Proof of Stickelberger's Congruence Via Jacobi Sums

relation between Gauss sums and Jacobi sums in order to introduce the factorials of the base p digits into Stickelberger's congruence in (essentially) one step:

LEMMA 1. *If* $\chi_1, \ldots, \chi_r$ *are multiplicative characters on* $\mathbf{F}_q$ *with nontrivial product* $\chi_1 \cdot \ldots \cdot \chi_r$, *then*

$$G(\chi_1 \cdot \ldots \cdot \chi_r) = \frac{G(\chi_1) \cdot \ldots \cdot G(\chi_r)}{J(\chi_1, \ldots, \chi_r)} .$$

Proof. See [6, Chapter 8, Theorem 3], noting that our weaker hypotheses than those of [6] are sufficient since we assume the trivial character vanishes at 0. □

PROOF OF STICKELBERGER'S CONGRUENCE VIA JACOBI SUMS

For $\chi_1, \ldots, \chi_r$ multiplicative characters on $\mathbf{F}_q = \mathbf{Z}[\zeta_{q-1}]/\mathfrak{p}$, it is easy to check that

$$J(\chi_1, \ldots, \chi_r)^p \equiv J(\chi_1, \ldots, \chi_r) \bmod \mathfrak{p} ,$$

so $J(\chi_1, \ldots, \chi_r) \equiv$ rational integer mod $\mathfrak{p}$. We will show below (Theorem 2) that when some χ_i is nontrivial, as an integer representative one can take a certain r-fold multinomial coefficient.

In the case $r = 2$ there is the following classical congruence: if $0 \leqslant k_1, k_2 < q - 1$ and not both k_1, k_2 are zero, then

$$J(\omega_{\mathfrak{p}}^{-k_1}, \omega_{\mathfrak{p}}^{-k_2}) \equiv \frac{(k_1 + k_2)!}{k_1! k_2!} \bmod \mathfrak{p} .$$

References for this congruence are given in the Notes in [6, Chapter 14]. We shall extend this congruence to Jacobi sums of any number of multiplicative characters of $\mathbf{F}_q$ as follows:

THEOREM 2. *For* $r \geqslant 1$ *and* $0 \leqslant k_1, \ldots, k_r < q - 1$ *with some* $k_j > 0$,

$$J(\omega_{\mathfrak{p}}^{-k_1}, \ldots, \omega_{\mathfrak{p}}^{-k_r}) \equiv \frac{(k_1 + \cdots + k_r)!}{k_1! \cdot \ldots \cdot k_r!} \bmod \mathfrak{p} .$$

The simplicity of the statement of this generalization makes it somewhat surprising that it does not seem to appear in the literature (such as that which is mentioned in the Notes in [8, Chapter 5]).

In our proofs of Theorems 1 and 2, we will view multinomial coefficients as special values of polynomials. For $t \geqslant 1$ and $n_1, \ldots, n_t \in \mathbf{N}$, define

$$\binom{X}{n_1, \ldots, n_t} = \frac{X(X-1) \cdot \ldots \cdot (X - n_1 - \cdots - n_t + 1)}{n_1! \cdot \ldots \cdot n_t!}.$$

In particular, $\binom{X}{0, \ldots, 0} = 1$.

When $t = 1$, this reduces (even in notation) to the binomial coefficient polynomial, so whereas many people would write (for $r \geqslant 2$ and $n_1, \ldots, n_r \in \mathbf{N}$)

$$\frac{(n_1 + \cdots + n_r)!}{n_1! \cdot \ldots \cdot n_r!}$$

as $\binom{n_1 + \cdots + n_r}{n_1, \ldots, n_r}$, we write it as $\binom{n_1 + \cdots + n_r}{n_1, \ldots, n_{r-1}}$; having one less integer in the bottom is convenient, as for binomial coefficients. The main advantage of this notation is that in $\mathbf{Z}[[X_1, \ldots, X_t]]$ one has

$$(1 + X_1 + \cdots + X_t)^m = \sum_{n_1, \ldots, n_t \geqslant 0} \binom{m}{n_1, \ldots, n_t} X_1^{n_1} \cdot \ldots \cdot X_t^{n_t}$$

for all integers m.

