DUALITÉ DES GROUPES ABÉLIENS FINIS ET COMPTAGE DE POINTS

Table des matières

Le but de ce document est d'expliquer un joli pont entre la dualité des groupes abéliens finis (en fait, seul le cas des groupes cycliques sera utilisé dans l'application concrète qu'on présente), et le fait de compter des zéros de polynômes sur un corps fini. Nous allons voir que les relations d'orthogonalité des caractères permettent ainsi de compter des points sur des « courbes » définies sur un corps fini. En écrivant le nombre de points sous la forme d'une somme exponentielle, on peut en déduire une estimation du nombre de points à l'aide de bornes connues sur ce type de sommes. J'ai appris ces résultats en M2, mais je me suis dit qu'ils pouvaient constituer un développement ou au moins un morceau de plan original sur les leçons du style *dualité des groupes abéliens finis, dénombrement, nombres complexes de module* 1... En plus, c'est très élémentaire !

Les applications au comptage de points et au théorème des deux carrés viennent du cours de M2 de Florent Jouve *Exponential sums over finite fields*, je ne saurais pas donner de référence papier... Cependant, une fois que l'on a les relations d'orthogonalité des caractères (qui, elles, se trouvent dans de nombreuses références), c'est vraiment juste un jeu d'écriture. On peut retrouver la plupart des résultats dans le chapitre 2 des notes d'Emmanuel Kowalski *Exponential sums over finite fields : elementary methods* disponibles ici : https://people.math.ethz.ch/~kowalski/exp-sums.pdf

Enfin, même si je ne l'ai pas encore lu en détail, je suis sûr que l'article de Nicolas Tosel : *Biais de Fourier et équations sur un groupe abélien fini*, paru dans la RMS 130-1, donne de belles applications arithmétiques un peu dans le même esprit que celles présentées ici.

1. Quelques rappels de dualité des groupes abéliens finis

Dans cette partie, on rappelle quelques généralités sur la dualité, mais c'est surtout pour se rafraîchir la mémoire et se souvenir dans quel ordre on peut prouver les choses. En réalité, il n'y a pas besoin de tout ceci pour l'application visée. J'essaierai de souligner quand une preuve plus courte est possible dans un cas particulier, ou quand un résultat peut être lu en diagonale si on ne s'intéresse qu'à l'application annoncée en introduction. La plupart des démonstrations se trouvent dans *L'algèbre discrète de la transformée de Fourier*, de Gabriel Peyré.

Définition 1.1 :
Soit G un groupe (pour l'instant, on ne le suppose pas abélien). On appelle *caractère* de G tout morphisme de groupes de G dans $\mathbf{C}^*$. On note $\widehat{G}$ l'ensemble des caractères du groupe G.

Remarque :
Si G est un groupe fini de cardinal n, alors tout caractère de G est à valeurs dans $\mathbf{U}_n := \{z \in \mathbf{C} \mid z^n = 1\}$. En particulier, pour tout $\chi \in \widehat{G}$ et tout $g \in G$, on a $\chi(g)^{-1} = \overline{\chi(g)}$. Cette remarque montre également que lorsque G est fini, $\widehat{G}$ l'est aussi, car ce dernier est inclus dans l'ensemble des fonctions de G dans $\mathbf{U}_n$.

Si $\chi_1, \chi_2 \in \widehat{G}$, on définit :

$$\begin{array}{rccc} \chi_1\chi_2 & : & G & \to & \mathbf{C}^* \\ & & g & \mapsto & \chi_1(g)\chi_2(g) \end{array}$$

et on vérifie que $\chi_1\chi_2$ est aussi un caractère.
Pour ce produit terme à terme, $\widehat{G}$ est un groupe abélien, appelé le *groupe dual* de G. (Pour montrer ce fait, on peut par exemple montrer que $\widehat{G}$ est un sous-groupe du groupe des fonctions de G dans $\mathbf{C}^*$). Puisque $\widehat{G}$ est toujours abélien, même quand G ne l'est pas, un groupe non-abélien n'a aucune chance d'être isomorphe à son dual. Qu'en est-il des groupes abéliens ?

Nous allons montrer que tout groupe abélien fini est isomorphe à son dual.

Pour cela, on commence par le cas des groupes cycliques.

Proposition 1.2 :
Tout groupe cyclique est isomorphe à son dual.

Démo : Soit G un groupe cyclique d'ordre n, et x un générateur de G. Alors se donner un caractère de G, c'est juste se donner l'image de x. Notons $\omega := \mathrm{e}^{\frac{2i\pi}{n}}$. Si $\chi \in \widehat{G}$, comme on sait que les caractères de G sont à valeurs dans $\mathbf{U}_n$, $\chi(x) = \omega^j$ pour un certain $j \in [\![0, n-1]\!]$. Dans ce cas, χ est le caractère $x^k \mapsto \omega^{jk}$, que l'on note χ_j. Donc un caractère de G est nécessairement l'un des χ_j, et réciproquement on vérifie que les χ_j sont bien des caractères. Donc $\widehat{G} = \{\chi_j, j \in [\![0, n-1]\!]\}$. Mais puisque pour tout $j \in [\![1, n-1]\!]$, le caractère χ_j n'est rien d'autre que $(\chi_1)^j$, on en déduit que $\widehat{G}$ est cyclique d'ordre n, engendré par χ_1. Une autre façon de le dire est de montrer que

$$\begin{array}{rccc} \phi & : & \mathbf{Z}/n\mathbf{Z} & \to & \widehat{G} \\ & & j & \mapsto & \chi_j \end{array}$$

est un isomorphisme de groupes. □

Pour les jolies applications de la fin du document, nous aurons seulement besoin du cas des groupes cycliques, donc la suite de cette partie peut être omise. Je l'inclus quand même pour rappeler une manière de montrer que $G \simeq \widehat{G}$ lorsque G est un groupe abélien fini.
Pour passer du cas des groupes cycliques au cas d'un groupe abélien fini quelconque, on aura besoin du résultat suivant :

