ON KUMMER'S MEMOIR OF 1857 CONCERNING FERMAT'S LAST THEOREM*

BY H. S. VANDIVER

1. *Introduction.* In a previous paper under the same title,† the writer considered an article by Kummer,‡ and pointed out that the argument there used for proving certain results regarding the equation

$$x^\lambda + y^\lambda + z^\lambda = 0, \tag{1}$$

where x, y and z are integers and λ is an odd prime, is deficient and incorrect in several respects. Kummer attempts to prove four theorems which in my first paper were numbered I to IV. I pointed out that the proofs of Theorems I and IV are incomplete, and that the proofs of Theorems II and III are inaccurate. In the present paper additions to and modifications of Kummer's arguments will be given, by means of which the demonstrations Theorems I and IV will be completed.

2. *Proof of Theorem* I. Assume that h_1 is the first factor of the class number of the field $\Omega(\alpha)$, α being a primitive λth root of unity,

$$E_\nu(\alpha) = \prod_{k=0}^{\mu-1} e(\alpha^{\gamma^k})^{\gamma^{-2k\nu}},$$

$$e(\alpha) = \sqrt{\frac{(1-\alpha^\gamma)(1-\alpha^{-\gamma})}{(1-\alpha)(1-\alpha^{-1})}},$$

where γ is a primitive root of λ, and $\mu = (\lambda - 1)/2$; then Kummer's Theorem I may be stated as follows.

If h_1 is divisible by λ but not by λ^2, then one and only one Bernoulli number B_ν in the set B_i, $i = 1, 2, \cdots, \mu - 1$, is divisible by λ. If under this assumption we also have $h_2 \equiv 0$ (mod λ), h_2 being the second factor of the class number of $\Omega(\alpha)$, then $E_\nu(\alpha)$ is the λth power of a unit in $\Omega(\alpha)$.

---

* Presented to the Society October 30, 1920.

† PROCEEDINGS OF THE NATIONAL ACADEMY, vol. 6, No. 5, pp. 266–69 (May, 1920).

‡ MATHEMATISCHE ABHANDLUNGEN OF THE BERLIN ACADEMY, 1857, pp. 41–74.

To complete Kummer's proof it is necessary to prove only the first statement in the theorem. Assume $\lambda > 5$. In another paper* the writer has given the relation

$$h_1 \equiv \prod_s \frac{\lambda(-1)^{(s\lambda^2-1)/2} B_{(s\lambda^2+1)/2}}{2^{(\lambda-3)/2}} \pmod{\lambda^2}, \quad (s = 1, 3, \cdots, \lambda - 2). \tag{2}$$

We shall now show that the assumption that two or more of the B's in the set B_i, $i = 1, 2, \cdots, \mu - 1$, are divisible by λ leads to the relation $h_1 \equiv 0 \pmod{\lambda^2}$ which will yield the proof desired, as it is known† that if $h_1 \equiv 0 \pmod{\lambda}$ then at least one of the B's is divisible by λ.

Assume that $B_a \equiv B_b \equiv 0 \pmod{\lambda}$, where a and b are each included in the set $1, 2, \cdots, \mu - 1$. Kummer‡ has shown that

$$\frac{B_c}{c} \equiv (-1)^{k\mu} \frac{B_{c+k\mu}}{c + k\mu} \pmod{\lambda},$$

where k is an integer and c is not a multiple of μ, and where $\lambda > 3$. This gives

$$(-1)^{s\mu(\lambda+1)} \frac{B_{(s\lambda^2+1)/2}}{(s\lambda^2 + 1)/2} \equiv \frac{B_{(s+1)/2}}{(s + 1)/2} \pmod{\lambda},$$

and if $(s + 1)/2 = a$, then s is included in the set $1, 3, 5, \cdots \lambda - 2$, and the above congruence gives

$$B_{(s\lambda^2+1)/2} \equiv 0 \pmod{\lambda}.$$

Similarly

$$B_{(s_1\lambda^2+1)/2} \equiv 0 \pmod{\lambda},$$

where $(s_1 + 1)/2 = b$. Applying these relations to (2), we obtain $h_1 \equiv 0 \pmod{\lambda^2}$, and Theorem I is proved, as the remainder of the proof as set forth by Kummer, is, in my opinion, rigorous.