Although the following two multinomial coefficient congruences are rather general, they will each be used only once, and in special cases.

C1. *For* $t \geqslant 1$, *choose* $n_1, \ldots, n_t \in \mathbf{N}$ *and* $d \in \mathbf{N}$ *with each* $n_i < p^d$. *For* $b \in \mathbf{Z}$,

$$\binom{b + p^d}{n_1, \ldots, n_t} \equiv \binom{b}{n_1, \ldots, n_t} \bmod p.$$

C2. *For* $d \geqslant 0$, $t \geqslant 1$, *and* $m_0, \ldots, m_t \geqslant 0$ *write*

$m_0 = c_0 + c_1 p + \cdots + c_d p^d$, $0 \leqslant c_i \leqslant p - 1$ *for* $i < d$;

$m_j = c_{0j} + c_{1j} p + \cdots + c_{dj} p^d$, $0 \leqslant c_{ij} \leqslant p - 1$ *for* $i < d$ *and* $1 \leqslant j \leqslant t$,

where $c_d, c_{dj} \geqslant 0$. *Then*

$$\binom{m_0}{m_1, \ldots, m_t} \equiv \binom{c_0}{c_{01}, \ldots, c_{0t}} \cdot \ldots \cdot \binom{c_d}{c_{d1}, \ldots, c_{dt}} \bmod p.$$

To prove C1, work in $\mathbf{F}_p[[X_1, \ldots, X_t]]$ and use the equation

$$(1 + X_1 + \cdots + X_t)^{b + p^d} = (1 + X_1 + \cdots + X_t)^b (1 + X_1^{p^d} + \cdots + X_t^{p^d}).$$

To prove C2, the condition on the leading "digits" $c_d, c_{d1}, \ldots, c_{dt}$ just being nonnegative reduces the proof to the case $d = 1$. Now look at the coefficient of $X_1^{m_1} \cdot \ldots \cdot X_t^{m_t}$ on both sides of the equation

$$(1 + X_1 + \cdots + X_t)^{m_0} = (1 + X_1 + \cdots + X_t)^{c_0}(1 + X_1^p + \cdots + X_t^p)^{c_1}$$

in $\mathbf{F}_p[X_1, \ldots, X_t]$. In the binomial case ($t = 1$), C2 is originally due to Lucas [9], and is also in [4]. The general result ($t > 1$) is due to Dickson [2, p. 76].

Proof of Theorem 2. For any χ, $J(\chi) = 1$, so we can assume $r > 1$. Since some $k_j > 0$ and a Jacobi sum is a symmetric function of its arguments, we choose $k_r > 0$. We will let $\alpha_1, \ldots, \alpha_{r-1}$ each run independently through representatives for the nonzero classes of $\mathbf{F}_q = \mathbf{Z}[\zeta_{q-1}]/\mathfrak{p}$, say the complex roots of $X^{q-1} - 1$. For s in $\mathbf{Z}$, $\omega_{\mathfrak{p}}^s(\alpha) \equiv \alpha^s \bmod \mathfrak{p}$ if $\alpha \not\equiv 0 \bmod \mathfrak{p}$ or $s \geqslant 0$ (we set $0^0 = 1$), so

