On a generalization of Jacobi sums

By Takashi ONO

To the memory of Takuro Shintani

1. Let K be a finite field with q elements: $K=\boldsymbol{F}_q$. Denote by $K^\times$ the multiplicative group of K. We extend, as usual, the domain of definition of a character χ of $K^\times$ to all of K by setting $\chi(0)=1$ if $\chi=1$, the trivial character, and $\chi(0)=0$ if $\chi\neq 1$. For characters χ, χ' of $K^\times$, the Jacobi sum is defined by

$$J(\chi, \chi')=\sum_{x\in K}\chi(x)\chi'(1-x). \tag{1.1}$$

When χ, χ', $\chi\chi'$ are all $\neq 1$, we have the equality

$$|J(\chi, \chi')|=\sqrt{q}. \tag{1.2}$$

This property of Jacobi sum is used to estimate the number of solutions in K of the equation of type

$$y^d=1-x^n. \tag{1.3}$$

The purpose of this paper is to generalize the definition (1.1) and the property (1.2) so that, among other things, we can estimate the number of solutions in K of the equation of type

$$y^d=x^a(1-tx^n)^b, \qquad t\in K^\times, \tag{1.4}$$

on the elementary level. Our proof of a generalization ((3.4) Theorem) of (1.2) does not use the additive character of K and so does not depend on the estimation of the Gauss sum as in the usual proof of (1.2).

2. Let A be a finite abelian group and K be the finite field with q elements. By a K-character of A, we shall mean a homomorphism of A into $K^\times$. Let ξ be a K-character of A. Let α be a character of A and β be a character of $K^\times$ in the ordinary sense. Consider the sum

$$J_\xi(\alpha, \beta; t)=\sum_{x\in A}\alpha(x)\beta(1-t\xi(x)), \qquad t\in K. \tag{2.1}$$

If $A=K^\times$, $\xi(x)=x$, $t=1$ and $\alpha\neq 1$, then (2.1) coincides with the Jacobi sum (1.1). (When $\alpha=1$ here, there is a slight discrepancy between (1.1) and (2.1) because $\alpha(0)=1$.) From the definition (2.1), we see easily that

(2.2) $$J_\xi(\alpha,\ \beta\,;\,t\xi(x))=\bar{\alpha}(x)J_\xi(\alpha,\ \beta\,;\,t)\,,\qquad x\in A\,,$$

where $\bar{\alpha}$ is the character of A which is the complex conjugate of α. It follows immediately from (2.2) that

(2.3) $$|J_\xi(\alpha,\ \beta\,;\,t\xi(x))|=|J_\xi(\alpha,\ \beta\,;\,t)|.$$

This means that the absolute value of (2.1) may be considered as a function on the cokernel: $\mathrm{Cok}\,\xi=K^\times/\mathrm{Im}\,\xi$. If, in particular, $\alpha(\mathrm{Ker}\,\xi)=1$, the sum

(2.4) $$J_\xi^*(\alpha,\ \beta\,;\,t)=\sum_{x\in A/\mathrm{Ker}\,\xi}\alpha(x)\beta(1-t\xi(x))$$

makes sense and we have

(2.5) $$J_\xi(\alpha,\ \beta\,;\,t)=[\mathrm{Ker}\,\xi]J_\xi^*(\alpha,\ \beta\,;\,t)\,,$$

where we write $[X]$ for the cardinality of a set X. Finally, in the general case, we put

(2.6) $$\sigma_\xi(\alpha,\ \beta)=\sum_{t\in K}|J_\xi(\alpha,\ \beta\,;\,t)|^2\,.$$

In the sequel, we shall often use the Kronecker delta $\delta_{x,y}=\delta(x,\ y)$ for elements $x,\ y$ of a set, in an obvious way. For example, we have

(2.7) $$J_\xi(\alpha,\ \beta\,;\,0)=[A]\delta_{\alpha,1}\,.$$

In view of (2.3), we can also write (2.6) as follows:

(2.8) $$\sigma_\xi(\alpha,\ \beta)=[A]^2\delta_{\alpha,1}+[\mathrm{Im}\,\xi]\sum_{t\in\mathrm{Cok}\,\xi}|J_\xi(\alpha,\ \beta\,;\,t)|^2\,.$$

If, in particular, $\alpha(\mathrm{Ker}\,\xi)=1$, we have, from (2.5),

(2.9) $$\sigma_\xi(\alpha,\ \beta)=[A]([A]\delta_{\alpha,1}+[\mathrm{Ker}\,\xi]\sum_{t\in\mathrm{Cok}\,\xi}|J_\xi^*(\alpha,\ \beta\,;\,t)|^2)\,,$$

since $[A]=[\mathrm{Ker}\,\xi][\mathrm{Im}\,\xi]$.

