BERNOULLI NUMBERS

JORDAN SCHETTLER

Department of Mathematics
The University of Arizona
P.O. Box 210089, 617 N. Santa Rita
Tucson, AZ 85721-0089, USA

1. Introduction

The Bernoulli numbers are a sequence of rational numbers with many interesting arithmetic properties. The appearances of Bernoulli numbers throughout mathematics are abundant and include finding a formula for the sum of the mth powers of the first n positive integers, values of L-functions, Euler-Macluarin summation formulas, and special cases of Fermat's Last Theorem. The denominators of the Bernoulli numbers are well understood, but the numerators are quite mysterious and the object of notable conjectures. We'll first review some elementary results, look at special values of L-functions, explain the connection to class numbers of cyclotomic fields, and finally explore two of the most celebrated theorems about Bernoulli numbers. The main reference here is the book [10] by Washington which is used throughout.

2. Basic Properties

The Bernoulli numbers $B_0, B_1, \ldots$ are defined by

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!},$$

so, in particular, $B_n \in \mathbb{Q}$ for all n and since

$$\frac{t}{e^t - 1} + \frac{t}{2} = \frac{t(1 + e^t)}{2(e^t - 1)}$$

is an even function, we have $B_1 = -1/2$, while $B_{2n+1} = 0$ for all $n \in \mathbb{N}$. We also have $B_0 = 1$ and for $n \geq 1$ the recurrence relation

$$B_n = \frac{-1}{n+1} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k \tag{1}$$

which can be seen by noting that

$$1 = \frac{t}{e^t - 1} \cdot \frac{e^t - 1}{t} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} \sum_{n=0}^{\infty} \frac{t^n}{(n+1)!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{B_k}{k!} \cdot \frac{1}{(n-k+1)!} \right) t^n.$$

In fact, the similarity of (1) to the binomial theorem can be used to define the Bernoulli numbers in a cute way as mentioned in [2]; i.e., one can use the formal relation

$$B^n = (B+1)^n$$

and then change exponents to subscripts after expanding the right hand side. We also define the Bernoulli polynomials

$$B_n(x) := \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k}$$

which are easily seen to satisfy

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!}; \tag{2}$$

moreover, fixing k and using (2) we find that

$$\begin{aligned} t\sum_{n=0}^{\infty} \frac{B_{n+1}(k) - B_{n+1}(0)}{n+1} \cdot \frac{t^n}{n!} &= \sum_{n=0}^{\infty} (B_n(k) - B_n(0)) \frac{t^n}{n!} = \frac{te^{xt}}{e^t-1} - \frac{t}{e^t-1} \\ &= \frac{t(e^{xt}-1)}{e^t-1} = t\sum_{j=0}^{k-1} e^{tj} = t\sum_{j=0}^{\infty} \left(\sum_{j=0}^{k-1} j^n \right) \frac{t^n}{n!}, \end{aligned}$$

so

$$1^n + 2^n + \cdots + k^n = \frac{B_{n+1}(k+1) - B_{n+1}(0)}{n+1}. \tag{3}$$

In fact, the formula (3) for the sum of the nth powers of the first k positive integers was the original impetus behind Jacob Bernoulli's introduction of the numbers named for him as noted in [5]. One can also show (3) by first proving

$$B_n(x+1) - B_n(x) = nx^{n-1}$$

and then telescoping, which is an easy exercise in [5].

In [9], it is commented that there are essentially only two facts we know about Bernoulli numbers themselves. One of these is contained in the following theorem.

Theorem 2.1. *Suppose $p \equiv 3 \pmod 4$ is prime. Then*

$$B_{\frac{p+1}{2}} \not\equiv 0 \pmod p.$$

Exercise 5.9 in [10] suggests how to give an analytic proof of (2.1) for sufficiently large p by using a theorem of Brauer and Siegel which in this case implies

$$\ln(h_{\mathbb{Q}(\sqrt{-p})}) \sim \ln(\sqrt{p}).$$

It is an exercise in [3] to prove for $d = -p^{-1} \bmod a$ that

$$\frac{B_{2m}}{2m}(a^{2m} - 1) \equiv \sum_{j=0}^{p-1} j^{2m-1}(dj \bmod a) \pmod p$$

from which it it follows that

$$\frac{B_{2m}}{2m}(2^{-2m} - 1) \equiv \frac{1}{2} \sum_{j=0}^{(p-1)/2} j^{2m-1} \pmod p$$

which can in turn be used to prove (2.1) for all primes $p \equiv 3 \pmod 4$ by setting $m = (p+1)/4$.