$$\begin{aligned}
&J(\omega_{\mathfrak{p}}^{-k_1}, \ldots, \omega_{\mathfrak{p}}^{-k_r}) \\
&= (-1)^{r-1} \sum_{\alpha_j} \omega_{\mathfrak{p}}^{-k_1}(\alpha_1) \cdot \ldots \cdot \omega_{\mathfrak{p}}^{-k_{r-1}}(\alpha_{r-1})\omega_{\mathfrak{p}}(1 - \alpha_1 - \cdots - \alpha_{r-1})^{q-1-k_r} \\
&\equiv (-1)^{r-1} \sum_{\alpha_j} \alpha_1^{-k_1} \cdot \ldots \cdot \alpha_{r-1}^{-k_{r-1}}(1 - \alpha_1 - \cdots - \alpha_{r-1})^{q-1-k_r} \bmod \mathfrak{p} \\
&\equiv \sum_{\substack{n_j \geqslant 0 \\ n_1 + \cdots + n_{r-1} \leqslant q-1-k_r}} \binom{q-1-k_r}{n_1, \ldots, n_{r-1}} (-1)^{r-1+n_1+\cdots+n_{r-1}} \prod_{1 \leqslant i \leqslant r-1} \Big(\sum_{\alpha_i} \alpha_i^{n_i - k_i}\Big).
\end{aligned}$$

The only time $\sum_{\alpha_i} \alpha_i^{n_i - k_i}$ isn't zero is when $(q-1) \mid (n_i - k_i)$, when the sum is $q - 1 \equiv -1 \bmod \mathfrak{p}$. From $0 \leqslant k_i < q - 1$ and

$$-(q-1) < -k_i \leqslant n_i - k_i \leqslant n_i \leqslant q - 1 - k_r < q - 1,$$

we see that $(q-1) \mid (n_i - k_i)$ if and only if $n_i = k_i$. Thus, if $k_1 + \cdots + k_{r-1} > q - 1 - k_r$, we have $J(\omega_{\mathfrak{p}}^{-k_1}, \ldots, \omega_{\mathfrak{p}}^{-k_r}) \equiv 0 \bmod \mathfrak{p}$, while if $k_1 + \cdots + k_{r-1} \leqslant q - 1 - k_r$,

$$\begin{aligned}
J(\omega_{\mathfrak{p}}^{-k_1}, \ldots, \omega_{\mathfrak{p}}^{-k_r}) &\equiv \binom{q-1-k_r}{k_1, \ldots, k_{r-1}} (-1)^{r-1+k_1+\cdots+k_{r-1}}(-1)^{r-1} \bmod \mathfrak{p} \\
&= \binom{q-1-k_r}{k_1, \ldots, k_{r-1}} (-1)^{k_1+\cdots+k_{r-1}} \\
&= \binom{k_1 + \cdots + k_r - q}{k_1, \ldots, k_{r-1}}.
\end{aligned}$$

If $k_1 + \cdots + k_{r-1} > q - 1 - k_r$, this last expression equals 0, so regardless of the value of $k_1 + \cdots + k_{r-1}$, we have by C1 that

$$J(\omega_\mathfrak{p}^{-k_1}, \ldots, \omega_\mathfrak{p}^{-k_r}) \equiv \binom{k_1 + \cdots + k_r}{k_1, \ldots, k_{r-1}} \bmod \mathfrak{p}$$

$$= \frac{(k_1 + \cdots + k_r)!}{k_1! \cdot \ldots \cdot k_r!}. \qquad \square$$

Remarks. 1. Theorem 2 is not true in general when all $k_j = 0$, since the Jacobi sum of the trivial character on $\mathbf{F}_q$ taken r times is $(1 - (1-q)^r)/q \equiv r \bmod p$.

2. It is reasonable to ask if Theorem 2 can be proven in general if it is just known for $r = 2$. After all, there are recursion formulas relating a multinomial coefficient to a product of binomial coefficients and a Jacobi sum of several characters to a product of Jacobi sums of two characters. However, this latter relation depends on hypotheses of nontriviality of certain characters which are not part of the hypotheses of Theorem 2 (for example, $J(\chi_1, \chi_2, \chi_3) = J(\chi_1, \chi_2) J(\chi_1\chi_2, \chi_3)$ precisely when $\chi_1\chi_2$ is nontrivial). Thus it would likely be cumbersome to use this approach to prove Theorem 2.

