T. Tao and the Syracuse conjecture

English translation of the paper "T. Tao et la conjecture de Syracuse" published in *La Gazette des Mathématiciens*, Number 168, April 2021

Jean-Paul Allouche

We will first recall the Syracuse conjecture (also known as the "3n + 1" problem) and give a very quick overview of the known results on the subject. Then we will attempt to give some of the ideas behind a remarkable recent result of T. Tao on this conjecture.

1   Introduction

The Syracuse conjecture, also called Collatz conjecture, Kakutani conjecture or $3x + 1$ problem (there is even a paper by B. Thwaites entitled *My conjecture*), is one of those extraordinarily attractive mathematical questions whose simplicity of statement is matched by their difficulty of proof, to the extent that many (most) of these questions are still open. One can think for example of the Goldbach conjecture or of the Fermat(–Wiles) theorem. A common characteristic of these very difficult conjectures is that their simple statements attract many amateurs, who are certainly well-intentioned, but who are sometimes difficult to convince that their approach is as naive as it is false. One can, however, hardly blame them, since even "professional mathematicians" are regularly seduced by these conjectures, and then realise that their own attempts towards a proof are unsuccessful. They probably do not know that P. Erdős once said to J. C. Lagarias, referring to this conjecture: "Hopeless. Absolutely hopeless", which is … not very encouraging.

Let us recall the statement of the Syracuse–Collatz–Kakutani–$(3x + 1)$–Thwaites problem.

Conjecture. *Let f be the function defined on the positive integers by*

$$f(n) = \begin{cases} \dfrac{3n + 1}{2} & \text{if } n \text{ is odd,} \\ \dfrac{n}{2} & \text{if } n \text{ is even.} \end{cases}$$

Then all the orbits of f are ultimately equal to $(1, 2, 1, 2, 1, 2, \ldots)$.

In other words, defining f^k as the k-th iterate of f, the orbit of every integer n under f, i.e., the sequence $(n, f(n), f^2(n), \ldots)$, contains the number 1, from which it alternatively takes the values 1 and 2.

Remark. (i) This conjecture is clearly equivalent to the following one: *Let* $v_2(n)$ *be the 2-adic valuation of the integer n (in other words, the largest integer a such that* 2^a *divides n) and g the function defined on the odd integers by* $\mathrm{Syr}(n) = (3n + 1)/2^{v_2(3n+1)}$. *Then, for any odd integer n, there exists an integer ℓ such that* $\mathrm{Syr}^\ell(n) = 1$.

(ii) The history of this conjecture and practically all the results before the one by T. Tao which is the subject of this paper can be found in the book by J. C. Lagarias [9].

We propose here to first recall the results that have been proved so far, and then to try to summarise Tao's result, which is stated as follows in [12].

Theorem (Tao). *Almost all orbits of f contain an almost bounded element.*

2   First steps

The simplicity of the statement of this conjecture is likely to impel us to "play" with it and to do experiments. If we compute the iterates of f on sufficiently small integers, we soon see that the orbits reach 1, and are therefore ultimately equal to $(1, 2, 1, 2, 1, 2, \ldots)$. We also see that the pre-period (i.e., the part before the periodic part) of the orbit of 27 is (curiously?) much longer than that of the smaller integers.

Another very general observation is that applying f to "half of all integers", namely the even ones, results in a value smaller than the starting one. Furthermore, if we consider integers of the form $(4m + 1)$, successive applications of f give $4m + 1 \to 6m + 2 \to 3m + 1$. Since $3m + 1 < 4m + 1$ (at least for $m > 0$), we thus obtain that "a quarter of the integers" leads, after just two iterations of f, to a number smaller than the starting one. So in fact at least "3/4 of all integers" belong to the set $S := \{n > 0 \mid \exists k, f^k(n) < n\}$. More precisely, the natural density of a set A of integers is by definition the limit, if it exists, of $\#\{n \in A \mid n \le x\}/x$ (and we define the upper and lower density by replacing the limit by the upper and lower limit). Thus what we just saw is equivalent to the statement

that the lower density of the set S is larger than or equal to $3/4$. If we now proceed by considering the integers of the form $8m + j$ with $j \in [0, 7]$, then the integers in the residual classes modulo 2^a for $a = 4, 5, \ldots$, we successively obtain lower bounds for the lower density of S which are closer and closer to 1. The author of this paper carried out these experiments in the second half of the 1970s with one of the first programmable pocket calculators (a TI 58): after keeping the machine running for forty-eight hours or more, we obtained values so close to 1 that it was tempting to try to prove this result, in the hope that it would be simpler than the initial conjecture.