Lemme 1.3 :
Si G et H sont deux groupes (quelconques), on a un isomorphisme naturel

$$\widehat{G \times H} \simeq \widehat{G} \times \widehat{H}.$$

Démo : Notons

$$\begin{array}{rccc} i_1 & : & G & \to & G \times H \\ & & g & \mapsto & (g, e_H) \end{array} \qquad \begin{array}{rccc} i_2 & : & H & \to & G \times H \\ & & h & \mapsto & (e_G, h) \end{array}$$

On vérifie sans souci que l'application

$$\begin{array}{rccc} \theta & : & \widehat{G \times H} & \to & \widehat{G} \times \widehat{H} \\ & & \varphi & \mapsto & (\varphi \circ i_1, \varphi \circ i_2) \end{array}$$

est un isomorphisme de groupes. Son inverse est donné par :

$$\begin{array}{rccc} \rho & : & \widehat{G} \times \widehat{H} & \to & \widehat{G \times H} \\ & & (\chi_1, \chi_2) & \mapsto & \chi_1 \otimes \chi_2 \end{array}$$

où $\chi_1 \otimes \chi_2$ est l'élément de $\widehat{G \times H}$ suivant :

$$\begin{array}{ccc} G \times H & \to & \mathbf{C}^* \\ (g, h) & \mapsto & \chi_1(g)\chi_2(h) \end{array}$$

□

On a également le petit lemme suivant.

Lemme 1.4 :
Si deux groupes sont isomorphes, leurs duaux le sont aussi.

Démo : Soit $\psi : G \to H$ un isomorphisme entre deux groupes G et H. Alors l'application

$$\begin{array}{ccc} \widehat{G} & \to & \widehat{H} \\ \chi & \mapsto & \chi \circ \psi^{-1} \end{array}$$

est un isomorphisme de groupes entre $\widehat{G}$ et $\widehat{H}$. Son inverse est tout simplement

$$\begin{array}{ccc} \widehat{H} & \to & \widehat{G} \\ \varphi & \mapsto & \varphi \circ \psi \end{array}$$

□

Finalement, pour conclure sur le fait qu'un groupe abélien fini est isomorphe à son dual, on utilise le théorème de structure suivant (admis ici).

Théorème 1.5 (Théorème de structure des groupes abéliens finis : existence) **:**
Si G est un groupe abélien fini, il existe une suite $d_1, \ldots, d_r$ d'entiers supérieurs ou égaux à 2, tels que $d_i \mid d_{i+1}$ et

$$G \simeq (\mathbf{Z}/d_1\mathbf{Z}) \times (\mathbf{Z}/d_2\mathbf{Z}) \times \cdots \times (\mathbf{Z}/d_r\mathbf{Z})$$

Démo : On peut le voir comme un cas particulier du théorème de structure des modules de type fini sur un anneau principal (ici, les $\mathbf{Z}$-modules, c'est-à-dire : les groupes abéliens). Cependant, on peut le démontrer en utilisant un lemme de prolongement de caractères, donc en restant dans le thème de la dualité des groupes abéliens finis. C'est d'ailleurs un bien joli développement (voir par exemple les développements de Benjamin Havret disponibles sur sa page http://www.normalesup.org/~havret/ (où la démonstration de l'unicité des d_i est aussi faite), ou la version d'Émilie, sur sa page : http://perso.eleves.ens-rennes.fr/people/emilie.tezenas-du-montcel/index.html). Le lemme en question est le suivant, que nous ne démontrerons pas non plus, mais c'est fait dans les deux versions citées ci-dessus.

Lemme 1.6 :
Soit G un groupe abélien fini, et soit H un sous-groupe de G. Tout caractère de H se prolonge en un caractère de G.

□

Corollaire 1.7 :
Tout groupe abélien fini est isomorphe à son dual.

Démo : En effet, si G est un groupe abélien fini, on peut utiliser le théorème de structure précédent pour l'écrire (à isomorphisme près) comme un produit de groupes cycliques :

$$G \simeq \prod_{i=1}^{r} \mathbf{Z}/d_i\mathbf{Z}$$

Ensuite, on utilise le lemme 1.4 pour « mettre des chapeaux » de chaque côté de l'isomorphisme :

$$\widehat{G} \simeq \widehat{\prod_{i=1}^{r} \mathbf{Z}/d_i\mathbf{Z}}$$

Puis on utilise le lemme 1.3 pour faire rentrer le chapeau dans le produit, on en déduit que :

$$\widehat{G} \simeq \prod_{i=1}^{r} \widehat{\mathbf{Z}/d_i\mathbf{Z}}$$

Enfin, d'après la proposition 1.2, on a : pour tout i,

$$\widehat{\mathbf{Z}/d_i\mathbf{Z}} \simeq \mathbf{Z}/d_i\mathbf{Z}$$

D'où :

$$\widehat{G} \simeq \prod_{i=1}^{r} \widehat{\mathbf{Z}/d_i\mathbf{Z}} \simeq \prod_{i=1}^{r} \mathbf{Z}/d_i\mathbf{Z} \simeq G$$

□

Remarques :
Dans cette preuve, on utilise seulement l'écriture d'un groupe abélien fini comme un produit de groupes cycliques, sans les conditions de divisibilité du théorème 1.5. Il existe sans doute des démonstrations plus simples de cette version affaiblie du théorème.

On peut aussi remarquer que dans la proposition 1.2, l'isomorphisme que l'on construit dépend du choix d'un générateur de notre groupe cyclique, donc l'isomorphisme entre G et $\widehat{G}$ que l'on construit dans la preuve précédente est loin d'être canonique. On retrouve le même phénomène que dans le cas des espaces vectoriels, un espace vectoriel de dimension finie est isomorphe à son dual, mais l'isomorphisme que l'on construit dépend du choix d'une base. L'analogie va plus loin car on peut montrer qu'on a un isomorphisme canonique entre G et son bidual lorsque G est abélien fini. C'est même exactement le même que dans la théorie des espaces vectoriels : on envoie $x \in G$ sur l'évaluation en x :

$$\begin{array}{rccc} \mathrm{ev}_x & : & \widehat{G} & \to & \mathbf{C}^* \\ & & \chi & \mapsto & \chi(x) \end{array}$$

qui est bien un élément du dual de $\widehat{G}$.

