Note

Young Gauss Meets Dynamical Systems

CONSTANTIN P. NICULESCU

Most people are convinced that doing mathematics is something like computing sums such as

$$S = 1 + 2 + 3 + \cdots + 100.$$

But we know that one who does this by merely adding terms one after another is not seeing the forest for the trees.

An anecdote about young Gauss tells us that he solved the above problem by noticing that pairwise addition of terms from opposite ends of the list yields identical intermediate sums. This famous story is well told by Hayes in [5], with references. A very convenient way to express Gauss's idea is to write down the series twice, once in ascending and once in descending order,

$$\begin{matrix} 1 & + & 2 & + & 3 & + & \cdots & + & 100 \\ 100 & + & 99 & + & 98 & + & \cdots & + & 1 \end{matrix}$$

and to sum columns before summing rows. Thus

$$\begin{aligned} 2S &= (1+100) + (2+99) + \cdots + (100+1) \\ &= \underbrace{101 + 101 + \cdots + 101}_{100 \text{ times}} \\ &= 10100, \end{aligned}$$

whence

$$S = 5050.$$

Of course, the same technique applies to any arithmetic progression

$$a_1, \quad a_2 = a_1 + r, \quad a_3 = a_1 + 2r, \ldots, \quad a_n = a_1 + (n-1)r, \tag{1}$$

and the result is the well-known summation formula

$$a_1 + a_2 + \cdots + a_n = \frac{n(a_1 + a_n)}{2}. \tag{2}$$

A similar idea can be used to sum up strings that are not necessarily arithmetic progressions. For example,

$$\binom{n}{0} a_0 + \binom{n}{1} a_1 + \cdots + \binom{n}{n} a_n = 2^{n-1}(a_0 + a_n),$$

for every arithmetic progression $a_0, a_1, \ldots, a_n$.

Seventy years ago, A. L. O'Toole [11] recommended that teachers avoid the above derivation of the formula (2), considering it a mere trick that offers no insight. Instead, he called attention to *the fundamental theorem of summation*, a discrete variant of the Leibniz-Newton theorem: If there is a function $f(x)$ such that $a_k = f(k+1) - f(k)$ for $k \in \{1, \ldots, n\}$, then

$$\sum_{k=1}^{n} a_k = f(n+1) - f(1) = f(k)\big|_1^{n+1}.$$

Indeed, this theorem provides a unifying approach for many interesting summation formulae (including those for arithmetic progressions and geometric progressions). However, determining the nature of the function $f(x)$ is not always immediate. In the case of an arithmetic progression (1) we may choose $f(x)$ as a second-degree polynomial, namely,

$$f(x) = \frac{r}{2}x^2 + (a_1 - \frac{3r}{2})x + C,$$

where C is an arbitrary constant.

Though more limited, "Gauss's trick" is much simpler, and besides, it provides a nice illustration of a key concept of contemporary mathematics, that of *measurable dynamical system*.

Letting $M = \{1, \ldots, n\}$, we may consider the measurable space $(M, \mathcal{P}(M), \mu)$, where $\mathcal{P}(M)$ is the power set of M and μ is the *counting measure* on M, defined by the formula

$$\mu(A) = |A| \quad \text{for every } A \in \mathcal{P}(M).$$

Every real sequence $a_1, \ldots, a_n$ of length n can be thought of as a function $f : M \to \mathbb{R}$, given by $f(k) = a_k$. Moreover, f is integrable with respect to μ, and

$$\int_M f(k) d\mu = a_1 + \cdots + a_n.$$

The main ingredient that makes possible an easy computation of the sum of an arithmetic progression is the existence of a nicely behaved map, namely,

$$T : M \to M, \quad T(k) = n - k + 1.$$

Indeed, the measure μ is *invariant* under the map T in the sense that

$$\int_M f(k) d\mu = \int_M f(T(k)) d\mu \tag{3}$$

regardless of the choice of f (for T is just a permutation of the summation indices).

When f represents an arithmetic progression of length n, then there exists a positive constant C such that

Published online: 21 January 2011

$$f(k)+f(T(k))=C, \quad \text{for all } k\in M, \tag{4}$$

and taking into account (3) we recover the summation formula (2) in the following equivalent form,

$$\int_M f(k)d\mu = \frac{1}{2}C|M|.$$

The natural generalization of the reasoning above is to consider arbitrary triples (M, T, μ), where M is an abstract space, μ is a finite positive measure defined on a σ-algebra Σ of subsets of M, and $T : M \to M$ is a measurable map that is invariant under the action of μ in the sense that (3) works for all $f \in L^1(\mu)$. Such triples are usually called *measurable dynamical systems*. In this context, if $f \in L^1(\mu)$ satisfies a formula like

$$f(T(x)) = \lambda f(x) + g(x) \tag{5}$$

with $\lambda \neq 1$, then the computation of $\int_M f(x)d\mu$, or rather of its *expectation*,

$$\mathcal{E}(f) = \frac{1}{\mu(M)}\int_M f(t)d\mu(t),$$

reduces to the computation of $\int_M g(x)d\mu$.