The other fact, which is the basis for a great deal of the study of Bernoulli numbers and their applications, is the subject of the following subsection.

2.1. **The von Staudt-Clausen Theorem.** The following theorem completely characterizes the denominators of Bernoulli numbers with even index.

Theorem 2.2 (von Staudt-Clausen)**.** *For each $n \in \mathbb{N}$ we have*

$$A_{2n} := B_{2n} + \sum_{\substack{p \text{ prime} \\ p-1|2n}} \frac{1}{p} \in \mathbb{Z}.$$

Sketch. Let q be an arbitrary prime. It suffices to show A_{2n} is in the ring of q-integers $\mathbb{Z}_{(q)} = \{x \in \mathbb{Q} | \mathrm{ord}_q(x) \geq 0\}$ since the intersection of all rings of q-integers is $\mathbb{Z}$. Using induction and formula (3) one can show $qB_{2n} \in \mathbb{Z}_{(q)}$ and

$$qB_{2n} \equiv 1^{2n} + 2^{2n} + \cdots + (q-1)^{2n} \pmod q.$$

If $q - 1|2n$, then

$$qB_{2n} \equiv 1 + 1 + \cdots + 1 = q - 1 \equiv -1 \pmod q,$$

so

$$\mathrm{ord}_q\left(B_{2n} + \frac{1}{q}\right) = \mathrm{ord}_q(qB_n + 1) - 1 \geq 0,$$

giving $A_{2n} \in \mathbb{Z}_{(q)}$. If $q - 1 \nmid 2n$, then for a primitive root g of p

$$qB_{2n} \equiv g^{1\cdot 2n} + g^{2\cdot 2n} + \cdots + g^{(q-1)\cdot 2n} = \frac{g^{2n(q-1)} - 1}{g^{2n} - 1} \equiv 0 \pmod q,$$

so

$$\mathrm{ord}_q(B_{2n}) = \mathrm{ord}_q(qB_{2n}) - 1 \geq 1 - 1 = 0,$$

again giving $A_{2n} \in \mathbb{Z}_{(q)}$. □

Corollary 2.3. *Let $n \in \mathbb{N}$ and write $B_{2n} = U_{2n}/V_{2n}$ with $U_{2n}, V_{2n} \in \mathbb{Z}, V_{2n} > 0$, and $(U_{2n}, V_{2n}) = 1$. Then*

$$V_{2n} = \prod_{\substack{p \text{ prime} \\ p-1|2n}} p$$

is squarefree and divisible by 6.

In fact, there are infinitely many n such that $V_{2n} = 6$. In particular, by Dirichlet's theorem on primes in arithmetic progression there are infinitely many primes of the form $p = 3m+1$, and for such a p the only divisors of $2p$ are $1, 2, p, 2p$, but $p+1$ isn't prime since p is an odd prime and $2p + 1 = 2(3m + 1) + 1 = 3(2m + 1)$ isn't prime since $m > 0$, so $V_{2p} = 6$.

There are quite a few analogues and generalizations of the von Stuadt-Clausen theorem. For example, Carlitz proved the following generalization of (2.2) in [2] for H-series (i.e., power series with rational coefficients) of the form

$$f(t) = \sum_{n=1}^{\infty} \frac{a_n}{n!} t^n$$

such that

$$f'(t) = \sum_{n=0}^{\infty} A_n f^n(t) \tag{4}$$

where $a_1 = 1 = A_0$ and $A_n \in \mathbb{Z}$ for all n.

Theorem 2.4. *Suppose $f(t)$ is an H-series satisfying (4) and set*

$$\frac{t}{f(t)} = \sum_{n=0}^{\infty} \frac{\beta_n}{n!} t^n.$$

Then for all $n \in \mathbb{N}$

$$\beta_{2n} + \sum_{\substack{p \text{ prime} \\ p-1 | 2n}} \frac{1}{p} e_p^{2n/(p-1)} \in \mathbb{Z}$$

where

$$t = \sum_{n=1}^{\infty} \frac{e_n}{n} f(t)^n.$$

3. L-functions and Kummer's Congruence's

3.1. **Dirichlet L-functions.** For a Dirichlet character $\chi : \mathbb{Z} \to \mathbb{C}$ with conductor f there are generalized Bernoulli numbers

$$B_{n,\chi} := f^{n-1} \sum_{k=1}^{f} \chi(k) B_n\left(\frac{k}{f}\right)$$

which satisfy

$$\sum_{k=1}^{f} \frac{\chi(k) t e^{kt}}{e^{ft} - 1} = \sum_{n=0}^{\infty} B_{n,\chi} \frac{t^n}{n!}, \tag{5}$$

so for the trivial character $\chi = 1$ we get back $B_{n,1} = B_n$ when $n > 1$.