It is now time to *do some mathematics* – simple for the moment – by stating the following result.

Theorem. *The lower density of S is equal to* 1*, and the same holds for the density of S.*

The proof is based upon the study of residual classes modulo 2^a. The above sketch for integers modulo 2 or 4 is easily generalised by induction on a.

Proposition. (i) *Let* $\alpha(a, j) := \#\{v \in [0, a - 1] \mid f^v(j) \text{ odd}\}$. *Then*

$$f^a(2^a n + j) = 3^{\alpha(a,j)} n + f^a(j).$$

(ii) *For all i in* $[0, a]$*, we have*

$$\#\{j \in [0, 2^a - 1] \mid \alpha(a, j) = i\} = \binom{a}{i}.$$

We see from this proposition that the study of $f(n)/n$ for $n \leq x$ will involve truncated sums of binomial coefficients. In order to estimate these sums, we use the following lemma.

Lemma. *For all* $\varepsilon \in (0, 1)$*, there exists an* $\eta \in (0, 1)$ *such that*

$$\frac{1}{2^a} \sum_{|i - a/2| > \varepsilon a} \binom{a}{i} \leq \eta^a$$

for all sufficiently large values of a.

A proof of this lemma suggested by G. Tenenbaum uses the relation

$$\sum_{i=0}^{m} \binom{a}{i} = (a - m)\binom{a}{m} \int_1^2 t^m (2 - t)^{a - m - 1} \, dt \quad \text{for } m \leq a - 1$$

which can be proved by finite induction on m and which is, by the way, one of the exercises in the beautiful book by L. Comtet (see [4, Exercise 12, p. 91]). For more details, one may refer to a paper by the author published in the *Séminaire de Théorie des Nombres de Bordeaux* [1]. The result on density 1 was also proved independently by C. J. Everett in 1977 [5], H. Möller in 1977 [10] and E. Heppner in 1978 [6], and also by R. Terras in his papers of 1976 and 1979 [13, 14], either for the initial problem or for

the generalisation due to H. Hasse, see the remark at the end of this section. These different papers written independently at about the same time suggest that the result on density 1 is not very difficult. It relies on a non-trivial limit, which is still a reasonably easy exercise. They also reveal the absence of electronic tools at that time (even if *Zentralblatt* existed in paper form as well as *MathSciNet* which was called *Mathematical Reviews*). I remember, however, that M. Mendès France told me about Heppner's and/or Möller's paper afterwards and put an "anonymous" letter with easily recognisable handwriting in my locker: *Forget about this problem. A friend who wishes you well* – which confirms Erdős' opinion as reported by Lagarias.

Why doesn't this density result provide a proof of the conjecture? It is the remaining term in the limit for the density computation that spoils the party – the result we actually obtain can be stated as follows:

$$\#\{n \leq x \mid \exists k, f^k(n) < n\} = x + O(x^{1-\delta})$$

for some δ in $(0, 1)$, and in fact even if we had $O(1)$, we would still only obtain a weak form of the conjecture, namely that any orbit will eventually "loop" (i.e., is ultimately periodic[1]), but there could be more than one loop. One can even refine this by making the above k depend on n (logarithmically), but this still remains far from the actual conjecture even in its weak form.

However, what the author had overlooked was that the result he had obtained in [1] (actually a bit more precise than that stated above) could yield something more, namely that the set $S_c := \{n \mid \exists k, f^k(n) < n^c\}$ is of density 1 for $c > 0.8691$. It was I. Korec who mentioned in 1994 in [7] that this was pointed out by the referee of his paper. Korec improves the 0.8691 value to $\log 3/\log 4$, which is about 0.7925 (see the review MR1290275 by Lagarias on *MathSciNet*).

Remark. A more general formulation of the conjecture, due to H. Hasse, is to replace "multiply by 3 and add 1, then divide by 2, or divide by 2, depending on the value modulo 2 of the starting integer", by "multiply by m and add a suitable residue in a complete system of fixed residues, then divide by d, or divide by d, depending on whether the starting integer is not or is divisible by d". The conjecture is then that there exists a function Φ such that if $m < \Phi(d)$, then all orbits are ultimately periodic and there is a finite number of possible periods, and if $m > \Phi(d)$, then there exists at least one non-ultimately periodic orbit. Möller [11] proposes the function $\Phi(d) = d^{d/(d-1)}$. (Note that $m = 3$ is "just" under the threshold $d^{d/(d-1)}$ for $d = 2$.) Some of the authors quoted above also give density results for this generalisation.