3. *Proof of Theorem* IV. We proceed to Kummer's fourth theorem:

THEOREM IV. *If h_1 is divisible by λ, but not by λ^2, and P is an ideal in $\Omega(\alpha)$, such that P^λ is the principal ideal $(F(\alpha))$,*

* This BULLETIN, vol. 25, p. 460, relation (8) for $a = 2$.

† Vandiver, ibid., p. 461.

‡ CRELLE, vol. 41, p. 368.

$B_\nu \equiv 0 \pmod{\lambda}$, $\nu < \mu$, *then* P *is or is not a principal ideal according as*

$$\frac{d_0^{\lambda-2\nu} \log F(e^v)}{dv^{\lambda-2\nu}} \equiv 0 \text{ or } \not\equiv 0 \pmod{\lambda}.$$

In this statement, the notation $d_0^{\lambda-2\nu}$ means that the function $F(e^v)$ is to be differentiated $\lambda - 2\nu$ times with respect to v, and zero substituted for v in the result. The letter e denotes the Napierian base.

In my first note two criticisms of Kummer's proof of this theorem were made. I shall here modify his argument so as to dispose of the difficulties in question, and consequently complete the demonstration. The first of my criticisms referred to formula (A) on page 53 of Kummer's memoir. The number $\psi_r(\alpha)$ which appears there is defined in the eleventh line from the bottom of the page as the product of certain ideal factors, but this decomposition is proved to hold only for the case where $\psi_r(\alpha)$ contains ideals of degree not higher than the first. In another paper,* Kummer gives the corresponding formula for the generalized function $\psi_r(\alpha)$ which contains ideals of higher degree, as follows:

$$(3) \quad \text{ind } E_n(\alpha) \equiv \frac{\gamma^{2n} - 1}{2(1 + r^{\lambda-2n} - (r+1)^{\lambda-2n})} \cdot \frac{d_0^{\lambda-2n} \log \psi_r(e^v)}{dv^{\lambda-2n}} \pmod{\lambda},$$

where r is an integer, $1 < r < \lambda - 1$, $E_n(\alpha)$ and γ are defined as before, and ind $(E_n(\alpha))$ is i in the relation

$$(E_n(\alpha))^{(q^t-1)/\lambda} \equiv \alpha^i \pmod{\mathfrak{P}},$$

$\mathfrak{P}$ being an ideal prime factor of the odd prime q, and t the exponent to which q belongs modulo λ. Further

$$\psi_r(\alpha) = \alpha^{-(r+1)h + \text{ind}(g^h+1)},$$

where g is a primitive root of $\mathfrak{P}$ such that $g^{(q^t-1)/\lambda} \equiv \alpha \pmod{\mathfrak{P}}$, h ranges over the integers $0, 1, 2, \cdots, q^t - 2$, excepting $(q^t - 1)/2$, ind $(g^h + 1)$ being defined as i in the relation

* CRELLE, vol. 44, p. 125.

$(g^h + 1) \equiv g^i \pmod{\mathfrak{P}}$. Although Kummer in defining q did not state that it was odd, his work is subject to that restriction, since his function $\psi_r(\alpha)$, as defined by him, has no meaning for $q = 2$, since in that case $(q^t - 1)/2$ is not integral. If, however, we take the function as defined by H. H. Mitchell * for the case $q = 2$, namely

$$\psi_r(\alpha) = \alpha^{-(r+1)h+\mathrm{ind}(g^h+1)}$$

where h ranges over the integers $1, 2, \cdots, 2^t - 2$, the formula (3) will also hold for this case, as Kummer's argument can be used without change except that it is necessary to note in proving formulas such as the following (loc. cit., middle of page 125)

$$(g^i - 1)(g^i - g^\lambda)(g^i - g^{2\lambda}) \cdots (g^i - g^{(\nu-1)\lambda}) \equiv 1 - g^{\nu i} \pmod{\mathfrak{P}},$$