Define the L-series attached to the Dirichlet character χ of conductor f by

$$L(s, \chi) = \sum_{n-1}^{\infty} \frac{\chi(n)}{n^s}$$

for $\operatorname{Re}(s) > 1$. Then

$$L(s, \chi) = \sum_{k=1}^{f} \chi(k) f^{-s} \zeta\left(s, \frac{k}{f}\right)$$

where for $b \in (0, 1]$

$$\zeta(s, b) := \sum_{n=0}^{\infty} \frac{1}{(b+n)^s}$$

is the Hurwitz zeta function.

Theorem 3.1. *For each $n \in \mathbb{N}$ and each $b \in (0, 1]$*

$$\zeta(1-n, b) = \frac{-B_n(b)}{n}.$$

In particular,

$$L(1-n, \chi) = -\sum_{k=1}^{f} \chi(k) f^{n-1} \frac{B_n(k/f)}{n} = -\frac{B_{\chi,n}}{n}.$$

Idea of Proof. Define complex valued functions

$$F(z) := \frac{ze^{(1-b)z}}{e^z - 1}$$

and

$$H(s) := \int_\gamma F(z) z^{s-2}\, dz$$

where γ is a path consisting of the positive real axis, a circle around the origin of "small" radius, and the negative real axis, with the orientation taking this order. The trick is to evaluate the integral in two different ways. One way uses Cauchy's residue theorem on the circle to find $H(1-n)$ in terms of a coefficient in the power series expansion of $F(z)$ around $z = 0$ which, after comparing the definition of $F(z)$ with (5), is not surprisingly given in terms of a value of the nth Bernoulli polynomial. The other way finds a formula for $H(s)$ when $\mathrm{Re}(s) > 1$ by shrinking the radius of the circle to make this piece vanish. Then we combine the remaining two pieces into an improper integral which is evaluated through rewriting part of the integrand as a geometric series, interchanging summation and integration, and finally recalling the integral defining the Γ function. The second way gives $H(s)$ in terms of the Hurwitz zeta function and the aforementioned Γ function. Taking the limit of this expression as $s \to 1-n$ gives the desired result. □

By setting $\chi = 1$ the trivial character, it follows immediately from (3.1) the well-known result that the Riemann zeta function $\zeta(s) = L(s, 1)$ has zeros at the negative even integers and is rational at the negative odd integers.

Also, if χ is an odd character (i.e., $\chi(-1) = -1$), then we have special function values at $s = 1$ for L-functions involving generalized Bernoulli numbers of the form $B_{1,\psi}$. Namely,

$$L(1, \chi) = \frac{\pi i}{f} B_{1,\overline{\chi}} \sum_{k=1}^{f} \chi(k) e^{2\pi i k/f}.$$

Now returning to the case of $\chi = 1$, we recall Euler's famous formula for the values of the Riemann zeta function at the positive even integers which can be shown with the functional equation.

Theorem 3.2. *For each $n \in \mathbb{N}$*

$$\zeta(2n) = (-1)^{n+1} \frac{(2\pi)^{2n}}{2(2n)!} B_{2n}.$$

Of course $\zeta(2n) > 0$ for $n \in \mathbb{N}$ and $\zeta(x) \to 1$ as $x \to \infty$ so we obtain the following result via Stirling's formula $n! \sim (n/e)^n \sqrt{2\pi n}$ as pointed out in [1].

Corollary 3.3. *The sequence $B_2, B_4, \ldots \in \mathbb{Q}^\times$ alternates in sign and*

$$|B_{2n}| \sim 4\pi\sqrt{e}\left(\frac{n}{\pi e}\right)^{2n+1/2}$$

as $n \to \infty$.