Proof of Theorem 1. It is obvious for $a = 0$, and see [11, pp. 96-97] for the case $a = 1$ (whose proof shows why one should expect the theorem to hold for positive powers of $\omega_\mathfrak{p}^{-1}$, not of $\omega_\mathfrak{p}$: $p^f - 1 = \#\mathbf{F}_q^\times$ is more closely related to $p^d - 1$ than to $p^d + 1$). Now we may assume $q > 3$. For $0 < a < q - 2$, we have by Lemma 1 that

$$G(\omega_\mathfrak{p}^{-(a+1)}) = \frac{G(\omega_\mathfrak{p}^{-a})\, G(\omega_\mathfrak{p}^{-1})}{J(\omega_\mathfrak{p}^{-a}, \omega_\mathfrak{p}^{-1})},$$

and $J(\omega_\mathfrak{p}^{-a}, \omega_\mathfrak{p}^{-1}) \equiv a + 1 \bmod \mathfrak{p}$ (hence also mod $\mathfrak{P}$) by Theorem 2, so by induction and the equation $\mathrm{ord}_\mathfrak{P}(\zeta_p - 1) = 1$,

$$G(\omega_\mathfrak{p}^{-a}) \equiv \frac{(\zeta_p - 1)^a}{a!} \bmod \mathfrak{P}^{a+1}$$

for $0 \leqslant a \leqslant p - 1$ (or $a < p - 1$ if $q = p$). If $q = p$ we're done, so assume $q > p$, i.e. $f \geqslant 2$. Going from $a = p - 1$ to $a = p$ is a problem because $\mathfrak{P} \mid p$ and we don't want to divide by p in our congruence modulo a power of $\mathfrak{P}$. We circumvent this with Jacobi sums.

For $1 \leqslant a < q - 1$, some digit a_i is > 0, so $\omega_\mathfrak{p}^{-a}, \omega_\mathfrak{p}^{-a_i p^i}$ are nontrivial.

Then by Lemma 1,

$$G(\omega_{\mathfrak{p}}^{-a}) = G(\omega_{\mathfrak{p}}^{-a_0} \cdot \ldots \cdot \omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}})$$
$$= \frac{G(\omega_{\mathfrak{p}}^{-a_0}) \cdot \ldots \cdot G(\omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}})}{J(\omega_{\mathfrak{p}}^{-a_0}, \ldots, \omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}})}$$
$$= \frac{G(\omega_{\mathfrak{p}}^{-a_0}) \cdot \ldots \cdot G(\omega_{\mathfrak{p}}^{-a_{f-1}})}{J(\omega_{\mathfrak{p}}^{-a_0}, \ldots, \omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}})},$$

the last equation holding since $G(\chi^p) = G(\chi)$ (see [7, p. 5]).

Since $\operatorname{ord}_{\mathfrak{P}}(a_i!) = 0$,

$$G(\omega_{\mathfrak{p}}^{-a_0}) \cdot \ldots \cdot G(\omega_{\mathfrak{p}}^{-a_{f-1}}) \equiv \frac{(\zeta_p - 1)^{a_0 + \cdots + a_{f-1}}}{a_0! \cdot \ldots \cdot a_{f-1}!} \bmod \mathfrak{P}^{a_0 + \cdots + a_{f-1} + 1}.$$

By Theorem 2 and C2,

$$J(\omega_{\mathfrak{p}}^{-a_0}, \ldots, \omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}}) \equiv \binom{a_0 + \cdots + a_{f-1}p^{f-1}}{a_0, \ldots, a_{f-2}p^{f-2}} \bmod \mathfrak{p}$$
$$\equiv \binom{a_0}{a_0, 0, \ldots, 0} \binom{a_1}{0, a_1, \ldots, 0} \cdot \ldots \cdot \binom{a_{f-1}}{0, \ldots, 0}$$
$$= 1.$$