that since $1 \equiv -1 \pmod{q}$ for $q = 2$,

$$g^{\nu i} - 1 \equiv 1 - g^{\nu i} \pmod{p}.$$

Kummer (loc. cit., page 120) gives the decomposition of the general $\psi_r(\alpha)$ into prime ideal factors. Instead of using this decomposition we shall examine the form of it given by Mitchell † and express the factorization of $\psi_r(\alpha)$, t arbitrary, in a form analogous to that given by Kummer for the case $t = 1$. Mitchell ‡ considers the Galois field of order q^t, where q is any prime, and $q^t - 1 = \lambda\nu$. Let the elements of this field other than zero each be represented as a power of a primitive root, and σ_i denote any element whose index is congruent to i, modulo λ. The symbol $\overset{j}{m}_i$ stands for the number of solutions of the relation $1 + \sigma_i = \sigma_j$ in the field. He then defines the function (page 167, relation 4),

$$\psi_{-a,-b}(\alpha) = \sum_{i,j} \overset{j}{m}_i \alpha^{-bi+(a+b)j},$$

where i and j each range over the integers $0, 1, \cdots, \lambda - 1$, and $a \not\equiv 0$, $b \not\equiv 0$, $a + b \not\equiv 0 \pmod{\lambda}$. For $b = -1$, $a = -r$,

* TRANSACTIONS OF THIS SOCIETY, vol. 17 (1916), p. 165.

† Loc. cit., p. 168.

‡ Loc. cit., p. 166, § 2.

this becomes

$$\psi_{r,1}(\alpha) = \sum_{i,\ j}^{i\ \ j} m_i \alpha^{-(r+1)j+i}.$$

Since there are $m_j = m_i$ numbers $h \equiv j$ and ind $(g^h + 1) \equiv i \pmod{\lambda}$, the preceding relation may be written in the form

$$\psi_r(\alpha) = \Sigma \alpha^{-(r+1)h+\text{ind}(g^h+1)}.$$

Now apply Mitchell's first theorem (loc. cit., page 173) to the particular function $\psi_r(\alpha)$. We have in this case $t = t_1$, where q belongs to the exponent t_1, modulo λ. We conclude that the number of times the ideal $\mathfrak{P}_i$, where this symbol designates the ideal obtained from $\mathfrak{P}$ by the substitution (α^k/α), $ki \equiv 1 \pmod{\lambda}$, is equal to the number of the expressions $|-riq^{t-j}|+|-iq^{t-j}|$, $j = 0, 1, \cdots, t-1$, whose values exceed λ, the integer i assuming $(\lambda - 1)/t$ values prime to λ such that the quotient of no two of them is congruent modulo λ to a power of q, and the symbol $|x|$ denoting the least positive residue of x, modulo λ. Now if we select the integers l in the set $1, 2, \cdots, \lambda - 1$, which satisfy the relation $|-rl| + |-l| > \lambda$, we may show from what precedes that

$$\prod_c \mathfrak{P}_c = \psi_r(\alpha), \tag{4}$$

where c ranges over the integers which satisfy $cl \equiv 1 \pmod{\lambda}$. For, the integers $1, 2, \cdots, \lambda - 1$, are congruent modulo λ to the integers

$$\begin{array}{llll} i_1, & i_1q, & i_1q^2, & \cdots, \ i_1q^{t-1}, \\ i_2, & i_2q, & i_2q^2, & \cdots, \ i_2q^{t-1}, \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ i_s, & i_sq, & i_sq^2, & \cdots, \ i_sq^{t-1}, \end{array}$$

in some order, where $s = (\lambda - 1)/t$, and $i_1, i_2, \cdots, i_s$ are the integers less than λ such that the quotient of no two of them is congruent, modulo λ, to a power of q, since this definition of the i's shows that no two elements of the above array are congruent modulo λ. According to the first conclusion regarding factors of $\psi_r(\alpha)$, we may write

$$\psi_r(\alpha) = \prod_a \mathfrak{P}_a^{d_a}, \tag{5}$$

where a ranges over the integers $i_1, i_2, \cdots, i_s$, and d_a is the number of expressions

