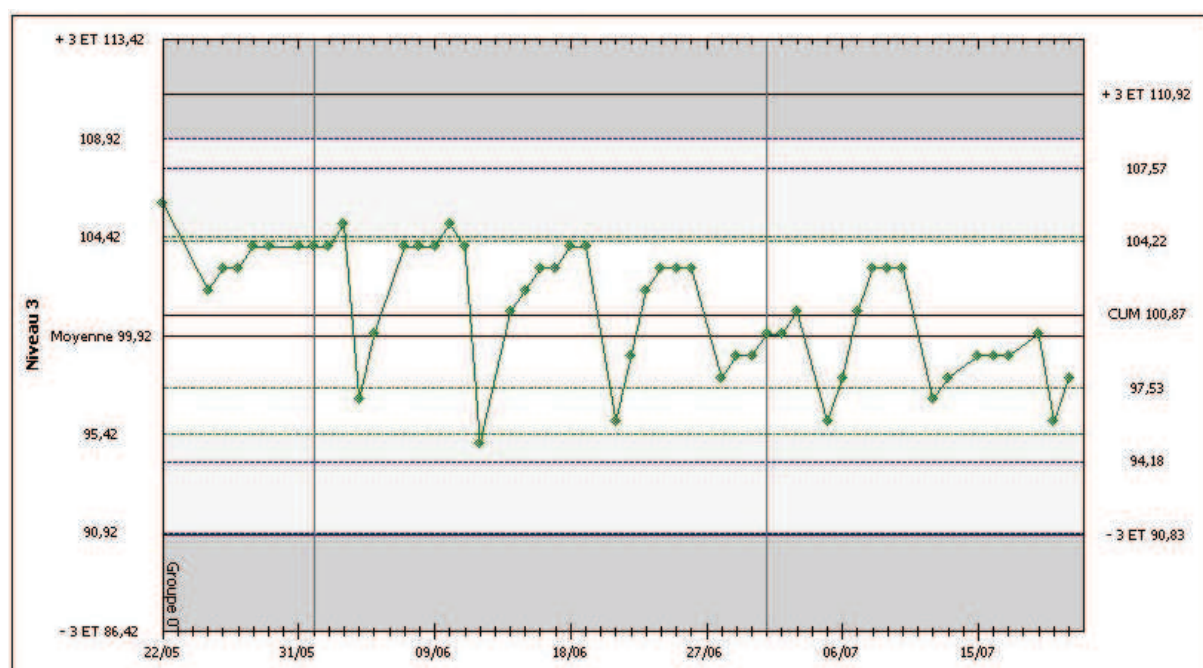
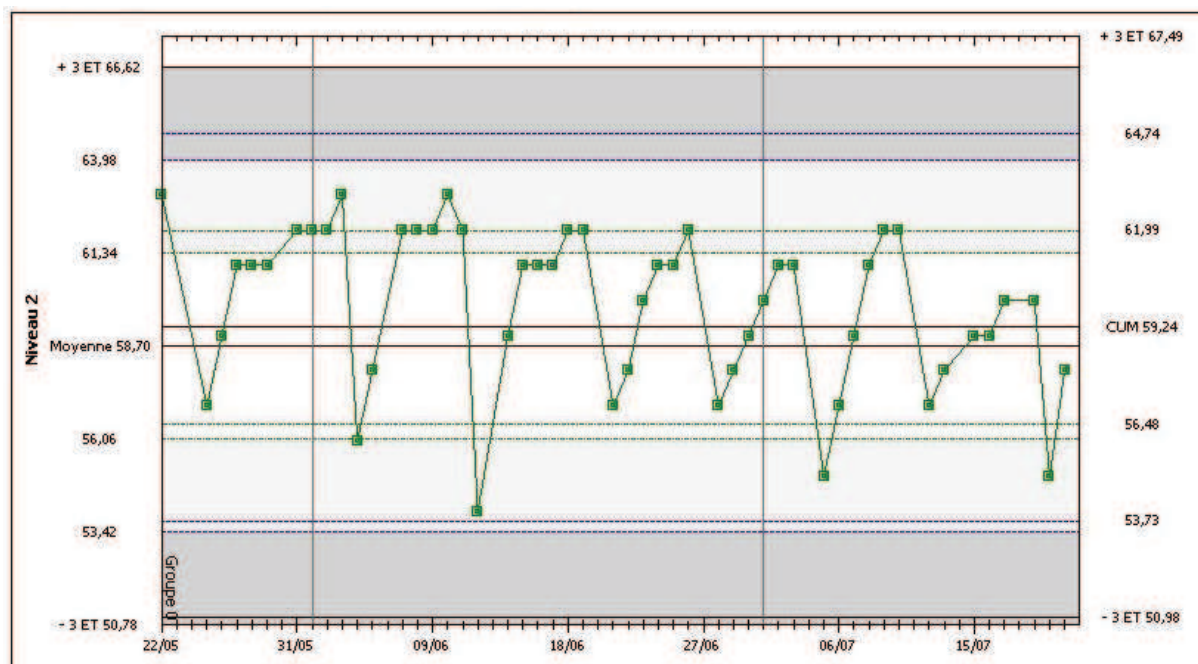
Il s'agit du problème de l'acide urique. Les graphiques qui suivent sont les représentations de Levey-Jennings pour les trois niveaux :



On peut constater un problème quant à la répartition des contrôles successifs.

On regarde de plus près chaque niveau :





On constate que pour tous les niveaux la fin de l'ascension s'effectue le même jour. Il peut s'agir d'une nouvelle calibration (nouvel étalonnage). Mais en regardant les dates de calibration de l'appareil ça ne correspond pas. Il s'agit donc d'un autre problème. Il peut s'agir d'un changement de lot de réactifs. Et là, les dates correspondent bien avec l'ouverture de nouveaux flacons du réactif. Cependant, l'acide urique se conserve très bien, donc l'augmentation des valeurs en s'éloignant de la valeur moyenne, après l'ouverture d'un flacon n'est pas expliquée. Il s'agit donc peut-être d'un problème d'évaporation du réactif. Pour remédier à ce problème le biologiste propose de faire des aliquots du réactif lors de l'ouverture du flacon, afin d'avoir la même concentration pour tous les contrôles. C'est peut-être la solution au problème.

Université Paris-sud 11
Faculté de mathématiques
Master 1 MFA - TER

Tuteur : Francis Choucroun

Les Nombres p -adiques et les Formes Quadratiques

Lucile Morelle

Orsay, le 3 juin 2011

Table des matières

1	Les nombres p-adiques	5
1.1	L'anneau \mathbb{Z}_p des entiers p -adiques	5
1.1.1	Définitions.	5
1.1.2	Les propriétés de \mathbb{Z}_p	6
1.2	Le corps \mathbb{Q}_p des nombres p -adiques.	7
1.2.1	Définition.	7
1.2.2	Les premières propriétés de \mathbb{Q}_p	8
1.2.3	L'espace métrique \mathbb{Q}_p	10
1.3	Equations p -adiques.	12
1.4	Les carrés de \mathbb{Q}_p	12
2	Les symboles de Legendre et Hilbert	15
2.1	Le symbole de Legendre.	15
2.1.1	Les carrés de \mathbb{F}_q	15
2.1.2	Définition et propriétés.	16
2.1.3	La loi de réciprocité quadratique.	16
2.2	Le symbole de Hilbert.	17
2.2.1	Définition et propriétés.	17
2.2.2	La bilinéarité du symbole de Hilbert.	19
2.2.3	Propriétés globales.	19
3	Formes quadratiques.	23
3.1	Définitions.	23
3.2	Orthogonalité.	24
3.3	Vecteurs isotropes.	25
3.4	Bases orthogonales.	26
3.5	Théorème de Witt.	28
3.6	Traductions.	29
3.7	Formes quadratiques sur \mathbb{F}_q	31
4	Formes quadratiques sur \mathbb{Q}_p.	32
4.1	Les deux invariants.	32
4.2	Représentation d'un élément de \mathbb{Q}_p par une forme quadratique.	33
4.3	Classification.	36
4.4	Le cas réel.	37

5	Formes quadratiques sur \mathbb{Q}.	38
5.1	Invariants d'une forme.	38
5.2	Représentation d'un nombre par une forme.	38
5.3	Classification.	42

Introduction

En mathématiques, et plus particulièrement en théorie algébrique des nombres et en géométrie algébrique, le principe local-global consiste à essayer de reconstituer une information sur un objet global à partir d'informations sur des objets locaux associés (ses localisations en tous les idéaux premiers), plus faciles à obtenir. L'un des premiers résultats (et peut-être le plus célèbre) de passage du local au global est le théorème de Hasse-Minkowski, démontré dans toute sa généralité par Helmut Hasse en 1924.

L'objet de ce travail est de démontrer le principe de Hasse, en passant par la classification des formes quadratiques sur \mathbb{Q}_p , pour aboutir à la classification des formes quadratiques sur \mathbb{Q} .

Ce TER se divise en cinq chapitres. Dans un premier temps, on introduit les nombres p -adiques. Ils jouent un rôle crucial dans la classification des formes quadratiques sur \mathbb{Q} , on en détaille donc les propriétés. Dans un second temps, on s'intéresse aux symboles de Legendre et Hilbert. Ensuite, on fait quelques rappels sur les formes quadratiques. C'est au chapitre 4 que la classification des formes quadratiques sur le corps des nombres p -adiques \mathbb{Q}_p a lieu. Enfin, le dernier chapitre est consacré à la démonstration du théorème de Hasse-Minkowski, dont découle la classification des formes quadratiques sur \mathbb{Q} .

Chapitre 1

Les nombres p -adiques

Dans tout ce chapitre, p va désigner un nombre premier.

1.1 L'anneau \mathbb{Z}_p des entiers p -adiques

1.1.1 Définitions.

Commençons par définir les entiers p -adiques :

Définition 1.1.1. *Un entier p -adique est une classe d'équivalence de suites $x := \{x_0, x_1, \dots, x_n, \dots\}$ d'entiers telle que $\forall n, x_n \equiv x_{n-1} \pmod{p^n}$; deux suites sont équivalentes si $x_n \equiv x'_n \pmod{p^n}, \forall n$.*

On peut également écrire la suite d'entiers sous la forme

$$\{a_0, a_0 + a_1p, a_0 + a_1p + a_2p^2, \dots\}$$

et si l'on veut, se borner à prendre des entiers a_i dans l'intervalle $[0, p-1]$. Cela suggère l'écriture suivante pour un entier p -adique " $x = \sum_{i=0}^{\infty} a_i p^i$ ".

À présent, intéressons nous à l'anneau des entiers p -adiques :

Pour tout $n \geq 1$, posons $\mathbb{A}_n = \mathbb{Z}/p^n\mathbb{Z}$; c'est l'anneau des classes d'entiers $(\text{mod } p^n)$. Un élément de \mathbb{A}_n définit de manière évidente un élément de \mathbb{A}_{n-1} ; on obtient ainsi un homomorphisme

$$\varphi_n : \mathbb{A}_n \longrightarrow \mathbb{A}_{n-1}$$

qui est surjectif, et de noyau $p^{n-1}\mathbb{A}_n$. La suite :

$$\dots \longrightarrow \mathbb{A}_n \longrightarrow \mathbb{A}_{n-1} \longrightarrow \dots \longrightarrow \mathbb{A}_2 \longrightarrow \mathbb{A}_1$$

forme un "système projectif", indexé par les entiers ≥ 1 .

Définition. *Un système projectif d'ensembles (resp. de groupes, resp. d'anneaux) indexé par \mathbb{N} est la donnée de :*

- i) une suite $(X_n)_{n \in \mathbb{N}}$ d'ensembles (resp. de groupes, resp. d'anneaux) ;*
- ii) une suite $(\varphi_n : X_{n+1} \rightarrow X_n)_{n \in \mathbb{N}}$ d'applications (resp. d'homomorphismes de groupes, resp. d'homomorphismes d'anneaux).*

Définition. Soit (X_n, φ_n) un système projectif. On appelle **limite projective** de ce système l'ensemble

$$X = \varprojlim (X_n, \varphi_n) = \{(x_n)_{n \in \mathbb{N}} \in \prod_{n \in \mathbb{N}} X_n \mid \forall n \in \mathbb{N}, \varphi_n(x_{n+1}) = x_n\}$$

Définition 1.1.2. On appelle **anneau des entiers p -adiques**, et on note \mathbb{Z}_p , la limite projective du système $(\mathbb{A}_n, \varphi_n)$ défini ci-dessus.

Par définition, un élément de $\mathbb{Z}_p = \varprojlim (\mathbb{A}_n, \varphi_n)$ est une suite $x = (\dots, x_n, \dots, x_1)$, avec :

$$x_n \in \mathbb{A}_n \text{ et } \varphi_n(x_n) = x_{n-1} \text{ si } n \geq 2.$$

L'addition et la multiplication de \mathbb{Z}_p sont définies "coordonnées par coordonnées"; autrement dit, \mathbb{Z}_p est un *sous-anneau* du produit $\prod_{n \geq 1} \mathbb{A}_n$. Si l'on munit les \mathbb{A}_n de la topologie discrète, et $\prod \mathbb{A}_n$ de la topologie produit, l'anneau \mathbb{Z}_p se trouve muni d'une topologie qui en fait un espace *compact* :

Proposition 1.1.3. L'anneau $\prod_{n \geq 1} \mathbb{A}_n$ (muni de la topologie produit) est compact.

Preuve :

Pour tout entier naturel n strictement positif, l'anneau A_n est compact car fini. La proposition résulte du théorème de Tychonoff. \square

Proposition 1.1.4. L'anneau \mathbb{Z}_p est compact.

Preuve :

Il suffit de montrer que \mathbb{Z}_p est fermé dans $\prod_{n \geq 1} \mathbb{A}_n$. Comme ce dernier est compact, \mathbb{Z}_p le sera aussi.

Soit $(x(n))_{n \in \mathbb{N}} \in \mathbb{Z}_p^{\mathbb{N}}$, une suite de \mathbb{Z}_p convergeant vers $x \in \prod_{n \geq 1} \mathbb{A}_n$. Soit $x_q(n) \in \mathbb{Z}/p^q\mathbb{Z}$ la q -ième coordonnée de $x(n)$ et x_q celle de x . Il s'agit de montrer que x est un élément de \mathbb{Z}_p , c'est-à-dire que pour tout entier naturel q strictement positif $\varphi_{q+1}(x_{q+1}) = x_q$. Pour tout $q > 0$, l'application φ_{q+1} est continue de $\mathbb{Z}/p^{q+1}\mathbb{Z}$ dans $\mathbb{Z}/p^q\mathbb{Z}$ (c'est une application entre deux espaces munis de la topologie discrète). Le résultat souhaité s'obtient par passage à la limite dans l'égalité $\varphi_{q+1}(x_{q+1}(n)) = x_q(n)$ (qui est vraie car pour tout n , on a $x(n) \in \mathbb{Z}_p$). \square

1.1.2 Les propriétés de \mathbb{Z}_p .

Soit $\varepsilon_n : \mathbb{Z}_p \rightarrow \mathbb{A}_n$ l'application qui associe à un entier p -adique x sa n -ième composante x_n .

Proposition 1.1.5. La suite $0 \rightarrow \mathbb{Z}_p \xrightarrow{p^n} \mathbb{Z}_p \xrightarrow{\varepsilon_n} \mathbb{A}_n \rightarrow 0$ est exacte.

(On peut donc identifier $\mathbb{Z}_p/p^n\mathbb{Z}_p$ à $\mathbb{A}_n = \mathbb{Z}/p^n\mathbb{Z}$.)

Preuve :

La multiplication par p (donc aussi par p^n) est injective dans \mathbb{Z}_p ; en effet, si $x = (x_n)$ est un entier p -adique tel que $px = 0$, on a $px_{n+1} = 0$ pour tout n , ce qui entraîne que x_{n+1} est de la forme $p^n y_{n+1}$, avec $y_{n+1} \in \mathbb{A}_{n+1}$; comme $x_n = \varphi_{n+1}(x_{n+1})$, on voit que x_n est également divisible par p^n , donc est nul. Il est clair que le noyau de ε_n contient $p^n\mathbb{Z}_p$; inversement, si $x = (x_m)$ appartient à $\text{Ker}(\varepsilon_n)$, on a $x_m \equiv 0 \pmod{p^n}$ pour tout $m \geq n$, ce qui signifie qu'il

existe un élément y_{m-n} bien déterminé de \mathbb{A}_{m-n} tel que $x_m = p^n y_{m-n}$. Les y_i définissent un élément y de $\mathbb{Z}_p = \varprojlim A_i$, et l'on vérifie tout de suite que $p^n y = x$, ce qui achève de démontrer la proposition. \square

Proposition 1.1.6. *i) Pour qu'un élément de \mathbb{Z}_p (resp. de \mathbb{A}_n) soit inversible, il faut et il suffit qu'il ne soit pas divisible par p .*

ii) Si \mathbb{U} désigne l'ensemble des éléments inversibles de \mathbb{Z}_p , tout élément de \mathbb{Z}_p différent de 0 s'écrit de façon unique sous la forme $p^n u$, avec $u \in \mathbb{U}$ et $n \geq 0$.

(Un élément de \mathbb{U} est appelé une *unité p -adique*.)

Preuve :

Il suffit de prouver *i)* pour les \mathbb{A}_n . Si $x \in \mathbb{A}_n$ n'appartient pas à $p\mathbb{A}_n$, son image dans $\mathbb{A}_1 = \mathbb{F}_p$ est non nulle, donc inversible ; il existe alors $y, z \in \mathbb{A}_n$ tels que $xy = 1 - pz$, d'où $xy(1 + pz + \dots + p^{n-1}z^{n-1}) = 1$, ce qui montre que x est inversible.

D'autre part, si $x \in \mathbb{Z}_p$ est non nul, il existe un plus grand entier n tel que $x_n = \varepsilon_n(x)$ soit nul ; on a alors $x = p^n u$, avec u non divisible par p , d'où $u \in \mathbb{U}$ d'après *i)*. L'unicité de cette décomposition est évidente. \square

Notation. Soit x un élément non nul de \mathbb{Z}_p ; écrivons x sous la forme $p^n u$, avec $u \in \mathbb{U}$. L'entier n est appelé la *valuation p -adique* de x , et notée $v_p(x)$. On pose $v_p(0) = +\infty$, et l'on a :

$$v_p(xy) = v_p(x) + v_p(y)$$

$$v_p(x + y) \geq \inf(v_p(x), v_p(y)).$$

Il en résulte aussitôt de ces formules que \mathbb{Z}_p est un anneau *intègre*.

Proposition 1.1.7. *La topologie de \mathbb{Z}_p peut être définie par la distance*

$$d(x, y) = e^{-v_p(x-y)}.$$

L'anneau \mathbb{Z}_p est un espace complet, dans lequel \mathbb{Z} est dense.

Preuve :

Les idéaux $p^n \mathbb{Z}_p$ forment une base de voisinages de 0 ; comme $x \in p^n \mathbb{Z}_p$ équivaut à $v_p(x) \geq n$, on voit bien que la topologie de \mathbb{Z}_p est définie par la distance

$$d(x, y) = e^{-v_p(x-y)}.$$

Comme \mathbb{Z}_p est compact, il est complet. Enfin, si $x = (x_n)$ est un élément de \mathbb{Z}_p , et si $y_n \in \mathbb{Z}$ est congru à $x_n \pmod{p^n}$, on a $\lim y_n = x$, ce qui prouve que \mathbb{Z} est dense dans \mathbb{Z}_p . \square

1.2 Le corps \mathbb{Q}_p des nombres p -adiques.

1.2.1 Définition.

Définition 1.2.1. *On appelle corps des nombres p -adiques, et on note \mathbb{Q}_p , le corps des fractions de l'anneau \mathbb{Z}_p .*

On voit tout de suite que $\mathbb{Q}_p = \mathbb{Z}_p[p^{-1}]$. Tout élément x de \mathbb{Q}_p^* s'écrit de façon unique sous la forme $p^n u$, avec $n \in \mathbb{Z}$, $u \in \mathbb{U}$; ici encore, n s'appelle la valuation p -adique de x , et se note $v_p(x)$. On a $v_p(x) \geq 0$ si et seulement si $x \in \mathbb{Z}_p$.

Remarques. 1) On aurait pu définir \mathbb{Q}_p (resp. \mathbb{Z}_p) comme le complété de \mathbb{Q} (resp. \mathbb{Z}) pour la distance p -adique d .

2) la distance d vérifie l'inégalité "ultramétrique" :

$$d(x, z) \leq \sup(d(x, y), d(y, z)).$$

On en déduit facilement qu'une suite u_n a une limite si et seulement si $\lim(u_{n+1} - u_n) = 0$; de même, une série converge si et seulement si son terme général tend vers 0.

1.2.2 Les premières propriétés de \mathbb{Q}_p .

Notons que l'anneau \mathbb{Z}_p est intègre.

Remarque. Les propriétés suivantes de \mathbb{Q}_p , se déduisent des propriétés de \mathbb{Z}_p .

Proposition 1.2.2. *Tout $x \in \mathbb{Q}_p^*$ admet une unique représentation $x = p^n u$, où $n \in \mathbb{Z}$ et u est une unité de \mathbb{Z}_p .*

Preuve :

En effet, si $x = a/b$ où a et b appartiennent à \mathbb{Z}_p , on sait que $a = p^h a'$ et $b = p^k b'$ où a' et b' sont des unités de \mathbb{Z}_p , donc $x = p^{h-k} a'/b'$. Si $p^n u = p^m v$ où u et v sont des unités de \mathbb{Z}_p , supposons par exemple que $n \geq m$, alors $p^{n-m} = vu^{-1}$ est une unité, et $n - m = v(p^{n-m}) = 0$, d'où $u = v$, ce qui prouve l'unicité. \square

Proposition 1.2.3. *La fonction v définie sur \mathbb{Q}_p par $x = p^{v(x)} u$, u unité de \mathbb{Z}_p , pour $x \neq 0$ et $v(0) = +\infty$, est une valuation sur \mathbb{Q}_p , c'est-à-dire une application de \mathbb{Q}_p dans $\mathbb{Z} \cup \{+\infty\}$ satisfaisant les propriétés suivantes :*

- i) $\forall x \in \mathbb{Q}_p, v(x) = +\infty \Leftrightarrow x = 0$;
- ii) $\forall x$ et $y \in \mathbb{Q}_p, v(xy) = v(x) + v(y)$;
- iii) $\forall x$ et $y \in \mathbb{Q}_p, v(x + y) \geq \inf(v(x), v(y))$.

Preuve :

C'est immédiat. \square

Proposition 1.2.4. *L'injection canonique de \mathbb{Z} dans \mathbb{Z}_p se prolonge de façon unique en un isomorphisme de \mathbb{Q} sur un sous-corps de \mathbb{Q}_p .*

Proposition 1.2.5. *Soient $a \in \mathbb{Q}_p$, $n \in \mathbb{Z}$, on a :*

$$V_n(a) = \{x \in \mathbb{Q}_p \mid v(x - a) \geq n\}.$$

La topologie pour laquelle $V_n(a)$ est une base de voisinages de a fait de \mathbb{Q}_p un corps topologique localement compact. L'application $y \mapsto p^{-n}(y - a)$ est un homéomorphisme de $V_n(a)$ sur $\mathbb{Z}_p = V_0(0)$.

Proposition 1.2.6. *Soit $x \in \mathbb{Q}_p$, il existe une unique suite d'entiers dans $[0, p-1]$, $(a_n)_{n \geq 0}$ et $n_0 \in \mathbb{Z}$, tels que la série $\sum_{n \geq n_0} a_n p^n$ converge vers x . Cette série est appelée **développement de Hensel** de x . Si $a_{n_0} \neq 0$, $n_0 = v(x)$.*

Preuve :

Commençons par montrer l'unicité :

Si $a = \sum_{n \geq n_0} a_n p^n$ et $b = \sum_{n \geq n_0} b_n p^n$ sont deux telles séries différentes, notons n_2 le plus petit entier relatif tel que $a_{n_2} \neq b_{n_2}$. Alors :

$$0 = |a - b|_p = |(a_{n_2} - b_{n_2})p^{n_2} + \sum_{n \geq n_2+1} (a_n - b_n)p^n|_p = p^{-n_2},$$

d'après le principe des triangles isocèles. En effet :

$$|(a_{n_2} - b_{n_2})p^{n_2}|_p = p^{-n_2} \quad \text{et} \quad \left| \sum_{n \geq n_2+1} (a_n - b_n)p^n \right|_p \leq p^{-(n_2+1)},$$

car $|\cdot|_p$ est ultramétrique. C'est absurde, on a donc l'unicité du développement. Montrons à présent l'existence :

On se ramène au cas où $|x|_p = 1$, en multipliant x par une puissance convenable de p ; en effet, l'ensemble des valeurs absolues p -adiques des rationnels est $\{p^n, n \in \mathbb{Z}\} \cup \{0\}$, qui est un fermé pour $(\mathbb{R}, |\cdot|)$ et est donc égal à l'ensemble des valeurs absolues de \mathbb{Q}_p . Ainsi, si $x \neq 0$, $|x|_p = p^m$, pour m bien choisi, et on multiplie x par p^{-m} . On montre alors que l'on peut approcher x par une suite d'entiers positifs $(n_i)_{i \geq 1}$ tels que $|x - n_i|_p \leq p^{-i}$ et $n_i \leq p^i - 1$.

Démontrons d'abord le résultat pour $x = \frac{a}{b} \in \mathbb{Q}$. Si l'on choisit a et b premiers entre eux, comme $|x|_p = 1$, $p \nmid b$ et donc p^i et b sont premiers entre eux, et on choisit une relation de Bezout $kp^i + lb = 1$ (on choisit l de telle sorte que al soit un entier positif). Dès lors, $|\frac{a}{b} - al|_p = |\frac{a}{b}(1 - lb)|_p = |kp^i|_p \leq p^{-i}$.

On peut alors écrire la décomposition de al en base p :

$$al = \sum_{j < i} c_j p^j + \sum_{j \geq i} c_j p^j.$$

On a bien :

$$d = \sum_{j < i} c_j p^j \leq p^i - 1$$

et

$$|x - d|_p = |x - al + (al - d)|_p \leq \max\{|x - al + (al - d)|_p, \left| \sum_{j < i} c_j p^j \right|_p\} = p^{-i}.$$

Maintenant, si $x \in \mathbb{Q}_p$, on choisit $y \in \mathbb{Q}$ tel que $|x - y|_p \leq p^{-i}$ et $n_i \leq p^i - 1$ tel que $|y - n_i|_p \leq p^{-i}$. On obtient que $|x - n_i|_p = |(x - y) + (y - n_i)|_p \leq p^{-i}$. Pour finir, il suffit de voir, si $x \in \mathbb{Q}_p$ et $(n_i)_{i \geq 1}$ la suite d'entiers associés, alors n_i et n_{i+1} ont leurs i premiers termes de la décomposition en base p , les termes en p^j , où $j < i$ se simplifient. Par conséquent, les n_i sont les sommes partielles d'une série de la forme $\sum a_n p^n$, avec $a_n \leq p - 1$, qui converge vers x . \square

Proposition 1.2.7. *L'image canonique de \mathbb{Q} dans \mathbb{Q}_p est dense.*

Preuve :

Soient en effet $a \in \mathbb{Q}_p$, $a = \sum_{n \geq n_0} a_n p^n$ son développement de Hensel et soit $k \in \mathbb{Z}$. Le voisinage $V_k(a)$ de a contient l'image du rationnel $\sum_{n_0 \leq n < k} a_n p^n$. D'où la densité de l'image canonique de \mathbb{Q} dans \mathbb{Q}_p . \square

Proposition 1.2.8. *Soient x un rationnel non nul, $v_p(x)$ l'exposant de p dans la décomposition de x en facteurs premiers : $v_p(x)$ est la valuation de l'image de x dans \mathbb{Q}_p .*

1.2.3 L'espace métrique \mathbb{Q}_p

Si p est un nombre premier, on note $v_p : \mathbb{Q}^* \rightarrow \mathbb{Z}$ la valuation p -adique : si a est un entier, $v_p(a)$ est le plus grand entier n tel que p^n divise a , et $v_p(b^{-1}a) = v_p(a) - v_p(b)$ si a et b sont des entiers. Il n'est pas difficile de vérifier que v_p est bien définie et est une valuation sur \mathbb{Q} (en posant $v_p(0) = +\infty$).

Définition 1.2.9. On définit sur \mathbb{Q}_p la valeur absolue p -adique par $|x|_p = p^{-v(x)}$ pour $x \neq 0$ et $|0|_p = 0$. Elle satisfait la relation suivante :

$$\forall x, y \in \mathbb{Q}_p, |x + y|_p \leq \max(|x|, |y|).$$

On dit qu'une suite (u_n) converge vers l si $\lim_n |u_n - l|_p = 0$.

Lemme 1.2.10. Pour la topologie p -adique, on a les propriétés suivantes :

- i) L'adhérence de \mathbb{Z} est \mathbb{Z}_p , lequel est compact. Le corps \mathbb{Q}_p est localement compact.
- ii) Une suite $(u_n) \in \mathbb{Z}_p$ converge si et seulement si $\lim_n (u_{n+1} - u_n) = 0$. De même, une série $\sum_n u_n$ converge si et seulement si $\lim_n u_n = 0$.

Preuve :

- i) Soit $x = (x_n)_n \in \mathbb{Z}_p$. Alors, $|x - x_n| \leq p^{-n}$, et la suite d'entiers converge donc vers x . Ensuite, $\mathbb{Q}_p = \cup_{m \geq 0} p^{-m} \mathbb{Z}_p$.

$$\begin{aligned} \text{L'application : } \mathbb{Z}_p &\longrightarrow \prod_{n \geq 1} \mathbb{Z}_p / p^n \mathbb{Z}_p \\ x &\longmapsto (x \bmod p^n)_{n \geq 1} \end{aligned}$$

est injective, continue et d'image fermée.

La compacité de \mathbb{Z}_p découle donc de la compacité du produit $\prod_{n \geq 1} \mathbb{Z}_p / p^n \mathbb{Z}_p$. On en déduit immédiatement la compacité de \mathbb{Q}_p .

- ii) Le second point provient de l'inégalité ultramétrique

$$|u_M + \dots + u_N|_p \leq \max_{M \leq n \leq N} (|u_n|_p). \quad \square$$

Lemme 1.2.11. Toute boule de l'espace métrique \mathbb{Q}_p en est une partie à la fois ouverte et fermée.

Preuve :

Notons $B(a, r)$ (resp. $B'(a, r)$) la boule ouverte (resp. fermée) de centre a et de rayon r . Si $x \in \mathbb{Q}_p : |x| \in \{p^n | n \in \mathbb{Z}\}$, soient donc n et n' définis par :

$$p^n < r \leq p^{n+1} \quad \text{et} \quad p^{n'} \leq r < p^{n'+1},$$

alors :

$$B(a, r) = B(a, p^{n+1}) = B'(a, p^n)$$

et :

$$B'(a, r) = B'(a, p^{n'}) = B(a, p^{n'+1}),$$

d'où le lemme. \square

Corollaire 1.2.12. L'espace \mathbb{Q}_p est totalement discontinu.

Preuve :

En effet, un espace topologique séparé est totalement discontinu si tout point possède une base de voisinages à la fois ouverts et fermés. \square

Théorème 1.2.13. (Ostrowski) Une norme non triviale sur \mathbb{Q} est équivalente à la norme usuelle $|\cdot|_\infty$ ou à la norme p -adique pour un nombre premier p .

Preuve :

Commençons par supposer qu'il existe $k \in \mathbb{N}$ tel que $\|k\| > 1$. Comme $\|1\| = 1$, l'inégalité triangulaire implique $\|k\| \leq k$ et il existe $\alpha \in]0, 1]$ tel que l'on ait $\|k\| = k^\alpha$. Soit $m \in \mathbb{N}$. On peut écrire m en base k sous la forme $m = \sum_{i=0}^n a_i k^i$ avec $a_i \in \{0, 1, \dots, k-1\}$ et $a_n \neq 0$ de telle sorte que l'on a $m \geq k^n$. Comme $\|a_i\| \leq a_i \leq k-1$ et $\|k^i\| = \|k\|^i$, on obtient la majoration :

$$\|m\| \leq (k-1) \sum_{i=0}^n k^{i\alpha} = \frac{k-1}{k^\alpha - 1} (k^{(n+1)\alpha} - 1) \leq \frac{k^\alpha(k-1)}{k^\alpha - 1} k^{n\alpha} \leq C m^\alpha,$$

où $C = k^\alpha(k-1)/(k^\alpha - 1)$ est indépendant de m .

On peut appliquer cette inégalité à m^m , ce qui nous donne $\|m\|^m \leq C m^{m\alpha}$ et, prenant la racine m -ième de cette inégalité et passant à la limite, l'inégalité $\|m\| \leq m^\alpha$. On a donc :

$$\frac{\log \|m\|}{\log m} \leq \frac{\log \|k\|}{\log k}$$

quel que soit $m \in \mathbb{N}$.

Par symétrie, on en déduit le fait que si $\|m\| > 1$, alors cette inégalité est une égalité.

Dans le cas général, il existe $n \in \mathbb{N}$ tel que l'on ait $\|k^n m\| > 1$, ce qui montre que l'on a égalité quel que soit $m \in \mathbb{N}$ puis, en utilisant la multiplicativité de la norme et le fait que $\|-1\| = 1$, que $\|x\| = |x|_\infty^\alpha$ quel que soit $x \in \mathbb{Q}$.

On a donc montré que s'il existe $k \in \mathbb{N}$ tel que $\|k\| > 1$, alors $\|\cdot\|$ est équivalente à la norme usuelle.

Dans le cas contraire, on a $\|l\| \leq 1$ pour tout nombre premier l . Comme on a supposé $\|\cdot\|$ non triviale, il existe au moins un nombre premier p tel que $\|p\| < 1$. Si il en existe un autre q , alors quel que soit $n \in \mathbb{N}$, on peut, d'après le théorème de Bézout, trouver $u_n, v_n \in \mathbb{Z}$ tels que l'on ait $u_n p^n + v_n q^n = 1$. On obtient donc :

$$1 = \|1\| = \|u_n p^n + v_n q^n\| \leq \|u_n\| \cdot \|p\|^n + \|v_n\| \cdot \|q\|^n \leq \|p\|^n + \|q\|^n,$$

ce qui est impossible pour n assez grand.

Il existe donc un et un seul nombre premier p tel que $\|p\| < 1$ et $\|\cdot\|$ est équivalente à la norme p -adique. Ceci termine la démonstration. \square

Remarque. Finalement, \mathbb{Q}_p est le complété de \mathbb{Q} pour la norme p -adique et \mathbb{Z}_p est l'anneau de ses entiers. Son idéal maximal est $p\mathbb{Z}_p$ car $|\mathbb{Q}_p^*|_p = |\mathbb{Q}^*|_p = p^{\mathbb{Z}}$ et donc, si $|x|_p < 1$, alors $|x|_p \leq p^{-1}$. De plus, on a que l'espace \mathbb{Q}_p est un corps ultramétrique complet et qu'une série converge dans \mathbb{Q}_p si et seulement si son terme général tend vers 0.

1.3 Equations p -adiques.

Lemme 1.3.1. *Soit $\cdots \rightarrow X_n \rightarrow X_{n-1} \rightarrow \cdots \rightarrow X_1$ un système projectif, et soit $X = \varprojlim X_n$ sa limite projective. Si les x_n sont finis et non vides, X est non vide.*

Preuve :

Le fait que X soit $\neq \emptyset$ est clair si les $X_n \rightarrow x_{n-1}$ sont surjectifs; on va se ramener à ce cas. Pour cela, notons $X_{n,p}$ l'image de X_{n+p} dans X_n ; pour n fixe, les $X_{n,p}$ forment une famille décroissante d'ensembles finis non vides; il en résulte que cette famille est stationnaire, i.e. que $X_{n,p}$ est indépendant de p pour p assez grand. Soit Y_n cette valeur limite des $X_{n,p}$. On vérifie immédiatement que $X_n \rightarrow X_{n-1}$ applique Y_n sur Y_{n-1} ; comme les Y_n sont non vides, on a $\varprojlim Y_n \neq \emptyset$; d'où *a fortiori*, $\varprojlim X_n \neq \emptyset$. \square

Notation. Si $f \in \mathbb{Z}_p[X_1, \dots, X_m]$ est un polynôme à coefficients dans \mathbb{Z}_p , et si n est un entier ≥ 1 , on note f_n le polynôme à coefficients dans A_n déduit de f par réduction (mod p^n).

Proposition 1.3.2. *Soient $f^{(i)} \in \mathbb{Z}_p[X_1, \dots, X_m]$ des polynômes à coefficients entiers p -adiques. Il y a équivalence entre :*

- i) Les $f^{(i)}$ ont un zéro commun dans $(\mathbb{Z}_p)^m$.*
- ii) Pour tout $n \geq 1$, les polynômes $f_n^{(i)}$ ont un zéro commun dans $(A_n)^m$.*

Preuve :

Soit X (resp. X_n) l'ensemble des zéros communs aux $f^{(i)}$ (resp. aux $f_n^{(i)}$). Les X_n sont finis, et l'on a $X = \varprojlim X_n$; d'après le lemme précédent, X est non vide si et seulement si les X_n sont non vides; d'où la proposition. \square

Définition 1.3.3. *Un point $x = (x_1, \dots, x_m)$ de $(\mathbb{Z}_p)^m$ est dit **primitif** si l'un des x_i est inversible, c'est-à-dire si les x_i ne sont pas tous divisibles par p ; on définit de manière analogue les éléments primitifs de $(A_n)^m$.*

Proposition 1.3.4. *Soient $f^{(i)} \in \mathbb{Z}_p[X_1, \dots, X_m]$ des polynômes homogènes à coefficients entiers p -adiques. Il y a équivalence entre :*

- i) Les $f^{(i)}$ ont un zéro commun non trivial dans $(\mathbb{Q}_p)^m$.*
- ii) Les $f^{(i)}$ ont un zéro commun primitif dans $(\mathbb{Z}_p)^m$.*
- iii) Pour tout $n \geq 1$, les $f_n^{(i)}$ ont un zéro commun primitif dans $(A_n)^m$.*

Preuve :

L'implication ii) \Rightarrow i) est évidente. Inversement, si $x = (x_1, \dots, x_m)$ est un zéro commun non trivial des $f^{(i)}$, posons :

$$h = \inf(v_p(x_1), \dots, v_p(x_m)), \text{ et } y = p^{-h}x.$$

Il est clair que y est un élément primitif de $(\mathbb{Z}_p)^m$, et que c'est un zéro commun des $f^{(i)}$; on a donc bien i) \Rightarrow ii).

En ce qui concerne l'équivalence de ii) et iii), elle résulte du lemme qui précède. \square

1.4 Les carrés de \mathbb{Q}_p .

Lemme 1.4.1. *Pour tout $n \in \mathbb{N}^*$, on a $v_p(n!) \leq \frac{n}{p-1}$.*

Preuve :

Soit $n \in \mathbb{N}^*$, il existe un plus grand entier k tel que $kp \leq n$. Le nombre d'entiers inférieurs ou égaux à n divisibles par p est alors $k = \lfloor \frac{n}{p} \rfloor$. Plus généralement, le nombre d'entiers inférieurs ou égaux à n divisibles par p^k est $\lfloor \frac{n}{p^k} \rfloor$. Ainsi, on a :

$$v_p(n!) = \sum_{k \in \mathbb{N}^*} \lfloor \frac{n}{p^k} \rfloor \leq \sum_{k \in \mathbb{N}^*} \frac{n}{p^k} = \frac{n}{p-1}.$$

□

Lemme 1.4.2. Soient p un nombre premier impair et $N \in \mathbb{N}$. On a :

$$\forall n \in \mathbb{N}^*, v_p \left(\frac{(p^{N+1})^n}{n!} \right) \geq n \frac{p-2}{p-1}.$$

Preuve :

Soit $n \in \mathbb{N}^*$. On a :

$$\begin{aligned} v_p \left(\frac{(p^{N+1})^n}{n!} \right) &= v_p((p^{N+1})^n) - v_p(n!) \\ &= n(N+1) - v_p(n!) \\ &\geq n - v_p(n!) \\ &\geq n - \frac{n}{p-1} \\ &\geq n \frac{p-2}{p-1} \end{aligned}$$

□

Remarque. Si $z \in \mathbb{Z}_p$, on a donc :

$$\lim_{n \rightarrow +\infty} v_p \left(\frac{(p^{N+1}z)^n}{n!} \right) = +\infty.$$

Lemme 1.4.3. Soit $N \in \mathbb{N}$. On a :

$$\forall n \in \mathbb{N}^*, v_2 \left(\frac{(2^{N+2})^n}{n!} \right) \geq n.$$

Preuve :

Soit $n \in \mathbb{N}^*$. On a :

$$\begin{aligned} v_2 \left(\frac{(2^{N+2})^n}{n!} \right) &= v_2((2^{N+2})^n) - v_2(n!) \\ &= n(N+2) - v_2(n!) \\ &\geq 2n - v_2(n!) \\ &\geq 2n - n \\ &\geq n \end{aligned}$$

□

Corollaire 1.4.4. (Structure de \mathbb{Q}_p^*) Soit p un nombre premier impair. On a :

$$\begin{aligned} \mathbb{Q}_p^* &\cong p^{\mathbb{Z}} \times \mathbb{Z}_p^* \cong p^{\mathbb{Z}} \times \mathbb{F}_p^* \times \mathbb{Z}_p. \\ \mathbb{Q}_2^* &\cong 2^{\mathbb{Z}} \times \mathbb{Z}_2^* \cong 2^{\mathbb{Z}} \times \{\pm 1\} \times \mathbb{Z}_2. \end{aligned}$$

(Si on considère l'application suivante :

$$\begin{aligned} \theta : p^{\mathbb{Z}} \times \mathbb{Z}_p^* &\rightarrow \mathbb{Q}_p^* \\ (p^n, u) &\mapsto p^n u \end{aligned}$$

L'application θ est trivialement un homomorphisme. D'après la proposition 1.2.2, il est surjectif : en effet, soit $x \in \mathbb{Q}_p^*$, alors il existe $n \in \mathbb{Z}$ et $u \in \mathbb{Z}_p^*$ uniques tels que $x = p^n u$, ainsi $\theta(p^n, u) = x$. C'est également une injection : soit x dans son noyau, alors $\theta(p^n, u) = 1$, c'est-à-dire $n = 0$ et $u = 1$. donc on a

$$\mathbb{Q}_p^* \cong p^{\mathbb{Z}} \times \mathbb{Z}_p^*.)$$

Corollaire 1.4.5. (Structure de \mathbb{Q}_p^{*2}) Soit p un nombre premier impair. On a :

$$\begin{aligned} \mathbb{Q}_p^{*2} &\cong p^{2\mathbb{Z}} \times \mathbb{F}_p^{*2} \times \mathbb{Z}_p. \\ \mathbb{Q}_2^{*2} &\cong 2^{2\mathbb{Z}} \times 2\mathbb{Z}_2. \end{aligned}$$

Preuve :

D'après le corollaire qui précède, $\mathbb{Q}_p^* \cong p^{\mathbb{Z}} \times \mathbb{F}_p^* \times \mathbb{Z}_p$ d'où $\mathbb{Q}_p^{*2} \cong p^{2\mathbb{Z}} \times \mathbb{F}_p^{*2} \times 2\mathbb{Z}_p \cong p^{2\mathbb{Z}} \times \mathbb{F}_p^{*2} \times \mathbb{Z}_p$ car 2 est inversible dans \mathbb{Z}_p .

De même, $\mathbb{Q}_2^{*2} \cong 2^{2\mathbb{Z}} \times \{1\} \times 2\mathbb{Z}_2 \cong 2^{2\mathbb{Z}} \times 2\mathbb{Z}_2$. \square

Remarque. Les formules du corollaire précédent prouvent que \mathbb{Q}_p^{*2} est un sous-groupe ouvert de \mathbb{Q}_p^* . De plus on a le théorème suivant :

Théorème 1.4.6. Soit p un nombre premier. Un entier p -adique est un carré si et seulement si c'est un carré modulo p (resp. modulo 8), pour $p \neq 2$ (resp. $p = 2$).

Corollaire 1.4.7. Soit p un nombre premier impair. On a :

$$\begin{aligned} \mathbb{Q}_p^*/\mathbb{Q}_p^{*2} &\cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}. \\ \mathbb{Q}_2^*/\mathbb{Q}_2^{*2} &\cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}. \end{aligned}$$

Preuve :

Les deux corollaire précédents nous donnent :

$$\begin{aligned} \mathbb{Q}_p^*/\mathbb{Q}_p^{*2} &\cong p^{\mathbb{Z}}/p^{2\mathbb{Z}} \times \mathbb{F}_p^*/\mathbb{F}_p^{*2}. \\ \mathbb{Q}_2^*/\mathbb{Q}_2^{*2} &\cong \mathbb{Z}/2^{2\mathbb{Z}} \times \{\pm 1\} \times \mathbb{Z}_2/2\mathbb{Z}_2. \end{aligned}$$

Or on sait que pour p premier impair, $\mathbb{F}_p^*/\mathbb{F}_p^{*2} \cong \{\pm 1\}$. La proposition 1.1.5 donne $\mathbb{Z}_2/2\mathbb{Z}_2 \cong \mathbb{Z}/2\mathbb{Z}$. Enfin, $p^{\mathbb{Z}}/p^{2\mathbb{Z}} \cong \mathbb{Z}/2\mathbb{Z}$, pour tout nombre premier. D'où le corollaire. \square

Remarque. En d'autres termes, on peut dire que :

- 1) Si $p \neq 2$, le groupe $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ est un groupe de type $(2, 2)$; il admet $\{1, p, u, up\}$ où u est une unité p -adique (et u n'est pas un carré modulo p).
- 2) Si $p = 2$, le groupe $\mathbb{Q}_2^*/\mathbb{Q}_2^{*2}$ est de type $(2, 2, 2)$; il admet pour représentants $\{\pm 1, \pm 5, \pm 2, \pm 10\}$.

Chapitre 2

Les symboles de Legendre et Hilbert

2.1 Le symbole de Legendre.

2.1.1 Les carrés de \mathbb{F}_q .

On pose $\mathbb{F}_q^2 = \{x \in \mathbb{F}_q \mid \exists y \in \mathbb{F}_q ; x = y^2\}$ et $\mathbb{F}_q^{*2} = \mathbb{F}_q^2 \cap \mathbb{F}_q^*$.

Remarque. Par Frobenius, dans un corps fini de caractéristique 2, tout élément est un carré.

Dans le cas où la caractéristique est supérieure à 2, soit \mathbb{F}_q un corps fini de cardinal $q = p^n$ (où p est un nombre premier impair). On a le résultat suivant :

Théorème 2.1.1. *Les carrés de \mathbb{F}_q^* forment un sous-groupe d'indice 2 de \mathbb{F}_q^* . Ce sous-groupe est le noyau de l'homomorphisme $x \rightarrow x^{\frac{(q-1)}{2}}$, à valeurs dans $\{\pm 1\}$.*

En d'autres termes, on a les deux suites exactes fondamentales suivantes :

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathbb{F}_q^* \xrightarrow{u} \mathbb{F}_q^{*2} \longrightarrow 1 \quad (2.1)$$

$$1 \longrightarrow \mathbb{F}_q^{*2} \longrightarrow \mathbb{F}_q^* \xrightarrow{v} \{\pm 1\} \longrightarrow 1 \quad (2.2)$$

avec $u(x) = x^2$ et $v(x) = x^{\frac{(q-1)}{2}}$.

Remarques. 1) La première suite donne le nombre de carrés non nuls : $\frac{(q-1)}{2}$.
2) La seconde suite donne un critère pour que x , différent de 0, soit un carré, à savoir :

$$x^{\frac{(q-1)}{2}} = 1.$$

Preuve du théorème :

• Soit $x \in \mathbb{F}_q^*$. On a $x^{q-1} = 1$, donc $(x^{\frac{q-1}{2}})^2 = 1$, donc $x = \pm 1$. Cela montre que v est à valeurs dans $\{\pm 1\}$.

- Si x est dans \mathbb{F}_q^{*2} , on a $x = y^2$, avec $y \in \mathbb{F}_q^*$, donc $x^{\frac{q-1}{2}} = y^{q-1} = 1$. Cela montre que $\mathbb{F}_q^{*2} \subset \text{Kerv}$. Comme l'équation $x^{\frac{q-1}{2}} = 1$ admet au plus $\frac{q-1}{2}$ racines, on a $\mathbb{F}_q^{*2} = \text{Kerv}$.
- Il en résulte que $|\text{Im}v| = (q-1)/|\text{Kerv}| = 2$, de sorte que v est surjective. \square

2.1.2 Définition et propriétés.

Définition 2.1.2. Soit p un nombre premier différent de 2 et soit $x \in \mathbb{F}_p^*$. Le symbole de Legendre de x , noté $\left(\frac{x}{p}\right)$, est défini par la formule suivante :

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}}.$$

Remarque. Le symbole de Legendre est à valeurs dans $\{\pm 1\}$. On a :

$$\left(\frac{x}{p}\right) = \begin{cases} 1 & \text{si } x \text{ est un carré modulo } p; \\ -1 & \text{sinon.} \end{cases}$$

On étend le symbole de Legendre à \mathbb{Z} tout entier en posant $\left(\frac{x}{p}\right) = 0$ si p divise x et $\left(\frac{x}{p}\right) = \left(\frac{\bar{x}}{p}\right)$ si p ne divise pas x ; où \bar{x} est l'image de x dans \mathbb{F}_p^* .

Nous allons admettre la proposition suivante :

Proposition 2.1.3. *i) $\forall x, y \in \mathbb{Z}^2$, $\left(\frac{x}{p}\right)\left(\frac{y}{p}\right) = \left(\frac{xy}{p}\right)$,
ii) $\forall y \in \mathbb{F}_p^*$, $\left(\frac{y}{p}\right) = \left(\frac{y^{-1}}{p}\right)$.*

2.1.3 La loi de réciprocité quadratique.

Nous admettons le résultat suivant :

Théorème 2.1.4. (Loi de réciprocité quadratique) Soient n et p deux nombres premiers impairs distincts. On a la formule :

$$\left(\frac{n}{p}\right) = (-1)^{\frac{(n-1)(p-1)}{4}} \left(\frac{p}{n}\right),$$

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

Remarque. On peut démontrer cette loi en partant le lemme de Gauss.

Soit p un nombre premier $\neq 2$, et soit S une partie de \mathbb{F}_p^* telle que \mathbb{F}_p^* soit la réunion disjointe de S et de $-S$; dans la suite, on prend $S = \{1, \dots, \frac{p-1}{2}\}$. Si $s \in S$ et $a \in \mathbb{F}_p^*$, on peut écrire as sous la forme

$$as = e_s(a)s_a \text{ avec } e_s(a) = \pm 1 \text{ et } s_a \in S.$$

Lemme 2.1.5. (Gauss) On a :

$$\left(\frac{a}{p}\right) = \prod_{s \in S} e_s(a).$$

Preuve :

On remarque d'abord que, si s et s' sont deux éléments distincts de S , on a $s_a \neq s'_a$ (car sinon, on aurait $s = \pm s'$, contrairement au choix de S). On en conclut que $s \mapsto s_a$ est une bijection de S sur lui-même. Faisons alors le produit des égalités $as = e_s(a)s_a$. On obtient :

$$a^{\frac{p-1}{2}} \prod_{s \in S} s = \left(\prod_{s \in S} e_s(a) \right) \prod_{s \in S} s_a = \left(\prod_{s \in S} e_s(a) \right) \prod_{s \in S} s$$

d'où

$$a^{\frac{p-1}{2}} = \prod_{s \in S} e_s(a),$$

ce qui démontre le lemme, puisque $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}}$ dans \mathbb{F}_p . \square

Traduction du signe $(-)$ en termes de congruences modulo 4 :

On écrit : $n\mathbf{R}p$ si n est un carré (mod p),
c'est-à-dire si n est "reste quadratique" (mod p),
 $n\mathbf{N}p$ sinon.

On a donc : $n\mathbf{R}p \Leftrightarrow p\mathbf{R}n$ si p ou $n \equiv 1 \pmod{4}$,

$n\mathbf{R}p \Leftrightarrow p\mathbf{N}n$ si p et $n \equiv -1 \pmod{4}$.

2.2 Le symbole de Hilbert.

On rappelle que V désigne l'ensemble des nombres premiers union $\{\infty\}$, et on pose $\mathbb{Q}_\infty = \mathbb{R}$.

2.2.1 Définition et propriétés.

Définition 2.2.1. Soient $a, b \in \mathbb{Q}_v^*$. On pose :

$(a, b)_v = 1$ si $z^2 - ax^2 - by^2 = 0$ a une solution différente de $(0, 0, 0)$
dans \mathbb{Q}_v^3

$(a, b)_v = -1$ sinon.

Le nombre $(a, b)_v = \pm 1$ s'appelle le **symbole de Hilbert** de a et b , relativement à \mathbb{Q}_v .

Remarque. Il est clair que $(a, b)_v$ ne change pas lorsqu'on multiplie a et b par des carrés ; le symbole de Hilbert définit donc une application de $\mathbb{Q}_v^*/\mathbb{Q}_v^{*2} \times \mathbb{Q}_v^*/\mathbb{Q}_v^{*2}$ dans $\{\pm 1\}$.

Soient Ω une clôture algébrique de \mathbb{Q}_v et ω une racine carrée de b dans Ω (où $b \in \mathbb{Q}_v$) ; on a alors :

$$\mathbb{Q}_v[\sqrt{b}] = \{\alpha + \omega\beta \mid \alpha, \beta \in \mathbb{Q}_v\} \subset \Omega$$

et on définit la norme N sur $\mathbb{Q}_v[\sqrt{b}]$ par :

$$N(\alpha + \omega\beta) = \alpha^2 - b\beta^2.$$

Elle vérifie $N(1) = 1$ et $N(xy) = N(x)N(y)$ pour tout $x, y \in \mathbb{Q}_v[\sqrt{b}]$. Cette formule munit $N(\mathbb{Q}_v[\sqrt{b}]) = \{N(x) | x \in \mathbb{Q}_v[\sqrt{b}]\}$ d'une structure de groupe, pour $b \notin \mathbb{Q}_p^2$. Ce groupe est appelé le groupe des normes des éléments de $\mathbb{Q}_v[\sqrt{b}]$.

Proposition 2.2.2. *Soient a et b deux éléments de \mathbb{Q}_v^* . Alors*

$$(a, b)_v = 1 \iff a \in N(\mathbb{Q}_v[\sqrt{b}]^*)$$

Preuve :

Si b est un carré, on a $b = c^2$ avec $c \in \mathbb{Q}_v^*$, et ainsi $\mathbb{Q}_v[\sqrt{b}] = \mathbb{Q}_v$ et $N(\mathbb{Q}_v[\sqrt{b}]^*) = \mathbb{Q}_v^*$. Par ailleurs, l'équation $z^2 - ax^2 - by^2 = 0$ admet $(0, 1, c)$ comme solution et donc $(a, b)_v = 1$, d'où la proposition dans ce cas. Supposons que b n'est pas un carré.