Therefore

$$J(\omega_{\mathfrak{p}}^{-a_0}, \ldots, \omega_{\mathfrak{p}}^{-a_{f-1}p^{f-1}}) \equiv 1 \bmod \mathfrak{P},$$

so we are done. □

Our method of proof shows that writing Stickelberger's congruence as

$$G(\omega_{\mathfrak{p}}^{-a}) \equiv \prod_{0 \leqslant i \leqslant f-1} \frac{(\zeta_p - 1)^{a_i}}{a_i!} \bmod \mathfrak{P}^{a_0 + \cdots + a_{f-1} + 1}$$

isolates terms in analogy with Lemma 1. This gives a new explanation for the appearance of base p digits in the denominator in Stickelberger's congruence. There are more sophisticated explanations, cf. the proof of Stickelberger's congruence via the Gross-Koblitz formula in [7, Chapter 15]. (Although both the original proof of the Gross-Koblitz formula in [5] and the proof in [7] are only done for finite fields of odd characteristic, the formula is also valid for characteristic 2 since Lemma 1.1 (ii) in [7, p. 333] is valid for all $\delta > 0$, not just for $\delta \geqslant 1/(p-1)$. Alternatively, in [1] Coleman gives a simple proof

which he explicitly points out is valid in all characteristics. Thus a proof of Stickelberger's congruence for all finite fields via the Gross-Koblitz formula is justified.)

Proof of Jacobi Sum Congruence Via Stickelberger

We now want to show that not only does Theorem 1 follow from Theorem 2, but Theorem 2 follows from Theorem 1, so the two theorems are equivalent. Some preliminary results will be required before the (tedious) proof is presented.

For $n \in \mathbf{N}$, write

$$n = c_0 + c_1 p + \cdots + c_d p^d, \ 0 \leqslant c_i \leqslant p - 1 .$$

From [3, Chapter IX],

$$\operatorname{ord}_p(n!) = \frac{n - (c_0 + \cdots + c_d)}{p - 1}, \ \frac{n!}{(-p)^{\operatorname{ord}_p(n!)}} \equiv c_0! \cdot \ldots \cdot c_d! \bmod p .$$

Note neither equation requires $c_d \neq 0$. We define

$$S_p(n) \stackrel{\text{def}}{=} c_0 + \cdots + c_d , \quad H_p(n) \stackrel{\text{def}}{=} c_0! \cdot \ldots \cdot c_d! ,$$

and note neither of these definitions requires $c_d \neq 0$. One sees easily that for any $n \in \mathbf{N}$, $n \equiv S_p(n) \bmod p - 1$, and for $n_1, \ldots, n_t \in \mathbf{N}$,

$$\operatorname{ord}_p \left(\frac{(n_1 + \cdots + n_t)!}{n_1! \cdot \ldots \cdot n_t!} \right) = \frac{S_p(n_1) + \cdots + S_p(n_t) - S_p(n_1 + \cdots + n_t)}{p - 1} .$$

For $x \in \mathbf{R}$, let $\langle x \rangle$ denote the fractional part of x. For $b \in \mathbf{Z}$, let $b \equiv b' \bmod q - 1$ where $0 \leqslant b' < q - 1$, so that $\left\langle \frac{b}{q-1} \right\rangle = \frac{b'}{q-1}$. Define

$$s_q(b) = S_p(b'), \quad h_q(b) = H_p(b') ,$$

so s_q and h_q are just the extensions of S_p and H_p from $\{b : 0 \leqslant b < q - 1\}$ by $(q - 1)$-periodicity. From [7, p. 10],

$$s_q(b) = (p - 1) \sum_{0 \leqslant i \leqslant f - 1} \left\langle \frac{p^i b}{q - 1} \right\rangle .$$

Since $\operatorname{ord}_{\mathfrak{P}}(\zeta_p - 1) = 1$, Stickelberger's congruence can be written for all a in $\mathbf{Z}$ as

$$\frac{G(\omega_{\mathfrak{p}}^{-a})}{(\zeta_p - 1)^{s_q(a)}} \equiv \frac{1}{h_q(a)} \bmod \mathfrak{P} .$$