For example, the integral of an odd function over an interval symmetric about the origin is zero; this corresponds to (5) for $T(x) = -x$, $\lambda = -1$, and $g = 0$. Among the many practical implications of this remark, the following two seem especially important:

a) the Fourier series of any odd function is a series of sine functions;
b) the barycenter of any body that admits an axis of symmetry lies on that axis.

Two other instances of the formula (5) are

$$\int_0^\infty \frac{\ln x}{1+x^2}dx = 0$$

and

$$\int_0^{\pi/4} \ln(1+\tan x)dx = \frac{\pi \ln 2}{8}. \tag{6}$$

In the first case, the measurable dynamical system under consideration is the triple consisting of the interval $M = (0, \infty)$, the map $T(x) = 1/x$, and the weighted Lebesgue measure $\frac{dx}{1+x^2}$. The invariance of this measure with respect to T is assured by the change of variable formula, while the formula (5) becomes $\ln(1/x) = -\ln x$.

In the second case, the measurable dynamical system is the triple $([0, \pi/4], \pi/4 - x, dx)$. For $f(x) = \ln(1 + \tan x)$, the formula (5) becomes

$$\ln(1+\tan(\pi/4-x)) = -\ln(1+\tan x) + \ln 2$$

and thus

$$\begin{aligned}\int_0^{\pi/4} \ln(1+\tan x)dx &= \int_0^{\pi/4} \ln(1+\tan(\pi/4-x))dx \\ &= \int_0^{\pi/4} [\ln 2 - \ln(1+\tan x)]dx \\ &= \frac{\pi \ln 2}{4} - \int_0^{\pi/4} \ln(1+\tan x)dx,\end{aligned}$$

whence (6). This formula admits a straightforward generalization:

$$\int_0^\theta \ln(1+\tan\theta\tan x)dx = -\theta\ln(\cos\theta),$$

for all $\theta \in (-\pi/2, \pi/2)$.

In the same manner we obtain the integral formulae

$$\int_0^\pi xf(\sin x)dx = \frac{\pi}{2}\int_0^\pi f(\sin x)dx,$$

$$\int_0^\pi f(\sin x)dx = 2\int_0^{\pi/2} f(\sin x)dx.$$

There is a relationship between the expectation of a function f and the values of the iterates of f under the action of T,

$$f, \quad f\circ T, \quad f\circ T^2, \ldots,$$

expressed in the *ergodic theorems*. A sample is Weyl's ergodic theorem; here M is the unit interval, μ is the restriction of Lebesgue measure to the unit interval, and $T : [0,1] \to [0,1]$ is the irrational translation defined by

$$T(x) = \{x+\alpha\};$$

here $\{\cdot\}$ denotes the fractional part and $\alpha > 0$ is some irrational number. The invariance of T is usually derived from the remark that the linear span of characteristic functions of subintervals of [0,1] is dense in $L^1([0,1])$. Thus the verification of the invariance formula (3) reduces to the (trivial) case where f is such a characteristic function.

The following result does not make use of the invariance property of T (but can be used to derive it).

AUTHOR

CONSTANTIN P. NICULESCU received his Ph.D. at the University of Bucharest; he has been teaching at the University of Craiova since 1976. He works on convex analysis (see his joint work with Lars-Erik Persson, *Convex Functions and Their Applications*), functional analysis, and dynamical systems. He also lectures on heuristic, the art and science of discovery and invention. His hobbies include reading, music, and gardening.

Department of Mathematics
University of Craiova
Craiova, RO-200585
Romania
e-mail: cniculescu47@yahoo.com

THEOREM 1. (*Weyl's Ergodic Theorem* [10]). *Suppose that* $\alpha > 0$ *is irrational. Then*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} f(\{x + k\alpha\}) = \int_0^1 f(t)dt \tag{7}$$

for all Riemann integrable functions $f : [0,1] \to \mathbb{R}$ *and all* $x \in [0,1]$.

PROOF. It is easy to check that the above formula holds for each of the functions $e^{2\pi int}$ ($n \in \mathbb{Z}$), and thus for linear combinations of them. By the Weierstrass approximation theorem (see [3]) it follows that the formula (7) actually holds for all continuous functions $f : [0,1] \to \mathbb{C}$ with $f(0) = f(1)$.