$$|-ri_aq^{t-j}|+|-i_aq^{t-j}|, \quad (j=0, 1, 2, \cdots, t-1),$$

whose values exceed λ. Since $\mathfrak{P}(\alpha)=\mathfrak{P}(\alpha^{q^\sigma})$, a an integer, the ideals $\mathfrak{P}(i_nq)$, $\mathfrak{P}(i_nq^2)$, $\cdots$, $\mathfrak{P}(i_nq^{t-1})$, where $\mathfrak{P}(\alpha^f)=\mathfrak{P}(f)$ are all equal, and the decomposition (5) is identical with the the decomposition (4), which was to be shown. Using Kummer's notation, the relation

$$|-ri_eq^{t-j}|+|-i_eq^{t-j}|>\lambda$$

may be written in the form

$$\gamma_{\mu-h}+\gamma_{\mu-h+\text{ind}r}>\lambda,$$

where γ_b is the least positive integer satisfying $\gamma_b\equiv\gamma^b \pmod{\lambda}$ and ind r is defined by $\gamma^{\text{ind}r}\equiv r \pmod{\lambda}$. In view of the above we may now use the relation*

$$\psi_r(\alpha)^{H_1\lambda}=\pm\alpha^d\prod_h F(\alpha^{\gamma^h}) \tag{5a}$$

where $H_1\lambda$ is the class number of $\Omega(\alpha)$, H_1 prime to λ.

Now, as pointed out in my first article, Kummer employs (page 54), without proof or reference the relation

$$\frac{d_0^{m\lambda^r}\log\varphi(e^v)}{dv^{m\lambda^r}}\equiv\frac{d_0^{m\lambda^r}\log\varphi_1(e^v)}{dv^{m\lambda^r}} \pmod{\lambda^{r+1}},$$

m not being a multiple of $\lambda-1$, and $\varphi(\alpha)\equiv\varphi_1(\alpha) \pmod{\lambda^{r+1}}$. We shall prove a special case of this relation and use of it will enable us to complete the proof of Theorem IV. Assume that $\varphi(x)$ and $\varphi_1(x)$ are two integral algebraic functions of x with rational integral coefficients, such that $\varphi(1)=\varphi_1(1)$ and $\varphi(\alpha)=\varphi_1(\alpha)$. Further let $\varphi(\alpha)$ be prime to λ, whence it follows that $\varphi(1)\not\equiv 0 \pmod{\lambda}$. Under these conditions, it will be shown that

$$\frac{d_0^{m\lambda}\log\varphi(e^v)}{dv^{m\lambda}}\equiv\frac{d_0^{m\lambda}\log\varphi_1(e^v)}{dv^{m\lambda}} \pmod{\lambda^2}. \tag{6}$$

* Kummer gives this relation in a form not containing the factor α^d, but this appears to be an error. Compare the reasoning in the writer's paper, ANNALS OF MATHEMATICS, (2), vol. 21, p. 74, on the determination of $e(\alpha)$ in relation (6). This does not affect Kummer's later arguments, however.

We have

$$\varphi(e^v) = \varphi_1(e^v) + VW_1$$

where $V = (e^{vp} - 1)/(e^v - 1)$ and W_1 is an integral function of e^v. If we divide VW_1 by $e^{vp} - 1$, we have a remainder which must be of the form cV where c is an integer. We then write

$$\varphi(e^v) = \varphi_1(e^v) + W(e^{vp} - 1) + cV.$$

Putting $v = 0$, we have

$$\varphi(1) = \varphi_1(1) + cp,$$

and by hypothesis $\varphi(1) = \varphi_1(1)$, so that $c = 0$. Hence

$$\varphi(e^v) = \varphi_1(e^v) + W(e^{vp} - 1)$$

and

$$\frac{\varphi(e^v)}{\varphi_1(e^v)} = 1 + \frac{W(e^{vp} - 1)}{\varphi_1(e^v)} = U, \text{ say.}$$

