WHENCE GAUSS SUMS?

Let $p$ be an odd prime, let $(\cdot/p)$ be the Legendre symbol, and let $\zeta_p = e^{2\pi i/p}$. Typically in a first number theory course the quadratic Gauss sum

$$\tau = \sum_{a=1}^{p-1} \left( \frac{a}{p} \right) \zeta_p^a$$

is pulled out of thin air, and its properties established by elementary calculations that appear to work for no discernible reason. More generally, for any Dirichlet character modulo $p$,

$$\chi : (\mathbb{Z}/p\mathbb{Z})^\times \longrightarrow \mathbb{C}^\times,$$

the corresponding Gauss sum

$$\tau(\chi) = \sum_{a=1}^{p-1} \chi(a) \zeta_p^a$$

satisfies many of the same properties. A person might wonder just what is going on and how anybody might conceive of such a thing. This writeup shows that the Gauss sum is a special case of a general symmetrizing device, the Lagrange resolvent, that has built-in equivariance and equation-solving properties that are easier to understand in general than in the confusingly overly-specific context of Gauss sums alone.

First we place the Gauss sum in the context of appropriate fields. The $p$th cyclotomic field is

$$K = \mathbb{Q}(\zeta), \quad \zeta = e^{2\pi i/p}.$$  

Also introduce the auxiliary field

$$F = \mathbb{Q}(\omega), \quad \omega = e^{2\pi i/(p-1)}$$

and the composite field

$$L = FK = \mathbb{Q}(\omega, \zeta).$$

Thus any Dirichlet character modulo $p$ in fact maps into $F^\times$,

$$\chi : (\mathbb{Z}/p\mathbb{Z})^\times \longrightarrow F^\times,$$

and so the corresponding Gauss sum lies in the composite field,

$$\tau(\chi) = \sum_{a\in(\mathbb{Z}/p\mathbb{Z})^\times} \chi(a) \zeta^a \in L.$$  

Next we work quite generally. Let $L/F$ be a Galois field extension with cyclic Galois group $G$. (If the characteristic is nonzero then assume that the order of $G$ is coprime to it.) Consider two data, an element of the larger field and a character of the Galois group into the multiplicative group of the smaller one,

$$\theta \in L, \quad \chi : G \longrightarrow F^\times.$$  

The Lagrange resolvent associated to $\theta$ and $\chi$ is the $\chi$-weighted average over the Galois orbit of $\theta$,

$$R = R(\theta, \chi) = \sum_{g \in G} \chi(g)\theta(g) \in L.$$ 

Since $R$ is a weighted average and since the character-outputs are fixed by the Galois group, the equivariance property of the Lagrange resolvent is immediate: for any $g \in G$,

$$g(R) = g(\sum_{\tilde{g}} \chi(\tilde{g})\tilde{g}(\theta)) = \sum_{\tilde{g}} \chi(\tilde{g})(g\tilde{g})(\theta) = \chi(g^{-1})\sum_{\tilde{g}} \chi(g\tilde{g})(\tilde{g})(\theta)) = \chi(g^{-1})R.$$ 

Consequently, letting $d = |\text{Gal}(L/F)|$,

$$g(R^d) = (g(R))^d = (\chi(g^{-1})R)^d = R^d$$

showing that $R^d$ lies in the smaller field $F$. Indeed, letting $m$ denote the order of $\chi$, this argument shows that $R(\theta, \chi)^m \in F$. However, the matter of finding a method to express $R(\theta, \chi)^m$ as an element of $F$ is context-specific.