Pour montrer que

$$\begin{array}{rccc} \varphi & : & G & \to & \widehat{\widehat{G}} \\ & & x & \mapsto & \mathrm{ev}_x \end{array}$$

est un isomorphisme, on peut noter qu'il suffit de montrer l'injectivité, puisque G et $\widehat{\widehat{G}}$ ont le même cardinal (en appliquant deux fois le corollaire 1.7). Or, montrer l'injectivité revient à montrer que si $x \in G \setminus \{e\}$, alors il existe un caractère $\chi \in \widehat{G}$ tel que $\chi(x) \neq 1$. Pour cela, on peut par exemple considérer le sous-groupe $\langle x \rangle$ engendré par x. Ce groupe étant cyclique et non trivial, il admet un caractère non trivial d'après la proposition 1.2. Il suffit ensuite de prolonger ce caractère en un caractère de G en utilisant le lemme 1.6, et on conclut.

2. L'orthogonalité des caractères

Dans cette partie, on ne traite que de l'orthogonalité des caractères au sens de la définition 1.1. L'expression « orthogonalité des caractères » peut aussi faire référence à un autre théorème célèbre, portant sur les caractères associés aux représentations irréductibles d'un groupe fini (voir par exemple le chapitre sur les représentations des groupes finis des *éléments d'analyse et d'algèbre* de Pierre Colmez). Les caractères que nous étudions ici, c'est-à-dire les morphismes de groupes à valeurs dans $\mathbf{C}^*$ sont un cas particulier, celui des représentations de dimension 1. Les démonstrations sont alors bien plus rapides.

Définition 2.1 :
Si G est un groupe fini, on notera $\mathbf{C}[G]$ l'ensemble des fonctions de G dans $\mathbf{C}$. Cet ensemble est naturellement muni d'une structure de $\mathbf{C}$-espace vectoriel.

La notation $\mathbf{C}[G]$ sera éclaircie dans quelques lignes.

Définition 2.2 :
Si G est un groupe fini, on peut munir $\mathbf{C}[G]$ du produit hermitien suivant : Si $\varphi, \psi \in \mathbf{C}[G]$,

$$\langle \varphi, \psi \rangle := \frac{1}{|G|} \sum_{x \in G} \varphi(x)\overline{\psi(x)}$$

On note $L^2(G)$ l'espace hermitien obtenu.

Pour $g \in G$, on notera δ_g la fonction de G dans $\mathbf{C}$ qui vaut 1 en g et 0 ailleurs (la fonction indicatrice du singleton $\{g\}$).

Proposition 2.3 :
Si G est un groupe fini, une base de $\mathbf{C}[G]$ est donnée par $(\delta_g)_{g \in G}$. En particulier, on a :

$$\dim_{\mathbf{C}} \mathbf{C}[G] = |G|$$

Démo : On vérifie sans souci que $(\delta_g)_{g \in G}$ est orthogonale, donc libre. Et elle est génératrice grâce à l'écriture de toute fonction $\varphi \in \mathbf{C}[G]$ sous la forme :

$$\varphi = \sum_{g \in G} \varphi(g)\delta_g.$$

C'est cette écriture qui justifie la notation $\mathbf{C}[G]$, via l'identification des élements $g \in G$ et de leur δ_g associé. □

Cependant, cette base $(\delta_g)_{g \in G}$ ne nous dit rien de plus que « se donner une fonction sur G, c'est se donner les valeurs $\varphi(g)$ ». Elle ne tient pas compte de la structure du groupe, elle est la même quel que soit le groupe considéré. Dans le cas des groupes abéliens finis, on a le théorème suivant, qui nous donne une autre base de $L^2(G)$, celle formée par les caractères de G. On obtient ainsi une base formée de fonctions sur G qui ne sont pas tout à fait quelconques : ce sont des morphismes de groupes.

Théorème 2.4 :
Soit G un groupe abélien fini. Le groupe dual $\widehat{G}$ est une base orthonormée de $L^2(G)$.

Démo : Tout repose sur l'assertion suivante : pour tout $\chi \in \widehat{G}$, on a

$$\sum_{x\in G} \chi(x) = \begin{cases} |G| \text{ si } \chi = 1 \text{ (i.e. si } \chi \text{ est le caractère trivial)} \\ 0 \text{ si } \chi \neq 1 \end{cases}$$

En effet, le cas du caractère trivial est clair et ensuite, si χ n'est pas le caractère trivial, soit $t \in G$ tel que $\chi(t) \neq 1$. On a alors :

$$\chi(t)\sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(t)\chi(g) = \sum_{g\in G} \chi(tg) = \sum_{h\in G} \chi(h)$$

en utilisant successivement le fait que χ est un morphisme de groupes, puis que dans un groupe, les translations sont bijectives. Donc

$$(\chi(t) - 1)\sum_{g\in G} \chi(g) = 0$$

et puisque $\chi(t) \neq 1$, c'est que la somme des $\chi(g)$ est nulle, c'est ce qu'on voulait.

Nous pouvons maintenant prouver le théorème : si χ_1, χ_2 sont deux caractères d'un groupe fini G, alors $\chi := \chi_1\overline{\chi_2}$ est aussi un élément de $\widehat{G}$, et donc l'assertion précédente nous dit que la somme

$$\sum_{x\in G} \chi_1(x)\overline{\chi_2(x)} \left(= \sum_{x\in G} \chi(x) \right)$$

est égale au cardinal de G si χ est le caractère trivial (c'est-à-dire si $\chi_1 = \chi_2$), et est nulle sinon (c'est-à-dire si $\chi_1 \neq \chi_2$). Ainsi, $\widehat{G}$ forme une famille orthonormée d'éléments de $L^2(G)$. Or cette famille a le bon cardinal pour être une base de $L^2(G)$, d'après le théorème 1.7 et la proposition 2.3. Donc $\widehat{G}$ forme bien une base orthonormée de $L^2(G)$. □

Une conséquence très utile de cette orthogonalité des caractères est le résultat suivant, qui est celui qui nous servira pour la partie comptage de points.