Now if $I \subset [0,1]$ is a subinterval, then for each $\varepsilon > 0$ one can choose continuous real-valued functions g, h with $g \le \chi_I \le h$ such that

$$g(0) = g(1), \quad h(0) = h(1) \quad \text{and} \quad \int_0^1 (h - g)dt < \varepsilon.$$

By the previous step we infer that

$$\overline{\lim_{N\to\infty}} \frac{1}{N}\sum_{k=0}^{N-1} \chi_I(\{x + k\alpha\}) - \underline{\lim_{N\to\infty}} \frac{1}{N}\sum_{k=0}^{N-1} \chi_I(\{x + k\alpha\})$$

lies in $(\int_0^1 \chi_I(t)dt - \varepsilon, \int_0^1 \chi_I(t)dt + \varepsilon)$. As $\varepsilon > 0$ was arbitrarily fixed, this shows that the formula (7) works for χ_I (and thus for all step functions on [0,1]).

The general case of a Riemann integrable function f can be settled in a similar way, by using Darboux integral sums.

The convergence provided by Weyl's ergodic theorem may be very slow.

In fact, we already noticed that

$$\int_0^1 \frac{\ln(t+1)}{t^2+1} dt = \int_0^{\pi/4} \ln(1 + \tan x)dx = 0.27220\ldots,$$

whereas the approximating sequence in (7) offers this precision only for $N > 10^4$.

However, Weyl's ergodic theorem has important arithmetic applications. A nice introduction is offered by the paper of P. Strzelecki [9]. Full details may be found in the monograph of R. Mañe [8].

An inspection of the argument of Weyl's ergodic theorem shows that the convergence (7) is uniform on [0,1] when $f : [0,1] \to \mathbb{C}$ is a continuous function with $f(0) = f(1)$.

It is worth mention that Gauss himself [4] was interested in the asymptotic behavior of dynamical systems involving the fractional part. In fact, in connection with the study of continued fractions he considered the dynamical system consisting of the map

$$G : [0,1) \to [0,1), \quad G(x) = \begin{cases} 0 & \text{if } x = 0 \\ \{\frac{1}{x}\} & \text{if } x \neq 0 \end{cases}, \tag{8}$$

and the invariant measure

$$d\mu(x) = \frac{1}{(\log 2)(1+x)} dx.$$

In the variant of Lebesgue integrability, the convergence defined by the formula (7) still works, but only *almost everywhere*. This was noticed by A. Ya. Khinchin [6], but can be deduced also from another famous result, Birkhoff's ergodic theorem, a large extension of Theorem 1. See [8] for details. It is Birkhoff's result that reveals the true nature of the Gauss map (8) and a surprising property of continued fractions (first noticed by A. Ya. Khinchin [7]). A nice account of this story (and many others) may be found in the book of K. Dajani and C. Kraaikamp [2].

REFERENCES

[1] G. D. Birkhoff, Proof of the ergodic theorem, *Proceedings of the National Academy of Sciences USA*, **17** (1931), 656-660

[2] K. Dajani and C. Kraaikamp, *Ergodic Theory of Numbers*, Carus Mathematical Monographs, The Mathematical Association of America, 2002.

[3] K. R. Davidson and A. P. Donsig, *Real Analysis with Real Applications*, Prentice-Hall. Inc., Upper Saddle River, 2002.

[4] C. F. Gauss, *Mathematisches Tagebuch 1796-1814*, Akademische Verlagsgesellschaft Geest & Portig K.G., Leipzig, 1976.

[5] B. Hayes, Gauss's Day of Reckoning. A famous story about the boy wonder of mathematics has taken on a life of its own, *American Scientist* **94** (2006), No. 3, pp. 200–205. Online: http://www.americanscientist.org/template/AssetDetail/assetid/50686?&print=yes

[6] A. Ya. Khinchin, Zur Birkhoffs Lösung des Ergodenproblems, *Math. Ann.* **107** (1932), 485-488.

[7] A. Ya. Khinchin, Metrische Kettenbruchprobleme, *Compositio Math.* **1** (1935), 361-382.

[8] R. Mañe, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, 1987.

[9] P. Strzelecki, On powers of 2, *Newsletter European Mathematical Society* No **52**, June 2004, pp. 7-8.

[10] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* **77** (1916), 313-352.

[11] A. L. O'Toole, Insights or Trick Methods?, *National Mathematics Magazine*, Vol. **15**, No. 1 (Oct., 1940), pp. 35-38.