---

[1] Let us recall that a sequence $(u_n)_{n \geq 0}$ is said to be ultimately periodic if it is periodic for large enough indices, i.e., if there exist $n_0 \geq 0$ and $T \geq 1$ such that, for all $n \geq n_0$ and for all $k \geq 0$, we have $u_{n+kT} = u_n$.

3   Going a bit further

We now give some brief indications on other results (the book by Lagarias mentioned above is very complete and includes in particular an impressive annotated bibliography). We will not discuss the numerical results which produce huge integers N such that the conjecture is true for all integers $n \leq N$, nor to those which show that the number of elements of any possible period apart from $(1, 2)$ for the orbits of the function f is necessarily fantastically large.

An interesting theoretical question is: can we say anything about the integers n for which there exists k such that $f^k(n) = 1$? In other words, can we ask "how many" pre-images of 1 exist under the iterates of f? Because our lack of understanding of this function f, we cannot even say that this set has density 1. The best result obtained thus far is a lower bound of the type

$$\#\{n \leq x \mid \exists k, f^k(n) = 1\} > x^d$$

for sufficiently large x, with one of the most recent values for d being $d = 0.84$, see [8].

4   The result obtained by T. Tao

4.1   Introductory remarks

As is the case for the classical conjectures recalled at the beginning of this overview, whenever we fail to obtain a desired result, we always try to obtain at least a weaker form. A typical example is Goldbach's conjecture, for which J. R. Chen proved a weaker form in [3]: *Any sufficiently large even number is the sum of a prime number and a number having at most two prime factors.* Or think of the Fermat–Wiles theorem, for which an attempt was made to restrict to the "first case" (if $x^p + y^p = z^p$, then $xyz \equiv 0 \bmod p$). Some results of this kind have been described above, in particular the one which states that the density of the set $\{n \leq x \mid \exists k, f^k(n) < n^c\}$ is equal to 1 for $c > 0.7925$.

This last result can be expressed as follows: *Almost all integers have in their orbit by f an element smaller than the (0.7925)-th power of the considered integer.* This is of course a (convenient) abuse of language since density is not a probability on the integers. Using this terminology, one can ask which "best function" B could be obtained to replace n^c in the set $\{n \mid \exists k, f^k(n) \leq n^c\}$, while keeping the natural density of this set equal to 1. In other words, B should be such that almost all integers n have an element $\leq B(n)$ in their orbits by f. A *caveat* is necessary: as pointed out for example in Tao's paper, one cannot "improve" $n \to n^c$ by "iterating". Indeed, even if it is true that, for almost all integers n, there exists an element $n' = f^k(n)$ with $n' \leq n^c$, one cannot apply the result of "almost all" to n' to obtain an element $n'' = f^\ell(n')$ such that $n'' \leq (n')^c$, and thus $f^{k+\ell}(n) = f^\ell(n') = n'' \leq (n')^c \leq (n^c)^c = n^{c^2}$, because n' could very well belong to the set of exceptions of the "almost all" and have no associated n''. Let us also note that we do

not know how to obtain $B(n) = 1$, since we do not know (end of the previous section) that the Syracuse conjecture is true for almost all integers. Tao's "tour de force" is to prove, *up to replacing the natural density by the logarithmic density* (see below), that one can take for B *any function tending to infinity, as slowly as one wants* for an infinitely large argument. For example, Tao writes, perhaps as a wink to estimates "à la Erdős", $B(n) = \log\log\log\log n$. He summarises this in a figurative way, stating that we can take an "almost bounded" B.

Definition. A set of integers $A \subset \mathbb{N}$ is said to have a logarithmic density equal to δ if the limit, when x tends to infinity, of

$$\frac{1}{\log x} \sum_{n \leq x,\, n \in A} \frac{1}{n}$$

exists and equals δ.

Remark. If the natural density of a set of integers exists, its logarithmic density also exists and is equal to the natural density. The converse is not true.

4.2   T. Tao's theorem

As we have seen above, Tao stipulates in his theorem a striking, even mediatic (in the non-pejorative sense of the term …) statement: *Almost all orbits of f contain an almost bounded element.* This means that, for any function B which tends to infinity, the *logarithmic* density of the set $\{n \mid \exists k, f^k(n) < B(n)\}$ is equal to 1. We will try to describe (as Tao himself does in the first pages of his paper) in a heuristic way, yet avoiding technical details (the paper has 49 pages), the steps of the proof and the small improvements that it suggests.