(\Rightarrow) Supposons que $(a, b)_v = 1$. Par définition, il existe $(x_0, y_0, z_0) \in \mathbb{Q}_v^3 \setminus \{0\}$ tel que $z_0^2 - ax_0^2 - by_0^2 = 0$. On a $x_0 \neq 0$ car sinon on aurait $z_0^2 - by_0^2 = 0$ (car sinon $z_0 = 0$, ce qui n'est pas) et $b = \frac{z_0^2}{y_0^2}$ ce qui n'est pas possible car b n'est pas un carré. On peut donc écrire $a = \frac{z_0^2}{x_0^2} - b \frac{y_0^2}{x_0^2} = N\left(\frac{z_0}{x_0} - \omega \frac{y_0}{x_0}\right)$ où N et ω sont définis comme précédemment. Ainsi $a \in N(\mathbb{Q}_v[\sqrt{b}]^*)$.

(\Leftarrow) Soit $a \in N(\mathbb{Q}_v[\sqrt{b}]^*)$. Il existe α et β dans \mathbb{Q}_v^* tels que $a = \alpha^2 - b\beta^2$. Le triplet $(1, \beta, \alpha)$ est une solution non triviale de l'équation $z^2 - ax^2 - by^2 = 0$ et $(a, b)_v = 1$. \square

Proposition 2.2.3. *Le symbole de Hilbert satisfait aux formules :*

- i) $(a, b)_v = (b, a)_v$ et $(a, c^2)_v = 1$;
- ii) $(a, -a)_v = 1$ et $(a, 1 - a)_v = 1$;
- iii) $(a, b)_v = 1 \Rightarrow (aa', b)_v = (a', b)_v$;
- iv) $(a, b)_v = (a, -ab)_v = (a, (1 - a)b)_v$;
- v) $(a, a)_v = (a, -1)_v$.

(Dans ces formules, a, a', b, c désignent des éléments de \mathbb{Q}_v^* ; on suppose $a \neq 1$ lorsque la formule contient le terme $1 - a$.)

Preuve :

i) L'équation $z^2 - ax^2 - c^2y^2 = 0$ admet pour solution $(0, 1, c)$ donc on a immédiatement $(a, c^2)_v = 1$.

Soit (x_0, y_0, z_0) une solution non triviale de $z^2 - ax^2 - by^2 = 0$ alors (y_0, x_0, z_0) est une solution non triviale de $z^2 - bx^2 - ay^2 = 0$ et réciproquement, d'où $(a, b)_v = (b, a)_v$.

ii) On a que $z^2 - ax^2 + ay^2 = 0$ admet $(1, 1, 0)$ pour solution, d'où $(a, -a)_v = 1$. On a aussi que $z^2 - ax^2 - (1 - a)y^2 = 0$ admet $(1, 1, 1)$ pour solution, d'où $(a, 1 - a)_v = 1$.

$$(a', b)_v = 1 \iff a' \in N(\mathbb{Q}_v[\sqrt{b}]^*)$$

iii) Si $(a, b)_v = 1$ alors on a :

$$\iff aa' \in N(\mathbb{Q}_v[\sqrt{b}]^*)$$

$$\iff (aa', b)_v = 1$$

Soit encore $(aa', b)_v = (a', b)_v$. iv) Puisque $(a, -a)_v = 1$ on a $(a, b)_v = (a, -ab)_v$ d'après iii). On raisonne de même en utilisant $(a, 1 - a)_v = 1$. v) On a $(a, a)_v = (a, -a^2)_v$ (d'après iv)), mais $(a, -a^2)_v = (a, -1)_v$. \square

2.2.2 La bilinéarité du symbole de Hilbert.

Nous allons admettre les résultats suivants :

Théorème 2.2.4. Soient a et b deux éléments de \mathbb{Q}_v^* .

Si $v = \infty$, $(a, b)_\infty$ vaut 1 si $a > 0$ ou $b > 0$ et vaut -1 sinon.

Si $v = p \neq \infty$, il existe deux entiers relatifs α et β , et deux éléments u_a et u_b dans \mathbb{Z}_p^* uniques tel que $a = p^\alpha u_a$ et $b = p^\beta u_b$.

$$\begin{aligned} (a, b)_p &= (-1)^{\alpha\beta\frac{p-1}{2}} \left(\frac{u_a}{p}\right)^\beta \left(\frac{u_b}{p}\right)^\alpha & \text{si } p \neq 2, \\ &= (-1)^{\frac{u_a-1}{2}\frac{u_b-1}{2} + \alpha\frac{u_b^2-1}{8} + \beta\frac{u_a^2-1}{8}} & \text{si } p = 2. \end{aligned}$$

Théorème 2.2.5. Le symbole de Hilbert est une forme bilinéaire non dégénérée sur le \mathbb{F}_2 -espace vectoriel $\mathbb{Q}_v^*/\mathbb{Q}_v^{*2}$ (pour les lois $+$ et \cdot).

La bilinéarité de $(a, b)_v$ correspond à la formule suivante :

$$(aa', b)_v = (a, b)_v (a', b)_v;$$

L'assertion " $(a, b)_v$ est non dégénérée" signifie que, si $b \in \mathbb{Q}_v^*$ est tel que $(a, b)_v = 1$ pour tout $a \in \mathbb{Q}_v^*$, on a $b \in \mathbb{Q}_v^{*2}$.

2.2.3 Propriétés globales.

Théorème 2.2.6. (Formule du produit) Si $a, b \in \mathbb{Q}^*$, on a $(a, b)_v = 1$ pour presque tout $v \in V$, et on a :

$$\prod_{v \in V} (a, b)_v = 1.$$

(L'expression "presque tout $v \in V$ " signifie "tous les éléments de V sauf un nombre fini".)

Preuve :

Puisque le symbole de Hilbert est bilinéaire, il suffit de prouver le théorème dans le cas où a et b sont égaux à -1 ou à un nombre premier.

Premier cas : $a = -1$ et $b = -1$.

On a $(-1, -1)_\infty = (-1, -1)_2 = -1$ et $(-1, -1)_p = 1$ pour tout $p \neq \infty$ et différent de 2 dans V .

En effet, on a $(-1, -1)_\infty = -1$, car l'équation $z^2 + x^2 + y^2 = 0$ n'admet pas de solution non triviale dans \mathbb{R}^3 . On a également $(-1, -1)_2 = -1$, car les carrés modulo 8 sont 0, 1 et 4 et de plus, $\forall k \in \mathbb{Z}$, on a :

$$\begin{aligned} (1 + 2k)^2 &= 1 + 4k + 4k^2 \\ &= 1 + 4 \underbrace{(k + k^2)}_{\text{pair}} \\ &\equiv 1 \pmod{8}. \end{aligned}$$

On en déduit que $z^2 + x^2 + y^2 \equiv 0 \pmod{8}$. Et comme les carrés de \mathbb{Q}_2^* correspondent aux carrés modulo 8, on a le résultat.

Pour tout p premier impair différent de l'infini, $(-1, -1)_p = 1$ découle du fait

que si on a $z^2 + x^2 + y^2 = 0$ alors $z^2 + x^2 = -y^2$. Si on prend $z = 1$ alors l'équation se ramène à $1 + x^2 = -y^2$. Si on note A l'ensemble des éléments de \mathbb{F}_p de la forme $1 + x^2$ et B ceux de la forme $-y^2$. On a immédiatement que A et B ont chacun $(p + 1)/2$ éléments; $A \cap B \neq \emptyset$, on en déduit que l'équation admet nécessairement une solution non triviale. D'où le théorème dans ce cas.

Deuxième cas : $a = l$ premier et $b = -1$.

Si $l = 2$, on a $(2, -1)_v = 1$ pour tout $v \in V$, d'où le théorème dans ce cas.

Si $l \neq 2$ on a $(l, -1)_v = 1$ pour tout $v \neq l$ impair dans V et $(l, -1)_l = (-1)^{\frac{l-1}{2}} = (l, -1)_2$. Le produit est bien égal à 1, d'où le théorème dans ce cas.

Troisième cas $a = l$ et $b = l'$ premiers.

Si $l = l'$, on a $(l, l)_v = (l, -1)_v$ pour tout V , et on est ramené au cas précédent. Si $l \neq l'$, deux cas se présentent :

Cas où $l' = 2$: on a $(l, 2)_v = 1$ pour tout $v \neq l$ impair dans V . De plus, $(l, 2)_2 = (-1)^{\frac{l^2-1}{8}}$ et $(l, 2)_l = \left(\frac{2}{l}\right) = (-1)^{\frac{l^2-1}{8}}$. D'où le théorème dans ce cas.

Cas où $l' \neq 2$: on a $(l, l')_v = 1$ pour tout $v \in V$ distinct de 2, l et l' . De plus, $(l, l')_2 = (-1)^{\frac{l-1}{2} \frac{l'-1}{2}}$, $(l, l')_l = \left(\frac{l'}{l}\right)$ et $(l, l')_{l'} = \left(\frac{l}{l'}\right)$ mais on a :

$$\begin{aligned} (l, l')_2 (l, l')_l (l, l')_{l'} &= (-1)^{\frac{l-1}{2} \frac{l'-1}{2}} \left(\frac{l'}{l}\right) \left(\frac{l}{l'}\right) \\ &= \left(\frac{l}{l'}\right) \left(\frac{l}{l'}\right) \\ &= 1 \end{aligned}$$

d'où le théorème dans ce cas. \square

Remarque. La formule du produit est essentiellement équivalente à la loi de réciprocité quadratique.

Théorème 2.2.7. Soit $(a_i)_{i \in I}$ une famille d'éléments de \mathbb{Q}^* et soit $(\varepsilon_{i,v})_{i \in I, v \in V}$ une famille de nombres égaux à ± 1 , avec $|I| < \infty$. Pour qu'il existe $x \in \mathbb{Q}^*$ tel que $(a_i, x)_v = \varepsilon_{i,v}$ pour tout $i \in I$ et tout $v \in V$, il faut et il suffit que les trois conditions suivantes soient satisfaites :

- i) Presque tous les $\varepsilon_{i,v}$ sont égaux à 1.
- ii) Pour tout $i \in I$, on a $\prod_{v \in V} \varepsilon_{i,v} = 1$.
- iii) Pour tout $v \in V$, il existe $x_v \in \mathbb{Q}_v^*$ tel que $(a_i, x_v)_v = \varepsilon_{i,v}$ pour tout $i \in I$.

Preuve :

La nécessité de i) et ii) résulte de la formule du produit ; celle de iii) est triviale (en prenant $x_v = x$).

La suffisance de ces conditions fait appel aux trois lemmes qui suivent :

Lemme 2.2.8. (Lemme chinois) Soient $a_1, \dots, a_n, m_1, \dots, m_n$ des entiers, les m_i étant premiers entre eux deux à deux. Il existe un entier a tel que $a \equiv a_i \pmod{m_i}$ pour tout i .

Preuve :

Soit m le produit des m_i . Le théorème de Bézout montre que l'homomorphisme

canonique

$$\mathbb{Z}/m\mathbb{Z} \rightarrow \prod_{i=1}^n \mathbb{Z}/m_i\mathbb{Z}$$

est un isomorphisme. Le lemme en résulte. \square

Lemme 2.2.9. (Théorème d'approximation) *Soit S une partie finie de V . L'image de \mathbb{Q} dans $\prod_{v \in S} \mathbb{Q}_v$ est dense dans ce produit (pour la topologie produit de celles des \mathbb{Q}_v).*

Preuve :

Quitte à agrandir S , on peut supposer que

$$S = \{\infty, p_1, \dots, p_n\}$$

où les p_i sont des nombres premiers distincts, et il s'agit de démontrer que \mathbb{Q} est dense dans $\mathbb{R} \times \mathbb{Q}_{p_1} \times \dots \times \mathbb{Q}_{p_n}$. Soit donc $(x_\infty, x_1, \dots, x_n)$ un point de ce produit, et montrons que ce point est adhérent à \mathbb{Q} ; quitte à faire une homothétie de rapport entier, on peut supposer que l'on a $x_i \in \mathbb{Z}_{p_i}$ pour $1 \leq i \leq n$; il s'agit alors de montrer que, pour tout $\varepsilon > 0$, et tout entier $N \geq 0$, il existe $x \in \mathbb{Q}$ tel que :

$$|x - x_\infty| \leq \varepsilon \text{ et } v_{p_i}(x - x_i) \geq N \text{ pour } i = 1, \dots, n.$$

D'après le "théorème chinois", appliqué aux $m_i = p_i^N$, il existe $x_0 \in \mathbb{Z}$ tel que $v_{p_i}(x_0 - x_i) \geq N$ pour tout i . Choisissons d'autre part un entier $q \geq 2$ qui soit premier à tous les p_i (par exemple un nombre premier). Il est facile de voir que les rationnels de la forme a/q^m , $a \in \mathbb{Z}$, $m \geq 0$, sont denses dans \mathbb{R} (cela provient simplement de ce que $q^m \rightarrow \infty$). On peut donc choisir un tel nombre $u = a/q^m$ tel que :

$$|x_0 - x_\infty + up_1^N \dots p_n^N| \leq \varepsilon.$$

Le nombre rationnel $x = x_0 + up_1^N \dots p_n^N$ répond alors à la question. \square

Lemme 2.2.10. (Théorème de Dirichlet) *Si a et m sont des entiers ≥ 1 premiers entre eux, il existe une infinité de nombres premiers p tels que $p \equiv a \pmod{m}$.*

Nous allons admettre ce résultat ici.

Revenons à présent au théorème 2.2.7, et soit $(\varepsilon_{i,v})$ une famille de nombres égaux à ± 1 , et satisfaisant aux conditions i), ii) et iii). Quitte à multiplier les a_i par le carré d'un entier, on peut supposer que tous les a_i sont entiers. Soit S le sous-ensemble de V formé de $\infty, 2$, et des facteurs premiers des a_i ; soit T l'ensemble des éléments $v \in V$ tels qu'il existe $i \in I$ avec $\varepsilon_{i,v} = -1$; ces deux ensembles sont finis. Distinguons deux cas :

1) Le cas où on a $S \cap T = \emptyset$.

Posons

$$a = \prod_{l \in T \text{ et } l \neq \infty} l \text{ et } m = 8 \prod_{l \in S \text{ et } l \neq 2, \infty} l.$$

Puisque $S \cap T = \emptyset$, les entiers a et m sont premiers entre eux, et, d'après le théorème de Dirichlet, il existe un nombre premier $p \equiv a \pmod{m}$ tel que

$p \notin S \cup T$. Nous allons voir que le nombre $x = ap$ répond à la question, autrement dit que $(a_i, x)_v = \varepsilon_{i,v}$, pour tout $i \in I$ et tout $v \in V$.

Si $v \in S$, on a $\varepsilon_{i,v} = 1$ puisque $S \cap T = \emptyset$, et il faut donc vérifier que $(a_i, x)_v = 1$. Si $v = \infty$, cela provient de ce que x est > 0 ; si v est un nombre premier l , on a $x \equiv a^2 \pmod{m}$, d'où $x \equiv a^2 \pmod{8}$ pour $l = 2$, $x \equiv a^2 \pmod{l}$ pour $l \neq 2$; comme x et a sont des unités l -adiques, cela montre que x est un carré dans \mathbb{Q}_l^* , et l'on a bien $(a_i, x)_v = 1$.

Si $v = l$ n'appartient pas à S , a_i est une unité l -adique. Comme $l \neq 2$, on a :

$$(a_i, b)_l = \left(\frac{a_i}{l}\right)^{v_l(b)} \quad \text{pour tout } b \in \mathbb{Q}_l^*,$$

en se référant au théorème 2.2.4. Si $l \notin T \cup \{p\}$, x est une unité l -adique, d'où $v_l(x) = 0$, et la formule ci-dessus montre que $(a_i, x)_l = 1$; d'autre part, on a $\varepsilon_{i,l} = 1$, puisque $l \notin T$. Si $l \in T$, on a $v_l(x) = 1$, d'autre part la condition iii) montre qu'il existe $x_l \in \mathbb{Q}_l^*$ tel que $(a_i, x_l)_l = \varepsilon_{i,l}$ pour tout $i \in I$; comme l'un des $\varepsilon_{i,l}$ est égal à -1 (puisque l appartient à T), on a $v_l(x_l) \equiv 1 \pmod{2}$, d'où :

$$(a_i, x)_l = \left(\frac{a_i}{l}\right) = (a_i, x_l)_l = \varepsilon_{i,l} \quad \text{pour tout } i \in I.$$

Reste enfin le cas $l = p$, que l'on ramène aux autres cas grâce à la formule du produit :

$$(a_i, x)_p = \prod_{v \neq p} (a_i, x)_v = \prod_{v \neq p} \varepsilon_{i,v} = \varepsilon_{i,p}.$$

Cela achève la démonstration du théorème 2.2.7 dans le cas $S \cap T = \emptyset$.

2) Cas général.

On sait que les carrés de \mathbb{Q}_v^* forment un sous-groupe de \mathbb{Q}_v^* . D'après le théorème d'approximation, il existe donc $x' \in \mathbb{Q}^*$ tel que x'/x_v soit un carré dans \mathbb{Q}_v^* pour tout $v \in S$. On a en particulier :

$$(a_i, x')_v = (a_i, x_v)_v = \varepsilon_{i,v} \quad \text{pour tout } v \in S.$$

Si l'on pose $\eta_{i,v} = \varepsilon_{i,v}(a_i, x')_v$, la famille $\eta_{i,v}$ vérifie les conditions i), ii) et iii), et de plus $\eta_{i,v} = 1$ si $v \in S$. D'après le 1) ci-dessus, il existe donc $y \in \mathbb{Q}^*$ tel que

$$(a_i, y)_v = \eta_{i,v} \quad \forall i \in I \text{ et } v \in V.$$

Si l'on pose $x = yx'$, il est clair que x répond à la question. \square

Chapitre 3

Formes quadratiques.

3.1 Définitions.

Rappelons pour commencer la notion générale de *forme quadratique* :

Définition. Un **module** sur A (ou A -module) est un ensemble V muni d'une loi de composition interne $+$, et d'une loi de composition externe (par A) \times telle que :

- i) V soit un groupe commutatif;
- ii) $\forall \lambda \in A, \forall x, y \in V, \lambda \times (x + y) = \lambda \times x + \lambda \times y$;
- iii) $\forall \lambda, \mu \in A, \forall x \in V, (\lambda + \mu) \times x = \lambda \times x + \mu \times x$;
- iv) $\forall \lambda, \mu \in A, \forall x, (\lambda\mu) \times x = \lambda \times (\mu \times x)$;
- v) $\forall x \in V, 1 \times x = x$.

Définition 3.1.1. Soit V un module sur un anneau commutatif A . Une application $Q : V \rightarrow A$ est appelée **forme quadratique** sur V si :

- 1) On a $Q(ax) = a^2Q(x)$ pour $a \in A$ et $x \in V$.
- 2) L'application $(x, y) \mapsto Q(x + y) - Q(x) - Q(y)$ est une forme bilinéaire.

Un tel couple (V, Q) est appelé un *module quadratique*.

Dans tout ce chapitre, nous nous limiterons au cas où A est un corps K de caractéristique différente de 2; le A -module V est alors un K -espace vectoriel; nous le supposons de dimension finie. On posera :

$$x.y = \frac{1}{2}[Q(x + y) - Q(x) - Q(y)],$$

ce qui a un sens puisque la caractéristique de K est différente de 2.

L'application $(x, y) \mapsto x.y$ est une *forme bilinéaire symétrique* sur V ; on l'appelle le **produit scalaire** associé à Q .

On a $Q(x) = x.x$. Cela établit une correspondance bijective entre *formes quadratiques* et *formes bilinéaires symétriques* (il n'en serait plus de même en caractéristique 2).

Si (V, Q) et (V', Q') sont deux modules quadratiques, on appelle **morphisme** (ou morphisme métrique) de (V, Q) dans (V', Q') toute application linéaire $f : V \rightarrow V'$ telle que $Q' \circ f = Q$; on a alors $f(x).f(y) = x.y$ pour $x, y \in V$.

Soit $(e_i)_{1 \leq i \leq n}$ une base de V . On appelle **matrice de Q** par rapport à cette base, la matrice $A = (a_{ij})$, où $a_{ij} = e_i.e_j$; c'est une matrice symétrique.

Si $x = \sum x_i e_i$ est un élément de V , on a :

$$Q(x) = \sum_{i,j} a_{ij} x_i x_j,$$

ce qui montre que $Q(x)$ est une "forme quadratique" en x_1, \dots, x_n au sens usuel.

Si on modifie la base (e_i) au moyen d'une matrice inversible X , la matrice A' de Q par rapport à la nouvelle base est $X.A.{}^tX$ où tX désigne la transposée de X .

On a en particulier :

$$\det(A') = \det(A). \det(X)^2$$

ce qui montre que $\det(A)$ est déterminé à la multiplication près par un élément de K^{*2} ; on l'appelle le **discriminant** de Q , et on le note $\text{disc}(Q)$.

3.2 Orthogonalité.

Soit (V, Q) un module quadratique sur K . Deux éléments x et y de V sont dits **orthogonaux** si $x.y = 0$.

L'ensemble des éléments orthogonaux à une partie H de V est noté H^0 ; c'est un sous-espace vectoriel de V .

Si V_1 et V_2 sont deux sous-espaces vectoriels de V , on dit que V_1 et V_2 sont **orthogonaux** si $V_1 \subset V_2^0$, i.e. si $x \in V_1, y \in V_2$ entraîne $x.y = 0$.

L'orthogonal V^0 de V tout entier est appelé le **radical** (ou le noyau) de V , et noté $\text{rad}(V)$. Sa codimension s'appelle le **rang** de Q .

Si $V^0 = 0$, on dit que Q est **non dégénérée**; cela équivaut à dire que le discriminant de Q est différent de 0 (auquel cas on peut le considérer comme un élément du groupe K^*/K^{*2}).

Soit U un sous-espace vectoriel de V , et soit U^* le dual de U . Soit $q_U : V \rightarrow U^*$ l'application qui associe à tout x de V la forme linéaire $(y \in U \mapsto x.y)$. Le noyau de q_U est U^0 . En particulier, on voit que Q est non dégénérée si et seulement si $q_V : V \rightarrow V^*$ est un isomorphisme.

Définition 3.2.1. Soient U_1, \dots, U_m des sous-espaces vectoriels de V . On dit que V est **somme directe orthogonale** des U_i si ceux-ci sont deux à deux orthogonaux et si V en est la somme directe. On écrit alors :

$$V = U_1 \overset{\wedge}{\oplus} \dots \overset{\wedge}{\oplus} U_m.$$

Remarque. Si $x \in V$ a pour composante x_i dans U_i , on a

$$Q(x) = Q_1(x_1) + \cdots + Q_m(x_m)$$

où $Q_i = Q|_{U_i}$ désigne la restriction de Q à U_i . Inversement, si (U_i, Q_i) est une famille de modules quadratiques, la formule ci-dessus munit $V = \bigoplus U_i$ d'une forme quadratique Q , dite *somme directe* des Q_i , et l'on a $V = U_1 \hat{\oplus} \cdots \hat{\oplus} U_m$.

Proposition 3.2.2. *Si U est un supplémentaire de $\text{rad}(V)$ dans V , on a :*

$$V = U \hat{\oplus} \text{rad}(V).$$

Preuve :

C'est clair. \square

Proposition 3.2.3. *Supposons (V, Q) non dégénéré. Alors :*

i) Tout morphisme métrique de V dans un module quadratique (V', Q') est injectif.

ii) Pour tout sous-espace vectoriel U de V , on a :

$$U^{00} = U, \quad \dim U + \dim U^0 = \dim V$$

$$\text{rad}(U) = \text{rad}(U^0) = U \cap U^0.$$

Pour que U soit non dégénérée, il faut et il suffit que U^0 le soit, auquel cas $V = U \hat{\oplus} U^0$.

iii) si V est somme directe orthogonale de deux sous-espaces, ceux-ci sont non dégénérés, et chacun d'eux est l'orthogonal de l'autre.

Preuve :

Si $f : V \rightarrow V'$ est un morphisme métrique, et si $f(x) = 0$, on a $x.y = f(x).f(y) = 0$ pour tout $y \in V$; d'où $x = 0$ puisque (V, Q) est non dégénéré.

Si U est un sous-espace vectoriel de V , l'homomorphisme $q_U : V \rightarrow U^*$ défini plus haut est surjectif; en effet, il s'obtient en composant $q_V : V \rightarrow V^*$ avec la surjection canonique $V^* \rightarrow U^*$, et l'on a que q_V est bijectif. On a donc une suite exacte :

$$0 \rightarrow U^0 \rightarrow V \rightarrow U^* \rightarrow 0$$

d'où $\dim V = \dim U^* + \dim U^0 = \dim U + \dim U^0$.

Ceci montre que U et U^{00} sont de même dimension; comme U est contenu dans U^{00} , on a $U = U^{00}$; la formule $\text{rad}(U) = U \cap U^0$ est évidente; en l'appliquant à U^0 , et en tenant compte de ce que $U^{00} = U$, on en déduit que $\text{rad}(U^0) = \text{rad}(U)$, d'où en même temps la dernière assertion de *ii)*. Enfin, *iii)* est triviale. \square

3.3 Vecteurs isotropes.

Définition 3.3.1. *Un élément x d'un module quadratique (V, Q) est dit **isotrope** si l'on a $Q(x) = 0$. Un sous-espace U de V est dit **isotrope** si tous ses éléments sont isotropes.*

On a évidemment le résultat suivant :

$$U \text{ isotrope} \Leftrightarrow U \subset U^0 \Leftrightarrow Q|_U = 0.$$

Définition 3.3.2. On appelle **plan hyperbolique** tout module quadratique ayant une base formée de deux éléments isotropes x, y tels que $x.y \neq 0$.

Quitte à multiplier y par $1/x.y$, on peut supposer que $x.y = 1$. La matrice de la forme quadratique par rapport à x, y est alors simplement $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; son discriminant est -1 (en particulier, elle est non dégénérée).

Proposition 3.3.3. Soit x un élément isotrope différent de 0 d'un module quadratique non dégénéré (V, Q) . Il existe alors un sous-espace U de V qui contient x et qui est un plan hyperbolique.

Preuve :

Puisque V est non dégénéré, il existe $z \in V$ tel que $x.z = 1$. soit $y = 2z - (z.z)x$, sachant que x est isotrope, on a :

$$\begin{aligned} Q(y) &= (2z - Q(z)x).(2z - Q(z)x) \\ &= 4Q(z) + Q(z)^2Q(x) - 4Q(z)(x.z) \\ &= 0 \end{aligned}$$

En d'autres termes y est un élément isotrope. De plus :

$$y.x = (2z - Q(z)x).x = 2z.x - Q(z)Q(x) = 2$$

qui est non nul car la caractéristique est différente de 2. Et le sous-espace $U = Kx + Ky$ est alors un plan hyperbolique. \square

Corollaire 3.3.4. Si (V, Q) est non dégénéré et contient un élément isotrope non nul, on a $Q(V) = K$.

Autrement dit, pour tout $a \in K$, il existe $v \in V$ tel que $Q(v) = a$.

Preuve :

Compte tenu de la proposition précédente, il suffit ici de montrer le résultat quand V est un plan hyperbolique, de base x, y avec x, y isotropes et $x.y = 1$. Si $a \in K$, on a alors $a = Q(x + \frac{a}{2}y)$, d'où $Q(V) = K$. \square

3.4 Bases orthogonales.

Définition 3.4.1. Une base (e_1, \dots, e_n) d'un module quadratique (V, Q) est dite **orthogonale** si elle est formée d'éléments deux à deux orthogonaux, i.e. si $V = Ke_1 \hat{\oplus} \dots \hat{\oplus} Ke_n$.

Remarque. Il revient au même de dire que la matrice de Q par rapport à cette base est une matrice diagonale :

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

Si $x = \sum x_i e_i$, on a alors $Q(x) = a_1 x_1^2 + \cdots + a_n x_n^2$.

Théorème 3.4.2. Tout module quadratique (V, Q) possède une base orthogonale.

Preuve :

Nous allons démontrer ce résultat par récurrence sur $n = \dim V$.

Le cas $n = 0$ est trivial.

Si V est isotrope, toute base de V est orthogonale. Sinon, on choisit un élément $e_1 \in V$ tel que $e_1.e_1 \neq 0$. L'orthogonal H de e_1 est un hyperplan, et comme e_1 n'appartient pas à H , on a $V = Ke_1 \hat{\oplus} H$; vu l'hypothèse de récurrence, H possède une base orthogonale (e_2, \dots, e_n) ; il est clair que (e_1, e_2, \dots, e_n) est une base orthogonale de V . Ce qui prouve le théorème. \square

Définition 3.4.3. *Deux bases orthogonales*

$$\mathbf{e} = (e_1, \dots, e_n) \quad \text{et} \quad \mathbf{e}' = (e'_1, \dots, e'_n)$$

de V sont dites **contiguës** si elles ont un élément en commun (i.e. s'il existe i et j tels que $e_i = e'_j$).

Théorème 3.4.4. *Supposons V non dégénéré de dimension ≥ 3 , et soient $\mathbf{e} = (e_1, \dots, e_n)$, $\mathbf{e}' = (e'_1, \dots, e'_n)$ deux bases orthogonales de V . Il existe une suite finie $\mathbf{e}^{(0)}, \mathbf{e}^{(1)}, \dots, \mathbf{e}^{(m)}$ de bases orthogonales de V telle que $\mathbf{e}^{(0)} = \mathbf{e}$, $\mathbf{e}^{(m)} = \mathbf{e}'$, et que $\mathbf{e}^{(i)}$ soit contiguë à $\mathbf{e}^{(i+1)}$ pour $0 \leq i < m$.*

On dit que $\mathbf{e}^{(0)}, \dots, \mathbf{e}^{(m)}$ est une **chaîne** de bases orthogonales contiguës reliant \mathbf{e} à \mathbf{e}' .

Preuve :

Nous distinguons trois cas :

i) On a $(e_1.e_1)(e'_1.e'_1) - (e_1.e'_1)^2 \neq 0$.

Cela revient à dire que e_1 et e'_1 ne sont pas proportionnels et que le plan $P = Ke_1 + Ke'_1$ est non dégénéré. Il existe alors $e_2, e'_2 \in P$ tels que :

$$P = Ke_1 \hat{\oplus} Ke_2 \quad \text{et} \quad P = Ke'_1 \hat{\oplus} Ke'_2.$$

Soit H l'orthogonal de P ; comme P est non dégénéré, on a $V = H \hat{\oplus} P$. Soit (e''_3, \dots, e''_n) une base orthogonale de H . On peut alors relier \mathbf{e} à \mathbf{e}' au moyen de la chaîne :

$$\mathbf{e} \rightarrow (e_1, e_2, e''_3, \dots, e''_n) \rightarrow (e'_1, e'_2, e''_3, \dots, e''_n) \rightarrow \mathbf{e}'$$

d'où le théorème, dans ce cas.

ii) On a $(e_1.e_1)(e'_2.e'_2) - (e_1.e'_2)^2 \neq 0$.

On raisonne de la même manière, en remplaçant e'_1 par e'_2 .

iii) On a $(e_1.e_1)(e'_i.e'_i) - (e_1.e'_i)^2 = 0$ pour $i = 1, 2$.

On démontre d'abord le lemme suivant :

Lemme 3.4.5. *Il existe $x \in K$ tel que $e_x = e'_1 + xe'_2$ soit non isotrope, et engendre avec e_1 un plan non dégénéré.*

Preuve :

On a $e_x.e_x = e'_1.e'_1 + x^2(e'_2.e'_2)$; on doit donc prendre x^2 distinct de $-(e'_1.e'_1)/(e'_2.e'_2)$. D'autre part, pour que e_x engendre avec e_1 un plan non dégénéré, il faut et il suffit que $(e_1.e_1)(e_x.e_x) - (e_1.e_x)^2 \neq 0$. \square

Si l'on explicite en tenant compte de l'hypothèse iii), on trouve que le premier membre est $-2x(e_1.e'_1)(e_1.e'_2)$. Or l'hypothèse iii) entraîne $e_1.e'_i \neq 0$ pour $i = 1, 2$. On voit donc que e_x vérifie les conditions du lemme si et seulement si l'on a à la fois $x \neq 0$ et $x^2 \neq -(e'_1.e'_1)/(e'_2.e'_2)$. Cela élimine au plus trois valeurs de x ; si K a au moins quatre éléments, on peut donc trouver un tel x . Reste le cas où $K = \mathbb{F}_3$ (le cas $K = \mathbb{F}_2$ est exclu puisque $\text{caract}(K) \neq 2$). Mais dans ce cas, tout carré non égal à 1, et l'hypothèse iii) s'écrit $(e_1.e_1)/(e'_2.e'_2)$ est donc égal à 1, et, pour réaliser la condition $x^2 \neq 0, -1$, il suffit de prendre $x = 1$. Ceci étant, choisissons $e_x = e'_1 + xe'_2$ vérifiant les conditions du lemme. Comme e_x n'est pas isotrope, il existe e''_2 tel que (e_x, e''_2) soit une base orthogonale de $Ke'_1 \hat{\oplus} Ke'_2$. Posons : $\mathbf{e}'' = (e_x, e''_2, e'_3, \dots, e'_n)$, c'est une base orthogonale de V . Comme $Ke_1 + Ke_x$ est un plan non dégénéré, la partie i) de la démonstration montre que l'on peut relier \mathbf{e} à \mathbf{e}'' par une chaîne de bases contiguës; d'autre part \mathbf{e}' et \mathbf{e}'' sont contiguës; d'où le théorème. \square

3.5 Théorème de Witt.

Soient (V, Q) et (V', Q') deux modules quadratiques non dégénérés; soit U un sous-espace vectoriel de V , et soit

$$s : U \longrightarrow V'$$

un morphisme métrique injectif de U dans V' . On cherche à prolonger s à un sous-espace plus grand que U , et si possible à V tout entier. On commence par le cas où U est dégénéré :

Lemme 3.5.1. *Si U est dégénéré, on peut prolonger s en un morphisme métrique injectif $s_1 : U_1 \rightarrow V'$, où U_1 contient U comme hyperplan.*

Preuve :

Soient x un élément non nul de $\text{rad}(U)$ et l une forme linéaire sur U telle que $l(x) = 1$. Comme V est dégénéré, il existe $y \in V$ tel que $l(u) = u.y$ pour tout $u \in U$; de plus, on peut supposer que $y.y = 0$ (remplacer y par $y - \lambda x$, avec $\lambda = \frac{1}{2}y.y$). Le sous-espace $U_1 = U \hat{\oplus} Ky$ contient U comme hyperplan. La même construction, appliquée à $U' = sU$, $x' = sx$, et $l' = l \circ s^{-1}$ donne $U'_1 = U' \hat{\oplus} Ky'$. Soit $s_1 : U_1 \rightarrow U'_1$ l'application linéaire qui coïncide avec s sur U et applique y sur y' ; il est immédiat que s_1 répond à la question. \square

Théorème 3.5.2. (Witt) *Si (V, Q) et (V', Q') sont isomorphes, et non dégénérés, tout morphisme métrique injectif*

$$s : U \longrightarrow V'$$

d'un sous-espace vectoriel U de V peut être prolongé en un isomorphisme de V sur V' .

Preuve :

Puisque V et V' sont isomorphes, on peut supposer que $V = V'$. D'autre part, en appliquant le lemme précédent, on voit que l'on peut se restreindre au cas où U est non dégénéré. On raisonne alors par récurrence sur la dimension de U . Si $\dim U = 1$, U est engendré par un élément x non isotrope; si $y = s(x)$,

on a $y.y = x.x$. On peut choisir $\epsilon = \pm 1$ tel que $x + \epsilon y$ ne soit pas isotrope ; sinon, en effet, on aurait $2x.x + 2x.y = 2x.x - 2x.y = 0$, ce qui entraînerait $x.x = 0$. Choisissons un tel ϵ , et soit H l'hyperplan orthogonal à $z = x + \epsilon y$; on a $V = Kz \hat{\oplus} H$. Soit σ la symétrie par rapport à H , i.e. l'automorphisme de V qui est l'identité sur H et qui change z en $-z$. Comme $x - \epsilon y$ appartient à H , on a : $\sigma(x - \epsilon y) = x - \epsilon y$ et $\sigma(x + \epsilon y) = -x - \epsilon y$, d'où $\sigma(x) = -\epsilon y$. L'automorphisme $-\epsilon\sigma$ prolonge donc s .

Si $\dim U > 1$, on décompose U sous la forme $U_1 \hat{\oplus} U_2$, avec $U_1, U_2 \neq 0$. D'après l'hypothèse de récurrence, la restriction s_1 de s à U_1 se prolonge en un automorphisme σ_1 de V ; quitte à remplacer s par $\sigma_1^{-1} \circ s$, on peut donc supposer que s est l'identité sur U_1 . Le morphisme s applique alors U_2 dans l'orthogonal V_1 de U_1 ; d'après l'hypothèse de récurrence, la restriction de s à U_2 se prolonge donc en un automorphisme σ_2 de V_1 ; l'automorphisme σ de V qui est l'identité sur U_1 et σ_2 sur V_1 répond à la question. \square

Corollaire 3.5.3. *Deux sous-espaces isomorphes d'un module quadratique non dégénéré ont des orthogonaux isomorphes.*

Preuve :

On prolonge un isomorphisme entre les deux sous-espaces en un automorphisme du module, et l'on restreint ce dernier aux orthogonaux, d'où le résultat. \square

3.6 Traductions.

Soit

$$f(X) = \sum_{i=1}^n a_{ii} X_i^2 + 2 \sum_{i < j} a_{ij} X_i X_j$$

une forme quadratique à n variables sur K ; on pose $a_{ij} = a_{ji}$ si $i > j$, de sorte que la matrice $A = (a_{ij})$ est symétrique. Le couple (K^n, f) est un module quadratique, dit **associé à f** (ou à la matrice A).

Définition 3.6.1. *Deux formes quadratiques f et f' sont dites équivalentes si les modules correspondants sont isomorphes.*

On écrit alors $f \sim f'$. Si A et A' sont les matrices respectives de f et f' , cela revient à dire qu'il existe une matrice inversible X telle que $A' = X.A.^t X$.

Soient $f(X_1, \dots, X_n)$ et $g(X_1, \dots, X_m)$ deux formes quadratiques ; on note $f \dot{+} g$ (ou simplement $f + g$ si aucune confusion n'est possible) la forme quadratique

$$f(X_1, \dots, X_n) + g(X_{n+1}, \dots, X_{n+m})$$

en $n + m$ variables. Cette opération correspond à celle de somme orthogonale. On écrit de même $f \dot{-} g$ (ou simplement $f - g$) pour $f \dot{+} (-g)$.

Voici quelques exemples de traductions :

Définition 3.6.2. *Une forme $f(X_1, X_2)$ à deux variables est dite hyperbolique si l'on a :*

$$f \sim X_1 X_2 \sim X_1^2 - X_2^2.$$

Cela signifie que le module (K^2, f) correspondant est un plan hyperbolique.

On dit qu'une forme $f(X_1, \dots, X_n)$ représente un élément a de K s'il existe $x \in K^n$, $x \neq 0$, tel que $f(x) = a$. En particulier, f représente 0 si et seulement si le module quadratique correspondant contient un élément isotrope non nul.

Proposition 3.6.3. *Si f représente 0, et est non dégénérée, on a $f \sim f_2 \dot{+} g$, où f_2 est hyperbolique. De plus, f représente tout élément de K .*

C'est la traduction de la proposition 3.3.3 et de son corollaire 3.3.4.

Corollaire 3.6.4. *Soit $g = g(X_1, \dots, X_{n-1})$ une forme quadratique non dégénérée, et soit $a \in K^*$. Les propriétés suivantes sont équivalentes :*

- i) g représente a .
- ii) On a $g \sim h \dot{+} aZ^2$, où h est une forme en $n - 2$ variables.
- iii) La forme $f = g \dot{-} aZ^2$ représente 0.

Preuve :

Il est clair que ii) \Rightarrow i). Inversement, si g représente a , le module quadratique V correspondant à g contient un élément x tel que $x.x = a$; si H désigne l'orthogonal de x , on a $V = H \hat{\oplus} Kx$ d'où $g \sim h + aZ^2$ où h désigne la forme quadratique attachée à une base de H .

L'implication ii) \Rightarrow iii) est immédiate.

Enfin, si la forme $f = g \dot{-} aZ^2$ a un zéro non trivial (x_1, \dots, x_{n-1}, z) , on a, soit $z = 0$, auquel cas g représente 0, donc aussi a , soit $z \neq 0$, auquel cas $g(x_1/z, \dots, x_{n-1}/z) = a$, d'où iii) \Rightarrow i). \square

Corollaire 3.6.5. *Soient g et h deux formes non dégénérées de rang ≥ 1 , et soit $f = g \dot{-} h$. Les propriétés suivantes sont équivalentes :*

- a) f représente 0.
- b) Il existe $a \in K^*$ qui est représenté par g et par h .
- c) Il existe $a \in K^*$ tel que $g \dot{-} aZ^2$ et $h \dot{-} aZ^2$ représentent 0.

Preuve :

L'équivalence b) \Leftrightarrow c) résulte du corollaire 3.6.4.

L'implication b) \Rightarrow a) est triviale.

Montrons que a) \Rightarrow b). Un zéro non trivial de f peut s'écrire sous la forme (x, y) , avec $g(x) = h(y)$. Si l'élément $a = g(x) = h(y)$ est $\neq 0$, il est clair que b) est vérifiée. Si $a = 0$, l'une des formes, g par exemple, représente 0, donc tout élément de K , et en particulier, toute valeur non nulle prise par h . \square

Le théorème 3.4.2 se traduit en la classique décomposition des formes quadratiques en "sommes de carrés" :

Théorème 3.6.6. *Soit f une forme quadratique en n variables. Il existe $a_1, \dots, a_n \in K$ tels que $f \sim a_1X_1^2 + \dots + a_nX_n^2$.*

Le rang de f est le nombre des indices i tels que $a_i \neq 0$. Il est égal à n si et seulement si le discriminant $a_1 \cdots a_n$ de f est $\neq 0$ (autrement dit, si f est non dégénérée).

Enfin, le corollaire au théorème de Witt donne le théorème de "simplification" suivant :

Théorème 3.6.7. Soient $f = g \dot{+} h$ et $f' = g' \dot{+} h'$ deux formes quadratiques non dégénérées. Si $f \sim f'$ et $g \sim g'$, on a $h \sim h'$.

Corollaire 3.6.8. si f est non dégénérée, on a

$$f \sim g_1 \dot{+} \cdots \dot{+} g_m \dot{+} h$$

où g_1, \dots, g_m sont hyperboliques, et h ne représente pas 0. Cette décomposition est unique, à équivalence près.

Remarque. Le nombre m des facteurs hyperboliques peut être caractérisé comme la dimension des sous-espaces isotropes maximaux du module quadratique correspondant à f .

3.7 Formes quadratiques sur \mathbb{F}_q .

Soit p un nombre premier différent de 2, et soit $q = p^f$ une puissance de p ; soit \mathbb{F}_q un corps à q éléments.

Proposition 3.7.1. Une forme quadratique sur \mathbb{F}_q de rang ≥ 2 (resp. de rang ≥ 3) représente tout élément de \mathbb{F}_q^* (resp. de \mathbb{F}_q).

Preuve :

Il s'agit ici de montrer que, si $a, b, c \in \mathbb{F}_q$ sont non nuls, l'équation $ax^2 + by^2 = c$ (*) a une solution.

Soit A (resp. B) l'ensemble des éléments de \mathbb{F}_q de la forme ax^2 (resp. de la forme $c - by^2$), avec $x \in \mathbb{F}_q$ (resp. avec $y \in \mathbb{F}_q$). On voit rapidement que A et B ont chacun $(q+1)/2$ éléments; on a donc $A \cap B \neq \emptyset$, d'où une solution de l'équation (*). \square

Rappelons d'autre part que le groupe $\mathbb{F}_q^*/\mathbb{F}_q^{*2}$ a deux éléments. Notons a un élément de \mathbb{F}_q^* qui n'est pas un carré.

Proposition 3.7.2. Toute forme quadratique non dégénérée de rang n sur \mathbb{F}_q est équivalente à

$$X_1^2 + \cdots + X_{n-1}^2 + X_n^2$$

ou

$$X_1^2 + \cdots + X_{n-1}^2 + aX_n^2$$

suivant que son discriminant est ou non un carré.

Preuve :

Nous allons démontrer ce résultat par récurrence sur n .

C'est clair si $n = 1$. La proposition précédente permet de montrer que la forme f représente 1. Cette forme est donc équivalente à $X_1^2 \dot{+} g$, où g est une forme à $n - 1$ variables, et l'on applique l'hypothèse de récurrence à g . D'où le résultat. \square

Corollaire 3.7.3. Pour que deux formes quadratiques non dégénérées sur \mathbb{F}_q soient équivalentes, il faut et il suffit qu'elles aient même rang et même discriminant.

(Le discriminant est alors considéré comme un élément du groupe quotient $\mathbb{F}_q^*/\mathbb{F}_q^{*2}$.)

Chapitre 4

Formes quadratiques sur \mathbb{Q}_p .

Dans cette partie, p désigne un nombre premier ; on note K le corps p -adique \mathbb{Q}_p .

Tous les modules quadratiques considérés sont relatifs à K , et sont supposés non dégénérés ; on fait les mêmes conventions pour les formes quadratiques.

4.1 Les deux invariants.

Soit (V, Q) un module quadratique de rang n ; nous noterons $d(Q)$ son discriminant ; c'est un élément de K^*/K^{*2} . Si $\mathbf{e} = (e_1, \dots, e_n)$ est une base orthogonale de V , et si l'on pose $a_i = e_i \cdot e_i$, on a :

$$d(Q) = a_1 \cdots a_n \quad (\text{dans } K^*/K^{*2}).$$

Dans tout ce qui suit, nous nous permettrons souvent de noter par la même lettre un élément de K^* et sa classe modulo K^{*2} .

Rappelons d'autre part que, si a et b sont des éléments de K^* , on a le symbole de Hilbert (a, b) , égal à ± 1 . Nous poserons :

$$\varepsilon(\mathbf{e}) = \prod_{i < j} (a_i, a_j).$$

On a $\varepsilon(\mathbf{e}) = \pm 1$. De plus, $\varepsilon(\mathbf{e})$ est un invariant de (V, Q) . En effet :

Théorème 4.1.1. *Le nombre $\varepsilon(\mathbf{e})$ ne dépend pas du choix de la base orthogonale \mathbf{e} .*

Preuve :

Si $n = 1$, on a $\varepsilon(\mathbf{e}) = 1$. Si $n = 2$, on a $\varepsilon(\mathbf{e}) = 1$ si et seulement si la forme $Z^2 - a_1 X^2 - a_2 Y^2$ représente 0, autrement dit (cf. corollaire 3.6.4) si et seulement si $a_1 X^2 + a_2 Y^2$ représente 1 ; mais cette dernière condition signifie qu'il existe $x \in V$ tel que $Q(x) = 1$, et cela ne dépend pas de \mathbf{e} . Pour $n \geq 3$, on raisonne par récurrence sur n . D'après le théorème 3.4.4, il suffit de prouver que $\varepsilon(\mathbf{e}) = \varepsilon(\mathbf{e}')$ sont contiguës. Mais, vu la symétrie du symbole de Hilbert, $\varepsilon(\mathbf{e})$ ne change pas de valeur lorsqu'on permute les e_i ; on peut donc supposer que $\mathbf{e}' = (e'_1, \dots, e'_n)$ est tel que $e'_1 = e_1$. Si l'on pose $a'_i = e'_i \cdot e'_i$, on a $a'_1 = a_1$. On peut écrire $\varepsilon(\mathbf{e})$ sous la forme :

$$\begin{aligned}\varepsilon(\mathbf{e}) &= (a_1, a_2 \cdots a_n) \prod_{2 \leq i < j} (a_i, a_j) \\ &= (a_1, d(Q)a_1) \prod_{2 \leq i < j} (a_i, a_j)\end{aligned}$$

puisque $d(Q) = a_1 \cdots a_n$.

De même

$$\varepsilon(\mathbf{e}') = (a_1, d(Q)a_1) \prod_{2 \leq i < j} (a'_i, a'_j).$$

Mais l'hypothèse de récurrence, appliquée à l'orthogonal de e_1 , montre que l'on a :

$$\prod_{2 \leq i < j} (a_i, a_j) = \prod_{2 \leq i < j} (a'_i, a'_j),$$

d'où le résultat recherché. \square

Nous écrirons désormais $\varepsilon(Q)$ à la place de $\varepsilon(\mathbf{e})$.

Traduction : Si f une forme quadratique à n variables, et si

$$f \sim a_1 X_1^2 + \cdots + a_n X_n^2,$$

les deux éléments

$$d(f) = a_1 \cdots a_n \text{ (dans } K^*/K^{*2}) \text{ et } \varepsilon(f) = \prod_{i < j} (a_i, a_j) \text{ (dans } \{\pm 1\})$$

sont des invariants de la classe d'équivalence de f .

4.2 Représentation d'un élément de \mathbb{Q}_p par une forme quadratique.

Lemme 4.2.1. *i) Le nombre d'éléments du \mathbb{F}_2 -espace vectoriel $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ est 2^r , avec $r = 2$ si $p \neq 2$ et $r = 3$ si $p = 2$.*

ii) Si $a \in \mathbb{Q}_p^/\mathbb{Q}_p^{*2}$ et $\varepsilon = \pm 1$, soit H_a^ε l'ensemble des $x \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ tels que $(x, a) = \varepsilon$. Si $a = 1$, H_a^1 a 2^r éléments et $H_a^{-1} = \emptyset$. Si $a \neq 1$, H_a^ε a 2^{r-1} éléments.*

iii) Soient $a, a' \in \mathbb{Q}_p^/\mathbb{Q}_p^{*2}$ et $\varepsilon, \varepsilon' = \pm 1$; on suppose que H_a^ε et $H_{a'}^{\varepsilon'}$ sont non vides. Pour que $H_a^\varepsilon \cap H_{a'}^{\varepsilon'} = \emptyset$, il faut et il suffit que $a = a'$ et $\varepsilon = -\varepsilon'$.*

Preuve :

On a que si $p \neq 2$, le groupe $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ est un groupe de type $(2, 2)$, autrement dit isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$. Et si $p = 2$, le groupe $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ est un groupe de type $(2, 2, 2)$, donc isomorphe à $(\mathbb{Z}/2\mathbb{Z})^3$. L'assertion i) en découle.

Pour l'assertion ii), le cas $a = 1$ est trivial; si $a \neq 1$, on a que l'homomorphisme $b \mapsto (a, b)$ applique $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ sur $\{\pm 1\}$; son noyau H_a^1 est donc un hyperplan de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ et a 2^{r-1} éléments; son complémentaire H_a^{-1} a 2^{r-1} éléments (c'est un hyperplan "affine" parallèle à H).

Enfin, si H_a^ε et $H_{a'}^{\varepsilon'}$ sont non vides et disjoints, ils ont nécessairement 2^{r-1} éléments chacun et sont complémentaires l'un de l'autre; cela entraîne $H_a^1 = H_{a'}^1$, d'où

$$(x, a) = (x, a') \text{ pour tout } x \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2};$$

comme le symbole de Hilbert est non dégénéré, on en déduit $a = a'$ et évidemment $\varepsilon = -\varepsilon'$; la réciproque est triviale. \square

Soit maintenant f une forme quadratique de rang n ; soit $d = d(f)$ et $\varepsilon = \varepsilon(f)$ ses deux invariants.

Théorème 4.2.2. *Pour que f représente 0, il faut et il suffit que :*

- i) $n = 2$ et $d = -1$ (dans $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$);
- ii) $n = 3$ et $(-1, -d) = \varepsilon$;
- iii) $n = 4$ et, soit $d \neq 1$, soit $d = 1$ et $\varepsilon = (-1, -1)$;
- iv) $n \geq 5$.

(En particulier, toute forme à au moins 5 variables représente 0.)

Avant de démontrer ce théorème, nous allons en donner une conséquence :

Soit $a \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$, et soit $f_a = f - aZ^2$ (cette forme est définie à équivalence près). On sait que f_a représente 0 si et seulement si f représente a . Or, on a :

$$d(f_a) = -ad, \quad \text{et} \quad \varepsilon(f_a) = (-a, d)\varepsilon.$$

En appliquant le théorème précédent à f_a et en tenant compte de ces formules, on obtient :

Corollaire 4.2.3. *Soit $a \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$. Pour que f représente a , il faut et il suffit que :*

- i) $n = 1$ et $a = d$;
- ii) $n = 2$ et $(a, -d) = \varepsilon$;
- iii) $n = 3$ et, soit $a \neq -d$, soit $a = -d$ et $(-1, -d) = \varepsilon$;
- iv) $n \geq 4$

(Précisons que, dans cet énoncé comme dans celui du théorème 4.2.2, a et d sont considérés comme des éléments de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$; ainsi, l'inégalité $a \neq -d$ signifie que a n'est pas égal au produit de $-d$ par un carré.)

Preuve du théorème :

On écrit f sous la forme $f \sim a_1X_1^2 + \dots + a_nX_n^2$, et l'on considère les cas $n = 2, 3, 4$ et ≥ 5 .

i) Le cas $n = 2$.

La forme f représente 0 si et seulement si $-a_1/a_2$ est un carré; mais $-a_1/a_2 = -a_1a_2 = -d$ dans $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$; on doit donc avoir $-d = 1$, soit $d = -1$.

ii) Le cas $n = 3$.

La forme f représente 0 si et seulement si la forme

$$-a_3f \sim -a_3a_1X_1^2 - a_3a_2X_2^2 - X_3^2$$

représente 0. Or par définition même du symbole de Hilbert, cette dernière forme représente 0 si et seulement si l'on a : $(-a_3a_1, -a_3a_2) = 1$.

En développant, on trouve :

$$(-1, -1)(-1, a_1)(-1, a_2)(a_3, a_3)(a_1, a_2)(a_1, a_3)(a_2, a_3) = 1.$$

Mais on a $(a_3, a_3) = (-1, a_3)$, on peut donc réécrire la condition ci-dessus sous la forme :

$$(-1, -1)(-1, a_1 a_2 a_3)(a_1, a_2)(a_1, a_3)(a_2, a_3) = 1$$

ou encore $(-1, -d)\varepsilon = 1$, donc $(-1, -d) = \varepsilon$.

iii) Le cas $n = 4$.