3.2. **p-adic L-functions.** In the last section we saw how special values of complex valued L-functions were related to Bernoulli numbers. It's also fruitful to consider L-functions arising from the non-archimedean valuations of $\mathbb{Q}$, namely the p-adic absolute values. We work in the completion $\mathbb{C}_p$ of the algebraic closure $\overline{\mathbb{Q}}_p$ of the completion $\mathbb{Q}_p$ of $\mathbb{Q}$ with respect to the p-adic norm. Here a function $f : \mathbb{C}_p \to \mathbb{C}_p$ is analytic at x when f may be expanded in a power series about x with some positive radius of convergence. We define p-adic meromorphic analogously.

Let p be an odd prime. Then for $a \in \mathbb{Z}_p$ with $p \nmid a$ let $\omega(a) \in \mathbb{Z}_p$ denote the unique $(p-1)$st root of unity such that

$$a \equiv \omega(a) \pmod{p}.$$

Theorem 3.4. *For each odd prime p and each Dirichlet character χ of conductor f, there exists a p-adic meromorphic L-function $L_p(s, \chi)$ on $\{s \in \mathbb{C}_p : |s| < p^{(p-2)/(p-1)}\}$ with*

$$L_p(1-n, \chi) = -(1 - \chi\omega^{-n}(p)p^{n-1})\frac{B_{n,\chi\omega^{-n}}}{n}$$

for all $n \in \mathbb{N}$. Actually, $L_p(s, \chi)$ is analytic whenever $\chi \neq 1$ and $L_p(s, 1)$ is analytic except for a simple pole at $s = 1$ with residue $1 - 1/p$.

Also, we have information about the coefficients when we expand about $s = 1$ for L-functions on nontrivial characters.

Theorem 3.5. *If $\chi \neq 1$ and $p^2 \nmid f$, then there are $a_0, a_1, \ldots \in \mathbb{Z}_p$ such that*

$$L_p(s, \chi) = a_0 + a_1(s-1) + a_2(s-1)^2 + \cdots$$

with $p | a_n$ for $n \in \mathbb{N}$.

Now we can combine the two previous results to arrive the following beautiful and useful congruences due to Kummer.

Corollary 3.6 (Kummer's Congruences)**.** *Let $a, m, n \in \mathbb{N}$ with $2m \equiv 2n \pmod{(p-1)p^{a-1}}$ and $2n \not\equiv 0 \pmod{p-1}$. Then*

$$(1 - p^{2m-1})\frac{B_{2m}}{2m} \equiv (1 - p^{2n-1})\frac{B_{2n}}{2n} \pmod{p^a}$$

Proof. Since $p - 1 \nmid 2n$ We have that the conductor of ω^{2n} is p, which isn't divisible by p^2, so

$$\begin{aligned} -(1 - p^{2m-1})\frac{B_{2m}}{2m} &= L_p(1-2m, \omega^{2m}) = a_0 + a_1(-2m) + a_2(-2m)^2 + \cdots \\ \equiv\ a_0 + a_1(-2n) + a_2(-2n)^2 + \cdots &= L_p(1-2n, \omega^{2n}) = -(1 - p^{2n-1})\frac{B_{2n}}{2n}. \end{aligned}$$

□

Actually, although using the coefficients in a power series expansion of a p-adic L-function to prove the Kummer congruences is slick and insightful, these congruences can be proved using purely elementary techniques. For example, the following congruences, credited to

Voronoi, can be used to prove (3.6), but to prove them requires only the von Staudt-Clausen theorem as demonstrated in [5].

Theorem 3.7. *Let $a, m, n \in \mathbb{N}$ and $B_{2m} = U_{2m}/V_{2m}$ with $U_{2m}, V_{2m} \in \mathbb{Z}$ and $(U_{2m}, V_{2m}) = 1$. Then $(a, n) = 1 \Rightarrow$*

$$(a^{2m} - 1)U_{2m} \equiv 2ma^{2m-1}V_{2m} \sum_{j=1}^{n-1} j^{2m-1} \left\lfloor \frac{ja}{n} \right\rfloor \pmod{n}.$$

4. Fermat's Last Theorem

Fermat's Last Theorem (FLT) asserts that $n \geq 3 \Rightarrow \nexists$solutions $x, y, z \in \mathbb{Z}$ to the equation

$$x^n + y^n = z^n$$

such that

$$0 \neq xyz.$$

The statement would follow if it were true for $n = 4$ and $n = p$ an odd prime. Euler handled the instance $n = 4$, and the instance $n = p$ is traditionally broken up into two cases: in the first case $p \nmid xyz$ and in the second case $p \nmid xy$ while $p|z$.