3. Now, we shall compute $\sigma_\xi(\alpha,\ \beta)$ by changing the order of summation. We begin with

(3.1) LEMMA. *Let χ be a non-trivial character of $K^\times$ and a, b be elements of $K^\times$. Then, we have*

$$s_{a,b}=\sum_{x\in K}\chi(1-ax)\bar{\chi}(1-bx)=q\delta_{a,b}-\chi(a)\bar{\chi}(b)\,.$$

PROOF. When $a=b$, we have $s_{a,a}=\sum_{x\neq a^{-1}}|\chi(1-ax)|^2=q-1$. When $a\neq b$, we have $s_{a,b}=\sum_{x\neq a^{-1},b^{-1}}\chi((1-ax)(1-bx)^{-1})$. Put $y=(1-ax)(1-bx)^{-1}$. Since $y=0,\ \infty,$

ab^{-1} correspond to $x=a^{-1}$, b^{-1}, ∞, respectively, under this transformation, we have $s_{a,b}=\sum_{y\neq 0, ab^{-1}}\chi(y)=-\chi(ab^{-1})$, q. e. d.

From now on, we assume that $\beta\neq 1$ since the case $\beta=1$ is trivial. Using the Lemma, the computation of $\sigma_{\xi}(\alpha, \beta)$ goes as follows:

$$\begin{aligned}\sigma_{\xi}(\alpha, \beta)&=\sum_{t\in K}\sum_{x,y\in A}\alpha(x)\beta(1-t\xi(x))\bar{\alpha}(y)\bar{\beta}(1-t\xi(y))\\ &=\sum_{x,y\in A}\alpha(x)\bar{\alpha}(y)\sum_{t\in K}\beta(1-t\xi(x))\bar{\beta}(1-t\xi(y))\\ &=\sum_{x,y\in A}\alpha(x)\bar{\alpha}(y)(q\delta(\xi(x), \xi(y))-\beta(\xi(x))\bar{\beta}(\xi(y)))\\ &=-\sum_{x,y\in A}\alpha(x)\beta(\xi(x))\bar{\alpha}(y)\bar{\beta}(\xi(y))+q\sum_{\xi(xy^{-1})=1}\alpha(xy^{-1})\\ &=-|\sum_{x\in A}\alpha(\beta\circ\xi)(x)|^2+q[A]\sum_{\xi(z)=1}\alpha(z)\\ &=-[A]^2\delta(\alpha(\beta\circ\xi), 1)+q[A][\mathrm{Ker}\,\xi]\delta(\alpha(\mathrm{Ker}\,\xi), 1)\,.\end{aligned}$$

Since $\alpha(\beta\circ\xi)=1$ implies $\alpha(\mathrm{Ker}\,\xi)=1$, we get the following

(3.2) THEOREM. *When* $\beta\neq 1$, *we have*

$$\sigma_{\xi}(\alpha, \beta)=\delta(\alpha(\mathrm{Ker}\,\xi), 1)(q[A][\mathrm{Ker}\,\xi]-[A]^2\delta(\alpha(\beta\circ\xi), 1))\,.$$

The definition (2.6) and (3.2) give:

(3.3) THEOREM. *If* $\beta\neq 1$ *and* $\alpha(\mathrm{Ker}\,\xi)\neq 1$, *then*

$$J_{\xi}(\alpha, \beta; t)=0 \qquad \textit{for all } t\in K.$$

Combining (2.9) with (3.2), we get:

(3.4) THEOREM. *If* $\beta\neq 1$ *and* $\alpha(\mathrm{Ker}\,\xi)=1$, *we have*

$$q=[\mathrm{Im}\,\xi](\delta(\alpha(\beta\circ\xi), 1)+\delta_{\alpha,1})+\sum_{t\in\mathrm{Cok}\,\xi}|J_{\xi}^{*}(\alpha, \beta; t)|^2\,.$$

(3.5) THEOREM. *If* $\beta\neq 1$, *then* $|J_{\xi}(\alpha, \beta; t)|\leqq[\mathrm{Ker}\,\xi]\sqrt{q}$, $t\in K^{\times}$.

This follows from (2.5), (3.3) and (3.4).