D'après le corollaire 3.6.5, on a que f représente 0 si, et seulement si, il existe un élément $x \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ qui est représenté par les deux formes

$$a_1 X_1^2 + a_2 X_2^2 \quad \text{et} \quad -a_3 X_3^2 - a_4 X_4^2.$$

D'après le cas ii) du corollaire 4.2.3, un tel x est caractérisé par les conditions suivantes :

$$(x, -a_1 a_2) = (a_1, a_2) \quad \text{et} \quad (x, -a_3 a_4) = (-a_3, -a_4).$$

Soit A la partie de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ définie par la première condition, et soit B celle définie par la seconde. Pour que f ne représente pas 0, il faut et il suffit que $A \cap B = \emptyset$. Or A et B sont évidemment non vides (on a $a_1 \in A$ et $-a_3 \in B$, par exemple). D'après la partie iii) du lemme 4.2.1, la relation $A \cap B = \emptyset$ équivaut donc à :

$$a_1 a_2 = a_3 a_4 \quad \text{et} \quad (a_1, a_2) = -(-a_3, -a_4).$$

La première condition signifie que $d = 1$. Si elle est réalisée, on a :

$$\varepsilon = (a_1, a_2)(a_3, a_4)(a_3 a_4, a_3 a_4) ;$$

en utilisant la relation $(x, x) = (-1, x)$, on obtient :

$$\begin{aligned} \varepsilon &= (a_1, a_2)(a_3, a_4)(-1, a_3 a_4) \\ &= (a_1, a_2)(-a_3, -a_4)(-1, -1). \end{aligned}$$

On voit ainsi que la seconde condition s'écrit :

$$\varepsilon = (-1, -1),$$

d'où le résultat cherché.

iv) Le cas $n \geq 5$.

Il suffit de traiter le cas $n = 5$. En utilisant le lemme 4.2.1 et la partie ii) du corollaire 4.2.3, on voit qu'une forme de rang 2 représente au moins 2^{r-1} éléments de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$, il en est *a fortiori* de même pour les formes de rang ≥ 2 . Comme $2^{r-1} \geq 2$, f représente au moins un élément $a \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ qui est distinct de d . On a :

$$f \sim aX^2 + g$$

où g est une forme de rang 4. Le discriminant de g est égal à d/a ; il est donc différent de 1, et d'après iii), la forme g représente 0. Il en est alors de même de f , ce qui achève la démonstration du théorème 4.2.2. \square

Remarque. Soit f une forme quadratique ne représentant pas 0. Les résultats ci-dessus montrent que le nombre d'éléments de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ qui sont représentés par f est égal à 1 si $n = 1$, à 2^{r-1} si $n = 2$, à $2^r - 1$ si $n = 3$, et à 2^r si $n = 4$.

4.3 Classification.

Théorème 4.3.1. *Deux formes quadratiques sur \mathbb{Q}_p sont équivalentes si et seulement si elles ont même rang, même discriminant, et même invariant ε .*

Preuve :

Que deux formes équivalentes aient mêmes invariants résulte de la définition de ceux-ci. La réciproque se démontre par récurrence sur le rang n des deux formes f et g considérées (le cas $n = 0$ étant trivial). Le corollaire du théorème du paragraphe précédent, montre que f et g représentent les mêmes éléments de $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$; on peut donc trouver $a \in \mathbb{Q}_p^*$ qui est représenté à la fois par f et par g ; cela permet d'écrire :

$$f \sim aZ^2 \dot{+} f' \quad \text{et} \quad g \sim aZ^2 \dot{+} g'$$

où f' et g' sont des formes de rang $n - 1$. On a :

$$\begin{aligned} d(f') &= ad(f) = ad(g) = d(g') \\ \varepsilon(f') &= \varepsilon(f)(a, d(f')) = \varepsilon(g)(a, d(g')) = \varepsilon(g') \end{aligned}$$

ce qui montre que f' et g' ont les mêmes invariants. compte tenu de l'hypothèse de récurrence, on a $f' \sim g'$, d'où $f \sim g$. \square

Corollaire 4.3.2. *A équivalence près, il existe une forme de rang 4 et une seule qui ne représente pas 0; si $(a, b) = -1$, c'est la forme $z^2 - ax^2 - by^2 + abt^2$.*

Preuve :

D'après le théorème du paragraphe précédent, une telle forme est caractérisée par $d(f) = 1$, $\varepsilon(f) = -(-1, -1)$, et un calcul immédiat montre que $z^2 - ax^2 - by^2 + abt^2$ a ces propriétés. \square

Remarque. On peut caractériser cette forme comme le norme réduite de l'unique corps non commutatif de degré 4 sur \mathbb{Q}_p .

Proposition 4.3.3. *Soient $n \geq 1$, $d \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ et $\varepsilon = \pm 1$. Pour qu'il existe une forme quadratique f de rang n telle que $d(f) = d$ et $\varepsilon(f) = \varepsilon$, il faut et il suffit que l'on ait :*

$$\begin{aligned} &n = 1, \varepsilon = 1; \\ \text{ou} &n = 2, d \neq -1; \\ \text{ou} &n = 2, \varepsilon = 1; \\ \text{ou} &n \geq 3. \end{aligned}$$

Preuve :

Le cas $n = 1$ est trivial.

Si $n = 2$, on a $f \sim aX^2 + bY^2$ et, si $d(f) = -1$, $\varepsilon(f) = (a, b) = (a, -ab) = 1$; on ne peut donc pas avoir simultanément $d(f) = -1$ et $\varepsilon(f) = -1$. Inversement, si $d = -1$, $\varepsilon = 1$, on prend $f = X^2 - Y^2$; si $d \neq -1$, il existe $a \in \mathbb{Q}_p^*$ tel que $(a, -d) = \varepsilon$, et l'on prend $f = aX^2 + adY^2$.

Si $n = 3$, on choisit $a \in \mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ distinct de $-d$; d'après ce qu'on vient de voir, il existe une forme g de rang 2 telle que $d(g) = ad$, $\varepsilon(g) = \varepsilon(a, -d)$; la forme $aZ^2 \dot{+} g$ convient.

Le cas $n \geq 4$ se ramène au cas $n = 3$ en prenant f de la forme $g(X_1, X_2, X_3) + X_4^2 + \dots + X_n^2$, où g a les invariants voulus. \square

Corollaire 4.3.4. *Le nombre des classes de formes quadratiques de rang n sur \mathbb{Q}_p , $p \neq 2$ (resp. $p = 2$) est égal à 4 (resp. 8) si $n = 1$, à 7 (resp. 15) si $n = 2$, et à 8 (resp. 16) si $n \geq 3$.*

Preuve :

On sait que $d(f)$ peut prendre 4 valeurs si $p \neq 2$ et 8 valeurs si $p = 2$. Comme $\varepsilon(f)$ peut prendre deux valeurs, on a le résultat. \square

4.4 Le cas réel.

Soit f une forme quadratique de rang n sur le corps \mathbb{R} des nombres réels. On sait que f est équivalente à

$$X_1^2 + \dots + X_r^2 - Y_1^2 - \dots - Y_s^2$$

où r et s sont deux entiers ≥ 0 tels que $r + s = n$; le couple (r, s) ne dépend que de f ; on l'appelle la **signature** de f .

On dit que f est **définie** si r ou s vaut 0, autrement dit si f a un signe constant; sinon, on dit que f est **indéfinie** (c'est le cas où f représente 0).

L'**invariant** $\varepsilon(f)$ se définit comme dans le cas de \mathbb{Q}_p ; comme $(-1, -1) = -1$, on a :

$$\varepsilon(f) = (-1)^{s(s-1)/2} = \begin{cases} 1 & \text{si } s \equiv 0, 1 \pmod{4} \\ -1 & \text{si } s \equiv 2, 3 \pmod{4}. \end{cases}$$

D'autre part :

$$d(f) = (-1)^s = \begin{cases} 1 & \text{si } s \equiv 0 \pmod{2} \\ -1 & \text{si } s \equiv 1 \pmod{2}. \end{cases}$$

On voit ainsi que la connaissance de $d(f)$ et $\varepsilon(f)$ équivaut à celle de la classe de s modulo 4; en particulier, $d(f)$ et $\varepsilon(f)$ déterminent f à équivalence près si $n \leq 3$.

On pourrait également vérifier que les parties i), ii) et iii) du théorème 4.2.2 et de son corollaire sont valables sur \mathbb{R} .

Chapitre 5

Formes quadratiques sur \mathbb{Q} .

Toutes les formes quadratiques considérées sont à coefficients dans \mathbb{Q} , et sont supposées non dégénérées.

5.1 Invariants d'une forme.

On note V la réunion de l'ensemble des nombres premiers et du symbole ∞ , et l'on convient que $\mathbb{Q}_\infty = \mathbb{R}$.

Soit $f \sim a_1 X_1^2 + \cdots + a_n X_n^2$ une forme quadratique de rang n . On lui associe les invariants suivants :

i) Le discriminant $d(f) \in \mathbb{Q}^*/\mathbb{Q}^{*2}$, égal à $a_1 \cdots a_n$.

ii) Soit $v \in V$. L'injection $\mathbb{Q} \rightarrow \mathbb{Q}_v$ permet de considérer f comme une forme quadratique sur \mathbb{Q}_v , que nous noterons f_v . Les invariants de f_v seront notés $d_v(f)$ et $\varepsilon_v(f)$.

Il est clair que $d_v(f)$ est l'image de $d(f)$ par $\mathbb{Q}^*/\mathbb{Q}^{*2} \rightarrow \mathbb{Q}_v^*/\mathbb{Q}_v^{*2}$; on a :

$$\varepsilon_v(f) = \prod_{i < j} (a_i, a_j)_v.$$

La formule du produit entraîne la relation :

$$\prod_{v \in V} \varepsilon_v(f) = 1.$$

iii) La signature (r, s) de la forme quadratique réelle f_∞ est un autre invariant de f .

Les invariants $d_v(f)$, $\varepsilon_v(f)$, et (r, s) sont parfois appelés les invariants **locaux** de f .

5.2 Représentation d'un nombre par une forme.

Théorème 5.2.1. (Hasse-Minkowski) *Pour que f représente 0, il faut et il suffit que, pour tout $v \in V$, la forme f_v représente 0.*

On appelle aussi ce théorème le principe local-global : f a un zéro "global" si et seulement si f a partout un zéro "local".

Remarque. En référence à ce théorème, on dit qu'une classe de variétés définies sur \mathbb{Q} "vérifie le principe de Hasse" si l'existence d'un zéro réel et d'un zéro p -adique pour tout p entraîne l'existence d'un zéro rationnel. On peut ainsi reformuler le théorème de Hasse-Minkowski en disant que les quadriques ou hypersurfaces de degré 2 vérifient le principe de Hasse.

Preuve : Soit f une forme quadratique sur \mathbb{Q} . Il suffit de traiter le cas où f est non dégénérée (dans le cas contraire, f représente déjà 0).

Il est immédiat que si f représente 0 dans \mathbb{Q} , elle représente 0 dans chacun des \mathbb{Q}_v .

Pour avoir la suffisance, on écrit f sous la forme :

$$f = a_1X_1^2 + \cdots + a_nX_n^2, \quad a_i \in \mathbb{Q}^*.$$

Quitte à remplacer f par a_1f , on peut en outre supposer que $a_1 = 1$. On considère séparément les cas $n = 2, 3, 4$ et ≥ 5 .

i) Le cas $n = 2$.

On a $f = X_1^2 - aX_2^2$; comme f_∞ représente 0, a est strictement positif. Si on écrit a sous la forme :

$$a = \prod_{p \in V \setminus \infty} p^{v_p(a)}$$

Le fait que f_p représente 0 montre que a est un carré dans \mathbb{Q}_p , donc que $v_p(a)$ est pair. Il en résulte que a est un carré dans \mathbb{Q} , et f représente bien 0.

ii) Le cas $n = 3$ (Legendre).

On a $f = X_1^2 - aX_2^2 - bX_3^2$; quitte à multiplier a, b par des carrés, on peut supposer que a et b sont des entiers sans facteurs carrés (i.e. $v_p(a), v_p(b)$ sont égaux à 0 ou à 1 pour tout nombre premier p). On peut aussi supposer que $|a| \leq |b|$. On raisonne alors par récurrence sur l'entier $m = |a| + |b|$. Si $m = 2$, on a :

$$f = X_1^2 \pm X_2^2 \pm X_3^2;$$

le cas de $X_1^2 + X_2^2 + X_3^2$ est exclu, puisque f_∞ représente 0; dans les autres cas, f représente bien zéro.

Supposons $m > 2$, i.e. $|b| \geq 2$, et écrivons b sous la forme :

$$b = \pm p_1 \cdots p_k$$

où les p_i sont des nombres premiers distincts. Soit p l'un des p_i ; nous allons voir que a est un carré modulo p . C'est évident si $a \equiv 0 \pmod{p}$. Sinon, a est une unité p -adique; par hypothèse, il existe $(x, y, z) \in (\mathbb{Q}_p)^3$ tel que

$$z^2 - ax^2 - by^2 = 0$$

et l'on peut supposer (x, y, z) primitif (cf. proposition 1.3.4). On a $z^2 - ax^2 \equiv 0 \pmod{p}$. On en conclut que, si $x \equiv 0 \pmod{p}$, il en est de même de z , et by^2 est

divisible par p^2 ; comme $v_p(b) = 1$, cela entraîne $y \equiv 0 \pmod{p}$ contrairement au fait que (x, y, z) est primitif. On a donc $x \not\equiv 0 \pmod{p}$, ce qui montre bien que a est un carré \pmod{p} . Ceci étant, comme $\mathbb{Z}/b\mathbb{Z} = \prod \mathbb{Z}/p_i\mathbb{Z}$, on voit que a est un carré modulo b . Il existe donc des entiers t, b' tels que :

$$t^2 = a + bb'$$

et l'on peut choisir t de telle sorte que $|t| \leq |b|/2$. La formule $bb' = t^2 - a$ montre que bb' est une norme de l'extension $K(\sqrt{a})/K$, où $K = \mathbb{Q}$ ou \mathbb{Q}_v ; on en conclut que f représente 0 dans K si et seulement si il en est de même de $f' = X_1^2 - aX_2^2 - b'X_3^2$. En particulier, f' représente 0 dans chacun des \mathbb{Q}_v . Mais l'on a :

$$|b'| = \left| \frac{t^2 - a}{b} \right| \leq \frac{|b|}{4} + 1 < |b| \quad \text{puisque } |b| \geq 2.$$

Ecrivons b' sous la forme $b''u^2$ avec b'' sans facteurs carrés; on a *a fortiori* $|b''| < |b|$. L'hypothèse de récurrence s'applique donc à la forme :

$$f'' = X_1^2 - aX_2^2 - b''X_3^2$$

qui est équivalente à f' . On voit ainsi que f'' représente 0 dans \mathbb{Q} , et il en est de même pour f .

iii) Le cas $n = 4$.

On a :

$$f = aX_1^2 + bX_2^2 - (cX_3^2 + dX_4^2).$$

Soit $v \in V$. Puisque f_v représente 0, le corollaire 3.6.5 montre qu'il existe $x_v \in \mathbb{Q}_v^*$ qui est représenté à la fois par $aX_1^2 + bX_2^2$ et par $cX_3^2 + dX_4^2$; d'après la partie ii) du corollaire 4.2.3 (qui s'applique également à $\mathbb{Q}_\infty = \mathbb{R}$), cela revient à dire que l'on a :

$$(x_v, -ab)_v = (a, b)_v \quad \text{et} \quad (x_v, -cd)_v = (c, d)_v.$$

Comme $\prod_{v \in V} (a, b)_v = \prod_{v \in V} (c, d)_v = 1$, on peut appliquer le théorème 2.2.7, et l'on en conclut qu'il existe $x \in \mathbb{Q}^*$ tel que :

$$(x, -ab)_v = (a, b)_v \quad \text{et} \quad (x, -cd)_v = (c, d)_v \quad \text{pour tout } v \in V.$$

La forme $aX_1^2 + bX_2^2 - xZ^2$ représente alors 0 dans chacun des \mathbb{Q}_v , donc dans \mathbb{Q} d'après ce que l'on vient de voir. On en conclut que x est représenté dans \mathbb{Q} par $aX_1^2 + bX_2^2$, et le même argument s'applique à $cX_3^2 + dX_4^2$; d'où le fait que f représente 0.

iv) Le cas $n \geq 5$.

On raisonne par récurrence sur n . On écrit f sous la forme

$$f = h - g$$

avec $h = a_1X_1^2 + a_2X_2^2$, $g = -(a_3X_3^2 + \dots + a_nX_n^2)$.

Soit S la partie de V formée de $\infty, 2$ et des nombres premiers p tels que $v_p(a_i) \neq 0$

pour un $i \geq 3$; c'est un ensemble fini. Soit $v \in S$. Puisque f_v représente 0, il existe $a_v \in \mathbb{Q}_v^*$ qui est représenté dans \mathbb{Q}_v par h et par g ; il existe donc $x_i^v \in \mathbb{Q}_v$, $i = 1, \dots, n$, tels que :

$$h(x_1^v, x_2^v) = a_v = g(x_3^v, \dots, x_n^v).$$

Mais les carrés de \mathbb{Q}_v^* forment un ensemble ouvert. On en conclut au moyen du théorème d'approximation, qu'il existe $x_1, x_2 \in \mathbb{Q}$ tels que, si $a = h(x_1, x_2)$, on ait $a/a_v \in \mathbb{Q}_v^{*2}$ pour tout $v \in S$. Considérons maintenant la forme :

$$f_1 = aZ^2 - g.$$

Si $v \in S$, g représente a_v dans \mathbb{Q}_v , donc aussi a puisque $a/a_v \in \mathbb{Q}_v^{*2}$; on en conclut que f_1 représente 0 dans \mathbb{Q}_v .

Si $v \notin S$, les coefficients $-a_3, \dots, -a_n$ de g sont des unités p -adiques; il en est de même de $d_v(g)$, et, puisque $v \neq 2$, on a $\varepsilon_v(g) = 1$. Comme le rang de g est ≥ 3 , le théorème 4.2.2 montre que g représente 0. Dans tous les cas, on voit que f_1 représente 0 dans \mathbb{Q}_v ; comme le rang de f_1 est $n - 1$, l'hypothèse de récurrence montre que f_1 représente 0 dans \mathbb{Q} , i.e. que g représente a dans \mathbb{Q} ; comme h représente a , on en déduit bien que f représente 0, ce qui achève la démonstration. \square

Corollaire 5.2.2. *Soit $a \in \mathbb{Q}^*$. Pour que f représente a dans \mathbb{Q} il faut et il suffit qu'il en soit ainsi dans chacun des \mathbb{Q}_v .*

Preuve :

Cela résulte du théorème, appliqué à la forme $aZ^2 - f$. \square

Corollaire 5.2.3. (Meyer) *Une forme quadratique de rang ≥ 5 représente 0 si et seulement si elle est indéfinie (c'est-à-dire si elle représente 0 dans \mathbb{R}).*

Preuve :

D'après le théorème 4.2.2, une telle forme représente 0 dans chacun des \mathbb{Q}_p . Ce qui prouve le résultat. \square

Corollaire 5.2.4. *Soit n le rang de f . Supposons que $n = 3$ (resp. $n = 4$ et $d(f) = 1$). Si f représente 0 dans tous les \mathbb{Q}_v sauf un au plus, alors f représente 0.*

Preuve :

Supposons que $n = 3$. D'après le théorème 4.2.2, f représente 0 dans \mathbb{Q}_v si et seulement si l'on a :

$$(-1, -d(f))_v = \varepsilon_v(f). \quad (*)_v$$

Mais les deux familles $\varepsilon_v(f)$, $(-1, -d(f))_v$ vérifient la formule du produit (chapitre 1). On en conclut que si $(*)_v$ est vraie pour tout v sauf un au plus, $(*)_v$ est vraie pour tout v ; d'après le théorème de Hasse-Minkowski, la forme f représente 0.

Lorsque $n = 4$ et $d(f) = 1$, on raisonne de la même manière, l'égalité $(*)_v$ étant remplacée par :

$$(-1, -1)_v = \varepsilon_v(f).$$

Le corollaire est alors démontré. \square

5.3 Classification.

Théorème 5.3.1. *Soient f et f' deux formes quadratiques sur \mathbb{Q} . Pour que f et f' soient équivalentes sur \mathbb{Q} , il faut et il suffit qu'elles le soient sur chacun des \mathbb{Q}_p .*

Preuve :

La nécessité est triviale. Pour prouver la suffisance, on raisonne par récurrence sur le rang n de f et f' . Si $n = 0$, il n'y a rien à démontrer. Sinon, il existe $a \in \mathbb{Q}^*$ représenté par f , donc aussi par f' (corollaire 5.2.2). On a alors $f \sim aZ^2 + g$, $f' \sim aZ^2 + g'$. D'après le théorème 3.6.7, on a $g \sim g'$ sur \mathbb{Q}_v pour tout $v \in V$. L'hypothèse de récurrence montre alors que $g \sim g'$ sur \mathbb{Q} , d'où $f \sim f'$.

Corollaire 5.3.2. *Soient (r, s) et (r', s') les signatures de f et de f' . Pour que f et f' soient équivalentes, il faut et il suffit que l'on ait :*

$$d(f) = d(f'), \quad (r, s) = (r', s') \quad \text{et} \quad \varepsilon_v(f) = \varepsilon_v(f') \quad \text{pour tout } v \in V.$$

Preuve :

En effet, ces conditions expriment simplement que f et f' sont équivalentes sur chacun des \mathbb{Q}_v . \square

Remarque. Les invariants $d = d(f)$, $\varepsilon_v = \varepsilon_v(f)$ et (r, s) ne sont pas arbitraires. Ils vérifient les relations suivantes :

- (1) $\varepsilon_v = 1$ pour presque tout $v \in V$, et $\prod_{v \in V} \varepsilon_v = 1$;
- (2) $\varepsilon_v = 1$ si $n = 1$, ou si $n = 2$ et si l'image d_v de d dans $\mathbb{Q}_v^*/\mathbb{Q}_v^{*2}$ est égale à -1 ;
- (3) $r, s \geq 0$ et $r + s = n$;
- (4) $d_\infty = (-1)^s$;
- (5) $\varepsilon_\infty = (-1)^{s(s-1)/2}$.

Inversement, on a la proposition suivante, que nous allons admettre :

Proposition 5.3.3. *Soient d , $(\varepsilon_v)_{v \in V}$ et (r, s) vérifiant les relations (1) à (5) ci-dessus. Il existe alors une forme quadratique de rang n sur \mathbb{Q} , ayant pour invariants d , $(\varepsilon_v)_{v \in V}$ et (r, s) .*

Conclusion

Dans ce TER, nous avons donc démontré le "principe de Hasse" et montré qu'en définissant le rang, le discriminant, les invariants ε_v et la signature, on aboutit aux théorèmes de classification suivants :

i) Soit q une puissance d'un nombre premier impair. Deux formes quadratiques sur \mathbb{F}_q sont équivalentes si et seulement si elles ont même rang et même discriminant.

ii) Soit p un entier premier. Deux formes quadratiques sont équivalentes sur \mathbb{Q}_p si et seulement si elles ont même rang, même discriminant et même invariant ε_p .

iii) Deux formes quadratiques sont équivalentes sur \mathbb{R} si et seulement si elles ont même rang et même signature.

iv) Deux formes quadratiques sont équivalentes sur \mathbb{Q} si et seulement si elles sont équivalentes sur chacun des \mathbb{Q}_v , $v \in V$.

Bibliographie

SERRE J.P., Cours d'arithmétique, P.U.F.

DESCOMBES R., Éléments de théorie des nombres, P.U.F.

AMICE Y., Les nombres p -adiques, P.U.F.

MORELLE Lucile

BEAUDUT Lucie

M2 Formation des professeurs agrégés

Février 2012



RAPPORT DE STAGE

Conseillère pédagogique : Mme. Delphine Cheymol

Lycée Parc-de-Vilgenis
80, rue de Versailles
91305 Massy cedex

PRESENTATION DU STAGE

Pendant dix semaines, nous avons effectué un stage pédagogique filé, au Lycée Parc de Vilgénis de Massy, avec pour conseillère pédagogique Mme Delphine Cheymol, Professeur agrégée de mathématiques. Nous avons donc assisté ou sommes intervenues, une demi-journée par semaine, au sein de quelques 'une de ses classes (Terminales S SI et BTS).

Pour un peu plus d'expérience, nous avons également eu la chance de pouvoir assister à un TP en classe de Seconde, avec Mme Anne Redouté-Metrich et à des cours en classes de Premières S et STG, avec Mr Jacques Tailleux.

Dans un premier temps, nous allons présenter l'établissement et plus particulièrement le département de mathématiques. Ensuite, nous poursuivrons avec le développement de notre phase d'observation, puis nous enchaînerons avec notre prise en charge de la classe. Nous insisterons personnellement, sur une de nos interventions, et donnerons nos conclusions respectives.

PRESENTATION DE L'ETABLISSEMENT

Sur un parc de 8 hectares achetés en 1963 par la commune de Massy à la compagnie Air France, le lycée Parc de Vilgénis ouvre ses portes en septembre 1967, avec un total de 295 élèves répartis en 12 classes. L'équipe éducative, comprenant une trentaine de professeurs, se met en place autour du proviseur M.BURLIN, et de son adjoint, M.SIMON.

Aujourd'hui, l'équipe du lycée Parc de Vilgénis, autour du proviseur M.BISET, a la volonté d'accompagner les élèves et les étudiants afin qu'ils expriment au mieux toutes leurs compétences. Le lycée Parc de Vilgénis offre à plus de 2000 élèves et étudiants, outre ses formations générales traditionnelles, un large éventail de sections spécialisées, du Bac Pro au BTS ainsi que des classes préparatoires aux grandes écoles et des licences professionnelles.

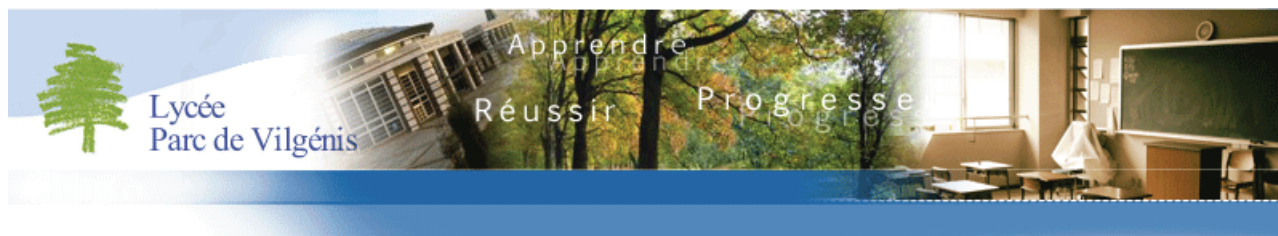
Le lycée Parc de Vilgénis propose une pédagogie innovante et personnalisée, un accompagnement éducatif attentif qui optimise la valorisation des capacités de chacun de ses élèves. Une approche par "projet personnel" vise à les rendre ambitieux en tenant compte de leurs souhaits et besoins. Les enseignants ont d'ailleurs particulièrement à cœur d'aider les élèves dans leurs choix d'orientation, tout au long de leur scolarité. En outre, la carte des formations est riche et ouverte sur le monde des entreprises afin de permettre à chacun une intégration professionnelle réussie.

De plus, le lycée Parc de Vilgénis possède des outils d'enseignement modernes et performants. Grâce à son réseau informatique, élèves et professeurs peuvent accéder aux ressources de l'Internet et de l'Intranet depuis environ 1000 postes de travail. L'établissement dispose également d'une salle de conférences dotée de réception numérique et de laboratoires : langues, multimédia, mathématiques. Un gymnase et un plateau d'évolution sportive se situent également à l'intérieur du campus.

Enfin, le cadre du lycée Parc de Vilgénis est très agréable. En effet, des bâtiments spacieux et rénovés situés dans un grand espace arboré accueillent avec aisance et confort ses élèves et étudiants, ainsi que les personnes travaillant sur le campus. Le lycée est par ailleurs d'un accès facile, au cœur d'un réseau de communications, à quelques minutes de la gare.

PRESENTATION DU DEPARTEMENT DE MATHÉMATIQUES

Le département de mathématiques est né en juin 2002, sous l'impulsion de Monsieur l'Inspecteur Pierre Michalack et de Monsieur le Proviseur Michel Bouchaud. Le lycée Parc de Vilgénis est un établissement important avec une bonne vingtaine de professeurs de mathématiques.



Au fil de leur expérience, les professeurs de mathématiques ont pu constater que cette structure a permis une meilleure assise, pour obtenir :

- du matériel auprès du conseil régional, 2 salles 20 postes et 3 salles vidéo entièrement dédiées aux mathématiques.
- des heures complémentaires élèves et professeurs : les professeurs de mathématiques travaillent dans d'excellentes conditions.
- une large reconnaissance de la part de la Mission à l'Initiative Pédagogique.

Par le caractère formel du département l'engagement de chacun des professeurs est devenu plus conséquent.

Les professeurs de mathématiques du lycée Parc de Vigénis travaillent essentiellement sur trois grands axes :

- le travail d'équipe : le travail en commun est très bénéfique et enrichissant pour la réflexion personnelle et pour la créativité. De plus, il apporte une certaine sécurité pour les nouveaux arrivants car devoir aborder pour la première fois le programme d'une classe n'est pas évident : le travail en commun permet de bénéficier des idées et des connaissances des autres. Il permet également aux professeurs d'effectuer un travail plus efficace pour les élèves.
- le calcul formel et les mathématiques expérimentales : à notre époque l'ordinateur associé à un calculateur formel couplé à un traitement de texte est notre outil d'expérience. Les professeurs utilisent pour leur part un logiciel nommé SWP (Scientific Work Place). Dans ce contexte expérimental, l'élève observe, interprète, formalise, échange et, dégagé des contraintes du calcul, laisse place à la réflexion. L'élève est plus impliqué.
- la promotion des mathématiques : plusieurs axes ont été développés afin de promouvoir les mathématiques au sein du lycée et à l'extérieur, comme la venue de chercheurs, la participation des élèves à des concours mathématiques (par exemple le "kangourou des mathématiques"), ou encore la création d'un club de mathématiques (dans le cadre du projet "maths en jean" (**M**éthode d'**A**pprentissage des **T**héories mathématiques en **J**umelant des **E**tablissements pour une **A**pproche **N**ouvelle du **S**avoir) en collaboration avec l'université d'Orsay).

DEROULEMENT DU STAGE FILE

Nous avons effectué ce stage avec Mme Cheymol, Professeur agrégée de mathématiques, pendant 10 semaines, à compter d'une demi-journée par semaine.

Nous avons commencé, les deux premières semaines, par une phase d'observation, au sein d'une classe de BTS et d'une classe de Terminale S SI. Celle-ci sera développée dans le paragraphe suivant.

Toutes les semaines qui ont suivies, nous avons eu la chance de pouvoir intervenir, en cours et TP, dans la classe de Terminale S (28 élèves) : une heure de cours en classe entière et deux heures de TP en demi-groupe par semaine.

Pour les TP :

Nous avons commencé notre intervention avec un TP, que nous avons préparé et proposé aux élèves. Nous sommes donc intervenues dans un demi-groupe chacune. Nous avons également rédigé une correction.

Les séances de TP qui ont suivies, nous n'avons pas eu à proposer de sujet, Mme Cheymol nous fournissait l'énoncé ainsi qu'un corrigé. Nous les préparions et faisons pour chacun d'eux, une grille de notation permettant d'évaluer les élèves. Nous intervenions à chaque séance, dans les deux demi-groupes (nous nous occupons de 4 à 5 élèves chacune).

Pour le cours :

Nous avons rédigé la première partie du cours sur les suites. Et l'une de nous a pu intervenir, pour le faire en classe entière.

Les séances qui ont suivies, nous n'avons pas eu à rédiger le cours. Nous sommes intervenues directement sur le cours de Mme Cheymol. Nous intervenions chacune une semaine sur deux. Nous préparions la partie du cours que nous devons exposer et les exercices que nous proposerions aux élèves.

Une correction de copies :

Dans le cadre du travail d'équipe, pendant la période des vacances scolaires, une feuille de révisions, avec correction, est fournie à tous les élèves de Terminale. A la rentrée, le professeur a pour habitude de donner un devoir d'une heure portant sur le thème de cette feuille d'exercices. Nous avons fait ce devoir pour la rentrée 2012, avec pour support la feuille d'exercices. Nous avons proposé un corrigé, en se référant aux modèles de rédaction (également fournis aux élèves), pour un maximum de rigueur. Les élèves ont fait ce devoir en notre absence, mais nous avons pu le corriger (à deux, nous en avons eu pour 4h de correction pour un paquet de 27 copies).

BILAN DU STAGE :

Mme Cheymol nous a donné un bon exemple du travail d'enseignant. Maintenant, nous avons une base solide, pour entrer dans le monde de l'enseignement. Nous pourrons aborder nos premiers cours plus sereinement : nous avons des idées sur la rédaction et l'exposé du cours, les séances d'exercices, les séances de TP, l'évaluation des élèves, la gestion de la classe... Cet éventail est beaucoup moins flou à présent.

PHASE D'OBSERVATION

GESTION DES BAVARDAGES (GRILLE D'ONSERVATION)

Nous avons rempli la grille d'observation concernant deux classes lors de nos deux premières séances de stage. Voici un condensé de nos observations.

- les BTS de 13h à 15h (en rouge)
- les terminales S de 15h à 17h (en bleu)

Quelle est l'attitude des élèves lorsqu'ils entrent dans la classe :

- ils restent debout ils s'assoient, et sortent leurs affaires ils bavardent

Comment le professeur attire-t-il l'attention de ses élèves avant de commencer le cours?

- rien Demande le silence dit : « chutttt »
- Mise en place d'un rituel : Commence le cours et les élèves comprennent
- autres : Parle plus fort

Comment le professeur gère-t-il les bavardages?

- laisse faire il s'interrompt et reprend l'élève il fini sa phrase et reprend l'élève ensuite
- s'interrompt, fixe les élèves concernés et attend qu'ils se taisent d'eux même
- dit : « chutttt » (de nombreuses fois)
- s'interrompt et reprend vivement l'élève responsable
-

Quelles phrases types le professeur utilise-t-il?

Pas de phrase type

S'il recourt à des sanctions, lesquelles?

- menace Pas de sanction : les élèves ne sont pas très bavards
→ Efficace : oui non
- change l'élève de place → Efficace : oui non
- demande le carnet de liaison à l'élève → Efficace : oui non
- le fait quitter la classe
-

Quand est-ce que les élèves bavardent le plus?

- au début du cours
- pendant que le professeur parle
- pendant qu'un élève est au tableau
- pendant qu'ils réfléchissent à un exercice -> les élèves sont le plus calme
- à la fin du cours
- après la pause

Le professeur adapte-t-il sa voix (ton, débit ?) selon l'agitation de la classe?

Mme Cheymol parle plus fort pour marquer une transition et lorsque les élèves bavardent et a tendance à parler plus bas lorsqu'il y a le silence.

Autres remarques : Mme Cheymol se déplace beaucoup dans la classe, se rapproche des élèves qui bavardent et vérifie que les élèves notent bien le cours ou la correction. De plus, elle attend que tous les élèves aient fini de noter pour passer à une remarque importante.

REMARQUES GENERALES

Nous avons pu observer lors d'une séance la façon de travailler de M. Taillez. Il écrit peu au tableau mais essaye de faire participer énormément les élèves à l'oral. Alors que Mme Cheymol vérifie que toutes les corrections au tableau et les remarques importantes soient notées par les élèves. Elle n'hésite pas à répéter plusieurs fois ce qui lui semble être difficile.

Concernant les méthodes de travail, Mme Cheymol fait des interrogations de cours très souvent mais les élèves sont prévenu. Une séance normale commence souvent par une correction d'exercice, un ou deux paragraphes de cours puis une application à l'aide d'exercices.

PRISE EN CHARGE DE LA CLASSE (LUCILE)

J'ai eu l'occasion d'intervenir seule, trois heures en classe entière pour faire le cours et une heure en demi-groupe pour un TP, dans une classe de terminale S SI.

Remarque : l'organisation du cours n'était pas à ma charge, en conséquence du stage filé.

J'ai choisi de développer ma première prise en charge de la classe, il s'agit d'une heure de TP sur ordinateur, en demi-groupe. J'ai préparé ce TP avec Lucie, et nous sommes intervenues pour un demi-groupe chacune, une heure.

LE CHOIX DU TP

Le TP devait porter sur la fonction logarithme. Nous avons donc sélectionné deux sujets que nous trouvions intéressants (sur Internet).

Le **premier sujet** choisi portait sur la courbe exponentielle, et une propriété particulière de ses tangentes, avec pour problématique : qu'est ce que le point où la tangente à la courbe $y = \exp(x)$ coupe l'axe des abscisses a de spécial ?

Le TP devant porter sur la courbe logarithme et non pas la courbe exponentielle, nous voulions modifier ce TP, qui aurait finalement porté sur la courbe logarithme, et une propriété de ses tangentes, avec pour nouvelle problématique : qu'est-ce que le point où la tangente à la courbe $y = \ln(x)$ coupe l'axe des ordonnées a de spécial ?

Le **deuxième sujet** choisi était un sujet d'annale de l'académie de Versailles, de l'année scolaire 2009-2010, trouvé sur le site de l'APMEP (Association des Professeurs de Mathématiques de l'Enseignement Public), et intitulé "Tangentes communes».

Mme Cheymol nous a guidées dans notre choix définitif, qui fut donc le deuxième sujet.

Au début, nous ne pensions pas modifier ce TP, mais pour faciliter le travail des élèves, le rappel d'une notion a été ajouté au sujet originel. Nous avons également ajouté une question théorique à la fin. De plus, nous avons rédigé une correction, que Mme Cheymol a envoyée aux élèves par e-mail, à la suite du TP.

DEROULEMENT DU TP

Mise en place : tous les élèves du demi-groupe entrent dans la salle, prennent un énoncé du TP (dans la pile pausée sur le bureau) et rejoignent un poste.

Remarque : les élèves ont un poste chacun.

Présentation : pour commencer, j'ai demandé l'attention de tous les élèves pour leur présenter rapidement le TP et évoquer, avant qu'ils débutent, le rappel concernant les tangentes communes, une notion qui n'était pas forcément acquise pour tous les élèves et nécessaire au bon déroulement du TP.

Le TP : la partie pratique du TP demandait l'utilisation du logiciel Geogebra, les élèves ont donc tous mis en route le logiciel, et ont répondu aux questions pratiques, sans trop de difficultés. Pendant cette période, je passais voir ce que faisait chacun des élèves.

C'est devenu plus compliqué pour eux et aussi pour moi lors de la résolution des questions théoriques. Tous les élèves ne parvenaient pas à répondre aux questions et je leurs ai donc donné des pistes sans leur donner la réponse tout de suite. Et cela à chaque élève en particulier, ce qui n'était pas vraiment la meilleure solution. Après avoir constaté que les questions étaient récurrentes, j'aurais dû penser à donner une piste pour y répondre, à tout le monde en même temps (en demandant leur attention au tableau, par exemple). Ce qui aurait permis de perdre moins de temps. J'ai finalement écrit une correction au tableau, mais un peu tard... Ensuite, un élève m'a pausé beaucoup de questions et j'ai eu du mal à le laisser pour m'occuper des autres. Si c'était à refaire, je ferais bien attention à ne pas rester trop longtemps avec un seul élève. Finalement, tous les élèves n'ont pas terminé la partie théorique du TP, mais toutes les questions pratiques ont été faites. Ils ont eu la correction dans les jours qui ont suivis.

Conclusion : la gestion de ce TP n'a pas été très facile. Je me suis rapidement rendue compte qu'il est difficile de répondre à toutes les demandes, d'être vraiment présent pour tous les élèves.

Remarque : suite à cette expérience, je me suis rendue compte qu'une évaluation complète et individuelle sur un TP est compliquée à mettre en place.

LES AUTRES SEANCES DE TP

Lors des séances de TP qui ont suivies, nous ne sommes plus intervenues seules. Nous avons travaillé en équipe (tutrice et stagiaires), notre présence permettant d'évaluer les élèves plus facilement (quatre à cinq élèves à évaluer par personne).

Ces évaluations ont été effectuées par le biais de grilles de notation, que Lucie et moi-même faisons pour chaque TP.

Pour le barème des TP, nous faisons attention à mettre au moins la moitié des points sur la partie pratique.

De plus, nous avons pu constater que les élèves vont rarement jusqu'au bout du TP proposé. C'est pourquoi, les dernières questions des TP étaient comptabilisées dans le barème soit par peu de points, soit en bonus.

Remarque : parfois il n'était finalement pas possible d'évaluer le TP, car les élèves rencontraient trop de difficultés.

BILAN

Globalement, les TP se sont bien passés.

La première séance, étant seule pour gérer le demi-groupe, était plus difficile que les autres, mais a été nécessaire, pour une vision réaliste du déroulement d'un TP.

Les autres séances m'ont permise de travailler sur l'évaluation des élèves.

Pendant les TP, les élèves ont été sérieux, ils travaillaient tous et étaient motivés (pour les TP qui ont été notés, tous les élèves avaient la moyenne).

Nous avons constaté que la plupart des élèves travaillent en binôme, bien qu'ils aient un poste par personne. Suite à cette constatation, une évaluation complète des TP, pour un enseignant, peut donc être mise en place, en faisant travailler les élèves par deux (par exemple).

Ce qui était bien : en ce qui concerne les mathématiques, je n'ai pas fait de grosses erreurs et j'étais à l'écoute des élèves.

Ce qu'il faut améliorer : il faudrait que je sois plus rigoureuse dans les corrections des exercices au tableau, que je parvienne à moduler d'avantage ma voix et à parler plus fort, pour animer un peu les cours.

PRISE EN CHARGE : LUCIE

J'ai choisi de m'attarder sur la première séance de cours : **Les Suites**

PREPARATION DU COURS :

Nous avons préparé en commun un polycopié sur les suites (cf en annexes) que j'ai pu projeter au tableau lors de la séance et distribuer aux élèves. Madame Cheymol avait au préalable fait quelques remarques et nous avait donné des conseils pour améliorer notre document, par exemple insister sur la différence entre u_n et (u_n) , ou mettre un guide d'utilisation de la calculatrice concernant le calcul des premiers termes d'une suite.

DEROULEMENT DE LA SEANCE :

La séance a eu lieu à 8h le mardi 3 janvier. J'ai suivi le polycopié projeté au tableau. Je suis allée très vite sur la définition d'une suite et les deux principaux modes de définition. J'ai traité trois exemples au tableau en les sollicitant. Je suis restée environ 10 min sur cette partie.

La phase sur l'utilisation de la calculatrice pour calculer les premiers termes fut plus difficile à gérer. Seulement quatre élèves possédaient une Casio, tous les autres élèves avaient une Texas Instrument 82. J'ai traité un exemple au tableau sur un simulateur de Texas Instrument 82. Comme prévu, je n'ai pas réussi à faire fonctionner tout de suite le logiciel, heureusement un élève a pu me dire quelle option changer pour que ça fonctionne. Je suis allée voir individuellement les élèves possédant une Casio. Au bout de 15 min, j'ai demandé aux élèves de ranger les calculatrices pour passer aux représentations graphiques.

Concernant la représentation d'une suite définie par son terme général, je leur ai expliqué comment tracer les premiers points et ils ont complété leur graphique en traçant un troisième point. Je suis allée assez vite aussi car je voulais passer plus de temps sur la représentation d'une suite définie par récurrence. Pour cela, je leur ai montré le principe de construction et ils ont eu à compléter le tracé de la première suite et faire deux autres tracés. Grâce aux conseils de Madame Cheymol, je leur ai fait remarquer qu'une suite définie par récurrence avec une fonction d'itération croissante n'est pas toujours croissante. Je leur ai laissé assez de temps pour tracer les graphiques sur leurs feuilles. De plus, avant d'afficher la correction je leur ai remontré plusieurs fois la construction.

Il restait 5 min avant la fin de la séance. J'ai fait quelques rappels sur les suites arithmétiques et géométriques. J'ai tracé deux colonnes au tableau et je leur ai demandé de me dire tout ce qu'ils se rappelaient sur ce sujet qu'ils ont vu en première. J'ai voulu que l'on vérifie si une suite donnée en exemple dans le cours était géométrique. Mais j'ai fait une erreur de raisonnement : après calcul on

trouvait : $\frac{w_{n+1}}{w_n} = \frac{1}{w_n}$, et j'ai conclut que la suite n'était pas géométrique. Heureusement, Mme Cheymol a pu corriger mon erreur lors de la séance suivante.

IMPRESSION SUR CETTE SEANCE :

Lors de cette séance, j'ai apprécié exposer le cours, je suis passée dans les rangs pour vérifier leur travail. Au tout début de l'heure, un élève m'a demandé de parler plus fort et d'écrire plus gros. Je pense que j'ai su répondre à sa demande durant l'heure et je pense m'être améliorée sur ce point lors des heures suivantes. Cependant je suis allée trop vite sur certaines notions, certains élèves étant réactifs et répondant à mes questions je ne me suis pas rendu compte que d'autres ont besoin de plus de temps pour comprendre les raisonnements. J'ai aussi effectué des abus de langage que l'on ne peut se permettre à leur niveau, cela s'explique car j'ai tendance à parler trop vite.

BILAN :

Lors des séances suivantes j'ai fait moins d'abus de langage en mathématiques et je parlais moins vite mais je dois encore être plus rigoureuse et parler plus posément.

Concernant la discipline, je constate que j'ai beaucoup de mal à reprendre un élève qui bavarde ou qui ne suit pas. J'ai fait quelques tentatives en reprenant toute la classe en même temps, mais qui sont restées pour la plupart inefficaces. Heureusement, les élèves de terminale S sont assez studieux, mais je dois m'améliorer sur ce point car ce n'est pas le cas de toutes les classes.

REMERCIEMENTS

Nous remercions tout d'abord M. Biset pour nous avoir accueillies chaleureusement dans son établissement, ainsi que Mme Issard Roch pour nous avoir trouvé ce stage et Mme Levi pour les séances de suivi de stage.

De plus nous souhaitons remercier vivement Mme Delphine Cheymol qui nous a consacré beaucoup de temps, nous a fait des comptes rendus détaillés de nos séances et nous a donné énormément de conseils. Nous remercions également les élèves de terminale S, avec qui nous avons passé la plupart de nos séances et qui ont été très agréables avec nous.

SOMMAIRE

Présentation du stage	2
Présentation de l'établissement	2
présentation du département de mathématiques	3
Déroulement du stage filé.....	4
Bilan du stage :	4
Phase d'observation.....	5
Gestion des bavardages (grille d'onservation).....	5
Remarques générales.....	6
Prise en charge de la classe (Lucile)	6
Le choix du TP.....	6
Déroulement du TP	7
Les autres séances de TP.....	7
Bilan.....	8
Prise en charge : Lucie.....	8
Préparation du cours :	8
Déroulement de la séance :	8
Impression sur cette séance :	9
Bilan :	9
Remerciements	10

Annexes.....	12
Enoncé du TP Geogebra Logarithme.....	12
Correction TP Geogebra Logarithme.....	14
3 tp et leurs grilles de notation.....	17
Cours sur les suites.....	23
Interrogation de rentrée : Enoncé.....	29
Interrogation de rentrée : Correction.....	30
Grille de notes.....	33

ANNEXES

ENNONCE DU TP GEOGEBRA LOGARITHME

La Courbe Logarithme

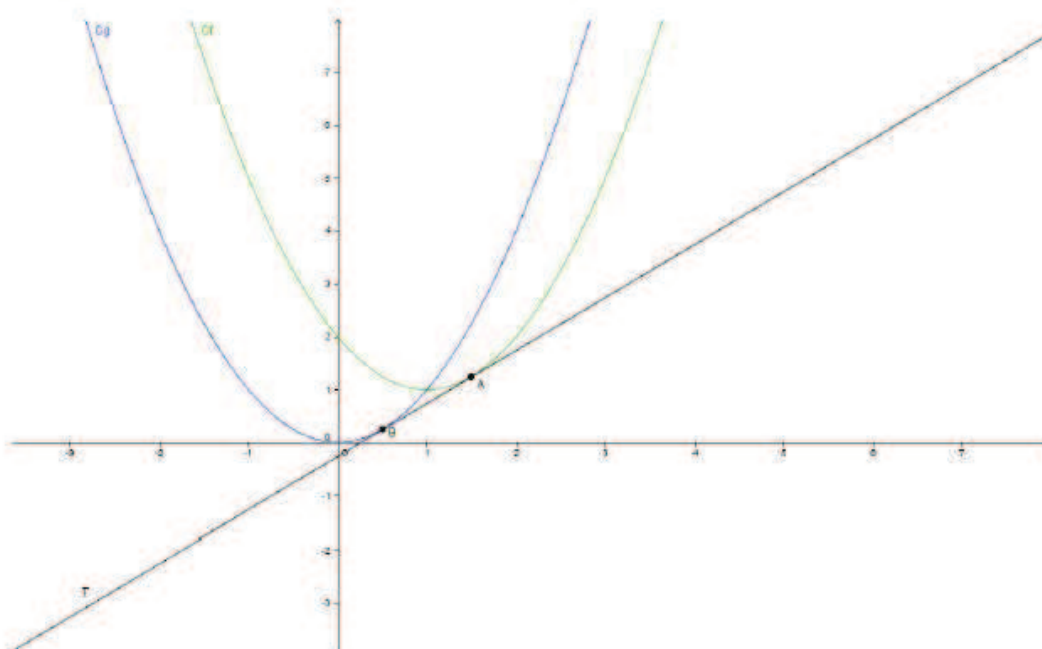
Pour répondre expérimentalement aux questions, vous allez utiliser le logiciel *Geogebra*.

A propos des tangentes communes.

Rappel :

Soit f et g deux fonctions définies et dérivables respectivement sur les ensembles D_f et D_g et soit C_f et C_g leurs courbes représentatives respectives. On dit que T est tangente commune à C_f et C_g si T est à la fois tangente à C_f en un point A de C_f et tangente à C_g en un point B de C_g .

Vous pouvez observer cette propriété sur l'image suivante :



Exercice :

Le plan est muni d'un repère orthonormé (O, \vec{i}, \vec{j}) . On considère les courbes C et Γ représentant respectivement les fonctions f et g définies par : $f(x) = \ln(x)$ sur \mathbb{R}^{*+} et $g(x) = x^2$ sur \mathbb{R} .

On se propose d'étudier le nombre de tangentes communes à ces deux courbes.

1. Représentez les courbes C et Γ . Placez sur la courbe C un point A . Estimez graphiquement le nombre de points d'intersection de la tangente T en A à la courbe C avec la courbe Γ .

2. Les courbes C et Γ semblent avoir combien de tangentes communes ? Pour quelles valeurs de l'abscisse de A cette situation se réalise-t-elle ?

3. Ecrivez une équation de la tangente T au point A d'abscisse a à la courbe C . Ecrivez ensuite une condition nécessaire et suffisante pour que cette droite ne possède qu'un point d'intersection avec la courbe Γ .

4. Sur le graphique, représentez graphiquement la fonction h définie sur \mathbb{R}^* par :

$$h(x) = 1 - \frac{1}{4x^2}.$$

5. a. Conjecturez graphiquement les solutions de l'équation $\ln(x) = h(x)$?
b. Démontrez vos conjectures.

CORRECTION TP GEOGEBRA LOGARITHME

Remarque : On admet dans ce TP qu'une tangente à la parabole d'équation $y=x^2$ est une droite qui coupe cette parabole en un unique point.

- 1) Graphiquement on constate que :
 - Si $x_A \in]-\infty; 0.34[$: Il y a deux points d'intersection entre la courbe Γ et la tangente T.
 - Si $x_A \approx 0.35$: Il y a un point d'intersection entre la courbe Γ et la tangente T.
 - Si $x_A \in]0.36; 2.56[$: Il n'y a pas de point d'intersection entre la courbe Γ et la tangente T.
 - Si $x_A \approx 2.56$: Il y a un point d'intersection entre la courbe Γ et la tangente T.
 - Si $x_A \in]2.58; \infty[$: Il y a deux points d'intersection entre la courbe Γ et la tangente T.

- 2) Il semble que C et Γ ont 2 tangentes communes aux points de C d'abscisses 0.35 et 2.56.

- 3) L'équation de la tangente T au point d'abscisse A est : $y = f'(a)(x - a) + y_A$.
 $f'(a) = (\ln(a))' = \frac{1}{a}$ et $y_A = \ln(a)$. Donc on obtient T : $y = \frac{1}{a}x + \ln(a) - 1$.
 La courbe Γ et la droite T possèdent un seul point d'intersection si et seulement si l'équation $\frac{1}{a}x + \ln(a) - 1 = g(x)$ admet une unique solution.
 C'est-à-dire si l'équation $x^2 - \frac{1}{a}x + \ln(a) - 1 = 0$ (*) admet une unique solution. On obtient un polynôme du second degré, il suffit de calculer son discriminant.
 $\Delta = \frac{1}{a^2} + 4(\ln(a) - 1)$. L'équation (*) admet une unique solution $\Leftrightarrow \Delta = 0 \Leftrightarrow \ln(a) = 1 - \frac{1}{4a^2}$.
 On est donc ramener à résoudre l'équation $\ln(x) = 1 - \frac{1}{4x^2}$

- 4) Après avoir tracé la courbe de h que l'on note H, on observe que H et C ont deux points d'intersection I_1 et I_2 .

- 5) a) Conjecture graphique : l'équation $\ln(x) = h(x)$ possède deux solutions, en $\alpha \approx 0.35$ et $\beta \approx 2.56$.

b) Montrons que l'équation $\ln(x) = h(x)$ possède deux solutions.

1^{ère} idée : Résoudre directement $\ln(x) = 1 - \frac{1}{4x^2}$. Mais on se rend compte que l'on ne sait pas résoudre.

2^{ème} idée : Etudier la fonction $F(x) = \ln(x) - 1 + \frac{1}{4x^2}$. C'est-à-dire dresser le tableau de variation de F et étudier les solutions de $F(x) = 0$.

Domaine de définition de F : $]0; +\infty[$.

F est dérivable sur $]0; +\infty[$ en tant que somme de fonctions dérivables, car les fonctions

$x \mapsto \ln(x)$ et $x \mapsto 1 + \frac{1}{4x^2}$ sont dérivables sur $]0; +\infty[$.

$$F'(x) = \frac{1}{x} - \frac{1}{2x^3} = \frac{2x^2 - 1}{2x^3}$$

$F'(x) = 0 \Leftrightarrow 2x^2 - 1 = 0 \Leftrightarrow x = \frac{1}{\sqrt{2}}$ ou $-\frac{1}{\sqrt{2}}$. Or d'après le domaine de définition, il reste $x = \frac{1}{\sqrt{2}}$.

Calcul utile pour dresser le tableau de variation :

On cherche le signe de F' sur $]0; +\infty[$: $F'(x) \geq 0 \Leftrightarrow \frac{2x^2 - 1}{2x^3} \geq 0 \Leftrightarrow 2x^2 - 1 \geq 0$ (car $x > 0$) $\Leftrightarrow x \leq -\frac{1}{\sqrt{2}}$ (impossible car $x > 0$) ou $x \geq \frac{1}{\sqrt{2}} \Leftrightarrow x \geq \frac{1}{\sqrt{2}}$.

$$F\left(\frac{1}{\sqrt{2}}\right) = \ln\left(\frac{1}{\sqrt{2}}\right) - 1 + \frac{1}{4\left(\frac{1}{\sqrt{2}}\right)^2} = -\frac{1}{2}\ln(2) - 1 + \frac{1}{2} = -\frac{1}{2(\ln(2) + 1)} < 0$$

Calcul de la limite de F en 0 :

$$\lim_{x \rightarrow 0} F(x) = \lim_{x \rightarrow 0} \left[\ln(x) - 1 + \frac{1}{4x^2} \right] \text{ Forme indéterminée de la forme } \infty - \infty.$$

$$\ln(x) - 1 + \frac{1}{4x^2} = \frac{1}{x^2 \left(x^2 \ln(x) - x^2 + \frac{1}{4} \right)} \text{ or } \lim_{x \rightarrow 0} x^2 \ln(x) = 0 \text{ par croissance comparée et}$$

$$\lim_{x \rightarrow 0} \left(x^2 \ln(x) - x^2 + \frac{1}{4} \right) = \frac{1}{4} \text{ donc } \lim_{x \rightarrow 0} \frac{1}{x^2 \left(x^2 \ln(x) - x^2 + \frac{1}{4} \right)} = \frac{1}{\frac{1}{4}} = 4 \text{ par limite de somme.}$$

$$\lim_{x \rightarrow 0} \frac{1}{x^2} = +\infty$$

Donc $\lim_{x \rightarrow 0} F(x) = +\infty$ par limite de produit.