Both of the above cases can be dealt with using minimal machinery in the special scenario when p is a regular prime, i.e. p does not divide the class number h of $\mathbb{Q}(\zeta_p)$ where ζ_p is a primitive pth root of unity:

Theorem 4.1.

$$p \text{ is a regular prime} \Rightarrow \text{FLT is true for } n = p$$

Thus it is of interest to determine which primes are regular and which are irregular, meaning not regular. It turns out that

$$p \text{ is irregular} \Leftrightarrow p|h^- := h/h^+$$

where h^+ is the class number of the maximal real subfield $\mathbb{Q}(\zeta_p + \zeta_p^{-1})$. In fact, it's a famous conjecture of Vandiver that $p \nmid h^+$ for all p. Moreover, in [10] the formula

$$h^- = 2p \prod_{\substack{j=1 \\ j \text{ odd}}}^{p-2} \left(-\frac{1}{2} B_{1,\omega^j} \right) \equiv \prod_{\substack{j=1 \\ j \text{ odd}}}^{p-4} \left(-\frac{1}{2} \frac{B_{j+1}}{j+1} \right) \pmod{p}$$

is derived with the help of the Kummer congruences where ω is as in the previous section, and this is further used to show the following theorem.

Theorem 4.2. *As above, for each $n \in \mathbb{N}_0$ write $B_n = U_n/V_n$ where $(U_n, V_n) = 1$. Then*

$$p|h^- \Leftrightarrow p|U_n \text{for some } n \in \{2, 4, \ldots, p-3\}.$$

Actually, it follows from (2.2) that $B_2, B_4, \ldots, B_{p-3}$ are in the ring of p-integers $\mathbb{Z}_{(p)} = \{q \in \mathbb{Q} | \text{ord}_p(q) \geq 0\}$, so combining the above statements gives

$$p \text{ is regular} \Leftrightarrow B_2, B_4, \ldots, B_{p-3} \in \mathbb{Z}_{(p)}^{\times}$$

as noted in [5]. The Kummer congruences can also be used to show that there are, in fact, infinitely many irregular primes, although it's still unknown whether or not there exist infinitely many regular primes. However, all is not lost from this direction of FLT. In particular, as long as a prime p is not "too" irregular, similar methods can be applied to

confirm FLT in that case. In particular, define the index of irregularity $i(p)$ to be the number of B_m with $m \in \{2, 4, \ldots, p-3\}$ which have numerators divisible by p. Then we have the following result.

Theorem 4.3. *We have $i(p) < \sqrt{p} - 2 \Rightarrow x^p + y^p = z^p$ has no solutions $x, y, z \in \mathbb{Z}$ with $p \nmid xyz$.*

5. Theorems of Herbrand and Ribet

Let p be an odd prime and $C = A/A^p$ where $A = \mathrm{Cl}(K)$ is the ideal class group of $K := \mathbb{Q}(\zeta_p)$ with $\zeta_p = e^{2\pi i/p}$. Take

$$\Delta := \mathrm{Gal}(K/\mathbb{Q}) \cong \mathbb{Z}/(p-1).$$

Then there is a group action of Δ on C given as follows: for the equivalence class $[I] \in A$ of a fractional ideal I in K take $\overline{[I]} = [I]A^p \in C$ and

$$g \cdot \overline{[I]} = \overline{[g(I)]}$$

for each $g \in G$. Also, C is an $\mathbb{F}_p$-vector space of dimension $n < \infty$ since $c^p = 0$ for all $c \in C$ and $|C| < \infty$. Moreover, for each $g \in G$ the map

$$L_g : C \to C$$

given by

$$c \mapsto g \cdot c$$

is an invertible $\mathbb{F}_p$-linear transformation. In this way, C is an $\mathbb{F}_p\Delta$-module.

Now consider the $\mathbb{F}_p$-character group

$$\mathrm{Irr}_{\mathbb{F}_p}(\Delta) = \langle \chi \rangle \cong \mathbb{Z}/(p-1).$$

Then (by say 2.1.7 in [8] or similar results) we have

$$C = \bigoplus_{k=1}^{p-1} C(\chi^{1-k})$$

as $\mathbb{F}_p\Delta$-modules where

$$C(\chi^k) = \epsilon_k \cdot C$$

are the χ^k-isotypical components with block idempotents

$$\epsilon_k = \frac{\chi^k(1)}{|\Delta|} \sum_{g \in \Delta} \chi^k(g^{-1}) g = \frac{1}{p-1} \sum_{g \in \Delta} \chi^{-k}(g) g.$$

By χ^k-isotypical we mean that for each $g \in \Delta$, $C(\chi^k)$ is the eigenspace of L_g corresponding to the eigenvalue $\chi^k(g)$.