(3.6) REMARK. When $A=K^{\times}$, $\xi(x)=x$, $t=1$, $\alpha\neq 1$, $\beta\neq 1$, $\alpha\beta\neq 1$, we have $\mathrm{Ker}\,\xi=1$, $\mathrm{Cok}\,\xi=1$ and hence $q=|J_{\xi}^{*}(\alpha, \beta; 1)|^2=|J(\alpha, \beta)|^2$. Therefore, (3.4) generalizes the classical formula (1.2). Note that here we did not use, as in the usual proof of (1.2), the relation $J(\alpha, \beta)G(\alpha\beta)=G(\alpha)G(\beta)$ and the estimation of the Gauss sum $G(\alpha)=\sum_{x\in K}\alpha(x)\psi(x)$, ψ being a fixed additive character $\neq 1$ of K.

(3.7) REMARK. When $K=\boldsymbol{F}_q$, q: odd, $A=K^\times$, $\alpha=\beta=\chi=$the character of order 2 and $\xi(x)=x^2$, we have $J_\xi(\alpha, \beta; t)=\sum_{x\in K^\times}\chi(x(1-tx^2))$ and $\alpha(\mathrm{Ker}\,\xi)=1$ if and only if $q\equiv 1 \pmod{4}$. We also have $[\mathrm{Cok}\,\xi]=[K^\times:(K^\times)^2]=2$ and $\alpha(\beta\circ\xi)=\chi\chi^2=\chi\neq 1$. Hence the equality in (3.4) becomes $q=A^2+B^2$ with $A=J_\xi^*(\alpha, \beta; 1)$, $B=J_\xi^*(\alpha, \beta; w)$, $w\in K^\times-(K^\times)^2$. This is essentially the formula of E. Jacobsthal [2]. (See also Chowla [1], Chapter IV.) In many cases, (3.4) provides explicit expressions of numbers as sum of certain number of squares. However, it does not seem to provide a constructive proof of the Lagrange's theorem: any natural number is a sum of 4 squares.

4. Some examples. Before giving the applica tion of above theorems to the estimation of number of solutions of equations over $K=\boldsymbol{F}_q$, we want to insert here some examples which are obtained directly from the theorems.

(4.1) *Example.* Let $\mathfrak{o}$ be the ring of integers of an algebraic number field, $\mathfrak{m}$ be an ideal of $\mathfrak{o}$ and $\mathfrak{p}$ be a prime factor of $\mathfrak{m}$. Put $A=(\mathfrak{o}/\mathfrak{m})^\times$, the group of invertible elements of the ring $\mathfrak{o}/\mathfrak{m}$ and $K=\mathfrak{o}/\mathfrak{p}=\boldsymbol{F}_q$, $q=N\mathfrak{p}$. Call ξ the natural K-character $A\to K^\times$. For non-trivial characters α, β of A, $K^\times$, respectively, we have the sum $J_\xi(\alpha, \beta)=\sum_{x\in A}\alpha(x)\beta(1-\xi(x))$ which coincides with the classical Jacobi sum when $\mathfrak{m}=\mathfrak{p}$. Since ξ is surjective, we have $[\mathrm{Cok}\,\xi]=1$, $[\mathrm{Ker}\,\xi]=[A]/[K^\times]=\varphi(\mathfrak{m})(q-1)^{-1}$. If $\alpha(\mathrm{Ker}\,\xi)\neq 1$, we have $J_\xi(\alpha, \beta)=0$ by (3.3). If $\alpha(\mathrm{Ker}\,\xi)=1$, (3.4) gives

$$q=(q-1)\delta(\alpha(\beta\circ\xi), 1)+[\mathrm{Ker}\,\xi]^{-2}|J_\xi(\alpha, \beta)|^2$$

and hence

$$|J_\xi(\alpha, \beta)|=\begin{cases}\varphi(\mathfrak{m})(q-1)^{-1} & \text{if } \alpha(\beta\circ\xi)=1\,,\\ \varphi(\mathfrak{m})(q-1)^{-1}\sqrt{q} & \text{if } \alpha(\beta\circ\xi)\neq 1\,.\end{cases}$$