Calcul de la limite de F en $+\infty$:

$$\lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} \left[\ln(x) - 1 + \frac{1}{4x^2} \right] = +\infty \text{ par limite de somme car}$$

$$\lim_{x \rightarrow +\infty} \ln(x) = +\infty \text{ et } \lim_{x \rightarrow +\infty} \left(x^2 \ln(x) - x^2 + \frac{1}{4} \right) = 0.$$

Tableau de variation :

x	0	α	$\frac{1}{\sqrt{2}}$	β	$+\infty$
F'		-	0	+	
F	$+\infty$	\searrow	0	\searrow	$-\frac{1}{2(\ln(2) + 1)} < 0$
			\nearrow	0	\nearrow
					$+\infty$

Conclusion de l'étude de F :

La fonction F est continue sur $\left]0; \frac{1}{\sqrt{2}}\right]$ et strictement décroissante donc F est une bijection sur $\left]0; \frac{1}{\sqrt{2}}\right]$ à valeur sur $\left[-\frac{1}{2}(\ln(2) + 1); +\infty\right[$. Or $0 \in \left[-\frac{1}{2}(\ln(2) + 1); +\infty\right[$, donc d'après le théorème de bijection, l'équation $F(x)=0$ admet une unique solution sur $\left]0; \frac{1}{\sqrt{2}}\right]$.

De même la fonction F est continue sur $\left[\frac{1}{\sqrt{2}}; +\infty\right[$ et strictement croissante donc F est une bijection sur $\left[\frac{1}{\sqrt{2}}; +\infty\right[$ à valeur sur $\left[-\frac{1}{2(\ln(2) + 1)}; +\infty\right[$. Or $0 \in \left[-\frac{1}{2}(\ln(2) + 1); +\infty\right[$, donc d'après le théorème de bijection, l'équation $F(x)=0$ admet une unique solution sur $\left[\frac{1}{\sqrt{2}}; +\infty\right[$.

Donc l'équation $F(x)=0$ admet deux solutions. C'est-à-dire que notre conjecture faite à la question précédente est vérifiée.

Conclusion de l'exercice :

On a donc montré que la condition nécessaire et suffisante obtenue à la question 3) est vérifiée pour deux valeurs de a. On ne sait pas calculer explicitement ces valeurs mais on sait

grâce à l'étude précédente que l'une appartient à l'intervalle $\left]0; \frac{1}{\sqrt{2}}\right]$ et l'autre appartient à l'intervalle $\left[\frac{1}{\sqrt{2}}; +\infty\right[$ (l'étude graphique permet de donner une valeur approchée à ces deux abscisses). Il existe donc deux points tels que les courbes C et Γ ont une tangente commune.

TS3 TP : expression du terme de rang n Fiche élève

d'une suite récurrente

On considère la suite récurrente (u_n) de premier terme $u_0 = 0$ et telle que, pour tout entier naturel n ,
 $u_{n+1} = u_n + 2n + 1$

1) En utilisant un tableur, SWP ou une calculatrice calculer et représenter graphiquement les 21 premiers termes de cette suite. Le nuage de points obtenus a-t-il une particularité? Si oui, laquelle ?

Appeler le professeur pour une vérification de la particularité trouvée

2) n étant un entier naturel donné, on peut calculer la valeur de u_n si on connaît la valeur de u_{n-1} .
 On voudrait à présent pouvoir calculer, pour n'importe quelle valeur de n , la valeur de u_n sans pour autant connaître u_{n-1} , pour cela il faudrait disposer d'une formule donnant u_n en fonction de n .

a) A l'aide des observations faites dans la première question, conjecturer une formule donnant, pour n'importe quelle valeur de n , u_n en fonction de n .

Appeler le professeur pour une vérification de la formule trouvée

b) Démontrer cette formule.

3)

a) Emettre une conjecture concernant la monotonie de la suite (u_n)

b) démontrer cette conjecture

4)

a) La suite (u_n) semble t-elle minorée ? majorée ? bornée ?

b) démontrer vos conjectures

Barème TP : expression du terme de rang n d'une suite récurrente

Elèves							
1) Formule de récurrence	1						
Aller jusqu'au rang 20	1						
Savoir tracer un nuage de points	2						
Trouver l'équation de la courbe : méthode 1							
Remarquer que la courbe est une parabole	1						
Trouver les coefficients de l'équation de la courbe	2						
Trouver l'équation de la courbe : méthode 2							
Savoir utiliser le logiciel pour tracer la courbe de tendance	2						
Trouver l'équation de la courbe	1						
2) a) Conjecturer la formule	1						
b) Bonne hypothèse de récurrence	0,5						
Initialisation	1						
Hérédité	2						
Conclusion	0,5						
3) a) Conjecture	1						
b) Justification	3						
4) a) Minorée	1						
Non majorée	1						
b) Justification de minorée	2						
Justification de non majorée	1						
Total	21						

TS TP : inversion Enoncé

Le plan complexe P est muni d'un repère orthonormé (O, \vec{u}, \vec{v})

$$F : P - \{O\} \rightarrow P - \{O\}$$

On considère l'application

$$M(z) \leftrightarrow M' \left(z' = \frac{-1}{\bar{z}} \right)$$

Le but du TP est de déterminer l'image d'une droite par cette application F

0.0.1 Partie A : avec géogébra

On veut conjecturer l'image de la droite D d'équation $x = 1$.

1)

a) Tracer D .

b) Comment s'écrit l'abscisse d'un point quelconque M de D ?

Créer un tel point M en tapant son abscisse dans la barre de saisie.

c) Créer le point M' image du point M par F .

2) Conjecturer l'image C de la droite D .

Demander au professeur de vérifier vos conjectures

0.0.2 Partie B : preuves

1) Déterminer la forme algébrique de l'abscisse z' du point M' image de M par F lorsque M appartient à D .

2) Démontrer que si $M \in D$ alors $M' \in C$. (Ainsi $F(D) \subset C$)

rappel : Un point de coordonnées $(x; y)$ appartient au cercle de centre $I(x_I, y_I)$ et de rayon $r \Leftrightarrow (x - x_I)^2 + (y - y_I)^2 = r^2$

3) Etude de la réciproque

a) Montrer que si $M'(z')$ appartient à C alors M a pour abscisse $z = \frac{2}{1 - e^{-i\theta}}$ où $\theta \neq 0(2\pi)$.

rappel : Un point s'abscisse z appartient au cercle de centre $I(z_I)$ et de rayon $r \Leftrightarrow z = z_I + re^{i\theta}$ avec $\theta \in \mathbb{R}$

b) Déterminer la partie réelle du complexe $\frac{2}{1 - e^{-i\theta}}$ où $\theta \neq 0(2\pi)$.

c) Conclure.

Bonus 1) La conjecture faite pour l'image de la droite D reste-t-elle valable pour toute droite parallèle à D ?

2) La conjecture faite pour l'image de la droite D reste-t-elle valable pour toute droite D du plan ?

Barème TP : INVERSION

A					
1) Tracer D			1		
Ecrire l'affixe de M			0,5		
Créer le point M sur géogebra			1		
Créer le point M' image de M par F sur géogebra			1		
2) Conjecturer l'image C de la droite D			1,5		
B					
1) Forme algébrique de z' du point M'			0,5		
2) Image de D par F est C			1		
3) Forme de l'affixe z de M si M' appartient à C			1		
Déterminer la partie réelle de z			1,5		
Conclure			1		
Bonus					
1) La conjecture reste vraie pour toute droite parallèle à D			0,5		
2) La conjecture reste vraie pour toute droite D du plan			0,5		
Total			11		

TS

TP : complexes et suites

Énoncé

Le plan complexe est muni d'un repère orthonormal direct (O, \vec{u}, \vec{v}) .

On considère la suite (A_n) des points du plan d'affixes (a_n) telles que
$$\begin{cases} \forall n \in \mathbb{N} \ a_{n+1} = \left(\frac{i}{2}\right) a_n \\ a_0 = 8 \end{cases}$$

1) Dessiner A_0, A_1, A_2, A_3, A_4 .

2) On considère pour tout entier naturel n , $d_n = A_n A_{n+1}$.

Conjecturer la nature de la suite (d_n) .

3) On considère pour tout entier naturel n , $L_n = \sum_{k=0}^n d_k$.

Ecrire un algorithme qui demande à l'utilisateur la valeur de n et qui renvoie la valeur L_n .

A l'aide de cet algorithme, que peut-on conjecturer quant à la convergence de la suite (L_n) ?

4) Démontrer les conjectures des questions 2 et 3.

5) Traduire géométriquement la limite de la suite (L_n) .

6) Quelle est la nature des triangles $OA_n A_{n+1}$? Justifier.

Barème TP : Complexes et suites

1) Dessiner A_0, \dots, A_4	0,5			
Utiliser la barre de saisi pour rentrer la formule	0,5			
2) Représenter les longueurs d_n sur geogebra	0,5			
Conjecturer de d_n géométrique	0,5			
3) Ecrire l'algorithme à la main	1			
Rentrer l'algorithme sur la calculatrice	1			
Conjecture de la convergence de L_n	0,5			
4) Calcul de la raison de d_n	1,5			
Calcul du premier terme	0,5			
Simplification de L_n	1			
Justification de la limite de L_n	1			
5) Traduction géométrique de la limite de L_n	1			
6) Nature des triangles $O A_n A_{n+1}$	0,5			
Justification de la nature des triangles	1			
Total	11			

Les suites

I. Généralités

1. Définition

Définition. Une suite réelle est une fonction de \mathbb{N} (ou d'une partie de \mathbb{N}) dans \mathbb{R} .

$$\begin{aligned} f : \mathbb{N} &\rightarrow \mathbb{R} \\ n &\longmapsto f(n) \end{aligned}$$

Notation. On note : $f(n) = u_n$, appelé terme général de la suite et on appelle n le rang du terme u_n .

La suite est alors notée u ou (u_n) ou $(u_n)_{n \in \mathbb{N}}$.

Remarque : Une suite peut être définie qu'à partir d'un certain rang p , on la note alors $(u_n)_{n \geq p}$.

Attention : $u_n \neq (u_n)$ (de la même façon, $f(x) \neq f$).

Il y a deux modes principaux de définition possibles d'une suite :

- La suite peut être définie par récurrence :

$$\begin{cases} u_0 \text{ est donné.} \\ \forall n \in \mathbb{N}, u_{n+1} = F(u_n) \text{ avec } F \text{ une fonction numérique donnée.} \end{cases}$$
- La suite peut être définie par son terme général :

$$\forall n \in \mathbb{N}, u_n = f(n), \text{ où } f \text{ est une fonction explicite.}$$

Exemples : Dire quel est le mode de définition de ces suites et calculer leurs premiers termes (à la main, puis à la calculatrice) :

Pour tout entier naturel n , on définit les suites u, v, w par :

$$\begin{cases} u_0 = -1/2 \\ u_{n+1} = 1/u_n \end{cases}$$

$$\begin{cases} v_n = \sqrt{n} \end{cases}$$


$$\begin{cases} w_0 = 1 \\ w_{n+1} = 2w_n + n \end{cases}$$

TI 83+


Accès au mode suites

<p>Touche MODE</p> <p>Choisir sur la troisième ligne Seq et appuyer sur ENTER</p> <p>Choisir sur la quatrième ligne Dot et appuyer sur ENTER</p>	
---	---


1) Utiliser le terme général

<p>On a $u_n = -4 + 0,8n$ et $v_n = 0,1 \times (-1,5)^n$</p> <ul style="list-style-type: none"> • Touche Y=. On obtient l'écran suivant (saisir éventuellement nMin = 0). Introduire la suite u. Pour la variable n, utiliser la touche X,T,θ,n. Valider avec la touche ENTER. Même opération pour la suite v. • Régler les paramètres de la table comme sur l'écran ci-contre Instruction TBL SET (touches 2nd et WINDOW). • Afficher la table de valeurs Instruction TABLE (touches 2nd et GRAPH). <p>→ Les suites u et v étant définies par une relation explicite, la donnée de u(nMin) et de v(nMin) n'est pas obligatoire.</p> <p>⇒ si des valeurs de u(nMin) et de v(nMin) sont saisies, elles apparaissent dans la table sans conséquences sur les autres valeurs de u.</p>	
--	--

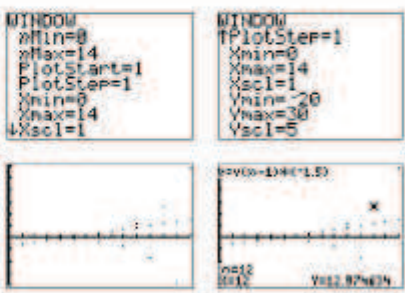
2) Utiliser la relation de récurrence

<p>Sur la calculatrice il faut exprimer u_n en fonction de u_{n-1}.</p> <p>Ainsi, $u_{n+1} = u_n + 0,8$ devient $u(n) = u(n-1) + 0,8$</p> <p>et $v_{n+1} = v_n \times (-1,5)$ devient $v(n) = v(n-1) \times (-1,5)$</p> <ul style="list-style-type: none"> • Touche Y= puis CLEAR pour effacer la suite déjà saisie. Introduire les deux relations de récurrence → n s'obtient avec la touche X,T,θ,n. → u et v s'obtiennent avec les touches 2nd 7 ou 2nd 8. Compléter u(nMin) et de v(nMin) par -4 et 0,1. Valider avec ENTER. • Régler les paramètres et afficher la table de valeurs la table comme ci-contre. 	
--	---

3) Afficher un terme de la suite

<p>Retour à l'écran de calcul. Instruction QUIT (touches 2nd et MODE).</p> <p>Saisir les séquences suivantes :</p> <p>2nd 7 (0 3 1) ENTER et 2nd 8 (1 2 5) ENTER.</p>	
--	---

4) Représentation graphique


<ul style="list-style-type: none"> • Ouvrir la fenêtre d'affichage : Touche WINDOW. Régler les paramètres comme sur les écrans ci-contre. Touches ▲ et ▼ pour passer d'une ligne à l'autre. Touche GRAPH pour obtenir la représentation ci-contre • La touche TRACE permet d'obtenir les coordonnées des points représentés Les touches ← et → permettent de passer d'un point à l'autre. Les touches ▲ et ▼ permettent de passer d'une suite à l'autre. 	
--	--

Casio Graph 35+

Accès au mode suites

Touche MENU icône Appuyer sur **EXE**

La calculatrice note a_n et b_n les deux suites au lieu de u_n et v_n .



1) En utilisant le terme général

On a $a_n = -4 + 0,8n$ et $b_n = 0,1 \times (-1,5)^n$

- On obtient l'écran suivant.

Sélectionner le sous-menu TYPE (touche **F3**) et choisir l'instruction an (touche **F1**).

Introduire la suite a . Pour la variable n , utiliser l'instruction \bar{n} (touche **F4**) Valider avec la touche **EXE**

Même opération pour la suite b Valider avec la touche **EXE**.

→ *Commentaire* : Les suites a et b sont ici définies par une relation explicite, la donnée de a_0 et b_0 n'est donc pas obligatoire

- Régler les paramètres de la table comme sur l'écran ci-contre

Instruction **RANG** (touche **F5**).

- Afficher la table de valeurs

Instruction **TABLE** (touche **F6**)

Recursion

an: -4+0,8n

bn: 0,1x(-1,5)^n

DEL CLR OPT F5/F6 OPT F5/F6

Select Type

Element: B

Column: 1Chan: Bnc

Column: 2Chan: CChan: Bnc

END (EXIT)

Recursion

an: -4+0,8n

bn: 0,1x(-1,5)^n

DEL CLR OPT F5/F6 OPT F5/F6

Table: Range n

Start: 1

End: 114

n	a _n	b _n
1	-3,2	0,15
2	-2,4	-0,15
3	-1,6	0,225
4	-0,8	-0,3375

DEL CLR OPT F5/F6 OPT F5/F6

TABLE

END (EXIT)

2) En utilisant la relation de récurrence

On a $u_{n+1} = u_n + 0,8$ soit $a_{n+1} = a_n + 0,8$
 et $v_{n+1} = v_n \times (-1,5)$ soit $b_{n+1} = b_n \times (-1,5)$

- Sélectionner le sous-menu TYPE (touche **F3**) et choisir l'instruction a_{n+1} (touche **F2**)

Introduire les deux relations de récurrence : utiliser l'instruction $n\Delta$, (touche **F4**) et choisir a_n (touche **F2**) et b_n (touche **F3**)

Valider avec la touche **EXE**.

- Régler les paramètres de la table comme ci-contre.

- Afficher la table de valeurs comme ci-contre.

Recursion

an: a_{n+1}=a_n+0,8

bn: b_{n+1}=b_nx(-1,5)

DEL CLR OPT F5/F6 OPT F5/F6

Table: Range an

Start: 1

End: 114

n	a _n	b _n
1	-4	0,15
2	-3,2	-0,15
3	-2,4	0,225

DEL CLR OPT F5/F6 OPT F5/F6

TABLE

END (EXIT)

3) Représentation graphique

- Régler la fenêtre d'affichage : instruction **V-Window** (touches **SHIFT F3**)

Régler les paramètres d'affichage comme sur les écrans ci-contre. Touches **▲** et **▼** pour passer d'une ligne à l'autre.

Puis touche **EXIT** puis instruction **TABL** (touche **F6**).

Puis choisir **G-PLT** (touche **F6**).

On obtient la représentation ci-contre

- L'instruction **TRACE** (touche **F1**) permet d'obtenir les coordonnées des points représentés.

Les touches **◀** et **▶** permettent de passer d'un point à l'autre.

Les touches **▲** et **▼** permettent de passer d'une suite à l'autre.

V-Window

Start: 1

End: 114

Start: 1

End: 114

Start: 1

End: 114


Start: 1


End: 114

DEL CLR OPT F5/F6 OPT F5/F6

n	a _n	b _n
1	-3,2	0,15
2	-2,4	-0,15
3	-1,6	0,225

DEL CLR OPT F5/F6 OPT F5/F6



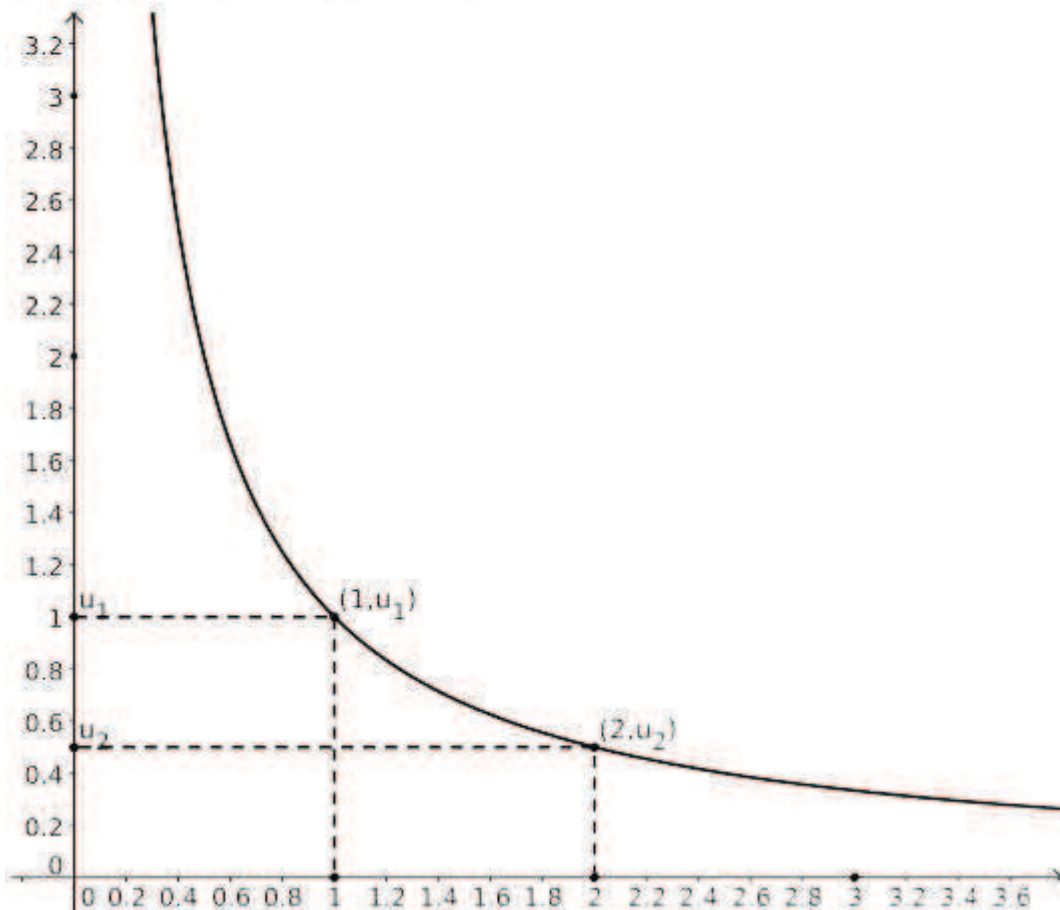


2. Représentation graphique d'une suite

- Représentation d'une suite définie par son terme général :

$$(u_n)_n \geq 1$$

$$u_n = f(n) \text{ où } f : x \mapsto \frac{1}{x} \text{ donc } u_n = \dots$$

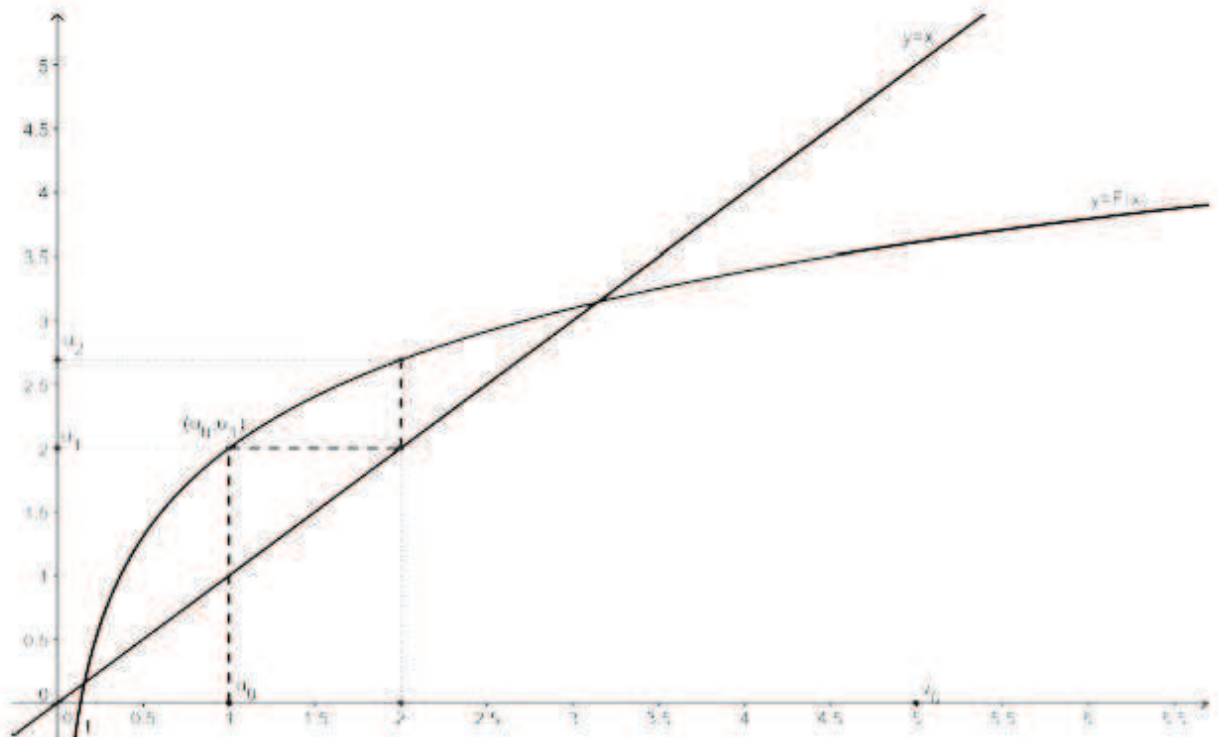


On peut lire les termes de la suite sur l'axe des ordonnées.

- Représentation d'une suite définie par récurrence :

$$\begin{cases} u_0 = 1 \\ u_{n+1} = F(u_n) \text{ où } F : x \mapsto 2 + \ln(x) \end{cases}$$

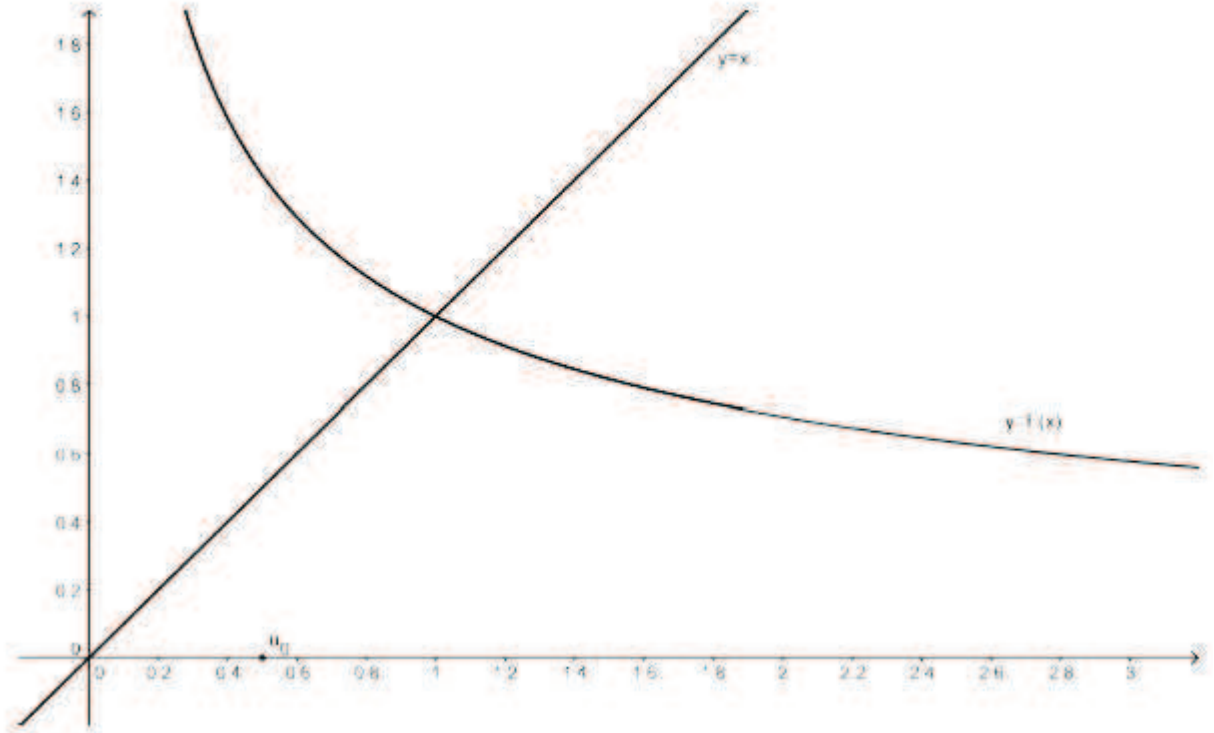
$$\begin{cases} v_0 = 5 \\ v_{n+1} = F(v_n) \text{ où } F : x \mapsto 2 + \ln(x) \end{cases}$$



On peut lire les termes de la suite sur l'axe des ordonnées et sur l'axe des abscisses.

$$\begin{cases} u_0 = \frac{1}{2} \\ u_{n+1} = 1/\sqrt{u_n} \end{cases}$$

La fonction d'itération est :



INTERROGATION DE RENTRÉE : ENNONCE

Interrogation de rentrée 2012

Exercice 1 :

Prérequis :

- 1) \exp et \ln sont des fonctions réciproques l'une de l'autre.
- 2) La fonction \ln est dérivable sur $]0; +\infty[$.
- 3) Soit u une fonction dérivable sur un intervalle I , alors e^u est dérivable sur I et $(e^u)' = u'e^u$.

Question : Montrer que pour tout $x \in]0; +\infty[$, on a $\ln'(x) = \frac{1}{x}$.**Exercice 2 :**

- 1) Déterminer les ensembles de définition des fonctions suivantes :

$$f(x) = \frac{\ln x}{1 - \ln x}, \quad g(x) = \ln(x^2 - 4) - \ln(x + 2), \quad h(x) = \ln(x - 2)$$

- 2) Parmi ces fonctions, certaines sont-elles égales ? Si oui, lesquelles ?
- 3) Calculer la fonction dérivée de g .
- 4) Quel est le sens de variation de la fonction g sur son domaine de définition ?
- 5) Résoudre dans \mathbb{R} l'équation :

$$f(x) = 2$$

- 6) Résoudre dans
- \mathbb{R}
- l'inéquation :

$$\ln(x^2 - 4) \leq 1$$

- 7) On considère les fonctions
- k
- et
- l
- définies sur l'intervalle
- $]0; +\infty[$
- par

$$k(x) = x^2(1 - \ln x) \text{ et } l(x) = x^2 - \ln x.$$

- a) Déterminer la limite de k en $+\infty$.
- b) Déterminer la limite de k en 0.
- c) Déterminer la limite de l en $+\infty$.

Exercice 3 :L'espace est muni d'un repère $(O, \vec{i}, \vec{j}, \vec{k})$.On considère les droites (D_1) et (D_2) d'équations paramétriques :

$$(D_1) \begin{cases} x = 1 + 2t \\ y = -3t \\ z = -2 + t \end{cases} \quad t \in \mathbb{R} \quad \text{et} \quad (D_2) \begin{cases} x = 1 - 2t \\ y = 5 + t \\ z = 1 + t \end{cases}$$

- 1) Donner 2 points appartenant à la droite (D_1) , et 2 vecteurs directeurs de la droite (D_2) .
- 2) Le point $A(-3, 6, -4)$ est-il un point de (D_1) ? de (D_2) ?
- 3) Étudier la position relative des droites (D_1) et (D_2) .

Corrigé de l'interrogation de rentrée 2012

Exercice 1 :

Si on pose $u = \ln$, par le prérequis (2), u est dérivable sur $]0; +\infty[$.
 Par le prérequis (3), on a donc e^{\ln} est dérivable sur $]0; +\infty[$ et $(e^{\ln})' = \ln' e^{\ln}$.
 De plus, par le prérequis (1), pour tout $x \in]0; +\infty[$, on a $e^{\ln x} = x$. On en déduit que la dérivée de la fonction e^{\ln} est constante égale à 1 sur $]0; +\infty[$.
 Pour tout $x \in]0; +\infty[$, on a donc $(e^{\ln})'(x) = 1$.
 Soit : $\ln'(x)e^{\ln(x)} = 1 \Leftrightarrow \ln'(x)x = 1 \Leftrightarrow \ln'(x) = 1/x$ (possible car $x \neq 0$).
 Finalement, pour tout $x \in]0; +\infty[$, on a $\ln'(x) = 1/x$.

Exercice 2 :

1) Ensembles de définition :

$$f(x) = \frac{\ln x}{1 - \ln x} \text{ définie si } \begin{cases} x > 0 \\ \ln x \neq 1 \end{cases} \Leftrightarrow \begin{cases} x > 0 \\ x \neq e \end{cases} \quad D_f =]0, e[\cup]e, +\infty[$$

$$g(x) = \ln(x^2 - 4) - \ln(x + 2) \text{ définie si } \begin{cases} (x^2 - 4) > 0 \\ x + 2 > 0 \end{cases} \Leftrightarrow \begin{cases} (x^2 - 4) > 0 \\ x > -2 \end{cases}$$

x	$-\infty$	-2	2	$+\infty$		
signe de $x^2 - 4$		+	0	-	0	+

$$D_g = (]-\infty; -2[\cup]2; +\infty[) \cap (]-2; +\infty[) \text{ donc } D_g =]2; +\infty[$$

$$h(x) = \ln(x - 2) \text{ définie si } x - 2 > 0 \Leftrightarrow x > 2 \quad D_h =]2; +\infty[$$

2) Les fonctions g et h sont égales car elles ont le même ensemble de définition et, pour tous a et b strictement positifs, $\ln a - \ln b = \ln \frac{a}{b}$. Donc, pour tout $x \in]2, +\infty[$,

$$\ln(x^2 - 4) - \ln(x + 2) = \ln\left(\frac{(x^2 - 4)}{(x + 2)}\right) = \ln\left(\frac{(x - 2)(x + 2)}{(x + 2)}\right) = \ln(x - 2)$$

3) La fonction g est dérivable sur son ensemble de définition.

$g(x) = \ln(x^2 - 4) - \ln(x + 2)$, or on a que g est égale à h donc leurs dérivées aussi sont égales. Et $h(x) = \ln(x - 2)$, on remarque alors que la dérivée de h est plus rapide à calculer : (dérivée de $\ln u$)

$$h'(x) = \frac{1}{x - 2}$$

Autrement, on peut calculer directement la dérivée de g (mais c'est plus long) :

$$g'(x) = \frac{2x}{x^2 - 4} - \frac{1}{x + 2} = \frac{2x}{(x + 2)(x - 2)} - \frac{1}{x + 2} = \frac{2x - (x - 2)}{(x + 2)(x - 2)} = \frac{x + 2}{(x + 2)(x - 2)}$$

$$\text{donc } g'(x) = \frac{1}{x - 2}$$

4) On résout $\frac{1}{x-2} > 0 \Leftrightarrow x - 2 > 0 \Leftrightarrow x > 2$. Donc g est croissante sur son ensemble de définition $]2; +\infty[$.

5) Equation :

Pour $x \in]0, e[\cup]e, +\infty[$, on a

$$f(x) = 2 \Leftrightarrow \frac{\ln x}{1 - \ln x} = 2 \Leftrightarrow \ln x = 2 - 2 \ln x \Leftrightarrow 3 \ln x = 2 \Leftrightarrow \ln x = 2/3 \Leftrightarrow x = e^{2/3}$$

$$\boxed{S = \{e^{2/3}\}}$$

6) Inéquation :

$\ln(x^2 - 4) \leq 1$; cette inégalité a un sens pour $x \in]-\infty; -2[\cup]2; +\infty[$

Pour $x \in]-\infty; -2[\cup]2; +\infty[$, $\ln(x^2 - 4) \leq 1 \Leftrightarrow x^2 - 4 \leq e \Leftrightarrow x^2 \leq e + 4$

Or $4 + e > 0$ donc $x^2 \leq e + 4 \Leftrightarrow -\sqrt{e + 4} \leq x \leq \sqrt{e + 4}$

On a aussi $\sqrt{e + 4} > 2$ d'où l'ensemble solution :

$$\boxed{S = [-\sqrt{e + 4}; -2[\cup]2; \sqrt{e + 4}]}$$

7) On considère la fonction k définie sur l'intervalle $]0; +\infty[$ par

$$k(x) = x^2(1 - \ln x)$$

a) Limite de k en $+\infty$:

$$\left. \begin{array}{l} \lim_{x \rightarrow +\infty} (1 - \ln x) = -\infty \\ \lim_{x \rightarrow +\infty} (x^2) = +\infty \end{array} \right\} \text{ donc } \lim_{x \rightarrow +\infty} k(x) = -\infty \text{ par limite de produit.}$$

b) Limite de k en 0 : pour tout $x \in]0; +\infty[$, $k(x) = x^2 - x^2 \ln x$.

$$\left. \begin{array}{l} \lim_{x \rightarrow 0} (x^2) = 0 \\ \lim_{x \rightarrow 0} (x^2 \ln x) = 0 \\ \text{par croissance comparée} \end{array} \right\} \text{ donc } \lim_{x \rightarrow 0} k(x) = 0 \text{ par limite de somme.}$$

c) Pour tout $x \in]0; +\infty[$, $l(x) = x^2(1 - \frac{\ln x}{x^2})$.

$$\left. \begin{array}{l} \lim_{x \rightarrow +\infty} (\frac{\ln x}{x^2}) = 0 \\ \text{par croissance comparée} \\ \lim_{x \rightarrow +\infty} (1 - \frac{\ln x}{x^2}) = 1 \\ \text{par limite de somme} \\ \lim_{x \rightarrow +\infty} (x^2) = +\infty \end{array} \right\} \text{ donc } \lim_{x \rightarrow +\infty} l(x) = +\infty \text{ par limite de produit.}$$

Exercice 3 :

$$(D_1) \begin{cases} x = 1 + 2t \\ y = -3t \\ z = -2 + t \end{cases} \quad t \in \mathbb{R} \quad \text{et} \quad (D_2) \begin{cases} x = 1 - 2t \\ y = 5 + t \\ z = 1 + t \end{cases}$$

1) Deux points appartenant à (D_1) : il suffit de prendre des valeurs différentes pour t .

Par exemple pour $t = 0$: $A(1, 0, -2)$, pour $t = 1$: $B(3, -3, -1)$.

Deux vecteurs directeurs de la droite (D_2) : par exemple $\vec{u} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ et

$$2\vec{u} \begin{pmatrix} -4 \\ 2 \\ 2 \end{pmatrix}$$

2) Le point $A(-3, 6, -4)$ est-il un point de (D_1) ? de (D_2) ?

$$\text{Pour } (D_1) : \text{ on résout } \begin{cases} -3 = 1 + 2t \\ 6 = -3t \\ -4 = -2 + t \end{cases} \Leftrightarrow \begin{cases} t = -2 \\ t = -2 \\ t = -2 \end{cases}$$

donc $A \in (D_1)$

$$\text{Pour } (D_2) : \text{ on résout } \begin{cases} -3 = 1 - 2t \\ 6 = 5 + t \\ -4 = 1 + t \end{cases} \Leftrightarrow \begin{cases} t = 2 \\ t = 1 \\ t = -5 \end{cases}, \text{ pas de solution,}$$

donc $A \notin (D_2)$

$$3) \text{ Un vecteur directeur de } (D_1) \text{ est } \vec{v} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$$

Les vecteurs \vec{u} et \vec{v} ne sont pas colinéaires car leurs coordonnées ne sont pas proportionnelles donc (D_1) et (D_2) ne sont pas parallèles.

$$\text{On cherche si elles sont sécantes : on résout } \begin{cases} 1 + 2t = 1 - 2u \\ -3t = 5 + u \\ -2 + t = 1 + u \end{cases} \Leftrightarrow \begin{cases} t = -u \\ u = 5/2 \\ t = 3/2 \end{cases}$$

Ce système n'a pas de solution donc les droites (D_1) et (D_2) ne sont pas sécantes.

Les droites (D_1) et (D_2) ne sont ni parallèles, ni sécantes donc elles sont non coplanaires.

GRILLE DE NOTES

		Notes
AUFFRAY	Sandrine	5,5
BEN AMADI	Amir	16
BEN AMADI	Wassim	17
BENELHADJ	Ilies	13
BENOIST	Morgane	13,5
CAMUS	Rémi	17
DE FREMICOURT	Noé	11
DE JOANNIS DE VERCLOS	Christophe	14,5
EL OUADRASSI	Billal	16
FITAN	Pierre	11,5
GUIGUE	Axel	13,25
HENRY	Gaëlle	8,25
JUPITER	Adrien	13,75
LECLERC	Armand	17,5
LUCIEN	Anthony	17
MAGRO	Rémi	17,5
MARION	François	14,25
MARTEU	Theophane	14
MATOUG	Camil	Abs
MONTES	Nicolas	17
PEIGNE	Sylvain	7,25
PERRIN	François-Xavier	9
PIERRON	Mathieu	18
PUJOS	Lucie	9,25
RIONDET	Lucas	12,5
SMAGGHE	Jérémy	9,5
TANGUY	Alexandre	5
VINCHON	Thibault	8
	Moyenne :	12,8148148

Dynamique de la dépression de consanguinité : effets combinés de la dominance et de la super-dominance

Etienne Adam, Lucile Morelle

Encadré par E. Porcher, A. Véber, V. Bansaye

1 Introduction

Une majorité des espèces de plantes à fleur possèdent les fonctions mâles et femelles sur le même individu, ce qui permet de l'autofécondation. Néanmoins on peut observer dans la nature une grande diversité des taux d'autofécondation : proche de 100% chez le blé, proche de 5% chez le maïs et des taux intermédiaires chez d'autres espèces. Comment expliquer une telle diversité ? Pour mieux comprendre l'évolution des taux d'autofécondation, il est judicieux d'étudier les forces évolutives contrôlant ce taux. La plus importante d'entre elles est la dépression de consanguinité. C'est la diminution de la valeur sélective des individus issus d'autofécondation. Elle peut être causée par deux mécanismes différents : les mutations délétères récessives et la superdominance. Les mutations délétères récessives sont des mutations qui ne s'expriment qu'à l'état homozygote. L'autofécondation augmente l'homozygotie donc diminue la valeur sélective moyenne. Néanmoins sur une longue échelle de temps les mutations ont tendance à être éliminées par la sélection naturelle (ce phénomène s'appelle la purge de la dépression). La superdominance a lieu quand les individus hétérozygotes ont une meilleure valeur sélective que les homozygotes.

L'effet de ces deux phénomènes séparés sur l'autofécondation a déjà été étudié mais pas encore conjointement. L'objectif de ce projet est d'étudier l'effet combiné de ces deux phénomènes. Pour cela nous avons commencé par poser un modèle probabiliste en population finie à deux loci : mutation délétère sur un locus et superdominance sur le second locus. Puis nous avons plutôt étudié la dynamique des génotypes à taux d'autofécondation constant du modèle déterministe limite en grande population. Nous avons dans un premier temps regardé les deux effets séparément puis nous avons fait des simulations pour voir la dynamique sur les deux loci. Ensuite nous avons cherché les taux d'autofécondation stables, c'est à dire qui ne peuvent être envahis par l'arrivée d'un nouveau taux. Nous n'avons traité que le cas à un seul locus pour les modèles d'invasion.

2 Modèle

2.1 Présentation

Nous nous plaçons dans une population diploïde de taille constante N , et un taux d'autofécondation σ . Nous allons nous intéresser à deux gènes di-alléliques, l'un avec mutation délétère récessive et l'autre avec superdominance.

Au premier locus : mutation délétère récessive, avec deux allèles a et A ;

Génotype	AA	Aa	aa
Valeur sélective	1	$1 - hs$	$1 - s$

où $s \simeq 1$ et h petit.

Au deuxième locus : superdominance, avec deux allèles B et C ;

Génotype	BB	BC	CC
Valeur sélective	$1 - t$	1	$1 - t$

Dans la population, on peut distinguer 10 génotypes :

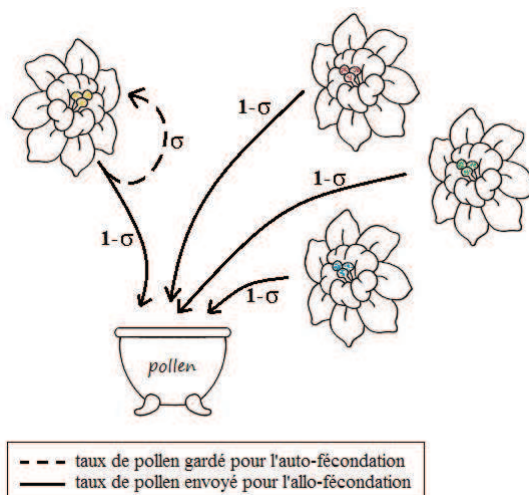
	AB/AB	AB/AC	AC/AC	AB/aB	AB/aC
Fréquence	x_1	x_2	x_3	x_4	x_5
Fitness w_i	$1 - t$	1	$1 - t$	$(1 - hs)(1 - t)$	$1 - hs$

	AC/aB	AC/aC	aB/aB	aB/aC	aC/aC
Fréquence	x_6	x_7	x_8	x_9	x_{10}
Fitness w_i	$1 - hs$	$(1 - hs)(1 - t)$	$(1 - s)(1 - t)$	$1 - s$	$(1 - s)(1 - t)$

Définition 2.1.1. On définit la valeur sélective moyenne par $W = \sum_{i=1}^{10} w_i x_i$

On fait évoluer la population de génération en génération en temps discret de la manière suivante :

- * Un individu pris au hasard dans la population meurt.
- * L'individu mort est remplacé par un nouvel individu obtenu selon le processus de reproduction suivant :
 - une mère est choisie aléatoirement proportionnellement à son fitness, elle est donc de génotype i avec probabilité $w_i x_i / W$,
 - chaque individu possède la même quantité de pollen, il en garde une proportion σ pour faire de l'auto-fécondation et en envoie une quantité $1 - \sigma$ pour faire de l'allofécondation
 - la mère reçoit donc une part σ de son propre pollen et une portion $1/N$ de l'ensemble du pollen envoyé pour l'allofécondation par tous les individus.
 - le père est donc de type $\begin{cases} i & \text{avec probabilité } \sigma + (1 - \sigma)x_i \\ j; j \neq i & \text{avec probabilité } (1 - \sigma)x_j \end{cases}$



Il reste donc à déterminer quel peut être le génotype de l'enfant sachant que sa mère est de type i et son père de type j . Pour le cas d'un seul locus, ceci est résumé dans le tableau suivant :

Couple mère × père	Fréquence des descendants AA	Fréquence des descendants Aa	Fréquence des descendants aa
$AA \times AA$	1		
$AA \times Aa$	1/2	1/2	
$AA \times aa$		1	
$Aa \times aa$	1/2	1/2	
$Aa \times Aa$	1/4	1/2	1/4
$Aa \times aa$		1/2	1/2
$aa \times AA$		1	
$aa \times Aa$		1/2	1/2
$aa \times aa$			1

2.2 Limite en grande population

On va regarder la limite de notre processus $X(n) \in [0, 1]^{10}$ lorsque N tend vers $+\infty$. On note M_i l'événement {mère de génotype i }, P_i l'événement {père de génotype i }, $Pa(i, j)$ l'événement $M_i \cap P_j$ et D_i l'événement {descendant de génotype i }.

Soit Φ_k la probabilité d'avoir un descendant de génotype k . On a :

$$\Phi_k = \sum_{i,j} \mathbb{P}(M_i) \mathbb{P}(P_j) \mathbb{P}(D_k | Pa(i, j)).$$

Théorème 2.2.1. *Lorsque N tend vers $+\infty$, le processus $(X(n), n \in \mathbb{N})$ tend vers le processus déterministe (x^n) défini par $x_i^{n+1} = \Phi_i(x^n)$. Si on change l'échelle de temps, $X(n)$ tend vers une solution de l'équation différentielle $x'_i = \Phi_i(x) - x_i$*

Démonstration : On considère le processus $(\tilde{X}(t), t \geq 0)$, défini par $\tilde{X}(t) = X([Nt])$, $X(0) = (x_1, \dots, x_{10})$.

On regarde le générateur de ce processus.

Soit $f \in \mathcal{C}^2(\mathbb{R}^{10}, \mathbb{R})$, regardons la quantité $\frac{\mathbb{E}(f(\tilde{X}(t))) - f(\tilde{X}(0))}{t}$ en $t = 1/N$.

On obtient : $N \sum_{i,j=1, i \neq j}^{10} x_i \Phi_j f(x_1, \dots, x_i - 1/N, \dots, x_j + 1/N, \dots, x_{10}) - f(\tilde{X}_0)$

En regroupant les termes, et en faisant tendre N vers $+\infty$ on obtient :

$$\mathcal{G}^\infty f(x) = \sum_{i=1}^{10} \partial_i f(x) (\Phi_i(x) - x_i)$$

D'où le résultat. \square

Nous allons maintenant étudier plus en détail ces équations différentielles.

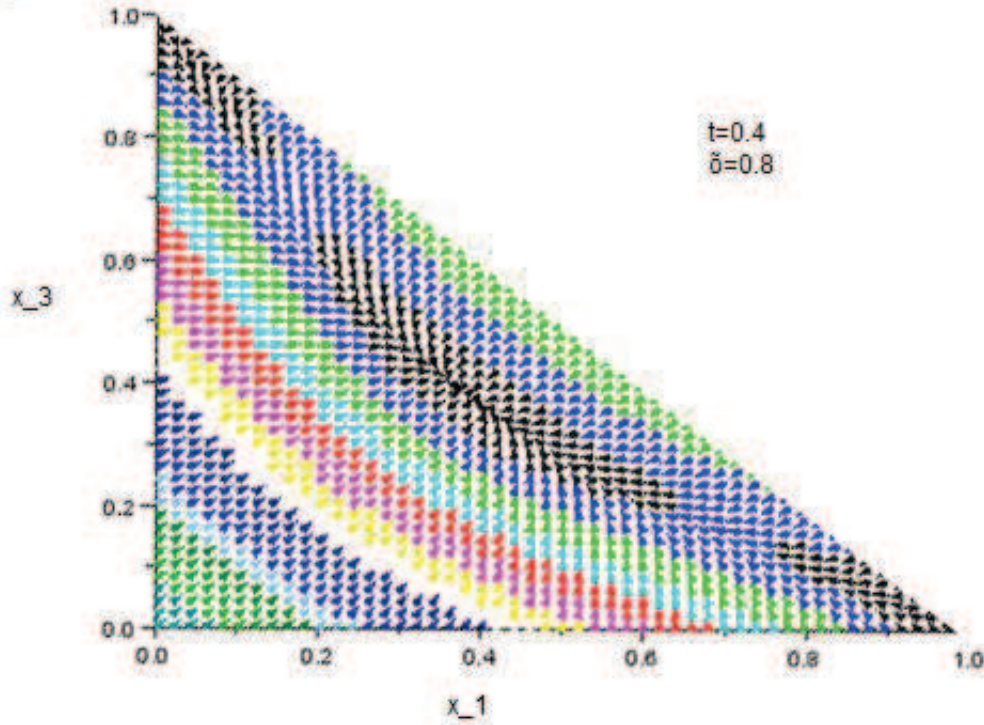
3 Dynamique des génotypes à taux d'autofécondation constant

3.1 Cas de la superdominance seule

Si l'on ne regarde que le locus avec superdominance, la limite en grande population donne l'équation différentielle suivante :

$$\begin{cases} \frac{dx_1}{dt} = \frac{(\sigma + (1-\sigma)x_1)(1-t)x_1 + (\sigma + (1-\sigma)x_2)x_2/4 + (1-\sigma)(2-t)x_1x_2/2}{W} - x_1 \\ \frac{dx_2}{dt} = \frac{(\sigma + (1-\sigma)x_2)x_2/2 + 2(1-\sigma)(1-t)x_1x_3 + (1-\sigma)(1-t)(x_1+x_3)x_2/2}{W} - x_2 \\ \frac{dx_3}{dt} = \frac{(\sigma + (1-\sigma)x_3)(1-t)x_3 + (\sigma + (1-\sigma)x_2)x_2/4 + (1-\sigma)(2-t)x_2x_3/2}{W} - x_3 \end{cases} \quad (1)$$

Ci-dessous le champ représenté pour $t = 0.4$ et $\sigma = 0.8$:



On voit la présence d'un point fixe qui est attractif. On peut calculer précisément la valeur d'un point fixe en fonction de t et σ :

Proposition 3.1.1. *Pour une condition initiale où $x_1(0) = x_3(0)$, on a un équilibre en $x_1 = x_3 = x$, $x_2 = 1 - 2x$, où x est une solution dans $[0, 1/2]$ de $2tX^2 - (1 + t/2 - \sigma/2 + \sigma t/2)X + 1/4 = 0$.*

Démonstration : Comme x_1 et x_3 ont un rôle symétrique, on admet qu'à l'équilibre $x_1 = x_3$.

On annule $\frac{dx_1}{dt}$:

$$Wx_1 = (\sigma + (1 - \sigma)x_1)(1 - t)x_1 + (\sigma + (1 - \sigma)x_2)x_2/4 + (1 - \sigma)(2 - t)x_1x_2/2$$

$$(1 - 2tx_1)x_1 = (\sigma + (1 - \sigma)x_1)(1 - t)x_1 + (\sigma + (1 - \sigma)(1 - 2x_1))^2/4 + (1 - \sigma)(2 - t)x_1(1 - 2x_1)/2$$

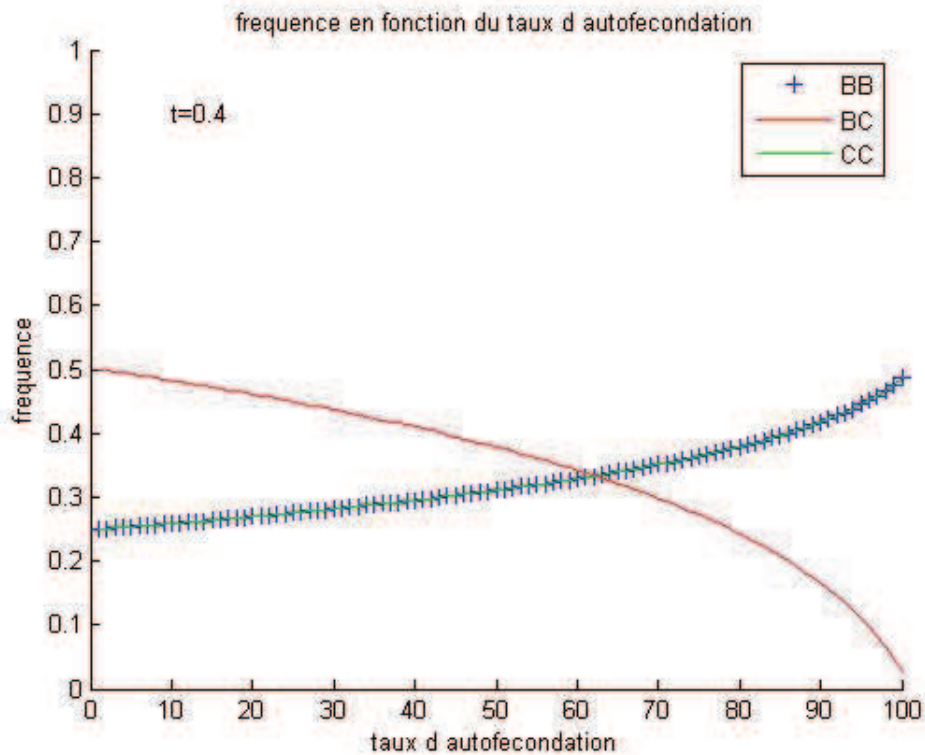
$$2tx_1^2 - (1 + t/2 - \sigma/2 + \sigma t/2)x_1 + 1/4 = 0$$

Le discriminant $\Delta = (1 + t/2 - \sigma/2 + \sigma t/2)^2 - 2t$ est toujours positif lorsque t et σ sont entre 0 et 1 (pour ces valeurs il atteint son minimum en $\sigma = 1$ et $t = 0.5$ et on a $\Delta = 0$).