Throughout the remainder of this section fix $k \in \{2, 4, \ldots, p-3\}$. It turns out that

$$C(\chi^{1-k}) \neq 0 \Leftrightarrow \mathrm{ord}_p(B_k) > 0.$$

The forward direction of this statement was proved by Herbrand in 1932, while the backward direction was proved by Ribet in 1976.

5.1. **Herbrand's Theorem.** In this subsection we'll state Stickelberger's theorem and give an indication of how this along with the Voronoi congruences is used to prove the following theorem of Herbrand as seen in [5].

Theorem 5.1 (Herbrand)**.** *We have*

$$C(\chi^{1-k}) \neq 0 \Rightarrow ord_p(B_k) > 0.$$

To prove the contrapositive of Herbrand's theorem assuming $\mathrm{ord}_p(B_k) = 0$ it's enough to show

$$C(\chi^{1-k}) = 0.$$

Define the Stickelberger element $\theta \in \mathbb{Q}\Delta$ by

$$\theta := \sum_{t=1}^{p-1} \mathrm{frac}\left(\frac{t}{p}\right) \sigma_t^{-1}$$

where $\sigma_t \in \Delta$ sends ζ_p to ζ_p^t and $\mathrm{frac}(x)$ denotes the fractional part of a real number x. Now define the Stickelberger ideal $S \subseteq \mathbb{Z}\Delta$ to be

$$S := \mathbb{Z}\Delta \cap \theta\mathbb{Z}\Delta.$$

Theorem 5.2 (Stickelberger)**.** *The Stickelberger ideal S annihilates the ideal class group A of $K = \mathbb{Q}(\zeta_p)$ (and therefore also annihilates C).*

For each $t \in \{1, \ldots, p-1\}$ take

$$r_t := (\sigma_t - t)\theta.$$

The following lemma is established in [5]

Lemma 5.3. *We have $r_t \in S$ for all t.*

We are now in a position to give a sketch of Herbrand's theorem as found in [5].

Sketch of Proof of (5.1). Suppose $\mathrm{ord}_p(B_k) = 0$. Check that for $s \in \{1, \ldots, p-1\}$ we have

$$r_s = -\sum_{t=1}^{p-1} \left\lfloor \frac{st}{p} \right\rfloor \sigma_t^{-1}.$$

Let $c \in C(\chi^{1-k})$. Then by (5.2) and (5.3)

$$1 = r_s \cdot c = c^{\sum_{t=1}^{p-1} \left\lfloor \frac{st}{p} \right\rfloor t^{p-2+k}} = c^{\sum_{t=1}^{p-1} \left\lfloor \frac{st}{p} \right\rfloor t^{k-1}}.$$

On the other hand, if $B_k = U_k/V_k$ with $U_k, V_k \in \mathbb{Z}$ and $(U_k, V_k) = 1$, then by the Voronoi congruences

$$c^{(s^k-1)U_k} = c^{ks^{k-1} \sum_{t=1}^{p-1} \left\lfloor \frac{st}{p} \right\rfloor t^{k-1}} = 1.$$

Thus choosing s to be a primitive root modulo p we get $p \nmid s^k - 1$, but by assumption $p \nmid U_k$, so

$$c = 1.$$

Therefore $C(\chi^{1-k}) = 0$.

□

5.2. **Ribet's Theorem.** The goal of this subsection is to sketch the ideas involved in proving the following converse to Herbrand's theorem as seen in [6].

Theorem 5.4 (Ribet)**.** *We have*

$$ord_p(B_k) > 0 \Rightarrow C(\chi^{1-k}) \neq 0.$$

Let E be the maximal unramified p-extension of K. Then there's a map $\phi_{E/K}$ from the fractional ideals $\mathcal{F}$ of K into $H := \mathrm{Gal}(E/K)$ given by sending a prime ideal $\mathfrak{p}$ in $\mathcal{O}_K$ to the Artin symbol $(E/K, \mathfrak{p}) \in H$. The kernel of this map is $\mathcal{F}^p\mathcal{P}$ where $\mathcal{P}$ is the set of principal fractional ideals. Thus there is a group isomorphism

$$\mathrm{Art} : C \to H$$

given by

$$\overline{[I]} \mapsto \prod_{\mathfrak{p}^\alpha || I} (E/K, \mathfrak{p})^\alpha$$

with

$$C = A/A^p = \frac{\mathcal{F}/\mathcal{P}}{(\mathcal{F}/\mathcal{P})^p} = \frac{\mathcal{F}/\mathcal{P}}{\ker(\phi_{E/K})/\mathcal{P}} \cong \mathcal{F}/\ker(\phi_{E/K}).$$

