An Introduction to Gauss and Jacobi Sums in Number Theory

Geoffrey Baring

February 7, 2023

Abstract

In this talk, I introduce the notion of Jacobi sums in the context of their application to the problem of counting solutions to diophantine equations over finite fields. We first introduce the notion of a Dirichlet character (a kind of generalization of the Legendre symbol) and use it to construct first Gauss and then Jacobi sums, and see a relation between the two. Finally, we use the latter to compute the numbers of solutions in a given finite field to certain diophantine equations.

1 Introduction

Last week, we were introduced to the Legendre symbol $\left(\frac{\cdot}{p}\right)$, which tells us when a certain integer is a *quadratic residue* modulo prime p, i.e. when it is a square in $\mathbb{F}_p^\times$. In this talk, we generalize the Legendre symbol to the notion of a "multiplicative" or "Dirichlet" character. From these characters we can construct Gauss and Jacobi sums, which we will see is an essential tool in number theory. We motivate this tool by considering the problem of counting solutions to Diophantine equations over finite fields.

2 Counting Solutions

Suppose we're given a Diophantine equation, such as $x^2 + y^2 = a$, and a finite field $\mathbb{F}_p$. A solution to this equation is a pair $(x, y) \in \mathbb{F}_p^2$ which satisfies the equation, or, more generally, a pair of congruence classes in $\mathbb{F}_p$ $[x]$ and $[y]$ such that $x^2 + y^2 \equiv a \pmod{p}$, considering $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$. Because $\mathbb{F}_p^2$ is finite, we can ask, how many such pairs of x and y satisfy this equation? and furthermore, does the number of solutions relate to the prime p we choose? For a given diophantine equation $\mathfrak{E}$, we denote the number of solutions $N(\mathfrak{E})$, with $\mathbb{F}_p$ implied. It is straightforward to write a computer program to compute $N(\mathfrak{E})$ for low primes and simple equations; here is a table for a few simple examples.

p	$x^2+y^2=0$	$x^2+y^2+1=0$	$x^2+2xy+y^2+1=0$	$y^2=x^3+1$
2	2	2	2	2
3	1	4	0	3
5	9	4	10	5
7	1	8	0	11
11	1	12	0	11
13	25	12	26	11
17	33	16	34	17
19	1	20	0	11
23	1	24	0	23
29	57	28	58	29
31	1	32	0	35
37	73	36	74	47
41	81	40	82	41
43	1	44	0	35
47	1	48	0	47
53	105	52	106	53

for equations

$$x^2+y^2=0 \tag{1}$$
$$x^2+y^2=-1 \tag{2}$$
$$x^2+2xy+y^2=-1 \tag{3}$$
$$y^2=x^3+1. \tag{4}$$

For the first three equations, a clear pattern in p emerges: there is a formula to compute $N(\mathfrak{E})$ when $p \equiv 1 \pmod 4$, and a different formula when $p \equiv 3 \pmod 4$. We will see what dictates this difference later, and why the cubic equation $y^2 = x^3 + 1$ is not so straightforward. To do so, we will need to build the language of Dirichlet characters and Gauss and Jacobi sums.

3 Dirichlet Characters

Recall that the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{F}_p \to \{0, \pm 1\}$ is defined by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & a = 0 \\ 1 & a = x^2, x \in \mathbb{F}_p \\ -1 & \text{otherwise.} \end{cases}$$

We also have (Euler's Criterion):

Proposition 3.1. *if $p > 2$ is prime, then $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \pmod p$.*

See Nico's talk notes from last week for a proof of Euler's Criterion. From this proposition (or, simply, from the definition of the Legendre symbol), it is easy to see that $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ for any $a, b \in \mathbb{F}_p$; in other words, $\left(\frac{\cdot}{p}\right)$ is *multiplicative*. With this in mind, let's generalize the notion.

Definition 3.1. *A multiplicative or Dirichlet character of order/modulus n on a finite field $\mathbb{F}_p$ is a map $\chi : \mathbb{F}_p \to \mathbb{C}$ such that*

1. *$\chi(0) = 0$, and*
2. *for all $a, b \in \mathbb{F}_p^\times$, $\chi(ab) = \chi(a)\chi(b)$.*

Note that the second condition implies that $\chi(1) = 1$, since $\chi(a) = \chi(1 \cdot a) = \chi(1) \cdot \chi(a)$ for any $a \in \mathbb{F}_p$. From the definition we then have that χ restricted to $\mathbb{F}_p^\times$ is a group homomorphism to $\mathbb{C}$, and since $\mathbb{F}_p^\times$ is cyclic this implies that its codomain must consist entirely of roots of unity. All elements of $\chi(\mathbb{F}_p^\times)$ must therefore be n-th roots of unity, giving the *order* of χ as the lowest $n \in \mathbb{N}$ such that this image is a subset of the n-th roots of unity in $\mathbb{C}$.