On peut vérifier que l'on a toujours une solution dans $[0, 1/2]$. \square

Remarque 3.1.2. *Les cas $\sigma = 0$ et $\sigma = 1$ donnent des solutions simples : pour $\sigma = 0$, $x = 1/4$ (x ne dépend pas de t !), pour $\sigma = 1$, si $t > 1/2$ alors $x = 1/4t$ ou $x = 1/2$, si $t < 1/2$, $x = 1/2$.*

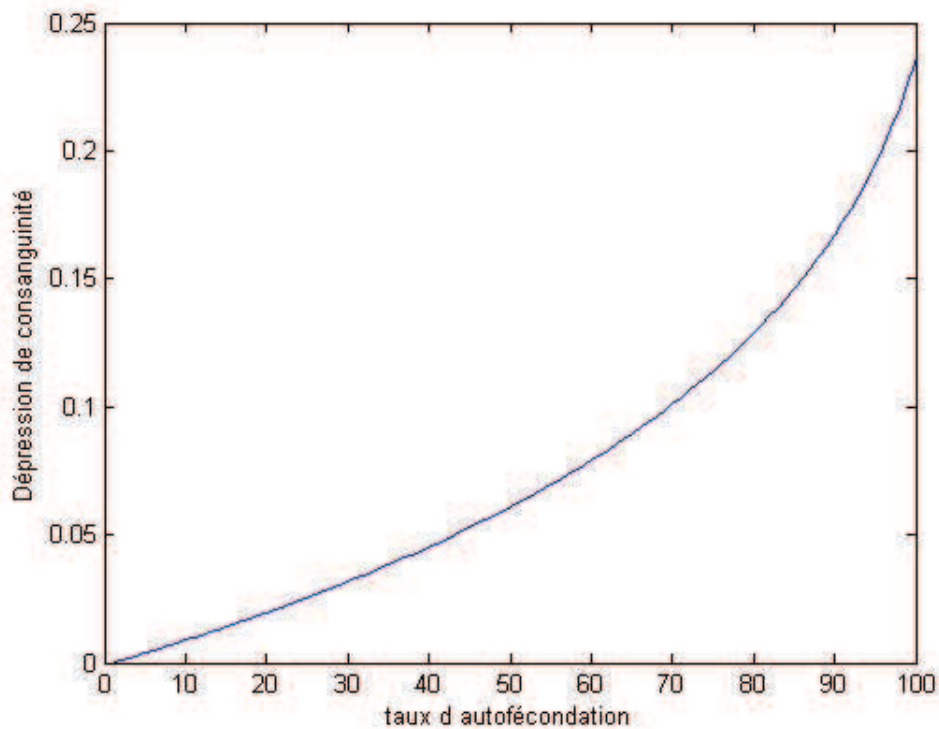
Ci-dessous on a représenté la fréquence des différents génotypes à l'équilibre en fonction du taux d'autofécondation. Comme attendu, un plus fort taux d'autofécondation diminue l'hétérozygotie, donc la valeur sélective moyenne.



Définition 3.1.3. On appelle *dépression de consanguinité* la diminution de la valeur sélective due à l'autofécondation. Si on note W_σ le fitness moyen à l'équilibre d'une population ayant un taux d'autofécondation σ , alors la dépression de consanguinité D vaut :

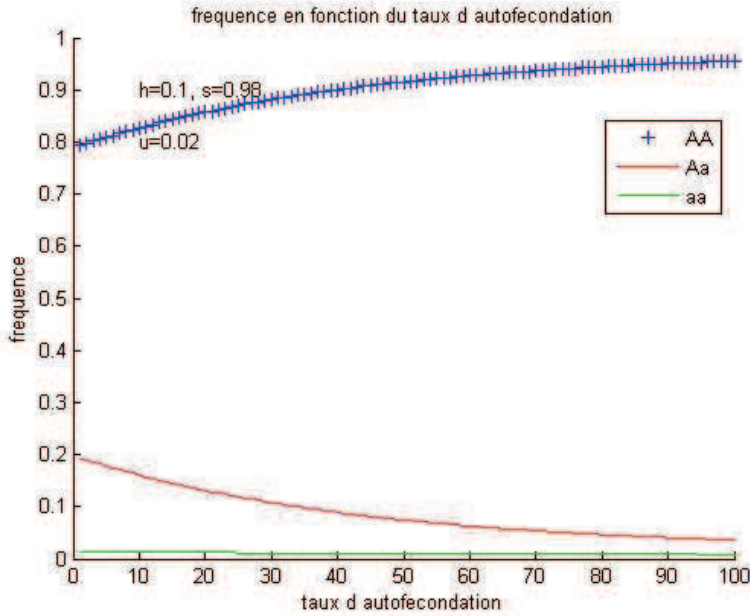
$$D = \frac{W_0 - W_\sigma}{W_0}$$

Ci-dessous la dépression de consanguinité pour le locus avec superdominance :

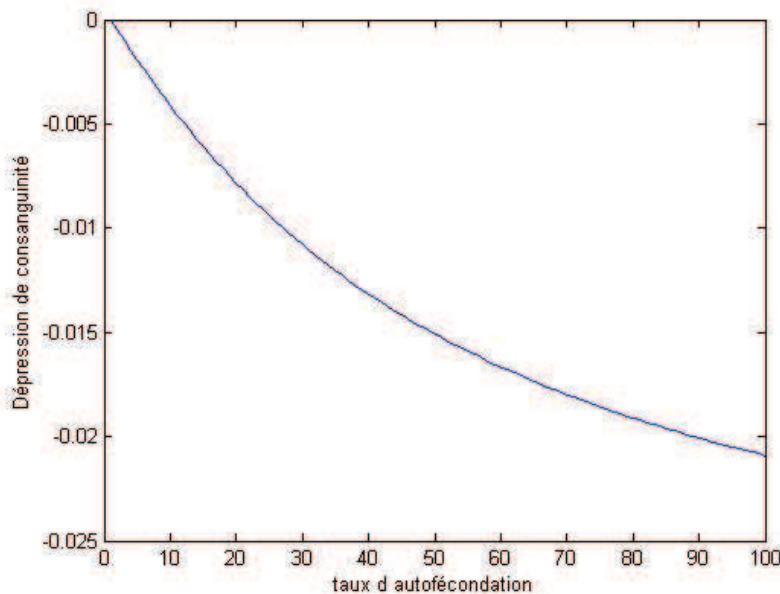


3.2 Cas de la mutation délétère

Sans mutation, l'allèle délétère récessive va être complètement éliminé par la sélection naturelle. Il faut donc ajouter un taux de mutation u à notre modèle. On considère qu'un individu de type AA a une probabilité u de muter en Aa , et un individu Aa a une probabilité u de muter en aa . On néglige les doubles mutations car le facteur u est très petit. A l'équilibre, la fréquence du génotype aa sera très faible, mais il restera encore des individus de type Aa , les mutations empêchant l'élimination totale de l'allèle a . On représente ci-dessous la fréquence des différents génotypes à l'équilibre en fonction du taux d'autofécondation et la dépression de consanguinité :



Proposition 3.2.1. On a l'approximation à l'équilibre suivante : $x_2 \sim \frac{u(2-\sigma)}{s\sigma+2h(1-\sigma)}$

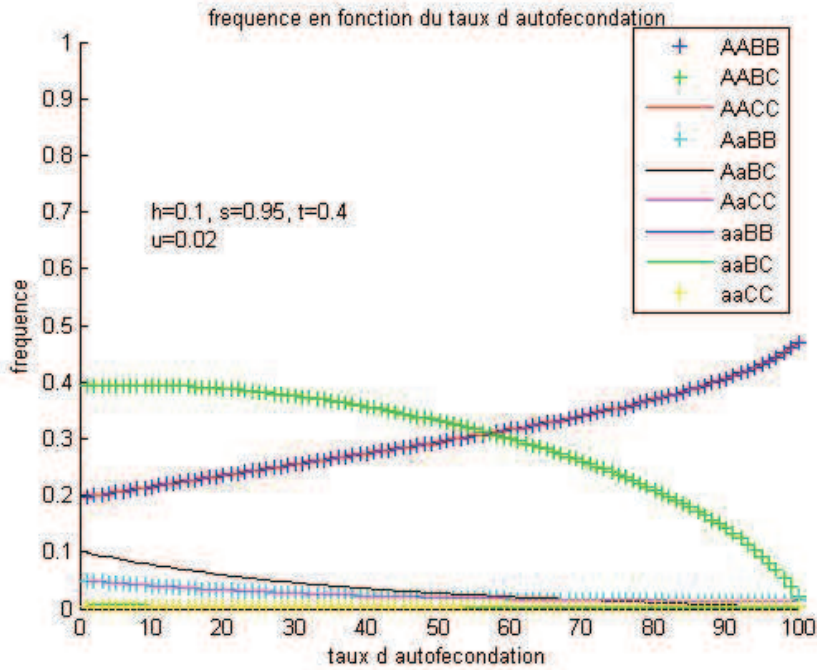


Le fitness moyen de la population augmente avec le taux d'autofécondation. C'est parce que l'on regarde juste la fréquence à l'équilibre donc il y a eu la purge de l'allèle délétère. Une autofécondation totale semble donc favorisée.

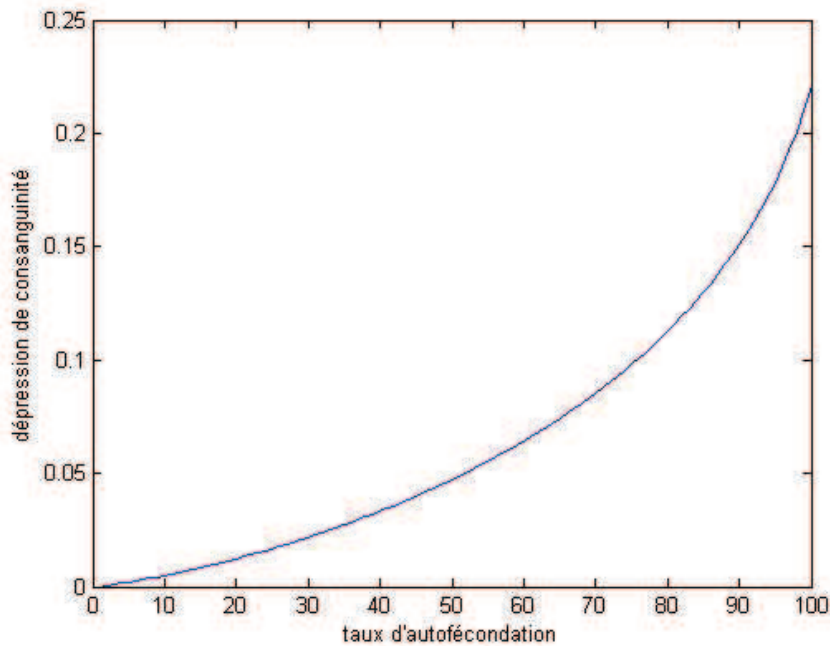
3.3 Cas des 2 loci

On va maintenant combiner les deux loci. Comme on a dix génotypes différents, il est difficile de faire une étude mathématique rigoureuse, on se contentera donc de faire des simulations numériques.

On a d'abord calculé les fréquences à l'équilibre des différents génotypes en fonction du taux d'autofécondation. On obtient le graphe suivant :



On remarque que l'allure des courbes pour les génotypes de type AA est la même pour le locus seul de la superdominance. Ceci laisse à penser que l'on aura la même courbe pour la dépression de consanguinité. C'est en effet le cas :



Plus le taux d'autofécondation est élevé, plus le fitness moyen est faible. Il semble donc que l'autofécondation ne soit pas favorisée.

4 Des taux intermédiaires ?

4.1 Dépression de consanguinité

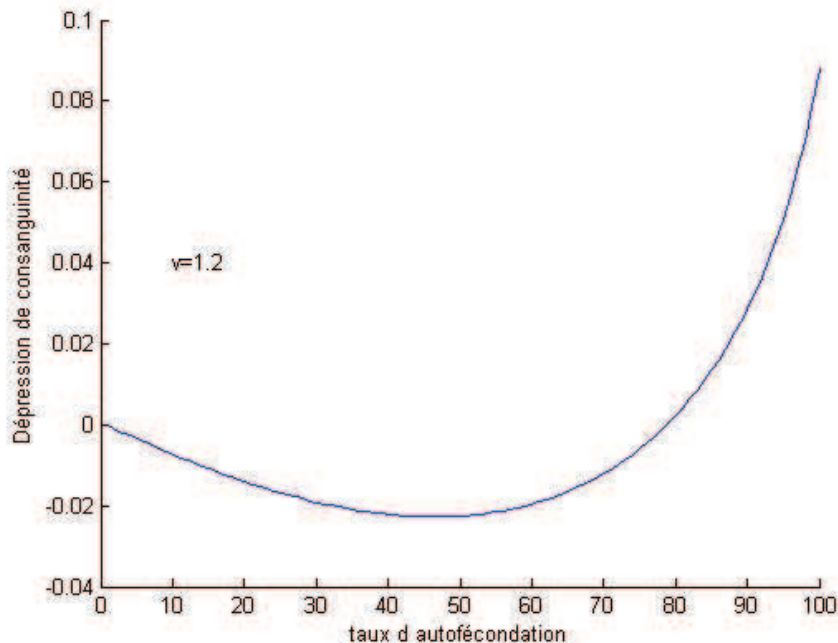
Sur les précédentes courbes, nous avons vu que la dépression de consanguinité était monotone avec le taux d'autofécondation. Ceci laisse penser que les taux d'autofécondation stables d'un point de vue évolutif sont de 0 ou 100%. Ce qui n'est pas satisfaisant au vu des données statistiques sur le taux d'autofécondation chez les plantes à fleur, car il y a une portion non négligeable à taux intermédiaire.

On va maintenant légèrement complexifier notre modèle. L'autofécondation peut être un avantage sélectif : en effet elle limite la perte de pollen et permet aussi la reproduction sans la présence proche d'un congénère. On va modéliser ceci en attribuant une valeur sélective $v > 1$ aux individus issus de l'autofécondation (et 1 pour ceux issus de l'allofécondation). Pour le cas de la superdominance seule cela donne :

$$\begin{cases} \frac{dx_1}{dt} = \frac{(v\sigma + (1-\sigma)x_1)(1-t)x_1 + (v\sigma + (1-\sigma)x_2)x_2/4 + (1-\sigma)(2-t)x_1x_2/2}{W^*} - x_1 \\ \frac{dx_2}{dt} = \frac{(v\sigma + (1-\sigma)x_2)x_2/2 + 2(1-\sigma)(1-t)x_1x_3 + (1-\sigma)(1-t)(x_1+x_3)x_2/2}{W^*} - x_2 \\ \frac{dx_3}{dt} = \frac{(v\sigma + (1-\sigma)x_3)(1-t)x_3 + (v\sigma + (1-\sigma)x_2)x_2/4 + (1-\sigma)(2-t)x_2x_3/2}{W^*} - x_3 \end{cases} \quad (2)$$

où $W^* = (v\sigma + 1 - \sigma)W$.

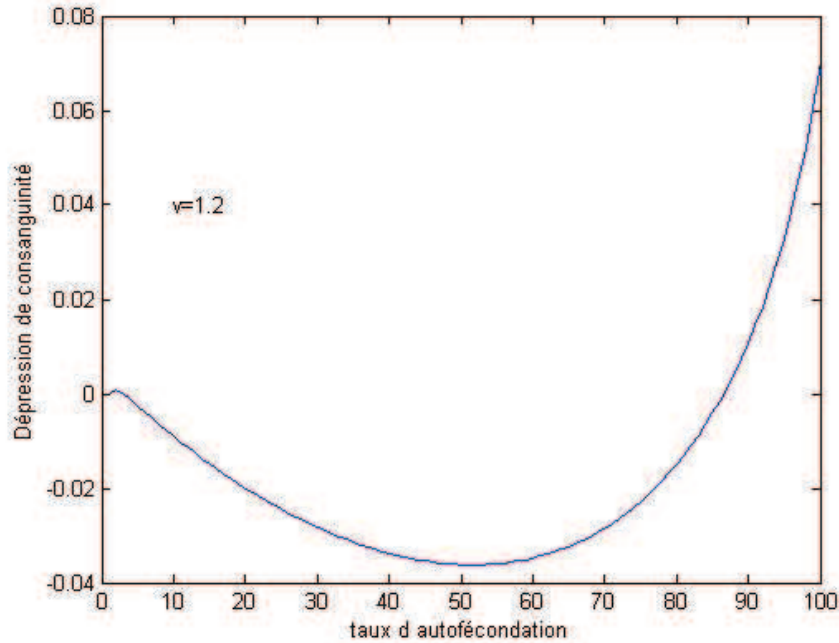
Observons la dépression de consanguinité obtenue pour le locus de la superdominance seul avec $v = 1.2$:



On a une dépression de consanguinité minimale pour $\sigma = 0.46$, ce qui pourrait expliquer la possibilité d'un taux d'autofécondation intermédiaire.

Pour le locus de la mutation délétère, on ne peut pas avoir de minimum de la dépression de consanguinité dans $]0, 1[$. On ne peut donc pas expliquer la présence de taux intermédiaires avec ce seul locus.

Pour les 2 loci ensembles, on a vu que les courbes des fréquences des génotypes étaient proches de celle du locus de la superdominance. On obtient là aussi un minimum intermédiaire pour la dépression de consanguinité :



On observe un minimum de la dépression de consanguinité pour un taux d'autofécondation $\sigma = 0.52$.

L'étude avec les deux loci ressemble beaucoup au cas de la superdominance seule.

4.2 Modèle d'invasion

On considère encore notre population de plantes, avec deux loci, sur le premier locus on a de la superdominance et le second locus détermine le taux d'autofécondation. On considère qu'au temps $t = 0$, la population n'est constituée que de l'allèle E_σ sur le second locus (donc a un taux d'autofécondation σ) la fréquence des génotypes BB, BC, CC étant à l'équilibre calculé précédemment. On fait apparaître par mutation un allèle $E_{\sigma'}$ sur le second locus (on introduit à une très faible fréquence des individus de type BC $E_\sigma E_{\sigma'}$) et on regarde l'évolution des génotypes. On considère que les individus de type $E_\sigma E_{\sigma'}$ ont un taux d'autofécondation $\frac{\sigma + \sigma'}{2}$ et les individus de type $E_{\sigma'} E_{\sigma'}$ ont un taux d'autofécondation σ' .

Il y a alors trois types de comportements possibles :

- la mutation est éliminée, il n'y a pas invasion
- la mutation persiste et cohabite avec l'allèle E_σ
- il y a invasion, l'allèle E_σ est éliminé, il ne reste plus que des $E_{\sigma'}$

On représente ensuite les résultats sur un graphe où en abscisse on a le taux d'autofécondation initial et en ordonnée le taux d'autofécondation mutant. On note par un \oplus les zones où il y a invasion et par un \ominus celles où il n'y a pas invasion. On peut alors lire graphiquement le taux d'autofécondation σ^* qui ne peut être envahi («stratégie évolutivement stable»), tous les points d'abscisse σ^* sont dans une zone \ominus . Ce taux n'existe pas forcément.

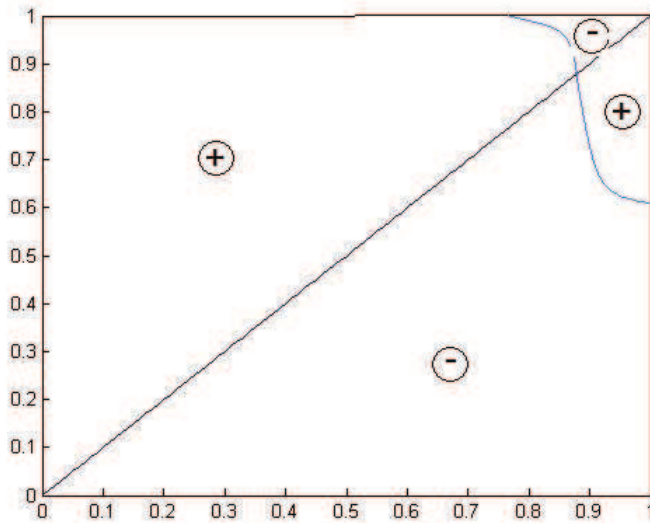
Pour $v \geq 1$, (on rappelle que v est la valeur sélective des individus issus de l'autofécondation), il y a invasion dès lors que $\sigma' > \sigma$. Donc seul un taux d'autofécondation à 100% est évolutivement stable.

Pourtant on a vu précédemment que pour $v = 1.2$ et $t = 0.4$, le fitness moyen atteignait son maximum en $\sigma = 0.46$. On peut donc avoir une invasion d'un taux d'autofécondation qui va diminuer le fitness moyen de la population. En fait on a invasion si la population mutante se développe plus vite que la population initiale. Pour $v \geq 1$, l'autofécondation est favorisée, donc la croissance est d'autant plus grande que le taux d'autofécondation est grand.

Peut-on considérer $v < 1$? Cela ne serait pas complètement absurde d'un point de vue

biologique : on peut considérer que nos individus sont proches, ce qui limite la difficulté de trouver un partenaire et la perte de pollen, aussi l'autofécondation peut, sur des loci non considérés, amener des mutations létales. Donc il est possible de considérer que la valeur sélective des individus issus de l'autofécondation est inférieure à 1.

On a étudié notre modèle d'invasion pour $v = 0.8$. On obtient le graphique suivant :



On a une stratégie évolutivement stable en $\sigma = 0.9$. Jusqu'à $\sigma = 0.8$, chaque taux se fait envahir par tous les taux qui lui sont supérieurs, tandis qu'il n'y a pas d'invasion pour les taux inférieurs. A partir de $\sigma = 0.8$, il y a des forts taux qui n'envahissent plus. Au delà de $\sigma = 0.9$, il y a invasion par quelques taux inférieurs.

5 Conclusion

A partir de notre modèle que nous avons petit à petit complexifié nous avons montré que l'on pouvait trouver des taux d'autofécondation intermédiaires qui maximisent le fitness moyen de la population. L'étude, pour chacun des locus pris séparément, des fréquences génotypiques à l'équilibre apporte des informations intéressantes pour les deux loci pris ensemble car dans ce dernier cas tout ressemble beaucoup au cas de la superdominance seule mais une étude mathématique est trop pénible à cause de la grande dimension du problème, c'est pourquoi nous avons surtout eu recours à des simulations numériques sous Matlab. Les résultats sur le modèle d'invasion nous montrent que la stabilité d'un taux n'est pas synonyme de grand fitness moyen de la population. En effet, lorsqu'un nouveau taux d'autofécondation apparaît, il y a invasion si la population mutante se développe plus vite que la population résidente. Cette vitesse de développement est certes liée à la valeur sélective mais pas seulement. D'où les résultats que l'on obtient lorsqu'on met une valeur sélective pour l'autofécondation supérieure à celle de l'allofécondation : on a un maximum du fitness moyen pour un taux intermédiaire mais l'équilibre évolutivement stable a lieu lorsqu'il y a autofécondation à 100%. Pour obtenir un équilibre évolutivement stable intermédiaire il faut considérer que la valeur sélective de l'autofécondation est inférieure à celle de l'allofécondation. Nous rappelons que ces résultats ne concernent que le locus de la superdominance seule. Nous n'avons pas eu le temps de traiter les deux loci ensemble (la procédure serait assez lourde car il y aurait une trentaine de génotypes à considérer!). Il y a aussi de nombreuses autres pistes que nous n'avons pas explorées : traiter le cas des populations finies, prendre en compte la limitation en pollen,...

6 Références

D. Charlesworth and J. H. Willis, **The genetics of inbreeding depression**, [Nature reviews, Genetics volume 10, november 2009].

E. Porcher and R. Lande, **The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation**, [J. Evol. Biol. 18 (2005) 497-508].

M. K. Uyenoyama and D. M. Waller, **Coevolution of Self-Fertilization and Inbreeding Depression, I. Mutation-Selection Balance at One and Two Loci**, [Theoretical Population Biology 40, 14-46 (1991)].

M. K. Uyenoyama and D. M. Waller, **Coevolution of Self-Fertilization and Inbreeding Depression, II. Symmetric Overdominance in Viability**, [Theoretical Population Biology 40, 47-77 (1991)].

S. Glemin, **How are deleterious mutations purged ? Drift versus nonrandom mating**, [Evolution, 57(12), 2003, pp. 2678-2687].

R. Durrett, **Probability Models for DNA Sequence Evolution**.

A. Etheridge, **Some Mathematical Models from Population Genetics**.

W. J. Ewens, **Mathematical Population Genetics**.



Stage de M2 MathSV

Prédictions théoriques pour la similarité entre deux
communautés écologiques séparées par une distance
géographique donnée

Sous la direction de Amandine Véber et Hélène Morlon

Lucile Morelle

16 juillet 2013

Table des matières

Introduction	3
1 Un modèle neutre de β-diversité	5
1.1 Introduction	5
1.2 Modèle	6
1.3 Equation de Malécot	9
1.4 Noyau de Gauss	11
1.4.1 En dimension deux	12
1.4.2 Noyaux gaussiens multidimensionnels	14
1.5 Noyaux « fat-tailed » (à queue lourde)	16
1.6 Modèle du votant et limite de non-spéciation	17
1.7 Expériences numériques	19
1.8 Conclusion	21
2 Inférence en deux dimensions basée sur les fréquences alléliques	23
2.1 Introduction	23
2.2 Fréquences alléliques	24
2.2.1 La formule de Wright-Malécot	24
2.2.2 Une statistique F appropriée	34
2.2.3 Inférence des paramètres à partir de F_{ST}	35
2.3 Inférence	36
2.3.1 Identité d'état et mesures de F_{ST}	36
2.3.2 Une approche du maximum de vraisemblance pour l'inférence	40
2.3.3 Résultats basés sur les fréquences alléliques	44
2.4 Conclusion	46
3 Un nouveau modèle neutre de β-diversité	47
3.1 Modèle	47
3.2 Fonction de similarité par paires	48
3.3 Etude de la loi de l'indice de dissimilarité	51
3.4 Estimations	58
3.5 Simulations	62
Conclusion	65
Bibliographie	67

Introduction

La similarité en composition des communautés écologiques décroît avec la distance géographique qui les sépare, ce patron est connu sous le nom de relation distance-décroissance.

D'où vient cette décroissance de la similarité écologique avec la distance géographique ? D'une part de la structure spatiale de l'environnement (dans le cas où le type d'habitat rencontré varie en fonction de l'endroit considéré). D'autre part de la limitation en dispersion des espèces (les individus d'une espèce se reproduisant de manière plus ou moins locale, celle-ci ne peut pas forcément se propager instantanément à toute la zone considérée).

La relation distance-décroissance est l'un des patrons les plus décrits et étudiés en écologie. Cependant, il en existe très peu de prédictions théoriques.

Notre objectif est d'obtenir des prédictions pour la relation distance-décroissance pour des modèles « neutres » de la biodiversité.

Un modèle « neutre » est un modèle qui ne prend pas en compte les différences d'adaptation des espèces à l'environnement. Les espèces seront donc considérées comme étant équivalentes du point de vue de leur démographie, de leur capacité de dispersion et de leur faculté à se diversifier pour former de nouvelles espèces.

Il existe un problème équivalent en génétique des populations. En considérant une seule espèce et un gène d'intérêt, les individus peuvent porter des allèles différents de ce gène et lorsque la population a une structure spatiale, les fréquences de chaque espèce deviennent une fonction de la coordonnée géographique.

Des outils existent déjà pour décrire les corrélations entre les diversités génétiques locales en fonction de la distance qui les sépare. Si on représente la population par un nuage de points dans \mathbb{R}^2 (par exemple) qui évolue au cours du temps grâce à des événements de reproduction, on peut décrire explicitement les arbres généalogiques reliant un échantillon d'individus.

Nous avons une relation de dualité entre le modèle de reproduction et ses généalogies. Ici, nous supposons la taille de population finie. Si nous ne nous intéressons plus aux allèles, mais seulement aux relations de parenté entre les individus, le coalescent peut alors être interprété comme la généalogie d'un échantillon d'individus dans le modèle. L'approche coalescente est rétrospective : au lieu de travailler en avant dans le temps, nous travaillons en arrière dans le temps, en retraçant les lignées des individus de l'échantillon jusqu'à ce qu'elles mutent ou coa-

lescent. Lorsque nous regardons en arrière dans le temps, les lignées fusionnent, ce qui signifie qu'en avant dans le temps, la coalescence est un cas où une lignée se divise en deux branches. Deux individus au temps t ont le même allèle si leurs ancêtres au temps 0 ont le même allèle et qu'il n'y a pas eu de mutation entre-temps. Nous sommes alors en mesure d'exprimer les corrélations spatiales entre les diversités locales en termes de probabilités de rencontre de deux marches aléatoires (les lignées ancestrales) avant qu'apparaisse une mutation.

Pour commencer, nous allons présenter l'étude de deux articles. Le premier chapitre sera consacré à l'article de Jérôme Chave et Egbert G. Leigh Jr., paru en 2002 dans la revue « Theoretical Population Biology », *A spatially Explicit Neutral Model of β -Diversity in Tropical Forests* et le chapitre suivant, à l'article de N. H. Barton, A. M. Etheridge, J. Kelleher et A. Véber, paru en 2013 dans la revue « Theoretical Population Biology », *Inference in two dimensions : allele frequencies versus lengths of shared sequence blocks*. Ensuite, dans un dernier chapitre, sous l'hypothèse de neutralité, nous nous intéresserons à un nouvel indice permettant de prédire la similarité en composition de communautés écologiques séparées par une distance donnée, il s'agira d'un nouveau modèle neutre de β -diversité.

Chapitre 1

Un modèle neutre de β -diversité

Dans ce chapitre, nous présentons l'étude de l'article de Jérôme Chave et Egbert G. Leigh Jr., paru en 2002 dans la revue « *Theoretical Population Biology* », *A spatially Explicit Neutral Model of β -Diversity in Tropical Forests*.

Pour représenter le renouvellement des espèces en forêt tropicale, nous allons utiliser le modèle « neutre » où le sort d'un arbre n'est pas affecté par l'espèce à laquelle il appartient, la dispersion des graines est limitée à partir des parents, et la spéciation est en équilibre avec l'extinction aléatoire. Nous allons calculer la fonction de similarité, c'est-à-dire la probabilité $F(r)$ que deux arbres, pris au hasard dans deux parcelles de forêt, appartiennent à la même espèce en fonction de la distance r qui sépare les parcelles. Nous supposerons pour cela que le noyau de dispersion des graines $P(r)$, c'est-à-dire la distribution des graines à partir de leurs parents, et les perspectives de mortalité et de reproduction, sont les mêmes pour tous les arbres, indépendamment de leur espèce. Une fois que nous aurons présenté le modèle et obtenu une formule générale en dimension deux pour la fonction de similarité, nous dériverons, de cette équation, la formule de Malécot. Nous regarderons plus particulièrement le cas des noyaux de dispersion de Gauss, à la fois en dimension deux et en dimension supérieure, ce qui nous permettra d'avoir une idée du comportement de $F_d(r)$, $d \leq 3$, pour r grand, pour tout noyau de dispersion $P(r)$, qui décroît assez rapidement pour r grand, puis nous nous intéresserons au cas des noyaux « fat-tailed » (à queue lourde). Ensuite, nous considérerons le cas limite de non-spéciation et nous mettrons en évidence la dualité de ce modèle neutre avec un modèle de marches aléatoires coalescentes. Enfin, pour vérifier la précision des prédictions du modèle, des simulations numériques seront utilisées.

1.1 Introduction

Pour commencer, nous définissons la β -diversité. Il s'agit d'une mesure de la biodiversité qui consiste à comparer la diversité des espèces entre écosystèmes ou le long de gradients environnementaux. Cela suppose de comparer le nombre de taxons qui sont uniques à chacun des écosystèmes. C'est le taux de variation en composition d'espèces dans l'ensemble des habitats ou parmi des communautés. Cela donne une mesure quantitative de la diversité des communautés des environnements changeants.

Au cours des 20 dernières années, la compréhension des processus maintenant la diversité

des arbres dans une région tropicale donnée a bien augmenté, grâce à l'établissement de grandes parcelles où la floristique a été répertoriée.

Nous avons moins d'informations sur les facteurs qui gouvernent la β -diversité, comme en témoigne la divergence en composition des espèces d'arbre entre parcelles de forêt à différents endroits. En effet, le renouvellement des espèces parmi les arbres tropicaux n'est qu'en voie d'être élucidé. En utilisant une extension de la théorie neutre de la biodiversité des arbres de Hubbell (2001), nous allons montrer comment la spéciation et la limitation de dispersion des graines peut contribuer au renouvellement des espèces.

Une mesure du renouvellement des espèces, ou β -diversité, est la relation distance-décroissance : comment l'augmentation rapide de r fait diminuer la probabilité $F(r)$ que deux arbres choisis au hasard séparés d'une distance r soient de la même espèce.

Ici, nous allons proposer une dérivation exacte et générale de $F(r)$ en utilisant tout noyau de dispersion symétrique et nous l'appliquerons à des noyaux particuliers.

Pour développer des prédictions du modèle, nous dériverons d'abord $F(r)$ quand les graines sont dispersées autour de leur parent suivant une distribution gaussienne symétrique. Cette formule est connue (Malécot, 1948) et des approximations (Nagylaki, 1976) et résultats asymptotiques (Sawyer, 1977a) ont déjà été obtenus.

Ensuite, nous utiliserons des simulations numériques de notre modèle afin de tester la précision des prédictions de notre modèle. A partir de la valeur moyenne de $F(r)$, pour r donné, à l'équilibre, nous évaluerons la variabilité de $F(r)$ et apprendrons comment les conditions initiales affectent l'approche de l'équilibre.

1.2 Modèle

On considère deux parcelles A et B , dont on connaît la floristique, avec S espèces en tout et nous notons :

$$\begin{cases} p_s^A & = \text{proportion d'individus appartenant à l'espèce } s \text{ sur la parcelle } A. \\ p_s^B & = \text{proportion d'individus appartenant à l'espèce } s \text{ sur la parcelle } B. \end{cases}$$

Remarque 1.2.0.1. p_s^A et p_s^B peuvent valoir 0.

Définition 1.2.0.2. On définit l'*indice de codominance* :

$$F(A, B) = \sum_{s=1}^S p_s^A p_s^B.$$

C'est la probabilité que deux arbres choisis au hasard, un dans A , l'autre dans B , soient de la même espèce.

Cet indice permet de mesurer la similitude en composition d'espèces des parcelles A et B .

Nous supposons que l'indice de codominance $F(A, B)$ ne dépend que de la distance qui sépare les parcelles A et B , que les graines se dispersent à une distance limitée autour de leur parent, et que les arbres sont distribués de façon homogène sur une surface, avec une densité ρ par unité d'aire.

Nous allons nous intéresser à la probabilité que deux arbres, pris au hasard dans chacune des parcelles appartiennent à la même espèce, en fonction de la distance r entre les parcelles : $F(r)$.

Dans un premier temps, nous nous plaçons en dimension 2 et nous nous intéressons à $F_2(x, y, t)$, la probabilité qu'un arbre choisi au hasard au temps t dans le voisinage $dS_x = dx_1 dx_2$ incluant le site $x = (x_1, x_2)$ soit de la même espèce qu'un arbre choisi au hasard dans un voisinage $dS_y = dy_1 dy_2$ incluant le site $y = (y_1, y_2)$.

Comment relier $F_2(x, y, t + dt)$ à $F_2(x, y, t)$?

Pour répondre à cette question nous faisons les hypothèses suivantes :

- (i) la distribution spatiale des arbres est un processus ponctuel de Poisson d'intensité ρ ,
- (ii) tout arbre en vie au temps t a une probabilité dt de mourir au temps $t + dt$ indépendamment du sort des autres arbres, tel que le temps de vie moyen de chaque arbre est une unité de temps, indépendamment de son espèce,
- (iii) une portion $P(x, y)dS_y$ des graines d'un arbre au site x , atterit dans dS_y , où $P(x, y)$ est le « noyau de dispersion »,
- (iv) chaque nouvel arbre a une probabilité ν d'être d'une espèce entièrement nouvelle.

Proposition 1.2.0.3. *Nous avons la relation suivante :*

$$F_2(x, y, t + dt) = (1 - 2dt)F_2(x, y, t) + 2dt(1 - \nu)Q(x, y, t), \quad (1.1)$$

où

$$Q(x, y, t) = \int_u F_2(u, y, t)P(u, x)dS_u + \frac{1 - F_2(y, y, t)}{\rho}P(y, x),$$

avec $dS_y > 1/\rho$.

Démonstration. Un des deux arbres meurt entre t et $t + dt$, avec probabilité $2dt$ et un seul arbre meurt par unité de temps. Si l'arbre mort se situe en x , il est immédiatement remplacé par le descendant d'un arbre voisin. On note $Q(x, y, t)$ la probabilité que le parent de ce descendant soit conspécifique avec l'arbre en y .

Les arbres en x et y sont conspécifiques au temps $t + dt$ si :

- aucun des deux arbres n'est mort et ils étaient conspécifiques au temps t ,
- un des deux arbres est mort, disons celui en x , il est remplacé par le descendant d'un arbre voisin, conspécifique avec l'arbre en y et il n'y a pas eu de mutation.

On obtient :

$$F_2(x, y, t + dt) = (1 - 2dt)F_2(x, y, t) + 2dt(1 - \nu)Q(x, y, t)$$

Déterminons à présent $Q(x, y, t)$.

La probabilité que l'arbre en x soit remplacé par le descendant d'un arbre dans le voisinage dS_u de $u = (u_1, u_2)$ est $P(u, x)dS_u$, et la probabilité que ce parent soit conspécifique avec l'arbre en y est $F_2(u, y, t)$.

Si l'arbre en x est remplacé par le descendant d'un arbre dans le voisinage dS_y de y , son parent a une probabilité $1/(\rho dS_y)$ d'être en y et $1 - 1/(\rho dS_y)$ d'être un autre arbre dans le voisinage, qui plus tard aura une probabilité $F_2(y, y, t)$ d'être conspécifique avec l'arbre en y .

On a $dS_y > 1/\rho$. (c'est-à-dire qu'il y a au moins un individu dans dS_y), donc $1 - 1/(\rho dS_y) > 0$; c'est bien défini. Donc on a :

$$Q(x, y, t) = \int_u F_2(u, y, t)P(u, x)dS_u - F_2(y, y, t)P(y, x)dS_y \\ + \left(\frac{1}{\rho dS_y} + \left(1 - \frac{1}{\rho dS_y}\right)F_2(y, y, t) \right) P(y, x)dS_y,$$

c'est-à-dire :

$$Q(x, y, t) = \int_u F_2(u, y, t)P(u, x)dS_u + \frac{1}{\rho dS_y} (1 - F_2(y, y, t)) P(y, x)dS_y,$$

donc finalement :

$$Q(x, y, t) = \int_u F_2(u, y, t)P(u, x)dS_u + \frac{1 - F_2(y, y, t)}{\rho} P(y, x).$$

□

En passant à la limite continue dans l'égalité (1.1), on obtient l'équation principale :

$$\frac{\partial F_2(x, y, t)}{\partial t} = -2F_2(x, y, t) + 2(1 - \nu)Q(x, y, t) \quad (1.2)$$

Si on suppose que la probabilité que deux arbres soient de la même espèce ne dépend que de leurs positions relatives :

$$\begin{cases} F_2(x, y, t) &= F_2(x - y, t) \\ P(x, y) &= P(x - y) \end{cases}$$

En faisant le changement de variable $z = y - x$, on a :

$$\frac{\partial F_2(z, t)}{\partial t} = -2F_2(z, t) + 2(1 - \nu) \left(\int_u F_2(u - z - x, t)P(u - x)dS_u + \frac{1 - F_2(0, t)}{\rho} P(z) \right).$$

En faisant le changement de variable $v = u - x$, on a :

$$\frac{\partial F_2(z, t)}{\partial t} = -2F_2(z, t) + 2(1 - \nu) \left(\int_v F_2(v - z, t)P(v)dS_v + \frac{1 - F_2(0, t)}{\rho} P(z) \right).$$

L'équation principale (1.2) devient donc :

$$\frac{\partial F_2(x, t)}{\partial t} = -2F_2(x, t) + 2(1 - \nu) \left(\int_u P(u)F_2(x - u, t)dS_u + \frac{1 - F_2(0, t)}{\rho} P(x) \right),$$

et à l'équilibre, on obtient :

$$F_2(x) = (1 - \nu) \left(\int_u P(u) F_2(x - u) du + \frac{1 - F_2(0)}{\rho} P(x) \right),$$

soit :

$$F_2(x_1, x_2) = (1 - \nu) \left(\int_{u_1, u_2} P(u_1, u_2) F_2(x_1 - u_1, x_2 - u_2) du_1 du_2 + \frac{1 - F_2(0, 0)}{\rho} P(x_1, x_2) \right). \quad (1.3)$$

1.3 Equation de Malécot

Nous allons dériver une équation formelle de l'équation précédente (1.3).

Définition 1.3.0.4. *La transformée de Fourier bi-dimensionnelle et son inverse, pour une fonction isotropique $Z(r)$ sont :*

$$z(q) = 2\pi \int_{r=0}^{\infty} Z(r) J_0(rq) r dr,$$

$$Z(r) = \frac{1}{2\pi} \int_{q=0}^{\infty} z(q) J_0(rq) q dq,$$

où J_0 est la fonction de Bessel de première espèce, d'ordre 0.

Théorème 1.3.0.5. *Pour $\nu > 0$, nous avons la formule suivante :*

$$F_2(r) = \frac{1 - F_2(0)}{\rho} \frac{1}{2\pi} \sum_{j=1}^{\infty} (1 - \nu)^j \int_{q=0}^{\infty} p(q)^j J_0(rq) q dq, \quad (1.4)$$

où J_0 est la fonction de Bessel de première espèce, d'ordre 0.

Démonstration. Nous allons prendre la transformée de Fourier de notre équation (1.3). On pose :

$$\begin{cases} f_2(k, m) & = \text{la transformée de Fourier de } F_2(x, y) \\ p(k, m) & = \text{la transformée de Fourier de } P(x, y) \end{cases}$$

Nous avons :

$$F_2(k, m) = (1 - \nu) \left(p(k, m) f_2(k, m) + \frac{1 - F_2(0)}{\rho} p(k, m) \right).$$

De plus, nous avons $|(1 - \nu)p(k, m)| < 1$, donc :

$$f_2(k, m) = \frac{1 - F_2(0)}{\rho} \underbrace{(1 - \nu)p(k, m)}_{1 - (1 - \nu)p(k, m)},$$

somme des termes d'une série géométrique de raison $(1 - \nu)p(k, m)$

d'où :

$$f_2(k, m) = \frac{1 - F_2(0)}{\rho} \sum_{j=1}^{\infty} (1 - \nu)^j p(k, m)^j.$$

On suppose que $P(x, y)$ et $F_2(x, y)$ sont des fonctions uniquement de $r = \sqrt{x^2 + y^2}$.

Alors, leurs transformées de Fourier $p(k, m)$ et $f_2(k, m)$ sont des fonctions uniquement de $q = \sqrt{k^2 + m^2}$.

Donc on obtient :

$$f_2(q) = \frac{1 - F_2(0)}{\rho} \sum_{j=1}^{\infty} (1 - \nu)^j p(q)^j. \quad (1.5)$$

Si on prend la transformée de Fourier inverse de (1.5), on a :

$$F_2(r) = \frac{1 - F_2(0)}{\rho} \frac{1}{2\pi} \int_{q=0}^{\infty} \sum_{j=1}^{\infty} (1 - \nu)^j p(q)^j J_0(rq) q dq.$$

Pour $\nu > 0$, les séries sont absolument convergentes, nous pouvons donc utiliser Fubini et intervertir somme et intégrale, d'où le résultat :

$$F_2(r) = \frac{1 - F_2(0)}{\rho} \frac{1}{2\pi} \sum_{j=1}^{\infty} (1 - \nu)^j \int_{q=0}^{\infty} p(q)^j J_0(rq) q dq.$$

□

On suppose maintenant que le noyau de dispersion $P(r)$ a une variance finie $2\sigma^2$, et $r \gg \sigma$.

Définition 1.3.0.6.

$$\sigma^2 = \int_{x,y} x^2 P(x, y) dx dy = \int_{x,y} y^2 P(x, y) dx dy$$

Théorème 1.3.0.7. Si $r \gg \sigma$ et $P(r)$ décroît avec r plus vite qu'une exponentielle négative, on a :

$$F_2(x, y) \simeq \frac{(1 - \nu)\sigma^2}{2\nu} \Delta F_2(x, y) \quad (1.6)$$

où Δ est l'opérateur Laplacien.

Démonstration. Comme $r \gg \sigma$, on peut négliger le terme multiplicatif $1 - F_2(0, 0)$, d'où :

$$F_2(x, y) \simeq (1 - \nu) \int P(w, z) F_2(x - w, y - z) dw dz.$$

Comme $P(r)$ décroît avec r plus vite qu'une exponentielle négative (comme c'est le cas pour le noyau gaussien par exemple), alors on peut développer $F_2(x - w, y - z)$ en série de Taylor au voisinage de (x, y) et négliger les termes d'ordre supérieur à deux. D'où :

$$\begin{aligned} F_2(x - w, y - z) \simeq & F_2(x, y) - \left(w \frac{\partial F_2(x, y)}{\partial x} + z \frac{\partial F_2(x, y)}{\partial y} \right) \\ & + \frac{1}{2} \left(w^2 \frac{\partial^2 F_2(x, y)}{\partial x^2} + wz \frac{\partial^2 F_2(x, y)}{\partial x \partial y} + z^2 \frac{\partial^2 F_2(x, y)}{\partial y^2} \right). \end{aligned}$$

Comme $P(x, y)$ est isotrope alors

$$\int_{x,y} xyP(x, y)dxdy = 0.$$

De plus, nous avons aussi :

$$\int_{x,y} P(x, y)dxdy = 1,$$

$$\int_{x,y} xP(x, y)dxdy = \int_{x,y} yP(x, y)dxdy = 0.$$

D'où :

$$F_2(x, y) \simeq (1 - \nu) \left(F_2(x, y) + \frac{1}{2}\sigma^2 \left[\frac{\partial^2 F_2(x, y)}{\partial x^2} + \frac{\partial^2 F_2(x, y)}{\partial y^2} \right] \right).$$

C'est-à-dire :

$$F_2(x, y) \simeq \frac{(1 - \nu)\sigma^2}{2\nu} \Delta F_2(x, y).$$

□

Si maintenant nous passons en coordonnées polaires et nous supposons que $\nu \ll 1$, nous avons immédiatement l'équation de Bessel :

$$\frac{d^2 F_2(r)}{dr^2} + \frac{1}{r} \frac{dF_2(r)}{dr} = (2\nu/\sigma^2)F_2(r).$$

Et nous savons que la seule solution positive telle que $\lim_{r \rightarrow \infty} F_2(r) = 0$ est :

$$F_2(r) = cK_0 \left(\frac{r\sqrt{2\nu}}{\sigma} \right)$$

où K_0 est la fonction de Bessel modifiée de seconde espèce, d'ordre 0 et c est une constante.

Cette équation de F_2 est vérifiée pour $r \gg \sigma$, mais ne peut pas être valable pour r petit car $K_0(r)$ diverge logarithmiquement quand r approche zéro. Pour r grand, nous allons pouvoir faire l'approximation :

$$K_0(r) \simeq \sqrt{\frac{\pi}{2r}} \exp(-r).$$

1.4 Noyau de Gauss

Nous nous plaçons dans le cas où $P(r)$ est la densité gaussienne radialement symétrique. Pour commencer, nous cherchons une expression de la fonction de similarité en dimension deux ; dans ce cas, nous savons calculer exactement (1.4). Ensuite, nous nous intéresserons au cas multi-dimensionnel.

1.4.1 En dimension deux

Proposition 1.4.1.1. *Si $P(r)$ admet pour transformée de Fourier $p(q)$, alors $p(q)^j$ est la transformée de Fourier de la densité de probabilité d'une somme de j variables aléatoires indépendantes, chacune de densité $P(r)$.*

$$P(r) = \frac{1}{2\pi\sigma^2} \exp(-r^2/2\sigma^2) \Rightarrow p(q) = \exp(-\sigma^2 q^2/2),$$

$$p(q)^j = \exp(-j\sigma^2 q^2/2) \Rightarrow \text{densité} = \frac{1}{2\pi j\sigma^2} \exp(-r^2/2j\sigma^2)$$

D'après ce qui précède, en prenant la transformée de Fourier inverse de l'équation (1.5), nous pouvons donc écrire :

$$F_2(r) = \frac{1}{2\pi\sigma^2} \frac{1 - F_2(0)}{\rho} \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right).$$

Proposition 1.4.1.2.

$$F_2(0) = \frac{\ln(1/\nu)}{2\rho\pi\sigma^2 + \ln(1/\nu)}.$$

Démonstration. Pour $r = 0$, nous avons :

$$F_2(0) = \frac{1}{2\pi\sigma^2} \frac{1 - F_2(0)}{\rho} \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j}.$$

Or, nous savons que :

$$\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} = \ln\left(\frac{1}{\nu}\right),$$

d'où :

$$F_2(0) = \frac{\ln(1/\nu)}{2\rho\pi\sigma^2 + \ln(1/\nu)}.$$

□

Remarque 1.4.1.3. *Pour les écologues, $1 - F_2(0)$ est l'indice de similarité de Simpson.*

Ce qui précède nous amène à

$$F_2(r) = \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right). \quad (1.7)$$

Théorème 1.4.1.4. *Si $r > \sigma$ et $\nu < 10^{-6}$:*

$$F_2(r) \simeq \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} 2K_0\left(\frac{r\sqrt{2\nu}}{\sigma}\right). \quad (1.8)$$

Si $r < \sigma$ et $\nu < 0.01$:

$$F_2(r) \simeq \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} \left[\ln\left(\frac{1}{\nu}\right) - \frac{r^2\pi^2}{12\sigma^2} \right]. \quad (1.9)$$

Démonstration. Démontrons d'abord (1.8) à partir de l'expression (1.7). Soient $r \gg \sigma$ et $\nu < 10^{-6}$, de telle sorte que nous pouvons remplacer $(1 - \nu)^j$ par $\exp(-j\nu)$:

$$\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right) \simeq \sum_{j=1}^{\infty} \frac{\exp(-j\nu)}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right).$$

La somme dans (1.7) est une série de Riemann ; nous avons finalement :

$$\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right) \simeq \int_{y=0}^{\infty} \exp\left(-y\nu - \frac{r^2}{2y\sigma^2}\right) \frac{dy}{y}.$$

Or, cette intégrale est connue (Gradshteyn and Ryzhik, 2000, (3.471.9)), elle vaut :

$$2K_0(r\sqrt{2\nu}/\sigma).$$

Par conséquent, nous obtenons (1.8) :

$$F_2(r) \simeq \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} 2K_0\left(\frac{r\sqrt{2\nu}}{\sigma}\right).$$

Démontrons à présent (1.9) à partir de l'expression (1.7). Soient $\nu < 0.01$ et $r \ll \sigma$, de telle sorte que $\exp(-r^2/2j\sigma^2) \simeq 1 - r^2/2j\sigma^2$:

$$\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \exp\left(-\frac{r^2}{2j\sigma^2}\right) \simeq \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} \left(1 - \frac{r^2}{2j\sigma^2}\right).$$

Ainsi, nous avons

$$F_2(r) \simeq \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} \left[\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j} - \frac{r^2}{2\sigma^2} \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j^2} \right].$$

Comme $\nu < 0.01$, alors nous avons

$$\sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{j^2} \simeq \frac{\pi^2}{6}.$$

Nous en déduisons (1.9) :

$$\begin{aligned} F_2(r) &\simeq \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} \left[\ln\left(\frac{1}{\nu}\right) - \frac{r^2\pi^2}{12\sigma^2} \right] \\ &\simeq F_2(0) - \frac{1}{2\rho\pi\sigma^2 + \ln(1/\nu)} \frac{r^2\pi^2}{12\sigma^2} \end{aligned}$$

□

1.4.2 Noyaux gaussiens multidimensionnels

Nous pouvons généraliser (1.7) à tout espace euclidien de dimension d .

Définition 1.4.2.1. *Le noyau de dispersion gaussien (symétrique, radial) en dimension d est défini par :*

$$P(x_1, \dots, x_d) = \frac{1}{(\sqrt{2\pi}\sigma)^d} \exp\left(-\frac{\sum_{i=1}^d x_i^2}{2\sigma^2}\right).$$

Dans la suite, nous noterons F_d la fonction de similarité en dimension d .

Théorème 1.4.2.2. *Pour $r > \sigma$ et $\nu < 10^{-6}$:*

$$F_d(r) \simeq \frac{1 - F_d(0)}{\rho} \frac{2}{(2\pi\sigma^2)^{d/2}} \left(\frac{r^2}{2\nu\sigma^2}\right)^{2-d/4} \times K_{d/2-1}\left(\frac{r\sqrt{2\nu}}{\sigma}\right)$$

où $K_{d/2-1}$ est la fonction de Bessel de seconde espèce d'ordre $d/2 - 1$.

Démonstration. Comme la transformée de Fourier $p(q)$ de la densité gaussienne $P(r)$ qui est $\exp(-\sigma^2 q)$ pour toute dimension d , nous trouvons

$$F_d(r) = \frac{1 - F_d(0)}{\rho} \sum_{j=1}^{\infty} \frac{(1 - \nu)^j}{(\sigma\sqrt{2j\pi})^d} \exp\left(-\frac{r^2}{2j\sigma^2}\right) \quad (1.10)$$

Comme dans le cas à deux dimensions, d'après $\nu < 10^{-6}$, on peut poser $(1 - \nu)^j \simeq \exp(-\nu j)$, on remplace la somme par l'intégrale et on utilise une formule de la fonction de Bessel de seconde espèce d'ordre $d/2 - 1$ (Gradshteyn et Ryzhik, 2000, (3.741.9)), pour obtenir finalement, pour $r > \sigma$:

$$F_d(r) \simeq \frac{1 - F_d(0)}{\rho} \frac{2}{(2\pi\sigma^2)^{d/2}} \left(\frac{r^2}{2\nu\sigma^2}\right)^{2-d/4} \times K_{d/2-1}\left(\frac{r\sqrt{2\nu}}{\sigma}\right)$$

□

Remarque 1.4.2.3. *Pour les problèmes biologiques, cette équation n'est utile qu'en dimension 1, 2 ou 3.*

Théorème 1.4.2.4.

$$F_1(r) \simeq \frac{1}{1 + \rho\sigma\sqrt{2\nu}} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right) \quad (1.11)$$

$$F_3(r) \simeq \frac{1 - F_3(0)}{2\pi\rho\sigma^2 r} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right) \quad (1.12)$$

Démonstration. On utilise le fait que

$$K_{1/2}(x) = K_{-1/2}(x) = \sqrt{\frac{\pi}{2x}} \exp(-x).$$

□

Remarque 1.4.2.5. $F_1(r)$ décroît de façon exponentielle avec r et ne diverge pas quand r approche 0.