Also, Δ acts on H as follows: for $g \in \Delta$ and $h \in H$, we have that

$$K \to K \hookrightarrow E$$

is a $\mathbb{Q}$-algebra homomorphism from K to E which extends uniquely to a $\mathbb{Q}$-algebra isomorphism $\widetilde{g}$ from E to E, so we may define a $\mathbb{Q}$-algebra isomorphism

$$g \cdot h = \widetilde{g} \circ h \circ \widetilde{g}^{-1},$$

but note that if $x \in K$, then $\widetilde{g}^{-1}(x) \in K$, giving $(g \cdot h)(x) = x$ and $g \cdot h \in H$. Moreover, Art preserves the Δ-actions, so to prove Ribet's theorem assuming $\mathrm{ord}_p(B_k) > 0$ it's enough to find an intermediate field $F \neq K$ in E/K such that

$$g \cdot a = \chi^{1-k}(g)a$$

whenever $g \in \Delta$ and $a \in \mathrm{Gal}(F/K)$. This suffices because $F \neq K$ implies $\mathrm{Gal}(E/F) \neq H$, so

$$\mathrm{Gal}(F/K) \cong \frac{H}{\mathrm{Gal}(E/F)} \ncong 0$$

by the fundamental theorem of Galois theory. Such an F can be constructed using the following theorem.

Theorem 5.5. *$ord_p(B_k) > 0 \Rightarrow$ there is a finite extension $\mathbb{F}/\mathbb{F}_p$ and representation*

$$\rho : G_{\mathbb{Q}} = Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{F})$$

such that

(i) ρ is unramified at all primes different than p,

(ii) ρ is a reducible non-semisimple representation (i.e, $p|im(\rho)$) of the form

$$\begin{pmatrix} 1 & * \\ 0 & \chi^{k-1} \end{pmatrix}$$

with the $*$ nontrivial, and

(iii) $\rho|_D$ is semisimple (i.e., $p \nmid \rho(D)$) where $D \leq G_{\mathbb{Q}}$ is a decomposition group of p.

To get F with this theorem let $F' \subseteq \overline{\mathbb{Q}}$ be the fixed field of $\ker(\rho)$ and $K^{\otimes(k-1)} \subseteq K$ be the fixed field of $\ker(\chi^{k-1})$. Then $\mathrm{Gal}(K^{\otimes(k-1)}/\mathbb{Q})$ acts on $H' := \mathrm{Gal}(E/K^{\otimes(k-1)})$ with

$$g \cdot a = \chi^{1-k}(g)a$$

whenever $g \in \mathrm{Gal}(K^{\otimes(k-1)}/\mathbb{Q})$ and $a \in H'$. Now $H' \leq H \cong C$ is abelian with $C^p \cong \{1\}$, so $H' \cong (\mathbb{Z}/p\mathbb{Z})^m$ for some $m \in \mathbb{N}_0$, but (ii) implies $m \neq 0$, so p is unramified in $E/K^{\otimes(k-1)}$ since the primes above p split in $E/K^{\otimes(k-1)}$ by (iii). Thus $E/K^{\otimes(k-1)}$ is unramified since (i) implies it's unramified away from p, so taking

$$F := KF'$$

gives the needed field since K and F' are linearly disjoint over $K^{\otimes(k-1)}$.

After the sufficiency of (5.5) is secured, the objective becomes to find a representation ρ satisfying the above properties (i)-(iii). To accomplish this one needs to use ideas from the study of modular forms. A modular form of weight k for a subgroup $\Gamma' \subseteq \Gamma := \mathrm{SL}_2(\mathbb{Z})$ with character

$$\epsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$$

is an analytic function f on the upper half-plane $\mathcal{H} \subseteq \mathbb{C}$ with a Fourier series of the form

$$\sum_{n=0}^{\infty} a_n(f)q^n$$

where $q = e^{2\pi iz}$ such that

$$f\left(\frac{az+b}{cz+d}\right) = \epsilon(d)(cz+d)^k f(z)$$

whenever

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma'.$$

Denote the set of all such functions by $M_k(\Gamma', \epsilon)$. If $f \in M_k(\Gamma', \epsilon)$ also has $a_0 = 0$, then f is called a cuspform. Denote the set of all cuspforms in $M_k(\Gamma', \epsilon)$ by $S_k(\Gamma', \epsilon)$. Now specialize Γ' to be the congruence subgroup