Corollaire 2.5 :
Soit G un groupe abélien fini. Pour tout $x \in G$, on a

$$\sum_{\chi\in\widehat{G}} \chi(x) = \begin{cases} |G| \text{ si } x = e \\ 0 \text{ si } x \neq e \end{cases}$$

Démo : Puisque $\widehat{G}$ est une base orthonormée de $L^2(G)$, on peut décomposer la fonction δ_e (fonction indicatrice de l'élément neutre) dans cette base :

$$\delta_e = \sum_{\chi\in\widehat{G}} \langle \delta_e, \chi\rangle \chi$$

Or

$$\langle \delta_e, \chi\rangle = \frac{1}{|G|}\sum_{x\in G} \delta_e(x)\overline{\chi(x)} = \frac{1}{|G|}\overline{\chi(e)} = \frac{1}{|G|}$$

D'où

$$\delta_e = \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \chi$$

ce qui prouve le corollaire. □

Remarque :
Une autre manière de montrer ce corollaire est de dire qu'il s'agit juste du théorème 2.4 appliqué au groupe abélien fini $\widehat{G}$. En effet, le théorème appliqué à $\widehat{G}$ nous dit que pour tout $\ell \in \widehat{\widehat{G}}$,

$$\sum_{\chi \in \widehat{G}} \ell(\chi) = \begin{cases} |G| \text{ si } \ell = 1 \text{ (i.e. si } \ell \text{ est le caractère trivial)} \\ 0 \text{ si } \ell \neq 1 \end{cases}$$

Or $\widehat{\widehat{G}} = \{\mathrm{ev}_x,\ x \in G\}$ (d'après l'isomorphisme canonique entre G et son bidual, que nous n'avons pas complètement prouvé ici), et donc en écrivant ℓ sous la forme ev_x, on obtient le corollaire précédent.
Pour résumer, on a démontré le théorème suivant :

Théorème 2.6 (Orthogonalité des caractères) **:**
Soit G un groupe abélien fini, dont on note e l'élément neutre.

(i) Pour tout $\chi \in \widehat{G}$, on a

$$\sum_{x \in G} \chi(x) = \begin{cases} |G| \text{ si } \chi = 1 \text{ (i.e. si } \chi \text{ est le caractère trivial)} \\ 0 \text{ si } \chi \neq 1 \end{cases}$$

(ii) Pour tout $x \in G$, on a

$$\sum_{\chi \in \widehat{G}} \chi(x) = \begin{cases} |G| \text{ si } x = e \\ 0 \text{ si } x \neq e \end{cases}$$

3. Module des sommes de Gauss

Soit $q = p^n$ une puissance d'un nombre premier, et $\mathbf{F}_q$ un corps fini à q éléments.

Définition 3.1 :
- On appelle *caractère additif* de $\mathbf{F}_q$ tout morphisme de groupes de $(\mathbf{F}_q, +)$ dans $\mathbf{C}^*$. Autrement dit, un caractère additif est un élément de $\widehat{(\mathbf{F}_q, +)}$.
- On appelle *caractère multiplicatif* de $\mathbf{F}_q$ tout morphisme de groupes de $(\mathbf{F}_q^\times, \times)$ dans $\mathbf{C}^*$. Autrement dit, un caractère multiplicatif est un élément de $\widehat{(\mathbf{F}_q^\times, \times)}$.

Si χ est un caractère multiplicatif de $\mathbf{F}_q$, on choisit de le prolonger par 0 en 0 si χ n'est pas le caractère trivial, et par 1 en 0 si χ est le caractère trivial. Avec cette convention, on peut voir les caractères multiplicatifs comme des fonctions définies sur tout $\mathbf{F}_q$, et pas seulement sur $\mathbf{F}_q^\times$.

Définition 3.2 :
Si χ est un caractère multiplicatif de $\mathbf{F}_q$ et ψ un caractère additif de $\mathbf{F}_q$, la *somme de Gauss* associée à ces deux caractères est :

$$\tau(\chi, \psi) := \sum_{x \in \mathbf{F}_q} \chi(x)\psi(x)$$

Comme $(\mathbf{F}_q, +)$ et $(\mathbf{F}_q^\times, \times)$ sont des groupes finis, les caractères multiplicatifs et additifs considérés sont à valeurs dans les racines de l'unité. En particulier, une somme de Gauss est une somme de nombres complexes de module 1. On a donc une majoration dite triviale du module de ces sommes :

$$|\tau(\chi, \psi)| \leqslant \sum_{x \in \mathbf{F}_q} \underbrace{|\chi(x)\psi(x)|}_{=1} = q$$

Cependant, nous allons voir que grâce aux relations d'orthogonalité des caractères, on peut dire beaucoup mieux que cette majoration brutale. Commençons par traiter les cas où au moins l'un des deux caractères est trivial :

- *Si χ est trivial mais ψ ne l'est pas :*
  Dans ce cas (puisqu'on a pris la convention de prolonger χ par 1 en 0), on a :
  $$\tau(\chi,\psi) = \sum_{x\in\mathbf{F}_q} \psi(x) = 0$$
  car ψ est un caractère additif non-trivial (point (i) du théorème 2.6).

- *Si ψ est trivial mais χ ne l'est pas :*
  Dans ce cas ψ est le caractère constant égal à 1, et χ est prolongé par 0 en 0. Donc :
  $$\tau(\chi,\psi) = \sum_{x\in\mathbf{F}_q} \chi(x) = \sum_{x\in\mathbf{F}_q^\times} \chi(x) = 0$$
  car χ est non trivial.

- *Si les deux caractères sont triviaux :*
  Dans ce cas
  $$\tau(\chi,\psi) = \sum_{x\in\mathbf{F}_q} 1 = q.$$

Remarquons que dans ces cas particuliers, on a une formule exacte pour $\tau(\chi,\psi)$, et pas seulement une majoration de son module. Il reste à traiter le cas de deux caractères non-triviaux. Dans ce cas, on ne connait pas de formule exacte pour $\tau(\chi,\psi)$, mais on connaît exactement le module de ces sommes de Gauss, qui s'avère être bien plus petit que la majoration triviale ($|\tau(\chi,\psi)| \leqslant q$).