Pour déterminer $F_3(0)$, nous devons partir de (1.10), car (1.12) n'a pas de limite évidente quand r approche 0.

Avec une erreur ne dépassant pas 1% pour $\nu \leq 10^{-4}$, nous pouvons écrire :

$$\sum_{j=1}^{\infty} \frac{(1-\nu)^j}{j^{3/2}} \simeq \sum_{j=1}^{\infty} \frac{1}{j^{3/2}} = \zeta(3/2)$$

où $\zeta(s) = \sum_{j=1}^{\infty} \frac{1}{j^s}$ est la fonction Zeta de Riemann.

Dans ce cas, nous obtenons :

$$F_3(0) \simeq \frac{\zeta(3/2)}{\rho(2\pi\sigma^2)^{3/2} + \zeta(3/2)}$$

Finalement, nous avons :

$$F_3(r) \simeq \frac{\sqrt{2\pi}\sigma}{\rho r(2\pi\sigma^2)^{3/2} + \zeta(3/2)r} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right) \quad (1.13)$$

Remarque 1.4.2.6. La quantité $F_d(0)$ représente en fait la probabilité que deux individus pris au hasard dans une même parcelle soient de la même espèce. C'est une mesure de l' α -diversité (diversité des espèces qui coexistent dans une région). Quand d augmente de 1 à 3, $F_d(0)$ devient progressivement moins sensible au taux de spéciation ν : $F_1(0)$ dépend de ν en $\sqrt{\nu}$, $F_2(0)$ en dépend logarithmiquement et $F_d(0)$, $d \geq 3$, ne dépend plus de ν , pour ν suffisamment petit. Par ailleurs, quand ν approche 0, pour r (suffisamment) grand, nous avons :

$$F_d(r) \sim \frac{1}{r^{d-2}}.$$

La similarité en une, deux et trois dimensions de l'espace est comparée dans la figure 1.1.

Théorème 1.4.2.7. (Sawyer, 1997a) Pour des noyaux de dispersion avec moments finis :

- En dimension 1, pour tout noyau de dispersion tel que $\int r^5 P(r) dr < \infty$, $F_1(r)$ a une expression proportionnelle à celle de l'équation (1.11) avec un terme de correction d'amplitude $\mathcal{O}(\nu(1+1/r))$.

- En dimension 2, pour tout noyau de dispersion tel que $\int r^3 P(r) dr < \infty$, $F_2(r)$ a une expression proportionnelle à celle de l'équation (1.9) à l'ordre dominant, avec un terme de correction d'amplitude $\mathcal{O}(\nu)$.

- En dimension 3, pour tout noyau de dispersion tel que $\int r^2 P(r) dr < \infty$, $F_3(r)$ a une expression proportionnelle à celle de l'équation (1.12) à l'ordre dominant, avec un terme de correction d'amplitude $\mathcal{O}(\nu)$.

Ici, nous disons qu'une fonction $f(r)$ est proportionnelle à une fonction $g(r)$ si la limite $\lim_{r \rightarrow \infty} (f(r)/g(r))$ est une constante indépendante de r . Par conséquent, pour tout noyau de dispersion $P(r)$ qui décroît assez rapidement pour r grand, les équations (1.11), (1.9) et (1.12) décrivent précisément les comportements de $F_d(r)$, $d \leq 3$, pour r grand.

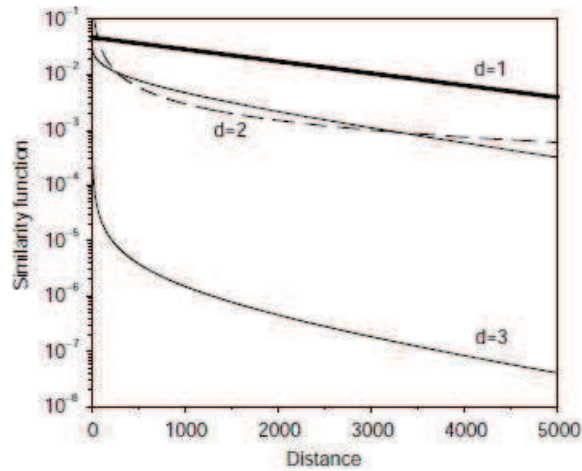


FIG. 1.1 – [Chave and Leigh, 2002] Prédictions théoriques pour la fonction de similarité en une, deux et trois dimensions de l'espace avec le noyau de dispersion Gaussien (lignes pleines) et en deux dimension pour le noyau de dispersion de Cauchy (ligne pointillée). Pour le noyau Gaussien, $\sigma = 8$ et $\nu = 10^{-6}$ et pour le noyau de Cauchy, $c = 1$. Pour une meilleure comparaison, la fonction correspondant au noyau de Cauchy a été multipliée par 3.

1.5 Noyaux « fat-tailed » (à queue lourde)

Lorsque le moment d'ordre trois du noyau de dispersion n'est pas fini, le théorème limite de Sawyer ne peut pas être appliqué. Bien que les noyaux de dispersion avec variance infinie aient longtemps été considérés comme inappropriés pour des problèmes écologiques, puisque les échantillons finis ont des variances finies, il a été montré que les noyaux de dispersion « fat-tailed » (à queue lourde) peuvent expliquer certains phénomènes écologiques.

Ici, nous allons nous intéresser aux noyaux « fat-tailed » de la forme suivante (Clark et al., 1999) :

$$P(r) = \frac{p}{\pi c^2 (1 + r^2/c^2)^{p+1}} \quad (1.14)$$

Le k -ième moment est fini si et seulement si $k < 2p$. Donc le théorème de Sawyer est valable si et seulement si $p > 3/2$.

Quelle est la forme de $F(r)$ quand le noyau de dispersion est à queue lourde ?

En dimension deux, la transformée de Fourier de (1.14) est :

$$\begin{aligned} p(q) &= 2\pi \int_0^\infty \frac{pc^{2p}}{\pi(c^2 + r^2)^{p+1}} J_0(rq) r dr \\ &= 2p \left(\frac{cq}{2}\right)^p \frac{K_p(cq)}{\Gamma(p+1)} \end{aligned}$$

où K_p est la fonction de Bessel modifiée de deuxième espèce d'ordre p .

Ici, la distance de dispersion médiane est $c\sqrt{3}$ et nous avons :

$$F_2(r) = \frac{1 - F_2(0)}{\rho} \frac{1}{2\pi} \sum_{j=1}^{\infty} (1 - \nu)^j \int_{q=0}^{\infty} \left(2p \left(\frac{cq}{2} \right)^p \frac{K_p(cq)}{\Gamma(p+1)} \right)^j J_0(rq) q dq.$$

Remarque 1.5.0.8. Si $p = 1/2$, le noyau de dispersion « fat-tailed » (1.14) correspond au noyau de Cauchy en dimension deux, $c/(2\pi(c^2 + r^2)^{3/2})$. Cette densité admet pour transformée de Fourier $p(q) = \exp(-cq)$. La densité de Cauchy en dimension d est $C_d c / (c^2 + r^2)^{(d+1)/2}$. Ici, nous avons $C_1 = 1/\pi$, $C_2 = 1/2\pi$, et $C_3 = 1/\pi^2$.

Théorème 1.5.0.9. En dimension d , la fonction de similarité correspondant au noyau de Cauchy est :

$$F_d(r) = C_d \frac{1 - F_d(0)}{\rho} \sum_{j=1}^{\infty} (1 - \nu)^j \frac{j c}{(r^2 + j^2 c^2)^{(d+1)/2}} \quad (1.15)$$

Théorème 1.5.0.10. Si le noyau de dispersion est de Cauchy, avec pour transformée de Fourier $p(q) = \exp(-cq)$, si $d > 1$, la fonction de similarité $F_d(r)$ est indépendante de r , lorsque $r \ll c$, alors que $F_d(r) \sim r^{1-d}$, lorsque $r \gg c$.

Remarque 1.5.0.11. Dans le cas de la dimension deux, la fonction de similarité correspondant au noyau de dispersion de Cauchy est comparé à la fonction de similarité correspondant au noyau de Gauss, à la figure 1.1 page 16. Pour $d \geq 2$, le comportement de $F_d(r)$ avec le noyau de Cauchy tend à ressembler au comportement de $F_{d+1}(r)$ avec le noyau de Gauss.

1.6 Modèle du votant et limite de non-spéciation

Maintenant, nous considérons le cas limite où le taux de spéciation approche zéro. On s'intéresse ici au cas où $\sigma < \infty$.

Quand $\sigma < \infty$, l'approximation de diffusion (1.6) s'applique avec $\Delta F(r) = 0$ quand $\nu = 0$. Dans ce cas, la seule solution stable est $F(r) = C$ et l'équation (1.3) nous donne $C = 1$. Ce qui implique qu'une unique espèce peut éventuellement pousser sur l'ensemble de la forêt.

Maintenant, nous ne nous plaçons plus dans le cas stationnaire. On suppose qu'il n'y a pas de spéciation, et qu'au temps $t = 0$, les N arbres de la forêt sont d'une espèce différente. Alors, après un grand nombre t de générations d'arbres, le nombre total d'espèces restantes serait $\sim N \ln(t)/t$ (résultat connu pour le modèle du votant).

De plus, d'après Cox et Griffeath (1986), pour t grand et r grand :

$$F_2(r) \sim 1 - \frac{\ln(\gamma r^2 / \sigma^2)}{\ln(t)}$$

où γ est la constante d'Euler.

Ainsi, si $\sigma < \infty$, $F_2(r)$ converge toujours vers 1, mais très lentement.

Quand $\sigma < \infty$, la variation de $F(r)$ dans le temps peut être simulée en utilisant une analogie avec le « modèle du votant » ; dans ce modèle, un « votant » est localisé à chaque site d'un réseau bidimensionnel. Au début, chaque « votant » préfère un candidat différent, et à chaque étape, chaque « votant » choisit au hasard un de ses 4 plus proches voisins et adopte sa préférence.

Définition 1.6.0.12. *Nous définissons le **modèle du votant pour une forêt** de la manière suivante :*

- un arbre est localisé à chaque site d'un réseau bidimensionnel qui représente la forêt,
- au temps $t = 0$, chaque arbre est d'une espèce différente,
- chaque arbre mort est remplacé par le descendant d'un de ses 4 plus proches voisins.

Nous en déduisons un modèle de marches aléatoires coalescentes de la manière suivante : au temps t , l'arbre au site x a d'une espèce transmise par exactement un antécédent, pour chaque temps $t - \tau < t$. Soit $A_x(\tau)$ la position de cet antécédent au temps $t - \tau$, tel que $A_x(0) = x$. Le processus $(A_x(\tau))_{\tau \geq 0}$ est une marche aléatoire, où la distance à chaque pas de temps $\|A_x(\tau) - A_x(\tau + 1)\|_2$ est égale à 1.

Comme τ augmente, $t - \tau$ diminue, de sorte que ces marches aléatoires voyagent en arrière dans le temps.

Si au temps $t = 0$ chaque arbre est d'une espèce différente, alors $F_2(x, y, t)$ représente la probabilité qu'au temps t les arbres en x et y soient de la même espèce, autrement dit, la probabilité que leurs espèces ont été transmises par le même ancêtre. C'est-à-dire :

$$F_2(x, y, t) = \mathbb{P}[\exists \tau, A_x(\tau) = A_y(\tau)]$$

Au temps $\min\{\tau; A_x(\tau) = A_y(\tau)\}$, les marches aléatoires de ces sites coalescent et pour tout τ plus grand, il n'y a plus qu'une marche aléatoire (qui correspond à leur ancêtre commun).

Si les préférences se transmettent sans spéciation, le modèle du votant prédit que, pour des systèmes en une ou deux dimensions, une préférence l'emportera finalement, tandis qu'en dimension trois ou plus, différentes préférences coexisteront. Nous avons cela pour les mêmes raisons que différentes marches aléatoires se rencontrent presque sûrement, en temps suffisant, en dimension un ou deux, mais pas en dimension trois ou plus.

La diffusion d'une espèce à toute la forêt est décrite par le résultat suivant :

Théorème 1.6.0.13. (Théorème de Cox, 1989)

- Si au temps $t = 0$ chacun des N arbres est d'une espèce différente, le temps moyen τ requis pour qu'une espèce envahisse la parcelle est :

$$\tau \simeq \frac{2N}{\pi} (\ln(N))^2.$$

- Si au temps $t = 0$ une proportion p_s des N arbres est de l'espèce s , le temps τ requis pour qu'une espèce envahisse la parcelle est :

$$\tau \simeq \frac{2H_s}{\pi} N \ln(N), \quad H_s = - \sum_{s=1}^S p_s \ln(p_s),$$

où H_s est l'entropie de l'état initial.

1.7 Expériences numériques

Les formules dérivées dans les sections précédentes s'appliquent pour des espaces géographiques infinis. Cependant, les systèmes réels sont finis. Le théorème de Cox montre que lorsque $\nu = 0$ le temps requis pour qu'une espèce envahisse la parcelle dépend de la taille de la parcelle et de la composition en espèces initiale. Lorsque $\nu > 0$, le fait que la forêt soit finie peut affecter la fonction de similarité $F(r)$ de diverses manières.

Ici, nous allons utiliser des simulations numériques pour aborder ces questions.

Pour simuler le modèle neutre de Hubbell, avec une dispersion limitée, nous considérons une forêt de $N = L^2$ arbres sur un réseau carré, où à chaque pas de temps (égal à une génération) :

- (i) un arbre choisi au hasard, indépendamment de son espèce, meurt ;
- (ii) l'arbre mort est immédiatement remplacé par le descendant d'un de ses voisins ;
- (iii) avec probabilité $2\pi P(r)rdr$, la distance de l'arbre qui vient de mourir au parent de son remplaçant est comprise entre r et $r + dr$, où $P(r)$, le noyau de dispersion, est le même pour tous les arbres ;
- (iv) avec probabilité ν , le remplaçant est un mutant (une espèce complètement nouvelle).

Pour évaluer la fonction de similarité, nous procédons de la manière suivante : on fait varier ν et $P(r)$, et pour chaque couple $(\nu, P(r))$, on fait 50 – 200 répétitions. Ensuite, on évalue $F(r)$ pour chaque répétition. Pour chacune des répétitions, quand l'équilibre est atteint, on choisit au hasard $100N \simeq 10^8$ paires d'individus du réseau et on demande s'ils sont conspécifiques.

Remarque 1.7.0.14. *En évaluant $F(r)$ de cette façon, il doit quand même y avoir des problèmes de corrélations au sein de chaque échantillon...*

En fait, nous simulons les marches aléatoires du modèle des votants. On commence au pseudo-temps $\tau = 0$ avec une marche aléatoire par site. Chaque marche aléatoire se déplace en moyenne une fois par pas de temps et les marches aléatoires coalescent toujours à la rencontre. Un événement de spéciation représente l'anéantissement d'une marche aléatoire. Par définition, la première représentation d'une nouvelle espèce n'a pas d'ancêtre conspécifique. Toutes les marches aléatoires qui ont précédemment coalescé avec l'anéantissement d'une marche définit un ensemble de sites (les positions initiales de ces marches aléatoires) auxquels on fixe un label d'espèce. On exécute des simulations tant que toutes les marches n'ont pas coalescé ou subi un anéantissement. A ce moment, chaque site dans le réseau a un label d'espèce, et on appelle une telle configuration, une distribution équilibrée d'espèces.

Remarque 1.7.0.15. *Dans l'algorithme des marches aléatoires, un pas de temps correspond à deux pas de temps de la version directe. Ainsi, le paramètre σ^2 du noyau de dispersion qu'on utilise dans les formules analytiques de la fonction de similarité devrait être multiplié par deux.*

Atteindre l'équilibre va prendre plus de temps quand ν est plus petit. Quand $\nu = 3.81 \times 10^{-6}$, le temps jusqu'à l'équilibre est 10^6 pas de temps ($\sim 2/\nu$ générations).

Les arbres morts vont être remplacés seulement par le descendant d'un de ses quatre plus proches voisins et on prend une distance entre les points du réseau égale à un ($\rho = 1$).

Nous comparons les prédictions du modèle avec la fonction de similarité $F(r)$ moyennée sur 50 répétitions lancées en utilisant le noyau de dispersion des plus proches voisins. Les résultats obtenus sont représentés figure 1.2. Nous constatons que pour $r > \sigma$ les résultats sont bons, alors que pour $r < \sigma$ ils sont moins satisfaisants.

Nous testons également les prédictions du modèle avec la fonction de similarité $F(r)$ moyennée sur 200 répétitions lancées en utilisant le noyau de dispersion gaussien, en supposant que la probabilité qu'un descendant remplace son parent est nulle. Les résultats obtenus sont représentés figure 1.3. Nous constatons que lorsque σ^2 est plus petit, la similarité décroît plus vite, et si θ augmente alors elle décroît moins vite.

La table en figure 1.4, nous permet de dire que pour le noyau de dispersion des plus proches voisins, la correspondance entre les valeurs prédites de ξ et c , et les valeurs ajustées de notre $F(r)$ simulée sont en accord avec la théorie et pour le noyau de dispersion gaussien, les valeurs prédites de ξ sous-estiment les valeurs observées de ξ juste quand ξ est plus grand que la taille du système et sinon les résultats sont en accord avec les prédictions théoriques.

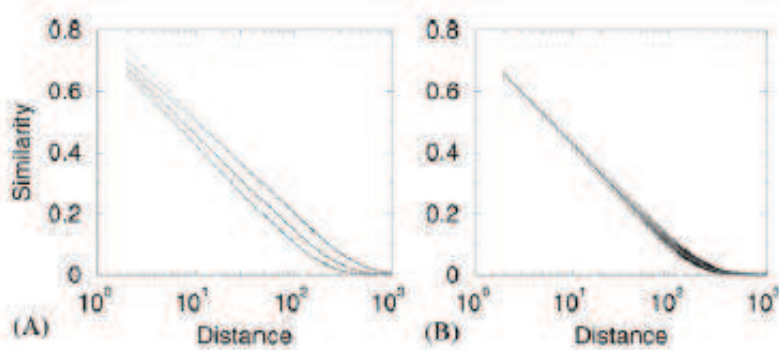


FIG. 1.2 – [Chave and Leigh, 2002] Fonction de similarité dans le modèle neutre avec dispersion des plus proches voisins. (A) Similarité moyennée sur 50 mesures, pour trois taux de spéciation différents, $\nu = \theta/N$, $\theta = 4, 8, 16$, $N = 1024^2$. Les lignes pointillées représentent les traces de la fonction $cK_0(x/\xi)$ pour des valeurs de c et ξ correspondant aux valeurs de θ (de haut en bas). (B) Similarité pour chacune des 50 mesures, pour $\theta = 16$, et similarité moyenne.

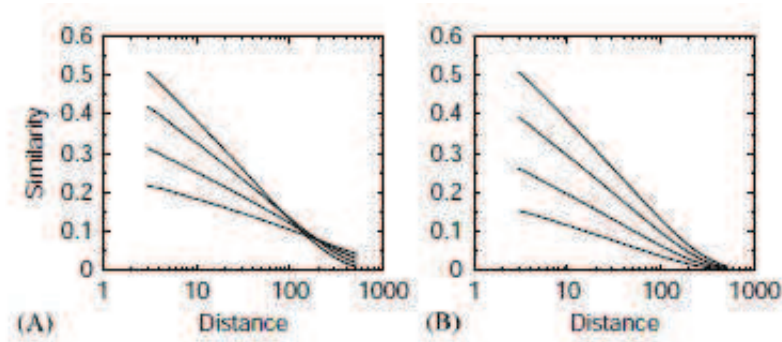


FIG. 1.3 – [Chave and Leigh, 2002] Fonction de similarité dans le modèle neutre avec dispersion gaussienne. Similarité moyennée sur 200 répétitions avec $\nu = \theta/N$, $N = 1024^2$. (A) Avec $\theta = 16$, $\sigma^2 = 2, 4, 8, 16$ (de haut en bas). (B) Avec $\xi = 256$, $\{\theta, \sigma^2\} = \{16, 1\}, \{32, 2\}, \{64, 4\}, \{128, 8\}$.

ν	σ^2	Observed S	Observed c	Predicted c	Observed ξ	Predicted ξ
3.81×10^{-6}	1	121	0.147	0.128	341	362
7.63×10^{-6}	1	222	0.150	0.134	215	256
1.53×10^{-5}	1	394	0.162	0.141	142	181
1.53×10^{-5}	2	264	0.121	0.115	239	256
1.53×10^{-5}	4	221	0.0953	0.0845	345	362
1.53×10^{-5}	8	198	0.0636	0.0552	585	512
1.53×10^{-5}	16	184	0.0390	0.0326	1295	724
3.05×10^{-5}	4	405	0.0970	0.0871	220	256
6.10×10^{-5}	8	670	0.0644	0.0574	215	256
1.22×10^{-4}	16	1165	0.0385	0.0337	209	256

FIG. 1.4 – [Chave and Leigh, 2002] Table de comparaison entre les valeurs prédites de ξ et c et celles trouvées en ajustant $cK_0(r/\xi)$ à la fonction $F(r)$ des simulations avec différents taux de spéciation $\nu = \theta/N$ et différents paramètres de dispersion σ ($N = 1024 \times 1024$). Le nombre moyen d'espèces dans le système, S , est également donné. Les trois premières lignes correspondent aux résultats avec le noyau de dispersion des plus proches voisins (figure 1.2) et les sept dernières à ceux avec le noyau gaussien (figure 1.3). Pour calculer les valeurs prédites de c , on a supposé que $C_0 = 0$.

1.8 Conclusion

Nous avons calculé la probabilité $F(r)$ que deux arbres séparés par une distance r soient de la même espèce, en supposant une concurrence équivalente, une dispersion limitée et un équilibre entre la spéciation et l'extinction, et nous avons pu étudier les simulations effectuées par Chave et Leigh (2002).

Nous avons appris comment la fonction de similarité $F(r)$ décroît avec l'augmentation de

la distance qui sépare les arbres comparés et quels sont les facteurs qui affectent la vitesse de la décroissance. $F(r)$ dépend de la dimension du système. En effet, si $d = 1$, $F(r)$ décroît plus lentement que lorsque $d = 2$. La fonction de similarité dépend également du noyau de dispersion considéré. Si le noyau est à queue lourde au lieu de gaussien, c'est comme si la forêt s'étendait à une dimension supérieure. En dimension d , avec un noyau de dispersion de Cauchy, nous obtenons des résultats semblables à ceux en dimension $d+1$, avec un noyau de dispersion de Gauss.

La théorie neutre suggère que la spéciation et la dispersion limitée influencent le renouvellement des espèces. En fait, on peut seulement résoudre le modèle neutre de β -diversité si $F(r)$ exprime un équilibre entre spéciation et extinction, mais cet équilibre est long à atteindre ($2/\nu$ pour le noyau des plus proches voisins, accélération de l'approche de l'équilibre avec un noyau de Cauchy). Finalement, la théorie neutre permet de faire des prédictions quantitatives et testables.

Chapitre 2

Inférence en deux dimensions basée sur les fréquences alléliques

Dans ce chapitre, nous présentons l'étude de l'article de N. H. Barton, A. M. Etheridge, J. Kelleher et A. Véber, soumis en 2013 à la revue « Theoretical Population Biology », *Inference in two dimensions : allele frequencies versus lengths of shared sequence blocks*.

Nous allons présenter une approche d'inférence de la taille de voisinage, \mathcal{N} , basée sur les fréquences alléliques. Sur des périodes intermédiaires (10-100 générations, par exemple), les populations qui vivent en deux dimensions approchent un quasi-équilibre qui est indépendant de leur structure locale et de leur histoire. Sur de telles échelles, la covariance normalisée des fréquences alléliques (ie. F_{ST} par paire) diminue avec le logarithme de la distance, et ne dépend que de la taille du voisinage, \mathcal{N} , et d'une « échelle locale », κ , de telle sorte que le taux de flux de gènes, σ^2 , ne peut pas être inféré. Nous allons montrer comment les corrélations spatiales peuvent être prises en compte, en supposant une distribution gaussienne des fréquences alléliques, ce qui nous permettra de donner des estimations du maximum de vraisemblance pour \mathcal{N} et κ .

2.1 Introduction

Au cours du dernier siècle, une grande partie des recherches en génétique des populations a été consacrée à donner un sens à la répartition spatiale. Les données génétiques peuvent être utilisées pour estimer les taux de flux de gènes et pour déduire l'histoire de la population. La structure de la population interagit également avec la sélection, empêchant l'adaptation et promouvant la divergence. Un modèle neutre de population structurée est essentiel si nous voulons détecter l'influence de la sélection à des loci spécifiques.

Ici, nous allons exploiter une caractéristique particulière des grandes populations spatialement structurées : les lignées ancestrales de deux gènes échantillonnés l'un à côté de l'autre vont soit coalescer rapidement, soit s'éloigner l'une de l'autre, et s'unir seulement dans un passé lointain (Wakeley, 2008). Cette répartition très variable des temps de coalescence reflète une séparation des échelles de temps entre dérive aléatoire locale et dérive aléatoire globale. Cela est considéré à la fois dans les modèles d'îles, et dans les populations qui sont réparties sur deux dimensions. Cette séparation permettra, en principe, l'inférence robuste de la structure

de la population locale, basée sur les fréquences alléliques à plusieurs loci.

Plusieurs modèles d'évolution en deux dimensions ont déjà été proposés. Dans tous ces modèles, sur des échelles de temps intermédiaires (suffisamment grandes pour que nous ne voyons pas les détails du mécanisme de reproduction locale, mais assez petites pour que nous ne voyons pas de nouvelles mutations ou les effets de sélection), la structure génétique n'est déterminée que par deux paramètres : le taux de diffusion des lignées ancestrales, σ^2 , et la taille de voisinage, \mathcal{N} . De plus, en utilisant le fait que la fonction génératrice du temps au plus récent ancêtre commun de deux individus puisse être interprétée comme la probabilité d'identité dans un modèle de mutation à infinité d'allèles, nous pourrions utiliser la formule de Wright-Malécot classique pour étudier le temps de coalescence des lignées ancestrales d'un échantillon de taille deux, pris sur de telles échelles spatiales intermédiaires. Dans tous ces modèles, la densité de la population locale est fixée.

Si nous nous intéressons à l'histoire profonde, seul un petit échantillon d'individus est nécessaire. Dans tous les cas, des grands échantillons vont rapidement fusionner et ainsi mener à quelques branches ancestrales. Cependant, de nombreux loci doivent être échantillonnés, puisque chaque locus aura une histoire singulière. En revanche, si nous nous intéressons à l'histoire récente, la coalescence est peu probable dans un passé récent et nous avons des informations négligeables sur la mutation dans une séquence courte. Il faut donc, par exemple, prendre de grands échantillons à de multiples loci (comme dans les sondages traditionnels sur F_{ST}). Dans notre contexte spatial, nous allons nous intéresser à l'estimation du paramètre \mathcal{N} . Notre objectif est d'illustrer comment, dans ce cadre, cette stratégie d'échantillonnage peut être utilisée.

Dans un premier temps, nous allons déterminer la formule de Wright-Malécot classique, qui déterminera la fonction génératrice du temps de coalescence de deux lignées. Nous mettrons en évidence la nécessité de déterminer une statistique F appropriée, puis nous décrirons comment estimer la taille de voisinage à partir des fréquences alléliques. Ensuite, nous montrerons que nous pouvons expliquer les corrélations spatiales, en ajustant une matrice de covariance, dans le cadre du modèle stepping stone. Nous utiliserons la dualité entre les fréquences alléliques et les généalogies, pour exprimer les corrélations entre les diversités génétiques locales et nous définirons une statistique F que nous utiliserons ensuite comme base pour une approche du maximum de vraisemblance afin d'estimer la taille de voisinage \mathcal{N} , à partir des fréquences alléliques dans la population, puis nous montrerons que cette statistique F peut être estimée par une quantité \mathcal{F} , que nous définirons. Enfin, après avoir estimé \mathcal{N} , nous présenterons quelques résultats numériques basés sur les fréquences alléliques.

2.2 Fréquences alléliques

2.2.1 La formule de Wright-Malécot

Dans tout ce qui suit, nous allons nous intéresser à une population qui est uniformément répartie sur un espace bi-dimensionnel (discret ou continu). Dans notre analyse, nous supposons toujours que l'espace est infini. Nous allons également supposer que le mouvement d'une lignée ancestrale peut être décrit par une marche aléatoire symétrique (ou mouvement brownien,

dans le cas continu) avec le paramètre de variance $\sigma^2 \in (0, \infty)$, et que deux lignées ne peuvent fusionner que localement, c'est-à-dire, quand elles sont raisonnablement proches l'une de l'autre.

Plus précisément, nous nous concentrons sur les modèles où la probabilité d'identité d'état est bien approchée par la formule de Wright-Malécot classique.

Théorème 2.2.1.1. *Si T est le temps de coalescence de deux lignées échantillonnées à une certaine distance x , et μ désigne la vitesse à laquelle les mutations tombent sur l'arbre généalogique, alors*

$$\mathbb{E}_x[e^{-2\mu T}] \simeq \begin{cases} \frac{K_0(x/l_\mu)}{\mathcal{N} + \log(l_\mu/\kappa)} & \text{pour } |x| > \kappa, \\ \frac{\log(l_\mu/\kappa)}{\mathcal{N} + \log(l_\mu/\kappa)} & \text{pour } |x| \leq \kappa, \end{cases}$$

où $l_\mu = \sigma/\sqrt{2\mu}$, $\mathcal{N} > 0$ est la taille de voisinage, κ est une échelle locale sur laquelle la probabilité d'identité d'état est à peu près constante et K_0 est la fonction de Bessel modifiée de seconde espèce d'ordre 0.

Remarque 2.2.1.2. *Il s'agit de la même formule de Malécot démontrée dans l'article de Chave et Leigh (2002). Ici, pour obtenir cette équation, nous allons utiliser la dualité du modèle neutre avec un modèle de marches aléatoires coalescentes.*

Démonstration. Pour démontrer ce résultat, nous travaillons avec un modèle stepping stone sur \mathbb{Z}^2 . Il y a $2N$ gènes dans chaque dème. Nous supposons que la population évolue en générations discrètes. Dans chaque génération, les premiers descendants sont générés par échantillonnage de Wright-Fisher dans chaque dème. Ensuite, une proportion $g_1(x - y)$ des descendants au dème x migre vers le dème y .

Plutôt que d'introduire un mécanisme de mutation, nous pensons à la formule de Wright-Malécot comme décrivant la fonction génératrice du nombre de générations en arrière jusqu'au plus récent ancêtre commun de deux individus échantillonnés à la séparation x (vecteur à deux dimensions) dans la population.

Soit $\psi_t(x)$ la probabilité que deux gènes échantillonnés de séparation x aient leur plus récent ancêtre commun exactement t générations dans le passé.

Pour $t > 1$, nous allons décomposer cette quantité en fonction de la séparation des ancêtres immédiats de ces deux gènes.

Si les deux gènes ont surgi comme des migrants du même dème, alors avec probabilité $1/2N$, ils ont un ancêtre commun à la génération précédente.

La probabilité que deux gènes échantillonnés de séparation x aient leur plus récent ancêtre commun à la génération précédente est donnée par $\psi_1(x)$. Il s'agit de la probabilité qu'ils aient un ancêtre commun à la génération précédente, multipliée par la probabilité que les deux gènes de séparation x aient surgi comme des migrants du même dème y (donnée par $G_1(x)$, le produit de convolution de deux copies de g_1). Ainsi,

$$\psi_1(x) = \frac{1}{2N}G_1(x),$$

En fait, $G_1(x)$ correspond à la loi de la séparation des deux lignées.

Si, d'autre part, ils ont des parents distincts, de séparation y , alors la probabilité que leur plus récent ancêtre commun soit t générations dans le passé est $\psi_{t-1}(y)$.

Pour $t > 1$, nous arrivons à la récurrence :

$$\psi_t(x) = \sum_y \left\{ G_1(x-y)\psi_{t-1}(y) - \frac{\mathbb{1}_{\{y=0\}}}{2N} G_1(x-y)\psi_{t-1}(0) \right\}. \quad (2.1)$$

En effet, comme $t > 1$, les deux gènes échantillonnés de séparation x n'ont pas leur plus récent ancêtre commun à la génération précédente. A cette génération, ils ont des parents distincts de séparation y avec probabilité $G_1(x-y)$, qui eux auront leur plus récent ancêtre commun $t-1$ générations dans le passé avec probabilité $\psi_{t-1}(y)$. Nous faisons donc une somme sur y . Cependant, lorsque $y = 0$, nous sommes dans le cas où les gènes sont des migrants du même dème et ils sont donc distincts avec probabilité $1 - 1/(2N)$, ce qui fait apparaître dans la somme le terme correctif $-\mathbb{1}_{\{y=0\}}G_1(x-y)\psi_{t-1}(0)/(2N)$.

Cela peut également s'écrire :

$$\psi_t(x) = \frac{1}{2N} \left(G_t(x) - \sum_{\tau=1}^{t-1} G_{t-\tau}(x)\psi_\tau(0) \right), \quad (2.2)$$

où G_t est la t -ème puissance de convolution de G_1 .

Cela peut se montrer par récurrence sur $t > 1$:

- La propriété est satisfaite pour $t = 2$:

$$\begin{aligned} \psi_2(x) &= \sum_y \left\{ G_1(x-y)\psi_1(y) - \frac{\mathbb{1}_{\{y=0\}}}{2N} G_1(x-y)\psi_1(0) \right\} \\ &= \sum_y \left\{ G_1(x-y)\frac{G_1(y)}{2N} - \frac{\mathbb{1}_{\{y=0\}}}{2N} G_1(x-y)\psi_1(0) \right\} \\ &= \frac{1}{2N} \left(G_2(x) - \sum_{\tau=1}^{2-1} G_{2-\tau}(x)\psi_\tau(0) \right) \end{aligned}$$

- On suppose que cette propriété est satisfaite pour $t-1$, c'est-à-dire :

$$\psi_{t-1}(x) = \frac{1}{2N} \left(G_{t-1}(x) - \sum_{\tau=1}^{(t-1)-1} G_{(t-1)-\tau}(x)\psi_\tau(0) \right).$$

- Montrons que cette propriété est vraie pour t :

$$\begin{aligned}
\psi_t(x) &= \sum_y \left\{ G_1(x-y)\psi_{t-1}(y) - \frac{\mathbb{1}_{\{y=0\}}}{2N} G_1(x-y)\psi_{t-1}(0) \right\} \\
&= \sum_y \left\{ G_1(x-y) \left[\frac{1}{2N} \left(G_{t-1}(y) - \sum_{\tau=1}^{(t-1)-1} G_{(t-1)-\tau}(y)\psi_\tau(0) \right) \right] \right. \\
&\quad \left. - \frac{1}{2N} G_1(x)\psi_{t-1}(0) \right\} \\
&= \frac{1}{2N} \left[\sum_y (G_1(x-y)G_{t-1}(y)) - G_1(x)\psi_{t-1}(0) \right. \\
&\quad \left. - \sum_y \sum_{\tau=1}^{t-2} (G_1(x-y)G_{(t-1)-\tau}(y)\psi_\tau(0)) \right]
\end{aligned}$$

Par Fubini, on peut intervertir les deux sommes :

$$\begin{aligned}
\psi_t(x) &= \frac{1}{2N} \left[\sum_y (G_1(x-y)G_{t-1}(y)) - G_1(x)\psi_{t-1}(0) \right. \\
&\quad \left. - \sum_{\tau=1}^{t-2} \left(\sum_y (G_1(x-y)G_{(t-1)-\tau}(y)) \psi_\tau(0) \right) \right] \\
&= \frac{1}{2N} \left[G_t(y) - G_{t-(t-1)}(x)\psi_{t-1}(0) - \sum_{\tau=1}^{t-2} (G_{t-\tau}(y)\psi_\tau(0)) \right] \\
&= \frac{1}{2N} \left[G_t(y) - \sum_{\tau=1}^{t-1} (G_{t-\tau}(y)\psi_\tau(0)) \right]
\end{aligned}$$

Ecrivons T pour le moment (aléatoire) auquel les deux gènes partagent leur plus récent ancêtre commun. La fonction génératrice de T , qui dépend bien sûr de la distance d'échantillonnage entre les deux gènes, est définie par $\phi(z, x) = \mathbb{E}_x[z^T]$ (l'indice x dans l'espérance étant utilisé pour indiquer que la distance d'échantillonnage est x).

Multiplions (2.2) par z^t et sommons sur t :

$$\begin{aligned}
\sum_{t=1}^{\infty} z^t \psi_t(x) &= \sum_{t=1}^{\infty} z^t \left[\frac{1}{2N} \left(G_t(x) - \sum_{\tau=1}^{t-1} G_{t-\tau}(x)\psi_\tau(0) \right) \right] \\
&= \frac{1}{2N} \sum_{t=1}^{\infty} z^t \left[G_t(x) - \sum_{\tau=1}^{t-1} G_{t-\tau}(x)\psi_\tau(0) \right] \\
&= \frac{1}{2N} \sum_{t=1}^{\infty} z^t G_t(x) - \frac{1}{2N} \sum_{t=1}^{\infty} z^t \sum_{\tau=1}^{t-1} G_{t-\tau}(x)\psi_\tau(0)
\end{aligned}$$

$$\sum_{t=1}^{\infty} z^t \psi_t(x) = \frac{1}{2N} \sum_{t=1}^{\infty} z^t G_t(x) - \frac{1}{2N} \sum_{t=1}^{\infty} \sum_{\tau=1}^{t-1} [z^{t-\tau} G_{t-\tau}(x)] [z^{\tau} \psi_{\tau}(0)]$$

Soit \tilde{G} , la Z -transformée (transformée de Laplace discrète) de G ;

$$\tilde{G}(z, x) = \sum_{t=1}^{\infty} G_t(x) z^t. \quad (2.3)$$

De plus :

$$\phi(z, x) = \mathbb{E}_x[z^T] = \sum_{t=1}^{\infty} z^t \psi_t(x)$$

D'où :

$$\begin{aligned} \phi(z, x) &= \frac{1}{2N} \sum_{t=1}^{\infty} z^t G_t(x) - \frac{1}{2N} \sum_{t=1}^{\infty} \sum_{\tau=1}^{t-1} [z^{t-\tau} G_{t-\tau}(x)] [z^{\tau} \psi_{\tau}(0)] \\ &= \frac{\tilde{G}(z, x)}{2N} - \frac{\tilde{G}(z, x) \phi(z, 0)}{2N} \end{aligned}$$

Et nous obtenons :

$$\phi(z, x) = \frac{\tilde{G}(z, x)}{2N} (1 - \phi(z, 0))$$

Posons $x = 0$ pour trouver une expression pour $\phi(z, 0)$:

$$\begin{aligned} \phi(z, 0) &= \frac{\tilde{G}(z, 0)}{2N} - \frac{\tilde{G}(z, 0) \phi(z, 0)}{2N} \\ &= \frac{\tilde{G}(z, 0)}{2N} \left[1 / \left(1 + \frac{\tilde{G}(z, 0)}{2N} \right) \right] \\ &= \frac{\tilde{G}(z, 0)}{2N} \frac{2N}{2N + \tilde{G}(z, 0)} \\ &= \frac{\tilde{G}(z, 0)}{2N + \tilde{G}(z, 0)} \end{aligned}$$

En substituant, on a :

$$\begin{aligned} \phi(z, x) &= \frac{\tilde{G}(z, x)}{2N} \left(1 - \frac{\tilde{G}(z, 0)}{2N + \tilde{G}(z, 0)} \right) \\ &= \frac{\tilde{G}(z, x)}{2N} \frac{2N + \tilde{G}(z, 0) - \tilde{G}(z, 0)}{2N + \tilde{G}(z, 0)} \end{aligned}$$

Finalement :

$$\phi(z, x) = \frac{\tilde{G}(z, x)}{2N + \tilde{G}(z, 0)} \quad (2.4)$$

Cela prend une forme particulièrement simple si g_1 est un noyau gaussien discrétisé que nous pouvons ensuite approcher par un noyau de dispersion strictement gaussien. Sur un ensemble infini,

$$\frac{1}{2N}G_t(x) = \frac{1}{2\mathcal{N}t} \exp\left(-\frac{|x|^2}{4\sigma^2 t}\right), \quad (2.5)$$

où $\mathcal{N} = 4N\pi\sigma^2$ est la taille de voisinage (ce qui correspond à la dispersion des lignées individuelles au taux $\sigma^2/2$, le facteur supplémentaire deux résultant du fait que G_1 régit la séparation entre les deux lignées).

En effet, si g_1 est gaussien alors, nous connaissons G_t , pour $t \geq 1$. Par récurrence sur $t \geq 1$, nous pouvons montrer que :

$$G_t(x) = \frac{1}{4\pi\sigma^2 t} \exp\left(-\frac{|x|^2}{4\sigma^2 t}\right)$$

- La propriété est satisfaite pour $t = 1$:

$$\begin{aligned} G_1(x) &= \int \exp\left(-\frac{|t|^2}{2\sigma^2}\right) \exp\left(-\frac{|x-t|^2}{2\sigma^2}\right) \frac{dt}{(2\pi\sigma^2)^2} \\ &= \int \exp\left(-\frac{|t|^2 + |x|^2 + |t|^2 - 2x \cdot t}{2\sigma^2}\right) \frac{dt}{(2\pi\sigma^2)^2} \\ &= \int \exp\left(-\frac{(\sqrt{2}|t| - |x|/\sqrt{2})^2 + |x|^2}{2\sigma^2}\right) \frac{dt}{(2\pi\sigma^2)^2} \\ &= \frac{\exp(-|x|^2/(4\sigma^2))}{(2\pi\sigma^2)^2} \int \exp\left(-\frac{(\sqrt{2}|t| - |x|/\sqrt{2})^2}{2\sigma^2}\right) dt \end{aligned}$$

On fait le changement de variable $u = \sqrt{2}|t| - |x|/\sqrt{2}$; $du = 2dt$, d'où :

$$G_1(x) = \frac{\exp(-|x|^2/(4\sigma^2))}{(2\pi\sigma^2)^2} \int \exp\left(-\frac{u^2}{2\sigma^2}\right) \frac{du}{2}$$

Or,

$$\int \exp\left(-\frac{u^2}{2\sigma^2}\right) \frac{du}{2} = \frac{1}{2} 2\pi\sigma^2 = \pi\sigma^2$$

D'où

$$G_1(x) = \frac{1}{4\pi\sigma^2} \exp\left(-\frac{|x|^2}{4\sigma^2}\right)$$

- On suppose que la propriété est satisfaite pour $t - 1$, c'est-à-dire :

$$G_{t-1}(x) = \frac{1}{4\pi\sigma^2(t-1)} \exp\left(-\frac{|x|^2}{4\sigma^2(t-1)}\right)$$

- Montrons que la propriété est vraie pour t :

$$\begin{aligned}
G_t(x) &= \int G_{t-1}(u)G_1(x-u)du \\
&= \int \exp\left(-\frac{|u|^2}{4\sigma^2(t-1)}\right) \exp\left(-\frac{|x-u|^2}{4\sigma^2}\right) \frac{du}{(4\pi\sigma^2)^2(t-1)} \\
&= \frac{1}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{|u|^2 + (t-1)|x-u|^2}{4\sigma^2(t-1)}\right) du \\
&= \frac{1}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{|u|^2 + (t-1)|x|^2 + (t-1)|u|^2 - 2(t-1)x \cdot u}{4\sigma^2(t-1)}\right) du \\
G_t(x) &= \frac{1}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{(\sqrt{t}|u| + (t-1)|x|/\sqrt{t})^2 + |x|^2(t-1 - (t-1)^2/t)}{4\sigma^2(t-1)}\right) du \\
&= \frac{1}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{(\sqrt{t}|u| + (t-1)|x|/\sqrt{t})^2 + |x|^2(t-1)/t}{4\sigma^2(t-1)}\right) du \\
&= \frac{\exp(-|x|^2(t-1)/(4t\sigma^2(t-1)))}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{(\sqrt{t}|u| + (t-1)|x|/\sqrt{t})^2}{4\sigma^2(t-1)}\right) du
\end{aligned}$$

On fait le changement de variable $v = \sqrt{t}|u| + (t-1)|x|/\sqrt{t}$; $dv = tdu$, d'où :

$$G_t(x) = \frac{\exp(-|x|^2/(4\sigma^2t))}{(4\pi\sigma^2)^2(t-1)} \int \exp\left(-\frac{|v|^2}{4\sigma^2(t-1)}\right) \frac{dv}{t}$$

Or,

$$\int \exp\left(-\frac{v^2}{4\sigma^2(t-1)}\right) \frac{dv}{t} = \frac{4\pi\sigma^2(t-1)}{t}$$

D'où

$$G_t(x) = \frac{1}{4\pi\sigma^2t} \exp\left(-\frac{|x|^2}{4\sigma^2t}\right)$$

Nous écrivons maintenant $\eta(x)$ pour la probabilité que deux lignées de séparation x fusionnent à la génération précédente. Ainsi :

$$\int_{\mathbb{R}^2} \eta(x)dx = \frac{1}{2N}.$$

En effet, il s'agit de la probabilité que deux lignées fusionnent à la génération précédente, or la probabilité que deux gènes de séparation x aient un ancêtre commun à la génération précédente est donnée par $G_1(x)/(2N)$ et

$$\int_{\mathbb{R}^2} \frac{1}{2N}G_1(x)dx = \frac{1}{2N} \int_{\mathbb{R}^2} G_1(x)dx = \frac{1}{2N},$$

car $G_1(x)$ est la loi de la séparation de deux lignée et son intégrale sur \mathbb{R}^2 est égale à 1.

Nous pouvons alors écrire :

$$\mathcal{N} = \frac{2\pi\sigma^2}{\int_{\mathbb{R}^2} \eta(x)dx}, \quad (2.6)$$

car nous savons que $\mathcal{N} = 4N\pi\sigma^2$.

Grâce à l'équation (2.3), avec l'approximation continue (2.5) pour G_t , et en utilisant le développement en série entière de $\log(1-z)$, $z < 1$, nous avons donc :

$$\frac{1}{2N}\tilde{G}(z, 0) = \frac{1}{2N} \sum_{t=1}^{\infty} \frac{z^t}{t} = \frac{1}{\mathcal{N}} \log\left(\frac{1}{\sqrt{1-z}}\right)$$

et

$$\frac{1}{2N}\tilde{G}(z, x) = \frac{1}{\mathcal{N}} \sum_{t=1}^{\infty} \frac{z^t}{2t} \exp\left(-\frac{|x|^2}{4\sigma^2 t}\right)$$

A condition que $|x|\sqrt{1-z}/\sigma$ n'est pas trop petit, $|x|/\sigma > 2$, par exemple, et z suffisamment proche de 1, nous avons :

$$\frac{1}{2N}\tilde{G}(z, x) \simeq \frac{1}{\mathcal{N}} K_0\left(\frac{|x|}{\sigma}\sqrt{1-z}\right),$$

où K_0 est la fonction de Bessel modifiée de deuxième espèce de degré zéro.

En effet, dans le cas où z suffisamment proche de 1, on peut remplacer z^t par $\exp(-t(1-z))$, d'où :

$$\begin{aligned} \frac{1}{2N} \sum_{t=1}^{\infty} \frac{z^t}{t} \exp\left(-\frac{|x|^2}{4\sigma^2 t}\right) &= \frac{1}{2N} \sum_{t=1}^{\infty} \frac{1}{t} \exp(-t(1-z)) \exp\left(-\frac{|x|^2}{4\sigma^2 t}\right) \\ &\simeq \frac{1}{2N} \int_0^{\infty} \exp\left(-t(1-z) - \frac{|x|^2}{4\sigma^2 t}\right) \frac{dt}{t} \\ &\simeq \frac{1}{\mathcal{N}} K_0\left(\frac{|x|}{\sigma}\sqrt{1-z}\right) \end{aligned}$$

Cependant, lorsque $|x| \downarrow 0$, $\mathcal{N}\tilde{G}(z, x)/(2N)$ tend vers $\log(1/\sqrt{1-z})$ alors que $K_0(|x|\sqrt{1-z}/\sigma)$ diverge.

Nous avons maintenant les ingrédients pour déterminer la fonction génératrice des temps de coalescence :

$$\phi(z, 0) = \mathbb{E}_0[z^T] = \frac{1}{1 - \frac{2\mathcal{N}}{\log(\sqrt{1-z})}} \quad (2.7)$$

et pour $|x|$ suffisamment grand et z suffisamment proche de 1 (au moins),

$$\phi(z, x) = \mathbb{E}_x[z^T] = \frac{\tilde{G}(z, x)}{2N + \tilde{G}(z, 0)} \simeq \frac{K_0\left(\frac{|x|}{\sigma}\sqrt{1-z}\right)}{\mathcal{N} - \log(\sqrt{1-z})} \quad (2.8)$$

Essentiellement la même dérivation peut être appliquée à toute distribution de dispersion, y compris la marche aléatoire du plus proche voisin. Bien sûr, l'expression (2.8) ne peut pas s'appliquer pour des $|x|$ très petits car il y a le problème de divergence en $x = 0$ de la fonction de Bessel modifiée de deuxième espèce d'ordre 0, K_0 . La solution exacte pour ces très petites distances d'échantillonnage dépendra des caractéristiques du mécanisme de dispersion.

Remarque 2.2.1.3. *Pour une dispersion gaussienne, Malécot (1948) trouve l'expression exacte comme une intégrale par rapport à une fonction de Bessel. Durrett (2008) considère le cas où la migration des lignées ancestrales est régi par marche aléatoire du plus proche voisin. On peut établir une récurrence similaire à (2.1) pour des populations qui sont distribuées à travers un continuum spatial (Barton et al., 2002).*

Lorsqu'on travaille dans un continuum spatial, pour contourner la divergence de la fonction de Bessel dans (2.8), il est souvent pratique de déclarer être à l'échelle locale κ pour laquelle la fonction génératrice est à peu près constante et égale à $\tilde{\phi}(z, 0)$ (Barton et al., 2002). C'est ce que nous allons faire ici, en écrivant l'équation (2.8) de la manière suivante :

$$\phi(z, x) = \frac{1 - \tilde{\phi}(z, 0)}{\mathcal{N}} K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right)$$

Assimilons $\phi(z, \kappa)$ à $\tilde{\phi}(z, 0)$, et posons $x = \kappa$ pour trouver une expression de $\phi(z, \kappa)$:

$$\begin{aligned} \phi(z, \kappa) &\simeq \frac{1 - \phi(z, \kappa)}{\mathcal{N}} K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right) \\ &\simeq \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N}} / \left[1 + \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N}} \right] \\ &\simeq \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N}} / \frac{\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N}} \\ &\simeq \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)} \end{aligned}$$

Toujours en assimilant $\phi(z, \kappa)$ à $\tilde{\phi}(z, 0)$, et en utilisant que $K_0(y) \simeq -\log y$ quand $y \downarrow 0$, nous obtenons :

$$\begin{aligned} \phi(z, x) &\simeq \frac{1 - \phi(z, \kappa)}{\mathcal{N}} K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right) \\ &\simeq \frac{1}{\mathcal{N}} \left(1 - \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)} \right) K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right) \\ &\simeq \left(\frac{1}{\mathcal{N}} - \frac{K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N} (\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right))} \right) K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right) \\ &\simeq \frac{\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right) - K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N} (\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right))} K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right) \\ &\simeq \frac{1}{\mathcal{N} + K_0 \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)} K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right) \end{aligned}$$

En utilisant que $K_0(y) \simeq -\log y$ quand $y \downarrow 0$, nous avons finalement :

$$\phi(z, x) \simeq \frac{K_0 \left(\frac{|x|}{\sigma} \sqrt{1 - z} \right)}{\mathcal{N} - \log \left(\frac{\kappa}{\sigma} \sqrt{1 - z} \right)} \quad (2.9)$$

Il est plus courant d'écrire $z = e^{-2\mu}$ avec μ un taux de mutation (par individu, par génération) pour un modèle de mutation à infinité d'allèles. La quantité $\phi(z, x)$, nous donne alors la probabilité que deux allèles échantillonnés à une séparation x soient identiques. En substituant dans

(2.9), on obtient la version plus familière de la formule de Wright-Malécot :

$$\phi(e^{-2\mu}, x) = \mathbb{E}_x[e^{-2\mu T}] \simeq \frac{K_0(x/l_\mu)}{\mathcal{N} + \log(l_\mu/\kappa)}, \text{ pour } |x| > \kappa \quad (2.10)$$

où $l_\mu = \sigma/\sqrt{2\mu}$, et

$$\phi(e^{-2\mu}, 0) = \frac{\log(l_\mu/\kappa)}{\mathcal{N} + \log(l_\mu/\kappa)}$$

□

En vertu de la migration symétrique en deux dimensions spatiales, nous voyons une dichotomie du comportement dans la séparation des deux lignées ancestrales échantillonnées près l'une de l'autre :

- soit elles coalescent rapidement,
- soit elles s'éloignent et ainsi prennent un certain temps pour se retrouver, au cours duquel elles sont très susceptibles de subir des mutations.

Si nous nous concentrons sur des distances d'échantillonnage intermédiaires x telles que $x^2/\sigma^2 \ll \mu^{-1}$, alors ce sera précisément ces lignées, qui parviennent à s'échapper loin les unes des autres, qui se différencieront avant de coalescer (notons que $\mathcal{O}(x^2/\sigma^2)$ est le temps minimum requis pour que les lignées se rencontrent et coalescent). Notre objectif est d'exploiter cette dichotomie pour déduire (un sous-ensemble) les paramètres σ^2, \mathcal{N} et κ .

Le dernier ingrédient clé dépend également de la séparation des échelles de temps. Bien qu'il soit peu probable que l'ensemble de la population atteigne un équilibre global en $\mathcal{O}(\mu^{-1})$ unités de temps, un quasi-équilibre sera atteint sur des échelles spatiales et temporelles intermédiaires. Par conséquent, si nous considérons une vaste zone d'échantillonnage A dont le diamètre satisfait $\text{Diam}(A) \ll \sigma\sqrt{\mu^{-1}}$, nous pouvons trouver un temps t^* tel que

$$\text{Diam}(A)^2/\sigma^2 \ll t^* \ll \mu^{-1}$$

et tel que dans les t^* dernières générations, la distribution allélique dans A est restée à peu près constante et des nouvelles mutations ne sont pas apparues. En particulier, les allèles présents dans l'échantillon sont tous distingués les uns des autres par mutation à un temps beaucoup plus lointain que t^* dans la généalogie.

En combinant les observations faites ci-dessus, nous pouvons fixer un certain temps de référence, $t^* \ll \mu^{-1}$ (il y a quelques centaines de générations, par exemple), et considérer une zone d'échantillonnage A telle que $\text{Diam}(A) \ll \sigma\sqrt{t^*}$. Une quantité importante de la coalescence locale « rapide » des lignées ancestrales aura lieu en temps t^* , mais ces lignées, qui réussissent à échapper à une autre, ne se rejoindront pas. Ces lignées vont accumuler des mutations avant la coalescence, c'est ce qui rend leurs longues excursions loin des autres observables.

Pour l'inférence, nous allons proposer une analyse basée sur un analogue de la statistique de Wright, F_{ST} . Nous nous appuyons fortement sur le quasi-équilibre des fréquences alléliques atteint sur la zone intermédiaire A et nous utiliserons également fortement la formule de Wright-Malécot.