(1) Studying the function f is classically equivalent to studying the function that Tao calls Syr. Let $v_2(n)$ be the 2-adic valuation of the integer n, that is $v_2(n) = a$ if 2^a divides n and 2^{a+1} does not divide n. We define the function Syr on odd integers by

$$\text{if } n \text{ is odd, then } \text{Syr}(n) = \frac{3n + 1}{2^{v_2(3n+1)}} = f^{v_2(3n+1)}(n).$$

Of course, the Syracuse conjecture is that, for any odd integer n, there exists an integer k such that $\text{Syr}^k(n) = 1$. And Tao's original statement is equivalent to: *Let B be a function defined on the odd integers that tends to infinity at infinity. Then, for almost all integers N, there exists an integer k such that $\text{Syr}^k(N) < B(N)$* (here "almost all integers" means that the set of odd integers for which the property is true, is of logarithmic density $1/2$ in the set of all integers).

(2) How do we compute the iterates of Syr for an odd integer N? Note that $a_1(N) := v_2(3N + 1)$, $a_2 := v_2(3\operatorname{Syr}(N) + 1)$, $a_3 := v_2(3\operatorname{Syr}^2(N) + 1)$, etc. Clearly,

$$\operatorname{Syr}(N) = (3N + 1)2^{-v_2(3N+1)}$$
$$= 3.2^{-a_1(N)}N + 2^{-a_1(N)},$$
$$\operatorname{Syr}^2(N) = (3\operatorname{Syr}(N) + 1)2^{-v_2(3\operatorname{Syr}(N)+1)}$$
$$= 3^2.2^{-a_1(N)-a_2(N)} + 3.2^{-a_1(N)-a_2(N)} + 2^{-a_2(N)},$$
etc.

and therefore,
$$\operatorname{Syr}^n(N) = 3^n 2^{-a_1(N)-a_2(N)-\cdots-a_n(N)}N$$
$$+ F_n(a_1(N), a_2(N), \ldots, a_n(N)),$$

where
$$F_n(a_1(N), a_2(N), \ldots, a_n(N))$$
$$:= 3^{n-1}2^{-a_1(N)-a_2(N)-\cdots-a_n(N)}$$
$$+ 3^{n-2}2^{-a_2(N)-a_3(N)-\cdots-a_n(N)}$$
$$+ \cdots + 3^1 2^{-a_{n-1}(N)-a_n(N)} + 2^{-a_n(N)}.$$

This formula can be compared with the one seen above for f: $f^a(2^a n + j) = 3^{a(a,j)}n + f^a(j)$, which essentially says that we can estimate the values of successive images of an integer in a class modulo 2^a from the images of a representative of this class, until a number of iterations equal to a (this number is thus of the order of the logarithm of the considered integer if we have chosen the representative in $[0, 2^a)$).

(3) Take as above $a_j(N) = v_2(3\operatorname{Syr}^{j-1}(N) + 1)$. Then, heuristically, *for a "typical" odd integer N large enough, and n much smaller than $\log N$, the vector $(a_1(N), a_2(N), \ldots, a_n(N))$ behaves like a geometric random vector of size n and parameter $1/2$, i.e., an n-tuple of independent random variables, all geometric with parameter $1/2$.* More precisely, the "behaves like" has to be taken in the sense of a small distance between random variables, where the distance between two discrete random variables X and Y taking their values in the same discrete space R is the total variation

$$\sum_{r \in R} |\mathbb{P}(X = r) - \mathbb{P}(Y = r)|.$$

A proposition proved in Tao's paper states that the heuristic property in italics at the beginning of step (3) is justified if N is uniformly distributed modulo 2^m for m slightly larger than $2n$. This gives a good control of $\operatorname{Syr}^n(N)$ for almost all N and for n of the order of $\gamma \log N$ with a small constant γ. Since one heuristically has an estimate like $\operatorname{Syr}^n(N) \approx (3/4)^n N$, in fact $\operatorname{Syr}^n(N) = e^{O(\sqrt{n})}(3/4)^n N$ (by the central limit theorem or by the Chernoff bound), one can in this way already obtain again Korec's result recalled above: the density of the set $S_c := \{n \mid \exists k, f^k(n) < n^c\}$ is 1 for $c > 0.8691$. As Tao points out, a result of this kind is somewhat analogous to

"almost sure local wellposedness results" for evolutionary partial differential equations, in which one has good short-time control for almost all initial conditions. Additionally, the theorem we want to prove is similar to an "almost sure almost global wellposedness" result. Now how to get from "local" to "global"?