Proposition 3.3 :
Si ψ est un caractère additif non trivial de $\mathbf{F}_q$ et χ un caractère multiplicatif non trivial de $\mathbf{F}_q$, alors :
$$|\tau(\chi,\psi)| = \sqrt{q}$$

Démo : On calcule le module au carré :
$$\begin{aligned}|\tau(\chi,\psi)|^2 &= \tau(\chi,\psi)\overline{\tau(\chi,\psi)} \\ &= \left(\sum_{x\in\mathbf{F}_q} \chi(x)\psi(x)\right)\left(\sum_{y\in\mathbf{F}_q} \overline{\chi(y)}\ \overline{\psi(y)}\right)\end{aligned}$$
Comme χ est non trivial, il a été prolongé par 0 en 0, donc les sommes qui apparaissent portent en fait sur $\mathbf{F}_q^\times$ (pas de contribution des termes où $x = 0$ ou $y = 0$).
Donc :
$$\begin{aligned}|\tau(\chi,\psi)|^2 &= \left(\sum_{x\in\mathbf{F}_q^\times} \chi(x)\psi(x)\right)\left(\sum_{y\in\mathbf{F}_q^\times} \overline{\chi(y)}\ \overline{\psi(y)}\right) \\ &= \sum_{y\in\mathbf{F}_q^\times}\sum_{x\in\mathbf{F}_q^\times} \chi(xy^{-1})\psi(x-y)\end{aligned}$$

Ici, on a utilisé le fait que χ est à valeurs dans les racines de l'unité (et même dans les racines $(q-1)$èmes de l'unité, puisque $|\mathbf{F}_q^\times| = q-1$), donc pour tout $y \in \mathbf{F}_q^\times$, $\overline{\chi(y)} = \chi(y)^{-1} = \chi(y^{-1})$.

De même, ψ étant à valeurs dans les racines de l'unité, $\overline{\psi(y)} = \psi(y)^{-1} = \psi(-y)$. En effet, ψ étant un caractère *additif*, on a $\psi(y)^{-1}$ qui est égal à « ψ de l'inverse de y » *dans le groupe* $(\mathbf{F}_q, +)$ *!* c'est-à-dire : l'opposé de y.

Maintenant, si $y \in \mathbf{F}_q^\times$ est fixé, la multiplication par y^{-1} est une bijection de $\mathbf{F}_q^\times$, donc on peut faire le changement de variables $u = xy^{-1}$ dans la somme intérieure. Celle-ci devient :

$$\sum_{u \in \mathbf{F}_q^\times} \chi(u)\psi(y(u-1))$$

et donc

$$\begin{aligned} |\tau(\chi,\psi)|^2 &= \sum_{y \in \mathbf{F}_q^\times} \sum_{u \in \mathbf{F}_q^\times} \chi(u)\psi(y(u-1)) \\ &= \sum_{u \in \mathbf{F}_q^\times} \chi(u) \sum_{y \in \mathbf{F}_q^\times} \psi(y(u-1)) \end{aligned}$$

Concentrons nous sur la somme intérieure :

- Si $u = 1$, alors :
$$\sum_{y \in \mathbf{F}_q^\times} \psi(y(u-1)) = \sum_{y \in \mathbf{F}_q^\times} \psi(0) = \sum_{y \in \mathbf{F}_q^\times} 1 = q - 1$$
- Si $u \neq 1$, alors : $y \mapsto y(u-1)$ est une permutation de $\mathbf{F}_q^\times$, et donc :
$$\sum_{y \in \mathbf{F}_q^\times} \psi(y(u-1)) = \sum_{v \in \mathbf{F}_q^\times} \psi(v) = \underbrace{\sum_{v \in \mathbf{F}_q} \psi(v)}_{=0 \text{ (orthogonalité)}} - \psi(0) = 0 - 1 = -1$$

  Finalement,

$$\begin{aligned} |\tau(\chi,\psi)|^2 &= \sum_{u \in \mathbf{F}_q^\times} \chi(u) \sum_{y \in \mathbf{F}_q^\times} \psi(y(u-1)) \\ &= \chi(1)(q-1) + \sum_{u \in \mathbf{F}_q^\times \setminus \{1\}} \chi(u)(-1) \\ &= (q-1) - \sum_{u \in \mathbf{F}_q^\times \setminus \{1\}} \chi(u) \end{aligned}$$

  Le caractère multiplicatif χ étant non trivial, on a d'après le point (i) du théorème 2.6 :

$$\sum_{u \in \mathbf{F}_q^\times \setminus \{1\}} \chi(u) = -\chi(1) = -1$$

  D'où $|\tau(\chi,\psi)|^2 = q$, ce qui conclut la preuve.

□

4. Estimation du nombre de points sur une sphère sur un corps fini

Avant de faire un exemple très concret, donnons les grandes lignes de la stratégie qui permet de transformer certains problèmes de comptage de points en des questions d'estimations de sommes exponentielles. Prenons à nouveau $q = p^n$ une puissance d'un nombre premier, et $\mathbf{F}_q$ un corps à q éléments.

On note $\widehat{\mathbf{F}_q}$ le dual de $(\mathbf{F}_q, +)$, c'est-à-dire le groupe des caractères additifs de $\mathbf{F}_q$. Alors pour tout $x \in \mathbf{F}_q$,

$$\sum_{\psi \in \widehat{\mathbf{F}_q}} \psi(x) = \begin{cases} q \text{ si } x = 0 \\ 0 \text{ si } x \neq 0 \end{cases}$$

Autrement dit,

$$\frac{1}{q} \sum_{\psi \in \widehat{\mathbf{F}_q}} \psi(x)$$

est la fonction indicatrice du singleton $\{0\}$ définie sur $\mathbf{F}_q$. C'est cela qui va permettre de compter les zéros de polynômes sur un corps fini.