2.2.2 Une statistique F appropriée

La théorie de base de l'inférence basée sur la fréquence allélique des populations spatialement structurées a été énoncée par Wright (1943b) et Malécot (1948) au milieu du siècle dernier, mais les tests statistiques couramment utilisés sont basés sur le modèle d'îles, qui ne permet pas d'intégrer la caractéristique de dispersion limitée de la plupart des espèces.

Le principe du modèle d'îles de Wright est le suivant : nous considérons D dèmes et N individus par île.

- Dans chaque île, chaque individu donne naissance à taux $1/2$ et son descendant remplace quelqu'un choisi au hasard dans cette même île.
- Chaque individu produit un migrant à taux m , qui remplace quelqu'un choisi au hasard sur une des $D - 1$ autres îles.

La statistique la plus couramment utilisée est la statistique F_{ST} de Wright qui exprime le déficit en hétérozygotes. C'est un indice de déviation par rapport aux proportions de Hardy-Weinberg. Elle est définie par :

$$F_{ST} = \frac{H_T - \overline{H_S}}{H_T}$$

où H_T désigne l'hétérozygotie attendue par individu, en supposant la population globale à l'équilibre de Hardy-Weinberg, H_S désigne l'hétérozygotie attendue par individu pour une sous-population supposée à l'équilibre de Hardy-Weinberg et $\overline{H_S}$ est la moyenne des H_S sur toutes les sous-populations.

La statistique F_{ST} est un indicateur de la cohésion de l'ensemble des sous-populations considérées. Elle donne l'effet de la subdivision entre les sous-populations et la population, et donne la réduction d'hétérozygotie dans les sous-populations, lié aux différences de fréquences alléliques. C'est un indice de diversification des populations. Si toutes les populations sont à l'équilibre et possèdent les mêmes fréquences alléliques, $F_{ST} = 0$, sinon si les fréquences alléliques moyennes sont différentes (dérivé génétique), il est positif. Cela reflète l'action de la dérive (diversification) et des migrations (homogénéisation).

Ici, elle correspond simplement à un écart standard des fréquences alléliques, elle ne contient aucune information sur la localisation spatiale.

Wright (1943b) a inclus la structure spatiale en calculant la variance sur différentes échelles, ce qui donne une hiérarchie des statistiques F . Cependant, toutes les études effectuées ont été limitées par la difficulté des calculs et l'absence de marqueurs génétiques.

Maintenant, nous avons une bonne puissance de calcul et des données génétiques, pourtant la plupart des analyses sont descriptives (auto-corrélation spatiale, par exemple) ou utilisent des méthodes de Monte Carlo pour ajuster les données, sans l'aide de la théorie explicite développée par Wright et Malécot.

De nombreuses statistiques F ont été proposées, mais la plupart d'entre-elles n'ont pas été justifiées pour l'application à une population bi-dimensionnelle.

Ici, la statistique proposée repose sur une variante de la statistique F_{ST} de Wright. Nous l'avons d'abord exprimée en termes de probabilité de coalescence de deux lignées ancestrales avant notre temps de référence t^* . Des prédictions théoriques pour ces probabilités peuvent ainsi être obtenues (numériquement) par inversion de la fonction génératrice du temps de coalescence (déterminée par la formule de Wright-Malécot classique).

Malgré tout, pour une grande taille de voisinage et sous notre hypothèse de séparation des échelles de temps, il y a une approximation analytique simple de la statistique F , que nous obtiendrons en (2.13) :

Théorème 2.2.2.1. *Si $r > 0$ désigne la distance entre deux points d'échantillonnage donnés dans A , alors la fonction $F(r)$ (probabilité que deux individus pris au hasard dans deux patches de séparation r soient identiques) vérifie :*

$$F(r) \simeq \frac{\log(\bar{r}/r)}{\mathcal{N} + \log(\bar{r}/\kappa)}$$

où \bar{r} est la moyenne géométrique de la séparation de toutes les paires possibles de certains individus échantillonnés à partir de A et \mathcal{N} , κ sont les paramètres de la population apparaissant dans la formule de Wright-Malécot.

Remarque 2.2.2.2. *Notons que σ^2 n'apparaît pas dans $F(r)$.*

Nous montrerons ensuite le résultat suivant :

Théorème 2.2.2.3. *La statistique $F(r)$ peut être estimée par*

$$\mathcal{F}(x, y) = \frac{1}{a-1} \sum_{i=1}^a \frac{(p_i(x) - \bar{p}_i)(p_i(y) - \bar{p}_i)}{\bar{p}_i}$$

où a est le nombre total d'allèles vus dans la région A , $p_i(x)$ est la fréquence locale de l'allèle i au site x et \bar{p}_i est la fréquence de l'allèle i dans la région A .

2.2.3 Inférence des paramètres à partir de F_{ST}

Sur la base des résultats décrits ci-dessus : comment devons-nous estimer la taille de voisinage à partir des fréquences alléliques ?

La méthode la plus simple est d'échantillonner les individus à partir de n endroits distincts et de régresser F_{ST} par paire, en fonction du logarithme de la distance (Rousset, 1997, 2003).

Rousset et Leblois (2007) ont mis en place une méthode de Monte Carlo qui utilise des simulations de coalescence pour tenir compte de la distribution complète des fréquences alléliques. Cependant, cela exige beaucoup de calculs et il est peu probable d'obtenir des informations en plus de celles concernant \mathcal{N} .

En principe, il y a des informations dans les relations d'ordre supérieur (par exemple, des taux de fusions multiples, qui peuvent être importants dans les populations à deux dimensions avec \mathcal{N} petit). Cependant, nous sommes concernés par des modèles locaux, pour lesquels la

mutation actuelle est négligeable : tout ce que nous pouvons observer, ce sont les fréquences des allèles qui se distinguent par des mutations qui se sont produites loin dans le passé.

A moins que F_{ST} soit anormalement élevée, la distribution des fréquences alléliques sera proche de la normale multivariée. Ainsi, nous ne pouvons pas aller au-delà des relations par paire, qui ne dépendent essentiellement que de \mathcal{N} .

Dans la partie 3, nous montrerons que, en supposant la normalité, nous pouvons expliquer les corrélations spatiales, en ajustant une matrice de covariance. D'ailleurs, si nous échantillons des régions assez petites de telle sorte que la similarité décroît avec $\log(1/r)$, cette matrice a une forme unique, qui dépend des endroits connus des gènes échantillonnés, ainsi que d'une unique « échelle locale », κ .

Remarque 2.2.3.1. *Cette approche est directe et plus transparente qu'une approche basée sur la simulation. Cependant, elle n'a guère été appliquée, même en théorie.*

2.3 Inférence

Dans cette partie, nous présentons une approche pour déduire la taille de voisinage à partir des fréquences alléliques dans la population, par le biais d'une statistique F appropriée. Nous présenterons les résultats dans le cadre du modèle stepping stone. Cependant, ils resteront valides chaque fois que la séparation des deux lignées ancestrales peut être décrite par une marche aléatoire symétrique en deux dimensions (ou son analogue continu : mouvement brownien) avec le paramètre de variance $2\sigma^2$ (correspondant au fait que chaque marche individuelle a pour variance σ^2), jusqu'à ce qu'elles fusionnent au moyen d'un mécanisme de coalescence local qui peut se résumer par un seul paramètre \mathcal{N} (le taux de coalescence « local »).

Dans la suite, nous utilisons la dualité entre les fréquences alléliques (à un seul locus) et les généalogies pour exprimer les corrélations entre les diversités génétiques locales en termes de probabilité de coalescence de deux lignées ancestrales en temps t^* . Nous définissons ensuite une statistique F que nous utilisons comme base pour une approche du maximum de vraisemblance afin d'estimer \mathcal{N} .

2.3.1 Identité d'état et mesures de F_{ST}

Supposons que nous sommes en mesure d'échantillonner parfaitement la diversité de la population locale à un moment donné (ou plutôt dans une petite zone autour d'un point donné).

Dans une analyse de données réelles, la variance d'échantillonnage peut être facilement incorporée dans la méthode d'inférence que nous dériverons plus tard.

Nous supposons que nous observons exactement a allèles distincts dans notre échantillon. Nous écrivons $p_i(x)$ pour la fréquence de l'allèle i dans le dème x (au moment de l'échantillonnage). Conformément à la statistique F classique, nous allons comparer les corrélations dans les fréquences alléliques entre les individus échantillonnés à distance spécifique, à celles observées sur une plus grande surface. Rappelons que cette surface, que nous désignons par A , est supposée être telle que son diamètre vérifie $\text{Diam}(A) \ll \sigma\sqrt{t^*}$, et que la mutation est supposée

être assez lente, de sorte que la diversité génétique dans A est entièrement due à des mutations qui ont eu lieu il y a plus de t^* générations dans le passé.

Ecrivons $\mathbb{P}_{x,y}$ pour la distribution des lignées ancestrales de deux individus échantillonnés à partir d'endroits $x, y \in A$ et T pour le temps aléatoire au cours duquel elles coalescent.

Conformément à notre hypothèse d'équilibre local dans A , nous supposons que si les deux lignées n'ont pas coalescé en temps t^* , alors la probabilité que les individus soient de types différents est indépendante de leur séparation initiale et est donnée par $H(t^*)$, l'hétérozygotie au temps t^* .

Si $|x - y| = r$, nous pouvons écrire :

$$1 - H_r := \mathbb{E} \left[\sum_{i=1}^a p_i(x)p_i(y) \right]$$

Soit :

$$1 - H_r = \mathbb{P}_{x,y}[T \leq t^*] + \mathbb{P}_{x,y}[T > t^*](1 - H(t^*))$$

A partir de cette expression, nous notons,

$$H_r = \mathbb{P}_r[T > t^*]H(t^*)$$

Nous écrirons également H_A pour l'hétérozygotie dans A au temps présent :

$$H_A = 1 - \frac{1}{|A|^2} \int_A \int_A \sum_{i=1}^a p_i(x)p_i(y) dx dy$$

En écrivant \mathbb{P}_A pour la distribution de deux lignées ancestrales de deux individus échantillonnés indépendamment et uniformément au hasard dans A , nous avons également :

$$H_A = \mathbb{P}_A(T > t^*)H(t^*)$$

Notre analogue à la statistique F_{ST} de Wright est alors définie par :

$$F(r) = \frac{H_A - H_r}{H_A} = \frac{\mathbb{P}_A[T > t^*] - \mathbb{P}_r[T > t^*]}{\mathbb{P}_A[T > t^*]} \quad (2.11)$$

Remarque 2.3.1.1. Notons que c'est indépendant de $H(t^*)$, d'où l'intérêt de renormaliser.

Si la taille de voisinage est grande, puisque nous échantillonnons sur une région A telle que $\text{Diam}(A) \ll \sigma\sqrt{t^*}$, $F(r)$ prend une forme particulièrement simple.

Pour commencer, nous déduisons de (2.11) :

$$1 - F(0) = \frac{\mathbb{P}_0[T > t^*]}{\mathbb{P}_A[T > t^*]} \quad (2.12)$$

Théorème 2.3.1.2.

$$F(r) \simeq \frac{1 - F(0)}{\mathcal{N}} \log \left(\frac{c}{r} \right),$$

où c est égal à la moyenne géométrique des séparations de toutes les paires possibles d'individus échantillonnés dans A .

Remarque 2.3.1.3. *Ce sera notre base pour l'inférence.*

Démonstration. Nous allons utiliser des estimations sur les marches aléatoires continues pour comparer $\mathbb{P}_0[T > t^*]$ à $\mathbb{P}_r[T > t^*]$. Cette comparaison va nous permettre de justifier l'approximation de $F(r)$ donnée par le théorème.

Nous écrivons $L_{t^*}(r)$ pour le temps total qu'une marche aléatoire de variance $2\sigma^2$, partant à une distance r de l'origine, passe à l'origine jusqu'en temps t^* .

Nous rappelons que nous écrivons T pour le temps aléatoire auquel deux lignées coalescent. Alors,

$$\mathbb{P}_r[T > t^*] = \mathbb{E}(\exp(-L_{t^*}(r)/N))$$

Ecrivons $T_0(r)$ pour le premier temps auquel une marche aléatoire partant de r , passe par 0.

En utilisant la propriété de Markov forte, nous obtenons :

$$\begin{aligned} \mathbb{P}_r[T > t^*] - \mathbb{P}_0[T > t^*] &= \mathbb{E}\left[e^{-L_{t^* - T_0(r)}(0)/N}\right] - \mathbb{E}\left[e^{-L_{t^*}(0)/N}\right] \\ &= \mathbb{E}\left[e^{-L_{t^* - T_0(r)}(0)/N} (1 - e^{-(L_{t^*}(0) - L_{t^* - T_0(r)}(0))/N})\right] \\ &= \mathbb{E}\left[e^{-L_{t^* - T_0(r)}(0)/N}\right] \mathbb{E}\left[1 - e^{-(L_{t^*}(0) - L_{t^* - T_0(r)}(0))/N}\right] \end{aligned}$$

Car le temps local à l'origine avant le temps $t^* - T_0(r)$ est indépendant du temps local en zéro dans l'intervalle $[t^* - T_0(r), t^*]$. Et finalement :

$$\mathbb{P}_r[T > t^*] - \mathbb{P}_0[T > t^*] \simeq \frac{1}{N} \mathbb{E}\left[e^{-L_{t^* - T_0(r)}(0)/N}\right] \mathbb{E}\left[L_{t^*}(0) - L_{t^* - T_0(r)}(0)\right]$$

Quand t^* tend vers ∞ , nous avons :

$$\mathbb{E}\left[e^{-L_{t^* - T_0(r)}(0)/N}\right] \simeq \mathbb{E}\left[e^{-L_{t^*}(0)/N}\right] = \mathbb{P}_0[T > t^*]$$

et

$$\mathbb{E}\left[L_{t^*}(0) - L_{t^* - T_0(r)}(0)\right] \rightarrow a(r)$$

où $a(r)$ est le « noyau potentiel » de la marche aléatoire. Cette fonction prend la forme

$$a(r) = 2C \log(r) + \mathcal{O}(1),$$

où $C = 1/(4\pi\sigma^2)$ (il y a un facteur supplémentaire deux dans le dénominateur ici, car nous nous intéressons à la marche aléatoire régissant la séparation de deux lignées, et non le mouvement d'une seule lignée).

Pour voir d'où ça vient, puisque nous supposons que $t^* \gg r^2$, au moment d'ordre r^2 lorsque la marche aléatoire partant de la séparation r atteint zéro pour la première fois, le marcheur aléatoire partant de l'origine a passé

$$\sim \sum_{n < r^2} \frac{1}{n} \sim 2C \log r$$

unités de temps en zéro, où $C = 1/(4\pi\sigma^2)$.

D'après cela, la marche partant de r se comporte comme une marche partant de zéro.

D'après les équations (2.11) et (2.12), nous avons :

$$\begin{aligned} F(r) &= \frac{\mathbb{P}_0[T > t^*]}{\mathbb{P}_A[T > t^*]} \frac{\mathbb{P}_A[T > t^*] - \mathbb{P}_r[T > t^*]}{\mathbb{P}_0[T > t^*]} \\ &= (1 - F(0)) \frac{\mathbb{P}_A[T > t^*] - \mathbb{P}_r[T > t^*]}{\mathbb{P}_0[T > t^*]} \\ &= (1 - F(0)) \left(\frac{\mathbb{P}_A[T > t^*] - \mathbb{P}_0[T > t^*]}{\mathbb{P}_0[T > t^*]} - \frac{\mathbb{P}_r[T > t^*] - \mathbb{P}_0[T > t^*]}{\mathbb{P}_0[T > t^*]} \right) \end{aligned}$$

En combinant ce qui précède, nous obtenons :

$$F(r) \simeq (1 - F(0)) \frac{2C}{N} (\log(c) - \log(r))$$

où c est égal à la moyenne géométrique des séparations de toutes les paires possibles d'individus échantillonnés dans A . Et finalement :

$$F(r) \simeq \frac{1 - F(0)}{N} \log\left(\frac{c}{r}\right)$$

□

Pour estimer $F(r)$ à partir des données, nous allons utiliser $\mathcal{F}(x, y)$, dont nous rappelons la définition.

Définition 2.3.1.4.

$$\mathcal{F}(x, y) = \frac{1}{a-1} \sum_{i=1}^a \frac{(p_i(x) - \bar{p}_i)(p_i(y) - \bar{p}_i)}{\bar{p}_i}.$$

où a est le nombre total d'allèles vus dans la région A , $p_i(x)$ est la fréquence locale de l'allèle i au site x et \bar{p}_i est la fréquence attendue de l'allèle i dans toute la région A .

Remarque 2.3.1.5. Notons que \bar{p}_i désigne également la probabilité qu'un individu échantillonné au hasard dans A soit de type i .

Théorème 2.3.1.6. En écrivant $r = |x - y|$,

$$\mathbb{E}[\mathcal{F}(x, y)] = 1 - \frac{\mathbb{P}_r[T > t^*]}{\mathbb{P}_A[T > t^*]} = F(r).$$

Démonstration. Pour montrer ce résultat, nous allons utiliser le fait que les fréquences alléliques dans A sont au quasi-équilibre depuis le temps t^* avant le présent.

Pour commencer, écrivons \bar{p}_{ij} pour la probabilité que deux individus échantillonnés au temps t^* soient de types i et j .

Notons que nous avons $\bar{p}_i \bar{p}_j = \mathbb{P}_A[T > t^*] \bar{p}_{ij}$, pour $i \neq j$.

Nous pouvons écrire :

$$\mathcal{F}(x, y) = \frac{1}{a-1} \sum_{i=1}^a \frac{p_i(x)p_i(y)}{\bar{p}_i} - \frac{1}{a-1}$$

En prenant l'espérance, nous obtenons :

$$\begin{aligned} \mathbb{E}[\mathcal{F}(x, y)] &= \frac{1}{a-1} \sum_{i=1}^a \frac{p_i(x)p_i(y)}{\bar{p}_i} - \frac{1}{a-1} \\ &= \frac{1}{a-1} \sum_{i=1}^a \left(\mathbb{P}_r[T < t^*] \frac{\bar{p}_i}{\bar{p}_i} + \mathbb{P}_r[T > t^*] \frac{\bar{p}_{ii}}{\bar{p}_i} \right) - \frac{1}{a-1} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\mathcal{F}(x, y)] &= 1 - \frac{a}{a-1} \mathbb{P}_r[T > t^*] + \frac{1}{a-1} \mathbb{P}_r[T > t^*] \sum_{i=1}^a \frac{\bar{p}_{ii}}{\bar{p}_i} \\ &= 1 - \frac{a}{a-1} \mathbb{P}_r[T > t^*] + \frac{1}{a-1} \mathbb{P}_r[T > t^*] \sum_{i=1}^a \frac{1}{\bar{p}_i} \left(\bar{p}_i - \sum_{j=1, j \neq i}^a \bar{p}_{ij} \right) \\ &= 1 - \frac{a}{a-1} \mathbb{P}_r[T > t^*] + \frac{1}{a-1} \mathbb{P}_r[T > t^*] \left(a - \sum_{i=1}^a \sum_{j=1, j \neq i}^a \frac{1}{\bar{p}_i} \bar{p}_{ij} \right) \\ &= 1 - \frac{1}{a-1} \mathbb{P}_r[T > t^*] \sum_{i=1}^a \sum_{j=1, j \neq i}^a \frac{1}{\bar{p}_i} \frac{\bar{p}_i \bar{p}_j}{\mathbb{P}_A[T > t^*]} \\ &= 1 - \frac{\mathbb{P}_r[T > t^*]}{\mathbb{P}_A[T > t^*]} \frac{1}{a-1} \sum_{i=1}^a \sum_{j=1, j \neq i}^a \bar{p}_j \\ &= 1 - \frac{\mathbb{P}_r[T > t^*]}{\mathbb{P}_A[T > t^*]} \\ &= F(r) \end{aligned}$$

□

Remarque 2.3.1.7. $\mathcal{F}(x, y)$ fournit ainsi une statistique sur laquelle effectuer la méthode du maximum de vraisemblance.

2.3.2 Une approche du maximum de vraisemblance pour l'inférence

En général, la méthode de la section précédente nous mène à :

$$F(r) \simeq \frac{\log(\bar{r}/r)}{\mathcal{N} + \log(\bar{r}/\kappa)} \quad (2.13)$$

où l'échelle « locale » κ est choisie telle que

$$F(0) \simeq \frac{\log(\bar{r}/\kappa)}{\mathcal{N} + \log(\bar{r}/\kappa)}$$

où \bar{r} est la moyenne géométrique de la séparation d'individus échantillonnés uniformément au hasard dans A , à l'exception que nous remplaçons les séparations inférieures à κ par κ , pour éviter que ça explose.

Ceci suggère immédiatement que nous pouvons obtenir une estimation pour \mathcal{N} et κ par une régression de $F(r)$ sur $\log r$.

Premièrement, si nous échantillonnons n emplacements distincts dans A (où chaque paire d'emplacements distincts est séparée par une distance au moins κ), pour calculer F , nous devons d'abord estimer la valeur correspondante à \bar{r} , il s'agit de la moyenne géométrique (modifiée) de la distance entre toutes les paires d'individus échantillonnées. Cela correspond à $n(n-1)$ paires de séparation non nulle et n paires de séparation nulle ; il s'agit des termes diagonaux, qui sont remplacés par κ . Ainsi, nous obtenons :

$$n^2 \log \bar{r} = n(n-1) \log(\bar{r}^*) + n \log \kappa$$

où \bar{r}^* correspond à la moyenne géométrique basée sur les $n(n-1)$ paires de séparation non nulle.

Proposition 2.3.2.1.

$$\log \left(\frac{\bar{r}}{\kappa} \right) = \left(1 - \frac{1}{n} \right) \log \left(\frac{\bar{r}^*}{\kappa} \right) = \frac{F(0)}{1 - F(0)} \mathcal{N}$$

et la régression de $F(r)$ sur $\log(r)$ a pour pente :

$$m = \frac{1}{\mathcal{N} + \log(\bar{r}/\kappa)} = \frac{1 - F(0)}{\mathcal{N}}$$

Démonstration. Comme

$$n^2 \log \bar{r} = n(n-1) \log(\bar{r}^*) + n \log \kappa,$$

alors

$$n^2 \log \left(\frac{\bar{r}}{\kappa} \right) = n(n-1) \log(\bar{r}^*) + (n - n^2) \log \kappa.$$

D'où

$$\log \left(\frac{\bar{r}}{\kappa} \right) = \frac{n-1}{n} \log(\bar{r}^*) + \frac{1-n}{n} \log \kappa,$$

donc

$$\log \left(\frac{\bar{r}}{\kappa} \right) = \frac{n-1}{n} \log \left(\frac{\bar{r}^*}{\kappa} \right) = \left(1 - \frac{1}{n} \right) \log \left(\frac{\bar{r}^*}{\kappa} \right).$$

De plus d'après (2.13),

$$F(0) = \frac{\log(\bar{r}/\kappa)}{\mathcal{N} + \log(\bar{r}/\kappa)}.$$

Donc

$$\log \left(\frac{\bar{r}}{\kappa} \right) = F(0) \mathcal{N} + F(0) \log \left(\frac{\bar{r}}{\kappa} \right),$$

autrement dit,

$$\log \left(\frac{\bar{r}}{\kappa} \right) = \frac{F(0)}{1 - F(0)} \mathcal{N}.$$

De (2.13), nous déduisons la pente de la régression de $F(r)$ sur $\log(r)$:

$$m = \frac{1}{\mathcal{N} + \log(\bar{r}/\kappa)},$$

et en substituant l'expression obtenue pour $\log(\bar{r}/\kappa)$, nous obtenons :

$$m = \frac{1 - F(0)}{\mathcal{N}}.$$

□

Théorème 2.3.2.2.

$$\mathcal{N} \sim \frac{1 - F(0)}{m}, \quad \kappa \sim \bar{r}^* \exp\left(-\frac{F(0)}{m} \frac{n}{n-1}\right).$$

Démonstration. L'équivalent pour \mathcal{N} découle immédiatement de l'expression de m dans la proposition précédente. En ce qui concerne κ , nous avons (par la proposition précédente) :

$$\left(1 - \frac{1}{n}\right) (\log(\bar{r}^*) - \log(\kappa)) = \frac{F(0)}{1 - F(0)} \mathcal{N},$$

donc

$$\frac{n-1}{n} \log(\kappa) = \frac{n-1}{n} \log(\bar{r}^*) - \frac{F(0)}{1 - F(0)} \mathcal{N},$$

d'où

$$\log(\kappa) = \log(\bar{r}^*) - \frac{n}{n-1} \frac{F(0)}{1 - F(0)} \mathcal{N},$$

et comme $m = (1 - F(0))/\mathcal{N}$,

$$\log(\kappa) = \log(\bar{r}^*) - \frac{nF(0)}{(n-1)m}.$$

Finalement

$$\kappa \sim \bar{r}^* \exp\left(-\frac{F(0)}{m} \frac{n}{n-1}\right).$$

□

Remarque 2.3.2.3. *Puisque la relation est juste logarithmique, il peut y avoir peu de puissance dans une méthode basée sur la régression. Nous nous attendons à obtenir une meilleure estimation en utilisant la méthode du maximum de vraisemblance basé sur $F(r)$ pour estimer le paramètre \mathcal{N} .*

Nous faisons l'approximation que les fluctuations dans les fréquences alléliques sont petites, de sorte que nous pouvons les approcher par une distribution gaussienne multivariée, dont nous rappelons la définition.

Définition 2.3.2.4. *Une distribution Gaussienne multivariée, en dimension d , de moyenne m et de matrice de covariance Σ est définie par :*

$$f_{\mathcal{N}(\mu, \Sigma)}(x) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x - m)^T \Sigma^{-1} (x - m)\right)$$

Remarque 2.3.2.5. Dans le cas isotropique, $\Sigma = \sigma^2 Id$.

Nous écrivons F_{ST} pour la matrice de covariance normalisée résultante.

Supposons que nous échantillonnons à partir d'un ensemble donné de n emplacements X_1, \dots, X_n de distribution gaussienne $\mathcal{N}_{(0, F^*)}$ avec F^* , matrice des covariances attendues, inconnue. Nous désignerons par \tilde{F} la matrice des covariances observées (empiriques) :

$$\tilde{F} = \frac{1}{n} \sum_{i=1}^n X_i X_i^T.$$

Alors la fonction de vraisemblance est donnée par :

$$L = \prod_{i=1}^n \frac{1}{\sqrt{(2\pi)^d \det(F^*)}} \exp\left(-\frac{1}{2} X_i^T F^{*-1} X_i\right)$$

En supprimant les termes ne dépendant pas de F^* , nous obtenons, à une constante près, que la fonction de log-vraisemblance est donnée par :

$$\begin{aligned} \log(L) &= \log\left(\left(\frac{1}{\sqrt{\det(F^*)}}\right)^n\right) + \log \exp\left(-\frac{1}{2} \sum_{i=1}^n X_i^T F^{*-1} X_i\right) \\ &= -\frac{n}{2} \log(\det(F^*)) - \frac{1}{2} \sum_{i=1}^n X_i^T F^{*-1} X_i \\ &= -\frac{n}{2} \log(\det(F^*)) - \frac{n}{2} \langle F^{*-1}, \tilde{F} \rangle_F \\ &\propto -\frac{1}{2} \left(\log(\det(F^*)) + \langle F^{*-1}, \tilde{F} \rangle_F \right) \end{aligned}$$

Finalement, la fonction de log-vraisemblance est de la forme :

$$\log(L) = -\frac{1}{2} \left(\log(\det(F^*)) + \sum_{j,k} \tilde{F}_{j,k} F_{j,k}^{*-1} \right)$$

Nous approchons F_{ST} , comme dans (2.13), en remplaçant r par κ sur la diagonale.

Comme nous avons estimé la fréquence d'allèle moyenne à partir des données, la covariance des déviations de la moyenne est

$$F_{j,k}^* = F_{STj,k} - F_{STj,*} - F_{ST*,k} + F_{ST**,}$$

où $*$ représente une moyenne sur un index.

Cette matrice est singulière, ayant une valeur propre nulle, et donc $\det(F^*)$ est calculé comme le produit de $n - 1$ valeurs propres strictement positives.

Pour un ensemble d'emplacements donné, F^* dépend uniquement de κ et l'estimateur du maximum de vraisemblance pour \mathcal{N} peut être trouvé explicitement :

$$\hat{\mathcal{N}} = \frac{n-1}{\sum_{j,k} \tilde{F}_{j,k} F_{j,k}^{*-1}} - \log(\bar{r}/\kappa).$$

2.3.3 Résultats basés sur les fréquences alléliques

La figure 2.1 montre la covariance normalisée des fréquences alléliques comme une fonction de la distance pour un modèle stepping stone sur \mathbb{Z}^2 avec la migration du plus proche voisin. Après 200 générations, pour des espacements de dèmes inférieurs à 7, c'est très proche de l'approximation logarithmique (2.13), comme le montre la figure 2.2.

La figure 2.3 montre le résultat de l'utilisation de la régression de $F(r)$ sur $\log r$ pour estimer \mathcal{N} , pour les simulations du modèle stepping stone. Enfin, la figure 2.4 montre la surface de vraisemblance pour les paramètres \mathcal{N} et κ obtenue par l'application du schéma présenté au paragraphe précédent, pour la population simulée en figure 2.1.

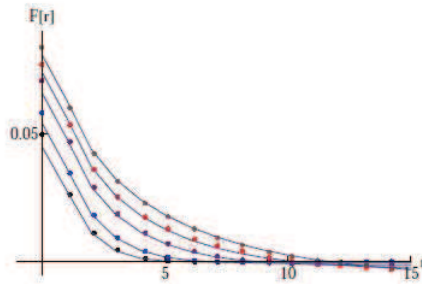


FIG. 2.1 – [Barton and al., 2013] Covariance normalisée des fréquences alléliques, F , en fonction de la distance pour un modèle stepping stone. Une grille de dèmes 40×40 , chacun avec $2N = 20$ individus haploïdes, simulée pour 200 générations ; il y avait migration entre chacun des quatre plus proches voisins, avec une fraction de migrants vers un voisin donné égale à $m/2 = 0.125$. Il y avait trois allèles, avec pour fréquences initiales $\{0.1, 0.4, 0.5\}$, et aucune mutation. Les points indiquent la moyenne de 10 répétitions indépendantes aux temps 10, 20, 50, 100, 200 (de bas en haut) ; les lignes rejoignent les prédictions théoriques pour ce modèle spatial discret, ce qui permet d'estimer la population moyenne à partir des valeurs réalisées. L'accord est proche, à l'exception d'une petite sous-estimation de $F(0)$. Cette simulation pourrait représenter l'échantillonnage de dix loci indépendants d'une population de 40×40 dèmes.

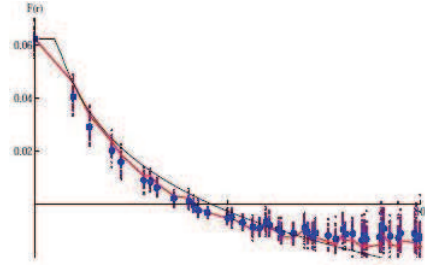


FIG. 2.2 – [Barton and al., 2013] Covariance en fonction de la distance pour un sous-échantillon de dèmes 10×10 , tiré de dix réalisations de la population 40×40 de la figure 2.1 ; ce qui peut être considéré comme un échantillonnage de dix loci indépendants, puisque les fréquences alléliques à chaque locus évoluent indépendamment. Chaque point est la moyenne sur dix réalisations ; différents points sont différents endroits de la grille 10×10 . La ligne rouge est la prédiction théorique, et la ligne noire, la simple prédiction en supposant que $F(r) = (1 - F(0))/\mathcal{N} \log(c/r)$, $F(0)$ étant estimé à partir de l'échantillon. Ceci est indépendant de la constante c . Cela peut être réécrit sous la forme $\log(\bar{r}/r)/(\mathcal{N} + \log(\bar{r}/\kappa))$ où $\kappa = 0.57$ est estimé à partir de $\bar{r} \exp(-\mathcal{N}F(0)/(1 - F(0)))$. La prédiction théorique correspond bien : c'est dentelé car $F(\{1, 1\})$ est légèrement inférieur à $F(\{0, 2\})$ même si $\{1, 1\}$ est plus proche de l'origine que $\{0, 2\}$. La courbe noire est la prédiction logarithmique naïve, ajustée à $F(0) = 0,062$ observé. Cela décroît un peu moins vite que prévu : nous obtenons $F(0) = 0,106$ à partir de la régression de F sur $\log r$, ce qui devrait avoir comme pente $(1 - F(0))/\mathcal{N}$.

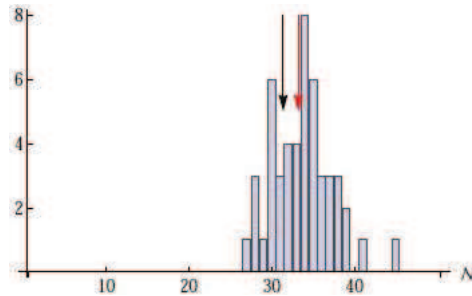


FIG. 2.3 – [Barton and al., 2013] Estimations de \mathcal{N} , chacune basée sur dix réalisations indépendantes (représentant dix loci indépendants). Ces estimations sont établies à partir de la pente de la régression de F_{ST} sur $\log r$. La répartition montre 49 répétitions de patches 10×10 , prélevés sur la même population 40×40 , mais à partir de $\{\{1, 1\}, \{6, 1\}, \dots, \{1, 6\}, \dots, \{31, 31\}\}$, ce qui représente la variation des estimations qui seraient obtenues par échantillonnage à partir d'endroits différents au sein d'une distribution stationnaire ; puisque les répétitions sont corrélées, cela sous-estime la variation entre les réalisations indépendantes. La flèche noire montre le vrai $\mathcal{N} = 4\pi \times 10 \times 0,25 = 31,4$ et la flèche rouge la moyenne sur les répétitions qui est une légère sur-estimation, 33.2. Ce biais est encore plus petit que la différence moyenne entre les estimations de \mathcal{N} obtenue en considérant différents groupes de patches 10×10 .

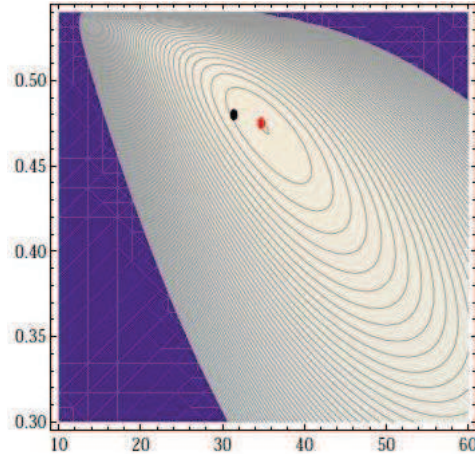


FIG. 2.4 – [Barton and al., 2013] Surface de vraisemblance basée sur dix loci échantillonnés à partir d’un patch 10×10 dans la population 40×40 de la figure 2.1 ; la log-vraisemblance est tracée en fonction de \mathcal{N} (en abscisse) et κ (en ordonnée). L’estimateur du maximum de vraisemblance (point rouge) est $\mathcal{N} = 34.75$, $\kappa = 0,475$; les vrais $\mathcal{N} = 31.4$, $\kappa = 0,48$ (point noir), avec log-vraisemblance plus faible que 1.5. Les contours sont espacés de deux unités de log-vraisemblance, de telle sorte que le cercle intérieur indique les limites d’appui pour chaque paramètre.

2.4 Conclusion

Nous nous sommes concentrés sur des populations réparties sur deux dimensions. L’argument central étant que les différentes approches sont appropriées à différentes échelles. Nous ne pouvons pas espérer déduire tous les détails de l’histoire de la population locale, mais nous ne pouvons pas supposer que les populations aient atteint l’équilibre sur de grandes échelles spatiales et temporelles. Cependant, en se concentrant sur des échantillons prélevés sur des patches convenables, nous pouvons faire des estimations robustes en supposant un quasi-équilibre local.

Les inférences fondées sur les fluctuations des fréquences alléliques dépendent de la distribution des moments où les gènes échantillonnés à une certaine distance r les uns des autres ont partagé un ancêtre commun. Si nous échantillonnons sur des patches locaux, alors nous pouvons seulement estimer la taille de voisinage, \mathcal{N} : l’information sur le taux de flux de gènes, σ^2 , est perdue. La relation logarithmique avec la distance doit cesser à une échelle suffisamment petite, $\sim \kappa$, puisque la fonction de similarité a une limite supérieure, $F(0) := F(\kappa)$.

Les résultats dépendent de la diminution de similarité avec $\log r$, ce qui est propre à l’isolement par la distance en deux dimensions. La théorie de l’isolement par la distance, originaire de Wright et Malécot, donne un cadre solide pour la compréhension de la répartition spatiale en deux dimensions. Elle fournit un modèle neutre clair grâce auquel nous pouvons détecter les effets de la sélection à des loci spécifiques. Les idées esquissées ici laissent ouvertes de nombreuses questions pour l’avenir.

Chapitre 3

Un nouveau modèle neutre de β -diversité

Dans ce chapitre, nous allons présenter un nouveau modèle neutre de β -diversité, basé sur les fréquences des espèces, inspiré des deux articles précédemment étudiés.

Pour étudier la β -diversité, nous proposons un nouveau modèle, toujours sous l'hypothèse de neutralité, le sort d'un arbre est indépendant de son espèce et la dispersion des graines est limitée à partir des parents. Nous définissons un indice de dissimilarité qui permettra d'obtenir des estimations concernant la fonction de similarité $F(r)$. Nous pouvons obtenir son espérance, mais pour avoir une expression de sa variance, nous devons introduire une nouvelle fonction de similarité, dite « par paires », $F^{(4)}$, qui désigne la probabilité que quatre arbres pris au hasard forment deux couples d'arbres conspécifiques. Enfin, nous déduisons des estimateurs pour $F(0)$ et $F(r)$, et nous regardons quelles informations nous pouvons réussir à obtenir en ce qui concerne les paramètres σ et ν .

3.1 Modèle

On considère A et B deux parcelles d'arbres disjointes séparées l'une de l'autre par une distance r . La distribution spatiale des arbres est un processus ponctuel de Poisson d'intensité ρ . Le nombre de points qui tombent dans A a une distribution de Poisson de paramètre $\rho|A|$, où $|A|$ désigne l'aire de la parcelle A et les nombres de points tombant dans des sous-ensembles disjoints sont indépendants.

Nous notons $p_i(A)$ pour la probabilité qu'un arbre de la parcelle A soit d'espèce i , où $i \in \mathbb{N}$. Nous avons :

$$p_i(A) = \frac{1}{N_A} \sum_{x \sim A} \mathbb{1}_{\{\text{arbre } x \text{ d'espèce } i\}},$$

où N_A désigne le nombre de points dans A .

Cette expression contient une variable aléatoire, N_A , qui suit une distribution de Poisson de paramètre $\rho|A|$. Pour faciliter les calculs, nous allons donc remplacer N_A par son espérance

dans l'expression de $p_i(A)$ et finalement nous avons :

$$p_i(A) = \frac{1}{\rho|A|} \sum_{x \sim A} \mathbb{1}_{\{\text{arbre } x \text{ d'espèce } i\}}$$

Remarque 3.1.0.1. $p_i(A)$ peut valoir 0.

Définition 3.1.0.2. On définit l'indice de dissimilarité :

$$\mathcal{F}(A, B) = \sum_i (p_i(A) - p_i(B))^2 \quad (3.1)$$

Cet indice va nous permettre d'estimer la similarité en composition d'espèces des deux parcelles A et B séparées par une distance donnée.

Nous supposons que l'indice de dissimilarité $\mathcal{F}(A, B)$ ne dépend que de la distance qui sépare les parcelles A et B et que les graines se dispersent à une distance limitée à partir de leurs parents, et que les arbres sont distribués de façon homogène sur une surface, avec une densité ρ par unité d'aire.

D'après l'article de Chave et Leigh (2002), nous avons une expression de la probabilité que deux arbres, pris au hasard dans chacune des parcelles appartiennent à la même espèce, en fonction de la distance r entre les parcelles : $F(r)$.

Nous allons nous intéresser à la loi de notre variable aléatoire $\mathcal{F}(A, B)$, calculer son espérance et sa variance. Pour cela nous commençons par nous intéresser à fonction de similarité par paires, la probabilité que quatre arbres, pris au hasard dans les parcelles A et B , forment deux paires d'arbres conspécifiques, en fonction de la distance r entre les parcelles : $F^{(4)}$.

3.2 Fonction de similarité par paires

Dans cette partie, nous allons nous intéresser à la fonction $F^{(4)}(a, b, c, d, t)$, où l'indice 4 désigne le nombre de parcelles considérées pour calculer la probabilité que les arbres en a et b soient conspécifiques et qu'en même temps, ceux en c et d soient conspécifiques.

La quantité $F^{(4)}(a, b, c, d, t)$ désigne la probabilité que deux arbres a et c choisis au hasard au temps t dans les voisinages respectifs dS_a (incluant le site $a = (a_1, a_2)$) et dS_c (incluant le site $c = (c_1, c_2)$), soient respectivement de la même espèce que deux arbres b et d choisis au hasard dans les voisinages respectifs dS_b (incluant le site $b = (b_1, b_2)$) et dS_d (incluant le site $d = (d_1, d_2)$).

A chaque pas de temps $\tau (= dt)$, un arbre meurt et il est immédiatement remplacé par le descendant d'un de ses voisins.

Comment relier $F^{(4)}(a, b, c, d, t + \tau)$ à $F^{(4)}(a, b, c, d, t)$?

Pour répondre à cette question nous faisons les hypothèses suivantes :

- (i) la distribution spatiale des arbres suit un processus ponctuel de Poisson d'intensité ρ ,
- (ii) chaque arbre meurt à taux 1, indépendamment de son espèce,
- (iii) chaque nouvel arbre a une probabilité ν d'être une espèce entièrement nouvelle,
- (iv) la fonction $b \mapsto P(a, b)$ désigne le « noyau de dispersion » des graines d'un arbre au site a .

Proposition 3.2.0.3. *Nous avons la relation suivante :*

$$F^{(4)}(a, b, c, d, t + \tau) = (1 - 4\tau)F^{(4)}(a, b, c, d, t) + 4\tau(1 - \nu)Q^{(4)}(a, b, c, d, t) \quad (3.2)$$

où

$$\begin{aligned} Q^{(4)}(a, b, c, d, t) &= \int_u F^{(4)}(u, b, c, d, t)P(u, a)dS_u + \frac{P(b, a)}{\rho} \left[F(c, d, t) - F^{(4)}(b, b, c, d, t) \right] \\ &+ \int_u F^{(4)}(a, u, c, d, t)P(u, b)dS_u + \frac{P(a, b)}{\rho} \left[F(c, d, t) - F^{(4)}(a, a, c, d, t) \right] \\ &+ \int_u F^{(4)}(a, b, u, d, t)P(u, b)dS_u + \frac{P(d, c)}{\rho} \left[F(a, b, t) - F^{(4)}(a, b, d, d, t) \right] \\ &+ \int_u F^{(4)}(a, b, c, u, t)P(u, d)dS_u + \frac{P(c, d)}{\rho} \left[F(a, b, t) - F^{(4)}(a, b, c, c, t) \right] \end{aligned}$$

où F désigne la fonction de similarité de Chave et Leigh (2002).

Remarque 3.2.0.4. *Pour obtenir ce résultat, le raisonnement est le même que celui conduisant à la proposition (1.2.0.3).*

Comme $P(a, b) = P(b, a)$ et $P(c, d) = P(d, c)$, nous pouvons écrire

$$\begin{aligned} Q^{(4)}(a, b, c, d, t) &= \int_u F^{(4)}(u, b, c, d, t)P(u, a)dS_u + \int_u F^{(4)}(a, u, c, d, t)P(u, b)dS_u \\ &+ \int_u F^{(4)}(a, b, u, d, t)P(u, b)dS_u + \int_u F^{(4)}(a, b, c, u, t)P(u, d)dS_u \\ &+ \frac{P(a, b)}{\rho} \left[2F(c, d, t) - F^{(4)}(b, b, c, d, t) - F^{(4)}(a, a, c, d, t) \right] \\ &+ \frac{P(c, d)}{\rho} \left[2F(a, b, t) - F^{(4)}(a, b, d, d, t) - F^{(4)}(a, b, c, c, t) \right] \end{aligned}$$

En passant à la limite continue dans l'équation (3.2), nous obtenons l'équation principale :

$$\frac{\partial F^{(4)}(a, b, c, d, t)}{\partial t} = -4F^{(4)}(a, b, c, d, t) + 4(1 - \nu)Q^{(4)}(a, b, c, d, t)$$

Si nous supposons que cette probabilité ne dépend que des positions relatives des différents arbres :

$$\begin{cases} P(a, b) &= P(a - b) \\ F^{(4)}(a, b, c, d, t) &= F^{(4)}(a - b, c - d, a - c, b - d, a - d, b - c, t) \end{cases}$$

En faisant les changements de variables : $i = a - b$, $j = c - d$, $k = a - c$, $l = b - d$, $m = a - d$, $n = b - c$, nous avons :

$$\begin{aligned}
\frac{\partial F^{(4)}(i, j, k, l, m, n, t)}{\partial t} &= -4F^{(4)}(i, j, k, l, m, n, t) + 4(1 - \nu) \times \\
&\left[\int_u F^{(4)}(u - i - a, j, u - k - a, l, u - m - a, n, t) P(u - a) dS_u \right. \\
&+ \int_u F^{(4)}(u - i - b, j, k, u - l - b, m, u - n - b, t) P(u - b) dS_u \\
&+ \int_u F^{(4)}(i, u - j - c, u - k - c, l, m, u - n - c, t) P(u - c) dS_u \\
&+ \int_u F^{(4)}(i, u - j - d, k, u - l - d, u - m - d, n, t) P(u - d) dS_u \\
&+ \frac{P(i)}{\rho} \left[2F(j, t) - F^{(4)}(0, j, n, l, l, n, t) - F^{(4)}(0, j, k, m, m, k, t) \right] \\
&\left. + \frac{P(j)}{\rho} \left[2F(i, t) - F^{(4)}(i, 0, m, l, m, l, t) - F^{(4)}(i, 0, k, n, k, n, t) \right] \right]
\end{aligned}$$

De nouveau par des changements de variables, nous obtenons :

$$\begin{aligned}
\frac{\partial F^{(4)}(i, j, k, l, m, n, t)}{\partial t} &= -4F^{(4)}(i, j, k, l, m, n, t) + 4(1 - \nu) \times \\
&\left[\int_v F^{(4)}(v - i, j, v - k, l, v - m, n, t) P(v) dS_v + \int_v F^{(4)}(v - i, j, k, v - l, m, v - n, t) P(v) dS_v \right. \\
&+ \int_v F^{(4)}(i, v - j, v - k, l, m, v - n, t) P(v) dS_v + \int_v F^{(4)}(i, v - j, k, v - l, v - m, n, t) P(v) dS_v \\
&+ \frac{P(i)}{\rho} \left[2F(j, t) - F^{(4)}(0, j, n, l, l, n, t) - F^{(4)}(0, j, k, m, m, k, t) \right] \\
&\left. + \frac{P(j)}{\rho} \left[2F(i, t) - F^{(4)}(i, 0, m, l, m, l, t) - F^{(4)}(i, 0, k, n, k, n, t) \right] \right]
\end{aligned}$$

A l'équilibre, nous avons :

$$\begin{aligned}
F^{(4)}(i, j, k, l, m, n) &= (1 - \nu) \times \\
&\left[\int_v F^{(4)}(v - i, j, v - k, l, v - m, n) P(v) dS_v + \int_v F^{(4)}(v - i, j, k, v - l, m, v - n) P(v) dS_v \right. \\
&+ \int_v F^{(4)}(i, v - j, v - k, l, m, v - n) P(v) dS_v + \int_v F^{(4)}(i, v - j, k, v - l, v - m, n) P(v) dS_v \\
&+ \frac{P(i)}{\rho} \left[2F(j) - F^{(4)}(0, j, n, l, l, n) - F^{(4)}(0, j, k, m, m, k) \right] \\
&\left. + \frac{P(j)}{\rho} \left[2F(i) - F^{(4)}(i, 0, m, l, m, l) - F^{(4)}(i, 0, k, n, k, n) \right] \right]
\end{aligned}$$

Remarque 3.2.0.5. *Il doit être possible de déterminer numériquement une expression pour $F^{(4)}$, mais nous ne l'avons pas fait ici...*

3.3 Etude de la loi de l'indice de dissimilarité

Nous commençons par calculer l'espérance de $\mathcal{F}(A, B)$ et ensuite nous allons calculer sa variance. Nous allons exprimer ces quantités à partir de nos fonctions de similarité F et $F^{(4)}$. Dans toute cette partie, nous allons considérer les arbres d'une même parcelle comme étant à distance κ (l'échelle « locale »).

Proposition 3.3.0.6. *L'espérance de $\mathcal{F}(A, B)$ est donnée par :*

$$\mathbb{E}[\mathcal{F}(A, B)] = 2[F(\kappa) - F(r)] + \frac{F(\kappa) |A| + |B|}{\rho |A||B|} \quad (3.3)$$

où κ désigne l'échelle « locale » et r la distance entre les parcelles A et B .

Remarque 3.3.0.7. *Le premier terme dépend de la distance qui sépare les parcelles A et B et le second terme dépend des aires de ces parcelles A et B ainsi que de l'intensité ρ selon laquelle les arbres sont répartis dans la forêt. Si la distance entre les parcelles augmente alors $F(r)$ diminue et donc $\mathbb{E}[\mathcal{F}(A, B)]$ augmente, donc la similarité diminue (plus les parcelles sont éloignées, moins les compositions en espèce des parcelles A et B sont similaires). Si ρ augmente ou si les aires des parcelles augmentent alors $\mathbb{E}[\mathcal{F}(A, B)]$ diminue, donc la similarité augmente (plus il y a d'arbres ou plus les parcelles sont grandes, plus les espèces sont représentées).*

Démonstration. Nous allons commencer par développer l'expression de $\mathcal{F}(A, B)$ pour faciliter le calcul de l'espérance.

$$\begin{aligned} \mathcal{F}(A, B) &= \sum_i (p_i(A) - p_i(B))^2 \\ &= \sum_i (p_i(A)^2 + p_i(B)^2 - 2p_i(A)p_i(B)) \\ &= \sum_i p_i(A)^2 + \sum_i p_i(B)^2 - 2 \sum_i p_i(A)p_i(B) \\ &= \sum_i \frac{1}{\rho^2 |A|^2} \left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 + \sum_i \frac{1}{\rho^2 |B|^2} \left(\sum_{y \sim B} \mathbb{1}_{\{y \text{ d'espèce } i\}} \right)^2 \\ &\quad - 2 \sum_i \frac{1}{\rho^2 |A||B|} \sum_{x \sim A} \sum_{y \sim B} \mathbb{1}_{\{x \text{ d'espèce } i\}} \mathbb{1}_{\{y \text{ d'espèce } i\}} \\ &= \frac{1}{\rho^2} \left[\frac{1}{|A|^2} \sum_i \left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 + \frac{1}{|B|^2} \sum_i \left(\sum_{y \sim B} \mathbb{1}_{\{y \text{ d'espèce } i\}} \right)^2 \right. \\ &\quad \left. - \frac{2}{|A||B|} \sum_i \sum_{x \sim A} \sum_{y \sim B} \mathbb{1}_{\{x \text{ et } y \text{ d'espèce } i\}} \right] \end{aligned}$$

Maintenant, nous passons à l'espérance :

$$\begin{aligned}
\mathbb{E}[\mathcal{F}(A, B)] &= \frac{1}{\rho^2} \mathbb{E} \left[\frac{1}{|A|^2} \sum_i \left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 + \frac{1}{|B|^2} \sum_i \left(\sum_{y \sim B} \mathbb{1}_{\{y \text{ d'espèce } i\}} \right)^2 \right. \\
&\quad \left. - \frac{2}{|A||B|} \sum_i \sum_{x \sim A} \sum_{y \sim B} \mathbb{1}_{\{x \text{ et } y \text{ d'espèce } i\}} \right] \\
\mathbb{E}[\mathcal{F}(A, B)] &= \frac{1}{\rho^2} \left[\frac{1}{|A|^2} \sum_i \mathbb{E} \left[\left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] + \frac{1}{|B|^2} \sum_i \mathbb{E} \left[\left(\sum_{y \sim B} \mathbb{1}_{\{y \text{ d'espèce } i\}} \right)^2 \right] \right. \\
&\quad \left. - \frac{2}{|A||B|} \sum_i \mathbb{E} \left[\sum_{x \sim A} \sum_{y \sim B} \mathbb{1}_{\{x \text{ et } y \text{ d'espèce } i\}} \right] \right] \\
&= \frac{1}{\rho^2} \left[\frac{1}{|A|^2} \sum_i \mathbb{E} \left[\left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] + \frac{1}{|B|^2} \sum_i \mathbb{E} \left[\left(\sum_{y \sim B} \mathbb{1}_{\{y \text{ d'espèce } i\}} \right)^2 \right] \right. \\
&\quad \left. - \frac{2}{|A||B|} \mathbb{E}[N_A] \mathbb{E}[N_B] \sum_i \mathbb{P}(x \in A \text{ et } y \in B \text{ d'espèce } i) \right]
\end{aligned}$$

D'une part, nous savons que

$$\sum_i \mathbb{P}(x \in A \text{ et } y \in B \text{ d'espèce } i) = F(r)$$

où r désigne la distance qui sépare les parcelles A et B .

D'où

$$\begin{aligned}
\frac{2}{|A||B|} \mathbb{E}[N_A] \mathbb{E}[N_B] \sum_i \mathbb{P}(x \in A \text{ et } y \in B \text{ d'espèce } i) &= \frac{2}{|A||B|} \rho^2 |A||B| F(r) \\
&= 2\rho^2 F(r)
\end{aligned}$$

D'autre part,

$$\begin{aligned}
\mathbb{E} \left[\left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] &= \mathbb{E} \left[\sum_{x_1 \sim A} \sum_{x_2 \sim A} \mathbb{1}_{\{x_1 \text{ et } x_2 \text{ d'espèce } i\}} \right] \\
&= \mathbb{E}[N_A N_A] \mathbb{P}(x_1 \in A \text{ et } x_2 \in A \text{ d'espèce } i)
\end{aligned}$$

Sommons sur i :

$$\begin{aligned}
\sum_i \mathbb{E} \left[\left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] &= \mathbb{E}[N_A^2] \sum_i \mathbb{P}(x_1 \in A \text{ et } x_2 \in A \text{ d'espèce } i) \\
&= \mathbb{E}[N_A^2] F(\kappa)
\end{aligned}$$

où κ très petit, désigne l'échelle « locale », la distance minimale pour laquelle l'expression de F est valide ($F(\kappa)$ correspond au $F(0)$ de Chave et Leigh (2002)).

