# MODULAR ARITHMETIC AND THE ENTROPY FUNCTION (AN UNDERGRADUATE COLLOQUIUM HELD ON 10.03.2010)

HERBERT GANGL

## 1. Modular arithmetic

In the following we fix an *odd* prime $p$.
⟦E.g. $p$ =3,5,7, 101, 691, 12421, 13931, 111...1 (19 digits), $34790! - 1, \ldots, 2^{43112609} - 1 \ldots$⟧
For such a $p$, we work in $\mathbb{Z}/p\mathbb{Z}$ :={residue classes modulo $p$}= $\{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\}$, where we can add and multiply, and even invert. We write $a \equiv b \pmod{p}$ or $a \equiv_p b$ if $a$ and $b$ have the same remainder when dividing by $p$.

(This structure, a field, is often also denoted $\mathbb{Z}_p$, a notation which alas clashes with the one for another standard object in arithmetic, the $p$-adic integers).

**Example 1.1.** *For $p = 7$ we have e.g. $\bar{5}^{-1} = \bar{3}$ (since $3 \cdot 5 \equiv 1 \pmod{7}$) and $\bar{4}/\bar{5} = \bar{4} \cdot \bar{3} = \bar{5}$ (as $5 \cdot 5 \equiv 4 \pmod{7}$).*

We will drop the bars over numbers from now on in our notation. We find the following little result:

**Claim 1.2.** $1 + \dfrac{1}{2} + \dfrac{1}{3} + \cdots + \dfrac{1}{p-1} \equiv 0 \pmod{p}$ *($p \geqslant 3$ prime).*

**Note 1.3.** *Each $a \in \mathbb{Z}/p\mathbb{Z}$, $a \neq 0$, has a unique inverse, hence*

$$\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{p-1}\} \pmod{p} \quad \overset{1:1}{\longleftrightarrow} \quad \{1, 2, \ldots, p-1\} \pmod{p}.$$

⟦*for $p = 7$:* $\{1, 2^{-1}, 3^{-1}, \ldots, 6^{-1}\} \equiv_7 \{1, 4, 5, 2, 3, 6\} \quad \overset{1:1}{\longleftrightarrow} \quad \{1, 2, 3, 4, 5, 6\}$⟧

From this we conclude

$$\sum_{n=1}^{p-1} n^{-1} \equiv \sum_{n=1}^{p-1} n \pmod{p}.$$

But we know how to sum the latter! ⟦Famous anecdote: late 18th century, the teacher of a noisy class among which was the $\sim$ 9-yr old Gauss gave the task to sum all the numbers from 1 to 100. Gauss promptly handed in his solution (on a tablet placed on the teacher's desk), without any calculation, while the other pupils kept on crunching numbers for much longer. In the end Gauss had given the only correct solution. He simply had spotted a symmetry of the problem that made it easy to solve it: he regrouped (1+100) + (2 +99) ...+ (50 +51), all of which partial sums are equal to 101, of which there are 50, hence overall 5050.⟧

More generally, we have

$$\sum_{n=1}^{p-1} n = \frac{(p-1)p}{2} \equiv 0 \pmod{p},$$

which implies the claim above.

The identity in Claim 1.2 is a special value of a function which looks like the logarithm: from Analysis we know that the power series for the logarithm around 1 can be written as

$$-\log(1-x)=\sum_{n=1}^{\infty}\frac{x^n}{n}\,.$$

Let us truncate this! (In fact, for a given odd prime $p$ we truncate before we get non-sensical terms modulo $p$, i.e. at $p-1$, as we are not allowed to have $p$ dividing the denominator—this would amount to dividing by 0.)

**Definition 1.4.** *The* **finite 1-logarithm** *attached to a given odd prime $p$ is defined as*

$$\mathcal{L}_1(x):=\mathcal{L}_1^{(p)}(x):=\sum_{n=1}^{p-1}\frac{x^n}{n}\pmod p \;=\; x+\frac{x^2}{2}+\cdots+\frac{x^{p-1}}{p-1}\pmod p\,.$$

We can now restate the above claim as our

**Property 1.5.** $\mathcal{L}_1(1)\equiv 0 \pmod p$.

A more general fact is

**Property 1.6.** $x^p\mathcal{L}_1\left(\frac{1}{x}\right)\equiv_p -\mathcal{L}_1(x)$ *as a polynomial.*

**Example:** $p=5$. LHS: $x^5(\frac{1}{x}+\frac{1}{2}\left(\frac{1}{x}\right)^2+\frac{1}{3}\left(\frac{1}{x}\right)^3+\frac{1}{4}\left(\frac{1}{x}\right)^4)=x^4+\frac{1}{2}x^3+\frac{1}{3}x^2+\frac{1}{4}$.
RHS: $-x-\frac{1}{2}x^2-\frac{1}{3}x^3-\frac{1}{4}x^4$. But $-\frac{1}{4}\equiv_5 1$, $-\frac{1}{3}\equiv_5 2$, ...