Par exemple, prenons un polynôme $P \in \mathbf{F}_q[X_1, \ldots, X_m]$. On peut se demander si l'équation

$$P(x_1, \ldots, x_m) = 0$$

a des solutions $(x_1, \ldots, x_m) \in \mathbf{F}_q^m$, et on peut vouloir estimer le nombre de solutions. Cette question peut se traduire en termes de sommes exponentielles. En effet, si $(x_1, \ldots, x_m) \in \mathbf{F}_q^m$, on a :

$$\frac{1}{q} \sum_{\psi \in \widehat{\mathbf{F}_q}} \psi(P(x_1, \ldots, x_m)) = \mathbb{1}_{P(x_1,\ldots,x_m)=0} = \begin{cases} 1 \text{ si } P(x_1, \ldots, x_m) = 0 \\ 0 \text{ sinon} \end{cases}$$

Ainsi, le nombre de points de $\mathbf{F}_q^m$ sur la courbe d'équations $P(x_1, \ldots, x_m) = 0$ est donné par :

$$N = \frac{1}{q} \sum_{x_1,\ldots,x_m \in \mathbf{F}_q} \sum_{\psi \in \widehat{\mathbf{F}_q}} \psi(P(x_1, \ldots, x_m))$$

On a donc exprimé le nombre de points sur cette courbe comme une somme de racines de l'unité (rappelons que ψ est à valeurs dans les racines q^{e} de l'unité). On peut avoir l'impression de n'avoir rien gagné, car nous n'avons pas décrit explicitement $\widehat{\mathbf{F}_q}$. Cependant, dans le cas d'un corps premier ($q = p$ un nombre premier), nous l'avons déjà fait !

En effet, dans ce cas $(\mathbf{F}_p, +)$ est un groupe cyclique, et nous avons décrit le dual de tels groupes à la proposition 1.2. On a montré que :

$$\widehat{\mathbf{F}_p} = \{\psi_h,\ h \in \mathbf{F}_p\}$$

où

$$\begin{array}{rccc} \psi_h & : & \mathbf{F}_p & \to & \mathbf{C}^* \\ & & x & \mapsto & \exp\left(\frac{2i\pi hx}{p}\right) \end{array}$$

Notation : Pour alléger un peu, on notera $e(z) := \exp(2i\pi z)$ pour tout complexe z.

Application : nombre de points sur les sphères dans $\mathbf{F}_p$

Soit p un nombre premier supérieur ou égal à 3. Pour $a \in \mathbf{F}_p$, on note

$$S(a,p) := \{(x,y,z) \in \mathbf{F}_p^3 \mid x^2+y^2+z^2 = a\}$$

On veut préciser le comportement asymptotique de $|S(a,p)|$ lorsque p devient grand.

D'après la stratégie expliquée au dessus, on peut écrire $|S(a,p)|$ comme une somme exponentielle :

$$|S(a,p)| = \frac{1}{p} \sum_{x,y,z \in \mathbf{F}_p} \sum_{\psi \in \widehat{\mathbf{F}_p}} \psi(x^2+y^2+z^2-a)$$

Ensuite, on utilise la description explicite de $\widehat{\mathbf{F}_p}$ pour obtenir :

$$\begin{aligned} |S(a,p)| &= \frac{1}{p} \sum_{x,y,z \in \mathbf{F}_p} \sum_{h \in \mathbf{F}_p} \psi_h(x^2+y^2+z^2-a) \\ &= \frac{1}{p} \sum_{x,y,z \in \mathbf{F}_p} \sum_{h \in \mathbf{F}_p} e\left(\frac{h}{p}(x^2+y^2+z^2-a)\right) \\ &= \frac{1}{p} \sum_{h \in \mathbf{F}_p} \tau(h)^3 e\left(\frac{-ah}{p}\right) \end{aligned}$$

où l'on a noté

$$\tau(h) := \sum_{x \in \mathbf{F}_p} e\left(\frac{hx^2}{p}\right)$$

Lemme 4.1 :
Pour tout $h \in \mathbf{F}_p^\times$, $|\tau(h)| = \sqrt{p}$.

Démo : Comme dans le calcul du module des sommes de Gauss (proposition 3.3), on calcule le carré du module :

$$|\tau(h)|^2 = \tau(h)\overline{\tau(h)} = \sum_{x,y \in \mathbf{F}_p} e\left(\frac{h}{p}(x^2-y^2)\right) = \sum_{x,y \in \mathbf{F}_p} e\left(\frac{h}{p}(x-y)(x+y)\right)$$

On effectue le changement de variables linéaire

$$\begin{cases} u = x-y \\ v = x+y \end{cases}$$

de matrice $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ inversible dans $\mathcal{M}_2(\mathbf{F}_p)$ car elle est de déterminant 2 et $p \geqslant 3$. D'où :

$$|\tau(h)|^2 = \sum_{u,v \in \mathbf{F}_p} e\left(\frac{h}{p}uv\right)$$

On sépare ensuite le terme en $u = 0$ des autres termes :

$$|\tau(h)|^2 = \sum_{v \in \mathbf{F}_p} 1 + \sum_{u \in \mathbf{F}_p^\times} \sum_{v \in \mathbf{F}_p} e\left(\frac{hu}{p}v\right)$$

Or pour tout $u \in \mathbf{F}_p^\times$, $v \mapsto e\left(\frac{hu}{p}v\right)$ est un caractère additif non-trivial de $\mathbf{F}_p$ (on utilise également ici le fait que h est non nul), donc par orthogonalité,

$$\sum_{v \in \mathbf{F}_p} e\left(\frac{hu}{p}v\right) = 0$$

Ainsi, la double somme dans l'expression de $|\tau(h)|^2$ est nulle, et il ne reste que le premier terme, égal à p. Donc $|\tau(h)|^2 = p$, d'où la conclusion. □

Remarque :
On peut aussi démontrer ce lemme en faisant apparaître une somme de Gauss, de la manière suivante : pour tout $h \in \mathbf{F}_p^\times$,

$$\tau(h) = \sum_{x \in \mathbf{F}_p} e\left(\frac{hx^2}{p}\right) = \sum_{y \in \mathbf{F}_p} \left(1 + \left(\frac{y}{p}\right)\right) e\left(\frac{h}{p}y\right)$$

où $\left(\frac{\cdot}{p}\right)$ est le symbole de Legendre :

$$\left(\frac{y}{p}\right) := \begin{cases} 1 \text{ si } y \text{ est un carré non nul dans } \mathbf{F}_p \\ 0 \text{ si } y = 0 \\ -1 \text{ si } y \text{ n'est pas un carré dans } \mathbf{F}_p \end{cases}$$

En séparant la somme en deux, on obtient :

$$\tau(h) = \sum_{y \in \mathbf{F}_p} e\left(\frac{h}{p}y\right) + \sum_{y \in \mathbf{F}_p} \left(\frac{y}{p}\right) e\left(\frac{h}{p}y\right)$$

La première somme est nulle car $e\left(\frac{h}{p}\cdot\right)$ est un caractère additif non-trivial, et la deuxième somme est de module $\sqrt{p}$ d'après la proposition 3.3, car c'est la somme de Gauss associée au caractère multiplicatif non-trivial donné par le symbole de Legendre et au caractère additif non-trivial $e\left(\frac{h}{p}\cdot\right)$.