We then have

$$(7) \quad \frac{d^{m\lambda} \log \varphi(e^v)}{dv^{m\lambda}} - \frac{d^{m\lambda} \log \varphi_1(e^v)}{dv^{m\lambda}}$$
$$= U^{-1} \frac{d^{m\lambda} U}{dv^{m\lambda}} + (m\lambda - 1) \frac{d^{m\lambda-1} U}{dv^{m\lambda-1}} \cdot \frac{d(U^{-1})}{dv}$$
$$+ \cdots + \frac{dU}{dv} \cdot \frac{d^{m\lambda-1}(U^{-1})}{dv^{m\lambda-1}}.$$

Now every derivative of U is divisible by λ, if zero is substituted for v. Hence every term of the right-hand member of the above relation consists of integers or fractions whose denominators are prime to λ (since $\varphi_1(1) \not\equiv 0 \pmod{\lambda}$) and whose numerators are divisible by λ^2, excepting possibly the first term. To examine this term, we write $X = W/\varphi_1(e^v)$, and, therefore

$$\frac{d^{m\lambda} U}{dv^{m\lambda}} = (e^{vp} - 1) \frac{d^{m\lambda} X}{dv^{m\lambda}} + m\lambda \frac{d(e^{v\lambda} - 1)}{dv} \cdot \frac{d^{m\lambda-1} X}{dv^{m\lambda-1}}$$
$$+ \cdots + X \frac{d^{m\lambda}(e^{v\lambda} - 1)}{dv^{m\lambda}}.$$

Since

$$\frac{d_0{}^k(e^{v\lambda} - 1)}{dv^k} \equiv 0 \pmod{\lambda^2},$$

for $k > 1$, it follows from these relations that

$$\frac{d_0{}^{m\lambda}U}{dv^{m\lambda}} \equiv 0 \pmod{\lambda^2};$$

and by using this with (7) the congruence (6) is obtained.

From (5a) we have

$$\psi_r(\alpha)^{H_1\lambda(\lambda-1)} = \alpha^{d(\lambda-1)}\prod_h F(\alpha^{\gamma^h})^{\lambda-1}, \tag{5b}$$

and we may write $F(\alpha^{\gamma^h})^{\lambda-1} = 1 + \omega(1-\alpha) = G(\alpha^{\gamma^h})$,

where ω is an integer in $\Omega(\alpha)$, since if

$$F(\alpha^{\gamma^h}) = a + \omega_1(1-\alpha),$$

where a is a rational integer, then

$$\begin{aligned} F(\alpha^{\gamma^h})^{\lambda-1} &= a^{\lambda-1} + \omega_2(1-\alpha) \\ &= 1 + \lambda k + \omega_2(1-\alpha) \\ &= 1 + \omega(1-\alpha), \end{aligned}$$

after using the known relation $\lambda = \epsilon(1-\alpha)^{\lambda-1}$ where ϵ is a unit in $\Omega(\alpha)$. The relation (5b) then becomes

$$\psi_r(\alpha)^{H_1\lambda(\lambda-1)} = \alpha^{d(\lambda-1)}\prod_h G(\alpha^{\gamma^h}), \tag{5c}$$

whence

$$\psi_r(1)^{H_1\lambda(\lambda-1)} = \prod_h G(1) = 1,$$

Since it is known that $\psi_r(1) = 1$. Hence we can apply (6) to (5c), and proceeding as Kummer did on pages 54–57 of his article, we obtain

$$\text{ind } E_\nu(\alpha) \equiv -\frac{(-1)^{\lambda+\mu}(\gamma^{2\nu}-1)B_{\nu\lambda-\mu}}{2H_1\lambda} \cdot \frac{d_0{}^{\lambda-2\nu}lG(e^v)}{dv^{\lambda-2\nu}},$$

modulo λ. By definition of G we also have

$$\frac{d_0{}^{\lambda-2\nu}lG(e^v)}{dv^{\lambda-2\nu}} \equiv -\frac{d_0{}^{\lambda-2\nu}lF(e^v)}{dv^{\lambda-2\nu}} \pmod{\lambda},$$

and the last two relations give the congruence obtained by Kummer (loc. cit., page 57). The remainder of his argument, up to page 61, appears to be correct, if we use the Dedekind definition of ideal. This establishes the theorem which I have designated IV.

CORNELL UNIVERSITY