(4) The last and most difficult step is to answer the above question by introducing a function that further "accelerates" the maps f and Syr seen above. This "first passage" function Pass is defined as follows: for $x \geq 1$ and any odd integer N, we write

$$T_x(N) := \inf\{n \in \mathbb{N} \mid \operatorname{Syr}^n(N) \leq x\}$$

with the usual convention that $T_x(N) := +\infty$ if $\operatorname{Syr}^n(N) > x$ for all n. The first passage function is then defined by

$$\operatorname{Pass}_x(N) := \operatorname{Syr}^{T_x(N)}(N).$$

Tao is then inspired by a work of J. Bourgain [2] who goes from a local almost everywhere to a global almost everywhere, thanks to the construction of an invariant probability measure. Alas! This is impossible here, but the author gets around this issue by introducing a *family* of probability measures which are approximately transported one to the other by iterating Syr a variable number of times. This is what will permit the use of an iterative argument, which was not feasible "directly" as we pointed out at the beginning of Section 4.1 with n^c and $(n^c)^c$.

We will not go any further in this attempt to demystify Tao's beautiful proof, whose high technicality, but above all its inventiveness, have been barely touched. To try to summarise it – too schematically – let us start by describing a temptation shared both by the professional mathematician who discovers the Syracuse conjecture and by the amateur: basically, iterating the application f from the beginning seems to consist *roughly* of replacing n every second time (when n is odd) by approximately $3/2 \cdot n$, and of replacing n every other time (when n is even) by $1/2 \cdot n$; in other words, applying f^2 amounts to multiplying n approximately by $3/4$. For example (with a "reasonably chosen" initial integer),

$$17 \xrightarrow{f} 26 \xrightarrow{f} 13 \xrightarrow{f} 20 \xrightarrow{f} 10 \xrightarrow{f} \cdots,$$

that is to say

$$17 \xrightarrow{f^2} 13 \xrightarrow{f^2} 10 \xrightarrow{f^2} \cdots.$$

Thus the orbit of a typical integer seems to be obtainable approximately by a sequence of multiplications by $3/4$. It is this temptation, which obviously does not constitute a proof, that Tao, at the cost of unprecedented effort and technicality for such an apparently innocent statement, has transformed into a proof for almost all integers. There should not be any misunderstanding about the purpose of this remark: to go from "we multiply roughly by $3/4$" to Tao's proof and its half a hundred pages is at least as difficult as transforming a frog or a toad into a charming princess or prince.

5 So, what now?

Now what can we expect for this conjecture? Tao indicates that, by further refining his approach, it should be possible to replace "almost all in logarithmic density" with "almost all in natural density". But he leaves little hope that the function tending to infinity as slowly as one likes in his statement can be replaced by a constant. In other words, even the statement "the orbit of almost any integer is ultimately periodic" is still totally out of reach.

References

[1] J.-P. Allouche, Sur la conjecture de "Syracuse–Kakutani–Collatz". In *Sém. Th. Nombres, Bordeaux*, CNRS, Talence, Exp. No. 9, 15 (1979)

[2] J. Bourgain, Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys.* **166**, 1–26 (1994)

[3] J. R. Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica* **16**, 157–176 (1973)

[4] L. Comtet, *Analyse combinatoire. Tome I*. Collection SUP: "Le Mathématicien", Presses Universitaires de France, Paris (1970)

[5] C. J. Everett, Iteration of the number-theoretic function $f(2n) = n$, $f(2n + 1) = 3n + 2$. *Adv. Math.* **25**, 42–45 (1977)

[6] E. Heppner, Eine Bemerkung zum Hasse-Syracuse-Algorithmus. *Arch. Math. (Basel)* **31**, 317–320 (1978/79)

[7] I. Korec, A density estimate for the $3x + 1$ problem. *Math. Slovaca* **44**, 85–89 (1994)

[8] I. Krasikov and J. C. Lagarias, Bounds for the $3x + 1$ problem using difference inequalities. *Acta Arith.* **109**, 237–258 (2003)

[9] J. C. Lagarias, *The ultimate challenge: the $3x + 1$ problem*. Amer. Math. Soc., Providence, RI (2010)

[10] H. Möller, *F*-Normalreihen. *J. Reine Angew. Math.* **289**, 135–143 (1977)

[11] H. Möller, Über Hasses Verallgemeinerung des Syracuse-Algorithmus (Kakutanis Problem). *Acta Arith.* **34**, 219–226 (1977/78)

[12] T. Tao, Almost all orbits of the Collatz map attain almost bounded values. arXiv:1909.03562 (2020)

[13] R. Terras, A stopping time problem on the positive integers. *Acta Arith.* **30**, 241–252 (1976)

[14] R. Terras, On the existence of a density. *Acta Arith.* **35**, 101–102 (1979)

———

Jean-Paul Allouche is *Directeur de recherche* emeritus at CNRS. He is working at IMJ-PRG, Sorbonne, Paris on subjects relating number theory and theoretical computer science, including the so-called *automatic sequences.*

jean-paul.allouche@imj-prg.fr