From this it is easy to see that the Legendre symbol is a multiplicative character of order 2 on $\mathbb{F}_p$.

Definition 3.2. *For a given p, we call the multiplicative character $\varepsilon : \mathbb{F}_p \to \mathbb{C}$ defined by $\varepsilon(a) = 1$ for all $a \in \mathbb{F}_p$ the trivial or principal character.*

Warning 3.1. *Note that because we define $\varepsilon(0) = 1$, this construction of the trivial character contradicts with our definition of a multiplicative character. However, we will proceed with this definition for reasons which will become clear given the next proposition.*

For a given p, we then have $\varepsilon\chi = \chi\varepsilon = \chi$ for any character on $\mathbb{F}_p$, where the multiplication of two characters is defined pointwise, i.e. $\chi\psi(a) = \chi(a)\psi(a)$. ε thus acts as a kind of identity on multiplicative characters out of $\mathbb{F}_p$.

Proposition 3.2. *Let p be an odd prime, n an integer such that $n|(p-1)$, and $a \in \mathbb{F}_p$. Then* $N(x^n = a) = \sum_{\chi^n = \varepsilon} \chi(a)$.

Proof. Note that $\chi^n = \varepsilon$ is the condition that χ has order dividing n. Let r be a generator of $\mathbb{F}_p^\times$. Since $n|(p-1)$, there are n characters of order dividing n, each determined by to which n-th root of unity it maps r, Denote the set of these characters $\mathcal{C}$. If $a = 0$, then $x^n = a$ has exactly one solution, namely $a = 0$, since $\mathbb{F}_p$ is an integral domain. In this case $\chi(0) = 0$ for all $\chi \in \mathcal{C}$, except for ε, which has $\varepsilon(0) = 1$[1]. In this case the given sum is $N(x^n = a)$.

If $a \neq 0$, suppose $x^n = a$ is solvable. Then there is a $b \in \mathbb{F}_p^\times$ such that $b^n = a$. Then for all $\chi \in \mathcal{C}$, $\chi(a) = \chi(b^n) = \chi^n(b) = \varepsilon(b) = 1$, so the given sum is n, which is the number of solutions to $x^n = a$. If $x^n = a$ is not solvable, then we must show the given sum is 0. Denote it by T. Since $\mathbb{F}_p^\times$ is generated by r, we have $a = r^k$ for some $k < p$, with $k \not| n$ by the assumption that $x^n = a$ is not solvable. Let λ be the multiplicative character defined by $\lambda(r^k) = e^{2\pi i(\frac{k}{p-1})}$, and consider $\chi = \lambda^{\frac{p-1}{n}}$. $\chi(a) = \chi^k(r) = e^{2\pi i(\frac{k}{n})} \neq 1$, and $\chi^n = \lambda^{p-1} = \varepsilon$. Then $\chi(a)T$ simply shuffles around the summands of T, so $\chi(a)T = T$, which can only be true if $T = 0$. □

Remark 3.1. *If p is odd and $n = 2$, we have $N(x^2 = a) = 1 + \left(\frac{a}{p}\right)$. Similarly, $N(x^3 = a) = 1 + \chi(a) + \chi^2(a)$, where $\chi : \mathbb{F}_p \to \mathbb{C}$ is the character of order 3 defined by $\chi(r) = e^{\frac{2\pi i}{3}}$.*

[1] See Warning **??** :)

4 Gauss Sums

Now that we have seen Dirichlet characters, we can dive right into the notions of Gauss and Jacobi sums and begin playing around with them. For the rest of this talk, we denote by ζ_n the primitive n-th root of unity $\zeta_n = e^{\frac{2\pi i}{n}}$.