Théorème 4.2 :
Pour tout p premier supérieur ou égal à 3, pour tout $a \in \mathbf{F}_p$, on a :

$$\left||S(a,p)| - p^2\right| \leqslant \frac{p-1}{p} p^{\frac{3}{2}}$$

En particulier,

$$|S(a,p)| \underset{p\infty}{=} p^2 + O\left(p^{\frac{3}{2}}\right)$$

où la constante implicite dans le O ne dépend pas de a.

Démo : On a déjà montré que

$$|S(a,p)| = \frac{1}{p} \sum_{h \in \mathbf{F}_p} \tau(h)^3 e\left(\frac{-ah}{p}\right)$$

En isolant le terme en $h = 0$, on obtient :

$$|S(a,p)| = \frac{\tau(0)^3}{p} + \frac{1}{p} \sum_{h \in \mathbf{F}_p^\times} \tau(h)^3 e\left(\frac{-ah}{p}\right)$$

Or $\tau(0) = p$, d'où :

$$|S(a,p)| = p^2 + \frac{1}{p} \sum_{h \in \mathbf{F}_p^\times} \tau(h)^3 e\left(\frac{-ah}{p}\right)$$

Ainsi,

$$\begin{aligned} \big||S(a,p)| - p^2\big| &= \left| \frac{1}{p} \sum_{h \in \mathbf{F}_p^\times} \tau(h)^3 e\left(\frac{-ah}{p}\right) \right| \\ &\leqslant \frac{1}{p} \sum_{h \in \mathbf{F}_p^\times} |\tau(h)|^3 \\ &\leqslant \frac{p-1}{p} p^{\frac{3}{2}} \end{aligned}$$

(où la dernière inégalité vient du lemme 4.1). Ceci conclut la preuve de la première assertion. En outre, comme $\frac{p-1}{p} \leqslant 1$, on a :

$$\big||S(a,p)| - p^2\big| \leqslant p^{\frac{3}{2}}$$

ce qui montre bien que

$$|S(a,p)| \underset{p\infty}{=} p^2 + O\left(p^{\frac{3}{2}}\right)$$

où la constante implicite dans le O est absolue. □

Remarque :
Expliquons avec les mains pourquoi il n'est pas étonnant de trouver p^2 comme terme dominant dans ce développement asymptotique :
Choisir $x, y, z \in \mathbf{F}_p$ tels que $x^2+y^2+z^2 = a$, c'est choisir x et y quelconques (p^2 choix possibles), puis choisir (si possible) z tel que $z^2 = a - x^2 - y^2$. Or

- si $a - x^2 - y^2$ n'est pas un carré dans $\mathbf{F}_p$, aucun z ne convient.
- si $a - x^2 - y^2$ est un carré dans $\mathbf{F}_p$, alors il y a deux valeurs de z qui conviennent (sauf si $a - x^2 - y^2 = 0$, mais bon, disons que la plupart du temps, il y aura deux racines carrées, et exceptionnellement une seule).

Comme il y a autant de carrés que de non-carrés dans $\mathbf{F}_p^\times$, et que les carrés nous donnent 2 valeurs possibles pour z, et les non-carrés aucune valeur possible, on peut se dire qu'en moyenne, cela se passe comme si pour chaque valeur de (x, y) il y avait un unique z convenable. Tout ceci est très heuristique, mais fait qu'on ne sursaute pas trop à la vue du résultat. Les estimations de sommes exponentielles permettent de montrer rigoureusement le résultat, et d'avoir une estimation de l'erreur commise en disant que $|S(a,p)|$ est environ égal à p^2.

5. Une autre application arithmétique : un point clef du théorème des deux carrés

Dans cette section on introduit les sommes de Jacobi, et on déduira de la proposition 3.3 le module de ces sommes. Une conséquence presque immédiate de la connaissance du module des sommes de Jacobi est un point important du théorème des deux carrés : le fait que tout nombre premier $p \equiv 1 \bmod 4$ est la somme de deux carrés d'entiers.

Définition 5.1 :
Soit $q = p^n$ une puissance d'un nombre premier, et $\mathbf{F}_q$ un corps fini à q éléments. Soit χ, φ deux caractères multiplicatifs de $\mathbf{F}_q$. On définit la *somme de Jacobi* associée à ces deux caractères :

$$J(\chi,\varphi) := \sum_{x\in\mathbf{F}_q} \chi(x)\varphi(1-x)$$

avec les conventions précédentes pour étendre les caractères multiplicatifs en 0.

Faisons le lien entre ces sommes et les sommes de Gauss introduites plus haut.

Proposition 5.2 :
Soit χ, φ deux caractères multiplicatifs non-triviaux de $\mathbf{F}_q$. Si $\chi\varphi$ est aussi non-trivial, alors pour tout ψ caractère *additif* non-trivial de $\mathbf{F}_q$ on a :

$$J(\chi,\varphi) = \frac{\tau(\chi,\psi)\tau(\varphi,\psi)}{\tau(\chi\varphi,\psi)}$$

où τ désigne la somme de Gauss (cf. définition 3.2)

Démo : En revenant aux définitions, on a :

$$J(\chi,\varphi)\tau(\chi\varphi,\psi) = \sum_{x\in\mathbf{F}_q} \chi(x)\varphi(1-x) \sum_{y\in\mathbf{F}_q} \chi(y)\varphi(y)\psi(y)$$