**Proof:** LHS $=x^p\sum_{n=1}^{p-1}\frac{1}{n}\left(\frac{1}{x}\right)^n=\sum_{n=1}^{p-1}\frac{x^{p-n}}{n}\equiv_p-\sum_{n=1}^{p-1}\frac{x^{p-n}}{p-n}=-\sum_{\ell=p-1}^{p-(p-1)}\frac{x^\ell}{\ell}=-\sum_{\ell=1}^{p-1}\frac{x^\ell}{\ell}$ =RHS. □

A bit more difficult to prove is

**Property 1.7.** $\mathcal{L}_1(1-x)\equiv_p\mathcal{L}_1(x)$.

This is a special case of a more elaborate identity:

**Theorem 1.8.** $(1-y)^p\mathcal{L}_1\Big(\dfrac{x}{1-y}\Big)-\mathcal{L}_1(x)$ *is symmetric in $x$ and $y$.*

**Corollary 1.9.** *(4-term identity)*

$$(1-y)^p\mathcal{L}_1\Big(\frac{x}{1-y}\Big)-\mathcal{L}_1(x)\equiv_p(1-x)^p\mathcal{L}_1\Big(\frac{y}{1-x}\Big)-\mathcal{L}_1(y)\,.$$

**Proof:** The leftmost term can be written

$$\sum_{r=1}^{p-1}\frac{1}{j}x^r(1-y)^{p-r}=\sum_{r=1}^{p-1}\frac{1}{r}x^r\sum_{\ell=0}^{p-r}\binom{p-r}{\ell}(-1)^\ell y^\ell$$

and we can take the bound for the inner sum to be $p-1$ instead, using the vanishing of binomial coefficients $\binom{p-r}{\ell}$ for $\ell>p-r$. Now the terms with $\ell=0$ add up to $\mathcal{L}_1(x)$, so the left hand side of the corollary can be readily reduced to the double sum $\sum_r\sum_\ell\frac{(-1)^\ell}{r}\binom{p-r}{\ell}x^ry^\ell$

with both sums ranging from 1 to $p-1$, and in order to show the symmetry with respect to swapping $x$ and $y$ we are left to show the straightforward binomial identity

$$\frac{(-1)^r}{\ell}\binom{p-\ell}{r} = \frac{(-1)^\ell}{r}\binom{p-r}{\ell}. \quad \square$$

## 2. The (binary) Entropy Function

Claude Shannon in 1948 when he (single-handedly) founded "Information Theory", in the process of which he was led to introduce the crucial "Entropy function" given by

$$H(x) = -x\log(x) \ - \ (1-x)\log(1-x)\,, \quad x \in (0,1)$$

and which can be extended to $x \in \mathbb{R}$ simply by replacing $\log(..)$ by $\log|..|$.

We find the following properties:

**Property 2.1.** $H(1) = 0$.

⟦Note that $\lim_{x\to 1}(1-x)\log(1-x) = 0$.⟧

**Property 2.2.** $H(1-x) = H(x)$ *(obvious).*

**Property 2.3.** $xH\big(\frac{1}{x}\big) = -H(x)$.

The latter is a consequence of a more elaborate identity:

**Theorem 2.4.** $(1-y)H\big(\frac{x}{1-y}\big) + H(y)$ *is symmetric in* $x$ *and* $y$.

This results in a very similar identity as for $\pounds_1$:

**Corollary 2.5.** *(4-term identity)*

$$(1-y)H\big(\frac{x}{1-y}\big) + H(y) = (1-x)H\big(\frac{y}{1-x}\big) + H(x)\,.$$

**Proof:** We first consider the leftmost term

$$\begin{aligned}
&(1-y)H\big(\frac{x}{1-y}\big)\\
&= \ (1-y)\Big(-\frac{x}{1-y}\log\big|\frac{x}{1-y}\big| \ - \ \frac{1-y-x}{1-y}\log\big|\frac{1-y-x}{1-y}\big|\Big)\\
&= \ -x\log|x| + x\log|1-y| - (1-y-x)\log|1-y-x| + (1-y-x)\log|1-y|\\
&= \ -x\log|x| + x\log|1-y| - (1-y-x)\log|1-y-x| + (1-y)\log|1-y| - x\log|1-y|
\end{aligned}$$

where the second and last term in the last sum cancel. Adding $H(y)$ gives the symmetric (in $x$ and $y$) expression