Definition 4.1. *Let χ be a character on $\mathbb{F}_p$, $a \in \mathbb{F}_p$. A sum of the form $\sum_{t\in\mathbb{F}_p} \chi(t)\zeta_p^{at}$ is called a Gauss sum on $\mathbb{F}_p$, and is denoted $g_a(\chi)$.*

In the case where $a = 1$ the Gauss sum is usually abbreviated $g(\chi)$, and if χ is fixed, the sum is often written simply as g_a or g. If χ is "quadratic", or of order 2, then $g_a(\chi)$ is a quadratic Gauss sum. For example, when χ is the Legendre symbol with respect to p, then we have

$$g(\chi) = \sum_{t\in\mathbb{F}_p} \left(\frac{t}{p}\right)\zeta_p^t = \sum_{t\in\mathbb{F}_p} \left(\frac{t}{p}\right)e^{\frac{2\pi i}{p}t}.$$

Proposition 4.1. *The above Gauss sum is given by*

$$g(\chi) = \begin{cases} \sqrt{p} & p \equiv 1 (mod\ 4) \\ i\sqrt{p} & p \equiv 3 (mod\ 4) \end{cases}$$

Proof. We will show first that $g^2 = \left(\frac{-1}{p}\right)p$, and then use the complex geometry of the summands to obtain the result.

$g^2 = \left(\sum_{k=1}^{p-1}\left(\frac{i}{p}\right)\zeta_p^i\right)\left(\sum_{j=1}^{p-1}\left(\frac{j}{p}\right)\zeta_p^j\right) = \sum_{i,j=1}^{p-1}\left(\frac{i}{p}\right)\left(\frac{j}{p}\right)\zeta_p^{i+j}$. Note that by grouping terms, g^2 can be written as a polynomial in ζ_p of degree $p-1$: $g^2 = a_0 + a_1\zeta_p + \cdots + a_{p-1}\zeta_p^{p-1}$, with coefficients $a_n = \sum_{i+j=n}\left(\frac{i}{p}\right)\left(\frac{j}{p}\right)$. Then $a_0 = \sum_{i\in\mathbb{F}_p}\left(\frac{i}{p}\right)\left(\frac{-i}{p}\right) = \left(\frac{-1}{p}\right)\sum_{i\in\mathbb{F}_p}\left(\frac{i}{p}\right)^2 = \left(\frac{-1}{p}\right)(p-1)$. Furthermore, $g_0 = 0$ by the fact that there are as many residues as non-residues in $\mathbb{F}_p$, so we must have $\sum_{i=0}^{p-1} a_i = 0$. Finally, for arbitrary a_n, the substitution $i = ni'$, $j = nj'$ gives $i' + j' \equiv 1 \pmod p$ and merely shuffles around the indices for the summands in the formula for a_n; hence $a_n = \sum_{i'+j'=1}\left(\frac{ni'}{p}\right)\left(\frac{nj'}{p}\right) = \sum_{i'+j'=1}\left(\frac{n^2}{p}\right)\left(\frac{i'}{p}\right)\left(\frac{j'}{p}\right) = \sum_{i'+j'=1}\left(\frac{i'}{p}\right)\left(\frac{j'}{p}\right)$. Therefore $a_n = a_1$ for all $1 \le n \le p$. Combining this with the fact that the a_i must sum to 0 gives $a_n = -\frac{a_0}{p-1} = -\left(\frac{-1}{p}\right)$. Thus $g^2 = \left(\frac{-1}{p}\right)\left[(p-1) - \zeta_p - \zeta_p^2 - \cdots - \zeta_p^{p-1}\right] = \left(\frac{-1}{p}\right)\left[(p-1) - (-1)\right] = \left(\frac{-1}{p}\right)p$, the desired result.

Now, we make use of the fact that $\left(\frac{\cdot}{p}\right)$ is *symmetric* when $p \equiv 1 \pmod 4$ and *antisymmetric* when $p \equiv 3 \pmod 4$; that is, $\left(\frac{a}{p}\right) = \left(\frac{a}{p}\right)$ when $p \equiv 1 \pmod 4$ and $\left(\frac{a}{p}\right) = -\left(\frac{a}{p}\right)$ when $p \equiv 3 \pmod 4$. This fact follows from Proposition **??** (Euler's Criterion) applied to $\left(\frac{-1}{p}\right)$, since $\frac{p-1}{2}$ is even if and only if $p \equiv 1 \pmod 4$. Then the summands of g can be paired up by $t \leftrightarrow -t$ in

$\mathbb{F}_p^\times$. When $p \equiv 1 \pmod 4$ the imaginary parts of each pair cancel, so g must be entirely real. Furthermore, g must be positive real by the even distribution of quadratic residues in $\mathbb{F}_p$, implying that $g = \sqrt{p}$. When $p \equiv 3 \pmod 4$, the same argument follows, only this time the paired terms' real parts cancel, yielding a sum along the positive imaginary axis; hence in this case $g = i\sqrt{p}$. $\square$

As a "bonus" application of Gauss sums (beyond the relation we will see they have with Jacobi sums), consider a proof of the law of quadratic reciprocity using quadratic Gauss sums:

Proposition 4.2. *Let p, q be two distinct, odd primes. Then $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.*

Proof. Let $g_p = g\left(\frac{\cdot}{p}\right) = \sum_{k=1}^{p-1}\left(\frac{k}{p}\right)\zeta_p^k$ as in Definition **??**. Denote $p^* = \left(\frac{-1}{p}\right)p$, which we have seen is equal to g_p^2. Then $g_p^{q-1} = (g_p^2)^{\frac{q-1}{2}} = \left[\left(\frac{-1}{p}\right)p\right]^{\frac{q-1}{2}} \equiv \left(\frac{-1}{p}\right)p \pmod p$, so $g_p^q \equiv \left(\frac{p^*}{q}\right)g_p \pmod p$. At the same time, the binomial theorem gives $g_p^q \equiv \sum_{k=1}^{p}\left(\frac{k}{p}\right)\zeta_p^{qk} \pmod q$ $= \left(\frac{q^{-1}}{p}\right)\sum_{t\in\mathbb{F}_p^\times}\left(\frac{t}{p}\right)\zeta_p^t = \left(\frac{q}{p}\right)g_p$. Thus $\left(\frac{q}{p}\right)\left(\frac{-1}{p}\right)p \equiv \left(\frac{p^*}{q}\right)p^* \pmod q$, implying that $\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right)$. Hence $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = \left(\frac{\left(\frac{-1}{p}\right)}{q}\right) = \left(\frac{-1}{p}\right)^{\frac{q-1}{2}} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$, and we have the law of quadratic reciprocity. $\square$

5 Jacobi Sums

Definition 5.1. *Let p be an odd prime. If χ and λ are Dirichlet characters on $\mathbb{F}_p$, then we denote the Jacobi sum of χ and λ by $J(\chi,\lambda) = \sum_{a+b=1}\chi(a)\lambda(b)$, where the sum is over all pairs $a, b \in \mathbb{F}_p$ such that $a + b = 1$.*

Theorem 5.1. *Suppose χ and λ are two multiplicative characters on $\mathbb{F}_p$ such that $\chi\lambda \neq \varepsilon$. Then $J(\chi,\lambda) = \dfrac{g(\chi)g(\lambda)}{g(\chi\lambda)}$.*

Proof. $g(\chi)g(\lambda) = \left(\sum_{a\in\mathbb{F}_p}\chi(a)\zeta_p^a\right)\left(\sum_{b\in\mathbb{F}_p}\lambda(b)\zeta_p^b\right) = \sum_{a,b\in\mathbb{F}_p}\chi(a)\lambda(b)\zeta_p^{a+b}$

$= \sum_{t\in\mathbb{F}_p}\left(\sum_{a+b=t}\chi(a)\lambda(b)\right)\zeta_p^t$. When $t = 0$, the inner sum is 0; when $t \neq 0$ let $x, y \in \mathbb{F}_p$ such that $a = tx$ and $b = ty$. Then when $a + b = t$, $x + y = 1$, so $\sum_{a+b=t}\chi(a)\lambda(b)$ can be rearranged as $\sum_{x+y=1}\chi(tx)\lambda(ty) = \chi\lambda(t)J(\chi,\lambda)$.

In total, we have $g(\chi)g(\lambda) = \sum_{t\in\mathbb{F}_p}\chi\lambda(t)J(\chi,\lambda)\zeta_p^t = J(\chi,\lambda)g(\chi\lambda)$. Since $\chi\lambda \neq \varepsilon$, $g(\chi\lambda) \neq 0$, so we can divide both sides by $g(\chi\lambda)$ to give the desired result. $\square$

Finally we can return to the original problem of counting solutions to diophantine equations, and see how Jacobi sums and similar structures arise in this computation.

For $x^2+y^2=0$, the number of solutions splits into $N(x^2+y^2=0)=\sum_{a+b=0} N(x^2=a)N(y^2=b)$. Substituting $b\mapsto b-1$ and using Proposition **??** gives $\sum_{a+b=1} N(x^2=a)N(y^2=b-1)=\sum_{a+b=1}\left[1+\left(\frac{a}{p}\right)\right]\left[1+\left(\frac{b-1}{p}\right)\right]=\sum_{a+b=1}\left[1+\left(\frac{a}{p}\right)+\left(\frac{b-1}{p}\right)+\left(\frac{a}{p}\right)\left(\frac{b-1}{p}\right)\right]=\sum_{a+b=1}1+\sum_{a+b=1}\left(\frac{a}{p}\right)+\sum_{a+b=1}\left(\frac{b-1}{p}\right)+\sum_{a+b=1}\left(\frac{a}{p}\right)\left(\frac{b-1}{p}\right)$. The second and third sums are simply summing over the Legendre symbol of all $t\in\mathbb{F}_p$, and are both therefore zero. The first sum is just p, since there are p pairs $a,b\in\mathbb{F}_p$ such that $a+b=1$, as choosing a subsequently fixes b, so the sum is over all $a\in\mathbb{F}_p$. The fourth sum reduces to $\left(\frac{-1}{p}\right)\sum_{a\in\mathbb{F}_p}\left(\frac{a}{p}\right)^2=(p-1)(-1)^{\frac{p-1}{2}}$; hence