As for the equation-solving properties of the Lagrange resolvent, begin by noting that the group of characters $\chi$ of the finite cyclic Galois group $G$ is again finite cyclic of the same order. Assume now that $F$ is large enough to contain the range of all such characters. Fix generators $g$ of the Galois group and $\chi$ of the character group. The expression of each Lagrange resolvent as a linear combination of the Galois orbit of $\theta$ encodes as an equality of column vectors in $L^d$ (with $d = |G|$ as before),

$$\begin{bmatrix} R(\theta, \chi^0) \\ R(\theta, \chi^1) \\ \vdots \\ R(\theta, \chi^{d-1}) \end{bmatrix} = V_\chi \begin{bmatrix} g^0(\theta) \\ g^1(\theta) \\ \vdots \\ g^{d-1}(\theta) \end{bmatrix},$$

where the matrix relating the vectors is the Vandermonde matrix,

$$V_\chi = \begin{bmatrix} \chi^0(g^0) & \chi^0(g^1) & \cdots & \chi^0(g^{d-1}) \\ \chi^1(g^0) & \chi^1(g^1) & \cdots & \chi^1(g^{d-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi^{d-1}(g^0) & \chi^{d-1}(g^1) & \cdots & \chi^{d-1}(g^{d-1}) \end{bmatrix} \in F^{d \times d}.$$ 

The top row and the left column of $V_\chi$ are all 1’s. As a very small case of Fourier analysis, orthogonality shows that the inverse of the Vandermonde matrix is essentially the transpose of another one,

$$V_\chi^T V_\chi = dI_d.$$ 

Thus we can invert the equality of column vectors in $L^d$ to solve for $\theta$ and its conjugates in terms of the resolvents,

$$\begin{bmatrix} g^0(\theta) \\ g^1(\theta) \\ \vdots \\ g^{d-1}(\theta) \end{bmatrix} = \frac{1}{d} V_\chi^T R(\theta, \chi^1) \begin{bmatrix} R(\theta, \chi^0) \\ R(\theta, \chi^1) \\ \vdots \\ R(\theta, \chi^{d-1}) \end{bmatrix}. $$

Especially, equate the top entries to see that \( \theta \) itself is the average of its resolvents,

\[
\theta = \frac{1}{d} \sum_{i=0}^{d-1} R(\theta, \chi^i).
\]

Since each resolvent is a \( d \)th root over \( F \), this expresses \( \theta \) in radicals.

Finally, to see that the Lagrange resolvent subsumes Gauss sums, specialize the environment back to \( F = \mathbb{Q}(\omega) \) (with \( \omega = e^{2\pi i/(p-1)} \)) and \( L = FK \) where \( K = \mathbb{Q}(\zeta) \) (with \( \zeta = e^{2\pi i/p} \)). Then \( \text{Gal}(L/F) \approx (\mathbb{Z}/p\mathbb{Z})^\times \), the automorphisms being

\[
g_a : \zeta \mapsto \zeta^a, \quad a \in (\mathbb{Z}/p\mathbb{Z})^\times.
\]

Also specializing the top-field element \( \theta \) to \( \zeta \), the Lagrange resolvent is indeed the Gauss sum if we view any character \( \chi : G \rightarrow F^\times \) as a character of \((\mathbb{Z}/p\mathbb{Z})^\times\) as well,

\[
R(\zeta, \chi) = \sum_{g \in G} \chi(g) g(\zeta) = \sum_{a \in (\mathbb{Z}/p\mathbb{Z})^\times} \chi(a) \zeta^a = \tau(\chi).
\]

The general reasoning has shown that if \( \chi \) has order \( d \) then \( \tau(\chi)^d \) lies in \( F \), and that \( \zeta \) can be expressed as an average of Gauss sums \( \tau(\chi) \). Since the order of each \( \chi \) divides \( p-1 \), this constructs \( \zeta \) from numbers whose \((p-1)\)st powers are rational numbers. While \( \zeta \) it has the rational power \( \zeta^p = 1 \), this power is higher than \( p-1 \). And while \( \zeta \) satisfies a polynomial of degree \( p-1 \), that polynomial does not take the form \( X^{p-1} - a \).

In particular, if \( p \) is a Fermat prime \( p = 2^n + 1 \) (where \( n = 2^e \) in turn) then the Gauss sums all satisfy \( \tau^{2^n} = 1 \) and so plausibly they can be constructed in turn by successions of square roots.