$$-x\log|x| \ - \ (1-y-x)\log|1-y-x| \ - \ y\log|y|\,. \quad \square$$

**Remarks 2.6.** (1) *The 4-term identity for* $H$ *is called the* fundamental equation of information theory*!*

(2) *We have the following "interpretation":*
$(1-y)H\left(\frac{x}{1-y}\right) \longleftrightarrow H(X \mid Y) =$ conditional *entropy (i.e. entropy of* $X$ *conditional on* $Y$*);*
$H(X,Y) =$ joint *entropy* $= H(Y) + H(X \mid Y)$.
*and the fundamental equation then says*

$$H(X,Y) = H(Y,X)\,,$$

*i.e., as the name "joint entropy" already suggests, this quantity should not depend on the order of* $X$ *and* $Y$.

- The analogy between $\mathcal{L}_1$ and $H$ was found by Kontsevich ("The $1\frac{1}{2}$-logarithm", Hirzebruch Retirement Volume '95), with a further interpretation of the 4-term identity above as a "2-cocycle equation".
- A different link between $H$ and the *derivative* of the *untruncated* higher sum

$$\sum_{n=1}^{\infty} \frac{x^n}{n^2} =: Li_2(x)$$

(more precisely of a variant of $x(1-x)\frac{d}{dx}Li_2(x)$) was found by Cathelineau ("On the homology of $\mathrm{SL}_2$", Math. Scand. '88) in connection with one of the famous Hilbert Problems (the third one, on scissors congruences); he also provided a homological interpretation.

An "explanation" of sorts of these related phenomena was given by Elbaz-Vincent and myself ("On Poly(ana)logs", Compositio Math. 2002). In that paper we also solved a follow-up question of Kontsevich:

**Q.:** How to find such identities ("functional equations") for the next such truncated series, i.e. for

$$\mathcal{L}_2(x) := \sum_{n=1}^{p-1} \frac{x^n}{n^2}\,?$$

**A.:** (Very roughly) Take a functional equation for the untruncated sum $Li_3(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^3}$, apply some "tangential procedure", very similar to one given by Cathelineau, to arrive at a functional equation for $\mathcal{L}_2$. (Similarly for even higher exponents.)

## 3. Connection to Fermat's Last Theorem

Mirimanoff (1905) gave a criterion for dismissing primes $p$ for the *first case* of FLT (i.e. for the possibility of an integer triple $(x,y,z)$ such that $x^p+y^p+z^p=0$ and $p \not| xyz$); he introduced polynomials

$$\varphi_j(T) = \sum_{n=1}^{p-1} n^{j-1} T^n\,, \qquad (j = 1, \ldots, p-1)\,.$$

We note that they are in fact our truncated sums in disguise (use "Little Fermat"):

$$\varphi_{p-j}(T) = \sum_{n=1}^{p-1} n^{p-j-1} T^n \equiv_p \sum_{n=1}^{p-1} \frac{T^n}{n^j} = \mathcal{L}_j(T)\,.$$

A crucial identity for Mirimanoff from which he deduced his FLT criteria was the following:

$$-\frac{1}{2}[\varphi_{p-1}(T)]^2 = \varphi_{p-2}(T) + (T-1)^{2p}\varphi_{p-2}\big(\frac{T}{T-1}\big)\,,$$

i.e. in our notation

$$-\frac{1}{2}[\pounds_1(T)]^2 = \pounds_2(T) + (T-1)^{2p}\pounds_2\big(\frac{T}{T-1}\big)\,.$$

This is a very special case of the following 2-variable identity.

**Challenge 3.1.** *Show that, for $p \geqslant 3$ prime we have*

$$\begin{aligned} -\pounds_1(a)\pounds_1(b) = -a^p\pounds_2\big(\tfrac{b}{a}\big) \quad & -(1-a)^p\,\pounds_2\big(\tfrac{1-b}{1-a}\big) && + \big(b(1-a)\big)^p\,\pounds_2\big(\tfrac{a(1-b)}{b(1-a)}\big) \\ & + \pounds_2(1-a) + \pounds_2(1-b) && + (ab)^p\,\pounds_2\big(1-\tfrac{1}{b}\big) + (ab)^p\,\pounds_2\big(1-\tfrac{1}{a}\big). \end{aligned}$$

(Here the grouping is intended and its meaning should become clear in a proof of the identity.)

**Research Challenge 3.2.**

(1) *Find possible generalisations of this equation (e.g. can one write $\pounds_1(a)\pounds_1(b)\pounds_1(c)$ in terms of $\pounds_3$?).*
(2) *Reinterpret other FLT criteria in terms of identities for $\pounds_j$ (or products thereof). Does the 2-variable equation above perhaps give improved FLT criteria?*
(3) *Consider nested sums and their equations and relate them to the above.*