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \middle| \, c \equiv 0 \pmod{N} \right\}.$$

There is an action of Hecke operators T_n for $n \in \mathbb{N}$ on $S_k(\Gamma_0(N), \epsilon)$. A form which is an eigenvector for some T_n is called an eigenform. The notion of an eigenform is a very useful one since it gives us information about the Fourier coefficients as noted in [7]. In our case, if f is an eigenform for almost all T_r with r prime, then the Fourier coefficients $a_n(f)$ lie in a number field and, more importantly, for each prime ℓ there is a representation

$$\rho_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{E})$$

where $\mathbb{E}$ is a finite extension of $\mathbb{Q}_\ell$. This ℓ-adic representation is absolutely irreducible (i.e., is remains irreducible whenever viewed as a representation over a larger field) and is unramified at almost all primes r with ℓ excluded from that list. However, ρ_f may or may not be semisimple. Now we carefully choose a basis such that with respect to this basis the

representation takes values in $\mathrm{GL}_2(\mathcal{O}_{\mathbb{E}})$ where $\mathcal{O}_{\mathbb{E}}$ is the ring of integers of $\mathbb{E}$. We can then reduce ρ_f modulo the maximal ideal of $\mathcal{O}_{\mathbb{E}}$ to get a representation

$$\overline{\rho}_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$$

for some finite extension $\mathbb{F}$ of $\mathbb{F}_\ell$. Now we can set $\ell = p$ and ask what choice of f will have a reduction $\overline{\rho}_f$ which satisfies the properties in (5.5). It's a critical proposition of Ribet's which explains that a reduction $\overline{\rho}_f$ satisfying (ii) exists when some reduction of ρ_f is reducible and its semisimplification is isomorphic to $1 \oplus \chi^{k-1}$. One can get the reducible condition if f looks modulo p like the Eisenstein series

$$-\frac{B_k}{2k} + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

where

$$\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}.$$

Thus under the assumption $\mathrm{ord}_p(B_k) > 0$ one would like to prove the existence of a cuspform $f \in S_k(\Gamma)$ with algebraic Fourier coefficients $a_n(f)$ satisfying the congruences

$$a_n(f) \equiv \sigma_{k-1}(n) \pmod{\wp}$$

where $\wp$ is a place above p. The University of Arizona's own Kirti Joshi suggests how to find such an f by considering a polynomial in the discriminant function. Using these ideas along with some other results like the Chebotarev density theorem, a lemma due to Deligne-Serre and a theorem of Brauer Nesbitt, one can find a representation ρ satisfying properties (i) and (ii) of (5.5). To get property (iii), however, one needs to use yet more machinery.

References

1. Tom M. Apostol, *Introduction to analytic number theory*, Springer, 1976.
2. L. Carlitz, *The Staudt-Clausen theorem*, Mathematics Magazine **34** (1961), no. 3.
3. Richard Crandall and Carl Pomerance, *Prime numbers: A computational perspective*, Springer, 2001.
4. Greg Fee and Simon Plouffe, *An efficient algorithm for the computation of Bernoulli numbers*, 2007.
5. Kenneth Ireland and Michael Rosen, *A classical introduction to modern number theory*, second ed., Springer, 1990.
6. Chandrashekhar Khare, *Notes on Ribet's converse to Herbrand*, Cyclotomic Fields and Related Topics (2000).
7. Neal Koblitz, *Introduction to elliptic curves and modular forms*, second ed., Springer, 1993.
8. K. Lux and H. Pahlings, *Representations of groups - theory and practice*, draft ed., 2007.
9. I. Sh. Slavutskii, *Real quadratic fields and the Ankeny-Artin-Chowla conjecture*, Journal of Mathematical Sciences **122** (2004), no. 6.
10. Lawrence C. Washington, *Introduction to cyclotomic fields*, second ed., Springer, 1996.
11. Jianqiang Zhao, *Bernoulli numbers, Wolstenholme's theorem, and p^5 variations of Lucas' theorem*, J. Number Theory (2007).