De plus, $\mathbb{E}[N_a^2] = \mathcal{V}[N_A] + \mathbb{E}[N_A]^2 = \rho|A| + \rho^2|A|^2$, d'où

$$\begin{aligned} \frac{1}{|A|^2} \sum_i \mathbb{E} \left[\left(\sum_{x \sim A} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] &= \frac{\rho|A| + \rho^2|A|^2}{|A|^2} F(\kappa) \\ &= \left(\frac{\rho}{|A|} + \rho^2 \right) F(\kappa) \end{aligned}$$

De la même façon, nous obtenons :

$$\frac{1}{|B|^2} \sum_i \mathbb{E} \left[\left(\sum_{x \sim B} \mathbb{1}_{\{x \text{ d'espèce } i\}} \right)^2 \right] = \left(\frac{\rho}{|B|} + \rho^2 \right) F(\kappa)$$

Nous en déduisons :

$$\begin{aligned} \mathbb{E}[\mathcal{F}(A, B)] &= \frac{1}{\rho^2} \left[\left(\frac{\rho}{|A|} + \rho^2 \right) F(\kappa) + \left(\frac{\rho}{|B|} + \rho^2 \right) F(\kappa) - 2\rho^2 F(r) \right] \\ &= 2[F(\kappa) - F(r)] + \left(\frac{1}{\rho|A|} + \frac{1}{\rho|B|} \right) F(\kappa) \end{aligned}$$

D'où le résultat :

$$\mathbb{E}[\mathcal{F}(A, B)] = 2[F(\kappa) - F(r)] + \frac{F(\kappa)}{\rho} \frac{|A| + |B|}{|A||B|}$$

□

Proposition 3.3.0.8. *La variance de $\mathcal{F}(A, B)$ est donnée par :*

$$\begin{aligned} \mathcal{V}[\mathcal{F}(A, B)] &= 2CF^{(4)}(\kappa, \kappa, r, r, r) + 4CF^{(4)}(r, r, \kappa, \kappa, r) \\ &\quad + (C_3 + 3C_2 + 3C_1 + 2)F^{(4)}(\kappa, \kappa, \kappa, \kappa, \kappa) \\ &\quad - (4C_2 + 12C_1 + 8)F^{(4)}(r, \kappa, r, \kappa, r) \\ &\quad - 4F(r)^2 + 4F(r)F(\kappa)(2 + C_1) - F(\kappa)^2(4 + C_1^2) \end{aligned}$$

où κ désigne l'échelle « locale » et r la distance entre les parcelles A et B .

Démonstration. Nous savons que $\mathcal{V}[\mathcal{F}(A, B)] = \mathbb{E}[\mathcal{F}(A, B)^2] - \mathbb{E}[\mathcal{F}(A, B)]^2$.

Pour commencer, calculons $\mathbb{E}[\mathcal{F}(A, B)]^2$:

$$\mathbb{E}[\mathcal{F}(A, B)]^2 = \left[2[F(\kappa) - F(r)] + \frac{F(\kappa)}{\rho} \frac{|A| + |B|}{|A||B|} \right]^2$$

d'après (3.3).

$$\begin{aligned}
\mathbb{E}[\mathcal{F}(A, B)]^2 &= 4[F(\kappa) - F(r)]^2 + 4[F(\kappa) - F(r)] \frac{F(\kappa)}{\rho} \frac{|A| + |B|}{|A||B|} \\
&\quad + \frac{F(\kappa)^2}{\rho^2} \left(\frac{|A| + |B|}{|A||B|} \right)^2 \\
&= 4F(\kappa)^2 - 8F(\kappa)F(r) + 4F(r)^2 + \frac{4F(\kappa)^2}{\rho} \frac{|A| + |B|}{|A||B|} \\
&\quad - \frac{4F(\kappa)F(r)}{\rho} \frac{|A| + |B|}{|A||B|} + \frac{F(\kappa)^2}{\rho^2} \left(\frac{|A| + |B|}{|A||B|} \right)^2
\end{aligned}$$

Finalement, nous avons :

$$\mathbb{E}[\mathcal{F}(A, B)]^2 = 4F(r)^2 - 4F(r)F(\kappa) \left[2 + \frac{|A| + |B|}{\rho|A||B|} \right] + F(\kappa)^2 \left[4 + \left(\frac{|A| + |B|}{\rho|A||B|} \right)^2 \right] \quad (3.4)$$

A présent, calculons $\mathcal{F}(A, B)^2$:

$$\begin{aligned}
\mathcal{F}(A, B)^2 &= \left[\sum_i p_i(A)^2 + \sum_i p_i(B)^2 - 2 \sum_i p_i(A)p_i(B) \right]^2 \\
&= \left[\sum_i p_i(A)^2 + \sum_i p_i(B)^2 \right]^2 + \left[2 \sum_i p_i(A)p_i(B) \right]^2 \\
&\quad - 2 \left(\sum_i p_i(A)^2 + \sum_i p_i(B)^2 \right) \left(2 \sum_i p_i(A)p_i(B) \right) \\
&= \left(\sum_i p_i(A)^2 \right)^2 + \left(\sum_i p_i(B)^2 \right)^2 + 2 \sum_i p_i(A)^2 \sum_j p_j(B)^2 \\
&\quad - 4 \left(\sum_i p_i(A)^2 \sum_j p_j(A)p_j(B) + \sum_i p_i(B)^2 \sum_j p_j(A)p_j(B) \right) \\
&\quad + 4 \sum_i p_i(A)p_i(B) \sum_j p_j(A)p_j(B) \\
&= \sum_{i,j} [p_i(A)^2 p_j(A)^2 + p_i(B)^2 p_j(B)^2 + 2p_i(A)^2 p_j(B)^2 \\
&\quad - 4p_i(A)^2 p_j(A)p_j(B) - 4p_i(B)^2 p_j(A)p_j(B) + 4p_i(A)p_i(B)p_j(A)p_j(B)]
\end{aligned}$$

Pour simplifier l'écriture, nous allons écrire $\mathbb{1}_{x,i}$ pour $\mathbb{1}_{\{x \text{ d'espèce } i\}}$.

$$\begin{aligned}
\mathcal{F}(A, B)^2 &= \frac{1}{\rho^4} \sum_{i,j} \left[\frac{1}{|A|^4} \left(\sum_{x \sim A} \mathbb{1}_{x,i} \right)^2 \left(\sum_{x \sim A} \mathbb{1}_{x,j} \right)^2 + \frac{1}{|B|^4} \left(\sum_{y \sim B} \mathbb{1}_{y,i} \right)^2 \left(\sum_{y \sim B} \mathbb{1}_{y,j} \right)^2 \right. \\
&\quad + \frac{2}{|A|^2 |B|^2} \left(\sum_{x \sim A} \mathbb{1}_{x,i} \right)^2 \left(\sum_{y \sim B} \mathbb{1}_{y,j} \right)^2 \\
&\quad + \frac{4}{|A|^2 |B|^2} \left(\sum_{x \sim A} \mathbb{1}_{x,i} \right) \left(\sum_{x \sim A} \mathbb{1}_{x,j} \right) \left(\sum_{y \sim B} \mathbb{1}_{y,i} \right) \left(\sum_{y \sim B} \mathbb{1}_{y,j} \right) \\
&\quad - \frac{4}{|A|^3 |B|} \left(\sum_{x \sim A} \mathbb{1}_{x,i} \right)^2 \left(\sum_{x \sim A} \mathbb{1}_{x,j} \right) \left(\sum_{y \sim B} \mathbb{1}_{y,j} \right) \\
&\quad \left. - \frac{4}{|A| |B|^3} \left(\sum_{y \sim B} \mathbb{1}_{y,i} \right)^2 \left(\sum_{x \sim A} \mathbb{1}_{x,j} \right) \left(\sum_{y \sim B} \mathbb{1}_{y,j} \right) \right] \\
&= \frac{1}{\rho^4} \sum_{i,j} \left[\frac{1}{|A|^4} \sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{x_3 \sim A} \sum_{x_4 \sim A} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_3, x_4 \text{ d'espèce } j\}} \right. \\
&\quad + \frac{1}{|B|^4} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \sum_{y_3 \sim B} \sum_{y_4 \sim B} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } i\}} \mathbb{1}_{\{y_3, y_4 \text{ d'espèce } j\}} \\
&\quad + \frac{2}{|A|^2 |B|^2} \sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } j\}} \\
&\quad + \frac{4}{|A|^2 |B|^2} \sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \mathbb{1}_{\{x_1, y_1 \text{ d'espèce } i\}} \mathbb{1}_{\{x_2, y_2 \text{ d'espèce } j\}} \\
&\quad - \frac{4}{|A|^3 |B|} \sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{x_3 \sim A} \sum_{y_1 \sim B} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_3, y_1 \text{ d'espèce } j\}} \\
&\quad \left. - \frac{4}{|A| |B|^3} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \sum_{y_3 \sim B} \sum_{x_1 \sim A} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_1, y_3 \text{ d'espèce } j\}} \right]
\end{aligned}$$

Maintenant nous allons prendre l'espérance de $\mathcal{F}(A, B)^2$:

$$\begin{aligned}
\mathbb{E}[\mathcal{F}(A, B)^2] &= \frac{1}{\rho^4} \mathbb{E} \left[\sum_{i,j} \left[\frac{1}{|A|^4} \sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{x_3 \sim A} \sum_{x_4 \sim A} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_3, x_4 \text{ d'espèce } j\}} \right] \right. \\
&\quad + \frac{1}{|B|^4} \mathbb{E} \left[\sum_{y_1 \sim B} \sum_{y_2 \sim B} \sum_{y_3 \sim B} \sum_{y_4 \sim B} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } i\}} \mathbb{1}_{\{y_3, y_4 \text{ d'espèce } j\}} \right] \\
&\quad + \frac{2}{|A|^2 |B|^2} \mathbb{E} \left[\sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } j\}} \right] \\
&\quad + \frac{4}{|A|^2 |B|^2} \mathbb{E} \left[\sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{y_1 \sim B} \sum_{y_2 \sim B} \mathbb{1}_{\{x_1, y_1 \text{ d'espèce } i\}} \mathbb{1}_{\{x_2, y_2 \text{ d'espèce } j\}} \right] \\
&\quad - \frac{4}{|A|^3 |B|} \mathbb{E} \left[\sum_{x_1 \sim A} \sum_{x_2 \sim A} \sum_{x_3 \sim A} \sum_{y_1 \sim B} \mathbb{1}_{\{x_1, x_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_3, y_1 \text{ d'espèce } j\}} \right] \\
&\quad \left. - \frac{4}{|A| |B|^3} \mathbb{E} \left[\sum_{y_1 \sim B} \sum_{y_2 \sim B} \sum_{y_3 \sim B} \sum_{x_1 \sim A} \mathbb{1}_{\{y_1, y_2 \text{ d'espèce } i\}} \mathbb{1}_{\{x_1, y_3 \text{ d'espèce } j\}} \right] \right] \\
&= \frac{1}{\rho^4} \sum_{i,j} \left[\frac{1}{|A|^4} \mathbb{E}[N_A^4] \mathbb{P}(x_1, x_2 \in A \text{ d'espèce } i \text{ et } x_3, x_4 \in A \text{ d'espèce } j) \right. \\
&\quad + \frac{1}{|B|^4} \mathbb{E}[N_B^4] \mathbb{P}(y_1, y_2 \in B \text{ d'espèce } i \text{ et } y_3, y_4 \in B \text{ d'espèce } j) \\
&\quad + \frac{2}{|A|^2 |B|^2} \mathbb{E}[N_A^2] \mathbb{E}[N_B^2] \mathbb{P}(x_1, x_2 \in A \text{ d'espèce } i \text{ et } y_1, y_2 \in B \text{ d'espèce } j) \\
&\quad + \frac{4}{|A|^2 |B|^2} \mathbb{E}[N_A^2] \mathbb{E}[N_B^2] \mathbb{P}(x_1 \in A, y_1 \in B \text{ d'espèce } i \text{ et } x_2 \in A, y_2 \in B \text{ d'espèce } j) \\
&\quad - \frac{4}{|A|^3 |B|} \mathbb{E}[N_A^3] \mathbb{E}[N_B] \mathbb{P}(x_1, x_2 \in A \text{ d'espèce } i \text{ et } x_3 \in A, y_1 \in B \text{ d'espèce } j) \\
&\quad \left. - \frac{4}{|A| |B|^3} \mathbb{E}[N_A] \mathbb{E}[N_B^3] \mathbb{P}(y_1, y_2 \in B \text{ d'espèce } i \text{ et } x_1 \in A, y_3 \in B \text{ d'espèce } j) \right]
\end{aligned}$$

Pour simplifier cette expression nous avons besoin des moments de la loi de Poisson. Nous savons que

$$\mathbb{E}(X^n) = M_X^{(n)}(0) = \left. \frac{d^n M_X(t)}{dt^n} \right|_{t=0}$$

où $M_X(t) = \exp(\lambda(e^t - 1))$ pour une variable aléatoire X qui suit une distribution de Poisson de paramètre λ .

Nous avons :

$$\begin{aligned}
M_X^{(1)}(t) &= \lambda M_X(t) e^t \\
M_X^{(2)}(t) &= \lambda \left[M_X(t) + M_X^{(1)}(t) \right] e^t \\
M_X^{(3)}(t) &= \lambda \left[M_X(t) + 2M_X^{(1)}(t) + M_X^{(2)}(t) \right] e^t \\
M_X^{(4)}(t) &= \lambda \left[M_X(t) + 3M_X^{(1)}(t) + 3M_X^{(2)}(t) + M_X^{(3)}(t) \right] e^t
\end{aligned}$$

De plus, nous avons $M_X(0) = 1$; pour une variable aléatoire de Poisson de paramètre λ , et nous en déduisons donc :

$$\mathbb{E}(X) = \lambda, \mathbb{E}(X^2) = \lambda + \lambda^2, \mathbb{E}(X^3) = \lambda + 3\lambda^2 + \lambda^3 \text{ et } \mathbb{E}(X^4) = \lambda + 3\lambda^2 + 3\lambda^3 + \lambda^4.$$

Ici, les variables aléatoires N_A et N_B suivent des lois de Poisson respectivement de paramètres $\rho|A|$ et $\rho|B|$. Nous obtenons donc :

$$\begin{aligned}
\mathbb{E}[N_A^4] &= \rho|A| + 3\rho^2|A|^2 + 3\rho^3|A|^3 + \rho^4|A|^4 \\
\mathbb{E}[N_B^4] &= \rho|B| + 3\rho^2|B|^2 + 3\rho^3|B|^3 + \rho^4|B|^4 \\
\mathbb{E}[N_A^2]\mathbb{E}[N_B^2] &= (\rho|A| + \rho^2|A|^2)(\rho|B| + \rho^2|B|^2) \\
&= \rho^2|A||B| + \rho^3|A||B|(|A| + |B|) + \rho^4|A|^2|B|^2 \\
\mathbb{E}[N_A^3]\mathbb{E}[N_B] &= (\rho|A| + 3\rho^2|A|^2 + \rho^3|A|^3)\rho|B| \\
&= \rho^2|A||B| + 3\rho^3|A|^2|B| + \rho^4|A|^3|B| \\
\mathbb{E}[N_A]\mathbb{E}[N_B^3] &= \rho^2|A||B| + 3\rho^3|A||B|^2 + \rho^4|A||B|^3
\end{aligned}$$

Donc finalement, avec $(x_i)_{i=1,\dots,4}$ des arbres de la parcelle A et $(y_i)_{i=1,\dots,4}$ des arbres de la parcelle B :

$$\begin{aligned}
\mathbb{E}[\mathcal{F}(A, B)^2] &= \frac{1}{\rho^4} \sum_{i,j} \\
&\left[\frac{1}{|A|^4} (\rho|A| + 3\rho^2|A|^2 + 3\rho^3|A|^3 + \rho^4|A|^4) \mathbb{P}(x_1, x_2 \text{ d'esp. } i \text{ et } x_3, x_4 \text{ d'esp. } j) \right. \\
&+ \frac{1}{|B|^4} (\rho|B| + 3\rho^2|B|^2 + 3\rho^3|B|^3 + \rho^4|B|^4) \mathbb{P}(y_1, y_2 \text{ d'esp. } i \text{ et } y_3, y_4 \text{ d'esp. } j) \\
&+ \frac{2}{|A|^2|B|^2} (\rho^2|A||B| + \rho^3|A||B|(|A| + |B|) + \rho^4|A|^2|B|^2) \mathbb{P}(x_1, x_2 \text{ d'esp. } i \text{ et } y_1, y_2 \text{ d'esp. } j) \\
&+ \frac{4}{|A|^2|B|^2} (\rho^2|A||B| + \rho^3|A||B|(|A| + |B|) + \rho^4|A|^2|B|^2) \mathbb{P}(x_1, y_1 \text{ d'esp. } i \text{ et } x_2, y_2 \text{ d'esp. } j) \\
&- \frac{4}{|A|^3|B|} (\rho^2|A||B| + 3\rho^3|A|^2|B| + \rho^4|A|^3|B|) \mathbb{P}(x_1, x_2 \text{ d'esp. } i \text{ et } x_3, y_1 \text{ d'esp. } j) \\
&\left. - \frac{4}{|A||B|^3} (\rho^2|A||B| + 3\rho^3|A||B|^2 + \rho^4|A||B|^3) \mathbb{P}(y_1, y_2 \text{ d'esp. } i \text{ et } x_1, y_3 \text{ d'esp. } j) \right]
\end{aligned}$$

En sommant sur i et j les probabilités, nous pouvons exprimer $\mathbb{E}[\mathcal{F}(A, B)^2]$ fonction de $F^{(4)}$.

Nous obtenons :

$$\begin{aligned}\mathbb{E}[\mathcal{F}(A, B)^2] &= - \left(\frac{4(|A|^2 + |B|^2)}{\rho^2|A|^2|B|^2} + \frac{12(|A| + |B|)}{\rho|A||B|} + 8 \right) F^{(4)}(r, \kappa, r, \kappa, r, \kappa) \\ &+ \left(\frac{|A|^3 + |B|^3}{\rho^3|A|^3|B|^3} + \frac{3(|A|^2 + |B|^2)}{\rho^2|A|^2|B|^2} + \frac{3(|A| + |B|)}{\rho|A||B|} + 2 \right) F^{(4)}(\kappa, \kappa, \kappa, \kappa, \kappa, \kappa) \\ &+ \left(\frac{2}{\rho^2|A||B|} + \frac{2(|A| + |B|)}{\rho|A||B|} + 2 \right) F^{(4)}(\kappa, \kappa, r, r, r, r) \\ &+ \left(\frac{4}{\rho^2|A||B|} + \frac{4(|A| + |B|)}{\rho|A||B|} + 4 \right) F^{(4)}(r, r, \kappa, \kappa, r, r)\end{aligned}$$

Pour simplifier, nous posons :

$$C_i = \frac{|A|^i + |B|^i}{\rho^i|A|^i|B|^i} \text{ et } C = \frac{1}{\rho^2|A||B|} + C_1 + 1$$

Nous pouvons ainsi écrire :

$$\begin{aligned}\mathbb{E}[\mathcal{F}(A, B)^2] &= 2CF^{(4)}(\kappa, \kappa, r, r, r, r) + 4CF^{(4)}(r, r, \kappa, \kappa, r, r) \\ &+ (C_3 + 3C_2 + 3C_1 + 2) F^{(4)}(\kappa, \kappa, \kappa, \kappa, \kappa, \kappa) - (4C_2 + 12C_1 + 8)F^{(4)}(r, \kappa, r, \kappa, r, \kappa)\end{aligned}$$

Et l'équation (3.4) s'écrit :

$$\mathbb{E}[\mathcal{F}(A, B)]^2 = 4F(r)^2 - 4F(r)F(\kappa)(2 + C_1) + F(\kappa)^2(4 + C_1^2)$$

Nous en déduisons la variance de $\mathcal{F}(A, B)$:

$$\begin{aligned}\mathcal{V}[\mathcal{F}(A, B)] &= 2CF^{(4)}(\kappa, \kappa, r, r, r, r) + 4CF^{(4)}(r, r, \kappa, \kappa, r, r) \\ &+ (C_3 + 3C_2 + 3C_1 + 2) F^{(4)}(\kappa, \kappa, \kappa, \kappa, \kappa, \kappa) \\ &- (4C_2 + 12C_1 + 8)F^{(4)}(r, \kappa, r, \kappa, r, \kappa) \\ &- 4F(r)^2 + 4F(r)F(\kappa)(2 + C_1) - F(\kappa)^2(4 + C_1^2)\end{aligned}$$

□

3.4 Estimations

Comme κ désigne la distance minimale pour laquelle l'expression de F est valide, $F(\kappa)$ correspond en fait au $F(0)$ de Chave et Leigh (2002). Nous écrirons désormais $F(0)$ pour désigner $F(\kappa)$.

Rappelons la fonction de similarité en dimension deux dans le cas général :

$$F(r) = \frac{1}{c_0} K_0 \left(\frac{r\sqrt{2\nu}}{\sigma} \right)$$

où c_0 est une constante (Chave et Leigh, 2002).

Nous avons les équivalents suivants pour la fonction de Bessel modifiée de seconde espèce d'ordre 0 :

$$K_0(r) \sim \sqrt{\frac{\pi}{2r}} \exp(-r) \text{ si } r \text{ grand}$$

$$K_0(r) \sim -\log(r) \text{ si } r \text{ petit}$$

Nous pouvons alors écrire, pour r grand :

$$F(r) \simeq \frac{1}{c_0} \left(\frac{\pi\sigma}{2r\sqrt{2\nu}} \right)^{1/2} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right),$$

et pour r petit :

$$F(r) \simeq -\frac{1}{c_0} \log\left(\frac{r\sqrt{2\nu}}{\sigma}\right).$$

Nous voulons estimer cette fonction de similarité. Pour cela, nous avons :

$$\mathbb{E}_r[\mathcal{F}(A, B)] = 2[F(0) - F(r)] + C_1 F(0),$$

où C_1 désigne la constante $(|A| + |B|)/(\rho|A||B|)$

Si nous sommes en possession d'un jeu de données, nous sommes alors en mesure d'estimer $\mathcal{F}(A, B)$ et nous notons $\hat{\mathcal{F}}(A, B)$ cette estimation.

Pour commencer, nous pouvons alors estimer $F(0)$ par :

$$\hat{F}_0 = \frac{1}{C_1} \lim_{r \rightarrow 0} \mathbb{E}_r[\hat{\mathcal{F}}(A, B)]$$

et en posant $\hat{c} = (2 + C_1)\hat{F}_0$, nous pouvons alors estimer $F(r)$ par :

$$\hat{F}(r) = \frac{\hat{c} - \mathbb{E}_r[\hat{\mathcal{F}}(A, B)]}{2}$$

Maintenant, nous allons nous intéresser aux paramètres σ et ν , inconnus.

Dans le cas où r est petit, nous avons :

$$\mathbb{E}_r[\mathcal{F}(A, B)] = c + \frac{2}{c_0} \log\left(\frac{r\sqrt{2\nu}}{\sigma}\right)$$

où $c = (2 + C_1)F(0)$.

Ainsi :

$$\frac{c_0}{2} (\mathbb{E}_r[\mathcal{F}(A, B)] - c) = \log(r) + \frac{\log(2)}{2} + \frac{\log(\nu/\sigma^2)}{2}$$

Nous pouvons alors écrire :

$$\frac{\nu}{\sigma^2} = \exp[c_0 (E_r[\mathcal{F}(A, B)] - c) - 2 \log(r) - \log(2)]$$

Dans le cas où r est grand, nous avons :

$$\mathbb{E}_r[\mathcal{F}(A, B)] = c - \frac{2}{c_0} \left(\frac{\pi\sigma}{2r\sqrt{2\nu}} \right)^{1/2} \exp\left(-\frac{r\sqrt{2\nu}}{\sigma}\right)$$

Ainsi :

$$\left(\frac{\sqrt{2\pi}}{r} \right)^{1/2} c_0 (\mathbb{E}_r[\mathcal{F}(A, B)] - c) = \left(\frac{\sqrt{\nu}}{\sigma} \right)^{-1/2} \exp\left(-\sqrt{2}r \left(\frac{\sqrt{\nu}}{\sigma} \right)\right)$$

Nous pourrions alors obtenir des informations sur $\sqrt{\nu}/\sigma$, en résolvant l'équation :

$$x^{-1/2} \exp(-A_r x) - B_r = 0$$

où $A_r = \sqrt{2}r$ et $B_r = (\sqrt{2\pi}/r)^{1/2} c_0 (\mathbb{E}_r[\mathcal{F}(A, B)] - c)$.

Remarque 3.4.0.9. *Il semble difficile d'obtenir des informations séparées en ce qui concerne les paramètres σ et ν , mais si nous avons des données, nous serons quand même en mesure d'estimer la quantité ν/σ^2 .*

Pour comparer des données expérimentales de similarité à notre modèle mathématique censé prédire ces données, nous pouvons également utiliser la méthode des moindres carrés. Notre modèle théorique est la famille de fonctions $F(r; \sigma, \nu)$ indexées par les paramètres σ et ν inconnus. La méthode des moindres carrés va nous permettre de sélectionner parmi ces fonctions, celle qui reproduit le mieux les données expérimentales, il s'agit de l'ajustement par la méthode des moindres carrés.

Si nous disposons de N mesures, $(F_i)_{i=1, \dots, N}$, les paramètres σ et ν « optimaux » au sens de la méthode des moindres carrés sont ceux qui minimisent la quantité :

$$S(\sigma, \nu) = \sum_{i=1}^N (F_i - F(r_i; \sigma, \nu))^2 = \sum_{i=1}^N R_i(\sigma, \nu)$$

où les $R_i(\sigma, \nu)$ sont les résidus au modèle, i.e. les écarts entre les points de mesure F_i et le modèle $F(r; \sigma, \nu)$.

La quantité $S(\sigma, \nu)$ peut être considérée comme une mesure de la distance entre les données expérimentales et le modèle théorique qui prédit ces données. La méthode des moindres carrés demande que cette distance soit minimale.

Dans le cas où r est petit, nous avons :

$$\begin{aligned} S(\sigma, \nu) &= \sum_{i=1}^N (F_i - F(r_i; \sigma, \nu))^2 \\ &= \sum_{i=1}^N \left(F_i + \frac{1}{c_0} \log\left(\frac{r_i\sqrt{2\nu}}{\sigma}\right) \right)^2 \\ &= \sum_{i=1}^N \left(F_i^2 + \frac{F_i}{c_0} \log(2r_i^2) + \frac{F_i}{c_0} \log\left(\frac{\nu}{\sigma^2}\right) + \frac{1}{4c_0^2} \log^2(2r_i^2) + \frac{1}{4c_0^2} \log^2\left(\frac{\nu}{\sigma^2}\right) \right. \\ &\quad \left. + \frac{1}{2c_0^2} \log(2r_i^2) \log\left(\frac{\nu}{\sigma^2}\right) \right) \end{aligned}$$

Posons $\theta = \nu/\sigma^2$, nous avons donc :

$$S(\theta) = \sum_{i=1}^N \left(F_i^2 + \frac{F_i}{c_0} \log(2r_i^2) + \frac{F_i}{c_0} \log(\theta) + \frac{1}{4c_0^2} \log^2(2r_i^2) + \frac{1}{4c_0^2} \log^2(\theta) + \frac{1}{2c_0^2} \log(2r_i^2) \log(\theta) \right)$$

Dérivons par rapport à θ :

$$\frac{\partial S(\theta)}{\partial \theta} = \frac{1}{\theta} \sum_{i=1}^N \left(\frac{F_i}{c_0} + \frac{\log(2r_i^2)}{2c_0^2} + \frac{1}{2c_0^2} \log(\theta) \right)$$

Ainsi :

$$\frac{\partial S(\theta)}{\partial \theta} = 0 \Leftrightarrow \theta = \exp \left(-\frac{1}{N} \sum_{i=1}^N (2c_0 F_i + \log(2r_i^2)) \right)$$

Dans le cas où r est grand, nous avons :

$$\begin{aligned} S(\sigma, \nu) &= \sum_{i=1}^N (F_i - F(r_i; \sigma, \nu))^2 \\ &= \sum_{i=1}^N \left(F_i - \frac{1}{c_0} \left(\frac{\pi \sigma}{2r\sqrt{2\nu}} \right)^{1/2} \exp \left(-\frac{r\sqrt{2\nu}}{\sigma} \right) \right)^2 \\ &= \sum_{i=1}^N \left(F_i^2 + \frac{1}{c_0^2} \left(\frac{\pi}{2\sqrt{2}r} \right) \left(\frac{\sqrt{\nu}}{\sigma} \right)^{-1} \exp \left(-2\sqrt{2}r \frac{\sqrt{\nu}}{\sigma} \right) - \frac{F_i}{c_0} \left(\frac{\sqrt{2}\pi}{r} \right)^{1/2} \left(\frac{\sqrt{\nu}}{\sigma} \right)^{-1/2} \exp \left(-r\sqrt{2} \frac{\sqrt{\nu}}{\sigma} \right) \right) \end{aligned}$$

Posons $\psi = \sqrt{\nu}/\sigma$, nous avons donc :

$$\begin{aligned} S(\psi) &= \sum_{i=1}^N \left(F_i^2 + \frac{1}{c_0^2} \left(\frac{\pi}{2\sqrt{2}r} \right) \psi^{-1} \exp \left(-2\sqrt{2}r\psi \right) - \frac{F_i}{c_0} \left(\frac{\sqrt{2}\pi}{r} \right)^{1/2} \psi^{-1/2} \exp \left(-r\sqrt{2}\psi \right) \right) \end{aligned}$$

Dérivons par rapport à ψ :

$$\begin{aligned} \frac{\partial S(\psi)}{\partial \psi} &= \sum_{i=1}^N \left(-\frac{\pi}{c_0^2} \left(\frac{\psi^{-2}}{2\sqrt{2}r} + \psi^{-1} \right) \exp \left(-2\sqrt{2}r\psi \right) + \frac{F_i}{c_0} \left(\frac{\sqrt{2}\pi}{r} \right)^{1/2} \left(\frac{\psi^{-3/2}}{2} + r\sqrt{2}\psi^{-1/2} \right) \exp \left(-r\sqrt{2}\psi \right) \right) \end{aligned}$$

Nous pourrions alors obtenir des informations sur ψ , en résolvant $\partial S(\psi)/\partial \psi = 0$.

Remarque 3.4.0.10. *Il semble difficile d'obtenir des estimateurs pour ν et σ séparément...*

3.5 Simulations

Pour commencer, dans tous les cas, nous constatons que nos prédictions théoriques donnent bien que la similarité en composition d'espèces décroît avec l'augmentation de la distance.

Pour terminer, nous avons simulé l'évolution en composition d'espèces d'une population au cours des générations et évalué la similarité en composition d'espèces en fonction de la distance entre les deux parcelles considérées pour notre modèle (voir l'annexe).

Nous avons simulé une forêt de taille $N = 5000$ (processus ponctuel homogène), composée initialement de deux espèces différentes à la même fréquence. Nous avons fait évoluer cette population de la façon suivante : à chaque génération, un arbre au hasard meurt et il est remplacé par le descendant d'un de ses quatre plus proches voisins, c'est-à-dire prend l'espèce d'un de ces voisins ou une nouvelle espèce (avec probabilité $\nu = 10^{-4}$). Après 10^9 générations, nous obtenons des résultats satisfaisants pour notre modèle. En effet, la similarité en composition d'espèce (en utilisant notre indice de dissimilarité) diminue bien lorsque la distance entre les parcelles considérées augmente et il en est de même lorsque nous évaluons la similarité de deux arbres pris au hasard dans chacune des parcelles, en faisant la moyenne sur 10^3 répétitions. Nous obtenons les figures suivantes :

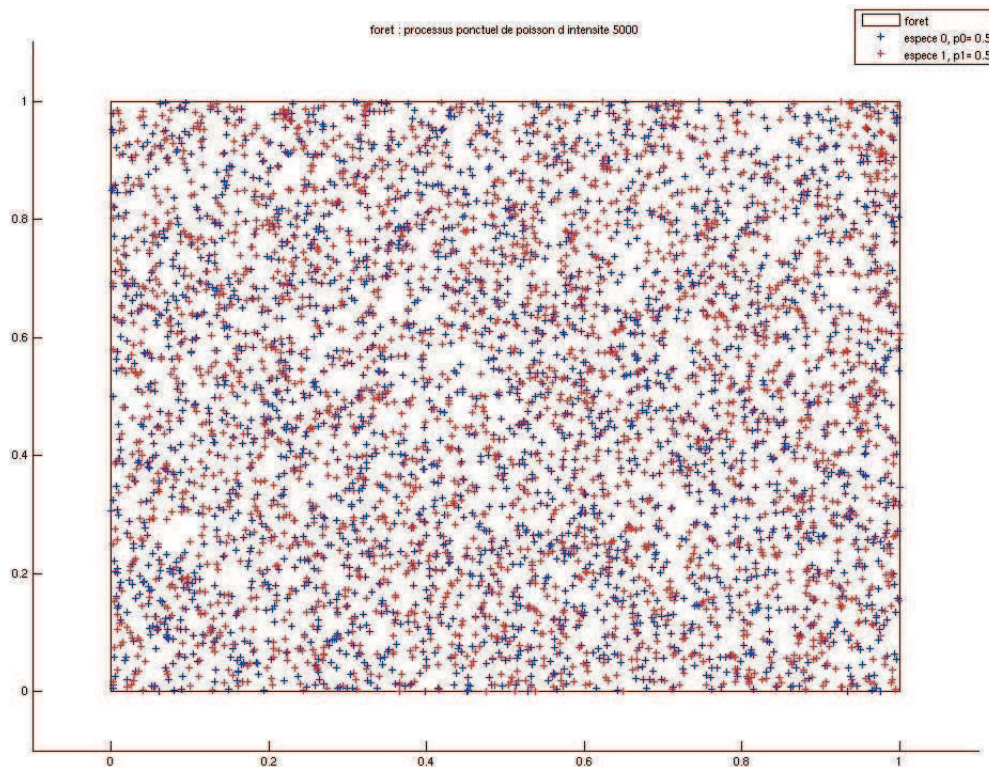


FIG. 3.1 – Forêt homogène simulée sur $[0, 1] \times [0, 1]$, de taille $N=5000$, avec deux espèces différentes à la même fréquence 0.5.

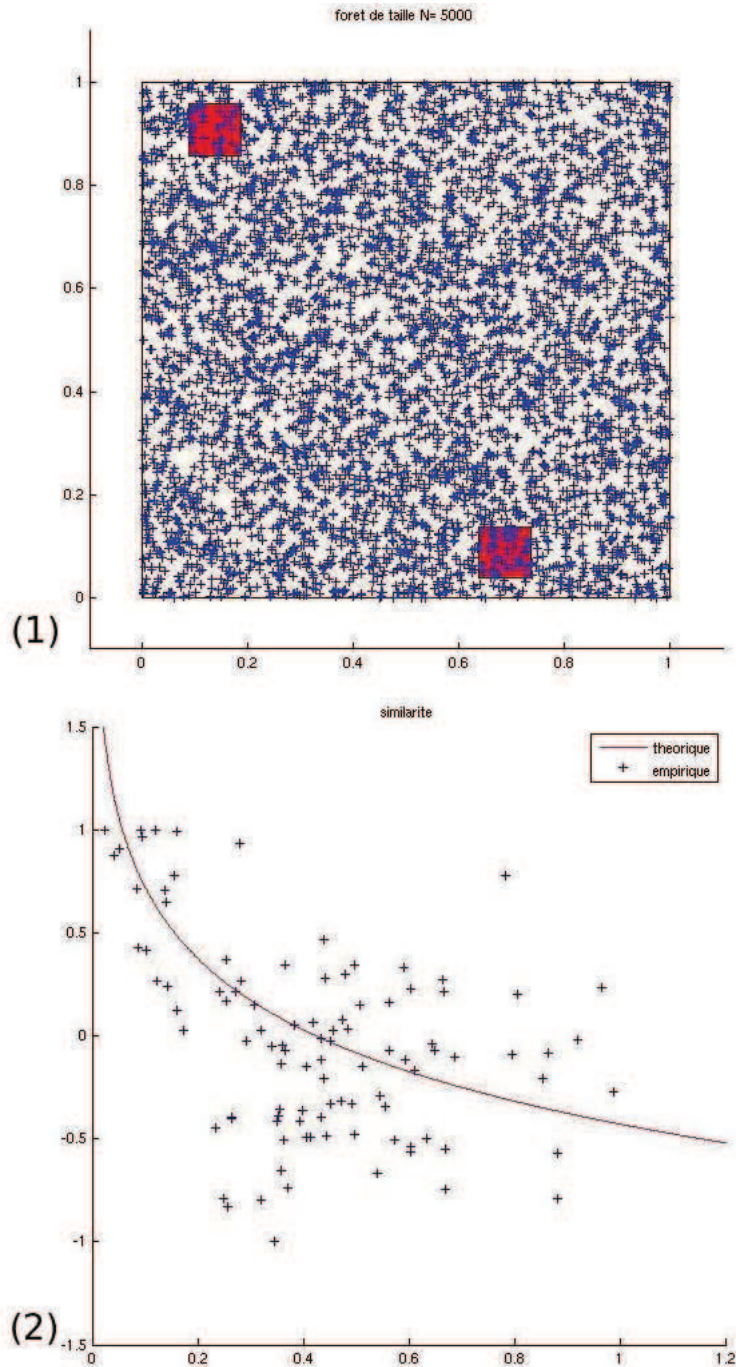


FIG. 3.2 – (1) Exemple de choix de deux parcelles à une certaine distance r l'une de l'autre. (2) [Nuage de points] Simulation de la similarité en composition d'espèces en fonction de la distance entre les deux parcelles considérées, en utilisant notre indice de dissimilarité, pour la forêt de taille $N = 5000$ simulée en figure 3.1, après 10^9 générations, en utilisant la noyau des 4 plus proches voisins, avec, le taux de mutation $\nu = 10^{-4}$ et l'échelle locale $\kappa = 0.001$. [Courbe rouge] Fonction théorique de $r \mapsto 1 - \mathbb{E}_r[\mathcal{F}(A, B)]$, où $\mathbb{E}_r[\mathcal{F}(A, B)] = c + 2 \log(r\sqrt{2\nu}/\sigma)/c_0$ (en prenant $\sigma = 1/\sqrt{N}$, $c_0 = 4$, $C_1 = 0.04$ et $c = (2 + C_1) * 0.7$).

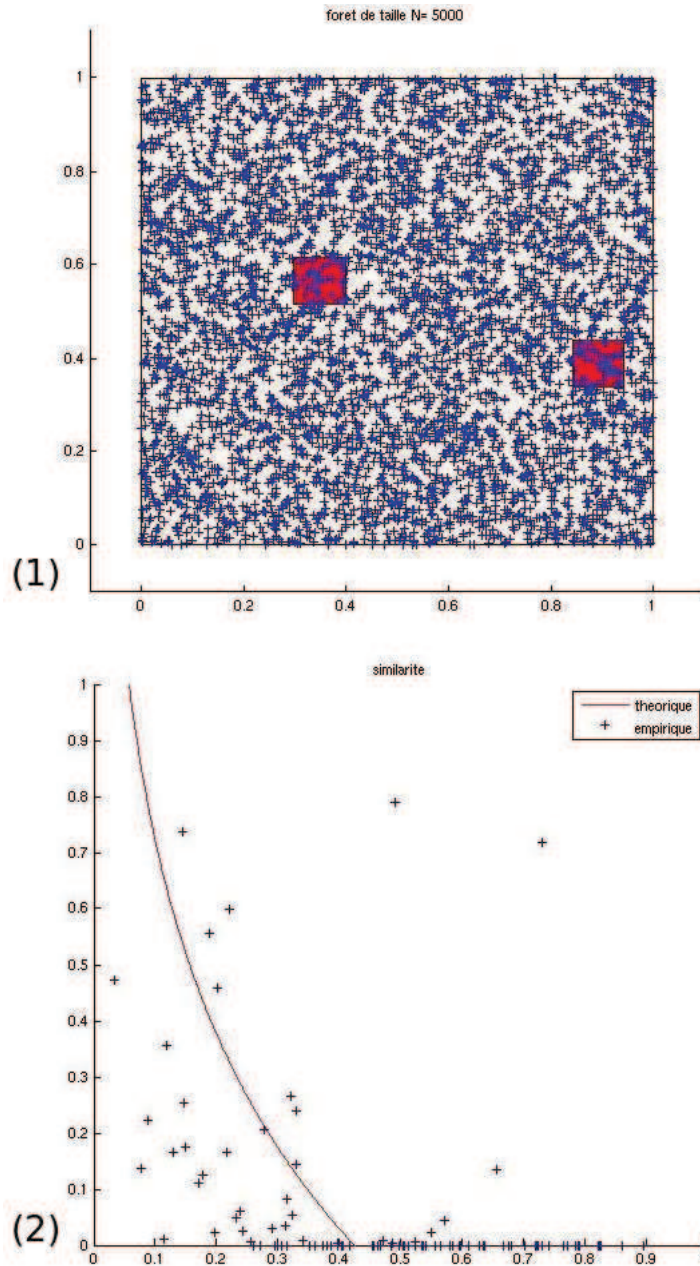


FIG. 3.3 – (1) Exemple de choix de deux parcelles à une certaine distance r l'une de l'autre. (2) [Nuage de points] Simulation de la similarité en composition d'espèces en fonction de la distance entre les deux parcelles considérées, en évaluant la similarité de deux arbres pris au hasard dans chacune des parcelles (en faisant la moyenne sur 10^3 répétitions), pour la forêt de taille $N = 5000$ simulée en figure 3.1, après 10^9 générations, en utilisant la noyau des 4 plus proches voisins, avec, le taux de mutation $\nu = 10^{-4}$ et l'échelle locale $\kappa = 0.001$. [Courbe rouge] Fonction théorique de $r \mapsto 1 - \mathbb{E}_r[\mathcal{F}(A, B)]$, où $\mathbb{E}_r[\mathcal{F}(A, B)] = c + 2 \log(r\sqrt{2\nu}/\sigma)/c_0$ (en prenant $\sigma = 1/\sqrt{N}$, $c_0 = 4$, $C_1 = 0.04$ et $c = (2 + C_1) * 0.7$).

Conclusion

L'article de Chave et Leigh (2002) nous donne une expression de la fonction de similarité en dimension deux, $F(r)$. Nous apprenons finalement comment $F(r)$ décroît avec l'augmentation de la distance qui sépare les communautés comparées. La théorie neutre donne alors des prédictions quantitatives et testables.

L'article de Barton et al. (2013) nous donne des estimations de la taille de voisinage \mathcal{N} et de l'échelle locale κ , dans le cadre du modèle stepping stone, sous l'hypothèse d'un quasi-équilibre local. La théorie de l'isolement par la distance donne alors un modèle neutre clair grâce auquel nous pouvons mettre en évidence les effets de la sélection à des loci spécifiques.

Des simulations nous ont permis de valider notre modèle neutre utilisant l'indice de dissimilarité que nous avons défini. Si nous avons accès à des données (la distribution spatiale, son intensité et les fréquences alléliques), grâce à ce nouveau modèle neutre de β -diversité, nous serions en mesure de proposer des estimations pour $F(0)$ et $F(r)$. Nous ne serions malheureusement pas capable de fournir des estimateurs pour le taux de flux de gènes σ^2 et le taux de mutation ν . Nous ne pourrions avoir que des informations sur le rapport ν/σ^2 , ce qui peut quand même être intéressant...

Bibliographie

Barton N., Depaulis F., Etheridge A., 2002, *Neutral evolution in spatially continuous populations*, Theor. Pop. Biol. 61, 31-48.

Barton N., Etheridge A., Véber A., 2010a, *A new model for evolution in a spatial continuum*, Electron. J. Probab. 15, 162-216.

Barton N., Etheridge A., Kelleher J., 2010b, *A new model for extinction and recolonization in two dimensions : quantifying phylogeography*, Evolution 64, 2701-2715.

Barton N., Etheridge A., Kelleher J. et Véber A., 2013, *Inference in two dimensions : allele frequencies versus lengths of shared sequence blocks*, Theor. Pop. Biol.

Chave J. et Leigh E. G., 2002, *A spatially Explicit Neutral Model of β -Diversity in Tropical Forests*, Theor. Pop. Biol. 62, 153-168.

Clark J. S., Silman M., Kern R., Macklin E., et HilleRisLambers J., 1999, *Seed dispersal near and far : Patterns across temperate and tropical forests*, Ecology 80, 1475-1494.

Condit R., Pitman N., Leigh E. G., Chave J., Terborgh J., Foster R. B., Nuñez V. P., Aguilar S., Valencia R., Villa G., Muller-Landau H., Losos E., Hubbell S. P., 2002, *Beta-diversity in tropical forest trees*, Science 295, 666-669.

Cox J. T., et Griffeath D., 1986, *Diffusive clustering in the 2- dimensional voter model*, Ann. Probab. 14, 347-370.

Cox J. T., 1989, *Coalescing random walks and voter model consensus times on the torus in \mathbb{Z}^d* , Ann. Probab. 17, 1333-1366.

Durrett R.T., 2008, *Probability models for DNA sequence evolution*, 2nd ed., Springer.

Etheridge A., 2008, *Drift, draft and structure : some mathematical models of evolution*, Banach Center Publ. 80, 121-144.

Felsenstein J., 1975, *A pain in the torus : some difficulties with the model of isolation by distance*, Amer. Nat. 109, 359-368.

Gradshteyn I. S. et Ryzhik I. M., 2000, *Tables of Integrals, Series and Products*, A. Jeffrey and D. Zwillinger Eds., 6th ed., Academic Press, San Diego.

Hubbell S. P., 2001, *A Unified Neutral Theory of Biodiversity and Biogeography*, Monographs in Pop. Biol., Princeton Univ. Press, Princeton, NJ.

Kimura M., Weiss G., 1964, *The stepping stone model of population structure and the decrease of genetic correlation with distance*, Genetics 49, 561-576.

Lawler G., Limic V., 2010, *Random walk : a modern introduction*, Cambridge University Press.

Leigh E. G., Wright S. J., Putz F. E., and Herre E. A., 1993, *The decline of tree diversity on newly isolated tropical islands : A test of a null hypothesis and some implications*, Evol. Ecol. 7, 76-102.

Malécot G., 1948, *Les mathématiques de l'hérédité*, Masson, Paris.

Nagylaki T., 1976, *The relation between distant individuals in geographically structured populations*, Math. Biosci. 28, 73-80.

Rousset F., 1997, *Genetic differentiation and estimation of gene flow from f-statistics under isolation by distance*, Genetics 145, 1219-1228.

Rousset F., 2003, *Genetic structure and selection in subdivided populations*, Princeton University Press.

Rousset F., Leblois R., 2007, *Likelihood and approximate likelihood analyses of genetic structure in a linear habitat : performance and robustness to model mis-specification*, Molecular Biology and Evolution 24, 2730-2745.

Sawyer S., 1977a, *Asymptotic properties of the equilibrium probability of density in a geographically structured population*, Adv. Appl. Probab. 9, 268-282.

Sawyer S., 1977b, *Rates of consolidation in a selectively neutral migration model*, Ann. Probab. 5, 486-493.

Wakeley J., 2008, *Coalescent theory : an introduction*, Roberts and Company, Englewood, Colorado.

Wright S., 1943a, *An analysis of local variability in flower color in *Linanthus parryae**, Genetics 28, 139-156.

Wright S., 1943b, *Isolation by distance*, Genetics 28, 114-138.

Annexe

Ici, nous proposons un code (Matlab) permettant de simuler notre modèle, l'évolution en composition d'espèces d'une population au cours des générations.

Modèle neutre de beta-diversité

rho=5000 ; (intensité du processus)
N=rho ; (nombre d'arbres dans la forêt)

F=zeros(3,N) ; (future forêt)
F(1 :2, :)=rand(2,N) ; (emplacements des arbres dans la forêt)

X=F(1, :) ; (abscisses)
Y=F(2, :) ; (ordonnées)

figure()
plot(X,Y,'+') (forêt, espèces non spécifiées)
title(['foret : processus ponctuel de poisson d intensite ', num2str(rho)])

Détermination des 4 plus proches voisin

PPV=zeros(N,4) ; (tableau des plus proches voisins pour tous les arbres de la forêt)

C=F(1 :2,2 :N) ; (cas du premier arbre)
d=zeros(1,N-1) ;
for j=1 :(N-1)
d(j)=[distance(F(1,1),F(2,1),C(1,j),C(2,1))] ;
end
[D,Ix] = sort(d) ;
PPV(1, :)=Ix(1 :4)+ones(1,4) ;

C=F(1 :2,1 :(N-1)) ; (cas du dernier arbre)
d=zeros(1,N-1) ;

```
for j=1 :(N-1)
d(j)=[distance(F(1,N),F(2,N),C(1,j),C(2,j))] ;
end
[D,Ix] = sort(d) ;
PPV(N, :)=Ix(1 :4) ;
```

```
for i=2 :(N-1)
C=F(1 :2,[1 :(i-1),(i+1) :N]) ;
d=zeros(1,N-1) ;
for j=1 :(N-1)
d(j)=[distance(F(1,i),F(2,i),C(1,j),C(2,j))] ;
end
[D,Ix] = sort(d) ;
for k=1 :4
I= Ix(k) ;
if I>=i,
I=I+1 ;
end
PPV(i,k)=I ;
end
end
```

On suppose deux espèces différentes dans la population :

p0=0.5 ; (fréquence de l'espèce 0)
p1=0.5 ; (fréquence de l'espèce 1)

F(3, :)=rand(1,N)>p0 ; (toutes les espèces)

```
E0=[] ; E1=[] ;
for i=1 :N
```

```

if F(3,i)==0,
E0=[E0,F(1 :2,i)];
else E1=[E1,F(1 :2,i)];
end
end

```

```

figure()
hold on
patch([0,1,1,0],[0,0,1,1],'w')
plot(E0(1, :),E0(2, :),'+b')
plot(E1(1, :),E1(2, :),'+r')
axis([-0.1 1.1 -0.1 1.1])
legend('foret',[ 'espece 0, p0= ',num2str(p0)],
[ 'espece 1, p1= ',num2str(p1)])
title(['foret : processus ponctuel de poisson d
intensite ', num2str(rho)])
hold off

```

A chaque pas de temps un arbre meurt :

```

nu=10^(-4); (taux de mutation)

```

```

for j=1 :10^9, (nombre de générations)
m=ceil(N*rand); (arbre mort)
r=ceil(4*rand); (arbre parent)
if m==0,
m=1;
end
if r==0,
r=1;
end
R=PPV(m,r); (remplacant)
if rand<nu
F(3,m)=max(F(3, :)) +1;
else
F(3,m)=F(3,R);
end
end

```

```

Ne=max(F(3, :))+1; (nombre maximum d'es-
pèces)
P=zeros(2,Ne);
P(1, :)=0 : (Ne-1); (P=proportions des es-
pèces)
for j=1 :Ne,
P(2,j)=mean(F(3, :)==(j-1));

```

```

end
Ner=sum(P(2, :)>0); (nombre d'espèces re-
présentées)
P2=P( :,P(2, :)>0); (proportions des especes
représentées)

```

Similarité entre deux parcelles A et B séparées
par une distance donnée ?

```

kappa=0.001; (échelle locale)

```

```

FR=zeros(2,100); (similarité)
for k=1 :100;
a=0.05+0.9.*rand(1,2); (centre de A)
b=0.05+0.9.*rand(1,2); (centre de B)
r=distance(a(1),a(2),b(1),b(2));
while r<kappa,
b=0.05+0.9.*rand(1,2);
r=distance(a(1),a(2),b(1),b(2));
end
FR(1,k)=r;

```

```

A=[a(1)-0.05,a(1)+0.05,a(2)-0.05,a(2)+0.05];
(parcelle A)
B=[b(1)-0.05,b(1)+0.05,b(2)-0.05,b(2)+0.05];
(parcelle B)

```

Composition des parcelles A et B :

```

IA=[]; IB=[];
for i=1 :N
if A(1)<F(1,i) & F(1,i)<A(2) & A(3)<F(2,i)
& F(2,i)<A(4)
IA=[IA i];
end
if B(1)<F(1,i) & F(1,i)<B(2) & B(3)<F(2,i)
& F(2,i)<B(4)
IB=[IB i];
end
end
end

```

```

FA=F( :,IA); (parcelle A)
FB=F( :,IB); (parcelle B)

```

```

NeA=max(FA(3, :))+1; (nombre maximum
d'espèces dans A)

```



```

PA=zeros(2,NeA);
PA(1, :)=0 :(NeA-1); (PA=proportions des
espèces dans A)
for j=1 :NeA,
PA(2,j)=mean(FA(3, :)==(j-1));
end
NerA=sum(PA(2, :)>0); (nombre d'espèces
représentées dans A)
PA2=PA( :,PA(2, :)>0) (proportions des es-
pèces représentées dans A)

```

```

NeB=max(FB(3, :))+1; (nombre maximum
d'espèces dans B)
PB=zeros(2,NeB);
PB(1, :)=0 :(NeB-1); (PB=proportions des
espèces dans B)
for j=1 :NeB,
PB(2,j)=mean(FB(3, :)==(j-1));
end
NerB=sum(PB(2, :)>0); (nombre d'espèces
représentées dans B)
PB2=PB( :,PB(2, :)>0) (proportions des es-
pèces représentées dans B)

```

Similarité (avec l'indice de dissimilarité)

```

FR(2,k)=1;
if min(NeA,NeB)==NeA
for l=1 :NeA
FR(2,k)=FR(2,k)-(PA(2,l)-PB(2,l))^2;
end
for l=(NeA+1) :NeB
FR(2,k)=FR(2,k)-PB(2,l)^2;
end
else
for l=1 :NeB
FR(2,k)= FR(2,k)-(PA(2,l)-PB(2,l))^2;
end
for l=(NeB+1) :NeA
FR(2,k)= FR(2,k)-PA(2,l)^2;
end
end
end

```

Similarité (en évaluant la similarité de 2 arbres pris au hasard dans chacune des parcelles, en faisant la moyenne sur 10^3 répétitions)

```
c=0;
```

```

for i=1 :1000;
mA=ceil(length(FA)*rand);
mB=ceil(length(FB)*rand);
if mA==0,
mA=1;
end
if mB==0,
mB=1;
end
if FA(3,mA)==FB(3,mB)
c=c+1;
end
FR(2,k)=c/1000;
end

```

Prédictions théoriques :

```

sigma=1/sqrt(N);
c0=4;
C1=0.2/(rho*0.001);
c=(2+C1)*0.7;

r=0.001 :0.001 :2;
L=ones(1,length(r));
E=1-c*L-2/c0*log(r*sqrt(2*nu)/sigma);

```

Figure :

```

subplot(2,1,2)
hold on
plot(r,E,'r')
plot(FR(1, :),FR(2, :),'+b')
title('similarite')
legend('theorique','empirique')
hold off

subplot(2,1,1)
hold on
patch([0,1,1,0],[0,0,1,1],'w')
patch([A(1),A(2),A(2),A(1)],[A(3),A(3),A(4),A(4)],'r')
patch([B(1),B(2),B(2),B(1)],[B(3),B(3),B(4),B(4)],'r')
plot(F(1, :),F(2, :),'+b')
axis([-0.1 1.1 -0.1 1.1])
title(['foret de taille N= ', num2str(rho)])
hold off

```