Comme χ et φ sont non-triviaux, ils sont prolongés par 0 en 0, donc on peut restreindre la somme aux $x \notin \{0,1\}$, et aux $y \neq 0$:

$$\begin{aligned} J(\chi,\varphi)\tau(\chi\varphi,\psi) &= \sum_{x\in\mathbf{F}_q\setminus\{0,1\}} \sum_{y\in\mathbf{F}_q^\times} \chi(x)\varphi(1-x)\chi(y)\varphi(y)\psi(y) \\ &= \sum_{x\in\mathbf{F}_q\setminus\{0,1\}} \sum_{y\in\mathbf{F}_q^\times} \chi(xy)\varphi((1-x)y)\psi(y) \end{aligned}$$

Ensuite, on peut vérifier que l'application :

$$(x,y) \mapsto (xy, y-xy)$$

établit une bijection entre $(\mathbf{F}_q\setminus\{0,1\})\times\mathbf{F}_q^\times$ et $\{(u,v)\in\mathbf{F}_q^\times\times\mathbf{F}_q^\times,\ u+v\neq 0\}$, dont la réciproque est donnée par

$$(u,v) \mapsto \left(\frac{u}{u+v}, u+v\right)$$

On peut donc réindexer la somme en posant $u := xy$ et $v := y - xy$. On obtient :

$$J(\chi,\varphi)\tau(\chi\varphi,\psi) = \sum_{\substack{u\in\mathbf{F}_q^\times \\ u+v\neq 0}} \sum_{v\in\mathbf{F}_q^\times} \chi(u)\varphi(v)\underbrace{\psi(u+v)}_{=\psi(u)\psi(v)}$$

En ajoutant puis en ôtant les termes en $u+v=0$, on peut réécrire cette somme :

$$J(\chi,\varphi)\tau(\chi\varphi,\psi) = \underbrace{\left(\sum_{u\in\mathbf{F}_q^\times} \chi(u)\psi(u)\right)}_{\tau(\chi,\psi)} \underbrace{\left(\sum_{v\in\mathbf{F}_q^\times} \varphi(v)\psi(v)\right)}_{\tau(\varphi,\psi)} - \sum_{u\in\mathbf{F}_q^\times} \chi(u)\varphi(-u)\psi(0)$$

(on utilise le fait que χ et φ sont non-triviaux pour identifier ces sommes sur $\mathbf{F}_q^\times$ aux sommes de Gauss, initialement définies sur tout $\mathbf{F}_q$. En effet, nous avons pris la convention d'étendre les

caractères multiplicatifs non-triviaux par 0 en 0). Pour conclure, il ne reste plus qu'à montrer que la dernière somme est nulle.
Or

$$\sum_{u\in\mathbf{F}_q^\times}\chi(u)\varphi(-u)\psi(0)=\sum_{u\in\mathbf{F}_q^\times}\chi(u)\varphi(-u)=\varphi(-1)\sum_{u\in\mathbf{F}_q^\times}(\chi\varphi)(u)$$

et

$$\sum_{u\in\mathbf{F}_q^\times}(\chi\varphi)(u)=0$$

car $\chi\varphi$ est un caractère multiplicatif non-trivial de $\mathbf{F}_q$. Finalement, on a bien montré l'égalité

$$J(\chi,\varphi)\tau(\chi\varphi,\psi)=\tau(\chi,\psi)\tau(\varphi,\psi)$$

(et on peut bien diviser par $\tau(\chi\varphi,\psi)$, car c'est un complexe de module $\sqrt{q}$ d'après la proposition 3.3, donc il est en particulier non nul). □

Corollaire 5.3 :
Si χ,φ sont deux caractères multiplicatifs non-triviaux de $\mathbf{F}_q$ tels que $\chi\varphi$ est aussi non-trivial, on a :

$$|J(\chi,\varphi)|=\sqrt{q}$$

Démo : Cela vient immédiatement de la proposition précédente et de la proposition 3.3 sur le module des sommes de Gauss. □

Application : Tout nombre premier $p \equiv 1 \bmod 4$ est somme de deux carrés

Si p est un nombre premier, le groupe $(\mathbf{F}_p^\times, \times)$ est cyclique d'ordre $p-1$. En particulier, dans le cas où $p \equiv 1 \bmod 4$, c'est un groupe cyclique d'ordre divisible par 4. Donc il en est de même de son dual $\widehat{\mathbf{F}_p^\times}$ d'après la proposition 1.2.

Maintenant, c'est un fait général que si G est un groupe cyclique d'ordre n, alors pour tout diviseur d de n, il existe un élément d'ordre d dans G (il suffit d'élever un générateur du groupe cyclique G à la puissance $\frac{n}{d}$).

Ainsi, si p est un nombre premier congru à 1 modulo 4, il existe un élément d'ordre 4 dans $\widehat{\mathbf{F}_p^\times}$, c'est-à-dire un caractère multiplicatif d'ordre 4. Notons χ un tel caractère. Alors $\chi_2 := \chi^2$ est d'ordre 2. En particulier, χ et χ_2 sont des caractères multiplicatifs non-triviaux de $\mathbf{F}_p$, et leur produit $\chi\chi_2$ est également non-trivial (χ_2 ne peut pas être l'inverse de χ car ils n'ont pas le même ordre).

On peut donc appliquer le corollaire 5.3 à la somme de Jacobi associée aux caractères χ et χ_2 :

$$|J(\chi, \chi_2)|^2 = p$$

Enfin, comme χ est d'ordre 4, il est à valeurs dans les racines 4^{e} de l'unité, c'est-à-dire dans $\{\pm 1, \pm i\}$. Pour cette même raison d'ordre, χ_2 est à valeurs dans $\{\pm 1\}$.

Mais si l'on revient à la définition des sommes de Jacobi :

$$J(\chi, \chi_2) = \sum_{x \in \mathbf{F}_p} \chi(x)\chi_2(1-x)$$

on voit que $J(\chi, \chi_2)$ appartient à $\mathbf{Z}[i]$. Autrement dit, il existe $(a, b) \in \mathbf{Z}^2$ tels que $J(\chi, \chi_2) = a + ib$. Ainsi, l'égalité

$$|J(\chi, \chi_2)|^2 = p = a^2 + b^2$$

nous donne une écriture de p comme somme de deux carrés d'entiers.