(4.2) *Example.* Let $\zeta\in\boldsymbol{C}$ be a primitive m-th root of 1 and $F=\boldsymbol{Q}(\zeta)$ be the cyclotomic field. Let $\mathfrak{p}$ be any prime ideal of the ring $\mathfrak{o}$ of integers of F prime to m and $q=N\mathfrak{p}$. Put $K=\mathfrak{o}/\mathfrak{p}=\boldsymbol{F}_q$. Let A be the cyclic group of order m generated by ζ and ξ be the K-character of A obtained by reducing numbers in A modulo $\mathfrak{p}$. Since $\mathfrak{p}$ is prime to m, we have $[\mathrm{Ker}\,\xi]=1$ and $[\mathrm{Cok}\,\xi]=(q-1)/m$. Therefore, from (3.5), we have

$$|J_\xi(\alpha, \beta; t)|\leqq\sqrt{q}, \qquad t\in K^\times.$$

Since $\alpha(\zeta)=\zeta^a$ for some $a\in\boldsymbol{Z}$, we can also write this as

$$\left|\sum_{i=1}^{m-1}\zeta^{ai}\beta(1-t\zeta^i)\right|\leqq\sqrt{q}, \qquad t\in\mathfrak{o}-\mathfrak{p},$$

where β is any non-trivial character of $(\mathfrak{o}/\mathfrak{p})^\times$.

5. Let A be a finite abelian group, K be the finite field with q elements and ω, ξ be K-characters of A. Let b, d be positive integers such that $q \equiv 1 \pmod{d}$ and $(b, d)=1$. Consider a function $f: A \rightarrow K$ defined by

$$f(x)=\omega(x)(1-t\xi(x))^{b}, \qquad t \in K^{\times}. \tag{5.1}$$

Put

$$E=\{(x, y) \in A \times K;\ y^{d}=f(x)\}. \tag{5.2}$$

Then we have

$$[E]=\sum_{\chi^{d}=1} \sum_{x \in A} \chi(f(x)), \tag{5.3}$$

where χ runs over all characters of $K^{\times}$ of exponent d. From (5.3) we have

$$|[E]-[A]| \leqq \sum_{\chi^{d}=1, \chi \neq 1} \left| \sum_{x \in A} \chi(f(x)) \right|. \tag{5.4}$$

Now, as we have $\chi(f(x))=\chi \circ \omega(x) \chi^{b}(1-t\xi(x))$, we get

$$J_{\xi}(\alpha, \beta; t)=\sum_{x \in A} \chi(f(x)), \tag{5.5}$$

with $\alpha=\chi \circ \omega$, $\beta=\chi^{b}$. Since $\beta \neq 1$, from (3.5), (5.4) and (5.5), it follows that

$$|[E]-[A]| \leqq (d-1)[\operatorname{Ker} \xi] \sqrt{q}. \tag{5.6}$$

Consider now the equation

$$y^{d}=f(x)=x^{a}(1-tx^{n})^{b}, \qquad t \in K^{\times}, \tag{5.7}$$

where a, b, d, n are positive integers such that $q \equiv 1 \pmod{d}$ and $(b, d)=1$. Put $A=K^{\times}$, $\omega(x)=x^{a}$, $\xi(x)=x^{n}$. Then, we have $f(x)=\omega(x)(1-t\xi(x))^{b}$. Call N the number of solutions (x, y) of (5.7) in $K \times K$. We have $N=[E]+1$ since $(0, 0)$ is the only solution of $y^{d}=f(x)$ outside E. Notice that $[E]-[A]=N-q$. Furthermore, we have $[\operatorname{Ker} \xi]=(n, q-1) \leqq n$. Hence we have

$$|N-q| \leqq (d-1) n \sqrt{q}. \tag{5.8}$$

If we assume that $(n, q)=1$, then, since $(d, b)=1$, the polynomial $Y^{d}-f(X) \in K[X, Y]$ becomes absolutely irreducible and the number m of distinct zeros of $f(x)=0$ in $\bar{K}$ is $n+1$. Hence, in this case, we can also write (5.8) as

$$|N-q| \leqq (d-1)(m-1) \sqrt{q}, \tag{5.9}$$

which fits the general theorem for curves over K. (See p. 43 (Theorem 2C) and p. 80 of Schmidt [3].)

References

[1] Chowla, S., Riemann Hypothesis and Hilbert's Tenth Problem, New York, Gordon & Breach, Science Publishers, Inc., 1965.
[2] Jacobsthal, E., Über die Darstellungen der Primzahlen der Form $4n+1$ als Summe zweier Quadrate, J. Reine Angew. Math. **132** (1907), 238-245.
[3] Schmidt, W. M., Equations over Finite Fields, an Elementary Approach, Lecture Notes in Math. Vol. 536, Springer, 1976.
[4] Weil, A., Numbers of solutions of equations in finite fields, Bull. Amer. Math. Soc. **55** (1949), 497-508.

(Received May 9, 1981)

Department of Mathematics
The Johns Hopkins University
Baltimore, Maryland, 21218
U. S. A.