$$N(x^2+y^2=0)=p+(p-1)(-1)^{\frac{p-1}{2}}=\begin{cases}2p-1 & p\equiv 1(\text{mod }4)\\ 1 & p\equiv 3(\text{mod }4).\end{cases}$$

A similar process yields $N(x^2+y^2=-1)=p+\left(\frac{-1}{p}\right)\sum_{a\in\mathbb{F}_p}\left(\frac{a}{p}\right)\left(\frac{a+1}{p}\right)$. To evaluate this sum, we split into the two cases modulo 4 and use the symmetry/antisymmetry relations. We can drop the cases $a=0$ and $a=-1$ from the sum, and note that $\left(\frac{a}{p}\right)\left(\frac{a+1}{p}\right)$ is 1 if and only if the two Legendre symbols are the same. We can form a (near, depending on the case modulo 4) bijection between pairs $(a,a+1)$ such that both Legendre symbols are -1 and those such that the first is -1 and the second is 1 by mapping $a\mapsto a^{-1}$ and applying the multiplicativity of the Legendre symbol. The total number of pairs in these sets will be the number of quadratic residues from 2 to $p-1$, which is $\frac{p-1}{2}-1$. A similar (near) bijection can be constructed between the other two sets of pairs of consecutive elements of $\mathbb{F}_p^\times$. In each case, there will always be one pair left over which is a "switch", and thus

$$N(x^2+y^2=-1)=p-\left(\frac{-1}{p}\right)=\begin{cases}p-1 & p\equiv 1(\text{mod }4)\\ p+1 & p\equiv 3(\text{mod }4).\end{cases}$$

Note that because maps such as $a\mapsto\left(\frac{a+1}{p}\right)$ are not technically characters, the sums above, while similar in structure, are not Jacobi sums. In general, it is more convenient to work with Jacobi sums since we can apply results such as Theorem **??**. In the above cases the substitutions for b which led to these components only serve to simplify the reduction of the problem when we eliminate b altogether; as such, they are not strictly necessary.

As another similar example, by the same process we have $N(x^2+y^2=1)=p+J\left(\left(\frac{\cdot}{p}\right),\left(\frac{\cdot}{p}\right)\right)$, which will yield the same result as $N(x^2+y^2=-1)$.

Continuing with the examples laid out at the beginning of the talk, we have for our simplest (to solve) equation

$$N((x+y)^2=-1)=pN(x^2=-1)=p\left[1+\left(\frac{-1}{p}\right)\right]=\begin{cases}2p & p\equiv 1(\text{mod } 4)\\ 0 & p\equiv 3(\text{mod } 4).\end{cases}$$

The reader can verify that all of these computations so far agree with the patterns we saw at the beginning.

Finally, the equation $N(y^2=x^3+1)$ presents a much more chaotic and difficult case, as can be seen in the table in section 2. Following a similar process to the above examples, we have $N(y^2=x^3+1)=\sum_{a+b=1}N(y^2=a)N(x^3=-b)=\sum_{a+b=1}\left[1+\left(\frac{a}{p}\right)\right]\left[1+\chi(-b)+\chi^2(-b)\right]=p+\sum_{a+b=1}\left(\frac{a}{p}\right)\chi(-b)+\sum_{a+b=1}\left(\frac{a}{p}\right)\chi^2(-b)=p+J\left(\left(\frac{\cdot}{p}\right),\chi\right)+J\left(\left(\frac{\cdot}{p}\right),\chi^2\right)$, with all other sums vanishing. Combining some of the previous results about Jacobi sums, we get the definite bound $|p-N(y^2=x^3+1)|<2\sqrt{p}$. Because the number of solutions includes two Jacobi sums, each has a different "pull" on $N(y^2=x^3+1)$, resulting in a more chaotic-seeming pattern in relation to p, as we saw from the initial table.